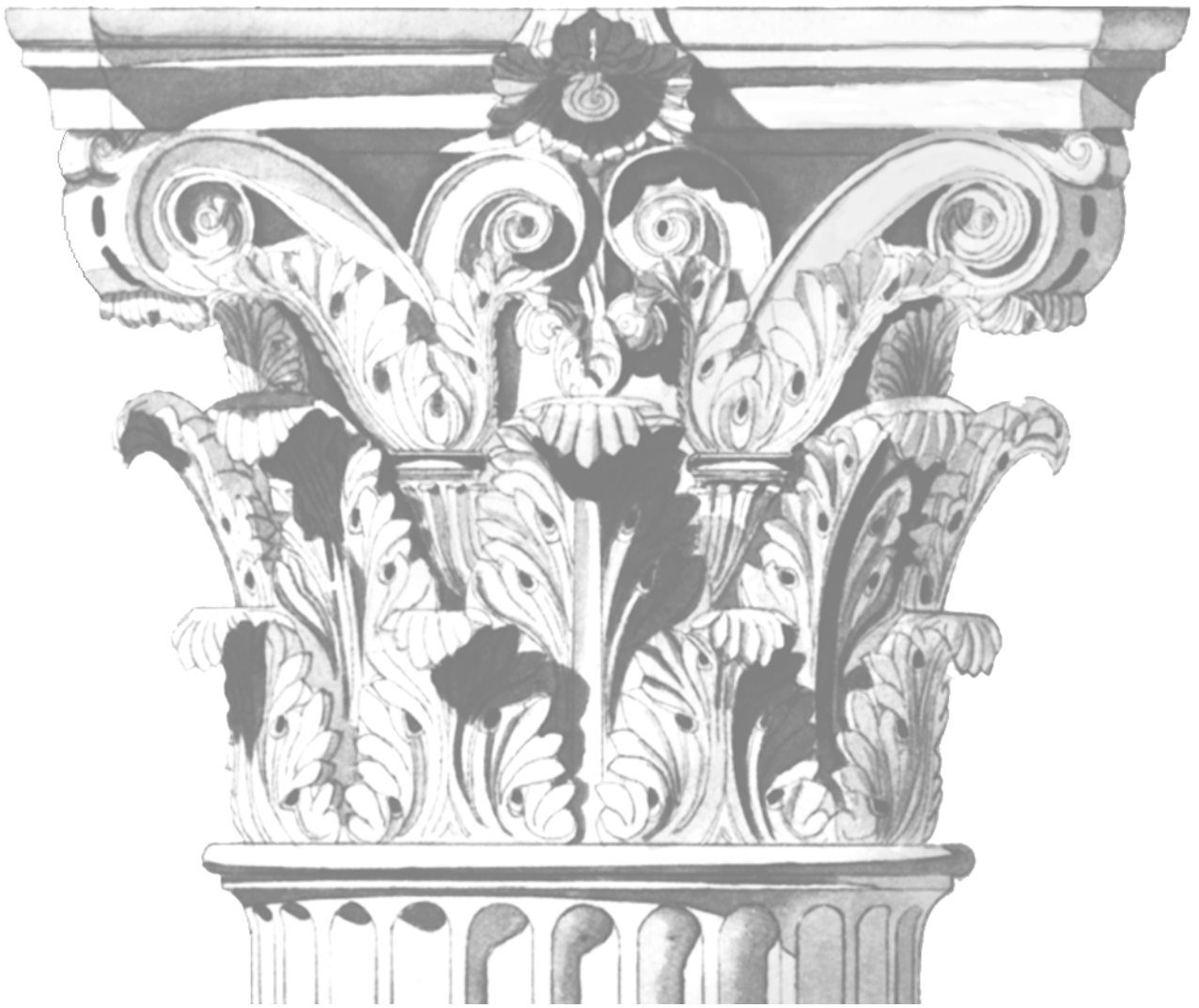
HUMAN LYSYL HYDROXYLASE

Identification of the residue involved in the binding of 2-oxoglutarate at the catalytic site and characterization of a novel isoenzyme, LH3, and its gene

KAISA PASSOJA

Collagen Research Unit, University of Oulu
Biocenter Oulu, University of Oulu
Department of Medical Biochemistry,
University of Oulu

OULU 2000



KAISA PASSOJA

HUMAN LYSYL HYDROXYLASE

Identification of the residue involved in the binding of 2-oxoglutarate at the catalytic site and characterization of a novel isoenzyme, LH3, and its gene

Academic Dissertation to be presented with the assent of the Faculty of Medicine, University of Oulu, for public discussion in the Auditorium of the Department of Medical Biochemistry, on September 14th, 2000, at 10 a.m.

OULUN YLIOPISTO, OULU 2000

Manuscript received: 14 August 2000
Manuscript accepted: 15 August 2000

Communicated by
Professor Heather N. Yeowell
Professor Juhani Syväoja

ISBN 951-42-5732-4 (URL: http://herkules.oulu.fi/isbn9514257324/)

ALSO AVAILABLE IN PRINTED FORMAT
ISBN 951-42-5731-6
ISSN 0355-3221(URL: http://herkules.oulu.fi/issn03553221/)

OULU UNIVERSITY PRESS
OULU 2000

Passoja, Kaisa, Human lysyl hydroxylase Identification of the residue involved in the binding of 2-oxoglutarate at the catalytic site and characterization of a novel isoenzyme, LH3, and its gene

Collagen Research Unit, University of Oulu, Kajaanintie 52 A, FIN-90220 Oulu, Finland, Biocenter Oulu, University of Oulu, P.O.B. 5000, FIN-90014 University of Oulu, Finland, Department of Medical Biochemistry, University of Oulu, P.O.Box 5000, FIN-90014 University of Oulu, Finland

(Manuscript received: 14 August 2000)

Abstract

Lysyl hydroxylase (E.C. 1.14.11.4, protocollagen-lysine 2-oxoglutarate 5-dioxygenase, PLOD) catalyses the formation of hydroxylysine in collagens and other proteins with collagen-like sequences. The hydroxylysine residues participate in the formation of collagen crosslinks and serve as attachment sites for carbohydrate units. The importance of lysine hydroxylation is demonstrated by the critical manifestations found in patients with the type VI variant of the Ehlers-Danlos syndrome, which is caused by a deficiency in lysyl hydroxylase activity.

Lysyl hydroxylase requires Fe^{2+}, 2-oxoglutarate, O_2 and ascorbate. The binding site for the C-5 carboxyl group of 2-oxoglutarate is characterized here by site-directed mutagenesis. Two conserved and one non-conserved amino acid residues at the possible binding site in human lysyl hydroxylase 1 were converted individually to alanine or lysine and the mutant polypeptides were expressed in insect cells. Mutation of arginine-700 to alanine inactivated the enzyme completely, whereas mutation of the other two residues had only a minor effect. In addition, the K_m of the arginine-700 to lysine mutant polypeptide for 2-oxoglutarate was increased 10-fold. The results thus indicate that this conserved arginine is the residue that binds the C-5 carboxyl group of 2-oxoglutarate in lysyl hydroxylases.

A novel human lysyl hydroxylase isoenzyme, termed lysyl hydroxylase 3, was identified, cloned and characterized here. The overall amino acid sequence identity between the novel human lysyl hydroxylase isoenzyme and the other human lysyl hydroxylase isoenzymes is about 60%. The highest expression levels of the mRNA for lysyl hydroxylase 3 among the tissues studied were found in the placenta, pancreas and heart. The novel isoenzyme was expressed as a recombinant protein in insect cells, and the protein was shown to function as a lysyl hydroxylase *in vitro* hydroxylation experiments using short synthetic peptides as substrates. No differences in catalytic properties were found between the recombinant lysyl hydroxylases 3 and 1.

The structure of the human gene for lysyl hydroxylase 3 was determined in the last part of this work. The gene is shown to be only 11.6 kb in size and to contain 19 exons. Transcription was found to be initiated at multiple sites, and the introns contained 15 full-length Alu retroposons or partial Alu fragments of more than 100 bp. The present characterization of the exon-intron organization of the gene will provide a basis for further studies to determine whether there is any genetic disease that is attributable to mutations in this gene.

Keywords: collagen, PLOD, oxygenase, gene structure.

Acknowledgements

This work was carried out at the Collagen Research Unit of the Department of Medical Biochemistry, University of Oulu, and Biocenter Oulu during the years 1995-2000.

My deepest gratitude is due to my supervisor, Professor of the Academy of Finland Kari Kivirikko, for introducing me to the field of basic scientific research and molecular biology. His broad knowledge of science and his great optimism have provided an excellent basis for my work. I would also like to express my sincere thanks to Professor Taina Pihlajaniemi for providing excellent research facilities during these years. I wish to thank Professor Ilmo Hassinen for his contribution in creating a good scientific atmosphere in the department. Docent Johanna Myllyharju and Docent Leena Ala-Kokko deserve my honest thanks for their collaboration and many helpful comments over the years.

I am grateful to Professor Heather Yeowell and Professor Juhani Syväoja for their constructive comments on the manuscript, and to Malcolm Hicks, M.A, for his careful revision of its language.

I wish to acknowledge Asta Pirskanen and Taru Kosonen for sharing their knowledge about science while we were working together with lysyl hydroxylases, and for their constant positive attitude. My thanks are also due to Tarja Helaakoski. My warmest thanks go to my colleague Kati Rautavuoma for her pleasant and sometimes hilarious collaboration, for her comments on all manner of subjects and for her friendship, all of which things have been of especial value to me. I wish to thank Minna Nokelainen and Anne Snellman for many cheerful and sometimes even desperate moments when organizing the dissertation party, and also Tero Pihlajamaa for being so much involved in it all, even though from a distance. Anne Latvanlehto deserves my special thanks for her friendship over the years and for sharing many unforgettable journeys. Pirkko Huhtala and Ritva Nissi are thanked for many scientific and non-scientific discussions, and companionship in our office together with Kati and Minna. I would also like to thank Raija Soininen, Joni Mäki, Laura Passi and all my other colleagues and friends at the Department.

I wish to thank Anu Myllymäki for her excellent and thorough technical assistance at all stages during this work and for many opinions exchanged while pipetting in the lab. Jaana Jurvansuu and Jaana Träskelin are also acknowledged. Sincere thanks go to Ari-

Pekka Kvist, Juha Näpänkangas and Marko Kervinen for their help with the computers, and to Marja-Leena Kivelä, Auli Kinnunen, Pertti Vuokila and Seppo Lähdesmäki for their kind help with many daily problems.

My parents have encouraged me enormously all the way during my work on this thesis, and I am very grateful for their trust and love. I also wish to thank my brother and my sister-in-law, and all my friends outside the department and at the medical school for "keeping me in touch with the real world". It is great you didn´t give up!

Finally, I wish to thank my fiancé Juha for his support over all these years and for his sometimes tricky comments on science and life. Without him all this would have been much simpler, but not nearly so productive. I would also like to thank him for his help in drawing the figures for the thesis.

The work has been supported by Suomalainen Lääkäriseura Duodecim, the Finnish Cultural Foundation, Kansan Sivistysrahasto and the Finnish Centre of Excellence Programme (2000-2005) of the Academy of Finland (44843).

Oulu, July 2000 Kaisa Passoja

Abbreviations

bp	base pair
C-	carboxy-
cDNA	complementary deoxyribonucleic acid
CLS	core-specific lectin
DDR1 and DDR2	discoidin domain receptors 1 and 2
DEAE	diethyl aminoethyl
DTT	dithiothreitol
dEST	database of expressed sequence tags
ER	endoplasmic reticulum
kb	kilobase(s)
kDa	kilodalton(s)
K_m	Michaelis constant
LH	lysyl hydroxylase
mRNA	messenger RNA
N-	amino-
N-linked	asparagine-linked
NMR	nuclear magnetic resonance
nt	nucleotide(s)
p	short arm of chromosome
PAGE	polyacrylamide gel electrophoresis
PCR	polymerase chain reaction
q	long arm of chromosome
RACE	rapid amplification of cDNA ends
RER	rough endoplasmic reticulum
SDS	sodium dodecyl sulphate
3´UTR	3´untranslated region
V_{max}	maximal velocity
X	any amino acid
Y	any amino acid

List of original articles

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

I Passoja K, Myllyharju J, Pirskanen A & Kivirikko KI (1998) 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.

II Passoja K, Rautavuoma K, Ala-Kokko L, Kosonen T & Kivirikko KI (1998) Cloning and characterization of a third human lysyl hydroxylase isoform. Proc. Natl. Acad. Sci. U. S. A. 95: 10482-10486.

III Rautavuoma K, Passoja K, Helaakoski T & Kivirikko KI (2000) Complete exon-intron organization of the gene for human lysyl hydroxylase 3 (LH3). Matrix Biol. 19: 73-79.

Contents

1 Introduction

The extracellular matrix is a network-like structure that surrounds the cells in every tissue. It varies greatly in its form and components between tissues, being mainly composed of collagens, elastin, proteoglycans and glycoproteins. Its primary role is to provide mechanical support and protection for the cells and to take part in cell attachment and migration.

Collagens are the major constituents of the extracellular matrix in all vertebrates, constituting one-third of the human body proteins. They form the majority of the organic material in bone and are also major components of the skin, cartilage, tendon, basement membranes, vascular walls and parts of the eye. The characteristic structural feature of a collagen is a triple helix of high mechanical strength. At least nineteen collagen types have been characterized so far, encoded by more than 30 genes dispersed through the genome.

A characteristic feature of collagen biosynthesis is the presence of a number of cotranslational and post-translational modifications, lysyl hydroxylase being one of the important enzymes participating in these intracellular modifications. The enzyme is a homodimer that requires Fe^{2+}, 2-oxoglutarate, O_2 and ascorbate for its catalytic activity. It functions in the endoplasmic reticulum before formation of the collagen triple helix by hydroxylating lysine residues in collagenous sequences. The resulting hydroxylysine residues participate in the formation of crosslinks between collagen molecules, thus stabilizing the extracellular matrix, and also serve as attachment sites for carbohydrates linked to collagens. This glycosylation is unique to collagenous proteins, but its functions are not fully understood. Different collagen types have variable amounts of hydroxylysine, and the hydroxylysine content of the same collagen type may vary between tissues. Inability to hydroxylate lysine residues on account of a deficiency in lysyl hydroxylase activity leads to a disease called the type VI variant of the Ehlers-Danlos syndrome.

Many previous studies have suggested that lysyl hydroxylase may have isoenzymes, and one such isoenzyme, termed lysyl hydroxylase 2, has recently been cloned. In the present work an additional, novel human lysyl hydroxylase isoenzyme, termed lysyl hydroxylase 3, is identified, cloned and characterized and the expression pattern of its mRNA is studied in a number of tissues. A recombinant protein is expressed in insect

cells and used to study the catalytic properties of the novel isoenzyme. Also the exon-intron organization and transcription initiation site of the gene for human lysyl hydroxylase 3 are determined. In order to obtain a better understanding of the function of the lysyl hydroxylases, the binding site of 2-oxoglutarate, an important substrate for the hydroxylation reaction, is characterized in human lysyl hydroxylase 1.

2 Review of the literature

2.1 Occurrence and functions of hydroxylysine

2.1.1 Hydroxylysine in collagens

The extracellular matrix is a support and repair component in most organs and tissues. Its main protein components are collagens, a family of closely related but chemically distinct macromolecules. Nowadays this superfamily in vertebrates contains at least 19 types with more than 30 distinct polypeptide chains. All collagen molecules consist of three polypeptide chains, called α chains. During folding of the macromolecule the three α chains first individually form left-handed helices and then join together to form the collagen triple helix by winding around each other in a right-handed manner. In some collagen types all three α chains in the molecule are identical, while in other types the molecule contains two or even three different chains. Typical of all α chains, except for short sequences at their ends, is the presence of a repeating triplet amino acid sequence, Gly-X-Y. In addition to the actual triple-helical collagenous domains, all collagen types also contain non-collagenous domains of varying sizes.

Collagen biosynthesis is characterized by the presence of an unusually large number of cotranslational and post-translational modifications, including intracellular modifications of precursor collagen chains, such as hydroxylation of lysine and proline residues, glycosylation of hydroxylysine residues and folding of the triple helix, and also extracellular modifications such as cleavage of the N- and C-propeptides and formation of supramolecular aggregates of various kinds. Fibrillar collagens form crosslinked fibrils, while other collagen types form network-like structures, for example, bind to the surface of collagen fibrils or form beaded filaments (see Hulmes 1992, Kielty *et al.* 1993, Kivirikko 1993, Prockop *et al.* 1994, Kivirikko 1995, Pihlajaniemi & Rehn 1995, Prockop & Kivirikko 1995, Bateman JF *et al.* 1996, Kivirikko & Pihlajaniemi 1998, Lamandé & Bateman 1999).

Hydroxylysine in collagens and other proteins with collagen-like sequences is almost exclusively found in the Y positions of the repeating Gly-X-Y triplet sequences. The exceptions include at least the sequences X-Hyl-Ser and X-Hyl-Ala, which are found in the short non-triple-helical regions (telopeptides) situated at the ends of the α chains (see Kivirikko 1993, Kivirikko & Pihlajaniemi 1998).

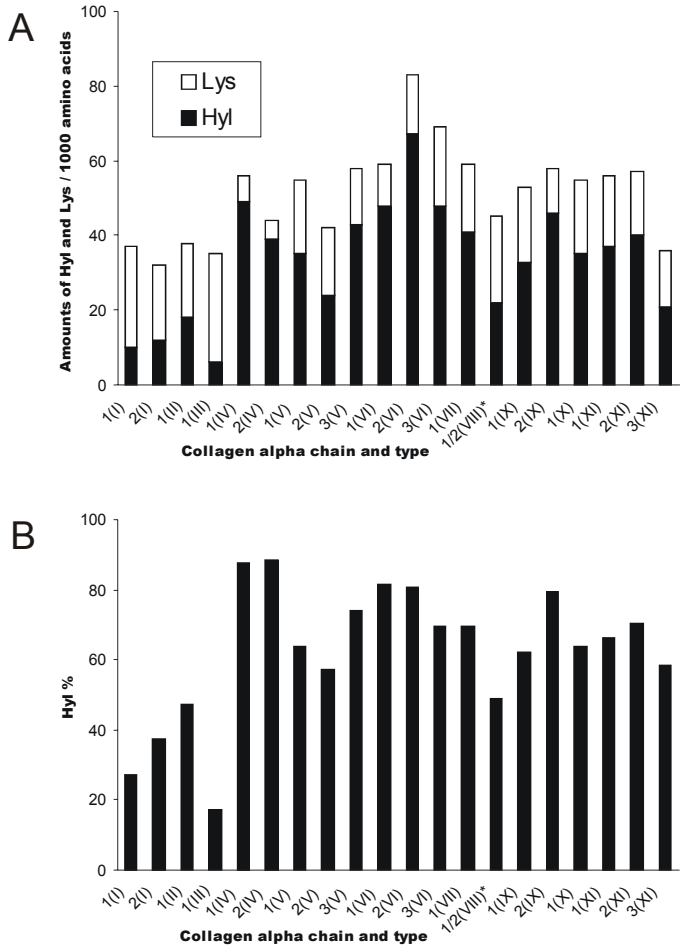


Figure 1. (A) Amounts of hydroxylysine and lysine residues in the triple-helical domains of various collagen chains. White columns indicate lysine (Lys) residues and black columns hydroxylysine (Hyl) residues / collagen chain. The values for Hyl and Lys residues on the y-axis are given per 1000 amino acids residues. (B) Percentages of lysine residues hydroxylated in different collagen α chains. The figure is modified from a table presented by Kivirikko et al. (1992).

*The value shown is the mean for two 50 000 kDa fragments of the triple-helical protein.

Great variation is found in the extent of lysine hydroxylation, not only between collagen types (Figure 1) (Reese *et al.* 1982, Jander *et al.* 1983, Schmid & Linsenmayer 1983, Kapoor *et al.* 1986, Burgeson RE & Morris NP 1987, Eyre & Wu JJ 1987, Kivirikko *et al.* 1992), but even within the same collagen type from different sources and the same tissue in different physiological and pathological conditions (Kivirikko & Myllylä 1980). This is at least partly explained by incomplete hydroxylation of lysine residues in the Y positions of the Gly-X-Y triplets, as a number of lysine residues in these positions in some collagens are not hydroxylated at all (Kivirikko & Pihlajaniemi 1998). As seen in Figure 1B, the highest amounts of hydroxylated lysines are found in the α chains of type IV collagen, where almost 90 % of the lysines are hydroxylated. The α1(VI), α2(VI) and α2(IX) chains are also effectively hydroxylated, to extents varying from 79 to 81 %. The lowest amount of hydroxylated lysines, only 17 %, is found in the α1(III) chain (Kivirikko *et al.* 1992).

The extent of lysine hydroxylation in collagens also varies as a function of age. Miller *et al* (1967) were the first to report that the hydroxylysine content of decalcified chick long bones decreased by nearly one-half from hatching to maturity. The hydroxylysine content of bone, skin and tendon collagens was examined in the chick from the 14-day embryo to the 18-month-old adult, and a significant fall in the extent of hydroxylation was seen immediately after hatching (Barnes *et al.* 1974). According to one study, the extent of lysine hydroxylation in collagen from embryonic chick long bones and the mandibular bone was about 3-fold relative to that in adult animals, whereas no significant difference was detected in the extent of hydroxylation in the collagen from the frontal bone synthesized during early embryonic development or during the postnatal period (Strawich & Glimcher 1983). Other reports also confirm the higher hydroxylysine content of collagen from embryonic tissues relative to that from adult tissues (Bailey & Shimokomaki 1971, Ryhänen & Kivirikko 1974a).

2.1.2 Hydroxylysine in other proteins

In addition to the collagens, hydroxylysine is also found in proteins with collagenous sequences but not defined as collagens. These include the subcomponent C1q of complement (Reid 1979). The complement system is the principal means by which antibodies defend vertebrates against bacterial infections (Alberts *et al.* 1994), and the subcomponent C1q consists of six A, six B and six C polypeptide chains each containing at their N-terminal ends about 80 amino acid residues of a collagen-like $(Gly\text{-}X\text{-}Y)_n$ sequence with hydroxylated lysine residues. Some hydroxylysine residues are glycosylated (Reid 1979, Reid & Porter 1981, Reid 1983). Another example of a non-collagenous protein containing hydroxylysine is acetylcholinesterase, a protein associated mainly with cells responsible for cholinergic synaptic transmission. This contains a triple-helical collagen-like tail structure which is associated with membranes and is rich in hydroxylysine (Rosenberry *et al.* 1982).

Additional proteins containing hydroxylysine residues in collagen-like sequences include collectins, members of the C-type lectin family (Drickamer 1999). Also belonging to this protein family are the pulmonary surfactant proteins SP-A and SP-D (Floros *et al.* 1986, Persson *et al.* 1989, Lu *et al.* 1992, Crouch *et al.* 1994), mannan-binding protein (Drickamer *et al.* 1986, Colley & Baenziger 1987), conglutinin (Davis & Lachmann 1984, Lee *et al.* 1991, Jensenius *et al.* 1994) and CL-40 (Jensenius *et al.* 1994, Lim *et al.* 1994). Collectins are oligomeric molecules composed of trimeric subunits containing a collagen-like sequence attached to a C-terminal globular carbohydrate-recognizing domain via alpha-coiled regions. They are believed to play an important role in many biological events such as cell-cell adhesion, serum glycoprotein turnover and innate immunity defence. According to some studies all collectins may have a single origin and be the result of ancestral gene rearrengements, and all collectin genes studied so far have actually been mapped to human chromosome 10 (Holmskov & Jensenius 1993, Kolble *et al.* 1993, Hansen & Holmskov 1998, Drickamer 1999). Other proteins containing a collagenous sequence and potential hydroxylysine residues include the type I and type II macrophage scavenger receptors (Kodama *et al.* 1990, Rohrer *et al.* 1990), the macrophage receptor MARCO (Elomaa *et al.* 1995, Elomaa *et al.* 1998), an adipose-specific collagen-like factor apM1 (Maeda *et al.* 1996) and a src-homologous, collagen-like (SHC) protein (Pelicci *et al.* 1992, Thomas *et al.* 1995) and ficolins (Ichijo *et al.* 1993, Lu *et al.* 1996).

In addition to proteins with a collagen-like-sequence, hydroxylysine residues are found in some proteins without a collagenous sequence. Anglerfish somatostatin-28 is a peptide hormone consisting of 28 residues, in which residue 23 is hydroxylysine in the sequence Trp-Hyl-Gly (Andrews *et al.* 1984). The function of this hydroxylysine residue remains uncertain (Andrews *et al.* 1984, Spiess & Noe 1985). Other examples include human tissue plasminogen activator (rtPA), human CD4 receptor (rCD4) and a related chimeric protein (rCD4-IgG), each of which contains one hydroxylysine residue in a X-Lys-Gly sequence at a surface-accessible site (Molony *et al.* 1995).

2.1.3 Functions of hydroxylysine

Hydroxylysine residues have at least two important functions in collagens: firstly their hydroxyl groups serve as sites of attachment for carbohydrate units, and secondly they are essential for the stability of the intermolecular collagen crosslinks (Kivirikko & Pihlajaniemi 1998).

2.1.3.1 Hydroxylysine-linked carbohydrate units

The carbohydrate units present in collagens are the monosaccharide galactose and the disaccharide glucosylgalactose. Two specific enzymes, hydroxylysyl galactosyltransferase (EC 2.4.1.50), transferring galactose to hydroxylysine residues, and

galactosylhydroxylysyl glucosyltransferase (EC 2.4.1.66), transferring glucose to galactosylhydroxylysine residues, are involved in the glycosylation reactions. The extent of glycosylation of hydroxylysine residues varies markedly between the collagen types and also within the same collagen type in various physiological and pathological states. A marked variation is also found in the ratio of the monosaccharide to the disaccharide, the predominant form being dependent on the collagen type (Kivirikko & Myllylä 1979, Kivirikko 1995).

The functions of the hydroxylysine-linked carbohydrate units are only partly understood. Becauce of their situation on the surface of the collagen molecules, it has been proposed that they may regulate the packing of collagen molecules into supramolecular assemblies (Kivirikko 1995). Experiments on *de novo* generation of type I collagen fibrils from type I procollagen *in vitro* have shown that collagen with an elevated content of hydroxylysine and hydroxylysine-linked carbohydrates forms thinner fibrils than the same protein with a normal degree of these modifications (Torre-Blanco *et al.* 1992). Similar results were reported by Notbohm *et al.* (1999), who studied the morphology of recombinant human type II collagen fibrils *in vitro,* and by Brinckmann *et al* (1999), who analyzed the fibrotic skin in lipodermatosclerosis. Opposite opinions also exist, however, as an electron microscope evaluation did not show any relationship between the level of collagen glycosylation and the diameter of fibrils formed *in vivo* or fibrillar aggregates (Batge *et al.* 1997). The rate of fibril formation with a highly hydroxylated and glycosylated type II collagen was found to be much slower than that with a collagen containing only very low amounts of hydroxylysine and carbohydrate units (Notbohm *et al.* 1999).

An additional hypothesis concerning the functions of glycosylated hydroxylysine residues has appeared recently, after the surprising finding that collagen molecules can bind to and activate receptor tyrosine kinases (Shrivastava *et al.* 1997, Vogel *et al.* 1997). These are a protein family needed in the control of cell growth, differentiation, metabolism and cell migration, the members of which probably share similar activation mechanisms (Schlessinger 1997, Vogel *et al.* 1999). The subgroup of discoidin domain receptors (DDR) is distinguished from the other members of this protein family by a discoidin homology repeat. Collagen types I-VI and XI are known to activate the discoidin domain receptor 1 (DDR1), whereas the discoidin domain receptor 2 (DDR2) seems to be activated only by fibrillar collagens, mainly types I and III (Schlessinger 1997, Vogel *et al.* 2000). Stimulation of the activity of both the DDR1 and DDR2 receptor tyrosine kinases requires the native triple-helical structure of collagen and occurs only with a certain delay in time (Schlessinger 1997, Vogel *et al.* 1999). The ability of a deglycosylated collagen to activate DDR2 is significantly reduced, thus indicating a central role for glycosylated hydroxylysine residues in the action of DDR2, whereas DDR1 seems to be activated normally even by deglycosylated collagen molecules (Vogel *et al.* 1997).

2.1.3.2 Collagen crosslinks

The formation of covalent crosslinks in fibrillar collagens occurs as an extracellular process after cleavage of the propeptides and fibril assembly (Kadler *et al.* 1996). The function of these crosslinks is to provide tensile strength and mechanical stability for the collagen fibres (Hulmes 1992). In spite of the important function of collagen crosslinks, their pattern varies greatly between tissues and also as a function of age (Eyre *et al.* 1988, Bailey *et al.* 1998). Two pathways of enzymatic crosslinking can be distinguished for fibrillar collagens, one based on lysine aldehydes and the other on hydroxylysine aldehydes (Kielty *et al.* 1993). The lysine aldehyde pathway predominates at least in adult skin, cornea and sclera and rat tail tendon, whereas the hydroxylysine aldehyde pathway predominates in bone, cartilage, ligaments, tendons, most internal organs and embryonic skin (Eyre 1987). Crosslinks formed in the hydroxylysine-derived aldehyde pathway are more stable than those formed in the lysine-derived pathway (Kielty *et al.* 1993).

Lysyl oxidase, a copper-dependent enzyme (for a review, see Smith-Mungo & Kagan 1998), initiates the crosslinking of collagens by catalyzing oxidative deamination of the ε-amino group in certain lysine and hydroxylysine residues situated in the telopeptide regions that remain after removal of the procollagen propeptides. This is the only step in the crosslinking process that is known to be under direct enzymatic control. The resulting aldehyde derivatives subsequently spontaneously react with lysine and hydroxylysine residues in the triple-helical regions on adjacent collagen molecules to form covalent crosslinks. The aldehydes can form two types of crosslinks, either by aldol condensation between two of the aldehydes or by condensation beetween one aldehyde and one ε-amino group of an unmodified lysine, hydroxylysine or glycosylated hydroxylysine residue. Bifunctional crosslinks can further spontaneously undergo intramolecular and intermolecular reactions to form several kinds of tri- and tetrafunctional crosslinks (see Eyre *et al.* 1984, Eyre 1987, Kielty *et al.* 1993, Kivirikko 1995, Prockop & Kivirikko 1995, Bailey *et al.* 1998, Knott & Bailey 1998).

2.1.3.3 Clinical aspects

Lysine residues in collagens have been reported to be overhydroxylated in diseases such as osteoporosis (Bailey *et al.* 1992, Knott *et al.* 1995), osteogenesis imperfecta (Kirsch *et al.* 1983, Lehmann *et al.* 1992, Lehmann *et al.* 1995a) and osteosarcoma (Shapiro & Eyre 1982, Lehmann *et al.* 1995b). Overmodification of lysine and hydroxylysine residues in type I collagen in osteogenesis imperfecta seems to be independent of the thermal denaturation temperature of this collagen (Rao *et al.* 1989). In agreement with this, Nokelainen *et al.* (1998) found that recombinant human type II collagens with either a high or a low hydroxylysine content had essentially identical denaturation temperatures, indicating that hydroxylysine and its glycosylated forms have no effect on the thermal stability of collagens. Overmodification of collagen polypeptides in osteogenesis imperfecta is caused by structural defects in the type I collagen α chains, resulting in

perturbation of helix formation (Bonadio & Byers 1985, Bateman *et al.* 1987). The defects present in the α chains increase the time spent by the polypeptide chains in a non-triple-helical conformation and thus lead to more extensive post-translational modification such as hydroxylation and glycosylation, and possibly to defective crosslinking (Hulmes 1992).

Transient increases in lysine hydroxylation also occur during bone fracture (Glimcher *et al.* 1980) and dermal wound healing and in fibrotic situations (Bailey *et al.* 1975, Reiser & Last 1986), leading to increased amounts of hydroxylysine-derived crosslinks. Other conditions which increase the hydroxylysine content of collagens include deficiencies in vitamin D (Toole *et al.* 1972, Barnes *et al.* 1973a, Dickson *et al.* 1979) and calcium ions (Barnes *et al.* 1973b).

Underhydroxylation of lysine residues is found in Ehlers-Danlos syndrome type VI (Krane *et al.* 1972), which is discussed in more detail in section 2.3.

2.2 Lysyl hydroxylase

Lysyl hydroxylase (protocollagen-lysine 2-oxoglutarate 5-dioxygenase, E.C. 1.14.11.4) catalyzes the formation of hydroxylysine by the hydroxylation of peptide-bound lysine residues in X-Lys-Gly triplets in collagens and more than 10 other proteins with collagenous domains (for reviews, see Kivirikko *et al.* 1992, Kivirikko & Pihlajaniemi 1998).

Lysyl hydroxylase is very similar in its catalytical properties to prolyl 4-hydroxylase, another collagen hydroxylase (for a review, see Kivirikko & Myllyharju 1998), and at the beginning of collagen research it was regarded as possible that the hydroxylation of lysine and proline residues might be catalyzed by a single enzyme (Kivirikko & Prockop 1967a, Kivirikko & Prockop 1967b). It was demonstrated in the late 1960´s, however, that these two enzymatic activities are located at different sites (Weinstein *et al.* 1969) and that purified proline hydroxylase does not act on lysine residues (Halme *et al.* 1970, Rhoads & Udenfriend 1970, Miller 1971). Kivirikko and Prockop (1972) partially purified lysyl hydroxylase from chick embryos by DEAE cellulose chromatography and gel filtration and demonstrated that lysyl hydroxylase and prolyl hydroxylase are indeed separate enzymes. After partial purification of lysyl hydroxylase by conventional protein purification methods (Ryhänen 1976), it was further purified by an affinity chromatography procedure involving concanavalin A-agarose (Turpeenniemi *et al.* 1977). It was first purified to homogeneity from chick embryo (Turpeenniemi-Hujanen *et al.* 1980) and then from human placenta by combining affinity chromatography on concanavalin A-agarose with affinity chromatography on collagen linked to agarose and with conventional methods (Turpeenniemi-Hujanen *et al.* 1981).

2.2.1 Molecular properties

The activity of lysyl hydroxylase purified from chick embryo and human placenta is recovered in two separate gel filtration peaks, with molecular weights corresponding to about 550 000 and 190 000 (Kivirikko & Prockop 1972, Ryhänen 1976, Turpeenniemi *et al.* 1977, Turpeenniemi-Hujanen *et al.* 1980, Turpeenniemi-Hujanen *et al.* 1981). The active enzyme of size 170 000 Da is a homodimer, and its subunit has a molecular weight of about 85 000 (Turpeenniemi-Hujanen *et al.* 1980, Turpeenniemi-Hujanen *et al.* 1981, Myllylä *et al.* 1988). The larger protein, corresponding to 550 000 Da, is also seen in SDS-PAGE analysis as one band with a molecular weight of 85 000, however, suggesting that it is probably an aggregate of the enzyme dimers (Turpeenniemi-Hujanen *et al.* 1980, Turpeenniemi-Hujanen *et al.* 1981).

Lysyl hydroxylase was cloned for the first time from chick embryo. A surprising finding was that no significant homology was detected between the primary structures of chick lysyl hydroxylase (Myllylä *et al.* 1991) and the two types of subunit of the prolyl 4-hydroxylases (Pihlajaniemi *et al.* 1987, Helaakoski *et al.* 1994, Annunen *et al.* 1997) in spite of the obvious similarities in catalytical properties between these two enzymes. Cloning and nucleotide sequencing were subsequently reported based on human (Hautala *et al.* 1992a, Yeowell *et al.* 1992), rat (Armstrong & Last 1995) and mouse (Ruotsalainen *et al.* 1999). As in the case of chick lysyl hydroxylase, the primary structures of the enzymes from all these sources differed from that of prolyl 4-hydroxylase.

The human lysyl hydroxylase polypeptide (lysyl hydroxylase 1, see 2.2.3) contains 709 amino acids and a signal peptide of 18 residues. The size of its mRNA is 3.2 kb in Northern hybridization. The coding sequence is 76 % identical to the chick sequence at the amino acid level, the C-terminal region being especially well conserved (Hautala *et al.* 1992a). The sequence of the 139 extreme C-terminal amino acids is 94 % identical between these two species (Hautala *et al.* 1992a) and is thought to be of functional significance as it contains many conserved and catalytically important amino acid residues (Pirskanen *et al.* 1996). The identity between the human (Hautala *et al.* 1992a) and rat (Armstrong & Last 1995) and also between the human (Hautala *et al.* 1992a) and mouse (Ruotsalainen *et al.* 1999) polypeptides is 91 %, and that between the chick (Myllylä *et al.* 1991) and rat (Armstrong & Last 1995) is 77 %.

Due to its high affinity for concanavalin A, lysyl hydroxylase was reported to be a glycoprotein (Ryhänen 1976, Turpeenniemi *et al.* 1977, Myllylä *et al.* 1988). Nucleotide sequencing of the chick, rat and human polypeptides revealed indeed that they all contain four potential attachment sites for asparagine-linked oligosaccharide units (Myllylä *et al.* 1991, Hautala *et al.* 1992a, Yeowell *et al.* 1992, Armstrong & Last 1995). A considerable heterogeneity is found in the extent of the glycosylation of the polypeptides, however, and only certain carbohydrate units seem be critical for maximal enzyme activity (Myllylä *et al.* 1988, Pirskanen *et al.* 1996).

2.2.2 Catalytic properties

Lysyl hydroxylase (EC 1.14.11.4) belongs to the family of 2-oxoglutarate-dependent dioxygenases (EC 1.14.11), an oxidoreductase family with over fifteen members that all have similar reaction mechanisms and catalytical properties. Prolyl 4-hydroxylase also belongs to this family (Kivirikko & Pihlajaniemi 1998).

The reaction catalyzed by lysyl hydroxylase requires Fe^{2+}, 2-oxoglutarate, O_2 and ascorbate and produces succinate and CO_2 (see Figure 2). One atom of the oxygen molecule is incorporated into the hydroxyl group formed on a lysine side chain in a peptide substrate and the other atom is incorporated into the 2-oxoglutarate to form succinate when the CO_2 is liberated (Kivirikko & Pihlajaniemi 1998).

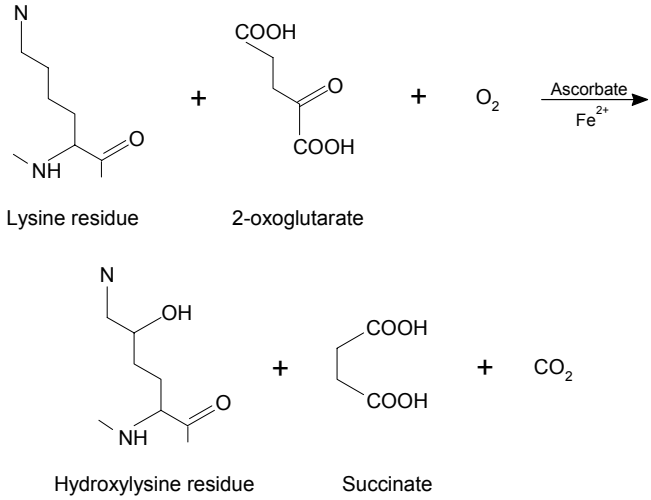


Figure 2. The hydroxylation reaction catalyzed by lysyl hydroxylase (Puistola *et al.* 1980a). The 2-oxoglutarate is decarboxylated during the hydroxylation of a lysine residue in a peptide linkage in the presence of Fe^{2+}, O_2 and ascorbate.

2.2.2.1 Peptide substrates

The minimum sequence requirement of lysyl hydroxylase is the tripeptide X-Lys-Gly, a common sequence in collagen molecules. The enzyme does not hydroxylate free lysine or the tripeptide Lys-Gly-Pro, whereas the tripeptide Ile-Gly-Lys is hydroxylated (Kivirikko *et al.* 1972). Polypeptides with the structure $(X\text{-}Lys\text{-}Gly)_n$ and lysine rich histones that contain X-Lys-Gly triplets and lysine vasopressin with the structure Cys-Tyr-Phe-Gln-Asn-Cys-Pro-Lys-Gly-NH_2 are all hydroxylated by the enxyme (Ryhänen 1975, Kikuchi & Tamiya 1982, Kivirikko *et al.* 1992). Surprisingly, Ryhänen *et al.* (1975) found that

lysyl hydroxylase also acts on arginine-rich histones, which do not contain any X-Lys-Gly sequence but do contain sequences such as X-Lys-Ser, X-Lys-Ala and X-Lys-Thr. The existence of hydroxylysine in the sequences X-Hyl-Ser and X-Hyl-Ala in the short non-collagenous domains of collagens is consistent with these results (Kivirikko & Myllylä 1980).

Critical determinants for the hydroxylation of a lysine residue are the amino acid sequence around the lysine residue, the peptide chain length and the peptide conformation (Kivirikko & Pihlajaniemi 1998). The role of the amino acid sequence around the lysine residue is well demonstrated by experiments performed with collagen from the cuticle of *Ascaris lumbricoides* worms. This collagen has a high content of lysine residues but no hydroxylysine, and it does not act as a substrate for lysyl hydroxylase *in vitro* (Kivirikko *et al.* 1972). It has also been found that a 99-amino-acid peptide fragment from the α1 chain of rat skin type I collagen, which contains several Ala-Lys-Gly sequences, is a very poor substrate for the enzyme (Kivirikko *et al.* 1973). Nevertheless, other collagens contain many hydroxylated Ala-Lys-Gly sequences. These findings suggest that some amino acid sequences around the lysine residue may inhibit hydroxylation (Kivirikko *et al.* 1992).

Experiments with peptide substrates of various sizes have shown that longer peptides have lower K_m values than shorter ones, whereas the reaction rate (V_{max}) is the same (Kivirikko *et al.* 1972). The K_m values for $(Ile\text{-}Lys\text{-}Gly)_2$, $(Ile\text{-}Lys\text{-}Gly)_3$-Phe and $(Ile\text{-}Lys\text{-}Gly)_5$-Phe, for example, are 4, 1 and 0.14 mM, respectively (Kivirikko *et al.* 1972, Kikuchi & Tamiya 1982).

The native triple-helical conformation of collagenous peptides entirely prevents lysine hydroxylation, and thus a non-triple-helical conformation is an absolute requirement (Kivirikko *et al.* 1973, Ryhänen & Kivirikko 1974b, Kivirikko *et al.* 1992). One typical feature of the conformation of lysyl hydroxylase substrates seems to be a folded "bent" structure like a β turn (Jiang & Ananthanarayanan 1991, Ananthanarayanan *et al.* 1992).

2.2.2.2 Cosubstrates and inhibitors

Lysyl hydroxylase requires Fe^{2+}, 2-oxoglutarate, O_2 and ascorbate (for reviews, see Kivirikko *et al.* 1992, Kivirikko & Pihlajaniemi 1998). The presence of Fe^{2+} is a specific requirement for the hydroxylation reaction, whereas other divalent cations such as Zn^{2+}, Cd^{2+}, Co^{2+}, Ca^{2+} and Cu^{2+} have an inhibitory effect (Ryhänen 1976). The K_m value of Fe^{2+} for human lysyl hydroxylase, 2 μM, is virtually identical to that for prolyl 4-hydroxylase, suggesting the presence of a similar Fe^{2+} binding site (Helaakoski *et al.* 1995, Pirskanen *et al.* 1996). The Fe^{2+} is probably located in a hydrophobic pocket and bound to these collagen hydroxylases by three side chains (Hanauske-Abel & Günzler 1982, Hanauske-Abel 1991, Roach *et al.* 1995). Modification of histidine residues by diethyl pyrocarbonate has been shown to inactivate both collagen hydroxylases, thus suggesting a function for histidine residues at the catalytic sites, probably at the Fe^{2+} binding sites (Myllylä *et al.* 1992). A search for conserved residues within the sequences of several 2-oxoglutarate dioxygenases and a related enzyme, isopenicillin N synthase, indicated weak

homology within two motifs containing histidine, termed His-1 and His-2, which are located about 50-70 amino acids apart towards the C-terminal regions of the enzymes. Site-directed mutagenesis of prolyl 4-hydroxylase demonstrated that conversion of either of the two conserved histidines in these motifs to serine inactivates the enzyme completely (Lamberg *et al.* 1995). Determination of the crystal structure of isopenicillin N-synthase indicated that two of the Fe^{2+} binding ligands are histidines present in the His-1 and His-2 motifs but the third ligand is an aspartate residue located in position +2 with respect to the histidine present in the His-1 motif (Roach *et al.* 1995). The corresponding amino acids in the lysyl hydroxylase sequence are His638, Asp640 and His690 (Hautala *et al.* 1992a, Myllylä *et al.* 1992, Pirskanen *et al.* 1996). Individual mutation of these amino acids to serine or alanine was shown to inactivate the human enzyme completely, suggesting that these three amino acids provide the ligands required for binding of the Fe^{2+} atom to its catalytic site (Pirskanen *et al.* 1996).

2-Oxogutarate is a highly specific requirement for the lysine hydroxylation reaction (Prockop *et al.* 1966, Kivirikko & Prockop 1967b). Although all the other kinetic constants of the cosubstrates for the lysyl and prolyl hydroxylases are practically identical, there is a difference in the K_m values for 2-oxoglutarate (Helaakoski *et al.* 1995, Pirskanen *et al.* 1996), that for human lysyl hydroxylase, 100 μM, being about four times higher than those for the human prolyl 4-hydroxylases (12-22 μM) (Helaakoski *et al.* 1995, Pirskanen *et al.* 1996), suggesting some differences in the structures of their 2-oxoglutarate binding sites (Kivirikko 1993, Kivirikko & Pihlajaniemi 1998). Also, the inhibition constants of 23 compounds with structures analogous to that of 2-oxoglutarate are higher for lysyl hydroxylase than for prolyl 4-hydroxylase (Majamaa *et al.* 1985). 2-Oxoadipinate was the only 2-oxoglutarate analogue studied which could act as a cofactor in the lysine hydroxylation reaction, but the V_{max} obtained with it was markedly lower than with 2-oxoglutarate (Majamaa *et al.* 1984, Majamaa *et al.* 1985).

The 2-oxoglutarate binding site of lysyl hydroxylase can probably be divided into two main subsites (Hanauske-Abel & Günzler 1982, Majamaa *et al.* 1984, Hanauske-Abel 1991). Subsite I is assumed to consist of a positively charged side-chain of the enzyme which binds the C-5 carboxyl group of 2-oxoglutarate, while subsite II consists of two *cis*-positioned coordination sites of the enzyme-bound Fe^{2+} and is chelated by the C1-C2 moiety (Hanauske-Abel & Günzler 1982, Hanauske-Abel 1991, Kivirikko & Pihlajaniemi 1998). Recent site-directed mutagenesis studies on prolyl 4-hydroxylase have demonstrated that the residue forming subsite I in the human enzyme is a lysine residue in position +10 with respect to the Fe^{2+} binding histidine present in the His-2 motif (Myllyharju & Kivirikko 1997).

Molecular oxygen is assumed to be bound to the Fe^{2+} end-on in an axillar position at the catalytic sites of the collagen hydroxylases (Hanauske-Abel & Günzler 1982, Hanauske-Abel 1991). The K_m values of oxygen for both enzymes are also quite similar, 40-50 μM (Myllylä *et al.* 1977, Puistola *et al.* 1980a, Turpeenniemi-Hujanen *et al.* 1981).

Ascorbate is a highly specific requirement for the lysine hydroxylation reaction, although it can be partly replaced by dithiothreitol or cysteine, but with a dramatically slower reaction rate (Puistola *et al.* 1980a). The ascorbate binding site appears to consist of two *cis*-positioned coordination sites of the enzyme-bound iron, and seems to ressemble the binding site of 2-oxoglutarate (Majamaa *et al.* 1986, Kivirikko & Pihlajaniemi 1998). The K_m values of lysyl and prolyl hydroxylases for ascorbate are

essentially identical, about 350 μM, suggesting a similar structure in the binding site (Helaakoski *et al.* 1995, Pirskanen *et al.* 1996).

Minoxidil, a piperidinopyrimidine nitrooxide that has an amino group on each side of the nitroxide oxygen, is an antihypertensive drug which also suppresses fibroblast proliferation and inhibits lysyl hydroxylase synthesis at the cell level (Murad & Pinnell 1987), its enzyme-suppressing effect apparently being dependent on one of the two amino groups (Murad *et al.* 1992). The drug reduces the steady-state mRNA level of lysyl hydroxylase (Hautala *et al.* 1992b, Yeowell *et al.* 1992) and also the amount of lysyl hydroxylase protein (Hautala *et al.* 1992b), but it has no effect on lysyl hydroxylase activity when added to the enzyme reaction *in vitro* (Murad & Pinnell 1987).

It has been reported that maximal lysyl hydroxylase activity is not observed *in vitro* even in the presence of saturating concentrations of cofactors unless bovine serum albumin, catalase and dithiothreitol are added to the reaction mixture. ρ-Mercuribenzoate is an inhibitor of the enzyme, but its effect can be partly reversed by the addition of dithiothreitol, indicating that free –SH groups are required for enzyme activity. Bovine serum albumin is likely to act in part by a ”protein effect” and in part by an effect related to the presence of many free thiol groups in this protein. Catalase destroys peroxide and acts through a non-specific protein effect (Kivirikko *et al.* 1992).

2.2.2.3 Reaction mechanism

Extensive kinetic studies indicate that the reaction mechanisms of the lysyl and prolyl hydroxylases are quite similar (Myllylä *et al.* 1977, Tuderman *et al.* 1977, Puistola *et al.* 1980a, Puistola *et al.* 1980b). The hydroxylation reaction occurs by an ordered binding of Fe^{2+}, 2-oxoglutarate, O_2 and the peptide substrate to the enzyme followed by an ordered release of the hydroxylated peptide, CO_2, succinate and Fe^{2+}, the Fe^{2+} not being released between catalytic cycles. One atom of the O_2 molecule is incorporated into the succinate and the other into the hydroxy group formed on the lysine or proline residue (for a reviews, see Kivirikko & Pihlajaniemi 1998).

The data concerning the reaction mechanism and structural features of the catalytic site of lysyl hydroxylase are mainly based on kinetic studies (Myllylä *et al.* 1977) and theoretical considerations (Hanauske-Abel & Günzler 1982, Hanauske-Abel 1991), and also on site-directed mutagenesis studies on both these hydroxylases (Lamberg *et al.* 1995, Pirskanen *et al.* 1996, Myllyharju & Kivirikko 1997) and observations with respect to isopenicillin N-synthase (Roach *et al.* 1995, Roach *et al.* 1997). The catalytic cycle of lysyl hydroxylase can be divided into two halves, generation of the reactive hydroxylating species and its subsequent utilization for hydroxylysine formation (Hanauske-Abel & Günzler 1982, Hanauske-Abel 1991). Molecular oxygen is assumed to be bound end-on in an axial position to the Fe^{2+}, which in turn is assumed to be located in a hydrophobic pocket at the catalytic site (Hanauske-Abel 1991). This produces a species characterized by doubly occupied orbitals, the dioxygen unit, one of its electron-rich orbitals being directed to the electron-depleted orbital at the C-2 position in the 2-oxoglutarate bound to the iron atom (Figure 3A). The C-2 then undergoes rehybridization

from its sp^2 hybridized planar oxo structure to an sp^3 hybridized tetrahedral transition state and forms a covalent bond with the non-coordinated atom of the dioxygen unit (Figure 3B). This weakens both the C-C bond in the 2-oxoglutarate and the O-O bond in the dioxygen. Decarboxylation will take place at the same time as cleavage of the O-O bond, and the original C-2 of the 2-oxoglutarate, which has become the C-1 of succinate, will return to the sp^2 hybridization. Simultaneously, a highly reactive ferryl ion is formed (Figure 3C) which hydroxylates a lysine residue in the peptide substrate in the second half of the reaction. This renews the enzyme-bound Fe^{2+} and concludes the catalytic cycle (Hanauske-Abel & Günzler 1982, Hanauske-Abel 1991).

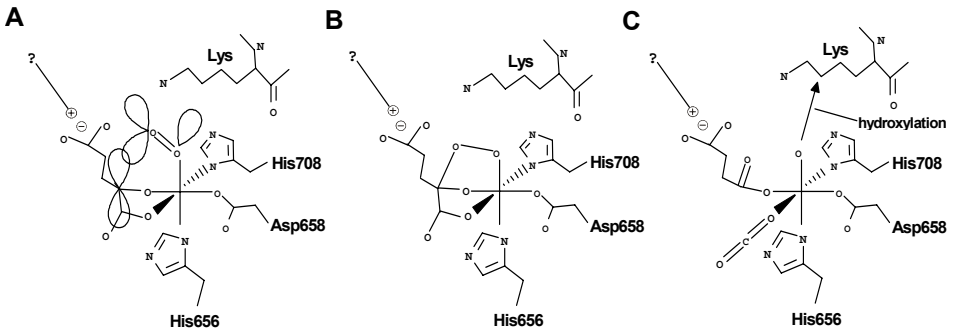


Figure 3. Schematic representation of the first half of the lysyl hydroxylase reaction at its catalytic site. The Fe^{2+} is coordinated with the enzyme by three residues, His656, Asp 658 and His708. The 2-oxoglutarate binding site consists of two main subsites: subsite I, which binds the C-5 carboxyl group of 2-oxoglutarate (marked with a question-mark) and subsite II, which consists of two *cis*-positioned equatorial coordination sites of enzyme-bound Fe^{2+} and is chelated by the C-1 carboxyl and C-2 oxo functions. Molecular oxygen is assumed to be bound end-on in an axial position, producing a dioxygen unit. (A) One of the electron-rich orbitals of the dioxygen is directed to the electron-depleted orbital at the C-2 position in the 2-oxoglutarate bound to the iron. (B) Nucleophilic attack on C-2 generates a tetrahedral intermediate, with loss of the double bond in the dioxygen unit and of double bond characteristics in the oxo-acid moiety. (C) Elimination of CO_2 coincides with the formation of succinate and a ferryl ion, which hydroxylates a lysine residue in the peptide substrate in the second half of the reaction. Modified from a figure presented by Kivirikko and Myllyharju (1998).

Ascorbate is not consumed stoichiometrically in the lysine hydroxylation reaction, and the enzyme can complete numerous catalytic cycles at a maximal rate in its absence (de Jong & Kemp 1984, Myllylä *et al.* 1984). The main function of ascorbate is to act as an alternative oxygen acceptor in the uncoupled decarboxylation of 2-oxoglutarate, a reaction occurring without subsequent hydroxylation of the peptide substrate. Ascorbate is consumed stoichiometrically in this reaction, which is needed for reactivation of the enzyme-bound Fe^{2+} (Myllylä *et al.* 1984). The rate of the uncoupled reaction taking place in the absence of the peptide substrate is only 1-4 % of that of the complete reaction. An

uncoupled reaction cycle can occur occasionally even in the presence of a saturating concentration of the peptide substrate, and certain non-hydroxylatable peptides even increase the rate of the uncoupled reaction (Counts *et al.* 1978, Myllylä *et al.* 1984).

The lysyl hydroxylase α_2 dimer and prolyl 4-hydroxylase $\alpha_2\beta_2$ tetramer have two catalytic sites (Kivirikko *et al.* 1992, Kivirikko & Pihlajaniemi 1998), and it has been proposed that the two substrate binding sites bind simultaneously to the same polypeptide substrate with several hydroxylatable sites (de Waal & de Jong 1988, de Jong *et al.* 1991). When hydroxylation occurs at one site and the polypeptide chain is released, it is still bound at the other site. This agrees with the finding that longer peptides have lower K_m values than shorter ones (Kivirikko *et al.* 1972). However, *C.elegans* prolyl 4-hydroxylase, an $\alpha\beta$ dimer with only one catalytic site, has K_m values for long peptides that are identical to that of the vertebrate prolyl 4-hydroxylase $\alpha_2\beta_2$ tetramer, a finding that is not consistent with this proposal (Veijola *et al.* 1996, Kivirikko & Pihlajaniemi 1998).

2.2.3 Isoenzymes

Isoenzymes are distinct forms of the same catalytic activity, differing in their physical and chemical properties. They can be present in different tissues, cell types, subcellular compartments or developmental stages of the same organism, or in different organisms.
Varying opinions have been expressed about the possible existence of lysyl hydroxylase isoenzymes over the years. Many findings have supported the view that such isoenzymes do exist. Large differences are found in the extent of lysine hydroxylation between collagen types and even within the same collagen type from different tissues and at different developmental stages (Barnes *et al.* 1974). Also, the deficiency in hydroxylysine in collagens of patients with Ehlers-Danlos syndrome type VI shows great variation between tissues and collagen types (Krane *et al.* 1972, Ihme *et al.* 1984). Ihme *et al.* (1984) reported a patient with this syndrome whose type I collagen was deficient in hydroxylysine in its triple-helical regions in skin and bone, but normally hydroxylated in tendon and lung. Moreover, the triple-helical domain of type V collagen from bone was normally hydroxylated, in contrast to type I collagen from the same origin, suggesting the presence of multiple isoenzymes.

Two novel isoenzymes, termed lysyl hydroxylase 2 and lysyl hydroxylase 3, have recently been cloned and characterized from human (Valtavaara *et al.* 1997, Valtavaara *et al.* 1998, paper II of the present study) and mouse (Ruotsalainen *et al.* 1999). Correspondingly the previously known isoenzyme is now called lysyl hydroxylase 1.

2.2.3.1 Lysyl hydroxylase 2

A cDNA coding for lysyl hydroxylase 2 was first cloned from human kidney and pancreas (Valtavaara et al. 1997). This encodes a polypeptide of 737 amino acids, including a signal peptide of 25 residues. The novel polypeptide shows an overall similarity of 75 % to the human and chick lysyl hydroxylase 1 polypeptides, the highest similarity, over 90 %, being found in the conserved C-terminus. The lysyl hydroxylase 2 polypeptide contains 7 putative N-glycosylation sites and nine conserved cysteine residues, in addition to which the two histidines and one aspartate thought to constitute the Fe^{2+} binding site have been conserved.

The mRNA for lysyl hydroxylase 2 is clearly a product of a different gene from that for lysyl hydroxylase 1, this being located in the region 3q23-q24 (Szpirer *et al.* 1997) while the lysyl hydroxylase 1 gene is located on chromosome 1 (see 2.2.4) (Hautala *et al.* 1992a). A new, alternatively spliced form of the human lysyl hydroxylase 2 cDNA has recently been cloned from human skin fibroblasts (Yeowell & Walker 1999a) and human kidney (Valtavaara, M., Risteli, M., Ruotsalainen H., Wang, C. and Myllylä, R., unpublished data). This novel lysyl hydroxylase 2b cDNA encodes a protein of 758 amino acids, of which 21 amino acids are encoded by a new exon, termed exon 13A, which is located between those coded by exons 13 and 14 in the originally described lysyl hydroxylase 2 cDNA, now designated lysyl hydroxylase 2a cDNA. Analysis of genomic DNA from different tissues has revealed that both transcripts, those of lysyl hydroxylase 2a and 2b, are generated from the same lysyl hydroxylase 2 gene (Yeowell & Walker 1999a).

Expression of the mRNAs for lysyl hydroxylase 2a and 2b has been studied by Northern hybridization and PCR analysis (Yeowell & Walker 1999a, Valtavaara, M., Risteli, M., Ruotsalainen H., Wang, C. and Myllylä, R., unpublished data). The lysyl hydroxylase 2a mRNA, of size 4.2 kb, is abundantly expressed in human heart, muscle, placenta and pancreas, whereas that of lysyl hydroxylase 2b, of size 4.4 kb, is well expressed in human skeletal muscle and heart but virtually undetectable in human placenta and pancreas (Valtavaara, M., Risteli, M., Ruotsalainen H., Wang, C. and Myllylä, R., unpublished data). PCR amplification studies indicated that the lysyl hydroxylase 2b mRNA was expressed as the major lysyl hydroxylase 2 transcript form in all tissues studied except kidney and spleen. Preliminary data further suggest that this alternative splicing event is not developmentally regulated (Yeowell & Walker 1999a).

Comparison of the catalytic properties of lysyl hydroxylases 2a and 2b after expression as recombinant proteins in insect cells has suggested a few differences between them (Valtavaara, M., Risteli, M., Ruotsalainen H., Wang, C. and Myllylä, R., unpublished data). The K_m of lysyl hydroxylase 2b for a synthetic peptide substrate has been reported to be three times higher than that of lysyl hydroxylase 2a, and that for ascorbate twice as high (see Table 4 in section 6.2). Recent measurements of mRNA levels and immunoblotting experiments have suggested that lysyl hydroxylase 2 is coregulated together with total collagen synthesis but not with specific collagen types (Wang *et al.* 2000).

2.2.3.2 Telopeptide lysyl hydroxylase(s)

The telopeptides are short non-triple-helical regions located at the ends of the collagen α chains, each α chain having an N-terminal and a C-terminal telopeptide region (Hulmes 1992, Kielty *et al.* 1993). Many studies have demonstrated the possible existence of one or more telopeptide-specific lysyl hydroxylases, which could also be responsible for the variations found in collagen crosslinking. Conclusive evidence is nevertheless lacking.

A detailed study of the extent of lysine hydroxylation in the N-terminal telopeptides of type I collagen as a function of age in different chick tissues indicated that variations at these sites were unrelated to variations in the corresponding triple-helical domains (Barnes *et al.* 1974). Moreover, a highly purified chick lysyl hydroxylase preparation was reported to fail to hydroxylate lysine residues in the telopeptide regions of type I collagen *in vitro* even though it did hydroxylate lysine residues in the triple-helical domain of collagen I (Barnes *et al.* 1974, Royce & Barnes 1985). It also seems that a specific increase in the hydroxylation of lysine residues of type I collagen in hypertrophic tendon occurs only in the N- and C-terminal telopeptide regions, whereas the extent of hydroxylation of lysine residues in the triple-helical regions is not affected (Gerriets *et al.* 1993). Similar results have been reported in rats with experimentally induced pulmonary fibrosis (Gerriets *et al.* 1996); suggesting that hydroxylation of telopeptide lysine residues may be under separate enzymatic control.

Bank *et al.* (1999) recently reported that lysine residues within the telopeptides of type I collagen in bone are underhydroxylated in patients with Bruck syndrome, an autosomal recessive disease with characteristics of fragile bones, joint contractures, scoliosis and osteoporosis (Brenner *et al.* 1993), leading to aberrant crosslinking. Lysine residues in the triple-helical region are normally modified in these patients, indicating that the hydroxylation defect is limited to the telopeptides. On the other hand, cartilage and ligament show an unaltered telopeptide hydroxylysine content (Bank *et al.* 1999). These results together with earlier observations thus provide evidence for the existence of a telopeptide lysyl hydroxylase with tissue-specific forms. According to Bank et al. (1999), the lysyl hydroxylase 1 and 2 genes can be excluded as candidate genes for Bruck syndrome because they are located on different chromosomes from its gene, which has been mapped to chromosome 17 (Bank *et al.* 1999).

Results partly differing from those of Bank *et al.* (1999) have been reported by Uzawa *et al.* (1999), who studied the expression patterns of different lysyl hydroxylase isoenzymes during the *in vitro* differentiation of human bone marrow stromal cells and skin fibroblasts. A 6-fold increase in the level of lysyl hydroxylase 2 mRNA was found in the course of progression from undifferentiated to fully differentiated bone marrow stromal cells, whereas no such increase was observed in skin fibroblasts. Since the increase in lysyl hydroxylase 2 mRNA level coincided with the onset of matrix mineralization and the increase in lysine hydroxylation of the non-triple-helical region in the α chains of type I collagen, the authors suggested that lysyl hydroxylase 2 may be involved in processing of the telopeptide domains of bone type I collagen (Uzawa *et al.* 1999).

2.2.4 Structure of the lysyl hydroxylase 1 gene

The lysyl hydroxylase 1 gene is located in the region 1p36.2→1p36.3 on human chromosome 1 and is not physically linked to the genes for prolyl 4-hydroxylase (Pajunen *et al.* 1988, Popescu *et al.* 1988, Pajunen *et al.* 1991, Hautala *et al.* 1992a, Helaakoski *et al.* 1994). It is 40 kb in size and contains 19 exons (Heikkinen *et al.* 1994). The 5´ flanking region and the first exon are rich in G+C nucleotides, the G+C content of the 5´ untranslated region of the mRNA being 75 % while that of the coding region is only 59 %. The promoter region of the gene lacks a TATAA box, which is a typical promoter element in highly regulated genes. Results obtained by S1 nuclease mapping and primer extension show the presence of multiple transcription initiation sites, the size of the 5´ untranslated region varying from 49 to 136 bp. The main transcription initiation sites are located 55-65 bp upstream of the ATG codon. A CCAAT sequence, which is a possible binding site for the CTF/NF transcription factor is present 370 bp upstream of the main transcription initiation site. This sequence is usually located 80 bp upstream of the transcription initiation site, but it can also function at different distances. The gene contains GGGCGG sequences, which are potential Sp1 transcription factor binding sites and typical of genes with multiple transcription initiation sites. Two such Sp1 binding sites are present in the coding strand of the promoter region, one potential Sp1 binding site also being found in the non-coding strand and one at the 5´ end of the first intron. Overall, the 5´flanking region of the human lysyl hydroxylase 1 gene ressembles the promoters of many housekeeping genes (Heikkinen *et al.* 1994).

The 19 exons of the human lysyl hydroxylase 1 gene represent only 7 % of its total size (see Table 1). The extreme 3´ exon is the largest one, being 887 bp in length and containing a 732 bp 3´ untranslated sequence and the polyadenylation signal, while exons 1 to 18 vary in size from 64 to 164 bp. Exons 2, 5 and 7 begin with the second base of a codon, exons 4,11,12 and 13 with the third base and all the others with the first base. The consensus sequence CAG-exon-GT(G/A) is found at the boundaries throughout, except for exon 13, as intron 13 begins with a sequence GTC instead of GT(G/A) (Heikkinen *et al.* 1994).

The sizes of the introns in the lysyl hydroxylase 1 gene vary more than those of the exons. The first intron, 12 500 bp, is by far the largest and intron 12, 350 bp, the smallest. The average size of an intron is nevertheless over 1 kb. Introns 9,15 and 16 have been sequenced completely, because of their involvement in gene rearrangements in patients with Ehlers-Danlos syndrome type VI (see section 2.3) (Heikkinen *et al.* 1994).

Table 1. Exon-intron structure of the human lysyl hydroxylase 1 gene

Exon/intron number	Exon size(bp)	Intron size(bp)
1	49-136	12,500
2	92	1,800
3	134	480
4	164	2,200
5	113	2,100
6	64	2,000
7	98	820
8	102	540
9	132	1,987
10	122	2,600
11	105	650
12	126	350
13	142	780
14	114	670
15	66	665
16	105	3,578
17	147	2,100
18	126	1,500
19	887	

(Heikkinen *et al.* 1994, OMIM 2000)

2.2.4.1 Alu sequences

Short interspersed elements (SINEs) are ubiquitous repetitive elements occurring in mammalian genomes. The most common human SINEs are 0.5-1.1 $\times 10^6$ Alu sequences, which represent about 6-13 % of the human genomic DNA. These occur at an average rate of one every 3-6 kb, although their distribution within the human genome is not uniform (Schmid & Maraia 1992, Makalowski *et al.* 1994, Mighell *et al.* 1997, Schmid 1998). Alu sequences are characterically found in the intronic regions of protein-coding genes in primates, and are typically located in clusters separated by up to a few hundred basepairs of non-Alu DNA. Consensus Alu sequences are approximately 280 bp in length and consist of two homologous but distinct monomers rich in G+C nucleotides, forming right and left subunits that are connected by an adenine-rich region. Alu sequences typically end in a 3´ A–rich region resembling a polyadenyl tail (Makalowski *et al.* 1994, Mighell *et al.* 1997).

It has been postulated that Alu sequences are inserted into the human genome by a retrotransposition mechanism involving a single-stranded RNA intermediate generated by RNA polymerase III. The progenitor of the human Alu elements is probably the 7SL RNA, an essential component of the signal recognition particle (SPR), a nucleoprotein

that targets certain proteins to the endoplasmic reticulum (Mighell *et al.* 1997). The rate of Alu amplification seems to have reached a maximum between 35 and 60 million years ago, and they are currently amplifying only at 1 % of the maximum rate (Deininger & Batzer 1999). Alu repeats may be further grouped into subfamilies by reference to the great variation found in their consensus sequences. At least 12 subfamilies have now been characterized, the Alu-J subfamily being probably the oldest one and the Alu-Y subfamily the youngest (Makalowski *et al.* 1994, Batzer *et al.* 1996). Alu sequences are involved in most human gene rearrengements that occur by homologous recombination, thus leading to genetic disorders such as haemophilia, but also to the regulatory variation needed in the course of evolution (Rüdiger *et al.* 1995, Britten 1996, Brosius 1999, Deininger & Batzer 1999).

At least introns 9 and 16 of the lysyl hydroxylase 1 gene contain Alu sequences. Intron 9, 2.0 kb in length, contains five of them, and intron 16, 3.6 kb in length, eight. The Alu repeats present in these introns seem to belong to the Alu-Sp, Alu-Sx or Alu-Sc subfamilies, or else they represent truncated forms of Alu repeats (Heikkinen *et al.* 1994). The large number of Alu repeats generates a marked homology between these two introns, the longest identical sequence being 44 nucleotides in length. This appears to explain the large gene duplication which is one of the most common mutation found in families with Ehlers-Danlos syndrome type VI (see section 2.3) (Hautala *et al.* 1993, Heikkinen *et al.* 1994, Pousi *et al.* 1994, Heikkinen *et al.* 1997).

2.2.5 Intracellular site of lysine hydroxylation

The lysine hydroxylation reaction occurs as an intracellular event during collagen biosynthesis. While nascent procollagen polypeptide chains are being synthesized on the ribozomes, the hydroxylation of lysine and proline residues and the glycosylation of hydroxylysine residues occur as cotranslational modifications and continue as post-translational modifications within the cisternae of the rough endoplasmic reticulum, until collagen triple helix formation prevents any further modification (Kivirikko & Myllylä 1984, Kivirikko 1993, Kivirikko & Pihlajaniemi 1998).

Results obtained in the 1970´s by assays of lysyl hydroxylase activity after subcellular fractionation of embryonic chick tendon and sternal cartilage cells indeed showed the highest enzyme activity to occur in the microsomal fraction. Subsequent fractionation of this microsomal fraction by discontinuous gradient centrifugation revealed that the enzyme is located almost exclusively in the rough endoplasmic reticulum fraction of both tendon and cartilage cells (Harwood *et al.* 1974). Lysyl hydroxylase was also shown by a variety of techniques to reside within the cisternae of the rough endoplasmic reticulum (Guzman *et al.* 1976, Ryhänen 1976, Peterkofsky & Assad 1979). Approximately 30 % of the total lysyl hydroxylase activity was found to exist freely within the cisternae, the remainder being associated with the membrane through ionic bonds (Peterkofsky & Assad 1979).

Cloning and sequencing of chick and human lysyl hydroxylase cDNAs indicated, however, that they lacked any ER-specific retention motifs in their primary structure

(Myllylä *et al.* 1991, Hautala *et al.* 1992a). Although the enzyme is thought to be located in the ER, it contains no KDEL–like sequence or double-lysine motif, signals necessary for the retention of a number of proteins in the lumen of the ER or its membrane (for reviews, see Pelham 1992, Pelham 1996, Teasdale & Jackson 1996). A 40 amino acid C-terminal peptide segment of lysyl hydroxylase has recently been identified as the determinant necessary for its membrane association and localization in the ER, as this segment seems to be able to convert cathepsin D, normally a soluble lysosomal protease, into a membrane-associated protein (Suokas *et al.* 2000). The immunofluorescence studies of Kellokumpu *et al.* (1994) indicated that lysyl hydroxylase is co-located with protein disulphide isomerase, a protein with the KDEL retention signal, and seems to reside in the same subcompartment of the ER in which the proteins containing KDEL reside. Moreover, the enzyme appears to be present only in association with the ER membrane, as a luminally-oriented peripheral membrane protein binding to it via weak electrostatic interactions. Stress treatment with tunicamycin seems to increase the enzyme activity and the amount of the enzyme protein (Kellokumpu *et al.* 1994).

2.3 Ehlers-Danlos syndrome type VI

Ehlers-Danlos syndrome is a group of connective tissue disorders with heterogeneous clinical characteristics such as hyperextensive skin, excessive bruising, muscular hypotonia and joint hypermobility. This autosomally recessively inherited syndrome has been divided traditionally into 10 subtypes (Krane *et al.* 1972, Kivirikko 1993, Steinmann *et al.* 1993, Yeowell & Pinnell 1993, Byers 1994). A more recent, simplified classification divides it into six major types, based primarily on the cause of each type (Beighton *et al.* 1998).

Ehlers-Danlos syndrome type VI, or the kyphoscoliotic type, was the first subtype in which the biochemical abnormality was characterized (Krane *et al.* 1972, Sussman *et al.* 1974). It is very uncommon, and only approximately two dozen patients have been identified (Byers 1994). These have the general characteristics of the Ehlers-Danlos syndrome but also some specific properties of this subtype, such as kyphoscoliosis and ocular manifestations: spontaneous retinal detachments, rupture of the globe after minor trauma and microcornea (Beighton 1970, Krane *et al.* 1972, Sussman *et al.* 1974, Steinmann *et al.* 1993). The biochemical abnormality involved in subtype VI in most but not all cases is a deficiency in lysyl hydroxylase activity (Eyre & Glimcher 1972, Krane *et al.* 1972, Pinnell *et al.* 1972, Quinn & Krane 1976, Beighton P 1993, Steinmann *et al.* 1993, Byers 1994, Prockop & Kivirikko 1995). The type VI variant has thus been divided biochemically into subtypes VIA and VIB, having similar clinical characteristics (McKusick 1992) but differing in that patients with subtype VIB have a normal level of lysyl hydroxylase activity while those with subtype VIA show clearly decreased activity levels, less than 25 % of normal when measured in skin fibroblasts (Krane *et al.* 1972, Judisch *et al.* 1976, McKusick 1983).

Molecular cloning of human lysyl hydroxylase 1 (Hautala *et al.* 1992a, Yeowell *et al.* 1992) and its gene (Heikkinen *et al.* 1994) has made it possible to characterize in detail

the mutations responsible for Ehlers-Danlos syndrome type VI. At least sixteen mutations have been characterized in patients with this disease, the defects in the mutant alleles being either heterozygous or compound homozygous. The first mutation reported in one family with two affected siblings was a homozygous single basepair substitution in the lysyl hydroxylase 1 gene, converting the codon of arginine (Arg319) to a translation termination codon and leading to an almost complete absence of lysyl hydroxylate activity in skin fibroblasts. The healthy parents and two of their three healthy siblings were heterozygous carriers of the mutation (Hyland *et al.* 1992). A common mutation, with a prevalence of 20 % among families studied, is a large homozygous duplication involving seven exons in the lysyl hydroxylase 1 gene (Hautala *et al.* 1993, Heikkinen *et al.* 1997) caused by an Alu-Alu recombination (Pousi *et al.* 1994). Pajunen *et al*. (1998) reported a patient with a homozygous insertion of two thymidines at the 5´splice site of intron 9 in the lysyl hydroxylase 1 gene, leading to an in-frame deletion of sequences coded by exon 9 in the mRNA. Cells extracted from the patient possessed only 12 % of the activity detected in unaffected cells (Pajunen *et al.* 1998). A homozygous single base substitution in exon 14, generating a premature termination codon (Tyr511X), was identified by Walker *et al* (1999), and a homozygous insertion of a single C nucleotide in a group of 4 C nucleotides in exon 16 has recently been described in a North American patient (Yeowell *et al.* 2000a).

The first example of a compound heterozygous patient was reported by Ha *et al* (1994). This patient had a triple base deletion which resulted in the loss of residue Glu532 in one allele and a single base substitution which converted Gly678 to arginine in the other allele. Lysyl hydroxylase activity had decreased to 24 % (Ha *et al.* 1994). Another patient has also been identified who is compound heterozygous for the same mutation of a premature stop codon (Tyr511X) in the exon 14 sequence in one allele and a splice site-mediated exon 5 deletion generating a premature stop codon in the other (Yeowell & Walker 1997). Recently, a heterozygous patient with a deletion of the penultimate adenosine from the 3´end of intron 15 followed by skipping of the exon 16 sequence in one allele has been identified by Pousi et al. (1998). The other allele had a deletion of exon 17 generated by an Alu-Alu recombination (Pousi *et al.* 1998). A novel point mutation resulting in the replacement of tryptophan (Trp612) by cysteine in the highly conserved C-terminal region has been reported in one allele of a new patient studied, the other being a functionless null allele (Brinckmann *et al.* 1998). A null-mutant lysyl hydroxylase gene has also been reported in a heterozygous British patient, with one nucleotide deletion in the acceptor splice site of intron 4 in one allele and an insertion of a C nucleotide in exon 2 in the other (Heikkinen *et al.* 1999), and the same study pointed to several exon-deleted splicing variants of lysyl hydroxylase mRNA not only in the affected cells but also in small amounts in normal human skin fibroblasts (Heikkinen *et al.* 1999). The previously reported non-sense mutation (Tyr511X) in exon 14, resulting in a reduction in the mRNA level as well as a skipping of exon 14 sequences in the mRNA, has now been described in one allele of a British patient, the second allele being operationally null (Pousi *et al.* 2000). A 15-bp deletion in exon 11 in one lysyl hydroxylase allele coding for amino acids 367-371 has been reported in two unrelated compound heterozygous patients with the syndrome. The mutations in their other alleles were a Cys1119Thr change in exon 10 and a predicted Gln49X in exon 2. It seems likely that the loss of cysteine 369 in the deleted sequence may have contributed to the

diminished enzyme activity, as individual mutation of Cys369 by site-directed mutagenesis followed by expression in a baculovirus system virtually eliminated lysyl hydroxylase activity (Yeowell *et al.* 2000b). Three novel point mutations that code for premature termination codons, Gln327X, Tyr142X and Arg670X, have also been identified recently in compound heterozygous patients. The Gln327X mutation in exon 10 was also reported as a homozygous mutation in another patient (Yeowell *et al.* 2000a).

Of the sixteen lysyl hydroxylase gene mutations characterized so far, four have been shown to occur in more than one unrelated patient. These include the large seven exon duplication (Hautala *et al.* 1993, Pousi *et al.* 1994, Heikkinen *et al.* 1997), the Tyr511X mutation in exon 14 (Yeowell & Walker 1997, Walker *et al.* 1999, Yeowell & Walker 1999b, Pousi *et al.* 2000), the 15 bp deletion in exon 11 (Yeowell *et al.* 2000b) and the Gln327X mutation in exon 10 (Yeowell *et al.* 2000a). These may therefore be considered common mutations in type VI of the Ehlers-Danlos syndrome.

3 Outline of the present research

Lysyl hydroxylase plays an important role in the biosynthesis of all collagens, by hydroxylating certain lysine residues to hydroxylysine. The resulting hydroxylysine residues are essential for the stability of collagen crosslinks and are also needed for the glycosylation of collagen molecules. The critical role of hydroxylysine residues in collagens is demonstrated by the marked changes in the mechanical properties of various tissues seen in patients with the type VI variant of the Ehlers-Danlos syndrome, a molecular disease caused by a deficiency in lysyl hydroxylase activity.

In spite of extensive research into lysyl hydroxylase and the Ehlers-Danlos syndrome, little was known about the catalytic and structural properties of human lysyl hydroxylase. Expression as a recombinant protein in insect cells by means of a baculovirus system had made it possible to study the amino acid residues that play critical roles at the catalytic site of the enzyme. Residues responsible for the binding of the Fe^{2+} atom to the catalytic site had previously been characterized by site-directed mutagenesis in our laboratory. It was therefore a natural choice to continue this work by studying the binding sites for the other cofactors. The specific goal was:

1. to identify the positively charged amino acid residue that binds the C-5 carboxyl group of 2-oxoglutarate, an essential cofactor for the hydroxylation reaction.

Many previous studies have suggested that lysyl hydroxylase may have isoenzymes, and one such isoenzyme, termed lysyl hydroxylase 2, had recently been cloned by another laboratory. The specific goals were thus:

2. to attempt to identify, clone and characterize an additional human lysyl hydroxylase isoenzyme. Subsequent goals were to study expression of the mRNA for this novel isoenzyme in human tissues and to produce a recombinant protein in insect cells to show the function of the cloned polypeptide as a lysyl hydroxylase and to characterize its catalytic properties.
3. to study the exon-intron organization of the novel human lysyl hydroxylase isoenzyme gene and the sequence of its promoter region and to compare these results with those for the gene for isoenzyme1.

4 Materials and methods

More detailed descriptions of the materials and methods are presented in the original papers I-III.

4.1 Site-directed mutagenesis studies of human lysyl hydroxylase 1 (I)

The cDNA for human lysyl hydroxylase 1 (Hautala *et al.* 1992a) was modified by converting the codons for Arg697, Arg700 and Ser705 individually to codons for alanine, and that for Arg700 also to one for lysine. The mutagenesis steps were performed in a pBluescript vector (Stratagene) containing the full-length lysyl hydroxylase 1 cDNA by means of an oligonucleotide-directed *in vitro* system based on the unique site elimination procedure (Pharmacia Biotech Inc.).

4.2 Isolation and characterization of cDNA and genomic clones (II, III)

In order to obtain a probe for cDNA library screening (II), PCR primers F1 and R3 were designed based on the finding of an EST sequence (AA 340 606) and used to obtain a 236-bp product from an adult human kidney λgt-10 cDNA library (Clontech). The purified PCR product was labelled with ^{32}P and used to screen foetal (λgt-11) and adult (λgt-10) human kidney and human placenta (λgt-11) cDNA libraries (Clontech). This resulted in five positive clones being obtained from the two kidney libraries and 16 from the placenta library. Seven of these were isolated and characterized in detail.

To obtain the 5' end of the cDNA, two new probes were amplified by PCR using the human placenta library as a template. These probes were further used for rescreening, as a result of which eight positive clones were obtained, four of which were characterized in detail. As an independent approach, rapid amplification of the 5′ cDNA ends was

performed from pooled human placenta cDNAs (Marathon-Ready cDNA, Clontech) using oligonucleotides generated from existing sequences as primers.

To characterize the structure of the human lysyl hydroxylase 3 gene (III), two P1 genomic clones, 18050 and 18051, were obtained from the P1 Human Library screening services (Genome Systems, Inc.) using a lysyl hydroxylase 3 cDNA specific sequence corresponding to nucleotides 2135-2516 as a probe. The QIAGEN Plasmid Midi Kit was used to isolate the P1 DNA. Sequencing was performed directly on the P1 DNA in the case of most of the introns, but the sizes of the three largest introns (6, 15 and 16) were determined by PCR with a forward primer corresponding to the 3' end of an exon and a reverse primer corresponding to the 5' end of the next exon. The sizes of the PCR products, and thus those of the introns, were further analyzed with a 1 % agarose gel. All the PCR products were also isolated, cloned into pUC18 (Pharmacia) and sequenced with vector and cDNA-specific primers.

4.3 Construction of baculovirus transfer vectors and generation of recombinant viruses (I, II)

After the generation of plasmids containing the mutant lysyl hydroxylase 1 cDNAs (I), these were digested with *Eag*I and *Eco*RI and the resulting fragments further cloned to the *Eag*I-*Eco*RI -digested baculovirus transfer vector pVL1392. All sequences were verified by dideoxynucleotide sequencing (Sanger *et al.* 1977).

Two pairs of sequence-specific primers were designed in order to construct the full-length lysyl hydroxylase 3 cDNA (II), LH3-5'B and LH3-5'N contained an artificial *Bam*HI site and a natural *Nar*I site, and correspondingly LH3-3'X and LH3-3'N contained an artificial *Xba*I site and a natural *Nar*I site (Figure 4). The 5' end of the cDNA (nucleotides 37-1360) was amplified with the first pair of primers and the 3' end

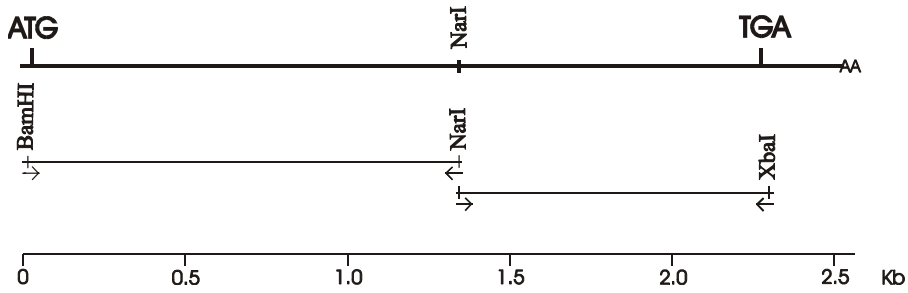


Figure 4. Amplification of the lysyl hydroxylase 3 cDNA in two parts by PCR. The 5´part contained an artificial *Bam*HI site and a natural *Nar*I site and the 3´part a natural *Nar*I site and an artificial *Xba*I site. The arrows indicate the locations of the oligonucleotide primers used in the amplification.

(nucleotides 1361-2317) with the second pair. The 5' PCR product was digested with *Bam*HI and *Nar*I and the 3' product with *Nar*I and *Xba*I, the resulting fragments were further ligated to the *Bam*HI-*Xba*I site of the baculovirus transfer vector pVL1393 (Luckow *et al.* 1993) and the whole construct was sequenced using Abi Prism 377.

Recombinant viruses were generated by culturing *Spodoptera frugiperda* Sf9 cells (Invitrogen) as monolayers in TNM-FH medium (Sigma) supplemented with 10 % foetal bovine serum (Gibco) at 27 °C. The recombinant baculovirus transfer vectors were then cotransfected into Sf9 cells with a modified *Autographa californica* nuclear polyhedrosis virus DNA (BaculoGold, Pharmingen) by calcium phosphate precipitation (Gruenwald & Heitz 1993). The resulting viral pool was collected four days later and amplified twice.

4.4 Expression and analysis of recombinant proteins (I, II)

High Five insect cells (Invitrogen) were cultured as monolayers in TNM-FH medium supplemented with 10 % foetal bovine serum (Life Technologies) at 27 °C, infected at a multiplicity of 5, harvested 24-72 h after infection and washed with a solution of 0.15 M NaCl and 0.02 M phosphate, pH 7.4. They were then homogenized in a solution of 1 % Nonidet P-40, 0.1 M glycine and 0.02 M Tris-HCl, pH 7.8 (termed Nonidet P-40 buffer) and centrifuged at 10.000 x g for 10 min. The insoluble pellets were further homogenized in a 50 % glycerol, 0.6 M NaCl, 1 % Nonidet P-40, 0.1 M glycine, 100 µM dithiothreitol and 0.06 M Tris-HCl buffer, pH 7.8 (glycerol buffer) and then incubated on ice for 30-60 min and centrifuged at 10,000 x g for 20 min. The pellets remaining after centrifugation were further solubilized in 1 % SDS.

Aliquots of the supernatants and remaining SDS pellets were analyzed by 8 % SDS-PAGE under reducing conditions using Coomassie blue staining.

4.5 Enzyme activity assays (I, II)

Lysyl hydroxylase activity was measured by a procedure based on the hydroxylation-coupled decarboxylation of 2-oxo[1-^{14}C]glutarate with 0.75 mg/ml of $(Ile\text{-}Lys\text{-}Gly)_3$ as the peptide substrate (Kivirikko & Myllylä 1982). Pooled soluble fractions of the cell homogenate were used as the source of enzyme. K_m values were determined by varying the concentrations of one substrate in the presence of a fixed concentration of the second substrate in the presence of an excess of the other substrates (Kivirikko & Myllylä 1982). Protein concentrations were determined with a Bio-Rad protein assay kit (Bio-Rad). The expression levels of the wild-type and mutant lysyl hydroxylases (I) were compared by densitometry of the Coomassie blue stained bands in SDS-PAGE using a BioImage instrument (BioImage, Millipore).

4.6 Nucleotide sequencing (I, II, III)

DNA sequencing was performed mostly using dRhodamine terminator or Big Dye Terminator cycle sequencing ready reaction kit (Perkin Elmer) with an Abi Prism 377 automated sequencer (Applied Biosystems). Some sequences were determined manually with a dideoxynucleotide sequencing system (Sanger *et al.* 1977) using a T7 sequencing kit (Pharmacia Biotech).

4.7 The BLAST program and other sequence analyses (I, II, III)

The BLAST (Basic local alignment search tool) program is a fast and readily available tool for searching protein and DNA databases for sequence similarities. It can detect distant sequence relationships and make it easier to distinguish real significance from statistical similarities (Altschul *et al.* 1990, Altschul *et al.* 1994, Heikkinen & Mattila 2000). The BLAST program was used here to search the database of human expressed sequence tags (EST) (Marra *et al.* 1998) with the lysyl hydroxylase 1 cDNA sequence (Hautala *et al.* 1992a) in order to obtain part of the cDNA sequence of human lysyl hydroxylase 3 (II). It was also used during the cDNA library screening to check the new sequence (II) and to detect Alu repeats when analyzing the gene for human lysyl hydroxylase 3 (III).

DNASIS and PROSIS version 6.00 software (Pharmacia Biotech) were used to analyze the nucleotide sequence data after sequencing (I, II, III). Transcription Factor Data Base using the TESS (Transcription Element Search Software on the WWW), Version 3.2, and the MatInspector, Version 2.2, programs were used when searching the consensus sites for the binding of transcription factors in the human lysyl hydroxylase 3 gene (III).

4.8 Isolation of RNA and Northern blot analysis (II)

Total RNA was isolated from cultured adult human skin fibroblasts and human Saos-2 osteosarcoma cells with the RNeasy Midi Kit (Qiagen). Samples of 15 µg of total RNA were separated out electrophoretically on a 1 % agarose gel containing 2 M formaldehyde, transferred to a nitrocellulose membrane and immobilized by baking.

Otherwise Northern blot analysis was carried out using the ready-made human multitissue Northern blots I and III and the foetal blot II (Clontech), each containing 2 µg of poly(A) RNA per sample. The tissues analyzed in these blots included heart, brain, placenta, lung, liver, skeletal muscle, kidney, pancreas, stomach, thyroid, spinal cord, lymph node, trachea, adrenal gland and bone marrow, and also foetal brain, lung, liver and kidney. The hybridization was carried out in ExpressHyb solution (Clontech) under

the stringent conditions recommended by the manufacturer, using a 352-bp ^{32}P-labelled PCR fragment from the 3' end of the novel cDNA as a probe.

4.9 Nuclease S1 assay (III)

In order to localize the transcription initiation site of the lysyl hydroxylase 3 gene, S1 protection experiments were performed as described previously (Pihlajaniemi & Myers 1987, Sambrook *et al.* 1989). A 637-bp PCR fragment (nucleotides –598 to +39 relative to ATG) was used as a probe after 5' end-labelling with T4 polynucleotide kinase (USB Corp.) and [γ-^{32}P]ATP (3000 Ci/mmol, Amersham Corp.). The double-stranded probe was hybridized to 50 µg of human placental total RNA (Clontech) in the appropriate buffer at 65 ^{o}C for 16 h. After the hybridization, 300 µl of S1 buffer was added and the mixture was digested with 800 units of S1 nuclease (Boehringer Mannheim) at room temperature for 15 min. The protected fragments were further analyzed on a 6 % polyacrylamide sequencing gel using yeast tRNA as a negative control. The exact sizes of the protected fragments were determined by comparing them with adjacent dideoxynucleotide sequencing reactions (Sanger *et al.* 1977) for lysyl hydroxylase 1 cDNA.

5 Results

5.1 Characterization of the 2-oxoglutarate binding site in human lysyl hydroxylase 1 (I)

5.1.1 Identification of amino acid residues playing a potential role in 2-oxoglutarate binding

As discussed in section 2.2.2.2, 2-oxoglutarate is known to be a highly specific requirement for the hydroxylation reaction (Kivirikko & Pihlajaniemi 1998). Its binding site in human lysyl hydroxylase 1 is assumed to consist of two main subsites (Hanauske-Abel & Günzler 1982, Majamaa *et al.* 1984, Hanauske-Abel 1991), of which subsite I is believed to comprise a positively charged side-chain of the enzyme which binds the C-5 carboxyl group of 2-oxoglutarate, while subsite II covers two coordination sites of the enzyme-bound Fe^{2+} (Hanauske-Abel & Günzler 1982, Majamaa *et al.* 1984, Hanauske-Abel 1991). In spite of the critical role of 2-oxoglutarate in the hydroxylation reaction, no detailed study has yet been made of its binding site in lysyl hydroxylases.

Recent site-directed mutagenesis studies of prolyl 4-hydroxylase have shown that the residue which binds the C-5 carboxyl group of 2-oxoglutarate is a lysine residue situated in position +10 with respect to the Fe^{2+} -binding histidine residue in the His-2 motif of the α subunit (Myllyharju & Kivirikko 1997). Meanwhile it has been demonstrated that the carboxyl group of valine in the tripeptide substrate of isopenicillin N-synthase becomes bound to Arg279 and Ser281, while the three Fe^{2+} binding ligands are His214, Asp216 and His270 (Roach *et al.* 1997). Comparison of the sequences of the His-2 motifs present in these enzymes with the three human lysyl hydroxylase isoenzymes shows a considerable degree of identity (Figure 1 in I and Figure 5). Two arginine residues and one serine are present in the lysyl hydroxylase 1 sequence in this region, but no lysine or histidine. The residue in lysyl hydroxylase 1 corresponding to Lys493 in the α subunit of human type I prolyl 4-hydroxylase is Arg700, which is conserved between the three lysyl

hydroxylase isoenzymes, and which was therefore chosen for the mutagenesis studies. Another arginine, Arg697, is not conserved between the members of the lysyl hydroxylase enzyme family, and was thus not likely to be involved in binding of the 2-oxoglutarate but could instead be used as a control. The only serine present between His690 and the C-terminus is the conserved residue 705, which was also studied further.

	690								697			**700**					**705**					
LH1	**H**	E	G	L	·	P	T	T	R	G	T	**R**	Y	I	A	V	**S**	F	V	D	P	*term*
LH2	**H**	E	G	L	·	P	V	K	N	G	T	**R**	Y	I	A	V	**S**	F	I	D	P	*term*
LH3	**H**	E	C	L	·	P	T	T	W	G	T	**R**	Y	I	M	V	**S**	F	V	D	P	*term*
α4PH1	**H**	A	A	C	·	P	V	L	V	G	N	**K**	·	·	W	V	**S**	N	K	W	L	
α4PH2	**H**	A	A	C	·	P	V	L	V	G	C	**K**	·	·	W	V	**S**	N	K	W	F	
IPNS	**H**	R	V	K	W	·	·	V	N	A	E	**R**	·	·	·	Q	**S**	L	P	F	F	

Figure 5. Comparison of the amino acid sequence extending from His690 (His 2 motif) to the C-terminus of human lysyl hydroxylase 1 (LH1) (Hautala *et al.* 1992a) with corresponding sequences in the other human lysyl hydroxylase isoenzymes (LH2-3) (Valtavaara *et al.* 1997, Valtavaara *et al.* 1998), two human prolyl 4-hydroxylase α subunit isoforms (α4PH1 and 2) (Helaakoski *et al.* 1989, Annunen *et al.* 1997) and isopenicillin N synthase from *Aspergillus nidulans* (IPNS) (Roach *et al.* 1995). Gaps (·) are introduced for alignment. Conserved amino acids are shown with black letters (modified from Fig.1 in paper I).

5.1.2 Identification of the amino acid residue involved in 2-oxoglutarate binding by site-directed mutagenesis

To study the function of the two conserved amino acid residues Arg700 and Ser705, these were converted individually to alanine. Arg697 was likewise changed to alanine as a conrol. After preliminary experiments, Arg700 was also converted to lysine. The effects of the various mutations were investigated by expressing the mutant polypeptides as recombinant proteins in High Five insect cells. The cells were harvested and homogenized as described in Materials and Methods (see 4.4) and the proteins soluble in the glycerol buffer were analyzed by SDS-PAGE.

No differences in mobility were found between the various mutant lysyl hydroxylase polypeptides in SDS-PAGE (Figure 2 in I). Only very minor differences were found in the amounts of the polypeptides in the glycerol buffer extracts when comparing the intensities of the Coomassie blue stained bands corresponding to the variably glycosylated lysyl hydroxylase polypeptides. These bands were further studied by densitometry. The amounts of the various mutant enzymes were calculated as percentages

of the amount of the wild-type enzyme and these percentages were then used to correct the enzyme activities of the various mutants for the very minor differences in expression levels.

Assays of lysyl hydroxylase indicated that mutation of Arg700 to alanine inactivated the enzyme completely, whereas mutation of Arg697 to alanine had only a minor effect (Table 1 in I). Mutation of Ser705 to alanine had a slightly larger effect than the Arg697Ala mutation, the activity level being 67 % of that of the wild-type enzyme. The effect of this mutation was still very minor relative to that of the Arg700Ala mutation, however. The Arg700Lys mutant enzyme was not completely inactive, but the activity level obtained in assays using 100 μM 2-oxoglutarate was only about 7 % of that of the wild-type enzyme. When the 2-oxoglutarate concentration in the *in vitro* hydroxylation reaction was increased to 300 μM, the relative activity of the Arg700Lys mutant increased to about 14 %, suggesting impaired binding of 2-oxoglutarate.

To study the effects of the Arg700Lys and Ser705Ala mutations in more detail, K_m values for 2-oxoglutarate were determined for the wild-type enzyme and these mutants. That of the Arg700Lys mutant enzyme was found to have increased from 100 to 1000 μM, i.e about 10-fold (Table 2), whereas that of the Ser705Ala mutant was found to be identical to that of the wild-type enzyme. The V_{max} obtained with the Arg700Lys enzyme was the same as with the wild-type enzyme.

Table 2. K_m values of the wild-type and Arg700Lys and Ser705Ala mutant lysyl hydroxylases for 2-oxoglutarate

Enzyme	K_m[a] (μM)
Wild-type	100±25
Arg700Lys	1000±160
Ser705Ala	100

[a] The values are given as mean ± S.D., n=8 for the wild-type lysyl hydroxylase, n=11 for Arg700Lys, and n=1 for Ser705Ala.

5.2 A novel isoenzyme, lysyl hydroxylase 3 (II, III)

5.2.1 Molecular cloning and characterization of the cDNA (II)

To find out whether human lysyl hydroxylase 1 has additional isoenzymes, the database of human expressed sequence tags (EST) was searched using the lysyl hydroxylase 1 sequence (Hautala *et al.* 1992a) as a probe. This resulted in the identification of a 343-bp sequence in a foetal kidney database showing 81 % identity to the 3' end of the coding region of lysyl hydroxylase 1 cDNA and 77 % identity to the 3' end of the coding region of lysyl hydroxylase 2a cDNA at the nucleotide level. PCR amplification from an adult

human kidney cDNA library gave a product corresponding to the EST sequence, and this product was then used to screen foetal and adult human kidney and placenta cDNA libraries. Seven of the 21 clones obtained were characterized further, but none of them contained the 5′ end of the cDNA. Finally the 5′ end was obtained by rescreening the placenta library with two new PCR probes and simultaneous 5'-RACE analysis with pooled human placental cDNAs.

The cDNA clones for human lysyl hydroxylase 3 cover 61 nucleotides of the 5' untranslated sequence, the whole coding region corresponding to 738 amino acid residues, and 295 nucleotides of the 3' untranslated sequence. The 3' untranslated sequence contains the canonical polyadenylation signal AATAAA, which is accompanied 11 bp downstream by a poly(A)+ tail of 46 nucleotides. The N-terminus of the coded polypeptide has a putative signal peptide, the size of which is 24 amino acid residues according to the computational parameters of von Heijne (von Heijne 1986). Thus the most likely first amino acid residue of the mature polypeptide 714 amino acids in size is a serine. The overall amino acid sequence identity between the human lysyl hydroxylases 3 and 1 is 59 %, and that between the human lysyl hydroxylases 3 and 2a 57 %, whereas the identity between the human lysyl hydroxylase 3 and the *C. elegans* lysyl hydroxylase is only 45 %. The 87 extreme C-terminal residues are 72 % identical between human lysyl hydroxylases 3 and 1, 69 % identical between lysyl hydroxylases 3 and 2, and 64 % identical between human lysyl hydroxylase 3 and the *C. elegans* lysyl hydroxylase, and thus the identity is clearly highest within the catalytically important C-terminal region (Pirskanen *et al.* 1996, Kivirikko & Pihlajaniemi 1998). All the critical amino acid residues required for the binding of the Fe^{2+} atom (Pirskanen *et al.* 1996) and the C-5 carboxyl group of 2-oxoglutarate are conserved in all these polypeptides, the residues being His643, Asp645, His695 and Arg705 in lysyl hydroxylase 3.

All lysyl hydroxylase isoenzymes have potential attachment sites for asparagine-linked oligosaccharides. Human lysyl hydroxylase 3 has two such sites, lysyl hydroxylases 1 and 2 have four and seven sites, respectively, whereas the *C. elegans* lysyl hydroxylase has only one site. Lysyl hydroxylase 3 cDNA contains ten cysteine residues, six of them being conserved in all five polypeptides (Figure 1 in II).

5.2.2 Expression in various human tissues (II)

Northern hybridization with a cDNA probe for lysyl hydroxylase 3 indicated the presence of only a single mRNA of about 3.0 kb in size, which is expressed in a variety of tissues (Figure 2 in II). This mRNA is slightly smaller than the 3.2 kb mRNA for isoenzyme 1 (Hautala *et al.* 1992a) and considerably smaller than the 4.2 kb and 4.4 kb mRNAs for isoenzymes 2a (Valtavaara *et al.* 1997) and 2b (Valtavaara, M., Risteli, M., Ruotsalainen H., Wang, C. and Myllylä, R., unpublished data). The highest expression levels among the tissues studied were found in the placenta, pancreas, spinal cord and heart. Some foetal tissues such as the lung, kidney and liver seemed to have higher expression levels than the corresponding adult tissues. High expression levels were also detected in human Saos-2 osteosarcoma cells and fibroblasts, the level being considerably higher in the

former. In earlier studies the mRNAs for lysyl hydroxylase 1 and 2 were found to be expressed in a variety of human tissues, the highest expression levels of the isoenzyme 1 mRNA being obtained in the liver and skeletal muscle (Heikkinen *et al.* 1994), whereas those for isoenzyme 2 mRNA were highest in the pancreas, placenta, heart and skeletal muscle (Valtavaara *et al.* 1997).

5.2.3 Expression in insect cells and characterization of the recombinant protein (II)

In order to study the catalytic properties of human lysyl hydroxylase 3, a baculovirus coding for this isoenzyme was generated and used to infect High Five cells. The cells were harvested 24-72 h after infection, homogenized as described in Materials and Methods (section 4.4) and analyzed in SDS-PAGE under reducing conditions followed by Coomassie blue staining. The infected cells expressed a new 80-85 kDa polypeptide not seen 24 h after infection (Figure 3 in II). This was present in all three fractions, i.e. the Nonidet P-40 and glycerol buffer extracts and the samples solubilized with 1 % SDS. The protein was most abundant in the SDS-soluble fraction. As has previously been shown for the recombinant isoenzyme 1 (Pirskanen *et al.* 1996), the novel polypeptide also migrated in multiple bands, probably due to heterogeneity in glycosylation.

The Nonidet P-40 and glycerol buffer extracts were analyzed for lysyl hydroxylase activity with an assay based on hydroxylation-coupled decarboxylation of 2-oxo[1-^{14}C]glutarate (Table 1 in II). The Nonidet P-40 buffer extract was found to give a high 2-oxo[1-^{14}C]glutarate decarboxylation rate even in non-infected cells. A slight increase in enzyme activity level was seen in both extracts 24 h after infection. The activity in the Nonidet P-40 extract reached its maximum value at 48 h, whereas that in the glycerol extract continued to increase throughout the 72 hours of the experiment.

K_m values for Fe^{2+}, 2-oxoglutarate, ascorbate and two peptide substrates with soluble cell extracts as sources of the enzyme were 2, 100, 300, 600 and 800 μM, respectively (Table 2 in II). All these values are essentially identical to those reported for recombinant human lysyl hydroxylase 1 (Pirskanen *et al.* 1996).

5.2.4 Organization of the gene (III)

Two P1 genomic clones for human lysyl hydroxylase 3 were obtained from the P1 Human Library screening services. Clone 18051 was found to contain only the 3' end of the gene, whereas clone 18050 spanned the entire gene. The exon-intron organization of the gene was studied by sequencing with cDNA-specific primers and comparing the sequence obtained with the previously published human lysyl hydroxylase 3 cDNA sequence. All the exons and introns, except for introns 6, 15 and 16, were characterized

by direct sequencing of clone 18050. Because introns 6, 15 and 16 were difficult to sequence, they were amplified using PCR before sequencing.

The overall size of the human lysyl hydroxylase 3 gene is only 11.6 kb, and it contains 19 exons, ranging in size from 64 to 454 bp (Table 1 in III). Exon 1 is 358-444 bp in size and contains 109 bp of a translated sequence. Exons 2-18 vary in size from 64 to 164 bp. The last exon, number 19, is clearly the largest, 454 bp, and contains 156 bp of the coding region and the entire 3' untranslated region. The sizes of the exons are well conserved as compared with those of the human lysyl hydroxylase 1 gene (Heikkinen *et al.* 1994), but the introns are markedly shorter. They range from 83 to 1883 bp, intron 6 being the largest. The boundaries of all the exons and introns follow the AG-exon-GT consensus rule.

In order to study the existence of Alu sequences in the human lysyl hydroxylase 3 gene, the sequence obtained was analyzed further using the BLAST program. A total of 15 full-length Alu retroposons, or partial Alu fragments of more than 100 bp, were found in the gene. The longest intron, number 6, of size 1883 bp, contains two full-length Alu repeats and one partial one in the coding strand and one additional full-length Alu repeat in the non-coding strand (Figure 1 in III). Intron 12, 916 bp in size, contains two full-length Alu repeats in the coding strand. Introns 5, 15 and 17, ranging in size from 673 to 1136 bp, each contain one full-length Alu repeat and three partial ones; all situated in the coding strand. Intron 16 contains one full-length Alu sequence in the coding strand and two partial ones, the corresponding intron in the human lysyl hydroxylase 1 gene having been shown earlier to contain eight Alu sequences (Heikkinen *et al.* 1994).

5.2.5 Determination of the transcription initiation site and analysis of the 5´ flanking region (III)

The transcription initiation site of the lysyl hydroxylase 3 gene was determined by S1 nuclease protection analysis as described in Materials and Methods (see section 4.9). After S1 nuclease digestion, nine protected bands ranging in size from 288 to 374 bp were found. Comparison of the protected fragments with the adjacent dideoxynucleotide sequencing reactions indicated the presence of one major transcription initiation site and several minor ones (Figure 2 in III). Thus the length of the 5'untranslated region in the human lysyl hydroxylase 3 mRNA ranges from 249 to 335 nucleotides, the 311-nucleotide form being the most abundant.

The 5' flanking region of the lysyl hydroxylase 3 gene was found to show no nucleotide sequence similarity to that of the human lysyl hydroxylase 1 gene (Heikkinen et al., 1994). Both genes lack a typical TATAA box (Figure 3 in III), and the lysyl hydroxylase 3 gene also lacks a CCAAT box in the 5´ flanking region. The sequence of the lysyl hydroxylase 3 gene at the transcription start site does not contain the loose consensus sequence of known promoter initiator (Inr) elements (Javahery *et al.* 1994), but two MAZ binding sites were found, one of them in the 5' flanking region and the other in the 5' untranslated sequence. The factors binding to the MAZ sites interact with Sp1 transcription factors and may function as important regulators of TATA-less promoters

(Parks & Shenk 1996, Parks & Shenk 1997). Three potential binding sites for the transcription factors Sp1, one for PEA-3, one for NF-1, one for USF and two for AP-1 were found in the 5' flanking region, while two additional Sp1, one NF-1, one AP-1, one AP-2 and one USF binding sites were present in the 5' untranslated sequence. The first intron of the gene contains one further Sp1 site, one AP-1 site and one MAZ site.

6 Discussion

6.1 Identification of the residue that binds the C-5 carboxyl group of 2-oxoglutarate in the lysyl hydroxylases

The reaction catalyzed by lysyl hydroxylase requires Fe^{2+}, 2-oxoglutarate, O_2 and ascorbate (see 2.2.2) (Kivirikko & Pihlajaniemi 1998). The binding site of 2-oxoglutarate in human lysyl hydroxylase 1 was characterized in paper I of the present work, thus extending the basic knowledge on the catalytic site of the enzyme reported from our laboratory (Pirskanen *et al.* 1996). A positively charged side-chain of the enzyme had been believed to play a central role in this binding site (Hanauske-Abel & Günzler 1982, Majamaa *et al.* 1984, Hanauske-Abel 1991), but no investigations into this residue had been carried out previously. One reason for the lack of information on the catalytic site of lysyl hydroxylase may be related to difficulties in the expression and purification of the enzyme, which easily forms insoluble aggregates. Recent site-directed mutagenesis studies on prolyl 4-hydroxylase have made it possible to locate the residue responsible for the binding of 2-oxoglutarate in that enzyme, too (Myllyharju & Kivirikko 1997).

The data reported in paper I strongly suggest that Arg700 is the positively charged residue that binds the C-5 carboxyl group of 2-oxoglutarate in human lysyl hydroxylase 1. Mutation of this residue to alanine inactivated the enzyme completely, while mutation to lysine increased the K_m for 2-oxoglutarate about 10-fold. Ser705 and Arg697, the two other residues studied, appeared to play no role in this binding.

An arginine residue corresponding to Arg700 is conserved in the sequences of almost all 2-oxoglutarate dioxygenases studied so far (Figure 1 in I), and when present, it is likely to have a similar function among the members of the enzyme group concerned here. However, the α subunits of prolyl 4-hydroxylases from all sources studied have a lysine in the corresponding position instead of an arginine (Helaakoski *et al.* 1989, Helaakoski *et al.* 1995, Annunen *et al.* 1997). Mutation of this lysine in prolyl 4-hydroxylase to arginine increased the K_m for 2-oxoglutarate about 15-fold with no change in the V_{max} (Myllyharju & Kivirikko 1997). The K_m values of the lysyl hydroxylases and prolyl 4-hydroxylases for 2-oxoglutarate (Helaakoski *et al.* 1995, Pirskanen *et al.* 1996)

differ considerably, as do the inhibition constants of these enzymes for 23 different 2-oxoglutarate analogues (Majamaa *et al.* 1985). The results presented in paper I thus indicate that these two collagen hydroxylases, which have very similar catalytic mechanisms, clearly have a different requirement with respect to the residue that binds the 2-oxoglutarate.

6.2 Human lysyl hydroxylase 3 – a new member of the lysyl hydroxylase family

A novel member of the lysyl hydroxylase family, termed lysyl hydroxylase 3, was cloned and characterized in this work. The processed novel polypeptide of 714 amino acid residues is very similar in size to the corresponding human lysyl hydroxylases 1 (Hautala *et al.* 1992a), 2a (Valtavaara *et al.* 1997) and 2b (Yeowell & Walker 1999a) and to that of *C. elegans* (Wilson *et al.* 1994), which consist of 709, 712, 733 and 714 amino acid residues, respectively (Fig. 1 in II and Table 3). The longest of these polypeptides is lysyl hydroxylase 2b, which contains 21 extra amino acid residues (Yeowell & Walker 1999a). Unlike the sizes of the processed lysyl hydroxylase polypeptides, the sizes of the corresponding mRNAs show major variation, ranging from the 3.0 kb transcript for lysyl hydroxylase 3 to the 4.4 kb transcript for lysyl hydroxylase 2b (Valtavaara, M., Risteli, M., Ruotsalainen H., Wang, C. and Myllylä, R., unpublished data). This is mainly due to variation in the lengths of the untranslated regions. The 3´UTR of an mRNA is known to play a role in the regulation of translation in some cases (for reviews, see Jackson 1993, Wilhelm & Vale 1993), and according to some studies there is a relationship between the length of the 3´ UTR and translational efficiency (Tanguay & Gallie 1996).

The overall amino acid sequence identity between polypeptides of all four human lysyl hydroxylases is about 47 %. The degree of identity is about 60 % and similarity about 85% when any two of the human isoenzymes are compared. The C-terminal region, which contains the amino acid residues involved in the binding of Fe^{2+} and 2-oxoglutarate, is particularly well conserved, the overall identity between the human isoenzymes in this region being about 68 % and the highest identity, between lysyl hydroxylases 1 and 3, reaching about 82 %. The recently characterized mouse isoenzymes are 91 % identical to the corresponding human enzymes (Ruotsalainen *et al.* 1999), the similarity between the corresponding isoenzymes in the two mammalian species thus being higher than that between two isoenzymes in the same species.

All human lysyl hydroxylase polypeptides contain 9 cysteine residues in conserved positions, but lysyl hydroxylases 1 and 2 also contain an additional cysteine conserved between them but not in lysyl hydroxylase 3. Still another unique cysteine is found in lysyl hydroxylases 2 and 3. The *C. elegans* lysyl hydroxylase contains 7 cysteine residues, 6 of which are conserved in the polypeptides of all four human enzymes, while its extreme N-terminal cysteine is conserved only in human lysyl hydroxylase 3. The 6 cysteines that are conserved in all human lysyl hydroxylases and that of *C. elegans* may form structurally important disulphide bonds (for reviews, see Freedman *et al.* 1998, Debarbieux & Beckwith 1999). Recent site-directed mutagenesis studies on the ten

cysteine residues in lysyl hydroxylase 1 showed that mutation of residues Cys369, Cys375, Cys552 and Cys687 eliminated enzyme activity (Yeowell *et al.* 2000b), these cysteines being also found in conserved positions in the sequences of all human lysyl hydroxylase isoenzymes. On the other hand, mutation of Cys267, Cys270 and Cys680 in lysyl hydroxylase 1 had an intermediate effect on the enzyme activity, and mutation of the other three cysteines did not cause any loss of activity. It thus seems that although disulphide bond formation may affect the relative contribution of each cysteine to the enzyme activity, catalytic activity does not appear to be directly related to dimerization of the enzyme (Yeowell *et al.* 2000b). The cysteine residues in α subunits of prolyl 4-hydroxylases form two intrachain disulphide bonds that are essential for the subunits to maintain the structure needed for tetramer assembly (John & Bulleid 1994, Lamberg *et al.* 1995).

Besides cysteine residues, the numbers of potential attachment sites for asparagine-linked oligosaccharides also vary between isoenzymes (see Table 3). Site-directed mutagenesis studies have indicated that glycosylation of the second site present in human lysyl hydroxylase 1 may be required for full enzyme activity (Pirskanen *et al.* 1996). This position also has a potential N-glycosylation site in lysyl hydroxylase 2, but surprisingly, not in human lysyl hydroxylase 3 or the *C. elegans* enzyme. It thus seems possible that the various lysyl hydroxylase isoenzymes may have different glycosylation requirements for their catalytic activity.

Table 3. Molecular properties of the human lysyl hydroxylase isoenzymes.

	LH1[1)]	LH2a[2)]	LH2b[3)]	LH3
size (aa)[a]	709	712	733	714
putative signal sequence (aa)	18	25	25	24
mRNA size (kb)	3.2	4.2	4.4	3.0
cysteine residues	10	11	11	10
N-glycosylation sites	4	7	7	2

[a] size of the polypeptide after cleavage of the signal peptide
[1)] (Hautala *et al.* 1992a)
[2)] (Valtavaara *et al.* 1997)
[3)] (Yeowell & Walker 1999a, Valtavaara, M., Risteli, M., Ruotsalainen H., Wang, C. and Myllylä, R., unpublished data)

The relative expression level of the lysyl hydroxylase 3 mRNA in the adult human liver was much lower than that of lysyl hydroxylase 1 mRNA, and its expression was much lower than that of isoenzyme 2 mRNA in the skeletal muscle (Heikkinen *et al.* 1994, Valtavaara *et al.* 1997). The level of lysyl hydroxylase 1 mRNA has previously been shown to be practically the same in several human tissues (Heikkinen *et al.* 1994), but expression of lysyl hydroxylases 2 and 3 seems to be more tightly regulated both in man and in the mouse (Valtavaara *et al.* 1997, Ruotsalainen *et al.* 1999). The lysyl hydroxylase 3 mRNA level has been shown to be high in mouse heart, lung, liver and testis (Ruotsalainen *et al.* 1999). Expression of lysyl hydroxylase 1 also seems to be

highly regulated in the mouse, suggesting that there may be some differences in regulation between the two species (Ruotsalainen *et al.* 1999). Expression of the different isoenzymes at the mRNA level has so far been studied mainly in tissues that are not very rich in collagens, as tissues containing large amounts of collagen, like bone, cartilage, skin and tendons, have not been analyzed. It should also be noted that comparison of mRNA levels in whole tissues does not take account of the possibility of major differences between certain cell types. Preliminary studies at the cell level indicate that lysyl hydroxylase 1 is the major form in skin fibroblasts, pancreatic adenocarcinoma cells and placental trophoblastic tumor cells, whereas lysyl hydroxylase 2 is the major lysyl hydroxylase form in fibrosarcoma, osteosarcoma and hepatoblastoma cells. Lysyl hydroxylase 3 seems to be a major isoenzyme in kidney adenocarcinoma cells (Wang *et al.* 2000).

Since its introduction in the 1980s, the baculovirus expression system has become established as an efficient and frequently used means of producing various recombinant proteins. It possesses many advantages over other expression systems, such as safety, high expression levels and the ability of insect cells to carry out most post-translational modifications (for reviews, see Miller 1993, Volkman 1995, Bonning & Hammock 1996, Jarvis *et al.* 1996). Human lysyl hydroxylase 1 has previously been successfully expressed in baculovirus systems (Krol *et al.* 1996, Pirskanen *et al.* 1996), making this system a natural choice for lysyl hydroxylase 3 as well. Our first purpose was to verify that the novel enzyme cloned by us does indeed function as a lysyl hydroxylase. This aspect was studied using synthetic peptides known to be hydroxylated by lysyl hydroxylase 1 *in vitro*. A definitive lysyl hydroxylase activity level was demonstrated in these experiments, considerably higher than the level obtained in another study on the same isoenzyme, probably due to differences in the expression methods used (Valtavaara *et al.* 1998). The recombinant protein was found to migrate in multiple bands in SDS-PAGE, which suggests utilization of the potential N-glycosylation sites found in the cDNA sequence. The expression data obtained in this work differ from those obtained in similar experiments with a virus coding for lysyl hydroxylase 1, in that little, if any, of the lysyl hydroxylase 1 could be extracted with the Nonidet P-40 buffer and only about 10 % was soluble in the glycerol buffer 72 h after infection, 90 % being found in the 1 % SDS-soluble portion at that point in time. By contrast, significant amounts of the lysyl hydroxylase 3 polypeptide were found in the Nonidet P-40 extracts, as has also been reported for the lysyl hydroxylase 2 polypeptide (Valtavaara *et al.* 1997). These results indicate the presence of definite differences in solubility between the isoenzymes, lysyl hydroxylase 3 being solubilized more readily than lysyl hydroxylase 1. This finding may partly explain the earlier data that show a marked heterogeneity in lysyl hydroxylase solubilization from various cells and tissues (Ryhänen 1976, Peterkofsky & Assad 1979, Kellokumpu *et al.* 1994).

The K_m values of lysyl hydroxylase 3 for the cosubstrates and peptide substrates determined here were essentially identical to those for the recombinant human lysyl hydroxylase 1 (Pirskanen *et al.* 1996). Recent preliminary studies have shown, however, that lysyl hydroxylase 3 mRNA levels do not correlate with those of the other lysyl hydroxylase isoenzymes or that of the α subunit of type I prolyl 4-hydroxylase (Wang *et al.* 2000), thus suggesting that differences may exist in the functions of lysyl hydroxylase 3 and the other isoenzymes in spite of the similarities in their K_m values. Examination of

the K_m values of all members of the human lysyl hydroxylase family characterized so far indicates that the largest differences are seen in those for the peptide substrate $(IKG)_3$, which vary from 420 to 1200 μM (Table 4). It thus appears that the catalytic properties of the lysyl hydroxylase isoenzymes operate to the same effect but are not completely identical.

The data in Table 4 do not exclude the possibility that there may be differences between lysyl hydroxylase isoenzymes with respect to the hydroxylation of various collagen types. Ihme et al. (1984) reported that lysine residues in type II, IV and V collagens are hydroxylated in patients with Ehlers-Danlos syndrome type VI to the same extent as in healthy subjects. This may be due to the existence of collagen type-specific isoenzymes (Ihme *et al.* 1983). Furthermore, the residual lysyl hydroxylase activity in cells from patients with this syndrome was reported to preferentially hydroxylate lysine residues in type IV collagen (Risteli *et al.* 1980). However, comparison of the mRNA levels of the different lysyl hydroxylase isoenzymes with those of the most abundant collagen types, I, III, IV and V, in various tissues nevertheless showed no significant correlations (Wang *et al.* 2000), the results thus indirectly arguing against any collagen type specificity in lysine hydroxylation.

Table 4. K_m values of the recombinant human lysyl hydroxylase isoenzymes for cosubstrates and two peptide substrates

Cosubstrate or peptide substrate	K_m, μM			
	LH1[1)]	LH2a[2)]	LH2b[3)]	LH3
Fe^{2+}	2	4	5	2
2-oxoglutarate	100	190	170	100
Ascorbate	350	470	230	300
Peptide substrate $(IKG)_3$	700	420	1200	800
Peptide substrate L-1[*]	500	n.r	n.r	500

[1)] (Pirskanen *et al.* 1996)
[2)] (Valtavaara *et al.* 1997)
[3)] (Valtavaara *et al.* 1998)
[*] Ala-Arg-Gly-Ile-Lys-Gly-Ile-Arg-Gly-Phe-Ser-Gly
n.r = not reported

In order to gain more information on the human lysyl hydroxylase 3 isoenzyme, its gene was isolated and the exon-intron structure determined. The gene had already been mapped to chromosome 7q36 (Valtavaara *et al.* 1998), while the lysyl hydroxylase 1 gene is located on chromosome 1p36.2-36.3 (Hautala *et al.* 1992a, Van Roy *et al.* 1993) and the lysyl hydroxylase 2 gene on chromosome 3q23-q24 (Szpirer *et al.* 1997). The lysyl hydroxylase isoenzyme genes are thus not linked to each other.

The human lysyl hydroxylase 3 gene was found to consist of 19 exons dispersed in a relatively small region of 11.6 kb, smaller than the average length of a gene in mammals, which is about 16.6 kb (Lewin 1994), and considerably smaller than the lysyl hydroxylase 1 gene, 40 kb (Heikkinen *et al.* 1994). However, both genes contain the same number of exons, 19, which is more than in a typical mammalian gene, which has

seven exons on average (Lewin 1994). The exons of the LH3 gene are well conserved in terms of size as compared with those of the lysyl hydroxylase 1 gene, the average size being about 150-160 bp in both. The exons represent altogether only about 7 % of the size of the lysyl hydroxylase 1 gene, however, but as much as about 25 % in the case of the lysyl hydroxylase 3 gene. The exon-intron boundaries are well conserved between the isoenzymes, the only boundary that differs by six nucleotides being the exon-intron boundary of intron 8. The consensus GT/AG sequences at the exon-intron boundaries are also found in both genes.

Both genes have multiple Alu sequences in their introns. The lysyl hydroxylase 3 gene contains at least 15 full-length or partial Alu repeats dispersed among six introns, while the lysyl hydroxylase 1 gene contains at least 13 Alu repeats in two introns. The Alu-Sx subfamily accounts for nearly half of the Alu copies present in DNA (Mighell *et al.* 1997), and Alu repeats belonging to this subfamily were also found in the lysyl hydroxylase 3 gene. Another Alu subfamily found extensively in the lysyl hydroxylase 3 gene is Alu-J, which is estimated to be about 55 million years old, while the Alu-Sx subfamily is about 40 million years old (Makalowski *et al.* 1994). Alu repeats resembling the consensus sequences for the Alu-Sp, Alu-Sq and Alu-Sc subfamilies were also found in the lysyl hydroxylase 3 gene, as were sequences resembling the Alu-Sb subfamily, which provides an example of a more recently amplified repeat found only in humans and the great apes (Makalowski *et al.* 1994). As discussed in section 2.2.5.1, Alu repeats are involved in many gene rearrangements that have occurred by homologous recombination (Rüdiger *et al.* 1995, Strout *et al.* 1998) leading to diseases such as neurofibromatosis, haemophilia and breast cancer (for a review, see Deininger & Batzer 1999). An Alu-mediated homologous recombination is the most commonly found mutation in the lysyl hydroxylase 1 gene in Ehlers-Danlos syndrome type VI, a disease that is attributable to lysyl hydroxylase deficiency (see 2.3) (Heikkinen *et al.* 1994, Pousi *et al.* 1994, Heikkinen *et al.* 1997). It is currently unknown whether any heritable disorder is brought about by recombination of any of the Alu repeats present in the human lysyl hydroxylase 3 gene.

The 5´ flanking region of the lysyl hydroxylase 3 gene shows no nucleotide sequence similarity to that of the lysyl hydroxylase 1 gene (Heikkinen *et al.* 1994). The length of the 5' untranslated region in its mRNA ranges from 249 to 335 nucleotides, a 311-nucleotide form being the major species, whereas this length in the lysyl hydroxylase 1 mRNA is shorter, 60-70 nucleotides. The lysyl hydroxylase 3 and 1 genes both lack a TATAA or related sequence and share a high G+C nucleotide content in their 5´ flanking regions, these features being characteristic of genes utilizing multiple transcription initiation sites, such as housekeeping genes. Both genes also lack a core promoter element, called an initiator (Inr), the function of which is analogous to that of TATAA in that it is capable of determining the precise site of transcription initiation (Kaufmann & Smale 1994). The 5´ flanking region of the lysyl hydroxylase 3 gene also lacks a CCAAT sequence, as is found in the isoenzyme 1 gene. The lysyl hydroxylase 3 gene contains binding sites for several other transcription factors, but the functional significances of these possible regulatory elements remain to be determined.

According to phylogenetic analyses, lysyl hydroxylase 3 is the oldest form of lysyl hydroxylase (Ruotsalainen *et al.* 1999), and our findings concerning the origin of its Alu sequences are consistent with this as the majority of its Alu sequences seem to belong to

the oldest Alu subfamilies. It seems likely that all lysyl hydroxylase isoenzymes are derived from an ancestral gene through two gene duplication events, lysyl hydroxylases 1 and 2 having been brought about by a more recent duplication than the less closely related lysyl hydroxylase 3. As the oldest lysyl hydroxylase enzyme, lysyl hydroxylase 3 has been thought to be evolutionarily less closely related to the others (Ruotsalainen *et al.* 1999). At least a considerable heterogeneity is seen when examining the overall similarity between the 5´ flanking regions of the human lysyl hydroxylase 1 and 3 genes, not only at the nucleotide level but also when comparing the potential binding sites for different transcription factors, which suggests the existence of major differences in the regulation of the expression of these two genes. Our results are also consistent with recent findings published by Wang *et al.* (2000) showing differences between the expression of lysyl hydroxylase 3 mRNA and that of mRNAs for other isoenzymes and prolyl 4-hydroxylase.

7 Future perspectives

Altogether three lysyl hydroxylase isoenzymes have been characterized in human tissues (Valtavaara *et al.* 1997, Valtavaara *et al.* 1998, paper II of the present study) and in the mouse (Ruotsalainen *et al.* 1999). Following this demonstration that lysyl hydroxylase indeed has multiple isoenzymes, the next important aim will be to point out the functional differences between them. The presence of lysyl hydroxylase isoenzymes probably explains the great variability in the extent of lysine hydroxylation between collagens from different tissues in healthy subjects and in patients with Ehlers-Danlos syndrome type VI. Preliminary investigations into the tissue distribution of mRNAs for the members of the lysyl hydroxylase enzyme family have already been published (Valtavaara *et al.* 1997, Valtavaara *et al.* 1998, Ruotsalainen *et al.* 1999, Wang *et al.* 2000, paper II of the present study), but no clear or significant differences between them have been reported so far. Monoclonal and polyclonal antibodies prepared against these isoenzymes should provide more information about their tissue distribution, and especially their expression in various cell types. Also, experiments are in progress to generate knock-out mice for studying the roles of the different isoenzymes. It will be of interest to see the phenotypes of the knock-out mice for the various isoenzymes and to learn the degrees of compensation that the other isoenzymes are capable of achieving. More research will also be needed to determine the differences in catalytic properties between the lysyl hydroxylase isoenzymes, at least with regard to their substrate specificity. It seems possible that none of the isoenzymes characterized so far is collagen type-specific (Wang *et al.* 2000), but instead it may be the amino acid sequences surrounding the individual lysine residues that are important for determining which lysyl hydroxylase isoenzyme catalyzes the hydroxylation reaction. It is also possible that one of the isoenzymes may turns out to be specific to non-collagenous proteins. Many findings suggest the presence of a telopeptide-specific lysyl hydroxylase, and it is possible that none of the isoenzymes characterized so far may be able to fulfil this function. One study has suggested that the gene for a telopeptide-specific lysyl hydroxylase isoenzyme may be located on chromosome 17 (Bank *et al.* 1999), whereas the genes for the three isoenzymes characterized so far have been mapped to chromosomes 1, 3 and 7 (Hautala *et al.* 1992a, Van Roy *et al.* 1993, Szpirer *et al.* 1997, Valtavaara *et al.* 1998). Thus a further search for additional lysyl hydroxylase isoenzymes would appear to be justified.

The recombinant human lysyl hydroxylase 1 produced in insect cells have a high tendency to form aggregates, and thus this recombinant protein does not seem to be an ideal target for crystallization experiments. Identification of distinct domains in the human lysyl hydroxylase polypeptides should provide a mean for structural studies in the future, whereupon it would seem possible to resolve the three-dimensional structures of these enzyme domains by means of NMR and crystallization.

The decreased lysyl hydroxylase activity observed in patients with Ehlers-Danlos syndrome type VIA is known to be due to mutations in the gene for lysyl hydroxylase 1, and it will be of interest to learn whether any heritable disorder is caused by mutations in the gene for lysyl hydroxylase 3. The large number of Alu sequences found in the human lysyl hydroxylase 3 gene should provide many possibilities for Alu-mediated homologous recombination. The exon-intron organization of the gene characterized in this work provides the necessary information for further considerations of this aspect and may thus contribute to the diagnosis of diseases and even to the devising of forms of gene therapy in the long run. Potential diseases associated with lysyl hydroxylase 2 are likewise of great interest.

References

Alberts B, Bray D, Lewis J, Raff M, Roberts K & Watson J (1994) Molecular biology of the cell, 3rd edn. Garland Publishing, Inc., New York, USA.

Altschul SF, Boguski MS, Gish W & Wootton JC (1994) Issues in searching molecular sequence databases. Nat Genet 6: 119-129.

Altschul SF, Gish W, Miller W, Myers EW & Lipman DJ (1990) Basic local alignment search tool. J Mol Biol 215: 403-410.

Ananthanarayanan VS, Saint-Jean A & Jiang P (1992) Conformation of a synthetic hexapeptide substrate of collagen lysyl hydroxylase. Arch Biochem Biophys 298: 21-28.

Andrews PC, Hawke D, Shively JE & Dixon JE (1984) Anglerfish preprosomatostatin II is processed to somatostatin-28 and contains hydroxylysine at residue 23. J Biol Chem 259: 15021-15024.

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.

Armstrong LC & Last JA (1995) Rat lysyl hydroxylase: molecular cloning, mRNA distribution and expression in a baculovirus system. Biochim Biophys Acta 1264: 93-102.

Bailey AJ, Paul RG & Knott L (1998) Mechanisms of maturation and ageing of collagen. Mech Ageing Dev 106: 1-56.

Bailey AJ & Shimokomaki MS (1971) Age related changes in the reducible cross-links of collagen. FEBS Letter 16: 86-88.

Bailey AJ, Sims TJ, Le L & Bazin S (1975) Collagen polymorphism in experimental granulation tissue. Biochem Biophys Res Commun 66: 1160-1165.

Bailey AJ, Wotton SF, Sims TJ & Thompson PW (1992) Post-translational modifications in the collagen of human osteoporotic femoral head. Biochem Biophys Res Commun 185: 801-805.

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: 1054-1058.

Barnes MJ, Constable BJ, Morton LF & Kodicek E (1973a) Bone collagen metabolism in vitamin D deficiency. Biochem J 132: 113-115.

Barnes MJ, Constable BJ, Morton LF & Kodicek E (1973b) The influence of dietary calcium deficiency and parathyroidectomy on bone collagen structure. Biochim Biophys Acta 328: 373-382.

Barnes MJ, Constable BJ, Morton LF & Royce PM (1974) Age-related variations in hydroxylation of lysine and proline in collagen. Biochem J 139: 461-468.

Bateman JF, Lamandé SR & Ramshaw JAM (1996) Collagen superfamily. In: Comper WD (ed) Extracellular matrix: Molecular components and interactions. Harwood Academic Publishers, Amsterdam, p 22-67.

Bateman JF, Mascara T, Chan D & Cole WG (1987) A structural mutation of the collagen alpha 1(I)CB7 peptide in lethal perinatal osteogenesis imperfecta. J Biol Chem 262: 4445-4451.

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 122: 109-115.

Batzer MA, Deininger PL, Hellmann-Blumberg U, Jurka J, Labuda D, Rubin CM, Schmid CW, Zietkiewicz E & Zuckerkandl E (1996) Standardized nomenclature for Alu repeats. J Mol Evol 42: 3-6.

Beighton P (1993) The Ehlers-Danlos syndrome. In: Beighton P (ed) McKusicks´s heritable disorders of connective tissue. Mosby-Year Book, Inc., St. Louis, pp 189-251.

Beighton P (1970) Serious ophthalmological complications in the Ehlers-Danlos syndrome. Br J Ophthalmol 54: 263-268.

Beighton P, De Paepe A, Steinmann B, Tsipouras P & Wenstrup RJ (1998) Ehlers-Danlos syndromes: revised nosology, Villefranche, 1997. Ehlers-Danlos National Foundation (USA) and Ehlers-Danlos Support Group (UK). Am J Med Genet 77: 31-37.

Bonadio J & Byers PH (1985) Subtle structural alterations in the chains of type I procollagen produce osteogenesis imperfecta type II. Nature 316: 363-366.

Bonning BC & Hammock BD (1996) Development of recombinant baculoviruses for insect control. Annu Rev Entomol 41: 191-210.

Brenner RE, Vetter U, Stoss H, Müller PK & Teller WM (1993) Defective collagen fibril formation and mineralization in osteogenesis imperfecta with congenital joint contractures (Bruck syndrome). Eur J Pediatr 152: 505-508.

Brinckmann J, Acil Y, Feshchenko S, Katzer E, Brenner R, Kulozik A & Kügler S (1998) Ehlers-Danlos syndrome type VI: lysyl hydroxylase deficiency due to a novel point mutation (W612C). Arch Dermatol Res 290: 181-186.

Brinckmann J, Notbohm H, Tronnier M, Acil Y, Fietzek PP, Schmeller W, Muller 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.

Britten RJ (1996) DNA sequence insertion and evolutionary variation in gene regulation. Proc Natl Acad Sci U S A 93: 9374-9377.

Brosius J (1999) RNAs from all categories generate retrosequences that may be exapted as novel genes or regulatory elements. Gene 238: 115-134.

Burgeson RE & Morris NP (1987) The collagen family of proteins. In: Uitto J & Perejda AJ (eds) Connective tissue disease, Molecular pathology of the extracellular martix. Marcel Dekker, New York.

Byers PH (1994) Ehlers-Danlos syndrome: recent advances and current understanding of the clinical and genetic heterogeneity. J Invest Dermatol 103: 47S-52S.

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

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

Crouch E, Chang D, Rust K, Persson A & Heuser J (1994) Recombinant pulmonary surfactant protein D. Post-translational modification and molecular assembly. J Biol Chem 269: 15808-15813.

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

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

de Jong L, van der Kraan I & de Waal A (1991) The kinetics of the hydroxylation of procollagen by prolyl 4- hydroxylase. Proposal for a processive mechanism of binding of the dimeric hydroxylating enzyme in relation to the high kcat/K_m ratio and a conformational requirement for hydroxylation of -X-Pro-Gly- sequences. Biochim Biophys Acta 1079: 103-111.

de Waal A & de Jong L (1988) Processive action of the two peptide binding sites of prolyl 4-hydroxylase in the hydroxylation of procollagen. Biochemistry 27: 150-155.

Debarbieux L & Beckwith J (1999) Electron avenue: pathways of disulfide bond formation and isomerization. Cell 99: 117-119.

Deininger PL & Batzer MA (1999) Alu repeats and human disease. Mol Genet Metab 67: 183-193.

Dickson IR, Eyre DR & Kodicek E (1979) Influence of plasma calcium and vitamin D on bone collagen. Effects on lysine hydroxylation and crosslink formation. Biochim Biophys Acta 588: 169-173.

Drickamer K (1999) C-type lectin-like domains. Curr Opin Struct Biol 9: 585-590.

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

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

Elomaa O, Sankala M, Pikkarainen T, Bergmann U, Tuuttila A, Raatikainen-Ahokas A, Sariola H & Tryggvason K (1998) Structure of the human macrophage MARCO receptor and characterization of its bacteria-binding region. J Biol Chem 273: 4530-4538.

Eyre D (1987) Collagen cross-linking amino acids. Methods Enzymol 144: 115-139.

Eyre D & Wu JJ (1987) Type XI or 1α2α3α collagen. In:Mayne, R & Burgeson RE (eds) Structure and function of collagen types. Academic Press, Orlando, p 115-139.

Eyre DR, Dickson IR & Van Ness K (1988) Collagen cross-linking in human bone and articular cartilage. Age- related changes in the content of mature hydroxypyridinium residues. Biochem J 252: 495-500.

Eyre DR & Glimcher MJ (1972) Reducible crosslinks in hydroxylysine-deficient collagens of a heritable disorder of connective tissue. Proc Natl Acad Sci U S A 69: 2594-2598.

Eyre DR, Paz MA & Gallop PM (1984) Cross-linking in collagen and elastin. Annu Rev Biochem 53: 717-748.

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

Freedman RB, Dunn AD & Ruddock LW (1998) Protein folding: a missing redox link in the endoplasmic reticulum. Curr Biol 8: R468-R470.

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.

Gerriets JE, Reiser KM & Last JA (1996) Lung collagen cross-links in rats with experimentally induced pulmonary fibrosis. Biochim Biophys Acta 1316: 121-131.

Glimcher MJ, Shapiro F, Ellis RD & Eyre DR (1980) Changes in tissue morphology and collagen composition during the repair of cortical bone in the adult chicken. J Bone Joint Surg 62: 964-973.

Gruenwald S & Heitz J (1993) Baculovirus Expression Vector System: Procedures & Methods Manual. Pharmingen, San Diego, CA.

Guzman NA, Rojas FJ & Cutroneo KR (1976) Collagen lysyl hydroxylation occurs within the cisternae of the rough endoplasmic reticulum. Arch Biochem Biophys 172: 449-454.

Ha VT, Marshall MK, Elsas LJ, Pinnell SR & Yeowell HN (1994) A patient with Ehlers-Danlos syndrome type VI is a compound heterozygote for mutations in the lysyl hydroxylase gene. J Clin Invest 93: 1716-1721.

Halme J, Kivirikko KI & Simons K (1970) Isolation and partial characterization of highly purified protocollagen proline hydroxylase. Biochim Biophys Acta 198: 460-470.

Hanauske-Abel HM (1991) Prolyl 4-hydroxylase, a target enzyme for drug development. Design of suppressive agents and the *in vitro* effects of inhibitors and proinhibitors. J Hepatol 13 Suppl 3: S8-S15.

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

Hansen S & Holmskov U (1998) Structural aspects of collectins and receptors for collectins. Immunobiology 199: 165-189.

Harwood R, Grant ME & Jackson DS (1974) Collagen biosynthesis. Characterization of subcellular fractions from embyonic chick fibroblasts and the intracellular localization of protocollagen prolyl and protocollagen lysyl hydroxylases. Biochem J 144: 123-130.

Hautala T, Byers MG, Eddy RL, Shows TB, Kivirikko KI & Myllylä R (1992a) 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.

Hautala T, Heikkinen J, Kivirikko KI & Myllylä R (1992b) Minoxidil specifically decreases the expression of lysine hydroxylase in cultured human skin fibroblasts. Biochem J 283 (Pt 1): 51-54.

Hautala T, Heikkinen J, Kivirikko KI & Myllylä R (1993) A large duplication in the gene for lysyl hydroxylase accounts for the type VI variant of Ehlers-Danlos syndrome in two siblings. Genomics 15: 399-404.

Heikkinen E & Mattila K (2000) CSC:n bio-opas. CSC - Tieteellinen laskenta Oy, Helsinki.

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, Pousi B, Pope M & Myllylä R (1999) A null-mutated lysyl hydroxylase gene in a compound heterozygote British patient with Ehlers-Danlos syndrome type VI. Hum Mutat 14: 351.

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 U S A 92: 4427-4431.

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

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

Holmskov U & Jensenius JC (1993) Structure and function of collectins: humoral C-type lectins with collagenous regions. Behring Inst Mitt 93: 224-235.

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

Hyland J, Ala-Kokko L, Royce P, Steinmann B, Kivirikko KI & Myllylä R (1992) A homozygous stop codon in the lysyl hydroxylase gene in two siblings with Ehlers-Danlos syndrome type VI. Nat Genet 2: 228-231.

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

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.

Ihme A, Risteli L, Krieg T, Risteli J, Feldmann U, Kruse K & Müller PK (1983) Biochemical characterization of variants of the Ehlers-Danlos syndrome type VI. Eur J Clin Invest 13: 357-362.

Jackson RJ (1993) Cytoplasmic regulation of mRNA function: the importance of the 3' untranslated region. Cell 74: 9-14.

Jander R, Rauterberg J & Glanville RW (1983) Further characterization of the three polypeptide chains of bovine and human short-chain collagen (intima collagen). Eur J Biochem 133: 39-46.

Jarvis DL, Weinkauf C & Guarino LA (1996) Immediate-early baculovirus vectors for foreign gene expression in transformed or infected insect cells. Protein Expr Purif 8: 191-203.

Javahery R, Khachi A, Lo K, Zenzie-Gregory B & Smale ST (1994) DNA sequence requirements for transcriptional initiator activity in mammalian cells. Mol Cell Biol 14: 116-127.

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

Jiang P & Ananthanarayanan VS (1991) Conformational requirement for lysine hydroxylation in collagen. Structural studies on synthetic peptide substrates of lysyl hydroxylase. J Biol Chem 266: 22960-22967.

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

Judisch GF, Waziri M & Krachmer JH (1976) Ocular Ehlers-Danlos syndrome with normal lysyl hydroxylase activity. Arch Ophthalmol 94: 1489-1491.

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

Kapoor R, Bornstein P & Sage EH (1986) Type VIII collagen from bovine Descemet's membrane: structural characterization of a triple-helical domain. Biochemistry 25: 3930-3937.

Kaufmann J & Smale ST (1994) Direct recognition of initiator elements by a component of the transcription factor IID complex. Genes Dev 8: 821-829.

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.

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

Kikuchi Y & Tamiya N (1982) Solid-phase synthesis of substrate model peptides and their hydroxylation with collagen lysyl hydroxylase. Bull Chem Soc Jpn 55: 1556.

Kirsch E, Glanville RW, Krieg T & Müller P (1983) Analysis of cyanogen bromide peptides of type I collagen from a patient with lethal osteogenesis imperfecta. Biochem J 211: 599-603.

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

Kivirikko KI (1995) Posttranslational processing of collagens. In: Bittar EE, Bittar N, & Greenwich CT (eds) Principles of Medical Biology. JAI Press, London, p 233-254.

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

Kivirikko KI & Myllylä R (1979) Collagen glycosyltransferases. Int Rev Connect Tissue Res 8: 23-72.

Kivirikko KI & Myllylä R (1980) The hydroxylation of prolyl and lysyl residues. In: Freedman BR & Hawkins HC (eds) The enzymology of post-translational modifications of protein 53-104. Academic Press, London.

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

Kivirikko KI & Myllylä R (1984) Biosynthesis of collagens. In: Piez KA & Reddi AH (eds) Extracellular Matrix Biochemistry 83. Elsevier, New York.

Kivirikko KI, Myllylä R & Pihlajaniemi T (1992) Hydroxylation of proline and lysine residues in collagens and other animal and plant proteins. In:Harding, J J & Crabbe, M J C (eds) Post-Translational Modifications of Proteins. CRC Press, Inc., Boca Raton, Florida, p 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 (1967b) Enzymatic hydroxylation of proline and lysine in protocollagen. Proc Natl Acad Sci U S A 57: 782-789.

Kivirikko KI & Prockop DJ (1967a) Purification and partial characterization of the enzyme for the hydroxylation of proline in protocollagen. Arch Biochem Biophys 118: 611-618.

Kivirikko KI & Prockop DJ (1972) Partial purification and characterization of protocollagen lysine hydroxylase from chick embryos. Biochim Biophys Acta 258: 366-379.

Kivirikko KI, Ryhänen L, Anttinen H, Bornstein P & Prockop DJ (1973) Further hydroxylation of lysyl residues in collagen by protocollagen lysyl hydroxylase in vitro. Biochemistry 12: 4966-4971.

Kivirikko KI, Shudo K, Sakakibara S & Prockop DJ (1972) Studies on protocollagen lysine hydroxylase. Hydroxylation of synthetic peptides and the stoichiometric decarboxylation of 2-ketoglutarate. Biochemistry 11: 122-129.

Knott L & Bailey AJ (1998) Collagen cross-links in mineralizing tissues: a review of their chemistry, function, and clinical relevance. Bone 22: 181-187.

Knott L, Whitehead CC, Fleming RH & Bailey AJ (1995) Biochemical changes in the collagenous matrix of osteoporotic avian bone. Biochem J 310 (Pt 3): 1045-1051.

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

Kolble K, Lu J, Mole SE, Kaluz S & Reid KB (1993) Assignment of the human pulmonary surfactant protein D gene (SFTP4) to 10q22-q23 close to the surfactant protein A gene cluster. Genomics 17: 294-298.

Krane SM, Pinnell SR & Erbe RW (1972) Lysyl-protocollagen hydroxylase deficiency in fibroblasts from siblings with hydroxylysine-deficient collagen. Proc Natl Acad Sci U S A 69: 2899-2903.

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: 11-16.

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.

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

Lee YM, Leiby KR, Allar J, Paris K, Lerch B & Okarma TB (1991) Primary structure of bovine conglutinin, a member of the C-type animal lectin family. J Biol Chem 266: 2715-2723.

Lehmann HW, Bodo M, Frohn C, Nerlich A, Rimek D, Notbohm H & Müller PK (1992) Lysyl hydroxylation in collagens from hyperplastic callus and embryonic bones. Biochem J 282 (Pt 2): 313-318.

Lehmann HW, Rimek D, Bodo M, Brenner RE, Vetter U, Worsdorfer O, Karbowski A & Muller PK (1995a) Hydroxylation of collagen type I: evidence that both lysyl and prolyl residues are overhydroxylated in osteogenesis imperfecta. Eur J Clin Invest 25: 306-310.

Lehmann HW, Wolf E, Roser K, Bodo M, Delling G & Muller PK (1995b) Composition and posttranslational modification of individual collagen chains from osteosarcomas and osteofibrous dysplasias. J Cancer Res Clin Oncol 121: 413-418.

Lewin B (1994) Genes V. Oxford University Press Inc., New York.

Lim BL, Willis AC, Reid KB, Lu J, Laursen SB, Jensenius JC & Holmskov U (1994) Primary structure of bovine collectin-43 (CL-43). Comparison with conglutinin and lung surfactant protein-D. J Biol Chem 269: 11820-11824.

Lu J, Tay PN, Kon OL & Reid KB (1996) Human ficolin: cDNA cloning, demonstration of peripheral blood leucocytes as the major site of synthesis and assignment of the gene to chromosome 9. Biochem J 313 (Pt 2): 473-478.

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

Luckow VA, Lee SC, Barry GF & Olins PO (1993) Efficient generation of infectious recombinant baculoviruses by site-specific transposon-mediated insertion of foreign genes into a baculovirus genome propagated in Escherichia coli. J Virol 67: 4566-4579.

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

Majamaa K, Günzler V, Hanauske-Abel HM, Myllylä R & Kivirikko KI (1986) Partial identity of the 2-oxoglutarate and ascorbate binding sites of prolyl 4-hydroxylase. J Biol Chem 261: 7819-7823.

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

Majamaa K, Turpeenniemi-Hujanen TM, Latipää P, Günzler V, Hanauske-Abel HM, Hassinen IE & Kivirikko KI (1985) Differences between collagen hydroxylases and 2-oxoglutarate dehydrogenase in their inhibition by structural analogues of 2- oxoglutarate. Biochem J 229: 127-133.

Makalowski W, Mitchell GA & Labuda D (1994) Alu sequences in the coding regions of mRNA: a source of protein variability. Trends Genet 10: 188-193.

Marra MA, Hillier L & Waterston RH (1998) Expressed sequence tags-ESTablishing bridges between genomes. Trends Genet 14: 4-7.

McKusick VA (1983) Mendelian inheritance in man. Catalogs of autosomal dominant, autosomal recessive, and X-linked phenotypes. Johns Hopkins University Press, Baltimore, 1378 pp.

McKusick VA (1992) Mendelian inheritence in man. Catalogs of autosomal dominant, autosomal recessive, and X-linked phenotypes. Johns Hopkins University Press, Baltimore, 2320 pp.

Mighell AJ, Markham AF & Robinson PA (1997) Alu sequences. FEBS Lett 417: 1-5.

Miller EJ, Martin GR, Piez KA & Powers MJ (1967) Characterization of chick bone collagen and compositional changes associated with maturation. J Biol Chem 242: 5481-5489. Miller LK (1993) Baculoviruses: high-level expression in insect cells. Curr Opin Genet Dev 3: 97-101.

Miller RL (1971) Chromatographic separation of the enzymes required for hydroxylation of lysine and proline residues of protocollagen. Arch Biochem Biophys 147: 339-342.

Molony MS, Wu S-L, Keyt LK & Harris RJ (1995) The unexpected precence of hydroxylysine in non-collagenous proteins. In: Techniques in protein chemistry VI 91-98.

Murad S & Pinnell SR (1987) Suppression of fibroblast proliferation and lysyl hydroxylase activity by minoxidil. J Biol Chem 262: 11973-11978.

Murad S, Tennant MC & Pinnell SR (1992) Structure-activity relationship of minoxidil analogs as inhibitors of lysyl hydroxylase in cultured fibroblasts. Arch Biochem Biophys 292: 234-238.

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

Myllylä R, Günzler V, Kivirikko KI & Kaska DD (1992) Modification of vertebrate and algal prolyl 4-hydroxylases and vertebrate lysyl hydroxylase by diethyl pyrocarbonate. Evidence for histidine residues in the catalytic site of 2-oxoglutarate-coupled dioxygenases. Biochem J 286 (Pt 3): 923-927.

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.

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.

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.

OMIM (2000) www.ncbi.nlm.nih.gov/Omim.

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

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

Pajunen L, Suokas M, Hautala T, Kellokumpu S, Tebbe B, Kivirikko KI & Myllylä R (1998) A splice-site mutation that induces exon skipping and reduction in lysyl hydroxylase mRNA levels but does not create a nonsense codon in Ehlers-Danlos syndrome type VI. DNA Cell Biol 17: 117-123.

Parks CL & Shenk T (1996) The serotonin 1a receptor gene contains a TATA-less promoter that responds to MAZ and Sp1. J Biol Chem 271: 4417-4430.

Parks CL & Shenk T (1997) Activation of the adenovirus major late promoter by transcription factors MAZ and Sp1. J Virol 71: 9600-9607.

Pelham HR (1992) The Florey Lecture, 1992. The secretion of proteins by cells. Proc R Soc Lond B Biol Sci 250: 1-10.

Pelham HR (1996) The dynamic organisation of the secretory pathway. Cell Struct Funct 21: 413-419.

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

Persson A, Chang D, Rust K, Moxley M, Longmore W & Crouch E (1989) Purification and biochemical characterization of CP4 (SP-D), a collagenous surfactant-associated protein. Biochemistry 28: 6361-6367.

Peterkofsky B & Assad R (1979) Localization of chick embryo limb bone microsomal lysyl hydroxylase at intracisternal and intramembrane sites. J Biol Chem 254: 4714-4720.

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

Pihlajaniemi T & Myers JC (1987) Characterization of a pro-α 2(I) collagen gene mutation by nuclease S1 mapping. Methods Enzymol 145: 213-222.

Pihlajaniemi T & Rehn M (1995) Two new collagen subgroups: membrane-associated collagens and types XV and XVII. Prog Nucleic Acid Res Mol Biol 50: 225-262.

Pinnell SR, Krane SM, Kenzora JE & Glimcher MJ (1972) A heritable disorder of connective tissue. Hydroxylysine-deficient collagen disease. N Engl J Med 286: 1013-1020.

Pirskanen A, Kaimio AM, 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.

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

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

Pousi B, Hautala T, Hyland JC, Schröter J, Eckes B, Kivirikko KI & Myllylä R (1998) A compound heterozygote patient with Ehlers-Danlos syndrome type VI has a deletion in one allele and a splicing defect in the other allele of the lysyl hydroxylase gene. Hum Mutat 11: 55-61.

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.

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

Prockop DJ, Kuivaniemi H & Tromp G (1994) Molecular basis of osteogenesis imperfecta and related disorders of bone. Clin Plast Surg 21: 407-413.

Prockop DJ, Weinstein E & Mulveny T (1966) Hydroxylation of lysine in a polypeptide precursor of collagen. Biochem Biophys Res Commun 22: 124-128.

Puistola U, Turpeenniemi-Hujanen TM, Myllylä R & Kivirikko KI (1980a) Studies on the lysyl hydroxylase reaction. I. Initial velocity kinetics and related aspects. Biochim Biophys Acta 611: 40-50.

Puistola U, Turpeenniemi-Hujanen TM, Myllylä R & Kivirikko KI (1980b) Studies on the lysyl hydroxylase reaction. II. Inhibition kinetics and the reaction mechanism. Biochim Biophys Acta 611: 51-60.

Quinn RS & Krane SM (1976) Abnormal properties of collagen lysyl hydroxylase from skin fibroblasts of siblings with hydroxylysine-deficient collagen. J Clin Invest 57: 83-93.

Rao VH, Steinmann B, de Wet W & Hollister DW (1989) Decreased thermal denaturation temperature of osteogenesis imperfecta mutant collagen is independent of post-translational overmodifications of lysine and hydroxylysine. J Biol Chem 264: 1793-1798.

Reese CA, Wiedemann H, Kühn K & Mayne R (1982) Characterization of a highly soluble collagenous molecule isolated from chicken hyaline cartilage. Biochemistry 21: 826-830.

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

Reid KB (1983) Proteins involved in the activation and control of the two pathways of human complement. Biochem Soc Trans 11: 1-12.

Reid KB & Porter RR (1981) The proteolytic activation systems of complement. Annu Rev Biochem 50: 433-464.

Reiser KM & Last JA (1986) Biosynthesis of collagen crosslinks: in vivo labelling of neonatal skin, tendon, and bone in rats. Connect Tissue Res 14: 293-306.

Rhoads RE & Udenfriend S (1970) Purification and properties of collagen proline hydroxylase from newborn rat skin. Arch Biochem Biophys 139: 329-339.

Risteli L, Risteli J, Ihme A, Krieg T & Müller PK (1980) Preferential hydroxylation of type IV collagen by lysyl hydroxylase from Ehlers-Danlos syndrome type VI fibroblasts. Biochem Biophys Res Commun 96: 1778-1784.

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

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

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

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

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.

Ryhänen L (1975) Hydroxylation of lysyl residues in lysine-rich and arginine-rich histones by lysyl hydroxylase in vitro. Biochim Biophys Acta 397: 50-57.

Ryhänen L (1976) Lysyl hydroxylase. Further purification and characterization of the enzyme from chick embryos and chick embryo cartilage. Biochim Biophys Acta 438: 71-89.

Ryhänen L & Kivirikko KI (1974a) Developmental changes in protocollagen lysyl hydroxylase activity in the chick embryo. Biochim Biophys Acta 343: 121-128.

Ryhänen L & Kivirikko KI (1974b) Hydroxylation of lysyl residues in native and denatured protocollagen by protocollagen lysyl hydroxylase in vitro. Biochim Biophys Acta 343: 129-137.

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

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

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

Schlessinger J (1997) Direct binding and activation of receptor tyrosine kinases by collagen. Cell 91: 869-872.

Schmid C & Maraia R (1992) Transcriptional regulation and transpositional selection of active SINE sequences. Curr Opin Genet Dev 2: 874-882.

Schmid CW (1998) Does SINE evolution preclude Alu function? Nucleic Acids Res 26: 4541-4550.

Schmid TM & Linsenmayer TF (1983) A short chain (pro)collagen from aged endochondral chondrocytes. Biochemical characterization. J Biol Chem 258: 9504-9509.

Shapiro FD & Eyre DR (1982) Collagen polymorphism in extracellular matrix of human osteosarcoma. J Natl Cancer Inst 69: 1009-1016.

Shrivastava A, Radziejewski C, Campbell E, Kovac L, McGlynn M, Ryan TE, Davis S, Goldfarb MP, Glass DJ, Lemke G & Yancopoulos GD (1997) An orphan receptor tyrosine kinase family whose members serve as nonintegrin collagen receptors. Mol Cell 1: 25-34.

Smith-Mungo LI & Kagan HM (1998) Lysyl oxidase: properties, regulation and multiple functions in biology. Matrix Biol 16: 387-398.

Spiess J & Noe BD (1985) Processing of an anglerfish somatostatin precursor to a hydroxylysine-containing somatostatin 28. Proc Natl Acad Sci U S A 82: 277-281.

Steinmann B, Royce PM & Superti-Furga A (1993) The Ehlers-Danlos syndrome. In:Royce PM & Steinmann B (eds) Connective tissue and its heritable disorders. Molecular, genetic and medical aspects. Wiley-Liss, Inc, New York, p 351-407.

Strawich E & Glimcher MJ (1983) Differences in the extent and heterogeneity of lysyl hydroxylation in embryonic chick cranial and long bone collagens. J Biol Chem 258: 555-562.

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

Suokas M, Myllylä R & Kellokumpu S (2000) A Single COOH-Terminal Peptide Segment Mediates both Membrane-Association and Localization of Lysyl Hydroxylase in the Endoplasmic Reticulum. J Biol Chem 275: 17863-17868.

Sussman M, Lichtenstein JR, Nigra TP, Martin GR & McKusick VA (1974) Hydroxylysine-deficient skin collagen in a patient with a form of the Ehlers-Danlos syndrome. J Bone Joint Surg 56: 1228-1234.

Szpirer C, Szpirer J, Riviére M, Vanvooren P, Valtavaara M & Myllylä R (1997) Localization of the gene encoding a novel isoform of lysyl hydroxylase. Mamm Genome 8: 707-708.

Tanguay RL & Gallie DR (1996) Translational efficiency is regulated by the length of the 3' untranslated region. Mol Cell Biol 16: 146-156.

Teasdale RD & Jackson MR (1996) Signal-mediated sorting of membrane proteins between the endoplasmic reticulum and the golgi apparatus. Annu Rev Cell Dev Biol 12: 27-54.

Thomas D, Patterson SD & Bradshaw RA (1995) Src homologous and collagen (Shc) protein binds to F-actin and translocates to the cytoskeleton upon nerve growth factor stimulation in PC12 cells. J Biol Chem 270: 28924-28931.

Toole BP, Kang AH, Trelstad RL & Gross J (1972) Collagen heterogeneity within different growth regions of long bones of rachitic and non-rachitic chicks. Biochem J 127: 715-720.

Torre-Blanco A, Adachi E, Hojima Y, Wootton JA, Minor RR & Prockop DJ (1992) Temperature-induced post-translational over-modification of type I procollagen. Effects of over-modification of the protein on the rate of cleavage by procollagen N-proteinase and on self-assembly of collagen into fibrils. J Biol Chem 267: 2650-2655.

Tuderman L, Myllylä R & Kivirikko KI (1977) Mechanism of the prolyl hydroxylase reaction. 1. Role of co-substrates. Eur J Biochem 80: 341-348.

Turpeenniemi-Hujanen TM, Puistola U & Kivirikko KI (1980) Isolation of lysyl hydroxylase, an enzyme of collagen synthesis, from chick embryos as a 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.

Turpeenniemi TM, Puistola U, Anttinen H & Kivirikko KI (1977) Affinity chromatography of lysyl hydroxylase on concanavalin A-agarose. Biochim Biophys Acta 483: 215-219.

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.

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 Roy N, Laureys G, Versteeg R, Opdenakker G & Speleman F (1993) High-resolution fluorescence mapping of 46 DNA markers to the short arm of human chromosome 1. Genomics 18: 71-78.

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

Vogel W, Brakebusch C, Fässler R, Alves F, Ruggiero F & Pawson T (2000) Discoidin domain receptor 1 is activated independently of β(1) integrin. J Biol Chem 275: 5779-5784.

Vogel W, Gish GD, Alves F & Pawson T (1997) The discoidin domain receptor tyrosine kinases are activated by collagen. Mol Cell 1: 13-23.

Vogel W, Kanz L, Brugger W, Berndt A & Kosmehl H (1999) Expression of laminin β_2 in normal human bone marrow. Blood 94: 1143-1145.

Volkman LE (1995) Baculovirus bounty. Science 269: 1834.

von Heijne G (1986) A new method for predicting signal sequence cleavage sites. Nucleic Acids Res 14: 4683-4690.

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.

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

Weinstein E, Blumenkrantz N & Prockop DJ (1969) Hydroxylation of proline and lysine in protocollagen involves two separate enzymatic sites. Biochim Biophys Acta 191: 747-750.

Wilhelm JE & Vale RD (1993) RNA on the move: the mRNA localization pathway. J Cell Biol 123: 269-274.

Wilson R, Ainscough R, Anderson K, Baynes C, Berks M, Bonfield J, Burton J, Connell M, Copsey T & Cooper J (1994) 2.2 Mb of contiguous nucleotide sequence from chromosome III of *C. elegans*. Nature 368: 32-38.

Yeowell HN, Allen JD, Walker LC, Overstreet MA, Murad S & Thai SF (2000b) 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, Walker LC, Murad S & Pinnell SR (1992) Characterization of a partial cDNA for lysyl hydroxylase from human skin fibroblasts; lysyl hydroxylase mRNAs are regulated differently by minoxidil derivatives and hydralazine. J Invest Dermatol 99: 864-869.

Yeowell HN & Pinnell SR (1993) The Ehlers-Danlos syndromes. Semin Dermatol 12: 229-240.

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 (1999b) Prenatal exclusion of Ehlers-Danlos syndrome type VI by mutational analysis. Proc Assoc Am Physicians 111: 57-62.

Yeowell HN & Walker LC (1999a) Tissue specificity of a new splice form of the human lysyl hydroxylase 2 gene. Matrix Biol 18: 179-187.

Yeowell HN, Walker LC, Farmer B, Heikkinen J & Myllylä R (2000a) 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.