TYPE XIII COLLAGEN

Organization of the mouse gene, generation of three genetically engineered mouse lines by homologous recombination, and biochemical studies on the molecular properties of the type XIII collagen protein

ANNE LATVANLEHTO

Faculty of Medicine,
Department of Medical Biochemistry
and Molecular Biology,
Biocenter Oulu,

University of Oulu

OULU 2004

ANNE LATVANLEHTO

TYPE XIII COLLAGEN

Organization of the mouse gene, generation of three genetically engineered mouse lines by homologous recombination, and biochemical studies on the molecular properties of the type XIII collagen protein

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 Pharmacology and Toxicology, on December 2nd, 2004, at 12 noon.

OULUN YLIOPISTO, OULU 2004

Supervised by
Professor Taina Pihlajaniemi

Reviewed by
Professor Donald Gullberg
Doctor Manuel Koch

ISBN 951-42-7592-6 (nid.)
ISBN 951-42-7593-4 (PDF) http://herkules.oulu.fi/isbn9514275934/
ISSN 0355-3221 http://herkules.oulu.fi/issn03553221/

OULU UNIVERSITY PRESS
OULU 2004

Latvanlehto, Anne, Type XIII collagen. Organization of the mouse gene, generation of three genetically engineered mouse lines by homologous recombination, and biochemical studies on the molecular properties of the type XIII collagen protein

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

2004

Oulu, Finland

Abstract

Genomic clones covering the entire mouse type XIII collagen gene (*Col13a1*) were isolated, and the complete exon-intron organization was characterized. The gene was found to be about 135 kb in size and to locate in the mouse chromosome 10. Comparison of gene structures and promoter regions between man and mouse indicated high conservation between the two species.

In order to understand the biological function of type XIII collagen, a mouse line that expresses type XIII collagen with replacement of the cytosolic and transmembrane domains by a short, non-descript sequence was generated using homologous recombination. Expression of this aminoterminally altered type XIII collagen led to mild but progressive muscular atrophy in mice. The integrity of muscle cells was disturbed and the basement membrane showed areas of detachment from the sarcolemma as well as clearly altered structure at myotendinous junctions. These phenotypical changes were, nevertheless, local, since the majority of the muscle was intact. The results show the importance of the membrane anchorage of the type XIII collagen protein in adhesion and, consequently in the maintenance of muscle integrity.

To study the significance of various regions of type XIII collagen, wild-type and mutant forms of the protein were produced recombinantly in insect cells. The transmembrane domain and the adjacent region of ectodomain were found to be crucial for the formation of type XIII collagen molecules with all of the three collagenous domains in trimeric conformation. A previously characterized conserved membrane-proximal region of the ectodomain was predicted to harbour a coiled-coil conformation. This was suggested to begin in the transmembrane domain of type XIII collagen and in several other collagenous transmembrane proteins. Type XIII collagen lacking this coiled-coil sequence was correctly folded with respect to its central COL2 and carboxylterminal COL3 domains. Between them, in the NC3 domain, a second coiled-coil sequence was found, and this was suggested to function as a second association region. The second coiled-coil sequence was found to be conserved in the two other type XIII collagen-like molecules as well.

To obtain precise information about the location and level of type XIII collagen expression, a reporter mouse line synthesizing a recombinant protein with the cytoplasmic and transmembrane portions of type XIII collagen linked in-frame with the β-galactosidase enzyme was generated. The reporter mice showed high expression of type XIII collagen at neuromuscular junctions and in the periosteum of bone. Interestingly, the growth of the reporter mice was reduced at puberty. Their long bones showed a decreased diameter and impaired mechanical properties. In addition, their peripheral nerves showed areas of detachment from muscle cells at neuromuscular junctions. These results provide further evidence for the role of type XIII collagen in cell adhesion. They also show the importance of proper adhesion conducted by type XIII collagen in signaling between the extracellular matrix and cells and in the cellular response.

Keywords: association, bone, cell adhesion, collagen, extracellular matrix, genetic structure, skeletal muscle, transgenic mice

Acknowledgements

This work was carried out at the Department of Medical Biochemistry and Molecular Biology, University of Oulu, during the years 1995-2004.

I wish to express my sincere gratitude to my supervisor Professor Taina Pihlajaniemi for offering me an opportunity to do research in her excellent research group. Her guidance and encouragement during the years has been vital for the successful accomplishment of this work. All group leaders in the Department of Medical Biochemistry and Molecular Biology, especially Research Professor emeritus Kari Kivirikko, Professor Leena Ala-Kokko, and Docent Johanna Myllyharju, deserve my thanks for contributing to the scientific athmosphere and Acting Professor Peppi Karppinen for contributing to the administration in our department.

I am very grateful to Professor Donald Gullberg and Doctor Manuel Koch for their prompt and thorough review of this thesis and their constructive suggestions to improve it. Sirkka-Liisa Leinonen is acknowledged for the revision of the language of the manuscript. Acting Professor Juha Tuukkanen, Docent Raija Sormunen and Mika Ilves, M.Sc., as well as all the other experienced collaborators who contributed to this work are acknowledged.

My warm thanks go to Ari-Pekka Kvist for teaching me the basics in laboratory work, and for sharing the exhilarating moments of science and inspiring argumentations. Anne Snellman and Hongmin Tu are acknowledged for their expertise, which they have willingly shared with me. I wish to thank Anu Koski for pleasant and fruitful co-operation. I owe my gratitude to my co-workers who became my friends, namely Kaisa Passoja, Merja Niemelä, Jaana Peters and Anne Tuomisto, for making my working days as well as my free time more pleasant. Joni Mäki, Lauri Eklund and Harri Elamaa are acknowledged for the enjoyable, almost decade-long sharing of office time and lunches.

I wish to thank all the skilful staff that I had a pleasure to work with in the Department of Medical Biochemistry and Molecular Biology, especially those who shared the challenge of studying type XIII collagen. I would like to express my sincere gratitude to Sirkka Vilmi for seamless, long-lasting co-operation and excellent technical assistance and to Aila White for excellent technical assistance. Pertti Vuokila, Auli Kinnunen, Marja-Leena Kivelä, Marja-Leena Karjalainen, Seppo Lähdesmäki and Risto Helminen are acknowledged for their often so essential help in everyday matters. The personnel in

the Laboratory Animal Center are acknowledged for good care of animals and smooth co-operation.

My warm thanks are due to my family, my parents Kaarina and Erkki Latvanlehto and my brother Aki Latvanlehto, for their support and genuine caring throughout of my life. I wish to present an armful of thanks to my dear friend Nina Mällinen, who has brought joy into my life in many ways, including my precious goddaughter Ida-Maria. Above all, I thank my fiancé Mikko Heikkinen for his love and encouragement. I also wish to thank my grandparents, other relatives and all friends for the support they have shown to me.

This work has been financially supported by the Medical Research Council of the Academy of Finland, the Sigrid Juselius Foundation, the Biocenter Oulu Graduate School and Farmoksen Tutkimus- ja Tiedesäätiö.

Oulu, November 2004 Anne Latvanlehto

Abbreviations

3D	three-dimensional
ADAM	a disintegrin and metalloproteinase
αx(y)	collagen polypeptide, where x stands for the number of chain and y for collagen type marked with Roman numerals
BM	basement membrane
bp	base pair
BP	bullous pemphigoid
BSA	bovine serum albumin
C-	carboxy-
C1q	collagen-like subcomponent of the first component of complement C1
Cbfa1	core-binding factor I
cDNA	complementary DNA
C. elegans	*Caenorhabditis elegans*
CLAC-P	collagen-like Alzheimer plaque component precursor
COL	collagenous
COLyAx	human collagen gene, where y stands for collagen type and x for the number of chain
Colyax	mouse collagen gene
dpc	day post coitum
D. melanogaster	*Drosophila melanogaster*
ECM	extracellular matrix
EDA	anhidrotic ectodermal dysplasia
EDTA	ethylenediaminetetraacetic acid disodium salt
ES	embryonic stem
FACIT	fibril-associated collagens with interrupted triple helices
FISH	fluorescence *in situ* hybridization
Gly	glycine
IGF	insulin-like growth factor
Ig	immunoglobulin
kb	kilobase
kDa	kilodalton

LDL	low-density lipoprotein
MARCO	macrophage receptor with collagenous sequences
MMP	matrix metalloproteinase
MOI	multiplicity of infection
mRNA	messenger RNA
MSR	macrophage scavenger receptor
MTJ	myotendinous junction
MuSK	muscle-specific kinase
Multiplexin	multiple triple-helix domains with interruptions
N-	amino-
NC	non-collagenous
NMJ	neuromuscular junction
nt	nucleotide
PAGE	polyacrylamide gel electrophoresis
PBS	phosphate-buffered saline
PCNA	proliferating cell nuclear antigen
PCR	polymerase chain reaction
pQCT	peripheral quantitative computed tomography
PECAM	platelet-endothelial cell adhesion molecule
RACE	rapid amplification of cDNA ends
rRNA	ribosomal RNA
RT-PCR	reverse transcription-polymerase chain reaction
SDS	sodium dodecyl sulphate
SRCL	scavenger receptor with C-type lectin
SSCP	single-strand conformation polymorphism
TNF	tumor necrosis factor
tRNA	transfer RNA
UNC	uncoordinated locomotion
UTR	untranslated region
Xaa	any amino acid in Gly-Xaa-Yaa
X-GAL	5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside
Yaa	any amino acid in Gly-Xaa-Yaa

List of original articles

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

I Kvist A-P, Latvanlehto A, Sund M, Horelli-Kuitunen N, Rehn M, Palotie A, Beier D, Pihlajaniemi T. (1999) Complete exon-intron organization and chromosomal location of the gene for mouse type XIII collagen (*col13a1*) and comparison with its human homologue. Matrix Biol. 18: 261-74

II Kvist A-P, Latvanlehto A, Sund M, Eklund L, Väisänen T, Hägg P, Sormunen R, Komulainen J, Fässler R, Pihlajaniemi T. (2001) Lack of cytosolic and transmembrane domains of type XIII collagen results in progressive myopathy. Am. J. Pathol. 159: 1581-92

III ∗Latvanlehto A, ∗Snellman A, Tu H, Pihlajaniemi T. (2003) Type XIII collagen and some other transmembrane collagens contain two separate coiled-coil motifs, which may function as independent oligomerization domains. J. Biol. Chem. 278: 37590-37599. ∗Equal contribution.

IV Latvanlehto A, Koski A, Sormunen R, Ilves M, Tuukkanen J, Kvist A-P, Pihlajaniemi T. Type XIII collagen is expressed at special adhesion sites and its deficiency leads to altered properties of muscle and bone. Manuscript.

Contents

1 Introduction

The cells of multicellular organisms are surrounded by extracellular matrix (ECM). This material gives structural support and protection to cells, but it is also involved in many active processes. Cells must communicate with each other to be able to function in a coordinated manner as a whole, living organism. The main constituents of the ECM are fibrils of elastin and collagen, proteoglycans, and structural glycoproteins, such as laminin and fibronectin. The typical feature of collagens is their triple-helical structure, which results in rigidity and strength. Collagens are traditionally classified as structural proteins of ECM, but they also have a wide variety of distinct functional motifs, and therefore they participate in conducting other functions, too. Moreover, there are a large variety of proteins in addition to collagens that contain collagenous structures. The molecular structure of type XIII collagen is special, since it is a type II transmembrane protein, a feature not utilized widely by many proteins. A small group of proteins share similarity of primary structure with type XIII collagen.

This work was started by characterizing the gene structure of mouse type XIII collagen. This information was needed in the following studies, where homologous recombination was used to generate targeted mouse lines. The mouse studies aimed at elucidating the function of type XIII collagen. The mouse was chosen as a model organism since it reproduces fast and resembles human in many ways. In order to understand the biochemical features of the mutant protein, we also studied type XIII collagen at the protein level using the insect cell expression system. Knowledge of the phenotypical changes in type XIII collagen transgenic mice can be expected to help in identifying human diseases caused by mutations in the type XIII collagen gene. Mouse models may also be beneficial for testing therapies to treat diseases.

2 Review of literature

This thesis describes the characterization of the mouse type XIII collagen gene, the generation of three targeted mouse lines, and molecular studies on the type XIII collagen protein. An overview of the collagen family of proteins, including the subgroup of transmembrane collagens and other collagenous proteins with transmembrane domains, is given. The mouse was used as a model organism in part of these studies, and mutations in type XIII collagen were found to affect the integrity of bone and muscle. Therefore, some aspects of these tissues are described. Moreover, a description of milestone mouse models for muscle and bone diseases supplemented with selected mouse models for collagen genes is given.

2.1 Superfamily of collagenous proteins

Collagens are the major components of loose and specialized connective tissue. In the human body, they can contribute to 30% of the protein mass. They are produced by cells in connective tissue, and they have traditionally been classified as structural proteins of ECM. Yet, collagens harbour diverse functions in addition to their structural ones by participating in a broad range of active processes in the body. (Myllyharju & Kivirikko 2004.) Moreover, several proteolytic non-collagenous fragments of various collagens also function as independent effectors (Ortega & Werb 2002). Collagens are numbered by roman numerals in the order of discovery, and 28 collagens have been discovered up till now (Koch *et al.* 2004). In addition, more than 20 other proteins include collagenous domains (Myllyharju & Kivirikko 2004). The types I to XIX have been known and studied for a relatively long time whereas types XX-XXVIII have not been known until recently.

The basic sequence classified as collagenous is the triplet of amino acids Gly-Xaa-Yaa. The glycine residue, being the smallest, is the only one that fits into the centre when three polypeptide chains wrap parallel around each other to form the collagen helix. While the glycine residues are buried inside the helix, the Xaa and Yaa residues are exposed to the solvent. They are often prolines and thus play a role in the rigidity of the collagen helix.

Many prolines in the Yaa position are hydroxylated, contributing to the stability of collagen molecules at body temperature. The collagen subunits are called α chains, and they adopt a left-handed helix. Three α chains form the right-handed collagen superhelix stabilized by a repertoire of molecular interactions, including intra- and interchain hydrogen and disulfide bonds. The three α chains in collagen molecules may be identical or present two or three distinct polypeptides, thus being homotrimers or heterotrimers, respectively. Both ends of all collagen molecules are non-collagenous. With the exception of transmembrane collagens, they bear a signal sequence to direct the α chains into the endoplasmic reticulum. There, α chains are exposed to post-translational modification and thereafter secreted outside the cells, where molecules can be further modified. The non-collagenous domains of collagens are referred to as NC domains, while the collagenous domains are cited as COL domains. In addition to having structural roles, the collagenous portions act as ligands for collagen receptors such as integrins. They can also be considered as stalks keeping the NC domains in the right position. Collagens can assemble with themselves and with other collagen types. In addition, they have affinity to a variety of other molecules due to their charged residues. Collagens can be classified into subfamilies according to the supramolecular assemblies they form in tissues. The basic classification can be done by dividing collagens into fibril-forming and non-fibrillar collagens. The latter is a heterogeneous group of collagens characterized by interruptions in their collagenous sequences. (For references, see van der Rest & Garrone 1991, Kivirikko 1993, Prockop & Kivirikko 1995, Pihlajaniemi & Rehn 1995, Brodsky & Shah 1995, Engel & Kammerer 2000, Exposito *et al.* 2002, Gelse *et al.* 2003.) Fibrillar collagens are the classical group of structural collagens which have been studied for decades. They were discovered first, since their long continuous collagen sequence renders them resistant to proteases, and thereby ensures their persistence in tissue preparations. Only some special matrix metalloproteinases (MMPs) can digest native collagen structure, but many more proteases are able to degrade denatured collagens, also called gelatins (Gelse *et al.* 2003). Many non-fibrillar collagens share similarity of functional non-collagenous motifs with other families of ECM proteins. They are thus multimodular proteins with diverse functions. (Brown & Timpl 1995, Exposito *et al.* 2002.) The non-collagenous domains of collagens offer targets for proteolytic cleavage. In addition to that, they are needed for chain selection and assembly, often also for higher-order structures. (Exposito *et al.* 2002, Gelse *et al.* 2003.)

2.1.1 Fibrillar collagens

Fibrillar collagens aggregate to form cross-striated fibers with high tensile strength. They contain a continuous collagenous sequence of about 1000 amino acids, a highly conserved C-propeptide, and a variable N-propeptide. (Pihlajaniemi & Rehn 1995, Exposito *et al.* 2002.) The group of fibrillar collagens includes the major collagen types I and II together with the minor types III, V, and XI. The two newer ones, types XXIV (Koch *et al.* 2003) and XXVII (Pace *et al.* 2003), are also classified as fibrillar collagens, although they have interruptions in their collagenous sequences and are more

homologous with each other than with the other fibrillar collagens. The assembly of collagen type I into fibers is regulated by the formation of heterotypic fibrils with collagen types III and V. The minor collagens make up the core of these fibrils. Type I-based fibers are the basic structural constituents of many tissues, such as bone. In bone, type I collagen accounts for up to 90% of the organic mass. Type II collagen constitutes 80% of the collagen in cartilage. The diameter of type II fibrils is regulated by collagen type XI. (Gelse *et al.* 2003.) Most probably, the new collagen types XXIV and XXVII resemble the other minor fibrillar collagens. Type XXIV collagen seems to exhibit more type I (Koch *et al.* 2003) and type XXVII collagen type II fibril-like array of expression (Boot-Handford *et al.* 2003, Pace *et al.* 2003).

2.1.2 FACIT and structurally related collagens

The largest subgroup of non-fibrillar collagens are the FACITs (*f*ibril-*a*ssociated *c*ollagens with *i*nterrupted *t*riple helices). As the name implies, they have interruptions in their collagenous sequences. FACIT collagens associate with a subset of type I or II collagen fibrils in a tissue-specific manner, exhibiting more restricted expression patterns than fibrillar collagens. (Prockop & Kivirikko 1995, Gelse *et al.* 2003.) This group contains the collagen types IX, XII, XIV, and XVI (Myllyharju & Kivirikko 2004) as well as the newer types XX (Koch *et al.* 2001) and XXI (Fitzgerald & Bateman 2001, Tuckwell 2002). FACIT collagens associate with the collagen fibrils by their C-terminal collagenous domain. The non-collagenous aminoterminus points out of the fiber interacting with ECM proteins as well as with other fibrils. (Olsen 1997.) The collagens IX (Olsen 1997) and XVI (Kassner *et al.* 2003) associate with type II fibrils in cartilage. Moreover, type XVI collagen has been shown to aggregate with microfibrils in a tissue-specific manner (Kassner *et al.* 2003). Type XII and XIV collagens participate in type I fibrils in dense connective tissues, where high tensile strength is required (Gelse *et al.* 2003). Type XX collagen is most likely also to be present in such structures (Koch *et al.* 2001), while type XXI collagen is expressed by smooth muscle cells in vascular wall (Chou & Li 2002).

Collagen types XIX, XXII, and XXVI share the common features of primary structure with FACIT collagens. However, type XIX collagen has not to date been proved to associate with fibrils, whereas purified collagen XIX self-aggregates into higher-order structures (Myers *et al.* 2003). Collagen XIX is involved in muscle differentiation and it localizes in special basement membrane (BM) zones (Myers *et al.* 1997, Sumiyoshi *et al.* 2001, Myers *et al.* 2003). Type XXII collagen does not show affinity to collagen fibril extracts but shows affinity to extrafibrillar microfibrils (Koch *et al.* 2004). Yet another new collagen, namely type XXVI collagen, shares similarity of primary structure with FACIT collagens, but it does not interact with collagen fibrils, either (Sato *et al.* 2002).

2.1.3 Other non-fibrillar collagens and their special features

One of the fundamental structural proteins of BM is type IV collagen. It forms a molecular network in BM. Another similar network is formed by laminins. There are six distinct type IV collagen α chains. The $\alpha1(IV)_2\alpha2(IV)$ heterotrimers are the most abundant type IV collagen isoforms in most basement membranes, while the other type IV collagen isoforms are expressed in a more tissue-specific manner. Other network-forming collagens comprise the structurally related short-chain collagens types VIII and X, which form hexagonal lattices. In addition to such networks, similarly to type II collagen fibrils, type X collagen has been localized to microfilaments, and its expression is highly restricted to the hyaline cartilage. Heterotrimeric type VI collagen molecules with extended non-collagenous ends form beaded microfilaments present in soft tissues. Type VII collagen forms anchoring fibrils that attach the epithelial basement membrane to the underlying stroma. (For references, see van der Rest & Garrone 1991, Kielty *et al.* 1993, Pihlajaniemi & Rehn 1995, Brown & Timpl 1995, Myllyharju & Kivirikko 2001, Exposito *et al.* 2002, Gelse *et al.* 2003.)

The type VX and XVIII collagens are classified as multiplexins (*multiple* triple heli*x* domains with *in*terruptions) due to the extensive interruptions in their collagenous sequences (Pihlajaniemi & Rehn 1995). They are concentrated in basement membrane zones with similar but not equal expression patterns (Muragaki *et al.* 1995, Pihlajaniemi & Rehn 1995, Saarela *et al.* 1998, Muona *et al.* 2002). The type XV and XVIII collagens are highly homologous proteoglycans and contain a C-terminal non-collagenous domain that can be proteolytically processed (Pihlajaniemi & Rehn 1995, Ramchandran *et al.* 1999). In fact, the C-terminal endostatin fragment of type XVIII collagen has become an intensively studied topic in cancer research since it has been shown to inhibit tumor growth (O'Reilly *et al.* 1997). The most recently reported collagen is type XXVIII collagen, but its molecular properties have not been published yet (Koch *et al.* 2004). In Table 1, the chain composition, classification, and major sites of expression of each collagen type are specified, and the references for them can be found in the text.

2.1.4 Transmembrane collagens

The integral transmembranous collagens, i.e. the types XIII (Hägg *et al.* 1998), XVII (Hopkinson *et al.* 1992), XXIII (Banyard *et al.* 2003), and XXV (Hashimoto *et al.* 2002), share some common features. They all are homotrimeric type II transmembrane proteins, thus lacking a classical signal sequence to guide them into the membranes. Instead, they are inserted into the membrane by the hydrophobic residues of their membrane-spanning regions. Only 5% of the transmembrane proteins possess the type II orientation. (Singer 1990.) All of them have been shown to be shed from the membrane, resulting in stable soluble proteins (Schäcke *et al.* 1998, Snellman *et al.* 2000b, Hashimoto *et al.* 2002, Banyard *et al.* 2003, Väisänen *et al.* 2004). Nevertheless, the processing enzyme varies. The transmembrane collagens and other known collagenous transmembrane proteins are schematically presented in Figure 1.

2.1.4.1 Biosynthesis and molecular properties of type XIII collagen

Type XIII collagen is a disulfide-bonded homotrimeric transmembrane protein with a short aminoterminal, non-collagenous end inside the cell (Hägg *et al.* 1998, Snellman *et al.* 2000a). The large extracellular ectodomain consists of three collagenous domains (COL1-COL3) interrupted and flanked by non-collagenous sequences (NC1-NC4) (Pihlajaniemi & Tamminen 1990). Complex alternative splicing affects the length of the type XIII collagen polypeptide (Pihlajaniemi *et al.* 1987, Tikka *et al.* 1988, Pihlajaniemi & Tamminen 1990, Juvonen *et al.* 1992, Juvonen & Pihlajaniemi 1992, Juvonen *et al.* 1993, Peltonen *et al.* 1997). Hägg *et al.* (1998) have reported that the longest possible human variant representing all known exons is 726 amino acid residues in length. The sizes of the collagenous domains COL1-COL3 are 104, 172, and 236 residues, respectively. The non-collagenous domains NC1-NC4 are 121, 53, 22, and 18 residues in length, respectively. More specifically, the human N-terminal non-collagenous domain NC1 is composed of a 38-residue intracellular, a 23-residue membrane-spanning, and a 60-residue extracellular portion.

Human type XIII collagen synthesized by HT-1080 cells and separated by SDS-PAGE under reducing conditions results in protein bands varying in size between 85 and 95 kDa (Hägg *et al.* 1998). The apparent molecular weight is slightly higher than expected based on the cDNA sequence. This is most probably due to a high imino acid composition typical of collagens (Furthmayr & Timpl 1971). Under non-reducing conditions, disulfide-bonded trimers more than 180 kDA in molecular weight can be detected (Hägg *et al.* 1998).

The molecular properties of human type XIII collagen have been studied as recombinant proteins produced in insect cells (Snellman *et al.* 2000a, Snellman *et al.* 2000b, Tu *et al.* 2002). Collagen biosynthesis requires a variety of enzymes for the production of fully stable molecules. The key enzyme is prolyl 4-hydroxylase, which adds hydroxyl groups in proline residues at the Yaa position of Gly-Xaa-Yaa triplets (Kivirikko & Pihlajaniemi 1998). These hydroxylated prolines are needed for stabilizing collagen molecules at body temperature. Stable, hydroxylated recombinant collagen molecules are obtained when co-expressed with prolyl 4-hydroxylase (Lamberg *et al.* 1996). Similarly, in the case of type XIII collagen, stable hydroxylated homotrimeric disulfide-bonded molecules that can be secreted into the cell culture medium have been produced (Snellman *et al.* 2000a, Tu *et al.* 2002). Native triple helices are resistant to proteases, such as pepsin, trypsin, or chymotrypsin, a property that has been extensively and successfully exploited in collagen research (Bruckner & Prockop 1981). As Snellman *et al.* (2000a) have shown, pepsin digestion of the recombinant type XIII collagen protein from insect cell extracts results in three triple-helical collagenous domains. The most stable one of them is the COL2 domain.

The chain association of fibrillar collagens is driven by their non-collagenous C-propeptides in the C- to N-terminus direction (Prockop & Kivirikko 1995). The trimerization of membrane collagens may differ from that of fibrillar ones. This is due to them being anchored in the membrane in type II orientation, the N-terminus lying inside and the C-terminus outside the cell. In fact, Snellman *et al.* (2000b) have shown that the trimerization of type XIII collagen proceeds from the N-terminus to the C-terminus. This

was concluded, since the deletion of cytosolic or C-terminal sequences has no effect on trimerization, while it is blocked by deleting a 21-residue sequence next to the plasma membrane. This region is conserved between collagenous transmembrane proteins. However, this region lacks cysteine residues that could be implicated in forming inter-chain disulfide bonds to trimers. According to the cDNA sequence resolved by Hägg *et al.* (1998), the human protein has altogether 8 cysteine residues. They are distributed in pairs in the transmembrane region, at the very end of the NC1 domain, at the border of the COL1 and NC2 domains, and in the NC4 domain. The sequence similarity between the human and mouse type XIII collagens is 90% at the amino acid level. The mouse type XIII collagen has otherwise the same cysteine residues as the human protein but lacks one of the residues in the NC2 domain and harbors an additional pair at the cytosolic-transmembrane junction. Most probably, the cysteines participating in the inter-chain disulfide-bond formation are the ones at the end of the NC1 domain or/and those at the border of the COL1 and NC2 domains (Snellman *et al.* 2000a). The other cysteines thus form intra-chain disulfide bonds.

Adjacent to the sequence needed for chain association in the NC1 domain, there is a furin-type endoprotease recognition sequence RRRR. This is at the positions 105-108 in the human type XIII collagen, separated from the COL1 domain by only 13 amino acid recidues. (Hägg *et al.* 1998.) Furin is a mammalian serine protease that belongs to the family of proprotein convertases that recognize a sequence (R/K)-X_n-(R/K)↓, where n=0, 2, 4, or 6 (Seidah & Chretien 1997). Furin is located in the *trans*-Golgi network and the plasma membrane, and it participates in the maturation of precursor proteins (Nakayama 1997). Recombinant type XIII collagen produced in insect cells is found partly in the medium, and the purified soluble protein possesses an N-terminus with amino acid recidues following the furin cleavage site (Snellman *et al.* 2000b). The shedding of the ectodomain into the medium is inhibited by the action of furin inhibitor (Snellman *et al.* 2000b) as well as by heparin due to the cluster of arginine residues at the cleavage site (Tu *et al.* 2002). In rotatory shadowing electron microscopy recombinant type XIII collagen ectodomain has been shown to have a 150-nm rod-like shape with bends at the non-collagenous domains (Tu *et al.* 2002). In cultured malignant and non-metastatic mammalian cells, the type XIII collagen ectodomain is released from the full-length protein into the medium by furin both in the *trans*-Golgi and at the cell surface. The amount of released ectodomain increases in the culture medium over time, while the amount in plasma membrane remains the same. This suggests that shedding is one of the mechanisms to maintain the amount of type XIII collagen molecules in membrane. The shed type XIII collagen ectodomain appears to retain some functions, since it transiently inhibits the initiation of cell adhesion and proliferation on a vitronectin substratum. (Väisänen *et al.* 2004.) Purified, recombinant ectodomain binds *in vitro* with immobilized fibronectin, nidogen-2, and perlecan. The binding site for fibronectin is localized to the collagenous sequences. On the contrary, the binding sites for nidogen-2 and perlecan have been mapped to non-collagenous domains. (Tu *et al.* 2002.)

2.1.4.2 Type XIII collagen; an adherence junction component

Type XIII collagen is a component of focal adhesions (Hägg *et al.* 2001), adhesive contacts by which cultured cells attach to the substratum. In tissues, adherence junctions occur between adjacent cells as well as between cells and matrix. Transmembrane proteins mediate the linkage between the actin cytoskeletons of adjacent cells or between the actin cytoskeleton and ECM proteins, such as laminins, fibronectin, vitronectin, and collagens. (Geiger *et al.* 2001.) The major family of cell-matrix adhesion receptors is the integrin family of proteins. The integrins are heterodimeric transmembrane glycoprotein receptors composed of α and β subunits. To date, 18 α and 8 β chains giving rise to at least 24 different integrin dimers have been characterized. The α4, α5, α8, αIIb, and αv subunit-containing integrins bind ECM ligands that incorporate an accessible RGD sequence, namely fibronectin and vitronectin. Laminin-binding integrins contain α3, α6, and α7 chains and collagen-binding integrins α1, α2, α10, and α11 subunits. The latter harbour an additional insertion of 200 amino acid residues called the I-domain, which is involved in collagen binding. The adhesion-dependent controlled growth of cells requires integrin signaling. In cancer, integrin expression is altered and leads to changes in cell adhesion, cell migration, and finally invasion. (van der Flier & Sonnenberg 2001.)

Hägg *et al.* (2001) have used antibody stainings to shown that type XIII collagen, together with vinculin and talin, appears in focal adhesions as soon as they start to form in freshly plated fibroblasts. The artificial disassembly of focal adhesions caused by the inactivation of a small GTPase Rho causes type XIII collagen and the other focal adhesion components to relocate to the cell center. The ectopic expression of type XIII collagen in insect cells makes them more adherent to various substrata, e.g. fibronectin and vitronectin. In differentiating cultured keratinocytes, type XIII collagen concentrates at intercellular contacts, co-localizing with the adherence junction component E-cadherin (Peltonen *et al.* 1999). The I-domains of the α1 (Nykvist *et al.* 2000) and α11 (Tu 2004) integrin subunits specifically recognize type XIII collagen *in vitro*. In addition, cells with α1β1 or α11β1 integrin as their only integrin receptor can spread on a type XIII collagen ectodomain substratum, while cells bearing just α2β1 integrin cannot (Nykvist *et al.* 2000, Tu 2004).

Type XIII collagen expression has been found in all tissues studied but in relatively small amounts. By antibody stainings, type XIII collagen has been localized to the adhesive structures between cells and matrix but also between cells. Intercalated discs of the myocardium (Hägg *et al.* 2001) and keratinocytes of the skin epidermis (Peltonen *et al.* 1999) are examples of cell-cell contacts where type XIII collagen is present. In striated muscle, myotendinous junctions and costameres represent cell-matrix contact sites where type XIII collagen can be found (Hägg *et al.* 2001). In muscle, type XIII collagen expression is also reported at neuromuscular junctions, possibly at the Schwann cells that cap the nerve terminal. By antibody stainings, the type XIII collagen protein has been localized on the basal side of various epithelia but also in mesenchyme and capillaries. (Hägg *et al.* 2001.) Co-localization with several adherence junction proteins, namely E-cadherin (Peltonen *et al.* 1999), vinculin (Hägg *et al.* 2001), and β1 integrin (Sund *et al.* 2001a), has been reported, as has also exclusive localization with desmosomal components (Peltonen *et al.* 1999). In development, human type XIII

collagen has been located by *in situ* hybridization to the reticulin network of calvarial bones and, to some extent, the periosteum. In long bones, type XIII collagen localizes both in cartilage and in ossifying bone. The hypertrophic and proliferating chondrocytes of the growth plate as well as the chondrocytes of the perichondrium and articular surfaces express type XIII collagen. In ossifying long bone, type XIII is found in the reticular network between spicules and periousteum, as it was also shown for calvarial bones. (Sandberg *et al.* 1989.) In development, type XIII collagen has been found to be expressed not only by the cells mentioned above, but also by neurons of the central and peripheral nervous system (Sund *et al.* 2001a). In summary, type XIII collagen is expressed by cells of both epithelial and mesenchymal origin.

Since collagens are trimers, mutant α chains still capable of associating with other chains may produce a more severe phenotype than a simple knock-out. This phenomenon, called dominant-negative or suicide effect, is due to degradation of even the intact chains that associate with mutant ones (Horton 2003). The function of type XIII collagen was elucidated by generation of transgenic mice. Transgenic expression of truncated α1(XIII) with a central 90-residue deletion in mice leads to embryonal lethality. Embryos die at two stages of development, at 10.5 dpc due to placental defects in the fusion of chorionic and allantois membranes or at 13.5 dpc due to poorly developed heart and defective angiogenesis. Adherence junctions in the heart are not normally developed, and vascular defects can be seen in the central nervous system as well as in the placenta. (Sund *et al.* 2001b.) Interestingly, a moderate load of the same truncated transgene expression leads to an increased susceptibility to develop lymphomas in mesentericum (Sund 2001). The results obtained from this transgenic mouse line suggest a role for type XIII collagen in cell adhesion and growth control.

2.1.4.3 Collagen types XXIII and XXV

Recently proteins homologous with type XIII collagen have been identified and named as collagen types XXIII (Banyard *et al.* 2003) and XXV (Hashimoto *et al.* 2002). All three proteins share the similarity of primary structure with three collagenous domains (COL1-COL3) and four non-collagenous domains (NC1-NC4). The NC4 domain is the most homologous one between the three proteins. Like type XIII collagen, the collagen types XXIII and XXV span the cell membrane in type II orientation. Similarly to type XIII collagen, the types XXIII and XXV can be shed from the cell membrane by furin-type endoprotease. The recognition signals KLRTVR and KIRIAR in the NC1 domains are separated from the COL1 domains by 10 amino acids both in type XXIII and XXV collagens, respectively. (Hashimoto *et al.* 2002, Banyard *et al.* 2003.)

Type XXIII collagen was found by Banyard *et al.* (2003) in a prostate carcinoma cell line. Later, it was shown to be expressed by another highly metastatic cell line as well, but not by poorly metastatic or non-metastatic cell lines. RT-PCR studies indicate that type XXIII collagen is expressed in rat lung and brain. Northern blot and RT-PCR analyses indicate that type XXIII collagen transcripts 3.1 kb in size are found in human heart and brain. The human polypeptide is 532 and the rat equivalent 540 residues in length (Figure

1). The human and rat polypeptides share 91% sequence similarity with each other. Type XXIII shows 54% and 56% sequence similarity with the type XIII and XXV collagens at the amino acid level, respectively. Type XXIII can be found as a 75-kDa protein in cell surface and as a 60 kDa-soluble protein in conditioned medium. Type XXIII has been found capable of forming homotrimers. The induced expression of type XXIII collagen in metastatic tumors makes it a potential marker in diagnosis and prognosis.

Type XIII and XXV collagens share about 43% sequence similarity at the amino acid level. Human type XXV collagen undergoes alternative splicing affecting all collagenous domains or, alternatively, results in an additional variant II with a totally different C-terminus. The longest human polypeptide is 654 residues in length (Figure 1). The mouse polypeptide is 666 residues in length, showing about 90% sequence similarity with the human protein. Murine type XXV collagen is expressed as a 4.4 kb transcript and found by RT-PCR at low levels in heart, testis, and eye and at high levels in brain, specifically in neurons. *In situ* hybridization shows localization in neurons of the neocortices and hippocampus as well as other subcortices. The shed form of type XXV collagen is also called collagen-like Alzheimer plaque component CLAC, and the full-length form is called CLAC precursor or CLAC-P. CLAC is found immobilized in the senile plaques in Alzheimer patients' brain. CLAC can be detected in extracts from diseased brains as 50-, 70-, and 100-kDa bands. Analyses of cell lysates have indicated homotrimers of 240 kDa in molecular weight. The glutamic acid at the N-terminus of the shed protein, following a furin cleavage site, is modified to become pyroglutamate. This modification makes the protein more stable. Both transmembrane and shed forms bind specifically to the fibrillized form of β-amyloid peptide by ionic interactions. Type XXV collagen may thus contribute to β-amyloidogenesis and neurotoxicity in Alzheimer's disease. (For reference, see Hashimoto *et al.* 2002.)

2.1.4.4 Type XVII collagen

Type XVII collagen was initially found in the blistering skin disease called bullous pemphigoid (BP) as a 180-kDa auto-antigen (BP180) (Diaz *et al.* 1990). Later, mutations of the *COL17A1* gene have been shown to cause junctional type epidermolysis bullosa (McGrath *et al.* 1996). The expression of type XVII collagen is, however, not restricted to skin but is also found at epithelial cell-matrix junctions in many tissues. Type XVII collagen is the largest collagenous transmembrane protein consisting of 1497 amino acid residues. Unlike the intracellular domains of the other proteins in this group, the one of collagen XVIII is large, spanning about one third of its size (see Figure 1). The extracellular part is composed of 15 collagenous domains in humans and 13 in mice, the one closest to the transmembrane domain being the largest. (Snellman & Pihlajaniemi 2003.) The sequence starting from the transmembrane domain and continuing to the extracellular part most likely forms a coiled-coil helix (Balding *et al.* 1997). This region has been suggested to be involved in triple helix formation in the N- to C-terminus direction (Areida *et al.* 2001).

Type XVII collagen is a component of skin hemidesmosomes, i.e. anchoring devices that link epithelial keratinocytes to underlying BM. Collagen XVII interacts with the α6 subunit of the α6β4 integrin (Hopkinson *et al.* 1995) and the intracellular PB230 auto-antigen (Hopkinson & Jones 2000). Type XVII collagen in cultured keratinocytes is subject to shedding to produce a 120-kDa soluble triple-helical ectodomain, the form that can also be detected in the skin basement membrane zone. In contrast, the 180-kDa membrane form is only found at the lateral face of basal epithelial cells. (Schäcke *et al.* 1998, Hirako *et al.* 1998.) The shedding was originally speculated to occur by the action of furin-type endoproteases. More recent data has shown that the effect of such proproteases is indirect, as they affect the activity of ADAMs, which are actually the enzymes that shed type XVII collagen (Franzke *et al.* 2002).

Table 1. Chain composition of collagen types, classification, and major sites of expression

Type	Constituent α chains	Structure/Group	Tissue distribution
I	α1(I), α2(I)	Fibrillar-major	Bone, tendon, ligament and skin
II	α1(II)	Fibrillar-major	Cartilage, intervertebral disc, and vitreous humor
III	α1(III)	Fibrillar-minor	Co-expressed with collagen I in vasculature and skin
IV	α1(IV), α2(IV),	3D network	Most basement membranes
	α3(IV) – α6(IV)		Glomerular and alveolar basement membranes
V	α1(V),α2(V), α3(V)	Fibrillar-minor	Co-expressed with collagen I in lung, cornea and bone
	α4(V)		Nervous system
VI	α1(VI), α2(VI), α3(VI)	Microfibril	Wide tissue distribution, not bone
VII	α1(VII)	Anchoring fibril	Squamous epithelium BM zone
VIII	α1(VIII), α2(VIII)	Hexagonal lattice	Many tissues, Descemet's membrane of cornea
IX	α1(IX), α2(IX), α3(IX)	FACIT	Associated with type II fibrils in cartilage and cornea
X	α1(X)	Hexagonal lattice	Hypertrophic cartilage
XI	α1(XI), α2(XI), α1(II)	Fibrillar-minor	Co-expressed with collagen II
XII	α1(XII)	FACIT	Associated with type I fibrils in perichondrium, ligament, and tendon
XIII	α1(XIII)	Transmembrane	Many tissues at low level
XIV	α1(XIV)	FACIT	Associated with type I fibrils in many tissues
XV	α1(XV)	Multiplexin	Many BM zones
XVI	α1(XVI)	FACIT	Associated with type II fibrils in hyaline cartilage and with microfibrils in skin
XVII	α1(XVII)	Transmembrane	Skin and intestinal epithelia
XVIII	α1(XVIII)	Multiplexin	Endothelial and epithelial BM zones
XIX	α1(XIX)	FACIT-like	Rare BM zones, in developing muscle
XX	α1(XX)	FACIT	Associated with type I fibrils (?) in sternal cartilage, cornea, and tendon
XXI	α1(XXI)	FACIT	Associated with type I fibrils in vessel wall
XXII	α1(XXII)	FACIT-like	Associated with microfibrils at tissue junctions
XXIII	α1(XXIII)	Transmembrane	Heart, lung and brain, metastatic cells
XXIV	α1(XXIV)	Fibrillar-minor	Co-expressed with collagen I in bone and cornea
XXV	α1(XXV)	Transmembrane	Neurons
XXVI	α1(XXVI)	FACIT-like	Testis and ovary
XXVII	α1(XXVII)	Fibrillar-minor	Co-expressed with collagen II in cartilage and epithelia

Compiled according to the references quoted in the text.

2.1.5 Other collagenous transmembrane proteins

In addition to collagen types XIII, XVII, XXIII, and XXV, certain other type II transmembrane proteins contain collagenous sequences. These other collagenous

transmembrane proteins belong to the superfamilies of class A macrophage scavenger receptor-like molecules, the C1q/TNF superfamily (Snellman & Pihlajaniemi 2003), and the colmedins (Loria *et al.* 2004), as discussed below. All collagenous transmembrane proteins are schematically presented in Figure 1.

2.1.5.1 Class A macrophage scavenger receptor-like molecules

Out of the proteins that belong to this superfamily, the structurally related glycoproteins macrophage scavenger receptors (MSRs), macrophage receptor with collagenous structures (MARCO), and scavenger receptor with C-type lectin (SRCL) are type II transmembrane proteins (Snellman & Pihlajaniemi 2003). MSRs and MARCO are expressed by macrophages. They are involved in host defence by binding polyanionic macromolecules, damaged or apoptotic cells, and pathogens. (Exposito *et al.* 2002.) MSR types I-III are homotrimeric proteins with a short intracellular domain, a transmembrane domain, a spacer domain, an extended coiled-coil domain, a short collagenous domain, and a type-specific variable domain. The latter results from alternative splicing, and this domain is cysteine-rich in type I and III MSRs and totally lacking in the type II variant. The size of the human type I MSR is 451 residues (Figure 1). (de Winther *et al.* 2000, Snellman & Pihlajaniemi 2003.) MSRs are involved in the formation of foam cells in atherosclerosis, the collagenous domain serving as a binding site for modified lipoproteins (Acton *et al.* 1993). The proteins function as integral membrane proteins since no shedding of their ectodomains has been reported (Snellman & Pihlajaniemi 2003). The primary structure of a 520-residue human MARCO polypeptide resembles the one for MSR type I, with the exception that MARCO lacks the elongated coiled-coil motif and its collagenous domain is extended (Figure 1) (Elomaa *et al.* 1995). The bacterial-binding region is localized proximal to the cysteine-rich domain (Elomaa *et al.* 1998). The primary structure of SRCL resembles that of MRSs as well. Differently from this, the C-terminal domain is replaced by a C-type lectin/carbohydrate recognition domain present in type I SRCL and lacking in type II SRCL. The length of the human SRCL I is 742 residues (Figure 1). (Nakamura *et al.* 2001.) SRCL type I is expressed by endothelial cells, and it can bind bacteria, yeast, and oxidized LDL (Ohtani *et al.* 2001).

2.1.5.2 Ectodysplasin A

Within the C1q/TNF superfamily, one family member includes collagenous sequences and possesses type II transmembrane protein orientation, namely ectodysplasin A (Ezer *et al.* 1999). In the X-linked disease anhidrotic ectodermal dysplasia (EDA), epidermal appendages, hair follicles, sweat glands, and teeth are abnormally developed due to mutations in the ectodysplasin gene (Kere *et al.* 1996). The equivalent gene was found to be affected in the *Tabby* mouse line (Ferguson *et al.* 1997, Srivastava *et al.* 1997). Ectodysplasin is involved in mesenchyme-epithelium communication in organogenesis,

and it is also expressed by cells of other than ectodermal origin affected in EDA (Montonen *et al.* 1998, Ezer *et al.* 1999). The ectodysplasin transcript undergoes alternative splicing (Srivastava *et al.* 1997, Bayes *et al.* 1998), EDA-A1 being the longest splice variant with 391 residues (Monreal *et al.* 1998) (Figure 1). Mutations causing EDA affect all of the three functional domains of ectodysplasin A, namely the TNF homology domain (Schneider *et al.* 2001), the collagenous domain (Bayes *et al.* 1998), and a putative furin cleavage site (Chen *et al.* 2001). The latter suggests that the active protein is a shed form. Ectopic expression of the EDA-A1 splice form can rescue the *Tabby* mouse phenotype, directly demonstrating the significance of EDA-A1 (Srivastava *et al.* 2001). Moreover, injection of recombinant EDA-A1 into the blood stream of pregnant *Tabby* mice can rescue the offspring from having the affected phenotype (Gaide & Schneider 2003). This is the first indication that such maternal treatment can lead to healing of offspring.

2.1.5.3 Colmedins

The *C. elegans* 598-residue protein UNC-122 belongs to a new group of type II transmembrane proteins including collagenous and olfactomedin domains (Figure 1). UNC-122 is expressed in coelomocytes, the *C. elegans* orthologue to scavenger cells as well as at post-synaptic sites of neuromuscular junctions. Deficiency of UNC-122 leads to neuroanatomical and locomotor defects in the nematode. The colmedin protein family also includes three more proteins, namely *C. elegans* COF-2, D. *melanogaster* CG6867, and rat gliomedin. (Loria *et al.* 2004.)

2.2 Non-collagenous motifs in trimerization of collagens

Trimerization of collagens is a multi-step process, which is best known in the case of the fibrillar collagens. (Pre)procollagen or collagen α chains are guided by a signal sequence or a transmembrane domain into the endoplasmic reticulum, where the maturation of collagens takes place. Nucleation and propagation of the collagen triple helix proceed very slowly, and it can thus be anticipated that other than the collagenous sequences are needed for α chain association. The chain association of fibrillar collagens is driven by their non-collagenous C-propeptides, which are highly homologous in size and sequence. Discontinuous recognition sequences involved in the chain selection initiate the trimerization of the proα chains. The formation of inter-chain disulfide bridges enhances trimerization, but the hydroxylation of residues is absolutely necessary for the chain association. Chain folding takes place in the C- to N-terminus direction in a zipper-like manner. (Prockop & Kivirikko 1995, Brodsky & Shah 1995, Bulleid *et al.* 1996, McLaughlin & Bulleid 1998, Beck & Brodsky 1998, Myllyharju & Kivirikko 2001.)

The collagens VIII (Kvansakul *et al.* 2003) and X (Dublet *et al.* 1999) contain C1q-like globular domains and have been shown to trimerize in the absence of the collagenous

parts. Recent studies indicate that several collagenous proteins contain three-stranded coiled-coil motifs, including the C-propeptides of fibrillar collagens (McAlinden *et al.* 2003). Among the collagenous transmembrane proteins, coiled-coils have been found in human type XVII collagen (Balding *et al.* 1997), MSRs (Emi *et al.* 1993), and SRCLs (Nakamura *et al.* 2001) as well as in mouse MARCO (Beck & Brodsky 1998). Coiled-coil is a highly versatile oligomerization motif, which is involved in many processes and may consist of two to five polypeptide chains varying in size from 21 up to 1500 residues. Left-handed, parallel coiled-coils are characterized by heptad repeats (abcdefg), where the a and d residues are hydrophobic, while the others are hydrophilic. (Lupas 1996, Beck & Brodsky 1998, Burkhard *et al.* 2001.) Nucleation and folding of coiled-coils is very fast when the chains adopt a helical form and associate with hydrogen bonds and hydrophobic interactions (Brodsky & Shah 1995). Heptad repeats alone are not sufficient for oligomerization but require triggering sites with a correct salt-bridge pattern, thus making the process pH-sensitive (Burkhard *et al.* 2001) as shown for MSR type I (Frank *et al.* 2000).

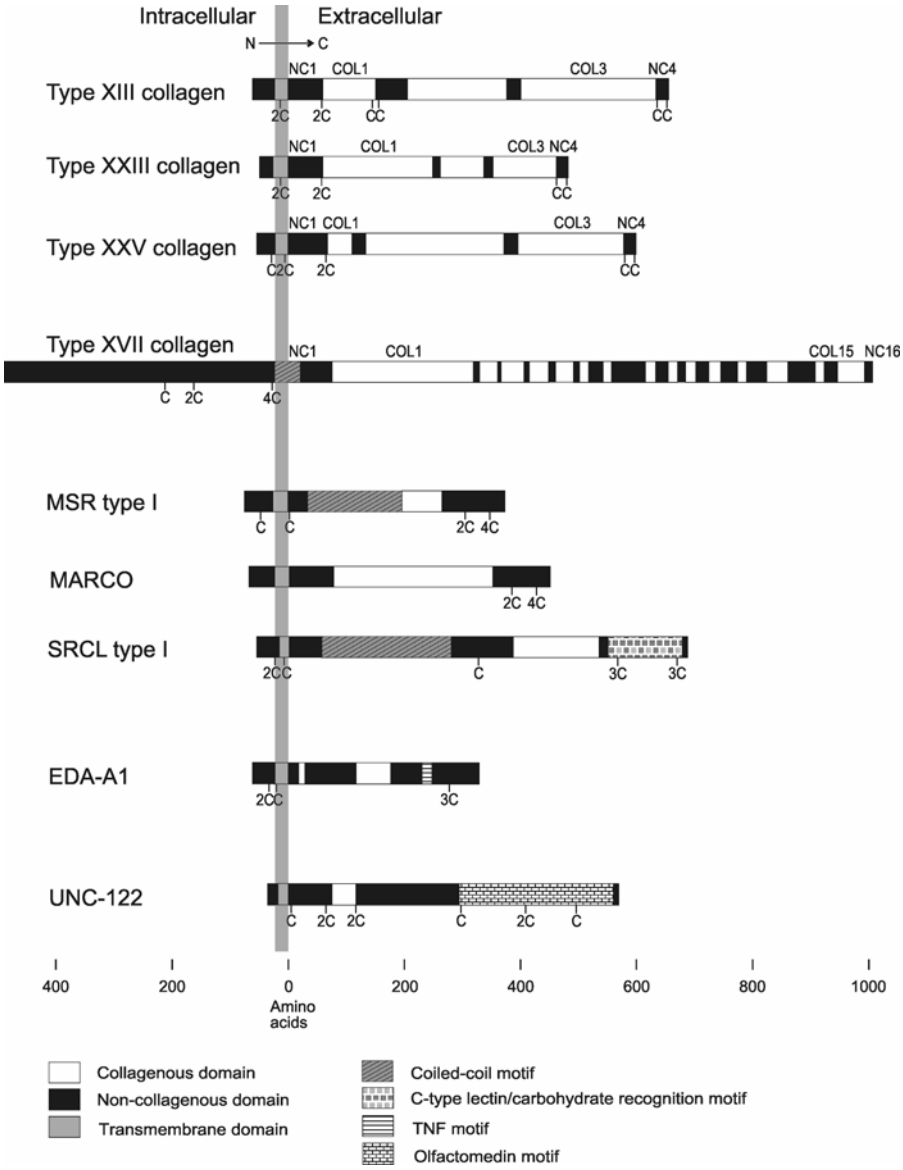

Fig. 1. Collagenous type II transmembrane proteins. The schematic representations have been outlined based on the sequences available in the National Center for Biotechnology Information (http://www.ncbi.nlm.nih.gov/) for the human type XIII collagen transcript variant I NM_005203, human type XXIII collagen NM_173465, human type XXV collagen transcript variant I NM_198721, human type XVII collagen long variant NM_000494, human macrophage scavenger receptor type I (MSR type I) NM_138715, human macrophage receptor with collagenous structure (MARCO) NM_006770, human scavenger receptor with C-type lectin type I (SRCL type I) NM_130386, human ectodysplasin-A1 (EDA-A1) NM_001399, and C. elegans uncoordinated locomotion protein (UNC-122) NM_061197.

2.3 Structure of collagen genes and transcripts

The 27 types of collagens published thus far are encoded by 42 genes (Myllyharju & Kivirikko 2004, Myllyharju 2004). The genes coding for fibrillar collagens are highly similar, with most of the exons being 54 bp, 54-9 bp, or multiples of those in length. The exons begin with a complete codon for glycine and end with a complete codon for Yaa. The ancestral exon coding for the collagenous sequence has most probably been a 54-bp exon coding for six Gly-Xaa-Yaa triplets. (Vuorio & de Crombrugghe 1990, Exposito *et al.* 2002.) Interruptions in collagenous sequences in non-fibrillar collagens are often coded by exon-intron boundaries. This is suggestive of the inactivation of the splice site signals followed by recruitment of intronic sequences into exons (Soininen *et al.* 1989, Buttice *et al.* 1990). The genes for collagens have evolved by duplication, and some of them can be found in pairs or clusters in the genome (Pihlajaniemi & Rehn 1995, Heikkilä & Soininen 1996, Sallinen *et al.* 2001). Many collagen genes have very large introns at their 5' end, exemplified by the about 600-kb first intron of *COL5A1* (Takahara *et al.* 1995). Most collagen genes are subjected to alternative splicing leading to tissue-specific products, which may have specialized functions (Boyd *et al.* 1993, Pihlajaniemi & Rehn 1995). At least six out of the collagen α chains are transcribed from two distinct promoters leading to specific regulation of the expression of the collagen variants (Nishimura *et al.* 1989, Sugimoto *et al.* 1994, Saitta & Chu 1994, Rehn & Pihlajaniemi 1995, Chou & Li 2002).

2.3.1 Human type XIII collagen gene and its products

The human type XIII collagen gene *COL13A1* spans more than 130 kb (Tikka *et al.* 1991) and is composed of at least 41, most likely 42 exons (Hägg *et al.* 1998). The last exon codes only the translation termination codon (Tikka *et al.* 1988). Tikka *et al.* (1991) have partly characterized the human gene and found that the first two introns are large, at least about 20 and more than 20 kb in size, respectively. Only some of the exons coding for the collagenous sequences are 54 or 45 bp in length, the most common exon length being 27 bp. The shortest one was found to be only 8 bp in length. All other exons begin and end with a complete amino acid codon, except the exons two and three, the junction of which codes for the first glycine in COL1. The human type XIII collagen gene was localized by *in situ* hybridization in the q22 region in the long arm of chromosome 10 (Shows *et al.* 1989).

The about 3-kb type XIII collagen mRNA undergoes complex alternative splicing. This affects the domain structures of COL1, NC2, and COL3 both in mouse (Peltonen *et al.* 1997) and in man (Pihlajaniemi *et al.* 1987, Tikka *et al.* 1988, Juvonen *et al.* 1992, Juvonen & Pihlajaniemi 1992, Juvonen *et al.* 1993) and additionally the structure of the NC4 domain in man (Pihlajaniemi & Tamminen 1990). Alternative splicing affects ten exons both in human and in mouse transcripts, although not all of the variable exons are exactly the same in the two species (Peltonen *et al.* 1997). Although alternative splicing of ten exons could, in theory, result in a large number of different transcripts, only a

limited range of variant transcripts occurring in significant amounts have been found in human cell lines and tissues (Juvonen *et al.* 1992) as well as in mouse tissues (Peltonen *et al.* 1997). Whether differentially spliced α chains fold to form stable trimers is unknown, as is also the importance of the alternative slicing of type XIII collagen transcripts.

2.4 Human diseases caused by mutations in collagen genes

Mutations in collagen genes leading to deficiency of a collagen or expression of a defective protein give rise to a wide spectrum of diseases. So far, more than 1300 mutations in altogether 23 human collagen genes have been reported, while no mutations leading to diseases have been found yet in 19 human collagen genes, *COL13A1* being one of these. (Myllyharju & Kivirikko 2004.) The majority of the mutations are glycine substitutions causing misfolding of the collagen triple helix. Most of the identified collagen mutations cause relatively rare, clinically heterogeneous diseases. One of these is osteogenesis imperfecta, which is caused by mutations in collagen I genes. Consequently, tissues rich in type I collagen are affected, bone being the major site with vulnerability to fractures due to decreased bone mass. The common disease osteoporosis may, in rare cases, result from mutations in collagen I genes. Mutations in the cartilage-specific genes coding for the collagens II, IX, and XI cause a spectrum of chondrodysplasias and osteoarthrosis. In addition, collagen IX mutations may predispose to intervertebral disc disease. (For references, see Kuivaniemi *et al.* 1991, Kivirikko 1993, Prockop & Kivirikko 1995, Myllyharju & Kivirikko 2001, Myllyharju & Kivirikko 2004, Myllyharju 2004.)

2.5 Skeletal muscle

The basic unit of contractile skeletal muscle is the sarcomere. The active contractile sarcomeric skeleton is composed of thin filaments of sarcomeric α-actin and thick filaments of myosin. Thin myofilaments are anchored at Z-discs to form myofibrils. The major intermediate filament in muscle, desmin, has been proposed to link myofibrils laterally together as well as to the muscle plasma membrane (sarcolemma). Additional sub-sarcolemmal cytoskeletons participate in linking peripheral myofibrils to the sarcolemma and, indirectly, to ECM. (Berthier & Blaineau 1997.)

Muscle tissue develops from mesoderm as cells migrate to their destination from somites, divide and fuse (Sanes & Lichtman 1999). There is a switch in integrin subtypes at the onset of terminal differentiation, α5β1 integrin giving way to α7β1 integrin concomitant with the fibronectin-rich ECM changing to a laminin-containing BM (Mayer 2003). Fibronectin promotes the adhesion of fibroblasts and is excluded from the site of myogenesis, where laminin plays a role in enhancing myogenesis. In addition to its involvement in development, BM is critical for the structure and function of muscle fibers. BM is also involved in muscle regeneration. The outer layer of the skeletal

basement membrane, the reticular lamina, incorporates fibrillar collagens. The inner layer, the basal lamina, contains two distinct networks formed by laminin and collagen IV. Both proteins are capable of self-assembly, and these independent networks are bound together by nidogens. (Sanes 2003.) Another group of important BM proteins consists of heparan sulphate proteoglycans, such as perlecan (Timpl & Brown 1996, Sanes 2003). Both the laminin and collagen IV networks are linked to the sarcolemma via interacting proteins (Sanes 2003). In muscle, the major laminin receptors are α7β1 integrin and the dystrophin-glycoprotein complex (DGC), which connect laminin indirectly to the actin cytoskeleton (Berthier & Blaineau 1997). DGC is a multiprotein complex composed of cytoplasmic, transmembranous, and extracellular proteins (Berthier & Blaineau 1997, Ervasti 2003). Laminin-2 is the major laminin in mature muscle, and it is a trimer composed of α2, β1, and γ1 chains (Gullberg *et al.* 1999).

In a motor unit, one branching nerve innervates several muscle cells, but only one nerve attaches to a muscle cell that is elongated symmetrically on both sides of the pretzel-shape synapse (Sanes & Lichtman 1999). When the muscle receives stimuli from terminal axons at the neuromuscular junction, an action potential initiates muscle contraction, finally transmitting the force into tendons. The myotendinous junction is a cell-ECM contact site, where the sarcolemma is extensively folded to increase the area to which the force is transmitted. (Berthier & Blaineau 1997.) At neuromuscular junctions (NMJ) and myotendinous junctions (MTJ), the basement membrane is structurally and functionally specialized to enable the assembly and maintenance of these structures (Timpl & Brown 1996, Berthier & Blaineau 1997, Sanes & Lichtman 1999). The basal lamina but not the reticular lamina penetrates the synaptic cleft in the NMJ and the invaginations in the MTJ (Sanes 2003).

2.5.1 Animal models for muscle diseases

In order to better understand the functions of proteins, mouse models can be used to study the consequences of mutations in a living, complex organism. Phenotypes in transgenic mice can provide ideas for identifying human diseases with a mutation in the corresponding gene. Mouse models can also be used to test potential therapies, since the phenotypes in mice often resemble those seen in affected humans. Mouse research was started by collecting spontaneous mutations. At the beginning of the transgenic era, foreign DNA was transferred into mice using viruses as carriers (Jaenisch 1976), and later micro-injection began to be used (Gordon *et al.* 1980). In 1989, a major technological enhancement was achieved by the generation of the first targeted mouse lines (Thompson *et al.* 1989). This was made possible by the ability to culture pluripotential embryonic stem cells (ES cells) and to maintain them as undifferentiated *in vitro* (Evans & Kaufman 1981, Williams *et al.* 1988). The genomic DNA in ES cells can be modified and, when returned to mouse embryos, ES cells can contribute to all tissues of the mouse. Differentiation of ES cells into germ cells thus makes the gene manipulation heritable, resulting in a generation of genetically modified mouse lines. (Thompson *et al.* 1989.)

Any disturbance in the molecular link between the cytoskeleton and the ECM may result in muscle instability. The effect depends on the importance and the unique nature of the molecule disrupted and thus on the capability for compensation by other proteins. For example, the deficiency of the major intermediate filament protein desmin leads to a muscle multisystem disorder but, surprisingly, not to embryonal lethality (Capetanaki *et al.* 1997). The results obtained by studying this mouse model led to a revised notion of the indispensability of desmin. The hallmark of muscle disorders is the mutations in the DGC protein dystrophin, which in man cause Duchanne muscular dystrophy, a severe muscle-wasting disease (Blake *et al.* 2002). The equivalent gene is mutated in *mdx* mice, leading to similar but milder symphtoms than what detected in human disease (Bulfield *et al.* 1984, Ryder-Cook *et al.* 1988). The phenotypical changes resulting from the lack of dystrophin directly demonstrates the importance of proper linkage between the cytoskeleton and ECM. In *mdx* mice, the dystrophin-related protein utrophin is up-regulated and redistributed extrajunctionally throughtout the sarcolemma, thus differing from its normal expression at neuromuscular and myotendinous junctions (Weir *et al.* 2002). In *mdx* muscle, transgenic overexpression of utrophin results in reduced muscle pathology (Tinsley *et al.* 1998). These phenomena have prompted studies to find ways to induce the expression of utrophin to treat Duchanne muscular dystrophy. In addition to structural defects in muscle, the lack of the cytoplasmic DGC protein α-dystrobrevin in mice leads to mislocation of a downstream signaling molecule (Grady *et al.* 1999). Therefore, this model shows that the function of DGC is not solely structural, but that it is also involved in signal transduction between ECM and cytosol.

Deficiency of the laminin α2 chain leads to a lack of the prominent BM component laminin-2 in muscle. Thereupon, BM becomes fragmented or absent. (Miyagoe *et al.* 1997.) Reduction in the amount of laminin-2 causes congenital muscular dystrophy in *dy/dy* mice (Xu *et al.* 1994). Laminin α2 chain null mice show congruent phenotypical changes with *dy/dy* mice, including waddling gait and twitching. In addition, these mice have notably shortened lifespan when compared to *dy/dy* mice. (Miyagoe *et al.* 1997.) Deficiency of the laminin receptor α7β1 integrin in mice causes progressive muscular dystrophy characterized by degeneration of MTJ (Mayer *et al.* 1997). β1 integrin subunit deficiency results in early embryonal lethality (Fässler & Meyer 1995), while muscle-specific knock-out of β1 leads to death at birth with highly under-developed muscles. The results are suggestive of an important role of β1-integrins in myoblast fusion (Schwander *et al.* 2003). Deficiency of the α5 integrin subunit leads to embryonal lethality as well. Studies with chimeric mice have revealed a congenital muscular dystrophy, which shows the importance of α5β1 over α7β1 integrin in muscle maturation (Taverna *et al.* 1998). Nevertheless, futher studies have shown that in addition to utrophin, α7β1 integrin is also up-regulated in dystrophin deficient muscle, and transgenic overexpression of the α7 subunit results in less symptoms (Vachon *et al.* 1997, Burkin *et al.* 2001). This thus provides a second potential approach to treat Duchanne muscular dystrophy.

Studies on mice have revealed the importance of several proteins in the formation, organization, and maintenance of the structure and function of neuromuscular junctions. The signaling pathway from the motor-neuron-derived agrin via the sarcolemmal MuSK to the muscle cell cytoplasmic rapsyn is required for neuromuscular junction formation, since deficiency of any of these proteins leads to disrupted synaptic differentiation and clustering of acetylcholine receptors (Gautam *et al.* 1999). DGC has also been shown to

be essential for the formation of the synaptic basement membrane and the clustering of acetylcholine receptors (Jacobson *et al.* 2001). Quite the contrary, deficiency of the dystrophin-related, junction-specific protein utrophin leads only to subtle derangements in the structure of neuromuscular junction (Grady *et al.* 1997, Deconinck *et al.* 1997). Three distinct α chains contribute to the formation of laminins in the basal lamina of the neuromuscular junction. They all play different roles, and similarly to its extrasynaptic function, α2 chain-containing laminin seems to be involved in the maintenance of structural intactness (Sanes 2003). Moreover, without another basement membrane component, perlecan, acetylcholinesterase does not accumulate at the synaptic cleft (Arikawa-Hirasawa *et al.* 2002).

Within the collagen family of proteins, the types VI, XIII, XV, and XIX have been found to affect muscle integrity. Type VI collagen participates in linking the basal lamina to the reticular lamina (Sanes 2003). Deficiency of type VI collagen leads to myopathy with variability in the fiber diameter and muscle regeneration (Bonaldo *et al.* 1998). The symptoms are similar to the human disease called Bethlem myopathy, known to be due to type VI collagen mutations (Jöbsis *et al.* 1996). The role of type XIII collagen in muscle is described in the articles II and IV of this thesis. Mice lacking type XV collagen as a basement membrane component develop a mild but progressive muscular disease. Knock-out mice are vulnerable to exercise-induced muscle injury, suggesting a role for collagen XV in muscle cell stabilization. (Eklund *et al.* 2001.) Another basement membrane collagen, type XIX collagen, is likely to be involved in muscle differentiation. The lack of collagen XIX results in perinatal lethality due to dysfunction of special smooth muscle and incomplete muscle transdifferentiation in the oesophagus. (Sumiyoshi *et al.* 2004.) These numerous mouse models together with the corresponding human diseases suggest a very important role for intact basement membrane as well as an operational connection of ECM with the cytoskeleton in the development muscle structure and in the maintenance of muscle integrity.

2.6 Bone

Bone is a specialized, hard, supportive connective tissue of vertebrates with an organic matrix mainly composed of type I collagen and a mineralized extracellular matrix composed primarily of calcium and phosphate salts. The mineralized portion also actively participates in the maintenance of calcium homeostasis in the body. There are two distinct major types of bone structure; spongy bone and compact cortical bone. Osteocytes are terminally differentiated osteoblast cells wrapped inside bone and responsible for the maintenance of bone structure. Osteoclasts are multinucleated cells derived from mononuclear precursor cells originating from various hemopoietic tissues, and they function in bone resorption. (Marks & Hermey 1996.) Bone marrow is involved in haematopoiesis, and it is a source of mesenchymal stem cells with a marked potential for use in various therapies (Bianco *et al.* 2001). The outer layer of cortical bone is unmineralized and called osteoid. It contains osteoblasts that are the cells typically responsible for bone matrix production. (Marks & Hermey 1996.) However, it has been

suggested that the direct transformation of a variety of tissues, including tendon, ligament, and cartilage, could also contribute to bone formation. Muscles attach to bones directly by fleshy fibres associated with connective tissue or via tendon entheses. (Benjamin *et al.* 2002.)

Skeletal structures can develop via two mechanisms; by direct intramembranous ossification or by an intermediate mechanism, endochondral ossification. Intramembranous ossification occurs during embryogenesis when mesenchymal cells differentiate directly into osteoblasts. This mechanism produces the flat bones of the skull, many facial bones, and part of the clavicle. Long bones, vertebrae, and ribs develop by endochondral ossification. When long bones grow laterally, mesenchymal cells differentiate into osteoprogenitor cells of perichondrium. Thereafter, they further mature into osteoblasts and form an osteoid and bony cover for the diaphysis. In the central part of the diaphysis, mesenchymal cells differentiate into chondrocytes and mature into hypertrophic chondrocytes. They make a framework for bone formation, after which the matrix around them mineralizes. At the ends of the long bones, there are secondary ossification centers called epiphyses. The outer surface of the epiphysis is not replaced by bone but forms articular cartilage. Long bones grow in length at the growth plates located between the diaphysis and the epiphyses and composed of chondrocytes at different stages of maturation. The diameter of long bones is regulated by the outer layer of cortical bone, periosteum, which forms new bone, and the inner layer, endosteum, that resorbs it. Cranial bones grow at the unossified borderlines of bones, called sutures. Bone homeostasis is maintained by strict regulation of ECM protein expression via neuronally or hormonally transmitted signaling molecules, growth factors, and transcription factors. (For references, see Aszódi *et al.* 2000, Olsen *et al.* 2000, de Crombrugghe *et al.* 2001, Horton 2003.)

2.6.1 Animal models for skeletal disorders

Since cartilage deposits the scaffold for endochondral bone formation, mutations in cartilage genes leading to inefficient ECM proteins naturally affect skeletogenesis. Mutations in the genes or ectopic expression of mutant α chains coding for the cartilage-specific collagens II, IX, and XI lead to chondrodysplasias and osteoarthrosis of variable severity in mice, showing phenotypes congruent with human diseases (Cremer *et al.* 1998, Aszódi *et al.* 2000, Aszódi *et al.* 2001, McLean & Olsen 2001, Helminen *et al.* 2002, Horton 2003). Simple and double knock-out studies have revealed that the collagens II and IX have distinct roles in cartilage. Type II collagen is essential for correct ECM formation and hence for normal endochondral ossification. In contrast, type IX collagen is needed for cartilage maintenance. (Aszódi *et al.* 2001.) The highly cartilage-specific type X collagen is needed for proper trabecular bone formation and is therefore essential for haematopoiesis (Gress & Jacenko 2000, Jacenko *et al.* 2002).

One of the first transgenic mouse lines established with the nuclear micro-injection technique led to embryonal lethality. The insertion site of the transgene was later found in the collagen α1(I) gene, resulting in a total transcriptional block (Schnieke *et al.* 1983).

As type I collagen is such a major component of ECM in several tissues, it is not surprising that embryos do not survive until birth. Likewise, it is obvious that in heterozygous mice the major site affected is bone, since the expression of type I collagen is highest in bone (Bonadio *et al.* 1990). The phenotypic consequences in lines transgenic or mutant with respect to the *Col1a1* or *Col1a2* genes mimic osteogenesis imperfecta with variable penetrance, and some of these mouse models can thus be used to test therapeutic strategies (Stacey *et al.* 1988, Khillan *et al.* 1991, Chipman *et al.* 1993, Pereira *et al.* 1993, Pereira *et al.* 1995, Forlino *et al.* 1999, Phillips *et al.* 2000). Mice overexpressing normal type XIII collagen develop massive bone overgrowth of bones formed by both endochondral and intramembranous ossification. Transgenic overexpression makes the mechanical properties of bone unusually good. This differs from the osteosclerotic situation in that trabecular bone is unaffected. Bone malformation develops at puberty as the mice start to move and grow. Therefore, a role for type XIII collagen in conducting signals of mechanical stress is suggested. Bone overgrowth is due to the high activity of osteoblasts, not to the decreased activity of osteoclasts. The expression of Cbfa1 (core-binding factor I) and IGF II (insulin growth factor II) is altered in type XIII collagen-overexpressing mice. (Ylönen *et al.* submitted.) The transcription factor Cbfa1 is absolutely required for osteoblast differentiation and hence for both intramembranous and endochondral bone formation in mice (Komori *et al.* 1997). Since the phenotype in type XIII collagen-overexpressing mice is so dramatic, it is not surprising that the expression of such a key player is changed, although the mechanism involved remains to be solved. IGFs are the most abundant growth factors stored in bone and produced by osteoblasts, and deficiencies in them result in delay in skeletal maturation and small bones (Mohan *et al.* 2003). Overexpression of IGF I in muscle results in large muscles, which leads to an increase in cortical bone density and area, but only the area of cancellous bone (Banu *et al.* 2003). This shows the way in which bone responds to the increased mechanical stress caused by abundant muscle mass. It has been proposed that not only the cells of sensory organs can sense mechanical forces but that adherent cells in general respond to mechanical stimuli. This is conducted in a manner similar to the response to chemical signals, *e.g.* cytokines and hormones. The response leads to changes in the size, dynamics, and signaling activity of cells as well as ECM remodeling (Geiger & Bershadsky 2001, Geiger & Bershadsky 2002). Thus, it is possible that type XIII collagen is involved in such processes.

3 Outline of the present study

When this study was started, type XIII collagen was studied mainly in man. The gene structure of the human type XIII collagen gene had been partly characterized, and its expression had been studied by *in situ* hybridization. The function of type XIII collagen was unknown, and it was therefore important to embark on studies using living organisms. Mouse was chosen as a model organism, since it has been used in biomedicine for decades and is hence well characterized. As a model organism, the mouse often presents similar phenotypic consequences of alterations in equivalent genes as man. Thus, the mouse is commonly used to characterize functions of genes and consequences of their mutations. Since no disease caused by mutations in the type XIII collagen gene was known, knowledge of phenotypical changes helps to identify such disorders. The resultant mouse models could also be useful as disease models in testing therapies for human diseases. The following goals were set for this doctoral project;

1. to isolate genomic clones to cover the entire mouse type XIII collagen gene and to characterize gene structure to enable comparisons of gene structure, chromosomal localization, and promoter region between man and mouse as well as to enable construction of targeting vectors for the generation of targeted, genetically modified mouse lines,
2. to generate mouse lines deficient of type XIII collagen, and to study the consequences of these mutations to elucidate the function of type XIII collagen in living organisms, and
3. to generate a mouse line harboring a reporter gene in-frame with type XIII collagen sequences to evaluate the expression pattern of type XIII collagen and simultaneously to study the consecuences of a lack of intact type XIII collagen protein.

Since the genetic modification introduced into the mouse line generated and reported in the original article II did not lead to a null allele but to the expression of aminoterminally altered type XIII collagen, one more goal was set, namely;

4. to study the molecular properties of the aminoterminally altered type XIII collagen at the protein level.

4 Materials and methods

A detailed description of the materials and methods used in the studies can be found in the original articles I-IV.

4.1 DNA and RNA analyses

4.1.1 Isolation and characterization of genomic clones (I)

A mouse genomic cosmid library (951303; Stratagene) was used to isolate genomic clones with a radiolabeled mouse type XIII collagen cDNA clone from the central area by screening according to standard protocols (Sambrook *et al.* 1989). To obtain isogenic clones for targeting vector construction, a 129/SvJ mouse genomic cosmid library (961301; Stratagene) was screened for the 5' and 3' ends with radiolabeled cDNA and oligonucleotide probes, respectively. Additional 129/SvJ genomic clones were obtained from the company Genome Systems, Inc. The clones were isolated by PCR screening of P1 libraries with the oligonucleotides corresponding to intron sequences adjacent to the exons 1, 7, and 15 of the mouse type XIII collagen gene. The 5' end genomic clone of human type XIII collagen was a generous gift from Dr. Karl Tryggvason (Tikka *et al.* 1991).

For restriction mapping, inserts were released from the cosmid vectors by *Not*I digestion and from the P1 vectors by *Not*I and *Sfi*I digestions. Inserts were subjected to both partial and complete digestion with the *Eco*RI, *Bam*HI, *Hind*III, and *Xba*I restriction enzymes, and the fragments were separated by CHEF-Mapper pulse field electrophoresis (Bio-Rad) using in-built algorithms for optimal DNA fragment resolution. DNA was subjected to Southern blotting and hybridization with insert-end-specific radiolabeled oligonucleotides (partial digestions) or exon-specific oligonucleotides (complete digestions) as probes. When generating the restriction map, the locations of exons were first evaluated roughly utilizing the Southern hybridization results of complete digestions,

after which the intron sizes were determined more accurately either by sequencing or by PCR using exon-specific oligonucleotides.

4.1.2 Nucleotide sequencing and sequence analysis (I-IV)

The intron-exon boundaries were sequenced directly from the cosmid and P1 clones manually using the ^{T7}Sequencing™ kit (Amersham Pharmacia Biotech) and the Cycle Sequencing kit (Amersham Pharmacia Biotech) or an ABI automatic sequencer (Perkin-Elmer). To obtain the proper sequence, some parts of the inserts in the P1 and cosmid clones were subcloned and sequenced with the ^{T7}Sequencing™ kit, as were also the clones from the 5' RACE in the original article I. In studies II-IV, all sequencing was performed automatically with an ABI-Prism DNA sequenator. (Sanger *et al.* 1977.)

The DNASIS (Amersham Pharmacia Biotech) and Chromas (Technelysium Pty Ltd) programs were used to analyze the nucleotide sequence data. Consensus sites for the binding of transcription factors were searched for in the Transcription Factor Database using the GCG Sequence Analysis Software Package, version 8.1 (Genetics Computer Group, Inc.) (Devereux *et al.* 1984). Alignments and homology searches were performed with GCG's GAP, BESTFIT, FASTA and PLOTSIMILARITY programs. The human and mouse promoters were predicted with the PROSCAN (http://bimas.dcrt.nih.gov/molbio/proscan/) and TSSG programs (http://dot.imgen.bcm.tmc.edu:1993/gene-finder/gf.html). About 5 kb of mouse, rat, and human genomic sequences covering and flanking the first exon were downloaded from the NCBI data bank (http://www.ncbi.nlm.nih.gov/Genomes/index.html) and analyzed for N- and E-box consensus sequences.

4.1.3 5' RACE and nuclease S1 protection (I)

To resolve the transcription start site for mouse type XIII collagen a blunt-ended cDNA covering the 5' untranslated region was generated using mRNA from a 18.5 dpc mouse fetus as a template with the Time-Saver cDNA synthesis kit (Amersham Pharmacia Biotech) as described (Rehn & Pihlajaniemi 1995). A linker was ligated to the blunt-ended cDNA, and PCR was performed using an antisense gene-specific primer and a sense linker-primer. Seminested PCR was carried out using the linker-primer and a nested gene-specific antisense primer. The resultant PCR products were cloned and selected by screening with a radioactively labeled probe corresponding to the 5' UTR of the mouse type XIII collagen gene. DNA extracted from positive clones was sequenced.

S1 nuclease protection was performed as described (Pihlajaniemi & Myers 1987, Sambrook *et al.* 1989). An 859-bp PCR fragment from the first exon to the 5' direction was 5' end-labeled and used as a probe in the hybridization of 20 µg of total RNA from mouse lung. After hybridization, the mixture was digested with 800 units of S1 nuclease (Boehringer Mannheim GmbH) at room temperature for 20 minutes, after which the

protected fragments were analyzed on a 6% polyacrylamide sequencing gel. Yeast tRNA was used as a negative control.

4.1.4 Determination of chromosomal location (I)

Radiolabeled primers were used to amplify the intron 25 sequences of C*ol13a1* by using genomic DNAs from a series of mouse strains as templates to test single-strand conformation polymorphism (SSCP) (Beier 1993). A portion of the reactions were denatured and separated electrophoretically on a 6% non-denaturing acrylamide sequencing gel. This was used to identify a series of polymorphisms between the *Mus musculus* (C57BL/6J) and *Mus spretus*, which were then used to analyze the DNA prepared from the BSS backcross panel (Rowe *et al.* 1994). The strain distribution pattern was analyzed using the Map Manager Program (Manly 1993).

To determine chromosomal localization by fluorescence *in situ* hybridization (FISH), the banding pattern was induced in the mouse fibroblast cells (DSM) (Lemieux *et al.* 1992, Matsuda *et al.* 1992). Biotin-labeled (Sigma Chemicals) genomic cosmid clones 19A and 6A1 (see article I) were used as probes in fluorescence *in situ* hybridization as described (Pinkel *et al.* 1988, Lichter *et al.* 1988, Tenhunen *et al.* 1995). Hybridization signals were visualized by fluorogenic avidin (Vector Laboratories). For acquisition, display, and quantification of hybridization signals, multicolor digital image analysis was used (Heiskanen *et al.* 1996).

4.1.5 Construction of targeting and expression vectors (II, III, IV)

Col13a1$^{N/N}$ and *Col13a1*$^{-/-}$ mice were generated with targeting vectors that originated from the 9.0-kb *Bam*HI fragment. This was sub-cloned from the cosmid clone 19A isolated from a 129/SvJ genomic library. It contained the promoter, the transcription initiation sequences, and the first protein-coding exon (see original article I). For the targeting construct used to generate the *Col13a1*^{N} allele, the *loxP* sequence was inserted into the unique *Not*I site in forward orientation in the 5' UTR. The selection cassette flanked by *loxP* sites was inserted into the genomic *Sfi*I site in the first intron of *Col13a1*, also in forward orientation. To construct the *Col13a1*$^{-}$ targeting vector the neomycin resistance gene was flanked by *loxP* sequences in reverse orientation. The newly made cassette was inserted to replace the 2.3-kb *Xba*I-*Sfi*I fragment containing the promoter, the transcription initiation sequences and the first protein-coding exon of *Col13a1* (see I in this thesis).

To generate the *Col13a1*^{LacZ} allele, the 7.4-kb *Bam*HI subclone containing the exon 2 and the surrounding intron sequences was used. This fragment was sub-cloned from the P1-K1 clone obtained by using commercial screening service (see original article I). A 1.2-kb *Nar*I fragment including most of the exon 2 coding sequences and the beginning of the second intron was released from the sub-clone. The *LacZ-neo*^{r} cassette was inserted

into the second exon in-frame with type XIII collagen sequences. This modification destroyed the furin endoprotease recognition sequence encoded by exon 2.

The cDNA clone E-26 (Pihlajaniemi & Tamminen 1990) covering the coding sequence for type XIII collagen, with the exception of the beginning of the translation, was used as a template for generating the human variant del1-61 α1(XIII) chain. A novel transcription start site was generated by oligonucleotides using a Site-Directed Mutagenesis kit (Stratagene). The novel cDNA was released and transported into the insect cell expression vector pVL1392 (Invitrogen).

Sequences coding for altered aminotermini of mouse type XIII collagen were obtained by reverse transcription of total RNA isolated from the skeletal muscle of a *Col13a1*$^{N/N}$ mouse using a gene-specific primer. A "long" altered aminoterminal sequence was amplified by PCR using a sense oligonucleotide primer corresponding to the nucleotides in the *loxP* sequence and a "short" altered aminoterminal sequence using a primer corresponding to the nucleotides in the first intron. The antisense primer used in both amplifications was chosen from an area coding for the COL1 domain. The long and short forms of the altered aminotermini were ligated to the mouse type XIII collagen cDNA moXIII(689) (Snellman *et al.* 2000b) to replace the original 5' sequences. The recombinant expression constructs XIII^{long} and XIII^{short} were generated by transferring the resultant inserts into the pVL1393 expression vector (Invitrogen).

4.1.6 Isolation of RNA, Northern blotting, RT-PCR, and quantitative real-time PCR (I, II, IV)

A mouse Multi-Tissue Northern blot (Clontech Laboratories, Inc.) prepared of mRNAs (2 μg/lane) isolated from various adult mouse tissues was hybridized with a radiolabeled 756-bp *Pvu*II fragment of mouse type XIII cDNA using the ExpressHyb™ (Clontech) hybridization solution.

To analyze whether *Col13a1*$^{N/N}$ mice express type XIII collagen, total RNA was isolated from cultured skin fibroblasts and muscles as described (Chomczynski & Sacchi 1987). RNA was reverse-transcribed with random oligohexamers and the moloney murine leukemia virus reverse transcriptase enzyme (Gibco BRL) at 42 °C. PCR was performed to amplify the products adjacent to the exons 16-26 (Sund *et al.* 2001b).

To resolve the 5' sequence of the type XIII collagen expressed in *Col13a1*$^{N/N}$ mice, total RNA was transcribed into cDNA using the type XIII collagen-specific antisense oligonucleotide primer chosen from the second exon with avian myeloblastosis virus reverse transcriptase (Finnzymes, Inc.) at 50°C. The 5' sequence was amplified using the same antisense oligonucleotide primer together with a sense oligonucleotide chosen from the region of the 5' UTR not excluded from the genome by gene modification. The products were sequenced.

For quantitative real-time PCR, 200 ng of total RNA from quadriceps muscles from *Col13a1*$^{N/N}$ and control mice were used as templates for a reverse transcription primed by random oligohexamers using the Taqman RT-PCR-Gold kit (PE Biosystems). The level of type XIII collagen transcripts was measured by quantitative real-time PCR analysis using

sense and antisense primers together with a fluorogenic probe, all chosen from the COL3-NC4 region (Sund *et al.* 2001a). The collagen type XIII mRNA expression levels were normalized to the levels of 18S rRNA (Majalahti-Palviainen *et al.* 2000).

Total RNA was extracted from quadriceps muscles, femoral bones, and lungs of one-month-old *Col13a1*$^{LacZ/LacZ}$ and 16- to 17-month-old *Col13a1*$^{-/-}$ mice and from 16.5 dpc fetuses originating from homozygous *Col13a1*$^{LacZ/LacZ}$ as well as *Col13a1*$^{-/-}$ matings and from wild-type C57BL/6J matings using the Tri Reagent protocol (Sigma). The total RNAs (0.5-2.5 μg) were used as templates for reverse transcription with random hexameric oligonucleotide primers and moloney murine leukemia virus reverse transcriptase (Promega) at 42 °C. The quantitative real-time PCR was performed as previously described (Sund *et al.* 2001a) with two distinct type XIII collagen-specific primer-probe pairs chosen from the exon 2 to 4 and the exon 25 to 26 areas, and the products were detected with the corresponding fluorogenic probes.

4.2 Cell culture

4.2.1 ES cell targeting and generation of Col13a1$^{N/N}$, Col13a1$^{-/-}$, and Col13a1$^{LacZ/LacZ}$ mouse lines (II, IV)

The following procedure for targeting applies to the generation of the *Col13a1*^{N} and *Col13a1*^{T} alleles, while the *Col13a1*^{LacZ} allele was generated in the Transgenic Core Facility of Biocenter Oulu according to standard methods. R1 ES cells (Nagy *et al.* 1993) were cultured as described (Fässler & Meyer 1995). The targeting vectors were linearized with appropriate restriction enzymes, and 20-50 μg was electroporated into 1.5-2 x 10^7 ES cells. The ES cells were cultured overnight, after which they were subjected to G418 sulphate selection (400 μg/ml, Life Technologies, Inc.) for 5-7 days. Individual colonies were picked, cultured, and stored. A minor portion of the cells were used for DNA extraction. DNA was analyzed for homologous recombination by restriction enzyme digestion and Southern blot hybridization with external and internal probes. Correctly targeted *Col13a1*^{N} ES cells (6.1 x 10^7) were electroporated as earlier with 40 μg supercoiled Cre plasmid to remove the selection cassette and the genomic sequences flanked by *loxP* sites. Ganciclovir selection (Syntex) was started 1-4 days afterwards, and cells were selected for five days. The DNA from the surviving ES clones was subjected to Southern blot hybridization for genotyping.

ES cells from one targeted clone were used to generate the *Col13a1*$^{T/T}$ and *Col13a1*$^{N/N}$ mouse lines, and two separate clones were used to generate the *Col13a1*$^{LacZ/LacZ}$ lines. ES cells were injected into C57BL/6J blastocysts that were then implanted into pseudopregnant foster mothers (Fässler & Meyer 1995). The ensuing chimeric mice were bred with C57Bl/6J females to establish the *Col13a*$^{T/T}$ and *Col13a1*$^{LacZ/LacZ}$ mouse lines with an inbred genetic background and with C57Bl/6J and 129/Sv females to produce *Col13a1*$^{N/N}$ mice with mixed and inbred genetic backgrounds, respectively. Homozygous

Col13a1$^{T/T}$ males were bred with C57Bl/6J females, which retain Cre recombinase activity in their mature oocytes irrespective of the *Cre* transgene transmission (Sakai & Miyazaki 1997) to result in the removal of the *loxP-neo*^{r}*-loxP* cassette. The resultant mice were bred with C57Bl/6J females to establish the *Col13a1*$^{-/-}$ mouse line with an inbred genetic background. For genotyping the mice with PCR, DNA was extracted from ear lobe or tail according to standard protocols (Sambrook *et al.* 1989).

4.2.2 Fibroblast cultures and cell adhesion studies (II)

Fibroblast cultures were established from 8-week-old mouse skin biopsies and 13.5 dpc embryos of both control 129/Sv and homozygous *Col13a1*$^{N/N}$ mice. The fibroblasts were cultured in Dulbecco's modified Eagle's medium (Biochrom KG seromed) supplemented with 10-15% heat-inactivated fetal bovine serum (Bioclear, Gibco BRL). Skin fibroblasts were used for immunofluorescent staining and embryonal fibroblasts for adhesion studies.

Cell adhesion was studied by seeding 30 000 embryonal fibroblasts in 0.2 ml of medium into the wells of 96-well plates (Beckton Dickinson Labware) and allowing them to attach. After 20, 40, or 60 minutes, the wells were emptied and washed, and the DNA content was assayed using the CyQuant Cell Proliferation Assay kit (Molecular Probes, Inc.). The amount of bound DNA from 7 to 11 replicates was measured with the Victor elisa plate reader (Wallac).

Adhesion to specific substrates was tested by plating embryonal fibroblasts in serum-free medium on Maxisorp 96-well plates (Nalge Nunc International) coated with 2 μg/cm^2 of mouse laminin, human fibronectin, mouse type IV collagen (Collaborative Biochemical Products), or heat-inactivated BSA (Boehringer Mannheim). The cells were allowed to attach for 5, 10, 20, or 40 minutes, and DNA content from five replicate measurements at each time point was assayed as above.

4.2.3 Generation of recombinant baculoviruses and recombinant protein production with baculovirus infection (III)

Recombinant baculoviruses were generated by transfecting the construct DNAs simultaneously with modified *Autographa californica* nuclear polyhedrosis virus DNA into *Spodoptera fugiperda* Sf9 insect cells using the BaculoGold transfection kit (Pharmingen). The recombinant viruses were plaque-purified and amplified as previously described (Gruenwald & Heitz 1993).

High Five insect cells were cultured as monolayers in TNM-FH (Sigma) insect cell medium supplemented with 10% fetal bovine serum (Bioclear), and when infected, in serum-free Express Five medium (Invitrogen) at 27 ˚C. The High Five cells in suspension were cultured in Express Five medium in a Certomat BS 4 shaker (B. Braun Biotech)

with 130 rpm horizontal agitation at 27 ˚C. The High Five cells were co-infected at a density of 1 x 10^6/ml in suspension and on plates of 1 x 10^5 cells/cm^2. Viruses encoding various type XIII collagen variants were used at MOI 5 together with the virus coding for both subunits of human prolyl 4-hydroxylase (4PHαβ) (Nokelainen *et al.* 1998) at MOI 1. During the 48-hour period of infection, ascorbate phosphate (80 μg/ml, Wako Pure Chemical Industries Ltd.) was added to the culture medium daily. As a control, cells were infected with the 4PHαβ virus alone.

4.3 Microscopy of cells and tissues

4.3.1 Immunofluorescence staining of fibroblasts (II)

Wild-type and *Col13a1*$^{N/N}$ mouse skin fibroblasts were cultured on glass coverslips for 1, 2, 3, 4, 9, and 12 hours followed by fixation of the cells and blocking of the background staining. The coverslips were incubated with rabbit anti-human type XIII collagen XIII/NC3-1 (Hägg *et al.* 1998), mouse anti-talin (Chemicon International, Inc.), and mouse anti-vinculin (Sigma) antibodies at appropriate dilutions, followed by extensive washing with PBS. Fluorogenic secondary antibodies were diluted according to the manufacturer's instructions (DAKO A/S) and allowed to bind to the specimens. After washing with PBS and shortly with distilled water, the coverslips were mounted on microscopic slides. Stained cells were viewed and photographed on a Leica Aristoplan microscope with appropriate filter units.

4.3.2 Immunofluorescence staining of tissues (II, IV)

Mice were sacrificed by carbon dioxide, and tissues or fetuses were embedded in TissueTec cryoprotectant (Sakura Finetek). Muscle tissues were frozen in isopentane cooled with liquid nitrogen and other tissues on carbon ice or in liquid nitrogen. Sections of 5-7 μm were cut and fixed appropriately in view of the antibodies to be used. Unspecific staining was reduced by blocking, and the sections were incubated with rabbit type XIII collagen antibodies anti-mouse moXIII/NC3 (see II in this thesis) or anti-human XIII/NC3-1 (Hägg *et al.* 1998). In addition, rabbit anti-collagen IV (Chemicon International, Inc.), rabbit anti-laminin-2 (ProGen), mouse anti-vinculin (Sigma), rat anti-tenascin-C (Sigma), rat anti-α5-integrin (Pharmingen), mouse anti-desmin (Sigma), mouse anti-PCNA (Santa Cruz), and rat anti-PECAM (Pharmingen) antibodies were used to stain muscle samples from the *Col13a1*^{N} line. For mouse monoclonal antibodies, a 1:50 dilution of goat anti-mouse IgG (DAKO A/S) was added to the blocking solution. Samples from the *Col13a1*^{LacZ} line were treated with the rabbit anti-β-galactosidase antibody (Rockland). After several washes in PBS, the sections were incubated with

fluorogenic secondary antibodies followed by an extensive wash with PBS, a short wash with distilled water, and finishing by mounting. For histological analysis, adjacent 10 μm sections were stained with hematoxylin-eosin according to standard protocols. The sections were viewed and photographed on a Leica Aristoplan microscope or on an Olympys BX51 microscope with the Olympus DP50 digital camera system.

4.3.3 β-galactosidase stainings (IV)

Mice from the *Col13a1*^{LacZ} line aged one month or at an appropriate pregnancy stage were sacrificed as described. The pieces of tissues, whole tissues, or cryo-sections were stained utilizing β-galactosidase enzyme activity as described (Gossler & Zachgo 1993). Briefly, the tissues were pre-fixed with 0.2% glutaraldehyde, 5 mM EGTA, 2 mM $MgCl_2$, 0.1 M potassium phosphate, pH 7.3, for 10 to 60 minutes at room temperature and, after extensive washing, stained overnight in washing solution supplemented with 1 mg/ml X-Gal (Duchefa) at slow rotation. The tissues were washed, post-fixed with 10% phosphate-buffered formalin, pH 7.0, and embedded in paraffin. Before embedding, bones were decalcified for 4-6 weeks in 0.5 M EDTA, pH 7.4, at room temperature. Paraffin sections of 10 μm were cut and stained for hematoxylin-eosin or eosin alone according to standard protocols. Sections were viewed on the Olympys BX51 microscope and photographed with the Olympus DP50 digital camera system.

4.3.4 Electron microscopy (II, IV)

For electron microscopy, quadriceps and gastrocnemius muscles from the *Col13a1*^{N} line were fixed with 2% glutaraldehyde in 0.1 M phosphate buffer, pH 7.4, and quadriceps muscles from the *Col13a1*^{LacZ} line with 1% glutaraldehyde, 4% paraformaldehyde in 0.1 M phosphate buffer, pH 7.3, postfixed in 1% osmium tetroxide, dehydrated in acetone, and embedded in Epon LX112 or Epon Embed 812 (Electron Microscopy Sciences), respectively. To identify neuromuscular junctions, *Col13*^{LacZ} muscle samples were first subjected to β-galactosidase staining as described above. Thin sections were examined in a Philips CM100 transmission electron microscope (Philips Electron Optics) at an acceleration voltage of 80 kV. Optical density measurements of the *Col13a1*^{N} samples were performed on the plasma membrane–basement membrane region using a CCD camera and an Electron Microscopy Menu version 2.1 from Tietz Video and Image Processing Systems GmbH. The values measured were divided by the plasma membrane density value for the same cell to normalize the results and to facilitate comparison. Moving averages of twenty data points were employed to reduce fluctuations in the measurements and to show the trend more clearly.

4.3.5 Immunoelectron microscopy (II)

Samples of gastrocnemius muscles from the *Col13a1*^{N} line were fixed with 4% paraformaldehyde in 0.1 M phosphate buffer, pH 7.4, containing 2.5% sucrose for 2 hours. Small tissue pieces were immersed in 2.3 M sucrose in PBS overnight and frozen in liquid nitrogen. Thin cryosections were first blocked and then incubated with the XIII/NC3-1 (Hägg *et al.* 1998) or moXIII/NC3 (see II in this thesis) antibodies for 60 minutes followed by incubation with protein A-gold complex (size 10 nm) for 30 minutes. (Slot & Geuze 1985.) All washings were performed in 0.1% BSA-C (Aurion) in PBS. The controls were prepared by carrying out the labeling procedure without the primary antibody. The sections were embedded in methylcellulose and examined in a Philips CM100 transmission electron microscope.

4.4 Protein analyses

4.4.1 Extraction of recombinant proteins produced in insect cells (III)

To study the type XIII collagen recombinant protein variants produced in insect cells, the cells were detached 48 hours post-infection and collected by centrifuging at 340 g for 10 minutes at room temperature. The cells were separated from the medium, and the latter was supplemented with 2 mM EDTA or Complete Protease Inhibitor Cocktail (Boehringer Mannheim). The cells were washed with PBS and collected as previously described. The cells were homogenized in 67 mM Tris-HCl, pH 7.5, 267 mM NaCl, 0.2% Triton X-100 supplemented with Complete Protease Inhibitor Cocktail. The cell lysates were incubated for 30 minutes on ice, after which they were centrifuged at 8000 g for 10 minutes at 4 ˚C. The resultant supernatants were recovered and the precipitates were dissolved in 1% SDS. Samples from the medium, the cell lysate, and the cell precipitate fractions were separated by SDS-PAGE analysis followed by Western blotting. The percentage of secreted protein was estimated using the Quantity One Quantitation Software (Bio-Rad).

4.4.2 Western blotting (III)

Recombinant protein samples were electrophoretically separated on 6 to 12% SDS-PAGE gels under reduced and/or non-reduced conditions followed by Western blotting. One or more of the following antibodies were used; the rabbit anti-mouse type XIII collagen antibodies moXIII/NC3 (see original article II) and XIII/NC1-Q610 (Hägg *et al.* 2001) and the rabbit anti-human type XIII collagen antibodies XIII/NC2-55 (Snellman *et al.*

2000a), XIII/NC3-1 (Hägg *et al.* 1998), and XIII/NC4-SO (Snellman *et al.* 2000a). The signals were detected with enhanced chemilumenescence (Amersham Biosciences).

4.4.3 Purification of type XIII collagen recombinant proteins (III)

To analyze the molecular properties of the recombinant type XIII collagen variant protein $XIII^{N\text{-}short}$, High Five cells in 500-ml suspension were separated from the medium 48 hours post-infection by centrifuging at 340 g for 10 minutes at room temperature. The conditioned medium was supplemented with 2 mM EDTA. To remove the debris and viruses, the medium was further centrifuged at 40 000 g for 45 minutes at 4 °C. The medium was then applied to a Resource S column (Amersham Biosciences) and eluted using a gradient program on ÄKTA explorer 10 (Amersham Biosciences). The fractions were analyzed by Western blotting using the antibody XIII/NC3-1 (Hägg *et al.* 1998), and those containing $XIII^{N\text{-}short}$ were concentrated to 1 ml. This was further purified on a Sephacryl S 500 column (Amersham Biosciences). The fractions containing most of $XIII^{N\text{-}short}$ were then concentrated to 0.5 µg/µl of total protein and subjected to N-terminal sequencing and pepsin digestions.

For the sequencing of $XIII^{N\text{-}long}$ proteins in cell lysate, infected insect cells from a 100 ml suspension culture were harvested as above. They were washed with 10 ml PBS twice, homogenized with 10 ml of PBS, and incubated for 30 minutes on ice. The debris and insoluble molecules were removed as described above, and the supernatant of the cell lysate was separated out in steps using a HiTrap Q column (Amersham Biosciences) by increasing the concentration of NaCl in PBS. The fractions containing $XIII^{N\text{-}long}$ were concentrated to 0.5 ml and subsequently precipitated with 75% cold ethanol for 30 minutes. The ethanol precipitates were dissolved in SDS-PAGE sample buffer for N-terminal sequencing.

For the sequencing of $XIII^{N\text{-}long}$ proteins in medium, the medium from the infection described above was centrifuged at 40 000 g for 45 minutes at 4 °C and loaded onto a HiTrap Q column (Amersham Pharmacia) followed by step-by-step elution with increasing concentration of NaCl. The fractions containing $XIII^{N\text{-}long}$ protein were combined and dialyzed against 20 mM Tris, 0.15 M NaCl, 2 mM EDTA, pH 7.4, and loaded onto a Resource Q column (Amersham Pharmacia) at 4 °C. The elution was performed by means of a programmed gradient with ÄKTA Explorer 10. The elution fractions containing $XIII^{N\text{-}long}$ protein were concentrated to 1ml and further separated by Superdex 200 (Amersham Pharmacia) in ÄKTA Explorer 10. The fractions containing $XIII^{N\text{-}long}$ protein were concentrated and subjected to N-terminal sequencing.

4.4.4 Pepsin digestions of recombinant proteins and N-terminal sequencing (III)

To test whether the recombinant proteins were in trimeric conformation, the crude cell lysate supernatants not supplemented with any protease inhibitor were freshly digested with 0.1 to 0.15 mg/ml pepsin (Boehringer Mannheim) for 2 to 5 minutes at room temperature. The digestion products were separated by 12% SDS-PAGE, followed by Western blotting with a broad range of the type XIII collagen-specific antibodies presented in the paragraph 4.4.2.

The XIII$^{N\text{-short}}$ protein (40 μg) partly purified from conditioned medium was digested with 0.8 μg pepsin (50:1 w/w) for 2 minutes at room temperature. The XIII$^{N\text{-short}}$ digestion products, 10 μg of the total XIII$^{N\text{-short}}$ protein, as well as protein samples from the purifications of XIII$^{N\text{-long}}$ in lysate and in medium were separated by 10 to 12% SDS-PAGE and electroblotted onto a ProBlott^{TM} membrane (Applied Biosystems). The resultant stained bands were cut for N-terminal protein sequencing analysis using a 492 Procise^{TM} protein sequencer (Applied Biosystems).

4.4.5 Generation of polyclonal antibodies against mouse type XIII collagen (II)

To generate a rabbit polyclonal antibody by a commercial supplier (Innovagen), a synthetic peptide (DYNGSINEALQEIRTL) corresponding to the mouse NC3 domain (Hägg *et al.* 1998) was used as an antigen. The resultant moXIII/NC3 peptide antibody was affinity-purified using the above peptide (Hägg *et al.* 1997).

4.5 Physiological studies

4.5.1 Treadmill experiment (II)

Wild-type and *Col13a1*$^{N/N}$ male mice aged 8 to 10 weeks were made to run on a motor-driven treadmill with 6 degree uphill tracks at a speed of 8.5 m x min^{-1} for six hours with two 20-minute resting periods. The runners were sacrificed together with unexercised control and *Col13a1*$^{N/N}$ male mice 48 hours after the exercise. β-glucuronidase activity was measured from the quadriceps muscle (Barrett 1972), and both gastrocnemius and quadriceps muscles were subjected to histological analyses.

4.5.2 Measurements of body weight and bone length (IV)

Mice from the *Col13a1*$^{-}$ and *Col13a1*^{LacZ} lines were weighed at the age of one, three, and 16 to 17 months. The lengths of the femoral and tibial bones in the female mice from the *Col13a1*^{LacZ} line were measured at the age of three months using a caliber.

4.6 Computational analyses

4.6.1 Statistical analyses (II, IV)

Statistical analyses were performed with either SigmaStat (Jandel Scientific, SPSS Science) or Excel (Microsoft Corp.). The data was analyzed for variation with the f-test and for statistical significance with Student's t-test. The p-values for difference were considered significant as follows; when $p < 0.05$ (∗), $p < 0.01$ (∗∗) and $p < 0.001$ (∗∗∗).

4.6.2 Amino acid sequence analysis and secondary structure prediction (III)

Sequence alignment of collagens XIII, XXIII, XXV, and XXVI was performed by the CLUSTAL W method (Thompson *et al.* 1994) and compiled into a figure using BOXSHADE at ch.EMBnet.org (http://www.ch.embnet.org/software/BOX_form.html). Coiled-coil predictions for the collagen types XIII, XVII, XXIII, XXV, and XXVI, MARCO, and EDA were made using the COILS program, version 2.1 (http://npsa-pbil.ibcp.fr/cgi-bin/npsa_automat.pl?page=npsa_lupas.html) (Lupas *et al.* 1991). The NCBI database accession numbers for the collagen types XIII, XVII, XXV, and XXVI, MARCO, and EDA are AJ293624, NM_000494, AF293340, AB085837, NM_006770, and NM_001399, respectively. The positions of heptad repeats in the NC1 and NC3 domains of type XIII collagen were predicted according to the COILS program.

4.7 Peripheral quantitative computed tomography measurements of bone (IV)

Female mice from the *Col13a1*^{LacZ} line aged three months were sacrificed, and their femoral and tibial bones were prepared in PBS. The bones were scanned with XCT 960 (Stratek) peripheral quantitative computed tomography (pQCT) at mid-diaphysis as

previously described (Jämsä *et al.* 1998). Bone mineral content, cortical bone mineral density, cross-sectional area, and polar moment of inertia were measured.

5 Results

5.1 Characterization of the mouse type XIII collagen gene (I)

A total of 7 overlapping genomic clones were obtained, covering the entire mouse type XIII collagen gene. Cosmid clones found from libraries covered the exons 14-42 and the 3' flanking sequences as well as the first exon, part of intron 1, and the 5' flanking sequences. The rest of the type XIII collagen gene was obtained by utilizing a commercial PCR screening service, where type XIII collagen gene-specific primers were used. The clone P1-K1 covered the exons 1-7, and the clone P1-K4 the exons 3-39.

To construct a restriction map for the type XIII collagen gene, inserts were released from the cosmid clones, after which the DNAs were partly or completely digested with the additional restriction enzymes. The size and order of the fragments as well as the approximate locations of the exons were determined by Southern blotting and hybridization with insert-end-specific and exon-specific oligonucleotides. More accurate intron sizes were determined by amplifying all the intron areas by PCR using exon-specific primers except for the introns 1, 2, 3, 7, 18, and 39, which were too large for amplification by PCR. The lengths of the introns were found to vary between 90 bp and 40 kb. The mouse type XIII collagen gene turned out to cover a 135-140 kb region in the genome.

The mouse type XIII collagen coding sequences were found to be distributed among 42 exons, the sizes of which range between 8 and 836 bp. The sizes of exons coding for purely collagenous sequences vary from 8 to 144 bp, but most of them are 27 bp in length. The exon-intron boundaries were sequenced, and they all conform to the consensus sequence AG-exon-GT. With the exception of the exons 2 and 3, all exons start and end with a complete codon. The last exon, exon 42, encodes only the translation termination codon. The human genomic clones described by Tikka *et al.* (1991) did not cover all exons, and the scanning of the NCBI data bank revealed a sequence that contained the lacking 5' region, its accession number being U82211.

5.2 Definition of the transcription start site of *Col13a1* and comparison of the promoter and 5´ flanking regions of the mouse and human genes (I)

As a result of the cloning of mouse cDNAs and the identification of a human cDNA clone, the initiation of the translation of type XIII collagen was found to begin at an ATG further upstream than it what was initially thought (Hägg *et al.* 1998). The transcription start site in the human EST clone R25685 locates 537 bp in the 5´ direction from the new upstream ATG. According to nuclease S1 mapping, the main site for the initiation of transcription locates 500 nucleotides upstream from this ATG codon in the mouse type XIII collagen gene, and several minor initiation sites are located between 470 and 525 bp upstream. Six 5' RACE cDNA clones were found to end at the same position, namely 548 bp before the initiation ATG.

A search for potential promoters with the PROSCAN v.1.6 and TSSG programs pointed at an identical position in the human and mouse genes. The combined data from both analyses suggest that the mouse promoter locates between -864 and -568 bp upstream of the initiation ATG. The 5' flanking region before the initiation of transcription contains no traditional TATAA motif, but there is a modified TATA-like box at the position –601 bp. The extreme 5’ cDNA clone found by 5’ RACE starts 53 bp downstream of the proposed TATA box. Furthermore, the 5' flanking region contains numerous consecutive GC boxes. The analyses suggest that the promoter in the human type XIII collagen gene lies between -905 to -605 bp in the 5´ direction from the initiation ATG with a TATA box at -636 bp. Comparison of the promoter regions with the Transcription Factor Data Base led to the identification of consensus motifs for several transcription factors in both genes.

To enable comparison of the human and mouse promoter and 5´ flanking areas, 4932 bp of the mouse and 2642 bp of the human type XIII collagen gene were sequenced and compared. The optimally aligned sequences plotted with PLOTSIMILARITY show the homology between the species to be high in the about 550 bp 5´ UTR and in the adjacent apparent promoter region of about 350 bp, but low further upstream. The predicted promoter areas and the 5’ untranslated sequences show 77.13% similarity between species.

5.3 Chromosomal localization of *Col13a1* (I)

To map the location of *Col13a1* in the mouse genome, SSCP and FISH analyses were performed. Primers corresponding to the intron 25 of *Col13a1* were found suitable for the identification of a SSCP between inbred mouse strains. *Col13a1* mapped to chromosome 10 with a likelihood score of 25 between microsatellite markers as follows: D10Mit5 – (2.3±1.6 cM) – *Col13a1* – (3.4±1.9 cM) – D10Mit15. To determine the chromosomal localization with FISH, two genomic cosmid probes, one from the 5’ (19A) and the other from the 3’ (6A1) end of the gene, were hybridized on metaphase chromosomes derived

from mouse fibroblast cells. Out of the 48 metaphase spreads 42 with the 19A probe and out of the 30 hybridizations 27 with the 6A1 probe showed a specific location on chromosome 10, band B4.

5.4 Generation of the *Col13a1*$^{N/N}$ mouse line (II)

When generating the *Col13a1*$^{N/N}$ mouse line, the aim was to create mice null for type XIII collagen. This was done by deleting most of exon 1, which encodes the 96 extreme N-terminal amino acid residues of the α1(XIII) chain, including the cytosolic and transmembrane domains. The targeting vector was constructed in such a way that one *loxP* site was inserted into the *Not*I site in the 5' UTR and the selection marker cassette flanked by *loxP* sites into the first intron of the mouse genomic sequences. Homologous recombination was identified in several ES cell clones, two of which were subjected to exclusion of all *loxP* flanked sequences by Cre-mediated recombination. One ES cell clone with a C*ol13a1*^{N} allele was used to generate chimeric males that transmitted the mutant allele to their progeny.

To confirm that the gene manipulation leads to a loss of type XIII collagen, mRNA and protein expression was studied. Surprisingly, the homozygous mouse tissues stained positively with a type XIII collagen antibody, and the central region of α1(XIII) mRNA was found to be amplified by RT-PCR. Based on these findings, the 5´ sequences of the type XIII collagen transcripts in *Col13a1*$^{N/N}$ mice were studied by RT-PCR. Sequencing of the products revealed a new 5' end that included the *loxP*-site and extended 155 bp to the first intron, after which the sequence continued into the second exon. The first intron contains a cryptic splice site capable of serving as a donor for splicing to the second exon. The new 5' sequence has an ATG codon in the *loxP* sequence and another in the first intron, both in-frame with the exon 2-derived sequences. Thus, contrary to our aims of generating a null allele, the transcripts derived from the *Col13a1*^{N} allele resulted in a synthesis of aminoterminally altered type XIII collagen molecules. The lacking 96 amino acid residues encoded by the first exon were replaced by either 65 or 11 new residues, depending on which of the two potential initiation codons was used. The aminoterminally altered type XIII collagen expressed in the *Col13a1*$^{N/N}$ mice lacked the cytosolic and transmembrane domains and the first 37 residues of the ectodomain. The *Col13a1*$^{N/N}$ mice were fertile and normal in appearance. They were born in a Mendelian ratio, and their life-span was normal.

5.5 Studies on fibroblasts expressing aminoterminally altered type XIII collagen (II)

Skin fibroblast cultures were established to explore the properties of aminoterminally altered type XIII collagen molecules. With antibody stainings, aminoterminally altered

type XIII collagen was found to be transported to the focal adhesions similarly to authentic collagen. The adhesion of fibroblasts derived from 13.5 dpc embryos was studied by plating them for 20, 40, and 60 minutes and measuring the number of adherent cells. Cells derived from *Col13a1*$^{N/N}$ embryos demonstrated an 8 to 15% lower adherence to culture plates than controls, a difference that was statistically significant. The adhesion of *Col13a1*$^{N/N}$ cells was reduced on a type IV collagen substratum, but not on fibronectin, laminin, or BSA.

5.6 Morphological changes in the skeletal muscle of the *Col13a1*$^{N/N}$ mice (II)

Systematic examination of gross anatomy and histological examination revealed changes only in the skeletal muscle of *Col13a1*$^{N/N}$ mice. Some of the muscle fibers were non-uniform in size, had a wavy sarcolemma, and were loosely attached to the adjacent fibers. Stainings with type IV collagen and laminin-2 antibodies suggested an infirm, fuzzy basement membrane in *Col13a1*$^{N/N}$ muscle. A more detailed electron-microscopic analysis was performed, and the mutant muscle showed vacuolization and disorganization in the myofilament and z-band architecture as well as accumulation and enlargement of mitochondria. The sarcolemma and the adjacent basement membrane showed a disorganized, fuzzy structure compared with that seen in corresponding wild-type samples. This was particularly evident at the myotendinous junctions. These abnormalities were progressive, since they were detected more frequently and were more pronounced in older mice. The optical density of the extracellular matrix immediately adjacent to the sarcolemma was measured in electron-microscopic samples. A transparent interphase of 10-20 nm was observed between the plasma membrane and the basement membrane in the controls, while in the *Col13a1*$^{N/N}$ mice this space was opaque. The various muscle findings were local, and most parts of the muscle were intact.

Immunofluorescence staining with antibodies detecting type XIII collagen revealed in the *Col13a1*$^{N/N}$ mice a clear staining that was somewhat fuzzier but otherwise comparable to that seen in the controls. In immunoelectron microscopy with type XIII collagen antibodies, the control samples contained gold particles mainly in close contact with plasma membranes. In contrast to that, most of the staining in *Col13a1*$^{N/N}$ muscle located in the extracellular space, and only some gold particles were detected at the plasma membrane or inside the cell. Furthermore, the mutant muscles showed areas where the basement membrane was detached and contained some gold particles. Immunoelectron microscopy confirmed the findings obtained in cultured fibroblasts. The aminoterminally mutant type XIII collagen molecules are transported out of the cell, and they are located in the matrix in the vicinity of the plasma membrane but not embedded in these structures. The amount of type XIII collagen transcripts in muscle was measured by quantitative real-time PCR analysis with the primers from the COL3-NC4 area. This measurement confirmed the equal levels of expression of the type XIII collagen alleles in wild-type and mutant mice. All these findings suggest that the aminoterminally altered type XIII collagen is expressed at a level comparable to intact molecules. They also imply

that the expression of aminoterminally altered type XIII collagen protein or a lack of intact molecules cause disturbances in the integrity of skeletal muscle.

5.7 Exercise-induced muscle damage in *Col13a1*$^{N/N}$ mice (II)

Due to the histopathological signs of a muscular disorder, the susceptibility of muscles to exercise-induced damage was tested. Male *Col13a1*$^{N/N}$ mice and age- and sex-matched wild-type mice were subjected to a running test on a motor-driven treadmill. Some of the *Col13a1*$^{N/N}$ mice began to show signs of exhaustion during the exercise, and the experiment had to be stopped somewhat prematurely. The mice were sacrificed two days later, and histological sections of the quadriceps and gastrocnemius muscles were examined for possible muscle damage. The extent of muscle damage varied between individuals, but the mutant runners generally had more numerous fibers undergoing degeneration and more intensive inflammation. Immunostaining for tenascin-C, a marker of inflammation, revealed stronger signals in the exercised mutant mice compared with the exercised controls. β-glucuronidase activity in the quadriceps muscle was used as a quantitative measure of injury. This activity was significantly higher in the *Col13a1*$^{N/N}$ runners than in the control runners, with absolute values of 2.76±0.91 vs. 1.39±0.32 μmol s^{-1} kg^{-1} protein (p<0.001). The increase in β-glucuronidase activity between the exercised and unexercised *Col13a1*$^{N/N}$ mice was 311%, compared with 190% between the control mice.

5.8 Secretion of aminoterminally altered and deleted type XIII collagen molecules expressed in insect cells (III)

To study the type XIII collagen molecules synthesized in *Col13a1*$^{N/N}$ mice, the altered mouse protein variants were produced in insect cells. In the XIII$^{N\text{-}long}$ and XIII$^{N\text{-}short}$ proteins, the first 96 residues of the mouse α1(XIII) chain were replaced with a 65- or 11-residue new aminoterminus, respectively. The human deletion variants del1-63 and del1-83, encoding human α1(XIII) chains lacking the first 61 or 83 residues, respectively, were also produced in insect cells. The conditioned medium was collected, and the proteins were extracted from infected insect cells, the remaining precipitates being solubilized in 1% SDS. The volumes of the cell and medium fractions were adjusted to correspond to the same cell number, and the samples were fractionated on denaturing SDS-PAGE gels under reducing or non-reducing conditions and analyzed by Western blotting. In serum-free medium, about 20% of the del1-38 protein used as a control was secreted into the medium through proteolytic cleavage by a furin-type protease. Slightly larger portions of the protein, namely about 50% of XIII$^{N\text{-}short}$ and 40% of del1-61, and somewhat smaller portions, 15% of XIII$^{N\text{-}long}$ and del1-83, were found to be secreted into the conditioned medium.

Sequencing of two bands detected with Western blotting in the $XIII^{N\text{-}long}$ cell fraction indicated a 90 kDa band having the aminoterminus MLYEVIRSLE predicted for intact $XIII^{N\text{-}long}$. The other, 80 kDa band represented a degradation product of the full-length $XIII^{N\text{-}long}$ α chain with the aminoterminus at residue 147 in the mouse type XIII collagen protein. Sequence analysis of the two bands observed in the $XIII^{N\text{-}long}$ medium indicated that the N-terminus of the higher molecular weight band represented full-length $XIII^{N\text{-}long}$. The smaller molecular weight band had an N-terminus of E^{107}APKMSPGCN (residue 107 of the mouse α1(XIII) chains) representing the sequences preceding the predicted furin cleavage site, indicating that the $XIII^{N\text{-}long}$ chains were cleaved at the predicted furin site in the manner previously described for the human del1-38 α chains. Only one band was observed in the medium of cells synthesizing the $XIII^{N\text{-}short}$ α chains, and identification of its N-terminal sequence indicated that it was also cleaved at the furin site.

5.9 Formation of disulfide-bonded trimers of type XIII collagen deletion variants (III)

The deletion variant del1-38 contains all cysteine residues found in the full-length human polypeptide, and it forms disulfide-bonded homotrimers similarly to full-length human α chains. There, interchain disulfide bonds are found in the NC1 domain and possibly in the COL1 and NC2 domains, while the cysteine residues occurring in the NC4 domain form intrachain bonds. (Snellman *et al.* 2000a.) Western blotting of non-reduced samples revealed disulfide-bonded trimers both in the cell supernatant and in the medium of the del1-38 variant. The $XIII^{N\text{-}short}$ and $XIII^{N\text{-}long}$ α1(XIII) chains lack the NC1 association domain and the cysteine residues in the transmembrane domain. Instead, $XIII^{N\text{-}long}$ contains two new cysteine residues in its unique sequences. The $XIII^{N\text{-}long}$ samples were found to contain trace amounts of trimers in the cell extracts, but only monomers could be visualized in the medium. No disulfide-bonded trimers could be detected in the cells or medium in the case of the $XIII^{N\text{-}short}$ α1(XIII) chains. Only minute amounts of trimers were detected in the cell extract of the del1-61 variant lacking the cytosolic and transmembrane domains, and only monomers were detected in the medium. Moreover, the del1-83 α chains, which practically corresponded to the $XIII^{N\text{-}long}$ and $XIII^{N\text{-}short}$ α chains in terms of type XIII collagen sequences, did not form disulfide-bonded trimers.

Although the modified α1(XIII) chains did not form disulfide-bonded trimers, they could form triple-helical trimers lacking interchain disulfide bonds. This may result from association and trimerization through other than the sequence shown to be important for the chain association (Snellman *et al.* 2000b). The ability of the modified chains to form stable triple-helical domains was assessed enzymatically. The pepsin-digested del1-38 α1(XIII) variant with three triple-helical collagenous domains was used as a control. The digested soluble proteins in the other cell supernatants were compared to it by Western blotting with various type XIII collagen-specific antibodies. The del1-38 sample resulted in four pepsin-resistant fragments presenting NC1 domains. Some faint bands could be detected during the digestion of the del1-61 and del1-83 samples, but once the digestion was complete, no bands representing the COL1 domain could be detected in any of the

$XIII^{N\text{-}long}$, $XIII^{N\text{-}short}$, del1-61, or del1-83 samples. On the contrary, almost identical patterns of pepsin-resistant fragments representing the COL2 and COL3 domains could be detected in all forms of α1(XIII) variants. Pepsin digestion of the partly purified $XIII^{N\text{-}short}$ protein revealed three pepsin-resistant fragments, which were subjected to N-terminal sequencing. The bands were found to represent the COL2 and COL3 domains, but not the COL1 domain. Altogether, the results indicate that the folding of COL1 into a pepsin-resistant triple helix is compromised when the NC1 association domain is missing, but the COL2 and COL3 domains are correctly folded. The results also suggest that, in addition to the association domain, the transmembrane domain is also needed for the formation of type XIII collagen molecules with all three collagenous domains in triple-helical conformation.

5.10 Prediction of potential coiled-coil structures in several collagenous transmembrane proteins (III)

For better understanding of the association sequence in the NC1 domain of type XIII collagen, a search of the possible coiled-coil elements was performed using the COILS program (Lupas *et al.* 1991). Type XIII collagen revealed in the COL1 domain a probability of forming coiled-coils with a window size of 21 and 28 residues, the coiled-coil domain already beginning within the transmembrane domain. A potential coiled-coil motif already starting within the transmembrane domain was found in several other collagenous transmembrane proteins as well, namely the collagen types XVII, XXIII and XXV, and MARCO as well as EDA, albeit with a very low probability. These coiled-coil regions of the collagen types XIII, XVII, XXIII, and XXV, MARCO, and EDA were 40, 37, 35, 22, 39, and 24 residues in length, respectively. Surprisingly, the COILS program also predicted that type XIII collagen would have a second coiled-coil sequence in the NC3 domain. The existence of a second coiled-coil sequence could explain the pepsin resistance of the COL2 and COL3 domains in the deletion variants. Similarly, strong coiled-coil predictions were obtained for the NC3 domain of the type XXIII and XXV collagens. Sequence comparisons indicated that a 21-residue stretch of the 40-residue NC1 coiled-coil domain of type XIII collagen is 48 and 38% identical and 52 and 57% homologous to the corresponding sequences in collagen types XXIII and XXV, respectively. In the case of the NC3 coiled-coil domain, 24 residues out of the 29-residue type XIII collagen sequence are 21 and 54% identical and 38 and 71% homologous to the corresponding sequences in the collagen types XXIII and XXV, respectively. Coiled-coil analysis of the recently identified type XXVI collagen, although not a membrane-spanning protein, revealed two potential coiled-coil motifs in the NC1 and NC3 domains similar to the collagen types XIII, XXIII, and XXV. Comparison of the NC1 and NC3 coiled-coil sequences between the collagen types XIII and XXVI revealed sequence homology between the respective NC1 coiled-coil domains, but not between the NC3 domains. More specifically, a 21-residue portion of the 40-residue type XIII collagen NC1 coiled-coil domain is 48% identical and 57% homologous to the corresponding sequence in type XXVI collagen.

5.11 Generation of *Col13a1*$^{-/-}$ and *Col13a1*$^{LacZ/LacZ}$ mouse lines (IV)

Exclusion of most of the 5' UTR, the first protein-coding exon, and part of the first intron did not lead to a null allele, as intended, but to expression of aminoterminally altered type XIII collagen (see original article II). Thus, the excised gene region was expanded to cover the predicted promoter region, in addition to the 5' UTR and the first protein-coding exon (see I in this thesis). The gene area was already excluded from the targeting vector, and the selection cassette was flanked by *loxP* sequences to enable its removal. Several cell lines were correctly targeted with the knock-out construct, but only one of them, following blastocyst injection, produced high chimeras that transmitted the targeted allele into the progeny. Two targeted knock-out mouse lines (*Col13a1*^{T}) were generated in such a way that they originated from two separate chimeric founders, but from the same ES cell line. Homozygous males from both targeted lines were bred with females that retain Cre activity irrespective of the *Cre* minigene transmission. All offspring went through the removal of the selection cassette, although they did not necessarily harbor the *Cre* transgene in their genome. Mice negative for the *Cre* transgene were chosen to start two knock-out lines (*Col13a1*$^{-}$). Homozygous mice were born in a Mendelian ratio, and no gross anatomical abnormalities could be detected. Moreover, the *Col13a1*$^{-/-}$ mice were fertile and had a normal life-span.

To obtain a marker directly linked to the expression of the type XIII collagen gene, a *LacZ*-reporter mouse line was generated (*Col13a1*^{LacZ}). The reporter gene was inserted in-frame with the exon 2 sequences at the furin proprotease cleavage site, resulting in inactivation of the recognition sequence. The *LacZ* gene did not contain its own translation start codon but was extended directly from the type XIII collagen sequences. However, it contained the translation stop codon, thus ensuring synthesis of polypeptides with the first 103 residues derived from the α1(XIII) chain encoding the cytosolic and transmembrane domains and the membrane-proximal ectodomain portion, and the rest encoding the enzyme β-galactosidase. Targeting of the reporter construct resulted in several correctly targeted ES cell lines, and two of them were successfully used to establish independent mouse lines. Homozygous, fertile *Col13a1*$^{LacZ/LacZ}$ mice were born in a Mendelian ratio, and they are viable at least up to 17 moths of age. Part of the *Col13a1*$^{LacZ/LacZ}$ mice appeared smaller than their littermate controls at the age of one month. The weights of mice from both lines were measured at the age of one, three, and 16 to 17 months. Although the *Col13a1*$^{-/-}$ mice were somewhat lighter than the control mice, the difference between the groups was not significant. On the contrary, the difference in weight between the one-month-old *Col13a1*$^{LacZ/LacZ}$ mice and the littermate controls was statistically significant, but did not remain significant in adulthood.

5.12 Expression of type XIII collagen in distinct tissues measured by quantitative real-time PCR (IV)

For quantitative real-time PCR analysis of type XIII collagen transcripts, lung, quadriceps muscle, and femoral bone samples from one-month-old *Col13a1*$^{LacZ/LacZ}$ and 16- to 17-month-old *Col13a1*$^{-/-}$ mice and littermate controls as well as 16.5 dpc fetuses from wild-type, homozygous *Col13a1*$^{LacZ/LacZ}$ and *Col13a1*$^{-/-}$ matings were used for RNA extraction. Two different primer pairs together with their probes, one from the 5' end and the other from the NC3 domain area, were chosen to measure the level of type XIII collagen transcripts in the samples. Use of the two primer pairs resulted in essentially identical results. The highest level of expression was detected in lung, and the expression increased with age. One-month-old femoral bone also exhibited strong type XIII collagen expression, which was considerably lower in the bones of old mice. The amount of type XIII collagen transcripts in muscle was very low, about 3% in one-month-old muscle and 2% in 16 to 17-month-old muscle compared with the expression in the lung of old mice. The level of type XIII collagen transcripts in 16.5 dpc fetuses was moderate.

The level of type XIII collagen transcripts in tissues from *Col13a1*$^{LacZ/LacZ}$ and *Col13a1*$^{-/-}$ mice was compared with that in the controls. The primer pair chosen from the exon 2 to 4 area exhibited very low to undetectable levels of type XIII collagen transcripts in all of the homozygous mutant tissues studied. The other primer pair showed low, but detectable levels of type XIII collagen transcripts irrespective of the mouse line or tissue studied, the level varying between 2 to 10% of the level in the corresponding control samples. We thus consider both mouse lines as type XIII collagen null phenotypes.

5.13 Expression of β-galactosidase in 16.5 dpc *Col13a1*^{LacZ} fetuses (IV)

No blue staining could be detected when fetuses from heterozygous matings between two *Col13a1*$^{+/LacZ}$ mice were subjected to whole-mount β-galactosidase stainings. Nevertheless, antibody staining with an anti-β-galactosidase antibody resulted in a specific staining pattern, and the most prominent staining was detected in all ossifying bone and in restricted areas of muscle presenting developing myotendinous junctions. This *Col13a1*-driven β-galactosidase staining coincides with type XIII collagen antibody staining of the same structures in wild-type fetuses. Both antibodies stain strongly certain other developing tissues, including pharynx, oesophagus, and intestine. Certain tissues showed differences in the relative staining intensities: the β-galactosidase antibody stained intensively meninges and lung, whereas the type XIII collagen antibody showed a more uniform staining pattern in these tissues. Type XIII collagen antibody staining was scant in *Col13a1*$^{LacZ/LacZ}$ fetuses in the structures that stained strongly with the anti-β-galactosidase antibody, and the same phenomenon was noticed in *Col13a1*$^{-/-}$ fetuses.

5.14 Type XIII collagen and neuromuscular junction in the *Col13a1*^{LacZ} line (IV)

The only tissues with detectable, specific blue staining indicative of β-galactosidase enzyme activity, were bone and muscle. The staining pattern in muscle was suggestive of a neuromuscular junction location. This was confirmed in further studies, which showed the β-galactosidase antibody staining to co-localize with the neuromuscular junction marker α-bungarotoxin, the β-galactosidase staining being somewhat broader. β-galactosidase staining was detected at the myotendinous junctions as well. At the electron-microscopic level, the nerve ends were found to be more loosely attached to the muscle cell in the *Col13a1*$^{LacZ/LacZ}$ samples compared with the *Col13a1*$^{+/LacZ}$ samples. In the light of the electron-microscopic results, a likely explanation for the twitching or gasping followed by shivering seen in some of the homozygous mutant mice is disturbance in the signaling between nerves and muscle cells. *Col13a1*$^{LacZ/LacZ}$ mice already showed shivering at the earliest time point studied, i.e. one month. Moreover, 27% of the 16- to 17-month-old *Col13a1*$^{-/-}$ mice shivered while none of the controls did.

5.15 Prediction of binding sites leading to muscle and post-synaptic expression in type XIII collagen sequences (IV)

The genomic 5’ sequences of three mammalian genes found in the NCBI data bank, namely man, mouse, and rat, were screened computationally for E- and N-box sequences known to lead to muscle and synapse-specific expression, respectively. The mouse type XIII collagen 5’ sequence appears to contain two N-box sequences at -1599 to -1594 and at +836 to +841 in the first intron. One N-box sequence in reverse orientation can be found in the human 5’ gene area at the position -472 to -467. In the rat sequence, there are two reverse N-boxes at the positions -445 to -440 and +2494 to +2499. Several E-box consensus sequences were found in all of the three genes both in the 5’ area and in the first intron.

5.16 Properties of bone in mice from the *Col13a1*^{LacZ} line (IV)

Bone was the other tissue in addition to muscle that showed clear specific staining in *Col13a1*$^{LacZ/LacZ}$ mice when β-galactosidase enzyme activity was utilized. Long bones and ribs showed strong periosteal staining. However, the staining did not cover the whole surface of the bone, but concentrated at places where muscle attaches to bone directly or via connective tissue. In calvarial bones, sutures as well as endosteal cells between spicules stained positive for β-galactosidase. The bones of three-month-old female *Col13a1*^{LacZ} mice were analyzed with peripheral quantitative computed tomography and

were found to be equally long but thinner than the bones of age- and sex-matched controls. Moreover, the parameters of the mechanical properties of bone were decreased. The total mineral content was slightly decreased, while mineral density was slightly increased.

6 Discussion

6.1 Comparison of mouse and human type XIII collagen genes indicate high homology

To study the function of type XIII collagen, it was considered important to generate transgenic mouse lines harboring targeted mutations. To enable generation of the targeting vectors and comparison of the sequences conserved between human and mouse species, the structure of the mouse gene was characterized. Some parts of the human type XIII collagen gene had not been fully elucidated, and uncompleted portions of the human gene at the 5' end were therefore also studied to enable comparison of these two genes.

The mouse type XIII collagen gene was found to contain 42 exons varying in size between 8 and 836 bp. Most of the exons in the fibrillar collagen genes are 54 or 45 bp or derivatives of these in length (Vuorio & de Crombrugghe 1990, Exposito *et al.* 2002). A typical feature of the mouse type XIII collagen gene is the abundance of exons coding for collagenous sequences 27 bp in length, half of the proposed ancestral 54-bp size. Another distinction from fibrillar collagen genes is the shared coding of a glycine residue in the collagenous sequence by two distinct exons. The exons in fibrillar collagens start with a complete codon for glycine and end with a complete codon for Yaa in the Gly-Xaa-Yaa sequence (Vuorio & de Crombrugghe 1990, Exposito *et al.* 2002). The exons in the 5' portion not found in the previously isolated human genomic clones (Tikka *et al.* 1991) were identified in the sequence found in the NCBI data bank. The identification of new exons led to a revision of the nomenclature for the exons in both mouse and human genes. The exon-intron organization of the mouse type XIII collagen gene was found to be identical with the human gene, with the exception of only three exons. In addition, the human gene appears to lack the exon corresponding to mouse exon 31, as this exon cannot be found in the completed human genomic sequences in the NCBI database. Thus, the human gene contains 41 exons.

Since not all introns could be amplified by PCR, the sizes were estimated according to restriction site mapping. Altogether, the mouse gene was found to cover 135-140 kb of genomic DNA. Since the entire mouse genome is now available in the NCBI data bank, the size of the mouse type XIII collagen gene can be given as 141 kb

(http://www.ncbi.nlm.nih.gov/LocusLink/LocRpt.cgi?l=12817), which is very close to that estimated in our study. The introns range between 90 bp and 40 kb in size, and the exons were found to be unevenly distributed within the mouse type XIII collagen gene. This is a very typical feature of many collagen genes (Takahara *et al.* 1995, Zhang *et al.* 1996, Chou & Li 2002), where the introns in the 5' portion are very large, as also shown for the type XIII collagen genes. All in all, one can conclude that the type XIII collagen gene structure is highly conserved between man and mouse. The restriction map obtained in the first study was used as a tool to produce the knock-in, knock-out, and deletion-type transgenic mouse constructs also presented in this thesis.

The characterization of mouse and human cDNA clones (Hägg *et al.* 1998) suggested the aminoterminal non-collagenous NC1 domain to be longer than had previously been thought (Pihlajaniemi & Tamminen 1990). The 5´ flanking sequences of the two species were compared, and the promoter previously suggested for the human type XIII collagen gene (Tikka *et al.* 1991) was here proposed to be part of the 5´ untranslated sequences. Thus, the data suggest a different interpretation of the promoter and 5´ untranslated sequences of the human type XIII collagen gene. Sequence similarity between the mouse and human 5' regions is quite low but increases just before the suggested promoter areas. The conserved proposed promoter region contains a modified TATAA motif, several GC boxes, and other conserved putative cis-acting elements. The absence of a traditional TATAA motif and the presence of multiple GC motifs make the type XIII collagen promoter similar to the promoters of genes transcribed widely but at low RNA levels.

The gene encoding mouse type XIII collagen was found to be located in band B4 in chromosome 10, between microsatellite markers as follows: D10Mit5 – (2.3±1.6 cM) – C*ol13al* – (3.4±1.9 cM) – D10Mit15. Today, according to the Mouse Genome Database (Blake *et al.* 2000) used by the NCBI Map Viewer (http://www.ncbi.nlm.nih.gov/mapview/), the microsatellite markers locate in respect to Col13a1 as follows: D10Mit5 – (0.4 cM) – *Col13al* – (4.6 cM) – D10Mit15, which is well compatible with the result obtained in this work. The sub-chromosomal linkage relationships in this region have been conserved on the human chromosome 10q21-22 according to the consensus genetic map data maintained in the Mouse Genome Database at the Jackson Laboratory (http://www.informatics.jax.org). This result is well in line with the mapping result of the *COL13A1* to chromosome 10q22 (Shows *et al.* 1989).

6.2 Aminoterminal sequences of type XIII collagen are required for muscle-basement membrane interaction

Deletion of the first protein-coding exon and the adjacent non-coding sequences in targeted mice did not lead to a deficiency of type XIII collagen but to a synthesis of aminoterminally altered molecules. Nevertheless, the resultant *Col13a1*$^{N/N}$ mice were not completely comparable to the controls, since some phenotypical changes were found in skeletal muscle. Detailed analyses revealed local changes in the plasma membrane-basement membrane interphase, which were especially evident at the myotendinous junction. Additional pathological findings, such as streaming of z-bands, and

disorganization of myofilaments as well as basement membrane detachment in older animals could result from the disturbance of the basement membrane-plasma membrane linkage under contraction. The changes in muscle integrity were local, but the condition was progressive with age. Regardless of that, no clear signs of necrosis or regeneration of abnormal fibers could be detected. A combination of only some of the muscle fibers being affected, a relatively short life-span and a sedentary lifestyle could prevent the development of severe myopathy. However, exercise induced more prominent damage in the muscles of young mutant mice than in the muscles of control mice. Since the condition is progressive, the difference between mutant and control mice would most likely have been more prominent if old mice had been subjected to running.

The major players in the linkage between the muscle cell cytoskeleton and ECM are laminin-2 and its receptor DGC (Berthier & Blaineau 1997, Gullberg *et al.* 1999, Ervasti 2003). Mutations in the genes coding for these proteins cause muscular dystrophies, which are characterized by progressive muscle wasting and weakness (Campbell 1995, Blake *et al.* 2002, Sanes 2003). Another laminin receptor and the major integrin in mature muscle is the α7β1 integrin. In mouse, the absence of the α7 subunit causes progressive muscular dystrophy with marked changes in the myotendinous junctions (Mayer *et al.* 1997). Since the *Col13a1*$^{N/N}$ mice showed changes in the basement membrane-plasma membrane interphase, which were especially evident at the myotendinous junctions, type XIII collagen can be suggested to participate in processes similar to those conducted by α7β1 integrin. The other major basement membrane component in muscle is collagen IV, which is an abundant protein in all basement membranes. The importance of this protein is difficult to study by the transgenic approach, since simultaneous deficiency of the α1(IV) and α2(IV) chains leads to early embryonal lethality (Pöschl *et al.* 2004). Although type IV collagen is such a major component of basement membranes, its receptors in muscle are not clearly defined. *Col13a1*$^{N/N}$ cells were found to harbor decreased adhesion to a type IV collagen substratum, which may implicate direct interaction between these two collagens. The disease called Bethlem myopathy is a muscle disease caused by mutations in the type VI collagen α1(VI) and α2(VI) chains (Jöbsis *et al.* 1996), and a knock-out mouse model for the α1(VI) gene shows histological features of myopathy similar to the human disease (Bonaldo *et al.* 1998). Type VI collagen plays a role in linking the basal lamina to the reticular lamina (Sanes 2003), and our findings indicate that type XIII collagen is also a molecule implicated in mediating the linkage between the muscle fiber and the ECM. *In vitro* studies indicate that the ectodomain of recombinant type XIII collagen binds with the basement membrane components nidogen-2 and perlecan (Tu *et al.* 2002), but the adhesion of *Col13a1*$^{N/N}$ cells to these basement membrane components was not tested. Altogether, the microscopic changes seen in *Col13a1*$^{N/N}$ mice and changes in cell adhesion suggest an important role for type XIII collagen in binding with type IV collagen or some other basement membrane components. Additionally, the anchorage of type XIII collagen to the plasma membrane is necessary for viable muscle cell-matrix linkage.

6.3 Aminoterminally altered type XIII collagen protein harbors a part of the features of the intact protein

It was found by RT-PCR that the aminoterminally altered molecules expressed in *Col13a1*$^{N/N}$ mice lack the cytosolic and transmembrane domains, which are replaced with short unrelated sequences. The promoter and 5´ flanking sequences were retained intact in the *Col13a1*^{N} allele and were thus most probably capable of driving the expression of the altered α1(XIII) chains. The new aminoterminal sequence is either 65 or 11 residues in length, the 11-residue nucleotide sequence surrounding the downstream ATG with somewhat better resemblance to the consensus sequence for the initiation of translation (Kozak 1987). Type XIII collagen is a type II transmembrane protein that lacks a signal sequence, and it is translocated into the membrane of the endoplasmic reticulum through a single transmembrane domain (Hägg *et al.* 1998). Both new sequences are devoid of any known functional motifs and signal sequences. To better understand the properties of the aminoterminally altered type XIII collagen, identical molecules were produced in insect cells. Purification of the mutant collagen from *Col13a1*$^{N/N}$ tissues was not attempted due to the low expression level of type XIII collagen. Instead, studies with insect cells indicated that the aminoterminally altered mouse α1(XIII) chains were secreted and processed by furin endoprotease in the same manner as previously shown for the human protein (Snellman *et al.* 2000a). Moreover, part of the longer altered α chains were secreted without proteolytic processing. The secretion of the recombinant mutant proteins is in line with the results obtained by immunoelectron microscopy, where the mutant molecules were found in the extracellular matrix adjacent to the sarcolemma in *Col13a1*$^{N/N}$ muscle. Furin proproteases are predominantly situated in the *trans*-Golgi network, but they may also be located in the plasma membrane (Seidah & Chretien 1997). Thus, a lack of the aminoterminal 96 residues of authentic type XIII collagen does not impair proteolytic processing at the residues 103-106. The results also suggest that the mutant molecules utilize the same secretory pathway as the authentic molecules. In the case of XIII$^{N\text{-long}}$, shortened, most likely proteolytic fragments utilising the aminoterminus at the residue 147 in the mouse sequence can be found in cell lysates, but these fragments are not secreted into the medium. It was thus proposed here that the sequences following residue 147 do not include information on translocation to the lumen of the endoplasmic reticulum, but such information may be included in the sequence between the residues 96 and 146 in the mouse type XIII collagen sequence. These residues code for the end of the NC1 domain and part of the COL1 domain. It is likely that the non-collagenous sequences encoded by exon 2, which also encodes the furin consensus sequence, are involved in directing type XIII collagen molecules into the endoplasmic reticulum rather than the collagenous ones. In addition to the results obtained with the recombinant proteins expressed in insect cells, type XIII collagen molecules lacking the cytosolic and transmembrane domains were found to be correctly transported to focal adhesions in *Col13a1*$^{N/N}$ cells. In view of the lack of the transmembrane domain, it is thus hypothesized that the altered molecules associate with some other focal adhesion components in the endoplasmic reticulum, the Golgi network or the subsequent secretion vehicles, and that transportation to the plasma membrane is assisted by the formation of such protein complexes.

Since the phenotypical changes in *Col13a1*$^{N/N}$ mice were rather small, the capacity of the mutant molecules to form disulfide-bonded trimers was studied in order to find out their potential to perform at least part of the functions of intact molecules. Neither of the aminoterminally altered molecules, the long or the short, was able to fold into disulfide-bonded trimers. Still, the mutant α chains exhibited the COL2 and COL3 domains in a pepsin-resistant triple-helical conformation. Nevertheless, the COL1 domains of the mutant trimers were found to be sensitive to pepsin implying that the transmembrane and association domain sequences are needed for correct folding of the COL1 domain. These sequences are not, however, included in the proteolytically processed ectodomain, although it is known to have a stable triple-helical COL1 domain (Tu *et al.* 2002). The studies suggest that the triple-helical conformation of the COL1 domain remains stable once it has been formed. All in all, the results obtained from studies with recombinant proteins produced in insect cells support the idea that the mutant molecules resemble intact molecules with respect to the carboxylterminal two thirds of their ectodomain structure. Thus, the aminoterminally altered type XIII collagen molecules may function partly like normal molecules in wild-type animals.

The impaired adhesion properties of cultured fibroblasts derived from *Col13a1*$^{N/N}$ mice indicate that the lacking 96 aminoterminal residues are of significance for the function of type XIII collagen molecules. The mutant fibroblasts showed a marked decrease in adhesion specifically to type IV collagen. This may result from the impaired direct interaction between collagens of types XIII and IV, but the effect may very well also be indirect. Fibroblasts express α1β1 integrin, a subtype known to have high affinity for type IV collagen (Kern *et al.* 1993, Kern & Marcantonio 1998). Mutant fibroblasts produce soluble type XIII collagen, which based on its partly correct triple-helical structure, is likely to retain some of the functional properties of intact type XIII collagen molecules. The amount of secreted type XIII collagen can be speculated to be higher in *Col13a1*$^{N/N}$ mice than in controls, which contain both secreted and membrane-bound forms of type XIII collagen. In the light of the reported interaction between the type XIII collagen ectodomain and the I domain of the α1 integrin subunit (Nykvist *et al.* 2000), this soluble type XIII collagen may act as a functional molecule capable of blocking the α1β1 integrin-type IV collagen interaction. Thus, we can conclude that the muscle phenotype is most likely due to the lack of the cytosolic and transmembrane domains of type XIII collagen, leading to removal of the anchorage characteristic of type XIII collagen and thereby to disturbances in the linkage between the plasma membrane and ECM. Nevertheless, it cannot be ruled out that the mutant soluble proteins interfere with an active process that may contribute to the phenotypical changes in the muscles of mutant mice. All in all, the intact type XIII collagen protein was found to be important for the integrity of muscle structure, but not essential for the development and viability of mice.

6.4 Two independent coiled-coil motifs are predicted in type XIII collagen-like molecules

While studying the molecular properties of aminoterminally altered mouse proteins, the importance of the association domain for proper folding of the α chains into stable disulfide-bonded trimers became evident. It has been shown that interchain disulfide bonds are formed by cysteines located in the NC1 domain or by cysteines at the junction of the COL1 and NC2 domains (Snellman *et al.* 2000a). Some human deletion variants were included in the studies, all of which retained the pair of cysteine residues at the NC1 domain and at the COL1/NC2 junction but lacked the cytosolic and transmembrane portions. Interestingly, none of the aminoterminally mutant or shortened α1(XIII) chains were able to associate efficiently into disulfide-bonded trimers. This was the case even with the del1-61 α1(XIII) chains, which contain the entire NC1 association domain, found to be conserved with other membrane-spanning collagenous proteins (Snellman *et al.* 2000b). This suggests that the previously identified transmembrane-proximal chain association region of about 20 residues is not sufficient to ensure proper association and disulfide-bonding of the α1(XIII) chains, but that the transmembrane domain is also required. Coiled-coil sequences are thought to have an important role in oligomerization (Beck & Brodsky 1998), and we predicted that the NC1 association domain will form a coiled-coil structure. This was found to start already in the transmembrane region, which explains the lack of formation of any proper disulfide-bonded trimers by the del1-61 α1(XIII) chains irrespective of the intact membrane-proximal association domain. Out of the collagenous transmembrane proteins, human type XVII collagen (Balding *et al.* 1997), human MSRs (Emi *et al.* 1993), human SRCLs (Nakamura *et al.* 2001) and mouse MARCO (Beck & Brodsky 1998) have previously been shown to possess coiled-coil structures. The short coiled-coil region in type XVII collagen and MARCO extends directly from the transmembrane region, whereas that in the MSRs and SRCL is long and separated from the transmembrane domain by a spacer domain. Here, coiled-coil structures were also identified in other collagenous transmembrane proteins, namely in the collagen types XIII, XXIII, and XXV and EDA. As with collagen type XVII and MARCO, the potential coiled-coils in the collagen types XIII, XXIII, and XXV and EDA also start from the transmembrane domain and are relatively short. Thus, there appear to be two types of coiled-coils within the collagenous transmembrane proteins, one being short and overlapping with the transmembrane domain, while the longer type is separated from the transmembrane domain by a spacer.

When studying the aminoterminally altered proteins, we found that chains lacking the aminoterminal sequences can still associate with respect to the COL2 and COL3 domains, although the COL1 cannot. Thus, it was thought that sequences residing in other parts of the type XIII collagen molecule could participate in the folding of the α chains into stable triple-helical collagen. This sequence was presumed to locate in the region extending from COL2 to NC4. Subsequently, we predicted the existence of a coiled-coil structure in the NC3 domain. Neither the NC1 nor the NC3 domain of type XIII collagen is affected by alternative splicing, whereas the COL1, NC2, COL3 and NC4 domains are subjected to such modification (Pihlajaniemi *et al.* 1987, Tikka *et al.* 1988, Pihlajaniemi & Tamminen 1990, Juvonen *et al.* 1992, Juvonen & Pihlajaniemi

1992, Juvonen *et al.* 1993, Peltonen *et al.* 1997). This further highlights the importance of the NC1 and NC3 domains. In addition to type XIII collagen, the second coiled-coil motif in the NC3 domain was found in the closely related collagen types XXIII and XXV. Altogether, the type XIII collagen family of molecules possess two coiled-coil domains, one located adjacent to the transmembrane domain and partly overlapping with it and the other in the NC3 domain region. Our protein expression studies suggest that these function as two independent oligomerization domains. Moreover, we also identified aminoterminal and internal coiled-coil domains in type XXVI collagen, which has been shown to have a signal sequence but not a transmembrane domain. Interestingly, this collagen has a consensus sequence for furin cleavage, RRRR, in its NC1 domain (Sato *et al.* 2002), although its involvement in the secretion of the protein has not been studied. The NC1 coiled-coil sequence of type XXVI collagen is homologous to that of type XIII collagen, but whether this acts as an oligomerization domain remains to be seen.

6.5 Reduced adhesion resulting from the lack of type XIII collagen leads to phenotypical changes in mice

We have generated novel knock-out (*Col13a1*$^{-/-}$) and knock-in (*Col13a1*$^{LacZ/LacZ}$) mouse lines by homologous recombination. The first mouse line was generated by deleting the entire promoter area and the first protein-coding exon, and the mutation was expected to lead to a null allele. The *Col13a1*$^{LacZ/LacZ}$ line was generated by inserting the marker gene with the translation stop codon in the exon 2, which is present in all type XIII collagen transcripts (Peltonen *et al.* 1997). Quantitative real-time PCR analyses showed that there are no type XIII collagen transcripts in homozygous mice from either of the lines when measured with primers chosen from the 5' portion of the gene. Some transcripts could be detected when performing the measurements using primers from an area coding for the NC3 domain. The latter results are suggestive of the occurrence of some type of leaky transcription of type XIII collagen in homozygous *Col13a1*$^{LacZ/LacZ}$ and *Col13a1*$^{-/-}$ mice. Such leaky transcription of a null allele is known to occur, and in the case of α7 integrin subunit knock-out mice, for example, 10% of the wild-type transcript levels were observed, but the transcripts were out-of-frame transcripts (Mayer *et al.* 1997). One can conclude that full-length, intact type XIII collagen is either totally lacking or considerably reduced in homozygous mice in both lines, and this deficiency leads to phenotypical changes.

In muscle, type XIII collagen localizes at the sarcolemmal membrane and the myotendinous and neuromuscular junctions, originating there either from terminal Schwann cells or from muscle cells (Hägg *et al.* 2001). Using the reporter gene, type XIII collagen expression was verified at the myotendinous and neuromuscular junctions. However, the marker gene signal was not detected at the sarcolemmal membrane, which is suggestive of a lack of type XIII collagen or its occurrence at low levels at these sites. The expression level at the neuromuscular junction was high enough to give a strong signal with β-galactosidase enzyme staining, while the expression at the myotendinous junction was detectable only by antibody staining. In double staining with α-

bungarotoxin, the *Col13a1*-driven β-galactosidase protein localized clearly at the neuromuscular junction itself, but also in the muscle cell sub-synaptically. Since the expression of type XIII collagen was found at the neuromuscular junction in the *Col13a1*^{LacZ} line, the exon 1-flanking sequences were searched for possible occurrence of N-box sequences. These are known to bind by Ets or related transcription factors, leading to synapse-specific gene expression (Koike *et al.* 1995, Duclert *et al.* 1996, de Kerchove *et al.* 2002). One consensus sequence was found in the 5' flanking region of the mouse gene and another in the first intron. Only one, in reverse orientation, was found in the human gene, and it located in the 5' UTR. Two consensus sequences in reverse orientation were found in the rat gene, one of these in the same position as the human one. The occurrence of N-box consensus sequences in all of the three genes nicely agrees with the detected expression of type XIII collagen at the neuromuscular junction in mice.

Electron-microscopic examination of neuromuscular junctions in the *Col13a1*$^{LacZ/LacZ}$ muscle revealed the structural disturbances, including partial detachment of the nerve terminal from the muscle cell. Thus, in addition to type XIII collagen being needed for the myotendinous junctions, as observed in the case of *Col13a1*$^{N/N}$ mutants, this collagen is also important for the maintenance of neuromuscular junctions. The twitching or gasping followed by shivering seen in some of the homozygous mutant mice may be due to dysfunction of neuromuscular junctions. Behaviour of the same kind has been detected in laminin α2 chain deficient mice (Miyagoe *et al.* 1997). In an analysis of mice expressing aminoterminally altered type XIII collagen lacking the membrane-bound form, changes were observed in muscle integrity, and abnormalities were especially evident at the plasma membrane-basement membrane interface. In light of the new results, some of the local phenotypical changes in the muscles of *Col13a1*$^{N/N}$ mice may result from the disturbance of neuromuscular junctions in these mice. The importance of proper basement membrane composition for the maintenance of neuromuscular junction structure is highlighted by knock-out studies of various laminin chains (Sanes 2003). In the muscles of homozygous mutant *Col13a1*$^{LacZ/LacZ}$ mice, basement membranes were fragmented at sites showing detachment of nerve terminals. Similarly to the implications of the findings on *Col13a1*$^{N/N}$ mice, the results obtained here propose the importance of type XIII collagen in the linkage between muscle cells and the adjacent basement membrane.

Since part of the *Col13a1*$^{LacZ/LacZ}$ mice were considerably smaller in size than their littermate controls, we compared their weights. The *Col13a1*$^{LacZ/LacZ}$ mice suffered a significant weight loss at puberty. A tendency toward reduced weight was detected in both the *Col13a1*^{LacZ} and *Col13a1*$^{-}$ lines in older mice, but this difference was not statistically significant. It is of interest to note that overexpression of type XIII collagen leads to massively thickened bones, suggesting that type XIII collagen participates in growth control in association with mechanical forces which become important when the mice start to move (Ylönen *et al.* submitted). The results obtained here provide further evidence for the involvement of type XIII collagen in bone physiology, since the *Col13a1*$^{LacZ/LacZ}$ mice had thin bones with reduced values for mechanical properties. The periosteum is known to contribute to the lateral growth of bones (Aszódi *et al.* 2000), and the expression of type XIII collagen in the periosteum is thus well in line with the proposed role for it. Namely, type XIII collagen is suggested to participate in the

conduction of signals of mechanical load from muscle to bone, and it is thereby involved in the cellular response to signals in the form of bone growth.

In view of the phenotypes seen in *Col13a1*$^{N/N}$, *Col13a1*$^{LacZ/LacZ}$, and *Col13a1*$^{-/-}$ mice, it is of interest to search for diseases located in the vicinity of this gene on chromosome 10. Only one potential disease affecting muscle or bone can be found, namely myofibrillar myopathy with arrhythmogenic right ventricular cardiomyopathy (Melberg *et al.* 1999) with a locus nearby, although at some distance from the type XIII collagen gene locus. Interestingly, however, the ectopic expression of α1(XIII) chains with a central deletion leads to early embryonal lethality in mice due to defects in cardiovascular development (Sund *et al.* 2001b), showing the involvement of type XIII collagen not only in skeletal muscle, but also in cardiac structure and function. Nevertheless, whether this disease results from mutations in the type XIII collagen gene remains to be seen. Based on mouse studies, mutations in the gene coding for type XIII collagen could result in such diseases as progressive muscular disorder, neuromuscular disorder, or bone disorder. The results obtained from this study may help to focus on human diseases with possible involvement of type XIII collagen.

7 Future perspectives

Our studies show that the lack of type XIII collagen leads to impaired properties of muscle and bone tissues. We have generated a new tool, namely the *Col13a1*^{LacZ} mouse line, to reliably study the expression of type XIII collagen in mice. This will facilitate further studies on the detailed expression pattern, including identification of expressing cell types. Several possibilities to study the function of type XIII collagen can be conceived, but the main focus in the near future will be on the tissues most prominently expressing type XIII collagen, to complement the present findings. In addition to expression profiling, the *Col13a1*^{LacZ} mouse line can simultaneously be used to study the consequences of a lack of type XIII collagen, since the expression of intact type XIII collagen molecules has been blocked by the insertion of the reporter gene. Here, we found delayed growth of young homozygous *Col13a1*^{LacZ} mice, and based on the muscle and bone phenotypes, suggested that type XIII collagen may have a role in the transduction of signals conducted by mechanical forces. Thus, the consequences of a lack of exercise due to immobilization would help to understand the role of collagen XIII in muscle and especially in bone. Other important aspects include the ability of mutant bone to withstand mechanical forces resulting in fractures. Thus, further studies will shed light on the role of type XIII collagen in normal bone physiology and in pathological situations. This suggested role does not rule out a role of type XIII collagen in cell adhesion, as it has been previously suggested. For example, the impaired adhesion of nerve cells at neuromuscular junctions points to the adhesive properties of type XIII collagen. In the muscles of homozygous *Col13a1*$^{-/-}$ and *Col13a1*$^{LacZ/LacZ}$ mice, no regeneration has thus far been noticed. The muscle may compensate for the impaired signaling from nerve terminals by changing its profile with respect to the proportion of fast and slow muscle cell types, which can be studied. Electroneuromyography can be used to measure the function of neuromuscular junctions and the conduction of nerve signals in muscles of homozygous mutant mice. Since the lack of type XIII collagen leads to changes in mouse behavior, it will be interesting to test the ability of mutant mice to preserve balance and coordination. Since disturbances in muscle integrity were also detected in the *Col13a1*$^{N/N}$ mice, they should be included in the studies to further evaluate the function of type XIII collagen when comparing phenotypical changes in these various models.

Studies with mutant mice should be complemented with other approaches. For example, gene expression profiling of mutant and control tissues and cells may reveal components that function together with type XIII collagen, and which may respond to the lack of type XIII collagen at the transcriptional level. Human diseases caused by mutations in type XIII collagen gene have not been reported yet. Knowledge of the roles of type XIII collagen in model organisms will yield better tools for identifying potential human diseases with involvement of type XIII collagen. Based on this study, candidate disorders include muscular and neuromuscular diseases and bone disorders, such as osteoporosis and bone overgrowth.

References

Acton S, Resnick D, Freeman M, Ekkel Y, Ashkenas J & Krieger M (1993) The collagenous domains of macrophage scavenger receptors and complement component C1q mediate their similar, but not identical, binding specificities for polyanionic ligands. J Biol Chem 268: 3530-3537.

Areida SK, Reinhardt DP, Müller PK, Fietzek PP, Köwitz J, Marinkovich MP & Notbohm H (2001) Properties of the collagen type XVII ectodomain. Evidence for N- to C-terminal triple helix folding. J Biol Chem 276: 1594-1601.

Arikawa-Hirasawa E, Rossi SG, Rotundo RL & Yamada Y (2002) Absence of acetylcholinesterase at the neuromuscular junctions of perlecan-null mice. Nat Neurosci 5: 119-123.

Aszódi A, Bateman JF, Gustafsson E, Boot-Handford R & Fässler R (2000) Mammalian skeletogenesis and extracellular matrix: what can we learn from knockout mice? Cell Struct Funct 25: 73-84.

Aszódi A, Hunziker EB, Olsen BR & Fässler R (2001) The role of collagen II and cartilage fibril-associated molecules in skeletal development. Osteoarthritis Cartilage 9 Suppl A: S150-S159.

Balding SD, Diaz LA & Giudice GJ (1997) A recombinant form of the human BP180 ectodomain forms a collagen-like homotrimeric complex. Biochemistry 36: 8821-8830.

Banu J, Wang L & Kalu DN (2003) Effects of increased muscle mass on bone in male mice overexpressing IGF-I in skeletal muscles. Calcif Tissue Int 73: 196-201.

Banyard J, Bao L & Zetter BR (2003) Type XXIII collagen, a new transmembrane collagen identified in metastatic tumor cells. J Biol Chem 278: 20989-20994.

Barrett AJ (1972) Lysosomal enzymes. In: Dingle, JT (ed) Lysosomes, a laboratory handbook 46-126. North-Holland, Amsterdam.

Bayes M, Hartung AJ, Ezer S, Pispa J, Thesleff I, Srivastava AK & Kere J (1998) The anhidrotic ectodermal dysplasia gene (EDA) undergoes alternative splicing and encodes ectodysplasin-A with deletion mutations in collagenous repeats. Hum Mol Genet 7: 1661-1669.

Beck K & Brodsky B (1998) Supercoiled protein motifs: the collagen triple-helix and the alpha-helical coiled coil. J Struct Biol 122: 17-29.

Beier DR (1993) Single-strand conformation polymorphism (SSCP) analysis as a tool for genetic mapping. Mamm Genome 4: 627-631.

Benjamin M, Kumai T, Milz S, Boszczyk BM, Boszczyk AA & Ralphs JR (2002) The skeletal attachment of tendons–tendon "entheses". Comp Biochem Physiol A Mol Integr Physiol 133: 931-945.

Berthier C & Blaineau S (1997) Supramolecular organization of the subsarcolemmal cytoskeleton of adult skeletal muscle fibers. Biol Cell 89: 413-434.

Bianco P, Riminucci M, Gronthos S & Robey PG (2001) Bone marrow stromal stem cells: nature, biology, and potential applications. Stem Cells 19: 180-192.

Blake DJ, Weir A, Newey SE & Davies KE (2002) Function and genetics of dystrophin and dystrophin-related proteins in muscle. Physiol Rev 82: 291-329.

Blake JA, Eppig JT, Richardson JE & Davisson MT (2000) The Mouse Genome Database (MGD): expanding genetic and genomic resources for the laboratory mouse. The Mouse Genome Database Group. Nucleic Acids Res 28: 108-111.

Bonadio J, Saunders TL, Tsai E, Goldstein SA, Morris-Wiman J, Brinkley L, Dolan DF, Altschuler RA, Hawkins JE, Jr. & Bateman JF (1990) Transgenic mouse model of the mild dominant form of osteogenesis imperfecta. Proc Natl Acad Sci USA 87: 7145-7149.

Bonaldo P, Braghetta P, Zanetti M, Piccolo S, Volpin D & Bressan GM (1998) Collagen VI deficiency induces early onset myopathy in the mouse: an animal model for bethlem myopathy. Hum Mol Genet 7: 2135-2140.

Boot-Handford RP, Tuckwell DS, Plumb DA, Rock CF & Poulsom R (2003) A novel and highly conserved collagen (pro(alpha)1(XXVII)) with a unique expression pattern and unusual molecular characteristics establishes a new clade within the vertebrate fibrillar collagen family. J Biol Chem 278: 31067-31077.

Boyd CD, Pierce RA, Schwarzbauer JE, Doege K & Sandell LJ (1993) Alternate exon usage is a commonly used mechanism for increasing coding diversity within genes coding for extracellular matrix proteins. Matrix 13: 457-469.

Brodsky B & Shah NK (1995) Protein motifs. 8. The triple-helix motif in proteins. FASEB J 9: 1537-1546.

Brown JC & Timpl R (1995) The collagen superfamily. Int Arch Allergy Immunol 107: 484-490.

Bruckner P & Prockop DJ (1981) Proteolytic enzymes as probes for the triple-helical conformation of procollagen. Anal Biochem 110: 360-368.

Bulfield G, Siller WG, Wight PA & Moore KJ (1984) X chromosome-linked muscular dystrophy (mdx) in the mouse. Proc Natl Acad Sci USA 81: 1189-1192.

Bulleid NJ, Wilson R & Lees JF (1996) Type-III procollagen assembly in semi-intact cells: chain association, nucleation and triple-helix folding do not require formation of inter- chain disulphide bonds but triple-helix nucleation does require hydroxylation. Biochem J 317: 195-202.

Burkhard P, Stetefeld J & Strelkov SV (2001) Coiled coils: a highly versatile protein folding motif. Trends Cell Biol 11: 82-88.

Burkin DJ, Wallace GQ, Nicol KJ, Kaufman DJ & Kaufman SJ (2001) Enhanced expression of the alpha 7 beta 1 integrin reduces muscular dystrophy and restores viability in dystrophic mice. J Cell Biol 152: 1207-1218.

Buttice G, Kaytes P, D'Armiento J, Vogeli G & Kurkinen M (1990) Evolution of collagen IV genes from a 54-base pair exon: a role for introns in gene evolution. J Mol Evol 30: 479-488.

Campbell KP (1995) Three muscular dystrophies: loss of cytoskeleton-extracellular matrix linkage. Cell 80: 675-679.

Capetanaki Y, Milner DJ & Weitzer G (1997) Desmin in muscle formation and maintenance: knockouts and consequences. Cell Struct Funct 22: 103-116.

Chen Y, Molloy SS, Thomas L, Gambee J, Bächinger HP, Ferguson B, Zonana J, Thomas G & Morris NP (2001) Mutations within a furin consensus sequence block proteolytic release of ectodysplasin-A and cause X-linked hypohidrotic ectodermal dysplasia. Proc Natl Acad Sci USA 98: 7218-7223.

Chipman SD, Sweet HO, McBride DJ, Jr., Davisson MT, Marks SC, Jr., Shuldiner AR, Wenstrup RJ, Rowe DW & Shapiro JR (1993) Defective pro α2(I) collagen synthesis in a recessive mutation in mice: a model of human osteogenesis imperfecta. Proc Natl Acad Sci USA 90: 1701-1705.

Chomczynski P & Sacchi N (1987) Single-step method of RNA isolation by acid guanidium thiocyanate-phenol-chloroform extraction. Anal Biochem 162: 156-159.

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

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

de Crombrugghe B, Lefebvre V & Nakashima K (2001) Regulatory mechanisms in the pathways of cartilage and bone formation. Curr Opin Cell Biol 13: 721-727.

Deconinck AE, Potter AC, Tinsley JM, Wood SJ, Vater R, Young C, Metzinger L, Vincent A, Slater CR & Davies KE (1997) Postsynaptic abnormalities at the neuromuscular junctions of utrophin-deficient mice. J Cell Biol 136: 883-894.

Devereux J, Haeberli P & Smithies O (1984) A comprehensive set of sequence analysis programs for the VAX. Nucleic Acids Res 12: 387-395.

Diaz LA, Ratrie H, Saunders WS, Futamura S, Squiquera HL, Anhalt GJ & Giudice GJ (1990) Isolation of a human epidermal cDNA corresponding to the 180-kD autoantigen recognized by bullous pemphigoid and herpes gestationis sera. Immunolocalization of this protein to the hemidesmosome. J Clin Invest 86: 1088-1094.

Dublet B, Vernet T & van der Rest M (1999) Schmid's metaphyseal chondrodysplasia mutations interfere with folding of the C-terminal domain of human collagen X expressed in Escherichia coli. J Biol Chem 274: 18909-18915.

Duclert A, Savatier N, Schaeffer L & Changeux JP (1996) Identification of an element crucial for the sub-synaptic expression of the acetylcholine receptor epsilon-subunit gene. J Biol Chem 271: 17433-17438.

Eklund L, Piuhola J, Komulainen J, Sormunen R, Ongvarrasopone C, Fässler R, Muona A, Ilves M, Ruskoaho H, Takala TE & Pihlajaniemi T (2001) Lack of type XV collagen causes a skeletal myopathy and cardiovascular defects in mice. Proc Natl Acad Sci USA 98: 1194-1199.

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

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.

Emi M, Asaoka H, Matsumoto A, Itakura H, Kurihara Y, Wada Y, Kanamori H, Yazaki Y, Takahashi E & Lepert M (1993) Structure, organization, and chromosomal mapping of the human macrophage scavenger receptor gene. J Biol Chem 268: 2120-2125.

Engel J & Kammerer RA (2000) What are oligomerization domains good for? Matrix Biol 19: 283-288.

Ervasti JM (2003) Costameres: the Achilles' heel of Herculean muscle. J Biol Chem 278: 13591-13594.

Evans MJ & Kaufman MH (1981) Establishment in culture of pluripotential cells from mouse embryos. Nature 292: 154-156.

Exposito JY, Cluzel C, Garrone R & Lethias C (2002) Evolution of collagens. Anat Rec 268: 302-316.

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

Ferguson BM, Brockdorff N, Formstone E, Ngyuen T, Kronmiller JE & Zonana J (1997) Cloning of Tabby, the murine homolog of the human EDA gene: evidence for a membrane-associated protein with a short collagenous domain. Hum Mol Genet 6: 1589-1594.

Fitzgerald J & Bateman JF (2001) A new FACIT of the collagen family: COL21A1(1). FEBS Lett 505: 275-280.

van der Flier A & Sonnenberg A (2001) Function and interactions of integrins. Cell Tissue Res 305: 285-298.

Forlino A, Porter FD, Lee EJ, Westphal H & Marini JC (1999) Use of the Cre/lox recombination system to develop a non-lethal knock-in murine model for osteogenesis imperfecta with an alpha1(I) G349C substitution. Variability in phenotype in BrtlIV mice. J Biol Chem 274: 37923-37931.

Frank S, Lustig A, Schulthess T, Engel J & Kammerer RA (2000) A distinct seven-residue trigger sequence is indispensable for proper coiled-coil formation of the human macrophage scavenger receptor oligomerization domain. J Biol Chem 275: 11672-11677.

Franzke CW, Tasanen K, Schäcke H, Zhou Z, Tryggvason K, Mauch C, Zigrino P, Sunnarborg S, Lee DC, Fahrenholz F & Bruckner-Tuderman L (2002) Transmembrane collagen XVII, an epithelial adhesion protein, is shed from the cell surface by ADAMs. EMBO J 21: 5026-5035.

Furthmayr H & Timpl R (1971) Characterization of collagen peptides by sodium dodecylsulfate-polyacrylamide electrophoresis. Anal Biochem 41: 510-516.

Fässler R & Meyer M (1995) Consequences of lack of β1 integrin gene expression in mice. Genes Dev 9: 1896-1908.

Gaide O & Schneider P (2003) Permanent correction of an inherited ectodermal dysplasia with recombinant EDA. Nat Med 9: 614-618.

Gautam M, DeChiara TM, Glass DJ, Yancopoulos GD & Sanes JR (1999) Distinct phenotypes of mutant mice lacking agrin, MuSK, or rapsyn. Brain Res Dev Brain Res 114: 171-178.

Geiger B & Bershadsky A (2001) Assembly and mechanosensory function of focal contacts. Curr Opin Cell Biol 13: 584-592.

Geiger B & Bershadsky A (2002) Exploring the neighborhood: adhesion-coupled cell mechanosensors. Cell 110: 139-142.

Geiger B, Bershadsky A, Pankov R & Yamada KM (2001) Transmembrane crosstalk between the extracellular matrix–cytoskeleton crosstalk. Nat Rev Mol Cell Biol 2: 793-805.

Gelse K, Poschl E & Aigner T (2003) Collagens–structure, function, and biosynthesis. Adv Drug Deliv Rev 55: 1531-1546.

Gordon JW, Scangos GA, Plotkin DJ, Barbosa JA & Ruddle FH (1980) Genetic transformation of mouse embryos by microinjection of purified DNA. Proc Natl Acad Sci USA 77: 7380-7384.

Gossler A & Zachgo J (1993) Gene and enhancer trap screens in ES cell chimeras. In: Joyner, AL (ed) Gene Targeting-A practical Approach 181-234. Oxford University Press, New York.

Grady RM, Grange RW, Lau KS, Maimone MM, Nichol MC, Stull JT & Sanes JR (1999) Role for alpha-dystrobrevin in the pathogenesis of dystrophin-dependent muscular dystrophies. Nat Cell Biol 1: 215-220.

Grady RM, Merlie JP & Sanes JR (1997) Subtle neuromuscular defects in utrophin-deficient mice. J Cell Biol 136: 871-882.

Gress CJ & Jacenko O (2000) Growth plate compressions and altered hematopoiesis in collagen X null mice. J Cell Biol 149: 983-993.

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

Gullberg D, Tiger CF & Velling T (1999) Laminins during muscle development and in muscular dystrophies. Cell Mol Life Sci 56: 442-460.

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

Heikkilä P & Soininen R (1996) The type IV collagen gene family. Contrib Nephrol 117: 105-129.

Heiskanen M, Kallioniemi O & Palotie A (1996) Fiber-FISH: experiences and a refined protocol. Genet Anal 12: 179-184.

Helminen HJ, Säämänen AM, Salminen H & Hyttinen MM (2002) Transgenic mouse models for studying the role of cartilage macromolecules in osteoarthritis. Rheumatology (Oxford) 41: 848-856.

Hirako Y, Usukura J, Uematsu J, Hashimoto T, Kitajima Y & Owaribe K (1998) Cleavage of BP180, a 180-kDa bullous pemphigoid antigen, yields a 120- kDa collagenous extracellular polypeptide. J Biol Chem 273: 9711-9717.

Hopkinson SB, Baker SE & Jones JCR (1995) Molecular genetic studies of a human epidermal autoantigen (the 180-kD bullous pemphigoid antigen/BP180): identification of functionally important sequences within the BP180 molecule and evidence for an interaction between BP180 and α6 integrin. J Cell Biol 130: 117-125.

Hopkinson SB & Jones JC (2000) The N terminus of the transmembrane protein BP180 interacts with the N-terminal domain of BP230, thereby mediating keratin cytoskeleton anchorage to the cell surface at the site of the hemidesmosome. Mol Biol Cell 11: 277-286.

Hopkinson SB, Riddelle KS & Jones JCR (1992) Cytoplasmic domain of the 180-kD bullous pemphigoid antigen, a hemidesmosomal component: molecular and cell biologic characterization. J Invest Dermatol 99: 264-270.

Horton WA (2003) Skeletal development: insights from targeting the mouse genome. Lancet 362: 560-569.

Hägg P, Rehn M, Huhtala P, Väisänen T, Tamminen M & Pihlajaniemi T (1998) Type XIII collagen is identified as a plasma membrane protein. J Biol Chem 273: 15590-15597.

Hägg P, Väisänen T, Tuomisto A, Rehn M, Tu H, Huhtala P, Eskelinen S & Pihlajaniemi T (2001) Type XIII collagen: a novel cell adhesion component present in a range of cell-matrix adhesions and in the intercalated discs between cardiac muscle cells. Matrix Biol 19: 727-742.

Hägg PM, Hägg PO, Peltonen S, Autio-Harmainen H & Pihlajaniemi T (1997) Location of type XV collagen in human tissues and its accumulation in the interstitial matrix of the fibrotic kidney. Am J Pathol 150: 2075-2086.

Jacenko O, Roberts DW, Campbell MR, McManus PM, Gress CJ & Tao Z (2002) Linking hematopoiesis to endochondral skeletogenesis through analysis of mice transgenic for collagen X. Am J Pathol 160: 2019-2034.

Jacobson C, Cote PD, Rossi SG, Rotundo RL & Carbonetto S (2001) The dystroglycan complex is necessary for stabilization of acetylcholine receptor clusters at neuromuscular junctions and formation of the synaptic basement membrane. J Cell Biol 152: 435-450.

Jaenisch R (1976) Germ line integration and Mendelian transmission of the exogenous Moloney leukemia virus. Proc Natl Acad Sci USA 73: 1260-1264.

Juvonen M & Pihlajaniemi T (1992) Characterization of the spectrum of alternative splicing of α1(XIII) collagen transcripts in HT-1080 cells and calvarial tissue resulted in identification of two previously unidentified alternatively spliced sequences, one previously unidentified exon, and nine new mRNA variants. J Biol Chem 267: 24693-24699.

Juvonen M, Pihlajaniemi T & Autio-Harmainen H (1993) Location and alternative splicing of type XIII collagen RNA in the early human placenta. Lab Invest 69: 541-551.

Juvonen M, Sandberg M & Pihlajaniemi T (1992) Patterns of expression of the six alternatively spliced exons affecting the structures of the COL1 and NC2 domains of the α1(XIII) collagen chain in human tissues and cell lines. J Biol Chem 267: 24700-24707.

Jämsä T, Jalovaara P, Peng Z, Väänänen HK & Tuukkanen J (1998) Comparison of three-point bending test and peripheral quantitative computed tomography analysis in the evaluation of the strength of mouse femur and tibia. Bone 23: 155-161.

Jöbsis GJ, Keizers H, Vreijling JP, de Visser M, Speer MC, Wolterman RA, Baas F & Bolhuis PA (1996) Type VI collagen mutations in Bethlem myopathy, an autosomal dominant myopathy with contractures. Nat Genet 14: 113-115.

Kassner A, Hansen U, Miosge N, Reinhardt DP, Aigner T, Bruckner-Tuderman L, Bruckner P & Grassel S (2003) Discrete integration of collagen XVI into tissue-specific collagen fibrils or beaded microfibrils. Matrix Biol 22: 131-143.

de Kerchove D'E, Cartaud J, Ravel-Chapuis A, Seroz T, Pasteau F, Angus LM, Jasmin BJ, Changeux JP & Schaeffer L (2002) Expression of mutant Ets protein at the neuromuscular synapse causes alterations in morphology and gene expression. EMBO Rep 3: 1075-1081.

Kere J, Srivastava AK, Montonen O, Zonana J, Thomas N, Ferguson B, Munoz F, Morgan D, Clarke A, Baybayan P, Chen EY, Ezer S, Saarialho-Kere U, de la Chapelle A & Schlessinger D (1996) X-linked anhidrotic (hypohidrotic) ectodermal dysplasia is caused by mutation in a novel transmembrane protein. Nat Genet 13: 409-416.

Kern A, Eble J, Golbik R & Kuhn K (1993) Interaction of type IV collagen with the isolated integrins $\alpha 1\beta$ 1 and $\alpha 2\beta 1$. Eur J Biochem 215: 151-159.

Kern A & Marcantonio EE (1998) Role of the I-domain in collagen binding specificity and activation of the integrins alpha1beta1 and alpha2beta1. J Cell Physiol 176: 634-641.

Khillan JS, Olsen AS, Kontusaari S, Sokolov B & Prockop DJ (1991) Transgenic mice that express a mini-gene version of the human gene for type I procollagen (COL1A1) develop a phenotype resembling a lethal form of osteogenesis imperfecta. J Biol Chem 266: 23373-23379.

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

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

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.

Koch M, Foley JE, Hahn R, Zhou P, Burgeson RE, Gerecke DR & Gordon MK (2001) alpha 1(Xx) collagen, a new member of the collagen subfamily, fibril-associated collagens with interrupted triple helices. J Biol Chem 276: 23120-23126.

Koch M, Laub F, Zhou P, Hahn RA, Tanaka S, Burgeson RE, Gerecke DR, Ramirez F & Gordon MK (2003) Collagen XXIV, a vertebrate fibrillar collagen with structural features of invertebrate collagens: selective expression in developing cornea and bone. J Biol Chem 278: 43236-43244.

Koch M, Schulze J, Hansen U, Ashwodt T, Keene DR, Brunken WJ, Burgeson RE, Bruckner P & Bruckner-Tuderman L (2004) A novel marker of tissue junctions, collagen XXII. J Biol Chem 279: 22514-22521.

Koike S, Schaeffer L & Changeux JP (1995) Identification of a DNA element determining synaptic expression of the mouse acetylcholine receptor delta-subunit gene. Proc Natl Acad Sci USA 92: 10624-10628.

Komori T, Yagi H, Nomura S, Yamaguchi A, Sasaki K, Deguchi K, Shimizu Y, Bronson RT, Gao YH, Inada M, Sato M, Okamoto R, Kitamura Y, Yoshiki S & Kishimoto T (1997) Targeted disruption of Cbfa1 results in a complete lack of bone formation owing to maturational arrest of osteoblasts. Cell 89: 755-764.

Kozak M (1987) An analysis of 5'-noncoding sequences from 699 vertebrate messenger RNAs. Nucleic Acids Res 15: 8125-8148.

Kuivaniemi H, Tromp G & Prockop DJ (1991) Mutations in collagen genes: causes of rare and some common diseases in humans. FASEB J 5: 2052-2060.

Kvansakul M, Bogin O, Hohenester E & Yayon A (2003) Crystal structure of the collagen alpha1(VIII) NC1 trimer. Matrix Biol 22: 145-152.

Labrousse AL, Buisson-Legendre N, Hornebeck W & Bernard P (2002) The metalloprotease-directed shedding of BP 180 (collagen XVII) from human keratinocytes in culture is unaffected by ceramide and cell- matrix interaction. Eur J Dermatol 12: 240-246.

Lamberg A, Helaakoski T, Myllyharju J, Peltonen S, Notbohm H, Pihlajaniemi T & Kivirikko KI (1996) Characterization of human type III collagen expressed in a baculovirus system. Production of a protein with a stable triple helix requires coexpression with the two types of recombinant prolyl 4-hydroxylase subunit. J Biol Chem 271: 11988-11995.

Lemieux N, Dutrillaux B & Viegas-Pequignot E (1992) A simple method for simultaneous R- or G-banding and fluorescence in situ hybridization of small single-copy genes. Cytogenet Cell Genet 59: 311-312.

Lichter P, Cremer T, Tang CJ, Watkins PC, Manuelidis L & Ward DC (1988) Rapid detection of human chromosome 21 aberrations by in situ hybridization. Proc Natl Acad Sci USA 85: 9664-9668.

Loria PM, Hodgkin J & Hobert O (2004) A conserved postsynaptic transmembrane protein affecting neuromuscular signaling in Caenorhabditis elegans. J Neurosci 24: 2191-2201.

Lupas A (1996) Coiled coils: new structures and new functions. Trends Biochem Sci 21: 375-382.

Lupas A, Van Dyke M & Stock J (1991) Predicting coiled coils from protein sequences. Science 252: 1162-1164.

Majalahti-Palviainen T, Hirvinen M, Tervonen V, Ilves M, Ruskoaho H & Vuolteenaho O (2000) Gene structure of a new cardiac peptide hormone: a model for heart-specific gene expression. Endocrinology 141: 731-740.

Manly KF (1993) A Macintosh program for storage and analysis of experimental genetic mapping data. Mamm Genome 4: 303-313.

Marks SCJr & Hermey DC (1996) The Sructure and Development of Bone. In: Bilezikian JP, Rajewsky K, & Rodan GA (eds) Principles of Bone Biology 3-14. Academic Press, Inc., San Diego.

Matsuda Y, Harada YN, Natsuume-Sakai S, Lee K, Shiomi T & Chapman VM (1992) Location of the mouse complement factor H gene (cfh) by FISH analysis and replication R-banding. Cytogenet Cell Genet 61: 282-285.

Mayer U (2003) Integrins: redundant or important players in skeletal muscle? J Biol Chem 278: 14587-14590.

Mayer U, Saher G, Fässler R, Bornemann A, Echtermeyer F, von der M, Miosge N & Poschl E (1997) Absence of integrin α7 causes a novel form of muscular dystrophy. Nat Genet 17: 318-323.

McAlinden A, Smith TA, Sandell LJ, Ficheux D, Parry DA & Hulmes DJ (2003) Alpha-helical coiled-coil oligomerization domains are almost ubiquitous in the collagen superfamily. J Biol Chem 278: 42200-42207.

McGrath JA, Gatalica B, Li K, Dunnill MG, McMillan JR, Christiano AM, Eady RA & Uitto J (1996) Compound heterozygosity for a dominant glycine substitution and a recessive internal duplication mutation in the type XVII collagen gene results in junctional epidermolysis bullosa and abnormal dentition. Am J Pathol 148: 1787-1796.

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

McLean W & Olsen BR (2001) Mouse models of abnormal skeletal development and homeostasis. Trends Genet 17: S38-S43.

Melberg A, Oldfors A, Blomstrom-Lundqvist C, Stalberg E, Carlsson B, Larsson E, Lidell C, Eeg-Olofsson KE, Wikstrom G, Henriksson KG & Dahl N (1999) Autosomal dominant myofibrillar myopathy with arrhythmogenic right ventricular cardiomyopathy linked to chromosome 10q. Ann Neurol 46: 684-692.

Miyagoe Y, Hanaoka K, Nonaka I, Hayasaka M, Nabeshima Y, Arahata K, Nabeshima Y & Takeda S (1997) Laminin alpha2 chain-null mutant mice by targeted disruption of the Lama2 gene: a new model of merosin (laminin 2)-deficient congenital muscular dystrophy. FEBS Lett 415: 33-39.

Mohan S, Richman C, Guo R, Amaar Y, Donahue LR, Wergedal J & Baylink DJ (2003) Insulin-like growth factor regulates peak bone mineral density in mice by both growth hormone-dependent and -independent mechanisms. Endocrinology 144: 929-936.

Monreal AW, Zonana J & Ferguson B (1998) Identification of a new splice form of the EDA1 gene permits detection of nearly all X-linked hypohidrotic ectodermal dysplasia mutations. Am J Hum Genet 63: 380-389.

Montonen O, Ezer S, Saarialho-Kere UK, Herva R, Karjalainen-Lindsberg ML, Kaitila I, Schlessinger D, Srivastava AK, Thesleff I & Kere J (1998) The gene defective in anhidrotic ectodermal dysplasia is expressed in the developing epithelium, neuroectoderm, thymus, and bone. J Histochem Cytochem 46: 281-289.

Muona A, Eklund L, Sormunen R, Väisänen T & Pihlajaniemi T (2002) Developmentally regulated expression of type XV collagen correlates with abnormalities in Col15a1-/- mice. Matrix Biol 21:89-102.

Muragaki Y, Timmons S, Griffith CM, Oh SP, Fadel B, Quetermous T & Olsen BR (1995) Mouse Col18a1 is expressed in a tissue-specific manner as three alternative variants and is localized in the basement membrane zone. Proc Natl Acad Sci USA 92: 8763-8767.

Myers JC, Li D, Amenta PS, Clark CC, Nagaswami C & Weisel JW (2003) Type XIX collagen purified from human umbilical cord is characterized by multiple sharp kinks delineating collagenous subdomains and by intermolecular aggregates via globular, disulfide-linked, and heparin-binding amino termini. J Biol Chem 278: 32047-32057.

Myers JC, Li D, Bageris A, Abraham V, Dion AS & Amenta PS (1997) Biochemical and immunohistochemical characterization of human type XIX defines a novel class of basement membrane zone collagens. Am J Pathol 151: 1729-1740.

Myllyharju J (2004) Molecular Biology and Biosynthesis of Collagens. In: Massaro, EJ & Rogers, JM (eds) The Skeleton - Biochemical, Genetic, and Molecular Interactions in Development and Homeostasis 77-87. Humana Press, Inc., Totowa.

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

Myllyharju J & Kivirikko KI (2004) Collagens, modifying enzymes and their mutations in humans, flies and worms. Trends Genet 20: 33-43.

Nagy A, Rossant J, Nagy R, Abramow-Newerly W & Roder JC (1993) Derivation of completely cell culture-derived mice from early-passage embryonic stem cells. Proc Natl Acad Sci USA 90: 8424-8428.

Nakamura K, Funakoshi H, Miyamoto K, Tokunaga F & Nakamura T (2001) Molecular cloning and functional characterization of a human scavenger receptor with C-type lectin (SRCL), a novel member of a scavenger receptor family. Biochem Biophys Res Commun 280: 1028-1035.

Nakayama K (1997) Furin: a mammalian subtilisin/Kex2p-like endoprotease involved in processing of a wide variety of precursor proteins. Biochem J 327: 625-635.

Nishimura I, Muragaki Y & Olsen BR (1989) Tissue-specific forms of type IX collagen-proteoglycan arise from the use of two widely separated promoters. J Biol Chem 264: 20033-20041.

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.

Nykvist P, Tu H, Ivaska J, Käpylä J, Pihlajaniemi T & Heino J (2000) Distinct recognition of collagen subtypes by alpha(1)beta(1) and alpha(2)beta(1) integrins. Alpha(1)beta(1) mediates cell adhesion to type XIII collagen. J Biol Chem 275: 8255-8261.

O'Reilly MS, Boehm T, Shing Y, Fukai N, Vasios G, Lane WS, Flynn E, Birkhead JR, Olsen BR & Folkman J (1997) Endostatin: an endogenous inhibitor of angiogenesis and tumor growth. Cell 88: 277-285.

Ohtani K, Suzuki Y, Eda S, Kawai T, Kase T, Keshi H, Sakai Y, Fukuoh A, Sakamoto T, Itabe H, Suzutani T, Ogasawara M, Yoshida I & Wakamiya N (2001) The membrane-type collectin CL-P1 is a scavenger receptor on vascular endothelial cells. J Biol Chem 276: 44222-44228.

Olsen BR (1997) Collagen IX. Int J Biochem Cell Biol 29: 555-558.

Olsen BR, Reginato AM & Wang W (2000) Bone development. Annu Rev Cell Dev Biol 16: 191-220.

Ortega N & Werb Z (2002) New functional roles for non-collagenous domains of basement membrane collagens. J Cell Sci 115: 4201-4214.

Pace JM, Corrado M, Missero C & Byers PH (2003) Identification, characterization and expression analysis of a new fibrillar collagen gene, COL27A1. Matrix Biol 22: 3-14.

Peltonen S, Hentula M, Hägg P, Ylä-Outinen H, Tuukkanen J, Lakkakorpi J, Rehn M, Pihlajaniemi T & Peltonen J (1999) A novel component of epidermal cell-matrix and cell-cell contacts: transmembrane protein type XIII collagen. J Invest Dermatol 113: 635-642.

Peltonen S, Rehn M & Pihlajaniemi T (1997) Alternative splicing of mouse α1(XIII) collagen RNAs results in at least 17 different transcripts, predicting α1(XIII) collagen chains with length varying between 651 and 710 amino acid residues. DNA Cell Biol 16: 227-234.

Pereira R, Khillan JS, Helminen HJ, Hume EL & Prockop DJ (1993) Transgenic mice expressing a partially deleted gene for type I procollagen (COL1A1). A breeding line with a phenotype of spontaneous fractures and decreased bone collagen and mineral. J Clin Invest 91: 709-716.

Pereira RF, Hume EL, Halford KW & Prockop DJ (1995) Bone fragility in transgenic mice expressing a mutated gene for type I procollagen (COL1A1) parallels the age-dependent phenotype of human osteogenesis imperfecta. J Bone Miner Res 10: 1837-1843.

Phillips CL, Bradley DA, Schlotzhauer CL, Bergfeld M, Libreros-Minotta C, Gawenis LR, Morris JS, Clarke LL & Hillman LS (2000) Oim mice exhibit altered femur and incisor mineral composition and decreased bone mineral density. Bone 27: 219-226.

Pihlajaniemi T & Myers JC (1987) Characterization of a pro- α2(I) collagen gene mutation by nuclease S1 mapping. Meth Enzymol 145: 213-222.

Pihlajaniemi T, Myllylä.R., Seyer J, Kurkinen M & Prockop DJ (1987) Partial characterization of a low molecular weight human collagen that undergoes alternative splicing. Proc Natl Acad Sci USA 84: 940-944.

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

Pihlajaniemi T & Tamminen M (1990) The α1 chain of type XIII collagen consists of three collagenous and four noncollagenous domains, and its primary transcript undergoes complex alternative splicing. J Biol Chem 265: 16922-16928.

Pinkel D, Landegent J, Collins C, Fuscoe J, Segraves R, Lucas J & Gray J (1988) Fluorescence in situ hybridization with human chromosome- specific libraries: detection of trisomy 21 and translocations of chromosome 4. Proc Natl Acad Sci USA 85: 9138-9142.

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

Pöschl E, Schlötzer-Schrehardt U, Brachvogel B, Saito K, Ninomiya Y & Mayer U (2004) Collagen IV is essential for basement membrane stability but dispensable for initiation of its assembly during early development. Development 131: 1619-1628.

Ramchandran R, Dhanabal M, Volk R, Waterman MJF, Segal M, Lu H, Knebelmann B & Sukhatme VP (1999) Antiangiogenic activity of restin, NC10 domain of human collagen XV: Comparison to endostatin. Biochem Biophys Res Commun 255: 735-739.

Rehn M & Pihlajaniemi T (1995) Identification of three N-terminal ends of type XVIII collagen chains and tissue-specific differences in the expression of the corresponding transcripts. The longest form contains a novel motif homologous to rat and Drosophila frizzled proteins. J Biol Chem 270: 4705-4711.

van der Rest M & Garrone R (1991) Collagen family of proteins. FASEB J 5: 2814-2823.

Rowe LB, Nadeau JH, Turner R, Frankel WN, Letts VA, Eppig JT, Ko MS, Thurston SJ & Birkenmeier EH (1994) Maps from two interspecific backcross DNA panels available as a community genetic mapping resource. Mamm Genome 5: 253-274.

Ryder-Cook AS, Sicinski P, Thomas K, Davies KE, Worton RG, Barnard EA, Darlison MG & Barnard PJ (1988) Localization of the mdx mutation within the mouse dystrophin gene. EMBO J 7: 3017-3021.

Saarela J, Rehn M, Oikarinen A, Autio-Harmainen H & Pihlajaniemi T (1998) The short and long forms of type XVIII collagen show clear tissue specificities in their expression and location in basement membrane zones in humans. Am J Pathol 153: 611-626.

Saitta B & Chu ML (1994) Two promoters control the transcription of the human alpha 2(VI) collagen gene. Eur J Biochem 223: 675-682.

Sakai K & Miyazaki J (1997) A transgenic mouse line that retains Cre recombinase activity in mature oocytes irrespective of the cre transgene transmission. Biochem Biophys Res Commun 237: 318-324.

Sallinen R, Latvanlehto A, Kvist AP, Rehn M, Eerola I, Chu ML, Bonaldo P, Saitta B, Bressan GM, Pihlajaniemi T, Vuorio E, Palotie A, Wessman M & Horelli-Kuitunen N (2001) Physical mapping of mouse collagen genes on chromosome 10 by high-resolution FISH. Mamm Genome 12: 340-346.

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

Sandberg M, Tamminen M, Hirvonen H, Vuorio E & Pihlajaniemi T (1989) Expression of mRNAs coding for the α1 chain of type XIII collagen in human fetal tissues: comparison with expression of mRNAs for collagen types I, II, and III. J Cell Biol 109: 1371-1379.

Sanes JR (2003) The basement membrane/basal lamina of skeletal muscle. J Biol Chem 278: 12601-12604.

Sanes JR & Lichtman JW (1999) Development of the vertebrate neuromuscular junction. Annu Rev Neurosci 22: 389-442.

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

Sato K, Yomogida K, Wada T, Yorihuzi T, Nishimune Y, Hosokawa N & Nagata K (2002) Type XXVI collagen, a new member of the collagen family, is specifically expressed in the testis and ovary. J Biol Chem

Schneider P, Street SL, Gaide O, Hertig S, Tardivel A, Tschopp J, Runkel L, Alevizopoulos K, Ferguson BM & Zonana J (2001) Mutations leading to X-linked hypohidrotic ectodermal dysplasia affect three major functional domains in the tumor necrosis factor family member ectodysplasin-A. J Biol Chem 276: 18819-18827.

Schnieke A, Harbers K & Jaenisch R (1983) Embryonic lethal mutation in mice induced by retrovirus insertion into the α1(I) collagen gene. Nature 304: 315-320.

Schwander M, Leu M, Stumm M, Dorchies OM, Ruegg UT, Schittny J & Muller U (2003) Beta1 integrins regulate myoblast fusion and sarcomere assembly. Dev Cell 4: 673-685.

Schäcke H, Schumann H, Hammami-Hauasli N, Raghunath M & Bruckner-Tuderman L (1998) Two forms of collagen XVII in keratinocytes. A full-length transmembrane protein and a soluble ectodomain. J Biol Chem 273: 25937-25943.

Seidah NG & Chretien M (1997) Eukaryotic protein processing: endoproteolysis of precursor proteins. Curr Opin Biotechnol 8: 602-607.

Shows TB, Tikka L, Byers MG, Eddy RL, Haley LL, Henry WM, Prockop DJ & Tryggvason K (1989) Assignment of the human collagen α1(XIII) chain gene (COL13A1) to the q22 region of chromosome 10. Genomics 5: 128-133.

Singer SJ (1990) The structure and insertion of integral proteins in membranes. Annu Rev Cell Biol 6: 247-296.

Slot JW & Geuze HJ (1985) A new method of preparing gold probes for multiple-labeling cytochemistry. Eur J Cell Biol 38: 87-93.

Snellman A, Keränen MR, Hägg PO, Lamberg A, Hiltunen JK, Kivirikko KI & Pihlajaniemi T (2000a) Type XIII collagen forms homotrimers with three triple helical collagenous domains and its association into disulfide-bonded trimers is enhanced by prolyl 4-hydroxylase. J Biol Chem 275: 8936-8944.

Snellman A & Pihlajaniemi T (2003) Collagenous transmembrane proteins. In: Pandalai, SG (ed) Recent Research Developments in Biological Chemistry 1-22. Research Signpost, Kerala, India.

Snellman A, Tu H, Väisänen T, Kvist AP, Huhtala P & Pihlajaniemi T (2000b) A short sequence in the N-terminal region is required for the trimerization of type XIII collagen and is conserved in other collagenous transmembrane proteins. EMBO J 19: 5051-5059.

Soininen R, Huotari M, Ganguly A, Prockop DJ & Tryggvason K (1989) Structural organization of the gene for the α1 chain of human type IV collagen. J Biol Chem 264: 13565-13571.

Srivastava AK, Durmowicz MC, Hartung AJ, Hudson J, Ouzts LV, Donovan DM, Cui CY & Schlessinger D (2001) Ectodysplasin-A1 is sufficient to rescue both hair growth and sweat glands in Tabby mice. Hum Mol Genet 10: 2973-2981.

Srivastava AK, Pispa J, Hartung AJ, Du Y, Ezer S, Jenks T, Shimada T, Pekkanen M, Mikkola ML, Ko MS, Thesleff I, Kere J & Schlessinger D (1997) The Tabby phenotype is caused by mutation in a mouse homologue of the EDA gene that reveals novel mouse and human exons and encodes a protein (ectodysplasin-A) with collagenous domains. Proc Natl Acad Sci USA 94: 13069-13074.

Stacey A, Bateman J, Choi T, Mascara T, Cole W & Jaenisch R (1988) Perinatal lethal osteogenesis imperfecta in transgenic mice bearing an engineered mutant pro-α1(I) collagen gene. Nature 332: 131-136.

Sugimoto M, Oohashi T & Ninomiya Y (1994) The genes COL4A5 and COL4A6, coding for basement membrane collagen chains α5(IV) and α6(IV), are located head-to- head in close proximity on human chromosome Xq22 and COL4A6 is transcribed from two alternative promoters. Proc Natl Acad Sci USA 91: 11679-11683.

Sumiyoshi H, Laub F, Yoshioka H & Ramirez F (2001) Embryonic expression of type XIX collagen is transient and confined to muscle cells. Dev Dyn 220: 155-162.

Sumiyoshi H, Mor N, Lee SY, Doty S, Henderson S, Tanaka S, Yoshioka H, Rattan S & Ramirez F (2004) Esophageal muscle physiology and morphogenesis require assembly of a collagen XIX-rich basement membrane zone. J Cell Biol 166: 591-600.

Sund M (2001) Type XIII collagen: regulation of cardiovascular development and malignant transformation in transgenic mice. Acta Univ. Oul. D 656: 48-58.

Sund M, Väisänen T, Kaukinen S, Ilves M, Tu H, Autio-Harmainen H, Rauvala H & Pihlajaniemi T (2001a) Distinct expression of type XIII collagen in neuronal structures and other tissues during mouse development. Matrix Biol 20: 215-231.

Sund M, Ylönen R, Tuomisto A, Sormunen R, Tahkola J, Kvist AP, Kontusaari S, Autio-Harmainen H & Pihlajaniemi T (2001b) Abnormal adherence junctions in the heart and reduced angiogenesis in transgenic mice overexpressing mutant type XIII collagen. EMBO J 20: 5153-5164.

Takahara K, Hoffman GG & Greenspan DS (1995) Complete structural organization of the human α1(V) collagen gene (COL5A1): divergence from the conserved organization of other characterized fibrillar collagen genes. Genomics 29: 588-597.

Taverna D, Disatnik MH, Rayburn H, Bronson RT, Yang J, Rando TA & Hynes RO (1998) Dystrophic muscle in mice chimeric for expression of α5 integrin. J Cell Biol 143: 849-859.

Tenhunen K, Laan M, Manninen T, Palotie A, Peltonen L & Jalanko A (1995) Molecular cloning, chromosomal assignment, and expression of the mouse aspartylglucosaminidase gene. Genomics 30: 244-250.

Thompson JD, Higgins DG & Gibson TJ (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res 22: 4673-4680.

Thompson S, Clarke AR, Pow AM, Hooper ML & Melton DW (1989) Germ line transmission and expression of a corrected HPRT gene produced by gene targeting in embryonic stem cells. Cell 56: 313-321.

Tinsley J, Deconinck N, Fisher R, Kahn D, Phelps S, Gillis JM & Davies K (1998) Expression of full-length utrophin prevents muscular dystrophy in mdx mice. Nat Med 4: 1441-1444.

Tikka L, Elomaa O, Pihlajaniemi T & Tryggvason K (1991) Human α1(XIII) collagen gene. Multiple forms of the gene transcripts are generated through complex alternative splicing of several short exons. J Biol Chem 266: 17713-17719.

Tikka L, Pihlajaniemi T, Henttu P, Prockop DJ & Tryggvason K (1988) Gene structure for the α1 chain of a human short-chain collagen (type XIII) with alternatively spliced transcripts and translation termination codon at the 5' end of the last exon. Proc Natl Acad Sci USA 85: 7491-7495.

Timpl R & Brown J (1996) Supramolecular assembly of basement membranes. BioEssays 18: 123-132.

Tu H (2004) Type XIII collagen: structural and functional characterization of the ectodomain and identification of the binding ligands. Acta Univ. Oul. D 782: 44-50.

Tu H, Sasaki T, Snellman A, Göhring W, Pirilä P, Timpl R & Pihlajaniemi T (2002) The type XIII collagen ectodomain is a 150-nm rod and capable of binding to fibronectin, nidogen-2, perlecan, and heparin. J Biol Chem 277: 23092-23099.

Tuckwell D (2002) Identification and analysis of collagen alpha 1(XXI), a novel member of the FACIT collagen family. Matrix Biol 21: 63-66.

Vachon PH, Xu H, Liu L, Loechel F, Hayashi Y, Arahata K, Reed JC, Wewer UM & Engvall E (1997) Integrins (alpha7beta1) in muscle function and survival. Disrupted expression in merosin-deficient congenital muscular dystrophy. J Clin Invest 100: 1870-1881.

Vuorio E & de Crombrugghe B (1990) The family of collagen genes. Annu Rev Biochem 59: 837-872.

Väisänen MR, Väisänen T & Pihlajaniemi T (2004) The shed ectodomain of type XIII collagen affects cell behaviour in a matrix-dependent manner. Biochem J 15: 685-693.

Weir AP, Burton EA, Harrod G & Davies KE (2002) A- and B-utrophin have different expression patterns and are differentially up-regulated in mdx muscle. J Biol Chem 277: 45285-45290.

Williams RL, Hilton DJ, Pease S, Willson TA, Stewart CL, Gearing DP, Wagner EF, Metcalf D, Nicola NA & Gough NM (1988) Myeloid leukaemia inhibitory factor maintains the developmental potential of embryonic stem cells. Nature 336: 684-687.

de Winther MP, van Dijk KW, Havekes LM & Hofker MH (2000) Macrophage scavenger receptor class A: A multifunctional receptor in atherosclerosis. Arterioscler Thromb Vasc Biol 20: 290-297.

Xu H, Christmas P, Wu XR, Wewer UM & Engvall E (1994) Defective muscle basement membrane and lack of M-laminin in the dystrophic dy/dy mouse. Proc Natl Acad Sci USA 91: 5572-5576.

Ylönen R, Kyrönlahti T, Sund M, Saarela J, Vuoristo J, Ilves M, Lehenkari P, Tuukkanen J & Pihlajaniemi T. Type XIII collagen strongly affects bone formation in transgenic mice. Submitted.

Zhang X, Zhou J, Reeders ST & Tryggvason K (1996) Structure of the human type IV collagen COL4A6 gene, which is mutated in Alport syndrome-associated leiomyomatosis. Genomics 33: 473-479.