

We are IntechOpen, the world's leading publisher of Open Access books Built by scientists, for scientists

4,800

Open access books available

122,000

International authors and editors

135M

Downloads

Our authors are among the

154

Countries delivered to

TOP 1%

most cited scientists

12.2%

Contributors from top 500 universities



WEB OF SCIENCE™

Selection of our books indexed in the Book Citation Index
in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?
Contact book.department@intechopen.com

Numbers displayed above are based on latest data collected.

For more information visit www.intechopen.com



Inherited Connective Tissue Disorders of Collagens: Lessons from Targeted Mutagenesis

Christelle Bonod-Bidaud and Florence Ruggiero

*Institut de Génomique Fonctionnelle de Lyon, ENS de Lyon, UMR CNRS 5242,
University Lyon 1,
France*

1. Introduction

The extracellular matrix (ECM) is the cell structural environment in tissues and organs. The ECM is a dynamic structure that it is constantly remodelled. It contributes to tissue integrity and mechanical properties. It is also essential for maintaining tissue homeostasis, morphogenesis and differentiation, which it does, through specific interactions with cells. The ECM is composed of a mixture of water and macromolecules classified into four main categories: collagens, proteoglycans, elastic proteins, and non-collagenous glycoproteins (also called adhesive glycoproteins). The nature, concentration and ratio of the different ECM components are all important factors in the regulation of the assembly of complex tissue-specific networks tuned to meet mechanical and biological requirements of tissues.

Collagens form a superfamily of 28 trimeric proteins, distinguishable from the other ECM components by their particular abundance in tissues (collagens represent up to 80-90% of total proteins in skin, tendon and bones) and their capacity to self-assemble into supramolecular organized structures (the best known being the banded fibers). The collagen superfamily is highly complex and shows a remarkable diversity in structure, tissue distribution and function (Ricard-Blum and Ruggiero, 2005).

The importance of collagens has been illustrated by the wide range of mutations in collagen genes that result in minor and severe human diseases. Various mutations (point, null or structural mutations, insertions, exon skipping, deletions) in genes encoding collagens are known to be responsible for a large spectrum of human disorders (*e.g.*, Ehlers-Danlos syndrome, epidermolysis bullosa, chondrodysplasia, osteogenesis imperfecta, Alport syndrome, Bethlem myopathy, Ulrich congenital muscular dystrophy, Fuchs' endothelial dystrophy, Knobloch syndrome) that affect different tissues and organs, such as skin, blood vessels, cartilage, bones, kidney, muscle, cornea and retina. Considering the variety of collagen-related diseases and the complexity of collagen biology, there is a clear need to understand how mutations alter collagen synthesis, cell trafficking, cell and molecular interactions to result in tissue dysfunction. In the eighties targeted mutagenesis emerged as a new approach to help establish the structure-function relationship of collagens. Along with the emergence of protein engineering and genetically modified mice, site-directed mutagenesis has become instrumental in understanding the physiopathology of diseases, as well as in developing new and specific therapies and drugs for the treatment of human

diseases. To date about 20 distinct genes encoding collagen chains have been ablated (by knock-out mutations) in mice or are involved in naturally occurring mutations. Only a few knock-in modified mice has been generated, in which a single point mutation or an exon deletion, for example, has been generated in a specific gene. This is likely due to the very large size of collagen genes. Site-directed knock-in mutations in mice have often proven to be more useful than knock-out mutations (which inactivate genes) for the analysis of the genotype-phenotype relationship, since small mutations represent the primary bases of inherited diseases.

The aim of this chapter is to describe the use of targeted mutagenesis in the understanding of the pathophysiology of inherited connective tissue disorders. Specifically we are concerned with mutations in collagen genes. We will focus on the use of site-directed mutagenesis to analyze the causative effects of human-identified collagen gene mutations. Recombinant molecules were used to analyze the effects of these mutations on collagen structure, biosynthesis, posttranslational modifications and interactions with binding partners and cells. This work has considerably improved our knowledge in development and in human disorders. These results will then be compared with the limited information about the introduction of subtle targeted mutations into murine collagen genes.

2. The collagen superfamily at a glance

The 28 members of the collagen superfamily exhibit considerable complexity and diversity in structure, assembly and function. However, collagens also share common features. (i) All members are modular proteins composed of collagenous (COL) domains flanked by non collagenous (NC) domains or linker regions. (ii) They are trimeric molecules formed by the association of three identical or different α -chains, which are characterized by repetitions of the G-x-y tripeptide (with the x and y positions often occupied by proline and hydroxyproline, respectively). (Abbreviations and single-letter codes for amino acids are given in Table 1 of the chapter by Figurski *et al.*) (iii) They are able to assemble into supramolecular aggregates in the extracellular space, although this property has not been proven for all recently identified collagen members. Collagens also undergo various post-translational modifications, including proteolytic processing, fibril formation, reticulation, shedding of transmembrane collagens and production of functional domains (also called matricryptins) (Ricard-Blum and Ruggiero, 2005). The mechanisms of collagen biosynthesis are far from being completely understood. Our knowledge is primarily based on the biosynthesis of fibril-forming collagens. Triple-helix formation commonly starts at the C-terminus (C-propeptide) of the α -pro-chain and proceeds toward the N-terminus (N-propeptide) in a zipper-like fashion. Prior to and simultaneously with triple-helix formation, specific prolines and lysines are chemically modified by addition of hydroxyl group. These modifications play a pivotal role in stabilization and resistance to temperature. Completed trimeric procollagens are secreted from the cells, proteolytically processed and assemble into collagen fibrils (Ricard-Blum and Ruggiero, 2005).

Based on their structure and supramolecular organization, collagens have been divided into several subfamilies (Myllyharju and Kivirikko, 2001). They are (i) the fibril-forming collagens I, II, III, V, XI, XXIV and XVII, which share the capacity to assemble into organized fibrils; (ii) the network-forming collagens IV, VIII and X and the FACIT (Fibril-Associated Collagen with Interrupted Triple-helix collagens) collagens IX, XII, XIV, XVI, XIX, XX, XI

and XXII, which are known to mediate protein-protein interactions; (iii) the basement membrane multiplexin (multiple triple-helix domains and interruptions) collagens XV and XVIII; (iv) the transmembrane collagens, including the neuronal XXV collagen and types XIII, XVII, XXIII; and finally (v) other unconventional collagens, such as the anchoring fibrils collagen VII and the ubiquitous collagen VI, which assembles into characteristic beaded filaments (Table 1).

The length of the triple helical domains varies noticeably among different collagen types. Fibril-forming collagens consist of a long central COL domain with about 1000 amino acids (330 G-x-y tripeptide repeats), flanked by small terminal globular extensions (NC domains). After proteolytic processing of the N and C-terminal extensions, the mature molecules aggregate into highly ordered fibrils with a banded pattern observable by transmission electron microscopy. In other collagens, the COL domains are shorter and/or contain interruptions. The NC domains can represent the main part of the molecule, as for the FACIT collagen XII. Most, if not all, collagen types are recognized by specific cell receptors, such as the major ECM integrin receptors, collagen-specific discoidin domain receptors (DDR) and the transmembrane proteoglycan syndecans (Humphries *et al.*, 2006; Xian *et al.*, 2010; Leitinger *et al.*, 2007). Through various interactions with these cell receptors, collagens can induce intracellular pathways directly or indirectly and regulate cell functions, such as migration, proliferation and differentiation. Certain collagens can also bind to growth factors and control their bioavailability by acting as reservoirs. The controlled release of growth factors by proteolytic activity or expression of a splice variant that does not contain the binding site controls morphogenesis, as described for the cartilage collagen II (Zhu *et al.*, 1999).

3. A large spectrum of mutations in collagen genes causes inherited disorders

A myriad of mutations has been characterized in collagen genes (Table 1). The function of the gene product and its tissue localization are criteria that lead to a number of inherited connective tissues disorders (reviewed in Bruckner-Tuderman and Bruckner, 1998; Bateman *et al.*, 2009). Typically mutations in collagen genes are null-mutations, *i.e.*, those resulting in the translation of an α -chain that cannot assemble into a triple helix and is consequently degraded intracellularly. Null mutations reduce the overall quantity of collagen in tissue and generally cause a human disorder. Small deletions and base substitutions can lead to synthesis of a mutated α -chain that is able to form a triple helix. The molecule is secreted, but its structure is compromised for supramolecular assembly, which normally occurs in the extracellular space. *In fine* collagen gene mutations result in defective matrix assembly and organization that in turn can affect cell function (Figure 1). In cases of large multimeric molecules, such as collagens, dominant-negative mutations can be more deleterious than null mutations. However, a growing body of evidence shows that the synthesis of a large quantity of abnormal collagen molecules in cells during development can induce endoplasmic reticulum stress, with consequences ranging from cell recovery to death (Tsang *et al.*, 2010). The correlation between phenotype severity and the location of a point mutation in the gene is not clear. However, a mutation located in the coding region for the amino-terminus of the fibrillar collagen triple helix generally results in a mild phenotype, whereas a mutation in the coding region for the carboxy-terminus of the molecule is often lethal. This

observation may be related to the C- to N-terminus directional propagation of the triple helix and the role of the C-propeptides in α -chain registration and triple helix nucleation. The nature of glycine substitution in the G-x-y repeats and the neighboring amino-acid sequence may have different biochemical and clinical consequences. These consequences include (i) delay of the triple-helix formation and over-glycosylation (Raghunath *et al.*, 1994); (ii) alteration of procollagen processing (Lightfoot *et al.*, 1994); (iii) retention of unfolded abnormal proteins intracellularly, leading to ER stress; and (iv) formation of abnormal unstable trimeric molecules, leading to disrupted fibrillogenesis.

The presence of a glycine in every third position is critical for triple-helix formation, since only glycine, the smallest amino acid, fits into the center of the triple helix. The majority of dominant-negative mutations in collagen genes are due to replacements of one of the glycines in the collagenous domains of the α -chains with a larger amino acid. Glycine substitution mutations in collagen genes underlie heritable connective tissue diseases, such as osteogenesis imperfecta (OI), chondrodysplasias, certain subtypes of Ehlers-Danlos syndrome (EDS), or Alport's syndrome (reviewed in Bruckner-Tuderman and Bruckner, 1998; Bateman *et al.*, 2009). Since a non-glycine amino acid does not easily fit into the interior space of the triple helix, helix formation is distorted, thereby affecting its structure and stability and impeding fibrillogenesis. Delay in triple-helix formation can result in over-modification and may affect collagen function.

Osteogenesis imperfecta (OI), also known as brittle bone disease, is caused by mutations in genes for collagen I, the most abundant collagen in organisms. OI is characterized by fragile bones that break easily and reduced bone mass. Most OI cases are believed to be associated with glycine substitution mutations in the *COL1A1* or *COL1A2* genes. Over 200 mutations have been reported for the *COL1A1* (located on chromosome 17) and *COL1A2* (located on chromosome 7) genes, which code for the collagen I pro- α 1 and pro- α 2 chains, respectively. This fact may explain the wide range of clinical characteristics and degrees of severity that are seen in the disease (Kuivaniemi *et al.*, 1991, Byers and Steiner 1992, Dalgleish, 1998). Because collagen I is found in other tissues of the body, OI has non-skeletal manifestations as well. People with OI may also suffer from muscle weakness, hearing loss, fatigue, joint laxity, distensible skin, or dentinogenesis imperfecta. The fibril-forming collagen I is mostly synthesized as the $[\alpha 1(I)]_2\alpha 2(I)$ heterotrimer chain, though a minor form $[\alpha 1(I)]_3$ is expressed in embryonic tissues. *COL1A1* and *COL2A1* are both susceptible to various mutations responsible for the production of quantitatively or qualitatively deficient fibrils. The clinical severity of OI relates to the extent of the conformational change in the collagen triple helix induced by the glycine substitution. These mutations result in altered fibrillogenesis. However, no general mechanism can be drawn from genotype/phenotype analyses.

Collagen VII, encoded by *COL7A1*, is the major component of the anchoring fibrils at the dermo-epidermal junction (Burgeson, 1993). *COL7A1* gene mutations cause dystrophic epidermolysis bullosa (DEB), a skin-blistering disorder (Bruckner-Tuderman, 1999). Approximately 200 mutations of *COL7A1* have been characterized, leading to a very high molecular heterogeneity of collagen VII defects (Dunnill *et al.*, 1996). Almost all cases of dominant DEB are caused by a glycine substitution in the triple helical region of collagen VII, and most of the mutations are unique to individual families. Some glycine substitutions in collagen VII interfere with biosynthesis of the protein in a dominant-negative manner, whereas others may lead to collagen VII retention within the rough endoplasmic reticulum.

Mutations in the *COL5A1* and *COL5A2* genes, encoding respectively the pro- α 1 and pro- α 2 chains of the fibril-forming collagen V, have been identified in approximately 50% of patients with a clinical diagnosis of classic Ehlers-Danlos syndrome (EDS) (Malfait *et al.*, 2010). Collagen V contains a third chain, pro α 3(V); but no mutation in *COL5A3* has been reported so far. Classic EDS is a heritable disorder of connective tissues characterized by skin hyperextensibility, fragile and soft skin, delayed wound healing with formation of atrophic scars, easy bruising, and generalized joint hypermobility. The majority of mutations lead to a non-functional *COL5A1* allele. One mutant *COL5A1* transcript showed a premature stop codon. A minority of mutations affect the structure of the central helical domain. In approximately one-third of patients, the disease is caused by a mutation leading to a non-functional *COL5A1* allele, resulting in collagen V haploinsufficiency. Structural mutations in *COL5A1* or *COL5A2*, resulting in the production of a functionally defective protein, account for a small proportion of patients.

Collagen V is a quantitatively minor fibril-forming collagen that co-polymerizes with collagen I to form heterotypic fibrils (Fichard *et al.*, 1995). Co-polymerisation has a critical role in the nucleation and growth of fibrils in tissues. A collagen V feature is to retain in the mature molecule a major part of the α 1(V) N-propeptide which projects beyond the surface of collagen fibrils. This domain was proposed to limit heterotypic fibril growth by steric hindrance and electrostatic interactions (Linsenmayer *et al.*, 1993). Skin biopsies revealed abnormalities in fibril formation (altered diameter, contour, or shape of dermal fibrils). However, abnormalities of fibril structure affected less than 5% of fibrils (reviewed in Fichard *et al.*, 2003). Moreover, the clinical phenotype of classical EDS supports an important role of collagen V in the biomechanical integrity of the skin, tendon and ligaments, although collagen V is only a minor component of the affected tissues. Thus, collagen V may be involved in functions other than the control of fibril growth in classical EDS. A likely hypothesis is that collagen V might be involved in the physiopathology of EDS through interactions with other fibril-associated components and/or with cell receptors. Along this line, it has been shown that mutations in the genes for the collagen V-binding partners, tenascin-X (*TNXB* gene) and collagen I (*COL1A1* gene), resulted in EDS (Lindor and Bristow, 2005).

Although mutant gene products are thought to impair matrix structure and assembly that eventually alters tissue function, growing evidence links ER stress and the unfolded protein response (UPR) to the initiation and progression of a broad repertoire of connective tissue disorders, including those caused by collagen gene mutations. Some mutant chains cannot be incorporated into procollagen molecules, consequently causing protein degradation with important downstream effects. Misfolded or slowly folding collagens are retained within the endoplasmic reticulum (ER) and ultimately targeted for degradation by a mechanism initially called "protein suicide." Because connective-tissue cells typically produce large quantities of collagens, the contribution of ER stress induced by misfolded collagens in disease pathogenesis has certainly been underrated. The current knowledge on the implications of unfolded protein response and ER stress in connective tissue diseases has been recently reviewed, and readers are referred to these reviews for further reading (Boot-Handford and Briggs, 2010; Tsang *et al.*, 2010). Notably, mutations in genes encoding collagen I (*COL1A1* and *COL1A2*) (osteogenesis imperfecta), collagen II (*COL2A1*) (spondyloepiphyseal dysplasia), and collagen X (*COL10A1*) (metaphyseal

Collagen subfamily	Collagen type	Associated diseases	Gene	Modified mice	References (mouse models)
Fibril forming collagens	I	osteogenesis imperfecta	COL1A1	knock-out	Löhler <i>et al.</i> , 1984
			COL1A2	single mutation	Forlino <i>et al.</i> , 1999
	II	chondrodysplasias	COL2A1	knock-out	Daley <i>et al.</i> , 2010
			COL3A1	knock-out	Aszodi <i>et al.</i> , 1998
			COL5A1	knock-out	Liu <i>et al.</i> , 1997
V	vascular Ehlers-Danlos syndrome	COL5A1	knock-out	Wenstrup <i>et al.</i> , 2004	
		COL5A2	knock-in	Andrikopoulos <i>et al.</i> , 1995	
FACIT	XI	nd	COL5A3	knock-out	Huang <i>et al.</i> , 2011
			COL11A1	knock-out	Wenstrup <i>et al.</i> , 2011
	IX	Stickler syndrome	COL11A2	knock-out	Mc Guirt <i>et al.</i> , 1999
			COL9A1	knock-out	Fässler <i>et al.</i> , 1994
			COL9A2	functional knock-out	Hagg <i>et al.</i> , 1997
Multiplexins	XII	nd	COL9A3	knock-out	Izu <i>et al.</i> , 2011
			COL12A1	knock-out	Ansorge <i>et al.</i> , 2009
	XIV	nd	COL14A1	knock-out	Sumiyoshi <i>et al.</i> , 2004
			COL19A1	knock-out	Eklund <i>et al.</i> , 2001
	XV	nd	COL15A1	knock-out	Seppinen <i>et al.</i> , 2008
XVIII	Knobloch syndrome	COL18A1	knock-out		
Network forming collagens	IV	nd	COL4A1	nd	
			COL4A2	nd	
	VIII	Alport syndrome	COL4A3	knock-out	Cosgrove <i>et al.</i> , 1996
			COL4A4	knock-out	Lu <i>et al.</i> , 1999
			COL4A5	knock-out	Rheault <i>et al.</i> , 2004
			COL4A6	nd	
X	corneal dystrophia	COL8A1	knock-out	Hopfer <i>et al.</i> , 2005	
		COL8A2	knock-out	Hopfer <i>et al.</i> , 2005	
Beaded-filament forming collagen	VI	Bethlem myopathy and Ullrich congenital muscular dystrophy	COL10A1	knock-out	Kwan <i>et al.</i> , 1997
			COL6A1	knock-out	Bonaldo <i>et al.</i> , 1998
	X	nd	COL6A2	nd	
			COL6A3	nd	
			COL6A4	nd	
			COL6A5	nd	
Anchoring collagens	VII	epidermolysis bullosa	COL7A1	knock-out	Heinonen <i>et al.</i> , 1999
			COL7A1	knock-out	Tanimura <i>et al.</i> , 2011
Transmembrane collagens	XVII	epidermolysis bullosa	COL17A1	knock-out	
			COL17A1	knock-out	

nd, not determined

Table 1. Collagen types, associated-diseases and mouse models.

chondrodysplasia) have been shown to induce ER dilatation in patient cells. Mutations that affect the triple helix, the C-propeptide for the fibril-forming collagens, and splice donor sites, as well as single amino-acid substitutions, were shown to cause ER stress. Recently, mutations that affected the signal peptide domain of the pro α 1(V)-collagen chain were shown to cause classic EDS. The signal peptides are the addresses of proteins destined for secretion. The mutant procollagen V is retained within the cell, leading to a collagen V haploinsufficiency and altered collagen fibril formation. It is probable that the signal peptide mutation also causes accumulation of the mutated protein within the ER and eventually to ER stress, as described for other collagen-related disorders (Symoens *et al.*, 2009).

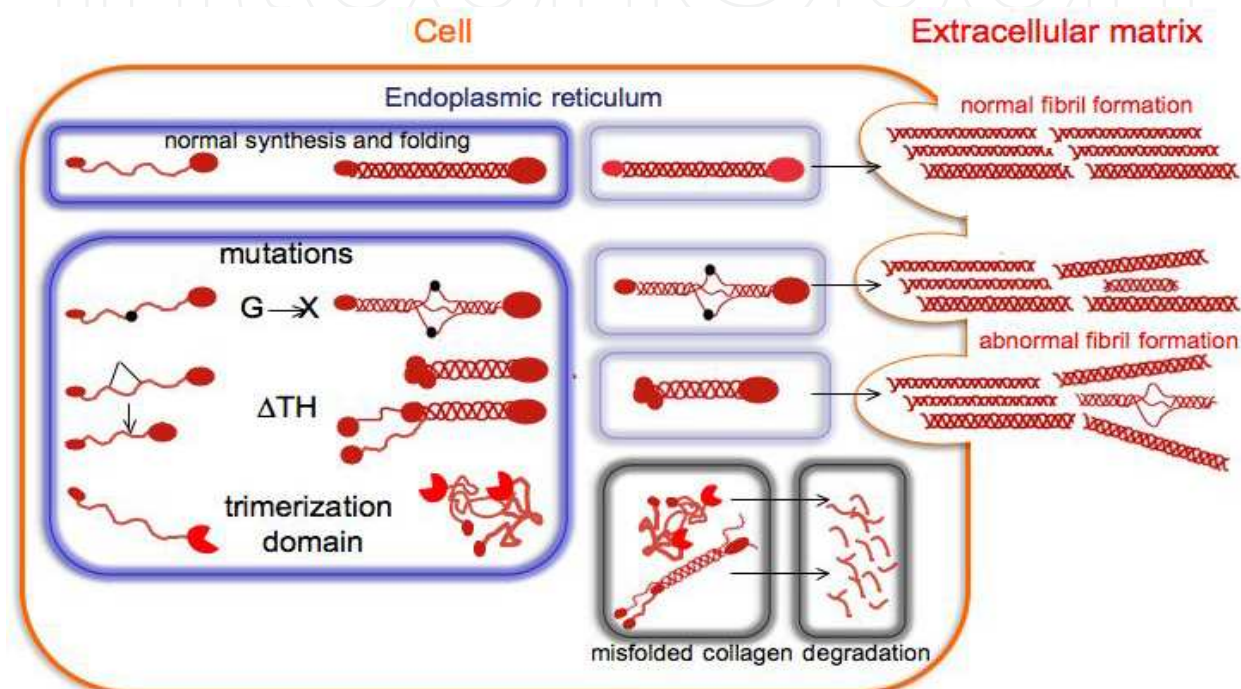


Fig. 1. Schematic diagram illustrating the biological consequences of point mutations or small deletions in collagen genes on chain synthesis, protein folding and subsequent fibril assembly in the extracellular matrix.

Mutations in the three major collagen VI genes (*COL6A1*, *COL6A2* and *COL6A3*) cause multiple muscle disorders, including the severe Ullrich congenital muscular dystrophy (UCMD) and the mild Bethlem myopathy, which is characterized by muscle weakness with striking joint laxity and progressive contractures. Three genetically distinct novel chains α 4(VI), α 5(VI), and α 6(VI) have recently been identified; but very little is known about their molecular assembly and biosynthesis and their possible involvement in human diseases (Gara *et al.*, 2011). Collagen VI biosynthesis is a complex multistep process. Monomer formation results from the heterotrimeric association of the three chains [α 1(VI), α 2(VI), and α 3(VI)] encoded by the *COL6A1*, *COL6A2* and *COL6A3* genes. Monomers first assemble into antiparallel dimers that associate laterally to form tetramers stabilized by disulphide bonds. The tetramers associate linearly to form the unique beaded filaments, the ultimate step of collagen VI biosynthesis. Dominant and recessive autosomal mutations in *COL6A1*, *COL6A2*, and *COL6A3* primarily result in dysfunctional microfibrillar collagen VI in muscle extracellular matrix. However they also affect other connective tissues, such as skin and

tendons. Different mutations have been shown to have variable effects on protein assembly, secretion, and its ability to form a functioning extracellular network. As observed in other collagen-related diseases, glycine-substitution mutations in *COL6A1*, *COL6A2*, or *COL6A3* that disrupt the triple-helix motif constitute a frequent pathogenic mechanism. Triple-helix distortion may exert a dominant-negative effect by reducing the ability of mutated monomers to form beaded filaments. Interestingly, mitochondrial dysfunction was implicated in the pathogenesis of a myopathic phenotype. Muscles lacking collagen VI are characterized by the presence of a dilated sarcoplasmic reticulum and dysfunctional mitochondria. This condition triggers apoptosis and leads to myofiber degeneration. Recently, it was shown that the persistence of abnormal organelles and apoptosis observed in some congenital muscular dystrophies are caused by defective activation of the autophagic machinery. Autophagy has a key role in the clearance of damaged organelles and in the turnover of cell components and is thus essential for tissue homeostasis. Recently, 56 novel mutations have been described, allowing a clinical classification and revealing the complexity of genotype-phenotype relationships (Briñas *et al.*, 2010).

The paucity of evidence-based data regarding correlations of genotype and phenotype is in part due to the large spectrum of mutations reported for the collagen genes [*e.g.*, about 200 mutations for the collagen I genes responsible for OI (Dagleish, 1998); 160 mutations in the *COL4A5* gene encoding collagen IV $\alpha 5$ chain responsible for Alport syndrome; 200 mutations in *COL7A1* responsible for EDB]. Things are not as simple as one gene-numerous mutations-one phenotype. Sometimes a combination of a mutation for a connective tissue disorder and a specific collagen gene mutation will result in another disease. Some patients with UCMD show clinical characteristics typical of classical disorders of connective tissue, such as EDS. Ultrastructure of skin biopsy samples from patients with UCMD showed alterations of collagen fibril morphology in skin that resemble those described in patients with EDS (Kirschner *et al.*, 2005). Recently, using the yeast two-hybrid approach, we showed a direct interaction between collagen V and collagen VI that may nicely explain the overlap of UCMD and classic EDS (Symoens *et al.*, 2011). Unexpectedly an arginine-->cysteine substitution localized at position 134 of the $\alpha 1(I)$ collagen chain resulted in classical EDS (Nuyntick *et al.*, 2000). This finding is indicative of genetic heterogeneity in collagen-related disorders.

A powerful approach to study the biochemical consequences of mutation and the protein structure/function relationship is to engineer a specific mutation into a functional domain of the molecule. Targeted mutagenesis approaches, including the use of alanine-scanning mutagenesis techniques, have led to important insights into the effects of collagen mutations on protein structure and function. A major limitation of mutagenesis strategies to investigate collagens is the large number of collagen gene mutations to be investigated in order to have a better understanding of the molecular mechanisms of "collagenopathies." Knowledge about the impact of collagen mutations has also been hampered by the technical difficulty of introducing targeted mutations of very large collagen genes into mice.

4. Lessons from site-directed mutagenesis of recombinant collagen genes and derived fragments

Production of a recombinant collagen gene represents a powerful technique to introduce a human mutation into the gene of interest by site-directed mutagenesis. It allows one to

analyze the impact of the mutation on collagen assembly and secretion. Collagen biosynthesis is a complex multistep process that takes place in the intracellular and extracellular space and includes various post-translational modifications, such as prolyl- and lysyl-hydroxylation, glycosylation, trimerization, proteolytic processing, polymerization and cross-links. Because of recombinant technology, these large multimeric proteins have been produced in large amounts in almost all existing expression systems (Ruggiero and Koch, 2008). This technological breakthrough enabled researchers to analyze in detail the effects of collagen mutations on biosynthesis, molecular and cell interactions, processing and, in some cases, self-assembly. Researchers can also address the question of the correlation of genotype, protein structure and function.

Mutations occurring in collagen I genes are the most extensively studied mutations among all collagen types. A first set of experiments substituted glycine 859 of the pro α 1(I) chain with cysteine or arginine by site-directed mutagenesis to reproduce two mutations identified in OI patients. In order to study the expression of the mutant molecule in the presence or absence of the wild-type pro α 1(I) chain, the mutated constructs were transfected into normal fibroblasts to look for a dominant-negative effect in the presence of the wild-type gene or in fibroblasts isolated from Mov13 homozygous mice (referred to as Mov13 fibroblasts hereafter), whose cells carry a provirus that prevents transcription initiation of the natural pro α 1(I) gene (Schnieke *et al.*, 1987). In agreement with observations of collagen I in OI patients, the mutated collagens were poorly secreted from the cells and exhibited reduced thermal stability and increased sensitivity to degradation. This supported the idea that the strict preservation of the G-x-y triplets is absolutely required for proper formation of the triple helix.

The integrity of the C-propeptide is pivotal for the trimerization of all fibril-forming collagens. The C-propeptides of the pro α 1(I) and pro α 2(I) chains contain an Asn-Ile-Thr sequence. That sequence fits a consensus sequence for the addition of N-linked oligosaccharides. To analyze the role of this post-translational modification, the asparagine residue of the pro α 1(I) chain was changed to glycine by site-directed mutagenesis. The expression of the corresponding molecule was analyzed in transfected normal and Mov13 fibroblasts (Lamandé and Bateman, 1995). The mutation did not impair heterotrimeric assembly and secretion of hybrid procollagen I into the extracellular space. Only a slight effect on C-proteinase cleavage efficiency was observed with the unglycosylated molecule. To circumvent the difficulty of producing a large repertoire of full-length mutated collagens I in order to undertake a genotype/phenotype analysis, a recombinant trimeric mini-collagen I was recently expressed in an *Escherichia coli* system. Recombinant mini-collagens can be obtained by fusing the sequence encoding a fragment of the pro α 1(I) chain triple-helix to the sequence encoding the C-terminal domain (called "foldon") of the bacteriophage T4 fibrin, which is capable of trimerization (Xu *et al.*, 2008). Two mutations (G901S and G913S), corresponding to mild and severe types of OI, respectively, were introduced into the recombinant mini-collagen I. Biophysical measurements and protease cleavage analysis revealed that the G913S mutant chain resulted in the formation of an unstable collagen I triple helix by disrupting salt bridges important for maintaining the chains in a triple-helix conformation (Yang *et al.*, 1997; Xu *et al.*, 2008). A very recent study utilized a recombinant bacterial collagen to develop a mutagenesis scheme in which a glycine residue within the triple-helix sequence is substituted with arginine or serine. The purpose was to analyze the

positional effect of glycine mutations on triple-helix formation and stability (Cheng *et al.*, 2011). Interestingly, all glycine mutations provoked a significant delay in the triple-helix formation. However, a more severe defect was observed when the mutation was located near the trimerization domain of the triple-helix where folding is initiated.

COL7A1 mutations cause dystrophic epidermolysis bullosa (DEB), a skin blistering disorder. Woodley and collaborators (2008) have used site-directed mutagenesis to elucidate the effect of human mutations on the function of collagen VII, which is the major component of the epidermal anchoring fibrils. To undertake a comprehensive analysis of the impact of human mutations in the formation, folding and stability of collagen VII and, particularly relevant to the DEB phenotype, its effect on cell attachment and migration, four distinct substitutions occurring in collagen VII (G2049E, R2063W, G2569R, and G2575R) were introduced using *COL7A1* cDNA. The authors demonstrated that the G2049E and R2063W mutants caused local destabilization of the triple helix and reduced the capability of collagen VII to elicit cell adhesion and migration. The G2569R and G2575R mutants interfered with triple-helix formation and stability. Alterations of protein stability and/or cell attachment to collagen VII mutants help explain the fragility of the dermal-epidermal junction observed in DEB patients. Naturally occurring *COL7A1* mutations were investigated in a separate study (Hammami-Hausli *et al.*, 1998). As commonly described for glycine-substitution mutants of collagens, the authors showed that three glycine substitutions located in the same triple-helix portion affected folding, stability and secretion of procollagen VII in a dominant-negative manner. However, the glycine substitution G1519D located in another segment of the triple helix had no effect on procollagen VII secretion or its ability to anchor fibril assembly. These data showed that the biological impact of glycine substitutions can depend on their position within the triple helix, as shown for collagen I (Cheng *et al.*, 2011).

Human collagen IV mutations, thought to affect the biosynthesis of this basement membrane collagen, were extensively investigated. These mutations were known to cause Alport syndrome, a severe renal disease leading eventually to kidney failure. Collagen IV chains, $\alpha 1(\text{IV})$ - $\alpha 6(\text{IV})$, are encoded by 6 genes, *COL4A1*-*COL4A6*, respectively. Although mutations have also been identified in *COL4A3* and *COL4A4*, about 30% of known missense mutations occur in the *COL4A5* gene, which encodes the human $\alpha 5(\text{IV})$ chain. Most of them are glycine substitutions. One glycine-substitution mutation in *COL4A5* could prevent correct α -chain folding or/and the association with other α -chains to form a stable triple helix. To address this question, the authors took advantage of the bacterial system. A DNA encoding a 22-kDa recombinant domain of the $\alpha 5(\text{IV})$ triple helix in its wild-type form or harboring the G1015V or G1030S mutations was expressed in *E. coli* (Wang *et al.*, 2004). The recombinant wild-type and mutant proteins were purified and assayed for changes in triple-helix assembly and stability by circular dichroism. The two different glycine-substitution mutants displayed different defects in the secondary structures of their protein products that matched with the severity of the patient phenotypes. However, the use of a bacterial system to analyze the effects of specific human mutations on mini-collagen assembly and stability presents several disadvantages. Because collagens are large multimeric proteins, full-length molecules cannot be produced in a bacterial host. Most importantly, the bacterial system is limited. Not all post-translational modifications needed for the triple-helix formation and stability, such as hydroxylation, glycosylation, and disulfide-bond formation, are present in bacteria. A few years later, the bacterial limitations were bypassed by the

development of the production of full-length recombinant collagen molecules in mammalian cells (Fichard *et al.*, 1997; Ruggiero and Koch, 2008). No less than eighteen human mutations (11 substitutions and 7 deletions) were introduced into the sequence encoding the trimerization NC1 domain of the $\alpha 5(\text{IV})$ chain gene. The constructs were transfected into cells together with constructs containing the wild-type sequences of $\alpha 3(\text{IV})$ and $\alpha 4(\text{IV})$ chains to analyze the impact of the mutations in the NC1 domain on the formation of the $\alpha 3\alpha 4\alpha 5$ collagen IV heterotrimer. Twelve out of 15 mutant chains did lose their capacity to assemble into heterotrimeric molecules. The three remaining mutants formed heterotrimers, but the mutations prevented their secretion into the extracellular space (Kobayashi *et al.*, 2008). The authors nicely demonstrated, using site-directed mutagenesis, that amino acid substitutions in the $\alpha 5(\text{IV})$ NC1 trimerization domain are specifically responsible for impairment of collagen IV heterotrimer assembly. This defect may be a main molecular mechanism for the pathogenesis of Alport syndrome. Interestingly, an interactome (a map of known and predicted molecular interactions, as well as phenotypic and structural landmarks) of collagen IV was recently constructed to identify functional and disease-associated domains and genotype-phenotype relationships (Parkin *et al.*, 2011). Construction of such interactomes will greatly improve our capacity to integrate all data from different site-directed mutagenesis experiments. This advance will greatly help our understanding of the molecular mechanisms underlying “collagenopathies”; and, consequently, it may lead to the development of specific treatments.

Collagens undergo a great variety of proteolytic modifications. The fate and functions of the released fragments derived from collagens are still under intensive investigation, but the consequences of mutations in the coding regions for the cleavage sites on collagen structure, self-assembly and function have not been investigated in detail. A large repertoire of proteinases is responsible for these processing interactions. Included among such enzymes are the ADAMTS (a disintegrin and metalloprotease with thrombospondin motifs) and the BMP-1/tolloid families of metalloproteinases and more recently the furin-like proprotein convertases (Ricard-Blum and Ruggiero, 2005). To investigate collagen processing, fastidious extraction and purification steps were often necessary to obtain limited amounts of unprocessed proteins and enzymes with full activity in order to undertake *in vitro* enzymatic assays. To circumvent this problem, we recently described a new cell system allowing a rapid and straightforward analysis of processing interactions. Our system relies on the use of site-directed mutagenesis. This strategy was particularly instrumental in analyzing the complex procollagen V processing during maturation. We showed it to be unique among the fibril-forming collagens (Bonod-Bidaud *et al.*, 2007). Collagen V is a minor fibrillar collagen that can be distinguished from the others by its capacity to control fibrillogenesis (Fichard *et al.*, 1995). In addition this molecule undergoes a particular form of processing; and it is involved in fundamental processes, such as development and human connective tissues disorders. The pro $\alpha 1(\text{V})$ N-terminus can be processed by the procollagen proteinases ADAMTS-2 and BMP-1 (Colige *et al.*, 2005; Bonod-Bidaud *et al.*, 2007), whereas the C-propeptide can be cleaved by furin and BMP-1 (Kessler *et al.*, 2001). The pro $\alpha 1(\text{V})$ C-propeptide furin cleavage site, which occurs immediately downstream of the recognition sequence RTRR, was double-mutated to alanine residues (R1584A/R1585A) to abolish furin cleavage. All constructs were introduced into cells, along with a BMP-1-expressing construct; and the cleavage products were directly analyzed in conditioned medium of the

transiently transfected cells. We were able to show that BMP-1 is capable of processing the $\alpha 1(V)$ C-propeptide in absence of furin activity (Bonod-Bidaud *et al.*, 2007). In the same way, the determinant for $\alpha 1(V)$ N-propeptide processing by BMP-1 activity was identified by introducing in the coding region for the cleavage site (S254/Q255-D256) three single mutations (S254A, Q255A and D256A), two double mutations (S254A/Q255N and Q255A/D256A) and one triple mutation (S254A/Q255A/D256A). The data highlighted the unexpected importance of the aspartic acid in the P2' position of the BMP-1 cleavage site (Bonod-Bidaud *et al.*, 2007). Processing, proteolytic release of functional domains and shedding of collagens are involved in fundamental processes. It is likely that substitutions located in the proteolytic cleavage sites may represent a molecular cause of connective tissues disorders. A reported mutation in the $\alpha 1(V)$ N-propeptide in one patient with classic EDS resulted in a protein product missing the sequence of exon 5 that encompasses the BMP-1 cleavage site. The abnormal-sized N-propeptide present in the mutated collagen V caused dramatic alterations in fibril structure (Takahara *et al.*, 2002).

5. Lessons from site-directed mutagenesis in mice

In vitro studies are useful and necessary approaches to understand the mechanisms of collagen biosynthesis and to establish structure-function relationship. However, they do not always reflect the normal and pathological *in vivo* situations. Genetically modified mice appear to be a powerful technique to better understand the physiopathology of connective tissue disorders. Several different genetically modified mice have been created during the last 10 years (reviewed in Aszódi *et al.*, 2006). This clearly opened doors to better understand collagen function in developing tissues and provide reliable mouse models for inherited collagen diseases. Along this line, a targeted disruption of *Col4a3* gene led to renal failure and eventually to the death of mice at 3-4 months of age (Cosgrove *et al.*, 1996; Miner and Sanes, 1996). This result is consistent with defects described for Alport disease.

In most cases, the gene of interest was disrupted and knock-out mice were preferably generated. Few transgenic mice harbouring point mutations or small deletion in collagen genes have been generated (Table 1). Naturally occurring mutations in mice disrupting collagen genes have also been identified and characterized. The *oim* mice present a spontaneously acquired deletion in the *Col1a2* gene that leads to an accumulation of $[\alpha(I)]_3$ collagen homotrimer in the extracellular matrix. These mice develop a phenotype similar to moderate OI in humans, providing a good model for this collagen disorder (Chipman *et al.*, 1993). It was shown that homozygous Mov13 embryos harboring an inactivated pro $\alpha 1(I)$ chain (due to the insertion of the Moloney murine leukaemia virus into the first exon of the *Col1a1* gene) died in utero around day 12 because of vascular failure (Löhler *et al.*, 1984). However, in 1999 Forlino *et al.* developed the first knock-in mouse model for human OI by introducing a G349C mutation into the *Col1a1* gene. Along this line, a knock-in mouse model for OI, harboring a point mutation (G610C) in *Col1a2* was recently created (Daley *et al.*, 2010). These mice had reduced body mass and bone strength and exhibited bone fracture susceptibility consistent with the clinical features of human OI. Thus, the G610C knock-in mouse represents a novel model for the study of OI pathogenesis and also for testing potential therapies for OI.

Another example concerns collagen V deficiency/dysfunction, which is responsible for Ehlers-Danlos syndrome (EDS). In the absence of the *Col5a1* gene, the mice died at the onset of organogenesis at approximately embryonic day 10 (Wenstrup *et al.*, 2004). Interestingly, a targeted deletion in the *Col5a2* gene, encoding the $\alpha 2(V)$ chain, recapitulated many of the clinical, biomechanical, morphologic, and biochemical features of the classical EDS. The deletion removes the sequence encoding the N-telopeptide (*pN*), a 20-residue region that confers flexibility to the N-terminal part of the molecule (Andrikopoulos *et al.*, 1995). A detailed study of the skin at the morphological, histological, ultrastructural and biochemical levels indicated that the *Col5a2* deletion impairs assembly and/or secretion of the $[\alpha 1(V)]_2\alpha 2(V)$ heterotrimer. Consequently, the $[\alpha 1(V)]_3$ homotrimer, and not the $[\alpha 1(V)]_2\alpha 2(V)$ heterotrimer, is the predominant species deposited into the matrix, which in turn severely impaired extracellular matrix organization (Chanut-Delalande *et al.*, 2004). These data underscored the importance of the collagen V $[\alpha 1(V)]_2\alpha 2(V)$ heterotrimer in dermal fibrillogenesis and can explain defects observed in the dermis of EDS patients.

6. Concluding remarks

Site-directed mutagenesis has been extensively used in collagen engineering and has shed light on collagen structure, expression, folding, secretion, interactions and self-assembly in the extracellular space. It also opened the way for the analysis of specific functional domains. It allowed the study of the wide variety of collagen types, including those expressed in trace amounts in tissues but nevertheless display pivotal functions. While it is true that site-directed mutagenesis has yielded important information on the functional consequences of a range of collagen mutations responsible for human diseases, only few studies have approached the consequences of collagen gene mutations on cell adaptation to ER stress. Collagen gene mutations affect protein synthesis, folding and secretion imbalance, which eventually induces ER stress. *In vitro* studies have been done on transfected cells, in which expression and trafficking of mutant collagen can be easily manipulated and analysed at the cellular level. The effects of gene manipulation can be studied *in vivo* using mice. The effect of collagen gene mutations on induction of an ER stress response could be straightforwardly addressed in the near future. It may be a key factor in pathogenesis (Boot-Handford and Briggs, 2010).

Mouse models are particularly useful for analysing the biological significance of collagens in pathological situations. Knock-out mice often lead to embryonic lethality, which hampers in-depth analysis of the phenotype. A few knock-in mice have been created with subtle mutations or small deletions that reproduce human mutations. The major reason for the paucity of knock-in mice is certainly that collagen genes are very large. Thus, they are difficult to manipulate. The introduction of a small deletion or a single point mutation in murine collagen genes still represents a considerable challenge. Nevertheless, the few examples of knock-in mouse lines tend to prove that mouse models can bring new information about *in vivo* consequences of collagen dysfunction that cannot be predicted by *in vitro* approaches. Knock-in mice are also indispensable models for assessing the effects of subtle mutations on tissue function, development, and aging. They are also valuable for developing specific gene therapy approaches to combat collagen-related disorders. The combination of site-directed mutagenesis in transfected cells and knock-in approaches in mice to address the impact of specific mutations will enable us to identify mechanisms

underlying the vast repertoire of collagen-related diseases. The implications may lead to the development of a specific therapy.

7. References

- Andrikopoulos, K., Liu, X., Keene, D.R., Jaenisch, R. & Ramirez, F. (1995). Targeted mutation in the *col5a2* gene reveals a regulatory role for type V collagen during matrix assembly. *Nat Genet.* 9(1):31-36.
- Aszodi, A., Legate, K.R., Nakchbandi, I. & Fässler, R. (2006). What Mouse Mutants Teach Us About Extracellular Matrix Function. *Annu. Rev. Cell Dev. Biol.* 22:591-621.
- Bateman, J.F., Boot-Handford, R.P. & Lamandé, S.R. (2009). Genetic diseases of connective tissues: cellular and extracellular effects of ECM mutations. *Nat Rev Genet.* 10(3):173-83.
- Bonaldo, P., Braghetta, P., Zanetti, M., Piccolo, S., Volpin, D. & Bressan, G.M. (1998). Collagen VI deficiency induces early onset myopathy in the mouse: an animal model for Bethlem myopathy. *Hum Mol Genet.* 7(13):2135-2140.
- Bonod-Bidaud, C., Beraud, M., Vaganay, E., Delacoux, F., Font, B., Hulmes, D.J. & Ruggiero, F. (2007). Enzymatic cleavage specificity of the pro α 1(V) chain processing analysed by site-directed mutagenesis. *Biochem J.* 405(2):299-306.
- Boot-Handford, R.P. & Briggs, M.D. (2010). The unfolded protein response and its relevance to connective tissue diseases. *Cell Tissue Res.* 339(1):197-211.
- Briñas, L., Richard, P., Quijano-Roy, S., Gartioux, C., Ledeuil, C., Lacène, E., Makri, S., Ferreira, A., Maugenre, S., Topaloglu, H., Haliloglu, G., Péniisson-Besnier, I., Jeannet, P.Y., Merlini, L., Navarro, C., Toutain, A., Chaigne, D., Desguerre, I., de Die-Smulders, C., Dunand, M., Echenne, B., Eymard, B., Kuntzer, T., Maincent, K., Mayer, M., Plessis, G., Rivier, F., Roelens, F., Stojkovic, T., Taratuto, A.L., Lubieniecki, F., Monges, S., Tranchant, C., Viollet, L., Romero, N.B., Estournet, B., Guicheney, P., Allamand, V. (2010). Early onset collagen VI myopathies: Genetic and clinical correlations. *Ann Neurol.* 68(4):511-520.
- Bruckner-Tuderman, L. & Bruckner, P. (1998). Genetic diseases of the extracellular matrix: more than just connective tissue disorders. *J Mol Med.* 76(3-4):226-237.
- Bruckner-Tuderman, L., Höpfner, B. & Hammami-Hauasli, N. (1999). Biology of anchoring fibrils: lessons from dystrophic epidermolysis bullosa. *Matrix Biol.* 18(1):43-54.
- Burgeson, R.E. (1993). Type VII collagen, anchoring fibrils, and epidermolysis bullosa. *J Invest Dermatol.* 101(3):252-255.
- Byers, P.H. & Steiner, R.D. (1992). Osteogenesis imperfecta. *Annu Rev Med.* 43:269-282.
- Chanut-Delalande, H., Bonod-Bidaud, C., Cogne, C., Malbouyres, M., Ramirez, F., Fichard, A. & Ruggiero, F. (2004). Development of a functional skin matrix requires deposition of collagen V heterotrimers. *Mol. Cell. Biol.*, 24(13):6049-6057
- Cheng, H., Rashid, S., Yu, Z., Yoshizumi, A., Hwang, E. & Brodsky, B. (2011). Location of glycine mutations within a bacterial collagen protein affects degree of disruption of triple-helix folding and conformation. *J Biol Chem.* 286(3):2041-2046.
- Chipman, S.D., Sweet, H.O., McBride, D.J. Jr, Davisson, M.T., Marks, S.C. Jr, Shuldiner, A.R., Wenstrup, R.J., Rowe, D.W. & Shapiro, J.R. (1993). Defective pro α 2(I) collagen synthesis in a recessive mutation in mice: a model of human osteogenesis imperfecta. *Proc Natl Acad Sci U S A.* 90(5):1701-1705.

- Colige, A., Ruggiero, F., Vandenberghe, I., Dubail, J., Kesteloot, F., Van Beeumen, J., Beschin, A., Brys, L., Lapière, C.M. & Nusgens B. (2005). Domains and maturation processes that regulate the activity of ADAMTS-2, a metalloproteinase cleaving the aminopropeptide of fibrillar procollagens types I-III and V. *J Biol Chem.* 280(41):34397-34408.
- Cosgrove, D., Meehan, D.T., Grunkemeyer, J.A., Kornak, J.M., Sayers, R., Hunter, W.J. & Samuelson, G.C. (1996). Collagen COL4A3 knockout: a mouse model for autosomal Alport syndrome. *Genes Dev.* 10(23):2981-92.
- Daley, E., Streeten, E.A., Sorkin, J.D., Kuznetsova, N., Shapses, S.A., Carleton, S.M., Shuldiner, A.R., Marini, J.C., Phillips, C.L., Goldstein, S.A., Leikin, S. & McBride D.J. Jr. (2010). Variable bone fragility associated with an Amish COL1A2 variant and a knock-in mouse model. *J Bone Miner Res.* 25(2):247-261.
- Dalgleish, R. (1998). The Human Collagen Mutation Database 1998. *Nucleic Acids Res.* 26(1):253-255.
- Dunnill, M.G., McGrath, J.A., Richards, A.J., Christiano, A.M., Uitto, J., Pope, F.M. & Eady, R.A. (1996). Clinicopathological correlations of compound heterozygous COL7A1 mutations in recessive dystrophic epidermolysis bullosa. *J Invest Dermatol.* 107(2):171-177.
- Eklund, L., Piuhola, J., Komulainen, J., Sormunen, R., Ongvarrasopone, C., Fässler, R., Muona, A., Ilves, M., Ruskoaho, H., Takala, T.E. & Pihlajaniemi, T. (2001). Lack of type XV collagen causes a skeletal myopathy and cardiovascular defects in mice. *Proc Natl Acad Sci U S A.* 98(3):1194-1199.
- Fässler, R., Schnegelsberg, P.N., Dausman, J., Shinya, T., Muragaki, Y., McCarthy, M.T., Olsen, B.R. & Jaenisch, R. (1994). Mice lacking alpha 1 (IX) collagen develop noninflammatory degenerative joint disease. *Proc Natl Acad Sci U S A.* 91(11):5070-5074.
- Fichard, A., Kleman, J.P. & Ruggiero F. (1995). Another look at collagen V and XI molecules. *Matrix Biol.* 14(7):515-531.
- Fichard, A., Tillet, E., Delacoux, F., Garrone, R. & Ruggiero, F. (1997). Human recombinant alpha1(V) collagen chain. Homotrimeric assembly and subsequent processing. *J Biol Chem.* 272(48):30083-7.
- Fichard, A., Chanut-Delalande, H. & Ruggiero, F. (2003). The Ehlers-Danlos syndrome: the extracellular matrix scaffold in question. *Med/Sci (Paris).* 19(4):443-452.
- Forlino, A., Porter, F.D., Lee, E.J., Westphal, H. & Marini, J.C. (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(53):37923-37931.
- Gara, S.K., Grumati, P., Squarzone, S., Sabatelli, P., Urciuolo, A., Bonaldo, P., Paulsson, M. & Wagener, R. (2011). Differential and restricted expression of novel collagen VI chains in mouse. *Matrix Biol.* 30(4):248-257.
- Hagg, R., Hedbom, E., Möllers, U., Aszódi, A., Fässler, R. & Bruckner, P. (1997). Absence of the alpha1(IX) chain leads to a functional knock-out of the entire collagen IX protein in mice. *J Biol Chem.* 272(33):20650-20654.
- Hammami-Hauasli, N., Schumann, H., Raghunath, M., Kilgus, O., Lüthi, U., Luger, T. & Bruckner-Tuderman, L. (1998). Some, but not all, glycine substitution mutations in

- COL7A1 result in intracellular accumulation of collagen VII, loss of anchoring fibrils, and skin blistering. *J Biol Chem.* 273(30):19228-19234.
- Heinonen, S., Männikkö, M., Klement, J.F., Whitaker-Menezes, D., Murphy, G.F. & Uitto, J. (1999). Targeted inactivation of the type VII collagen gene (Col7 α 1) in mice results in severe blistering phenotype: a model for recessive dystrophic epidermolysis bullosa. *J Cell Sci.* 112 (Pt 21):3641-3648.
- Humphries, J.D., Byron, A. & Humphries, M.J. (2006). Integrin ligands at a glance. *J Cell Sci.* 119(Pt 19):3901-3903.
- Huang, G., Ge, G., Wang, D., Gopalakrishnan, B., Butz, D.H., Colman, R.J., Nagy, A. & Greenspan, D.S. (2011). α 3(V) collagen is critical for glucose homeostasis in mice due to effects in pancreatic islets and peripheral tissues. *J Clin Invest.* 121(2):769-783.
- Hopfer, U., Fukai, N., Hopfer, H., Wolf, G., Joyce, N., Li, E. & Olsen, B.R. (2005). Targeted disruption of Col8a1 and Col8a2 genes in mice leads to anterior segment abnormalities in the eye. *FASEB J.* 19(10):1232-1244.
- Izu, Y., Sun, M., Zwolanek, D., Veit, G., Williams, V., Cha, B., Jepsen, K.J., Koch, M. & Birk D.E. (2011). Type XII collagen regulates osteoblast polarity and communication during bone formation. *J Cell Biol.* 193(6):1115-1130.
- Kessler, E., Fichard, A., Chanut-Delalande, H., Brusel, M. & Ruggiero, F. (2001). Bone morphogenetic protein-1 (BMP-1) mediates C-terminal processing of procollagen V homotrimer. *J Biol Chem.* 276(29):27051-27057.
- Kirschner, J., Hausser, I., Zou, Y., Schreiber, G., Christen, H.J., Brown, S.C., Anton-Lamprecht, I., Muntoni, F., Hanefeld, F. & Bönnemann, C.G. (2005). Ullrich congenital muscular dystrophy: connective tissue abnormalities in the skin support overlap with Ehlers-Danlos syndromes. *Am J Med Genet A.* 132A(3):296-301.
- Kobayashi, T., Kakihara, T. & Uchiyama, M. (2008). Mutational analysis of type IV collagen alpha5 chain, with respect to heterotrimer formation. *Biochem Biophys Res Commun.* 366(1):60-65.
- Kuivaniemi, H., Tromp, G. & Prockop, D.J. (1991). Mutations in collagen genes: causes of rare and some common diseases in humans. *FASEB J.* 5(7):2052-2060.
- Kwan, K.M., Pang, M.K., Zhou, S., Cowan, S.K., Kong, R.Y., Pfordte, T., Olsen, B.R., Silience, D.O., Tam, P.P. & Cheah, K.S. (1997). Abnormal compartmentalization of cartilage matrix components in mice lacking collagen X: implications for function. *J Cell Biol.* 136(2):459-471.
- Lamandé, S.R. & Bateman, J.F. (1995). The type I collagen pro alpha 1(I) COOH-terminal propeptide N-linked oligosaccharide. Functional analysis by site-directed mutagenesis. *J Biol Chem.* 270(30):17858-17865.
- Leitinger, B. & Hohenester, E. (2007). Mammalian collagen receptors. *Matrix Biol.* 26(3):146-155.
- Lightfoot, S.J., Atkinson, M.S., Murphy, G., Byers, P.H. & Kadler, K.E. (1994). Substitution of serine for glycine 883 in the triple helix of the pro alpha 1 (I) chain of type I procollagen produces osteogenesis imperfecta type IV and introduces a structural change in the triple helix that does not alter cleavage of the molecule by procollagen N-proteinase. *J Biol Chem.* 269(48):30352-30357.
- Lindor, N.M. & Bristow, J. (2005). Tenascin-X deficiency in autosomal recessive Ehlers-Danlos syndrome. *Am J Med Genet A.* 135(1):75-80.

- Linsenmayer, T.F., Gibney, E., Igoe, F., Gordon, M.K., Fitch, J.M., Fessler, L.I. & Birk, D.E. (1993). Type V collagen: molecular structure and fibrillar organization of the chicken alpha 1(V) NH2-terminal domain, a putative regulator of corneal fibrillogenesis. *J Cell Biol.* 121(5):1181-1189.
- Liu, X., Wu, H., Byrne, M., Krane, S. & Jaenisch, R. (1997). Type III collagen is crucial for collagen I fibrillogenesis and for normal cardiovascular development. *Proc Natl Acad Sci U S A.* 94(5):1852-1856.
- Löhler, J., Timpl, R. & Jaenisch, R. (1984). Embryonic lethal mutation in mouse collagen I gene causes rupture of blood vessels and is associated with erythropoietic and mesenchymal cell death. *Cell.* 38(2):597-607.
- Lu, W., Phillips, C.L., Killen, P.D., Hlaing, T., Harrison, W.R., Elder, F.F., Miner, J.H., Overbeek, P.A. & Meisler, M.H. (1999). Insertional mutation of the collagen genes Col4a3 and Col4a4 in a mouse model of Alport syndrome. *Genomics.* 61(2):113-124.
- Malfait, F., Wenstrup, R.J. & De Paepe, A. (2010). Clinical and genetic aspects of Ehlers-Danlos syndrome, classic type. *Genet Med.* 12(10):597-605.
- Miner, J.H., Sanes, J.R. (1996). Molecular and functional defects in kidneys of mice lacking collagen alpha 3(IV): implications for Alport syndrome. *J Cell Biol.* 135(5):1403-1413.
- Myllyharju, J. & Kivirikko, K.I. (2001). Collagens and collagen-related diseases. *Ann Med.* 33(1):7-21.
- Nuytinck, L., Freund, M., Lagae, L., Pierard, G.E., Hermanns-Le, T. & De Paepe, A. (2000). Classical Ehlers-Danlos syndrome caused by a mutation in type I collagen. *Am J Hum Genet.* 66(4):1398-1402.
- Parkin, J.D., San Antonio, J.D., Pedchenko, V., Hudson, B., Jensen, S.T. & Savige, J. (2011). Mapping structural landmarks, ligand binding sites, and missense mutations to the collagen IV heterotrimers predicts major functional domains, novel interactions, and variation in phenotypes in inherited diseases affecting basement membranes. *Hum Mutat.* 32(2):127-143.
- Raghunath, M., Bruckner, P & Steinmann, B. (1994). Delayed triple helix formation of mutant collagen from patients with osteogenesis imperfecta. *J Mol Biol.* 236(3):940-949.
- Rheault, M.N., Kren, S.M., Thielen, B.K., Mesa, H.A., Crosson, J.T., Thomas, W., Sado, Y., Kashtan, C.E. & Segal, Y. (2004). Mouse model of X-linked Alport syndrome. *J Am Soc Nephrol.* 15(6):1466-1474.
- Ricard-Blum, S. & Ruggiero, F. (2005). The collagen superfamily: from the extracellular matrix to the cell membrane. *Pathol Biol.* 53(7):430-442.
- Ruggiero, F. & Koch, M. (2008). Making recombinant extracellular matrix proteins. *Methods.* 45(1):75-85.
- Schnieke, A., Dziadek, M., Bateman, J., Mascara, T., Harbers, K., Gelinis, R. & Jaenisch, R. (1987). Introduction of the human pro alpha 1(I) collagen gene into pro alpha 1(I)-deficient Mov-13 mouse cells leads to formation of functional mouse-human hybrid type I collagen. *Proc Natl Acad Sci U S A.* 84(3):764-768.
- Seppinen, L., Sormunen, R., Soini, Y., Elamaa, H., Heljasvaara, R. & Pihlajaniemi, T. (2008). Lack of collagen XVIII accelerates cutaneous wound healing, while overexpression of its endostatin domain leads to delayed healing *Matrix Biol.* 27(6):535-546.
- Sumiyoshi, H., Mor, N., Lee, S.Y., 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(4):591-600.
- Symoens, S., Malfait, F., Renard, M., André, J., Hausser, I., Loeys, B., Coucke, P. & De Paepe, A. (2009). COL5A1 signal peptide mutations interfere with protein secretion and cause classic Ehlers-Danlos syndrome. *Hum Mutat.* 30(2):E395-403.
- Symoens, S., Renard, M., Bonod-Bidaud, C., Syx, D., Vaganay, E., Malfait, F., Ricard-Blum, S., Kessler, E., Van Laer, L., Coucke, P., Ruggiero, F. & De Paepe, A. (2011). Identification of binding partners interacting with the $\alpha 1$ -N-propeptide of type V collagen. *Biochem J.* 433(2):371-381.
- Takahara, K., Schwarze, U., Imamura, Y., Hoffman, G.G., Toriello, H., Smith, L.T., Byers, P.H. & Greenspan, D.S. (2002). Order of intron removal influences multiple splice outcomes, including a two-exon skip, in a COL5A1 acceptor-site mutation that results in abnormal pro- $\alpha 1(V)$ N-propeptides and Ehlers-Danlos syndrome type I. *Am J Hum Genet.* 71(3):451-465.
- Tanimura, S., Tadokoro, Y., Inomata, K., Binh, N.T., Nishie, W., Yamazaki, S., Nakauchi, H., Tanaka, Y., McMillan, J.R., Sawamura, D., Yancey, K., Shimizu, H. & Nishimura, E.K. (2011). Hair follicle stem cells provide a functional niche for melanocyte stem cells. *Cell Stem Cell.* 8(2):177-187.
- Tsang, K.Y., Chan, D., Bateman, J.F. & Cheah, K.S. (2010). In vivo cellular adaptation to ER stress: survival strategies with double-edged consequences. *J Cell Sci.* 123(Pt 13):2145-54.
- Wang, Y.F., Ding, J., Wang, F. & Bu, D.F. (2004). Effect of glycine substitutions on $\alpha 5(IV)$ chain structure and structure-phenotype correlations in Alport syndrome. *Biochem Biophys Res Commun.* 316(4):1143-1149.
- Wenstrup, R.J., Florer, J.B., Brunskill, E.W., Bell, S.M., Chervoneva, I. & Birk, D.E. (2004). Type V collagen controls the initiation of collagen fibril assembly. *J Biol Chem.* 279(51):53331-53337.
- Wenstrup, R.J., Smith, S.M., Florer, J.B., Zhang, G., Beason, D.P., Seegmiller, R.E., Soslowsky, L.J. & Birk, D.E. (2011). Regulation of collagen fibril nucleation and initial fibril assembly involves coordinate interactions with collagens V and XI in developing tendon. *J Biol Chem.* 286(23):20455-204565.
- Woodley, D.T., Hou, Y., Martin, S., Li, W. & Chen, M. (2008). Characterization of molecular mechanisms underlying mutations in dystrophic epidermolysis bullosa using site-directed mutagenesis. *J Biol Chem.* 283(26):17838-17845.
- Xian, X., Gopa, S. & Couchman, J.R. (2010). Syndecans as receptors and organizers of the extracellular matrix. *Cell Tissue Res.* 339(1):31-46.
- Xu, K., Nowak, I., Kirchner, M. & Xu, Y. (2008). Recombinant collagen studies link the severe conformational changes induced by osteogenesis imperfecta mutations to the disruption of a set of interchain salt bridges. *J Biol Chem.* 283(49):34337-34344.
- Yang, W., Battineni, M.L. & Brodsky B. (1997). Amino acid sequence environment modulates the disruption by osteogenesis imperfecta glycine substitutions in collagen-like peptides. *Biochemistry.* 36(23):6930-6935.
- Zhu, Y., Oganessian, A., Keene, D.R. & Sandell, L.J. (1999). Type IIA procollagen containing the cysteine-rich amino propeptide is deposited in the extracellular matrix of prechondrogenic tissue and binds to TGF- $\beta 1$ and BMP-2. *J Cell Biol.* 144(5):1069-1080.