



Published in final edited form as:

*Cell Tissue Res.* 2010 January ; 339(1): 247–257. doi:10.1007/s00441-009-0844-4.

## Collagens

**Marion K. Gordon** and

Department of Pharmacology and Toxicology, Ernest Mario School of Pharmacy, Rutgers University, Piscataway, N.J., USA

**Rita A. Hahn**

Department of Pharmacology and Toxicology, Ernest Mario School of Pharmacy, Rutgers University, Piscataway, N.J., USA

Marion K. Gordon: magordon@eohsi.rutgers.edu

## Abstract

The collagens represent a family of trimeric extracellular matrix molecules used by cells for structural integrity and other functions. The three  $\alpha$  chains that form the triple helical part of the molecule are composed of repeating peptide triplets of glycine-X-Y. X and Y can be any amino acid but are often proline and hydroxyproline, respectively. Flanking the triple helical regions (i.e., Col domains) are non-glycine-X-Y regions, termed non-collagenous domains. These frequently contain recognizable peptide modules found in other matrix molecules. Proper tissue function depends on correctly assembled molecular aggregates being incorporated into the matrix. This review highlights some of the structural characteristics of collagen types I-XXVIII.

## Keywords

Collagens; Extracellular matrix; Fibrils; FACITs; Basement membrane

## Introduction

Collagens are extracellular matrix molecules used by cells for structural integrity and a variety of other functions. Recent reviews offer insights into the collagen family members and important aspects of their structures, functions, or associated disease states (Kadler et al. 2007; Myllyharju and Kivirikko 2001, 2004; Brodsky and Persikov 2005; Gelse et al. 2003; Ortega and Werb 2002). Table 1 lists the  $\alpha$  chains of the various collagens and the National Center for Biotechnology Information reference numbers that provide sequence information and cite some published literature for the 28 different collagen types. These  $\alpha$  chains are used to build trimeric molecules, which are woven together into a triple helix in at least one region. The triple helical domains form as a result of glycine (gly) being used every third residue (i.e., repeating peptide triplets of gly-X-Y). In this triplet, X is often proline, and Y is frequently hydroxyproline. In each  $\alpha$  chain, triple helical regions, termed Col domains, are flanked by non-collagenous (non-gly-X-Y) regions, termed NC domains. These NC domains often contain recognizable peptide modules found in other matrix molecules. The function of a collagen depends on the proper supramolecular assembly of molecules into an aggregate that becomes incorporated into the matrix. This review attempts to highlight some of the structural characteristics of the numbered collagens, types I-XXVIII. Space precludes us from including the growing number of triple-helix-containing molecules that are not among

the numbered collagens, many of which (e.g., the collectin family) are involved in innate immunity. Schematic diagrams of the aggregated forms of the numbered collagens are provided where known.

## Fibrillar collagens

The ways that  $\alpha$  chains form trimeric molecules, which in turn assemble into fibrils, are shown in Fig. 1. Some types, such as collagen II, form homotrimeric molecules. Others, such as types I and V, form heterotrimers. With the exception of collagens XXIV and XXVII, the major triple helical domain is slightly longer than 1000 amino acid residues and has a perfect gly-X-Y triplet structure. All non-fibrillar collagens have at least one imperfection or interruption in their triple helices. The amino end of a fibrillar collagen, called the N-peptide or N-propeptide, usually contains at least one small triple helical domain, designated the minor helix. Once the major triple helix has formed, the amino and carboxyl ends undergo processing. Processed molecules are aligned in a quarter stagger arrangement in the growing fibril. In electron micrographs, fibrils have an alternating light and dark pattern, which has led to them being called banded fibrils. Pure single-type fibrils are unlikely to exist. Types V and XI collagen nucleate fibrils of types I and II collagen, respectively (Kadler et al. 2008; Wenstrup et al. 2006). The portion of the V and XI N-peptides that are retained after processing serve to regulate fibril diameter (Birk 2001; Blaschke et al. 2000; Linsenmayer et al. 1993; Birk et al. 1990). Fibril diameter control is crucial to proper tissue function. In the fibrillar collagen group, collagens XXIV and XXVII are unique because their major triple helices are shorter than those of the other members, and they have one or two interruptions. The number depends on interpretation: one can view the amino-most interruption (yellow triangle in Fig. 1) as the separation between the minor and major helices, or it can be viewed as an additional interruption in the major triple helix of a collagen that contains no minor helix (Koch et al. 2003; Pace et al. 2003; Boot-Handford et al. 2003). Interestingly, the C-propeptides of both collagens XXIV and XXVII contain chain selection sequences that resemble those used in invertebrate fibrillar collagens (Lees et al. 1997). This and the analysis of the genes suggest that these two newly identified fibrillar collagens are of ancient origin (Boot-Handford et al. 2003).

## Collagens associated with banded fibrils

### FACITs

Several non-fibrillar types of collagen associate with the surface of collagen fibrils. One group has been named "FACITs" (fibril-associated collagens with interrupted triple helices). The prototype is collagen IX. This collagen is a product of cartilage, where it is cross-linked to the surface of type II collagen fibrils (Eyre et al. 1987). Types XII and XIV collagen are present in tissues containing either type I or type II collagen fibrils (Walchli et al. 1994; Eyre 2002). Proof of their association with banded fibrils has been demonstrated in tendon (Keene et al. 1991). Fibril association is schematicized in Fig. 2 for types IX, XII, and XIV. These three collagens can be alternatively spliced to yield short and long variants (Svoboda et al. 1988; Gerecke et al. 1997; Imhof and Trueb 2001) that might confer different properties to the fibrils in tissues. Type XIV collagen is also a fibril diameter regulator in early stages of fibrillogenesis (Ansorge et al. 2009). The way that diameter regulation may occur is shown in Fig. 2. Type XX collagen has been cloned from chicken (Koch et al. 2001). The human gene is disrupted in the region encoding the Col 1 domain, so the molecule is probably not used as a collagen in humans. Many lines of evidence suggest that, in addition to a fibril-associated form, type XII collagen has an additional aggregation that involves interaction with basement membrane components (Wessel et al. 1997; Ljubimov et al. 1996; Anderson et al. 2000; Cheng et al. 2001; Bader et al. 2009). This underscores an

important point: many collagens can participate in more than one kind of supramolecular aggregate.

The FACIT-like collagens, viz., types XVI (Pan et al. 1992), XIX (Yoshioka et al. 1992; Myers et al. 1994, 1997), XXI (Chou and Li 2002; Li et al. 2005), and XXII (Koch et al. 2004), are primarily localized to basement membrane zones or junctions between tissues. Type XVI collagen has 10 collagenous domains flanked by non-collagenous domains (Pan et al. 1992; Yamaguchi et al. 1992), and one of its supramolecular aggregated forms involves association with dermal fibrillin 1 near the epidermal basement membrane. Another form, found in cartilage, is associated with banded fibrils, but only fibrils that do not have collagen IX on their surface (Kassner et al. 2003). The aggregate form of collagen XIX is unknown, but noteworthy is the finding that Col19a1 null mice undergo a transdifferentiation of smooth muscle to skeletal muscle in the abdominal part of the esophagus (Sumiyoshi et al. 2004). Little is known about collagen type XXI other than its primary structure and potential chain association from recombinant  $\alpha$  chains (Chou and Li 2002; Li et al. 2005). Type XXII collagen, a close relative in sequence to domains of collagen XXI, is found in the basement membrane of the myotendinous junction. Another form of it is associated with cartilage microfibrils (Koch et al. 2004).

Much has been written about collagen VII, since it is crucial to the attachment complex that rivets the epidermis to the dermis (for reviews, see Bruckner-Tuderman 2009; Uitto 2008; Aumailley et al. 2006; Uitto and Pulkkinen 1996). The attachment complex is comprised of hemidesmosomes in the basal epithelial cells, anchoring filaments in the basement membrane, and anchoring fibrils that reach from the basement membrane down into the dermis. Type VII collagen is the major component of anchoring fibrils and has recently been shown to merge into the banded collagen fibrils of the dermis (Villone et al. 2008). A representation of the antiparallel dimer form of collagen VII is shown in Fig. 2, where one end of the aggregate is localized to the lamina densa (bound to anchoring filaments) of the epithelial basement membrane, and the other end interacts with a dermal collagen fibril.

## Network-forming collagens

Collagen types IV, VI, VIII, and X form networks of various kinds. Type IV collagen, otherwise known as basement membrane collagen, makes a three-dimensional structure (Yurchenco and Ruben 1988; Barge et al. 1991) from laterally associating chicken wire-like two-dimensional structures (Fig. 3). There are six  $\alpha$  chains that can be used to make the type IV collagen trimers (for reviews on structure and disease states, see Hudson et al. 1993; Khoshnoodi et al. 2008). The most common chain composition, however, is  $[\alpha 1(\text{IV})]_2\alpha 2(\text{IV})$ . Only two representative  $\alpha$  chains are shown in Fig. 3, one to represent the  $\alpha 1$  and  $\alpha 2$  chains, and the other to represent the  $\alpha 3$ ,  $\alpha 4$ ,  $\alpha 5$ , and  $\alpha 6$  chains. The N-terminal domains of the  $\alpha 1$  and  $\alpha 2$  chains are composed of 143 and 167 amino acid residues, respectively, whereas the other four chains have small amino terminal domains ranging from 13 to 29 amino acids. The triple helical domains of the six chains range from 1271 amino acids to 1416 residues, and all have more than 20 interruptions in the gly-X-Y triplet structure. The carboxyl termini all have approximately 228 amino acid residues.

Type VI collagen is another network-forming molecule. It has a ubiquitous tissue distribution. Once thought to make heterotypic trimers by using the smaller-sized  $\alpha 1$  and  $\alpha 2$  chains plus the larger  $\alpha 3$  chain, it is now known that three additional chains,  $\alpha 4$ – $\alpha 6$ , exist in mouse. All three are similar to the  $\alpha 3$  chain. Orthologs of  $\alpha 5$  and  $\alpha 6$  have been found in human (Fitzgerald et al. 2008; Gara et al. 2008). In mouse, the  $\alpha 4$  chain is used to make trimers far more frequently than the  $\alpha 5$  and  $\alpha 6$  chains, so why do humans lack the  $\alpha 4$  chain? Spatial expression patterns indicate that the tissue distribution of mouse  $\alpha 4$  is similar to that

of human  $\alpha 5$ , suggesting that human  $\alpha 5$  is its equivalent. The lack of the  $\alpha 4$  chain in humans is attributable to a disrupted COL4A4 gene. This disruption is probably a relatively recent gene alteration, since rhesus monkeys (Old World monkey lineage) have an intact  $\alpha 4$  (VI) chain gene, whereas chimpanzees and humans do not (Fitzgerald et al. 2008).

In the type VI  $\alpha$  chains, the short ~330-amino-acid collagenous domain is flanked on each side by von Willebrand factor A domains (see Fig. 3). To make the type VI collagen supramolecular aggregate, molecules form dimers, the dimers form tetramers, and the tetramers form end to end associations that yield the ultrastructural appearance of beads on a string (Engel et al. 1985). The structure is represented in Fig. 3. Type VI collagen has a crucial role in the function in muscle, since mutations cause Bethlem myopathy and Ulrich congenital muscular dystrophy (for a review, see Lampe and Bushby 2005).

Types VIII and X are related short chain collagens (for reviews, see Schmid et al. 1990; Shuttleworth 1997; Plenz et al. 2003). Type VIII is expressed by many cells, especially by endothelial cells, whereas type X is restricted to hypertrophic chondrocytes. The type VIII  $\alpha 1$  and  $\alpha 2$  chains are slightly longer than the  $\alpha 1(X)$  collagen chain because of an additional exon in the type VIII genes encoding an extra polypeptide sequence for the amino terminal domain. This is shown as a yellow box in Fig. 3. One of the supramolecular aggregates that type VIII collagen can make is a hexagonal lattice (Fig. 3). This structure is beautifully demonstrated by Descemet's membrane of the cornea (Sawada et al. 1990). Recombinant type VIII collagen molecules allowed to aggregate also make hexagonal lattices (Stephan et al. 2004). Additional macro-molecular forms must exist, since type VIII collagen made by vascular endothelial cells has never been visualized in a hexagonal lattice. With regard to type X collagen, no in vivo lattice that reacts with a collagen X antibody has been observed. However, type X collagen molecules form hexagonal lattices when allowed to assemble into aggregates in vitro (Kwan et al. 1991). What one does observe in vivo is that type X collagen has a fibril-associated form, and that it also forms fine filaments in the pericellular matrix of hypertrophic chondrocytes (Schmid and Linsenmayer 1990). The latter may be hexagonal lattices that have collapsed during preparation for ultrastructural analysis. In individual molecules, type X collagen, having only one  $\alpha$  chain, forms homotrimers. Although two  $\alpha$  chains are known for type VIII collagen, the preponderance of evidence suggests each chain forms homotrimers (Illidge et al. 1998; Greenhill et al. 2000; Stephan et al. 2004).

## Transmembranous collagens

Types XIII, XVII, XXIII, and XXV collagens are transmembrane molecules inserted in the plasma membrane in a type II orientation (for a review, see Franzke et al. 2003). All members of the group are also shed from the cell surface, generating soluble forms (for reviews, see Franzke et al. 2003; Veit et al. 2007). Type XVII collagen is unlike the other three members of the group, being much larger and having many more collagenous domains plus a large intracellular domain (Fig. 4). As a component of the hemidesmosomes, collagen XVII, together with  $\alpha 6\beta 4$  integrin, extends from the cell membrane into the basement membrane to bind with laminin 332, the anchoring filament component of the attachment complex. As such, collagen XVII is indirectly assembled into a supramolecular complex with collagen VII, which binds to laminin 332 from the dermal side.

Collagen types XIII, XXIII, and XXV are similar to each other. Types XIII and XXV have a conserved linear pattern of domains. Collagen XXIII has conserved domains, but they are rearranged compared with the other two (Koch et al. 2006). This is indicated by the geometric patterns superimposed on the domain structures in Fig. 4. The initiation of chain selection, followed by zippering of the triple helix (i.e., folding the molecule properly), is

not so straight forward for a transmembrane collagen. Coiled coils have been postulated to play a role in molecular assembly (McAlinden et al. 2003). For collagen XIII, this has been established by examining NC3 domain mutations (Snellman et al. 2007). Our understanding of collagens XIII, XXIII, and XXV is still in its infancy. Collagen XXV is known to be a component of the amyloid plaques characteristic of Alzheimer's disease (Hashimoto et al. 2002). The molecule appears to assemble amyloid beta protein fibrils into bundles that are resistant to proteases (Soderberg et al. 2005). With regard to collagens XIII and XXIII, the level of these molecules is elevated in some tumors (Vaisanen et al. 2005; Banyard et al. 2003; Koch et al. 2006). A most interesting finding is that a mutation in collagen XIII has been associated with altering an animal's immune response, thereby increasing its susceptibility to bacteria. The presence of the mutant collagen XIII has been correlated with a predisposition toward developing B cell lymphomas (Tuomisto et al. 2008).

## Endostatin precursor collagens

The carboxyl terminal domains of collagens XV and XVIII can be cleaved to generate antiangiogenic peptides. (This also occurs for type IV collagen.) These are called either endostatin and restin, or endostatin-XVIII and endostatin-XV. These cleaved fragments have some distinct differences in properties (Sasaki et al. 2000). As full-length molecules, collagen XV and XVIII are basement membrane collagens with similar features. There is high homology between triple helical domains and the amino and carboxyl non-collagenous domains (Rehn et al. 1994). Moreover, both molecules are proteoglycan core proteins. Collagen XV has a chondroitin sulfate glycosaminoglycan side chain (Li et al. 2000), and collagen XVIII has a heparin sulfate side chain (Halfter et al. 1998). However, some important differences exist in the tissue distribution of these collagens. The most striking is that collagen XV is expressed in heart and skeletal muscle, whereas collagen XVIII is expressed by smooth muscle (Sasaki et al. 2000). Because of this pattern, collagen XV null mice unsurprisingly have skeletal myopathies and cardiovascular defects (Eklund et al. 2001). Mice null for both genes demonstrate that the collagens do not functionally compensate for each other (Ylikarppa et al. 2003). Another interesting feature about collagen XVIII is that it is like fibronectin in the sense that it, too, has a plasma form. The long form of the molecule synthesized by the liver circulates in the blood (Musso et al. 2001).

## Other collagens

Collagens XXVI and XXVIII do not easily fit into any category. Collagen XXVI is expressed in testis and ovaries, especially in neonates. The molecule is small for a collagen, having only 438 amino acids in total, including two short collagenous domains (69 aa and 33 aa). It does, however, undergo modification processes expected for a collagen, such as prolyl hydroxylation, trimer formation, and secretion (Sato et al. 2002).

The collagen XXVIII triple helix is flanked by von Willebrand factor A domains, and the molecule has some sequence relationship with type VI chains based on phylogenetic analyses, but the triple helical domain is longer than that of collagen VI chains. The molecule is expressed predominantly by dorsal root ganglia and peripheral nerves. Newborn, but not adult, sciatic nerve expresses collagen XXVIII mRNA, although the protein is detected in adult sciatic nerve, suggesting a long half life for the molecule.

## Concluding remarks

There are many collagens, generating many questions. Because the family is so diverse, we can look forward with pleasure to the exciting answers that will emerge.

## Acknowledgments

The work of the authors is supported by NIH EY009056 and U54AR055073.

## References

- Anderson S, SundarRaj S, Fite D, Wessel H, SundarRaj N. Developmentally regulated appearance of spliced variants of type XII collagen in the cornea. *Invest Ophthalmol Vis Sci* 2000;41:55–63. [PubMed: 10634601]
- Ansorge HL, Meng X, Zhang G, Veit G, Sun M, Klement JF, Beason DP, Soslowsky LJ, Koch M, Birk DE. Type XIV collagen regulates fibrillogenesis: premature collagen fibril growth and tissue dysfunction in null mice. *J Biol Chem* 2009;284:8427–8438. [PubMed: 19136672]
- Aumailley M, Has C, Tunggal L, Bruckner-Tuderman L. Molecular basis of inherited skin-blistering disorders, and therapeutic implications. *Expert Rev Mol Med* 2006;8:1–21. [PubMed: 17040578]
- Bader HL, Keene DR, Charvet B, Veit G, Driever W, Koch M, Ruggiero F. Zebrafish collagen XII is present in embryonic connective tissue sheaths (fascia) and basement membranes. *Matrix Biol* 2009;28:32–43. [PubMed: 18983916]
- Banyard J, Bao L, Zetter BR. Type XXIII collagen, a new transmembrane collagen identified in metastatic tumor cells. *J Biol Chem* 2003;278:20989–20994. [PubMed: 12644459]
- Barge A, Ruggiero F, Garrone R. Structure of the basement membrane of corneal epithelium: quick-freeze, deep-etch comparative study of networks deposited in culture and during development. *Biol Cell* 1991;72:141–147. [PubMed: 1756303]
- Birk DE. Type V collagen: heterotypic type I/V collagen interactions in the regulation of fibril assembly. *Micron* 2001;32:223–237. [PubMed: 11006503]
- Birk DE, Fitch JM, Babiarz JP, Doane KJ, Linsenmayer TF. Collagen fibrillogenesis in vitro: interaction of types I and V collagen regulates fibril diameter. *J Cell Sci* 1990;95:649–657. [PubMed: 2384532]
- Blaschke UK, Eikenberry EF, Hulmes DJ, Galla HJ, Bruckner P. Collagen XI nucleates self-assembly and limits lateral growth of cartilage fibrils. *J Biol Chem* 2000;275:10370–10378. [PubMed: 10744725]
- Boot-Handford RP, Tuckwell DS, Plumb DA, Rock CF, Poulson R. A novel and highly conserved collagen (pro(α)1 (XXVII)) with a unique expression pattern and unusual molecular characteristics establishes a new clade within the vertebrate fibrillar collagen family. *J Biol Chem* 2003;278:31067–31077. [PubMed: 12766169]
- Brodsky B, Persikov AV. Molecular structure of the collagen triple helix. *Adv Protein Chem* 2005;70:301–339. [PubMed: 15837519]
- Bruckner-Tuderman L. Can type VII collagen injections cure dystrophic epidermolysis bullosa? *Mol Ther* 2009;17:6–7. [PubMed: 19116634]
- Cheng EL, Maruyama I, SundarRaj N, Sugar J, Feder RS, Yue BY. Expression of type XII collagen and hemidesmosome-associated proteins in keratoconus corneas. *Curr Eye Res* 2001;22:333–340. [PubMed: 11600933]
- Chou MY, Li HC. Genomic organization and characterization of the human type XXI collagen (COL21A1) gene. *Genomics* 2002;79:395–401. [PubMed: 11863369]
- Eklund L, Pihola J, Komulainen J, Sormunen R, Ongvarrasopone C, Fassler R, Muona A, Ilves M, Ruskoaho H, Takala TE, Pihlajaniemi T. Lack of type XV collagen causes a skeletal myopathy and cardiovascular defects in mice. *Proc Natl Acad Sci USA* 2001;98:1194–1199. [PubMed: 11158616]
- Engel J, Furthmayr H, Odermatt E, von der Mark H, Aumailley M, Fleischmajer R, Timpl R. Structure and macromolecular organization of type VI collagen. *Ann N Y Acad Sci* 1985;460:25–37. [PubMed: 3938630]
- Eyre D. Collagen of articular cartilage. *Arthritis Res* 2002;4:30–35. [PubMed: 11879535]
- Eyre DR, Apon S, Wu JJ, Ericsson LH, Walsh KA. Collagen type IX: evidence for covalent linkages to type II collagen in cartilage. *FEBS Lett* 1987;220:337–341. [PubMed: 3609327]

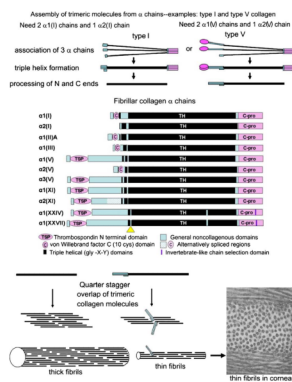
- Fitzgerald J, Rich C, Zhou FH, Hansen U. Three novel collagen VI chains, alpha4(VI), alpha5(VI), and alpha6(VI). *J Biol Chem* 2008;283:20170–20180. [PubMed: 18400749]
- Franzke CW, Tasanen K, Schumann H, Bruckner-Tuderman L. Collagenous transmembrane proteins: collagen XVII as a prototype. *Matrix Biol* 2003;22:299–309. [PubMed: 12935815]
- Gara SK, Grumati P, Urciuolo A, Bonaldo P, Kobbe B, Koch M, Paulsson M, Wagener R. Three novel collagen VI chains with high homology to the alpha3 chain. *J Biol Chem* 2008;283:10658–10670. [PubMed: 18276594]
- Gelse K, Poschl E, Aigner T. Collagens—structure, function, and biosynthesis. *Adv Drug Deliv Rev* 2003;55:1531–1546. [PubMed: 14623400]
- Gerecke DR, Olson PF, Koch M, Knoll JH, Taylor R, Hudson DL, Champliand MF, Olsen BR, Burgeson RE. Complete primary structure of two splice variants of collagen XII, and assignment of alpha 1(XII) collagen (COL12A1), alpha 1(IX) collagen (COL9A1), and alpha 1(XIX) collagen (COL19A1) to human chromosome 6q12-q13. *Genomics* 1997;41:236–242. [PubMed: 9143499]
- Greenhill NS, Ruger BM, Hasan Q, Davis PF. The alpha1(VIII) and alpha2(VIII) collagen chains form two distinct homotrimeric proteins in vivo. *Matrix Biol* 2000;19:19–28. [PubMed: 10686422]
- Halfter W, Dong S, Schurer B, Cole GJ. Collagen XVIII is a basement membrane heparan sulfate proteoglycan. *J Biol Chem* 1998;273:25404–25412. [PubMed: 9738008]
- Hashimoto T, Wakabayashi T, Watanabe A, Kowa H, Hosoda R, Nakamura A, Kanazawa I, Arai T, Takio K, Mann DM, Iwatsubo T. CLAC: a novel Alzheimer amyloid plaque component derived from a transmembrane precursor, CLAC-P/collagen type XXV. *EMBO J* 2002;21:1524–1534. [PubMed: 11927537]
- Hudson BG, Reeders ST, Tryggvason K. Type IV collagen: structure, gene organization, and role in human diseases. Molecular basis of Goodpasture and Alport syndromes and diffuse leiomyomatosis. *J Biol Chem* 1993;268:26033–26036. [PubMed: 8253711]
- Illidge C, Kielty C, Shuttleworth A. The alpha1(VIII) and alpha2(VIII) chains of type VIII collagen can form stable homotrimeric molecules. *J Biol Chem* 1998;273:22091–22095. [PubMed: 9705353]
- Imhof M, Trueb B. Alternative splicing of the first F3 domain from chicken collagen XIV affects cell adhesion and heparin binding. *J Biol Chem* 2001;276:9141–9148. [PubMed: 11098058]
- Kadler KE, Baldock C, Bella J, Boot-Handford RP. Collagens at a glance. *J Cell Sci* 2007;120:1955–1958. [PubMed: 17550969]
- Kadler KE, Hill A, Canty-Laird EG. Collagen fibrillogenesis: fibronectin, integrins, and minor collagens as organizers and nucleators. *Curr Opin Cell Biol* 2008;20:495–501. [PubMed: 18640274]
- Kassner A, Hansen U, Miosge N, Reinhardt DP, Aigner T, Bruckner-Tuderman L, Bruckner P, Grassel S. Discrete integration of collagen XVI into tissue-specific collagen fibrils or beaded microfibrils. *Matrix Biol* 2003;22:131–143. [PubMed: 12782140]
- Keene DR, Lunstrum GP, Morris NP, Stoddard DW, Burgeson RE. Two type XII-like collagens localize to the surface of banded collagen fibrils. *J Cell Biol* 1991;113:971–978. [PubMed: 2026656]
- Khoshnoodi J, Pedchenko V, Hudson BG. Mammalian collagen IV. *Microsc Res Tech* 2008;71:357–70. [PubMed: 18219669]
- Koch M, Foley JE, Hahn R, Zhou P, Burgeson RE, Gerecke DR, Gordon MK. Alpha 1(XX) collagen, a new member of the collagen subfamily, fibril-associated collagens with interrupted triple helices. *J Biol Chem* 2001;276:23120–23126. [PubMed: 11274142]
- Koch M, Laub F, Zhou P, Hahn RA, Tanaka S, Burgeson RE, Gerecke DR, Ramirez F, Gordon MK. Collagen XXIV, a vertebrate fibrillar collagen with structural features of invertebrate collagens: selective expression in developing cornea and bone. *J Biol Chem* 2003;278:43236–43244. [PubMed: 12874293]
- Koch M, Schulze J, Hansen U, Ashwodt T, Keene DR, Brunken WJ, Burgeson RE, Bruckner P, Bruckner-Tuderman L. A novel marker of tissue junctions, collagen XXII. *J Biol Chem* 2004;279:22514–22521. [PubMed: 15016833]

- Koch M, Veit G, Stricker S, Bhatt P, Kutsch S, Zhou P, Reinders E, Hahn RA, Song R, Burgeson RE, Gerecke DR, Mundlos S, Gordon MK. Expression of type XXIII collagen mRNA and protein. *J Biol Chem* 2006;281:21546–21557. [PubMed: 16728390]
- Kwan AP, Cummings CE, Chapman JA, Grant ME. Macromolecular organization of chicken type X collagen in vitro. *J Cell Biol* 1991;114:597–604. [PubMed: 1860888]
- Lampe AK, Bushby KM. Collagen VI related muscle disorders. *J Med Genet* 2005;42:673–685. [PubMed: 16141002]
- Lees JF, Tasab M, Bulleid NJ. Identification of the molecular recognition sequence which determines the type-specific assembly of procollagen. *EMBO J* 1997;16:908–916. [PubMed: 9118952]
- Li D, Clark CC, Myers JC. Basement membrane zone type XV collagen is a disulfide-bonded chondroitin sulfate proteoglycan in human tissues and cultured cells. *J Biol Chem* 2000;275:22339–22347. [PubMed: 10791950]
- Li HC, Huang CC, Chen SF, Chou MY. Assembly of homotrimeric type XXI minicollagen by coexpression of prolyl 4-hydroxylase in stably transfected *Drosophila melanogaster* S2 cells. *Biochem Biophys Res Commun* 2005;336:375–385. [PubMed: 16115607]
- Linsenmayer TF, Gibney E, Igoe F, Gordon MK, Fitch JM, Fessler LI, Birk DE. 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* 1993;121:1181–1189. [PubMed: 8501123]
- Ljubimov AV, Burgeson RE, Butkowski RJ, Couchman JR, Zardi L, Ninomiya Y, Sado Y, Huang ZS, Nesburn AB, Kenney MC. Basement membrane abnormalities in human eyes with diabetic retinopathy. *J Histochem Cytochem* 1996;44:1469–1479. [PubMed: 8985139]
- McAlinden A, Smith TA, Sandell LJ, Ficheux D, Parry DA, Hulmes DJ. Alpha-helical coiled-coil oligomerization domains are almost ubiquitous in the collagen superfamily. *J Biol Chem* 2003;278:42200–42207. [PubMed: 12920133]
- Musso O, Theret N, Heljasvaara R, Rehn M, Turlin B, Champion JP, Pihlajaniemi T, Clement B. Tumor hepatocytes and basement membrane-producing cells specifically express two different forms of the endostatin precursor, collagen XVIII, in human liver cancers. *Hepatology* 2001;33:868–876. [PubMed: 11283851]
- Myers JC, Yang H, D'Ippolito JA, Presente A, Miller MK, Dion AS. The triple-helical region of human type XIX collagen consists of multiple collagenous subdomains and exhibits limited sequence homology to alpha 1(XVI). *J Biol Chem* 1994;269:18549–18557. [PubMed: 8034603]
- Myers JC, Li D, Bageris A, Abraham V, Dion AS, Amenta PS. Biochemical and immunohistochemical characterization of human type XIX defines a novel class of basement membrane zone collagens. *Am J Pathol* 1997;151:1729–1740. [PubMed: 9403723]
- Myllyharju J, Kivirikko KI. Collagens and collagen-related diseases. *Ann Med* 2001;33:7–21. [PubMed: 11310942]
- Myllyharju J, Kivirikko KI. Collagens, modifying enzymes and their mutations in humans, flies and worms. *Trends Genet* 2004;20:33–43. [PubMed: 14698617]
- Ortega N, Werb Z. New functional roles for non-collagenous domains of basement membrane collagens. *J Cell Sci* 2002;115:4201–4214. [PubMed: 12376553]
- Pace JM, Corrado M, Missero C, Byers PH. Identification, characterization and expression analysis of a new fibrillar collagen gene, COL27A1. *Matrix Biol* 2003;22:3–14. [PubMed: 12714037]
- Pan TC, Zhang RZ, Mattei MG, Timpl R, Chu ML. Cloning and chromosomal location of human alpha 1(XVI) collagen. *Proc Natl Acad Sci USA* 1992;89:6565–6569. [PubMed: 1631157]
- Plenz GA, Deng MC, Robenek H, Volker W. Vascular collagens: spotlight on the role of type VIII collagen in atherogenesis. *Atherosclerosis* 2003;166:1–11. [PubMed: 12482545]
- Rehn M, Hintikka E, Pihlajaniemi T. Primary structure of the alpha 1 chain of mouse type XVIII collagen, partial structure of the corresponding gene, and comparison of the alpha 1(XVIII) chain with its homologue, the alpha 1(XV) collagen chain. *J Biol Chem* 1994;269:13929–13935. [PubMed: 8188673]
- Sasaki T, Larsson H, Tisi D, Claesson-Welsh L, Hohenester E, Timpl R. Endostatins derived from collagens XV and XVIII differ in structural and binding properties, tissue distribution and anti-angiogenic activity. *J Mol Biol* 2000;301:1179–1190. [PubMed: 10966814]

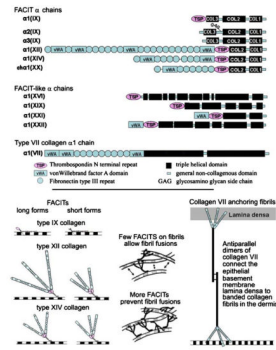


- Sato K, Yomogida K, Wada T, Yorihazi T, Nishimune Y, Hosokawa N, Nagata K. Type XXVI collagen, a new member of the collagen family, is specifically expressed in the testis and ovary. *J Biol Chem* 2002;277:37678–37684. [PubMed: 12145293]
- Sawada H, Konomi H, Hirosawa K. Characterization of the collagen in the hexagonal lattice of Descemet's membrane: its relation to type VIII collagen. *J Cell Biol* 1990;110:219–227. [PubMed: 2104858]
- Schmid TM, Linsenmayer TF. Immunoelectron microscopy of type X collagen: supramolecular forms within embryonic chick cartilage. *Dev Biol* 1990;138:53–62. [PubMed: 2307289]
- Schmid TM, Popp RG, Linsenmayer TF. Hypertrophic cartilage matrix. Type X collagen, supramolecular assembly, and calcification. *Ann N Y Acad Sci* 1990;580:64–73. [PubMed: 2186696]
- Shuttleworth CA. Type VIII collagen. *Int J Biochem Cell Biol* 1997;29:1145–1148. [PubMed: 9438378]
- Snellman A, Tuomisto A, Koski A, Latvanlehto A, Pihlajaniemi T. The role of disulfide bonds and alpha-helical coiled-coils in the biosynthesis of type XIII collagen and other collagenous transmembrane proteins. *J Biol Chem* 2007;282:14898–14905. [PubMed: 17344215]
- Soderberg L, Dahlqvist C, Kakuyama H, Thyberg J, Ito A, Winblad B, Naslund J, Tjernberg LO. Collagenous Alzheimer amyloid plaque component assembles amyloid fibrils into protease resistant aggregates. *FEBS J* 2005;272:2231–2236. [PubMed: 15853808]
- Stephan S, Sherratt MJ, Hodson N, Shuttleworth CA, Kielty CM. Expression and supramolecular assembly of recombinant alpha1(VIII) and alpha2(VIII) collagen homotrimers. *J Biol Chem* 2004;279:21469–21477. [PubMed: 14990571]
- Sumiyoshi H, Mor N, Lee SY, Doty S, Henderson S, Tanaka S, Yoshioka H, Rattan S, Ramirez F. Esophageal muscle physiology and morphogenesis require assembly of a collagen XIX-rich basement membrane zone. *J Cell Biol* 2004;166:591–600. [PubMed: 15302855]
- Svoboda KK, Nishimura I, Sugrue SP, Ninomiya Y, Olsen BR. Embryonic chicken cornea and cartilage synthesize type IX collagen molecules with different aminoterminal domains. *Proc Natl Acad Sci USA* 1988;85:7496–7500. [PubMed: 3050996]
- Tuomisto A, Sund M, Tahkola J, Latvanlehto A, Savolainen ER, Autio-Harmainen H, Liakka A, Sormunen R, Vuoristo J, West A, Lahesmaa R, Morse HC 3rd, Pihlajaniemi T. A mutant collagen XIII alters intestinal expression of immune response genes and predisposes transgenic mice to develop B-cell lymphomas. *Cancer Res* 2008;68:10324–10332. [PubMed: 19074901]
- Uitto J. Epidermolysis bullosa: prospects for cell-based therapies. *J Invest Dermatol* 2008;128:2140–2142. [PubMed: 18695685]
- Uitto J, Pulkkinen L. Molecular complexity of the cutaneous basement membrane zone. *Mol Biol Rep* 1996;23:35–46. [PubMed: 8983017]
- Vaisanen T, Vaisanen MR, Autio-Harmainen H, Pihlajaniemi T. Type XIII collagen expression is induced during malignant transformation in various epithelial and mesenchymal tumours. *J Pathol* 2005;207:324–335. [PubMed: 16110459]
- Veit G, Zimina EP, Franzke CW, Kutsch S, Siebolds U, Gordon MK, Bruckner-Tuderman L, Koch M. Shedding of collagen XXIII is mediated by furin and depends on the plasma membrane microenvironment. *J Biol Chem* 2007;282:27424–27435. [PubMed: 17627939]
- Villone D, Fritsch A, Koch M, Bruckner-Tuderman L, Hansen U, Bruckner P. Supramolecular interactions in the dermo-epidermal junction zone: anchoring fibrilcollagen VII tightly binds to banded collagen fibrils. *J Biol Chem* 2008;283:24506–24513. [PubMed: 18599485]
- Walchli C, Koch M, Chiquet M, Odermatt BF, Trueb B. Tissue-specific expression of the fibril-associated collagens XII and XIV. *J Cell Sci* 1994;107:669–681. [PubMed: 8207089]
- Wenstrup RJ, Florer JB, Davidson JM, Phillips CL, Pfeiffer BJ, Menezes DW, Chervoneva I, Birk DE. Murine model of the Ehlers-Danlos syndrome. col5a1 haploinsufficiency disrupts collagen fibril assembly at multiple stages. *J Biol Chem* 2006;281:12888–12895. [PubMed: 16492673]
- Wessel H, Anderson S, Fite D, Halvas E, Hempel J, SundarRaj N. Type XII collagen contributes to diversities in human corneal and limbal extracellular matrices. *Invest Ophthalmol Vis Sci* 1997;38:2408–2422. [PubMed: 9344363]

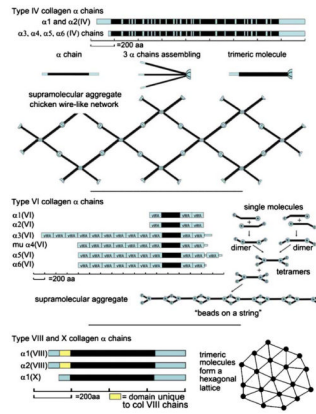
- Yamaguchi N, Kimura S, McBride OW, Hori H, Yamada Y, Kanamori T, Yamakoshi H, Nagai Y. Molecular cloning and partial characterization of a novel collagen chain, alpha 1(XVI), consisting of repetitive collagenous domains and cysteine-containing noncollagenous segments. *J Biochem* 1992;112:856–863. [PubMed: 1284248]
- Ylikarppa R, Eklund L, Sormunen R, Muona A, Fukai N, Olsen BR, Pihlajaniemi T. Double knockout mice reveal a lack of major functional compensation between collagens XV and XVIII. *Matrix Biol* 2003;22:443–448. [PubMed: 14614990]
- Yoshioka H, Zhang H, Ramirez F, Mattei MG, Moradi-Ameli M, van der Rest M, Gordon MK. Synteny between the loci for a novel FACIT-like collagen locus (D6S228E) and alpha 1 (IX) collagen (COL9A1) on 6q12-q14 in humans. *Genomics* 1992;13:884–886. [PubMed: 1639419]
- Yurchenco PD, Ruben GC. Type IV collagen lateral associations in the EHS tumor matrix. Comparison with amniotic and in vitro networks. *Am J Pathol* 1988;132:278–291. [PubMed: 3400773]



**Fig. 1.** Assembly of  $\alpha$  chains into trimeric collagen molecules and then of molecules into fibrils. *Top* The folding of three chains into a single molecule. *Middle* Fibrillar collagen  $\alpha$  chains showing domain structures. Scale is approximate (C-pro C-propeptide domain, TH gly-X-Y repeating region). The yellow triangle is a short non-gly-X-Y domain in collagens XXIV and XXVII and has been interpreted in two ways: as a separation between the minor and major helices in collagen XXIV, and as an interruption in the major (and only) triple helix in collagen XXVII. *Bottom* Perfect quarter stagger overlaps result when no bulky groups protrude from the fibril surface. *Bottom right* Electron micrograph of collagen fibrils in the rabbit cornea. The thin diameters (~25 nm) are the result of fibrils being heterotypic structures composed of ~80% type I and 20% type V collagen.  $\times 60,000$

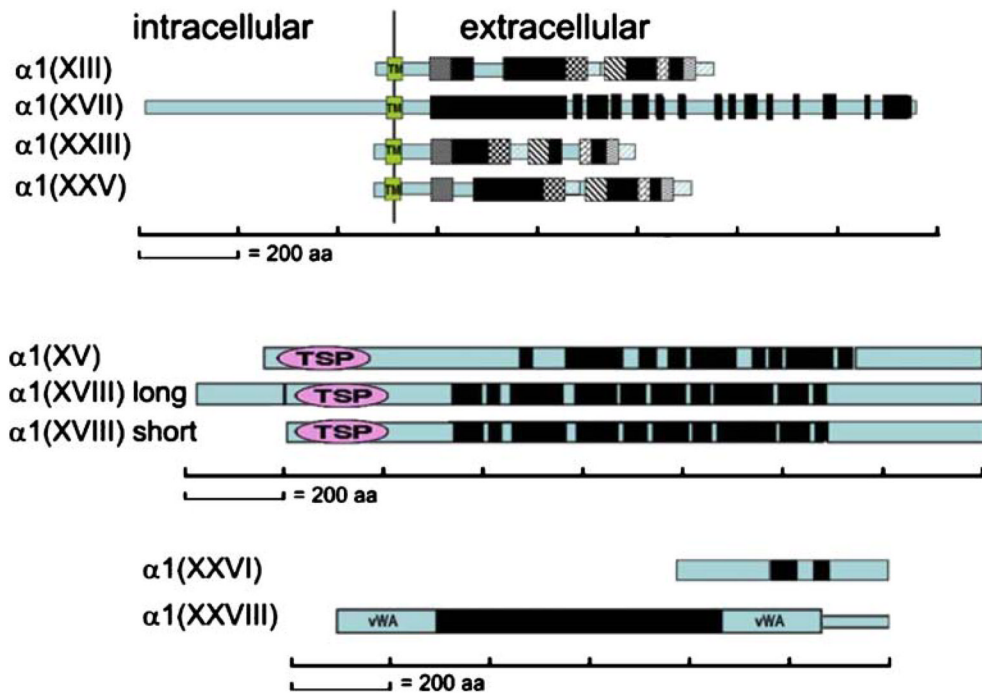


**Fig. 2.** Collagens associated with banded fibrils. *Top* FACIT  $\alpha$  chains, FACIT-like  $\alpha$  chains, and the type VII  $\alpha 1$  chain. Domains are indicated *below*. Scale is approximate. *Bottom left* Long and short forms of types IX, XII, and XIV collagen on the surface of a fibril. *Bottom middle* The presence of a few FACITs allows the collagen fibrils to come into contact (*double-headed arrows*) and to fuse forming larger diameter structures. More FACITs associated with the fibril surface hinders fibril fusions. This has been shown recently for type XIV collagen. *Bottom right* A collagen VII antiparallel dimer functioning as a “rivet” linking the epithelial basement membrane with the banded fibrils of the dermis



**Fig. 3.**

Network-forming collagens. *Top* Representative type IV collagen chains. The chicken wire supramolecular aggregate of type IV collagen is shown *below*. *Middle* Type VI collagen chains, including the murine  $\alpha 4$ , which does not have a human counterpart. Single molecules forming dimers, and dimers interacting to become tetramers are presented *right*. The tetramers aggregate end to end. *Bottom* The  $\alpha$  chains for short chain collagens, types VIII and X (*yellow boxes* domains unique to the collagen VIII  $\alpha$  chains, *blue boxes* conserved non-collagenous domains, *black boxes* collagenous domains that are also conserved between collagen VIII and X). A drawing of a hexagonal lattice, a supramolecular aggregate that these collagens can assume, is shown *right*



**Fig. 4.** Transmembrane, endostatin precursor, and other collagens. The  $\alpha$  chains for each collagen are shown (black bars collagenous domains, blue bars NC domains, TM in green boxes transmembrane domains). In the transmembrane group, matching geometric patterns in collagens XIII, XXIII, and XXV indicate conserved regions (TSP thrombospondin N-terminal module, vWA von Willebrand factor A domains)

**Table 1**

Collagen  $\alpha$  chains, number of amino acids (*aa*), and National Center for Biotechnology Information (*NCBI*) reference numbers (*SP* signal peptide, *vWC* von Willebrand factor C domain, *FNIII* fibronectin type III domain)

Collagen $\alpha$ chain	Number of amino acids	NCBI reference number
$\alpha 1(I)$	1464 (includes 22 aa SP)	NP_000079
$\alpha 2(I)$	1366 aa (includes SP)	NP_000080
$\alpha 1(II)A$	1487 aa (includes 25 aa SP)	NP_001835
$\alpha 1(II)B$	Same as $\alpha 1(II)A$ but lacks <i>vWC</i> domain	NP_149162
$\alpha 1(III)$	1466 aa (includes 23 aa SP)	NP_000081
$\alpha 1(IV)$	1669 aa (includes 27 aa SP)	NP_001836
$\alpha 2(IV)$	1712 aa (includes 25 aa SP)	NP_001837
$\alpha 3(IV)$	1670 aa (includes 28 aa SP)	NP_000082
$\alpha 4(IV)$	1690 aa (includes 38 aa SP)	NP_000083
$\alpha 5(IV)$	1685 aa (includes 26 aa SP)	NP_000486
$\alpha 6(IV)$	1691 aa (includes 21 aa SP)	NP_001838 (B isoform NP_378667)
$\alpha 1(V)$	1838 aa (includes SP)	NP_000084
$\alpha 2(V)$	1499 aa (includes SP)	NP_000384
$\alpha 3(V)$	1745 aa (includes 29 aa SP)	NP_056534
$\alpha 1(VI)$	1028 aa (includes 19 aa SP)	NP_001839
$\alpha 2(VI)$	1019 aa (includes 20 aa SP)	NP_001840
	2C2a and 2C2a' isoforms	NP_478054 and NP_478055
$\alpha 3(VI)$	3177 (with 25 aa SP)	NP_004360 - multiple splicings
$\mu \alpha 4(VI)$	Not in human; mouse=2309 aa	Swiss-Prot A2AX52
$\alpha 5(VI)$	2611 (includes SP)	NP_694996 (partial-shows 2526 aa)
	Isoforms 2 and 3	NP_203699 and NP_203700
$\alpha 6(VI)$	2263 aa (includes SP)	NP_001096078
$\alpha 1(VII)$	2944 aa (includes 16 aa SP)	NP_000085
$\alpha 1(VIII)$	744 aa (includes 28 aa SP)	NP_065084; NP_001841
$\alpha 2(VIII)$	703 aa (includes SP)	NP_005193
$\alpha 1(IX)$	921 aa (includes 23 aa SP)	NP_001842
	Short form 678 aa (includes 23 aa SP)	NP_511040
$\alpha 2(IX)$	689 aa (includes SP)	NP_001843
$\alpha 3(IX)$	684 aa (includes SP)	NP_001844
$\alpha 1(X)$	680 aa (includes 18 aa SP)	NP_000484
$\alpha 1(XI)A$	1806 aa (includes 36 aa SP)	NP_001845
$\alpha 1(XI)B$	1818 aa (includes 36 aa SP)	NP_542196
$\alpha 1(XI)C$	1767 aa (includes 36 aa SP)	NP_542197
$\alpha 2(XI)$	1736 aa (includes 22 aa SP)	NP_542411
	Isoforms 2 and 3	NP_542412 and NP_542410
$\alpha 3(XI)$	Same as $\alpha 1(II)A$	NP_001835
$\alpha 1(XII)$	3063 aa (includes 23 aa SP)	NP_004361 (has NC1 variants)
	Short form 1899 aa, includes same SP as long form	NP_542376 (has NC1 splice variants)

Collagen $\alpha$ chain	Number of amino acids	NCBI reference number
$\alpha$ 1(XIII)	717 aa (transmembranous)	NP_005194 (20+ splice variants)
$\alpha$ 1(XIV)	1796 aa (includes SP)	NP_066933 (has NC1 variants)
	Short form without N-terminal FNIII domain; also has NC1 variants)	
$\alpha$ 1(XV)	1388 aa (includes 25 aa SP)	NP_001846
$\alpha$ 1(XVI)	1604 aa (includes SP)	NP_001847
$\alpha$ 1(XVII)	1497 aa (transmembranous)	NP_000485
$\alpha$ 1(XVIII)	1516 aa (includes 23 aa SP)	NP_085059
	Short form 1336 aa (includes 33 aa SP)	NP_569712
$\alpha$ 1(XIX)	1142 aa (includes 23 aa SP)	NP_001849
ch $\alpha$ 1(XX)	Not in human; ch=1472 aa (without SP)	NP_001004392
$\alpha$ 1(XXI)	957 aa (includes 22 aa SP)	NP_110447
$\alpha$ 1(XXII)	1626 aa (includes SP)	NP_690848
$\alpha$ 1(XXIII)	540 aa (transmembranous)	NP_775736
$\alpha$ 1(XXIV)	1714 aa (includes SP)	NP_690850
$\alpha$ 1(XXV)	654 aa (transmembranous)	NP_942014
	Isoform 2 is 642 aa	NP_115907
$\alpha$ 1(XXVI)	439 aa (includes SP)	NP_597714
$\alpha$ 1(XXVII)	1860 aa (includes 41 aa SP)	NP_116277
$\alpha$ 1(XXVIII)	1125 aa (includes SP)	NP_001032852