INTERNATIONAL JOURNAL OF Experimental Pathology

Int. J. Exp. Path. (2017)

REVIEW ARTICLE

Fell Muir Review: Collagen fibril formation *in vitro* and *in vivo*

Karl E. Kadler

Faculty of Biology, Medicine and Health, Wellcome Trust Centre for Cell-Matrix Research, Manchester Academic Health Science Centre, University of Manchester, Manchester, UK

INTERNATIONAL JOURNAL OF EXPERIMENTAL PATHOLOGY

doi: 10.1111/iep.12224

Received for publication: 9 April 2016
Accepted for publication: 21 January 2017

Correspondence:
Karl E. Kadler
Faculty of Biology, Medicine and Health
Wellcome Trust Centre for Cell-Matrix Research
Manchester Academic Health Science Centre
University of Manchester
Michael Smith Building
Oxford Road
Manchester M13 9PT
UK
E-mail: karl.kadler@manchester.ac.uk

SUMMARY

It is a great honour to be awarded the Fell Muir Prize for 2016 by the British Society of Matrix Biology. As recipient of the prize, I am taking the opportunity to write a minireview on collagen fibrillogenesis, which has been the focus of my research for 33 years. This is the process by which triple helical collagen molecules assemble into centimetre-long fibrils in the extracellular matrix of animals. The fibrils appeared a billion years ago at the dawn of multicellular animal life as the primary scaffold for tissue morphogenesis. The fibrils occur in exquisite three-dimensional architectures that match the physical demands of tissues, for example orthogonal lattices in cornea, basket weaves in skin and blood vessels, and parallel bundles in tendon, ligament and nerves. The question of how collagen fibrils are formed was posed at the end of the nineteenth century. Since then, we have learned about the structure of DNA and the peptide bond, understood how plants capture the sun's energy, cloned animals, discovered antibiotics and found ways of editing our genome in the pursuit of new cures for diseases. However, how cells generate tissues from collagen fibrils remains one of the big unsolved mysteries in biology. In this review, I will give a personal account of the topic and highlight some of the approaches that my research group are taking to find new insights.

Keywords
collagen, electron microscopy, fibril, fibripositor, procollagen, tendon

A brief introduction to collagen

There are several excellent reviews on the collagen family and collagen structure [two such examples are Mienaltowski and Birk (2014) and Bella (2016)], and therefore, only a brief account will be given here.

Collagens are a large family of proteins that have three left-handed polyproline II-like helices wound into a right-handed supercoiled triple helix. The chains have a repeating Gly-X-Y triplet in which glycine is located at every third residue position and X and Y are frequently occupied by the imino acids proline and hydroxyproline [see Bella *et al.* 1994; Brodsky & Ramshaw 1997; Brodsky & Persikov 2005; reviewed by Bella (2016)]. The first 20 or so collagens were identified in animal tissues at the protein level and were assigned Roman numerals [reviewed by Myllyharju and Kivirikko (2004)]. However, with the advent of genome sequencing it became apparent that many more collagens exist. We now know that there are 28 distinct collagens in vertebrates [Huxley-Jones *et al.* 2007; reviewed by Kadler *et al.* (2007) and Mienaltowski and Birk (2014)], almost 200 in *Caenorhabditis elegans* [reviewed by Johnstone (2000)], and further collagens in marine invertebrates (Trotter & Koob 1989; Thurmond & Trotter 1994; Exposito *et al.* 2010), bacteria [see Ghosh *et al.* (2012) and references therein] and viruses (e.g. see Rasmussen *et al.* 2003; Legendre *et al.* 2011). It has become clear that the triple helix is an important motif that is not restricted to collagens (Brodsky & Shah 1995) but which occurs in a wide range of proteins including asymmetric acetylcholinesterase (Johnson *et al.* 1977), macrophage scavenging receptors (Kodama *et al.* 1990), complement component C1q (Reid & Day 1990), ectodysplasin (Ezer *et al.* 1999), and the mannose-binding lectin, collectins and ficolins in the lectin pathway (Garred *et al.* 2016) that are involved in mediating host–pathogen interactions (Berisio & Vitagliano 2012).

The polypeptide chains in collagens are termed α-chains. Because there are numerous collagen genes and their protein products trimerize in a specific combination to produce a collagen 'type', a nomenclature has evolved to specify a particular α-chain based on the collagen type in which it is found. The nomenclature involves the α symbol followed by an Arabic number followed by a Roman numeral, in brackets (parentheses). The α symbol, Arabic number and Roman numeral are read together to indicate the gene that encodes that particular α-chain. Thus, α1(I) and α2(I) denote that these chains are found in type I collagen and are encoded by the genes COL1A1 and COL1A2.

Collagens can be homotrimers and heterotrimers. Moreover, some collagens of the same type can be homotrimeric or heterotrimer, e.g. type I collagen can exist as a homotrimer of three α1(I) chains (i.e. $[\alpha1(I)]_3$) or a heterotrimer of two α1(I) chains and a single α2(I) chain (i.e. $[\alpha1(I)]_2$, α2(I)). Furthermore, heterotrimeric collagens can have three different α-chains (e.g. α1(IX), α2(IX) and α3(IX)) that are encoded by three different genes (i.e. COL9A1, COL9A2 and COL9A3 respectively), and some collagen types contain specific combinations of a family of six chains [e.g. $[\alpha1(IV)]_2$, α2(IV) and α3(IV), α4(IV), α5(IV) and α5(IV), α5(IV), α6(IV) (see Hudson *et al.* (2003) for a review)]. There is chain selection specificity such that of the 45 different collagen α-chains in vertebrates, only 28 different types occur (Table 1). For fibrillar collagens (Figure 1), the chain selection mechanism resides in the non-collagenous sequences at the C-terminal end of each pro-α-chain (Lees *et al.* 1997; Bourhis *et al.* 2012). The chain selection mechanism in other collagens is less well understood.

Table 1 Collagen types and their chain compositions

Type	Gene	α-chain	Molecules
I	COL1A1	α1(I)	$[\alpha1(I)]_3$
	COL1A2	α2(I)	$[\alpha1(I)]_2\alpha2(I)$
II	COL2A1	α1(II)	$[\alpha1(II)]_3$
III	COL3A1	α1(III)	$[\alpha1(III)]_3$
IV	COL4A1	α1(IV)	$[\alpha1(IV)]_2\alpha2(IV)$
	COL4A2	α2(IV)	α3(IV), α4(IV), α5(IV)
	COL4A3	α3(IV)	α5(IV), α5(IV), α6(IV)
	COL4A4	α4(IV)	
	COL4A5	α5(IV)	
	COL4A6	α6(IV)	
V	COL5A1	α1(V)	$[\alpha1(V)]_2\alpha2(V)$
	COL5A2	α2(V)	$\alpha1(V), \alpha2(V), \alpha3(V)]_3$
	COL5A3	α3(V)	$[\alpha3(V)]_3$
VI	COL6A1	α1(VI)	α1(VI), α2(VI) and any of
	COL6A1	α2(VI)	α3(VI), α4(VI), α5(VI) and
	COL6A1	α3(VI)	α6(VI) (Maass *et al.* 2016)
	COL6A1	α4(VI)	
	COL6A1	α5(VI)	
	COL6A1	α6(VI)	
VII	COL7A1	α1(VII)	$[\alpha1(VII)]_3$
VIII	COL8A1	α1(VIII)	$[\alpha1(VIII)]_3$
	COL8A2	α2(VIII)	$[\alpha2(VIII)]_3$
IX	COL9A1	α1(IX)	α1(IX), α2(IX), α3(IX)
	COL9A2	α2(IX)	
	COL9A3	α3(IX)	
X	COL10A1	α1(X)	$[\alpha1(X)]_3$
XI	COL11A1	α1(XI)	$[\alpha1(XI)]_2\alpha2(XI)$;α1(XI), α2(XI), α3(XI)*
	COL11A2	α2(XI)	$[\alpha2(XI)]_3$
XII	COL12A1	α1(XII)	$[\alpha1(XII)]_3$
XIII	COL13A1	α1(XIII)	$[\alpha1(XIII)]_3$
XIV	COL14A1	α1(XIV)	$[\alpha1(XIV)]_3$
XV	COL15A1	α1(XV)	$[\alpha1(V)]_3$
XVI	COL16A1	α1(XVI)	
XVII	COL17A1	α1(XVII)	
XVIII	COL18A1	α1(XVIII)	
XIX	COL19A1	α1(XIX)	
XX	COL20A1	α1(XX)	
XXI	COL21A1	α1(XXI)	
XXII	COL22A1	α1(XXII)	
XXIII	COL23A1	α1(XXIII)	
XXIV	COL24A1	α1(XXIV)	
XXV	COL25A1	α1(XXV)	
XXVI	COL26A1	α1(XXVI)	
XXVII	COL27A1	α1(XXVII)	$[\alpha1(XXVII)]_3$
XXVIII	COL28A1	α1(XXVIII)	$[\alpha1(XXVIII)]_3$

*The α3(XI) chain is encoded by the COL2A1 gene.

Fibrillar collagens

The 28 collagen types that occur in vertebrates can be classified according to domain structure, function and supramolecular assembly [for a review, see Mienaltowski and Birk (2014)]. The most abundant are the fibrillar collagens that form the basis of the fibrils in bony, cartilaginous, fibrous and tubular structures and will be the focus of the remainder of this review. The fibril-forming collagens are types I, II, III, V, XI, XXIV and XXVII. They have uninterrupted triple helices of approximately 300 nm in length and have globular domains (propeptides) at each terminus of each α-chain. Types XXIV and XXVII were identified by genome sequencing and were added to this group on the basis of protein domain structure (Koch *et al.* 2003) and the presence of type XXVII collagen in thin fibrils (Plumb *et al.* 2007).

Collagen fibrils are complex macromolecular assemblies that comprise different fibrillar collagen types (Hansen & Bruckner 2003). The fibrils are either 'predominately type I collagen' or 'predominately type II collagen'. Predominately type I collagen fibrils occur in bony, tubular and fibrous tissues, whereas cartilaginous tissues contain predominately type II collagen fibrils. Collagen fibrils range in length from a few microns to centimetres (Craig *et al.* 1989) and therefore have molecular weights in the tera Dalton range [based on calculations described by Chapman (1989)]. The fibrils provide attachment sites for a broad range of macromolecules including fibronectin, proteoglycans and cell surface receptors such as integrins, discoidin domain-containing receptors and mannose receptors (Di Lullo *et al.* 2002; Jokinen *et al.* 2004; Sweeney *et al.* 2008; Orgel *et al.* 2011). Furthermore, the fibrils vary in diameter depending on species, tissue and stage of development (Parry *et al.* 1978; Craig *et al.* 1989) and in response to injury and repair (Pingel *et al.* 2014). Collagen fibrils are arranged in exquisite

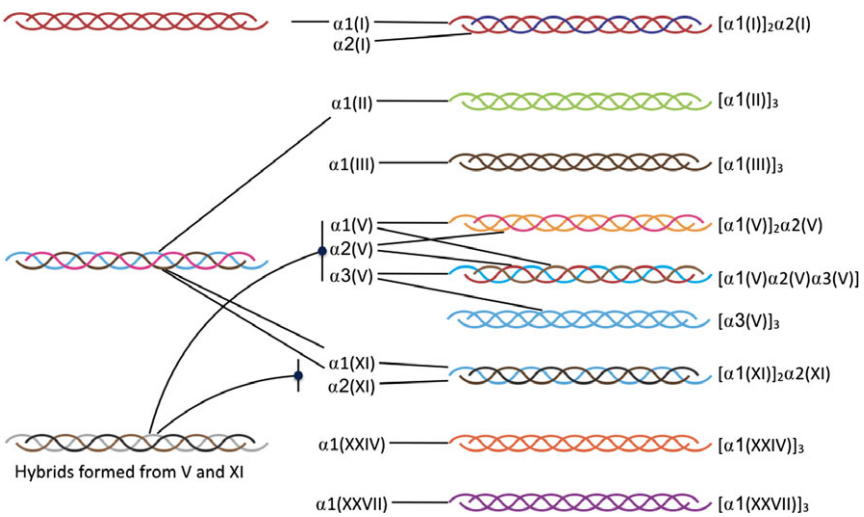

Figure 1 Schematic diagram of the chain composition of the fibril-forming collagens.

three-dimensional architectures *in vivo* including parallel bundles in tendon and ligament, orthogonal lattices in cornea, concentric weaves in bone and blood vessels and basket weaves in skin. How the fibrils assemble, how length and diameter are regulated, how molecules attach to fibril surfaces and how the multiscale organization is achieved are questions for which answers are either sketchy or not available. The extreme size and compositional heterogeneity of collagen fibrils mean that they are extremely difficult to study by conventional molecular, genetic and biochemical approaches. Collagen molecules are also hydroxylated at specific prolyl residues (by prolyl hydroxylases), lysyl residues (by lysyl hydroxylases or PLODs, procollagen-lysine 5-dioxygenases) and are non-reducibly cross-linked [by lysyl oxidase (LOX) enzymes; Eyre *et al.* 1984; Gistelinck *et al.* 2016], which adds to the difficulty of studying the fibrils.

Collagen fibrils in tendon (Heinemeier *et al.* 2013) and cartilage (Heinemeier *et al.* 2016) are extremely long lived with estimates exceeding hundreds of years. Therefore, the collagen in the fibrils is particularly prone to modification by advanced glycation end products (Verzijl *et al.* 2000; Thorpe *et al.* 2010). Thus, the two major experimental approaches used in the study of collagen fibril assembly have been electron microscopy of tissues to describe the organization of the fibrils *in vivo* (explained below) and reconstitution of fibrils *in vitro* using collagen extracted from tissues (explained below) or recombinant collagens (e.g. Fertala *et al.* 1996).

Collagen fibril assembly in vitro

Studies by Gross (Gross & Kirk 1958), Wood & Keech (Wood & Keech 1960), Hodge & Petruska (Hodge 1989), Silver (Silver & Trelstad 1980) and Chapman (Bard & Chapman 1968), to name a few, showed that exposure of animal tissues (typically skin and tendon) to weak acidic solutions (typically acetic acid) or neutral salt buffers yielded a solution of collagen molecules that when neutralized and warmed to approximately 30°C, produced elongated fibrils that had the same alternating light and dark transmission electron microscope banding appearance as fibrils occurring *in vivo* (Holmes & Chapman 1979) (Figure 2). The characteristic banding pattern of the fibrils arises from *D*-staggering of triple helical collagen molecules that are 4.4 × *D* in length (where *D* is 67 nm, to a close approximation). The electron-dense stain used at neutral pH penetrates more readily into regions of least protein packing (the 'gaps') between the N- and C-termini of collagen molecules that are aligned head-to-tail along the long axis of the fibril. The fact that fibrils with *D*-periodic banding could be formed *in vitro* from purified collagen showed that all the information required to form a collagen fibril was contained within the amino acid sequence and triple helical structure of the collagen molecule (Hulmes *et al.* 1973).

Subsequent studies showed that collagen fibrils from embryonic tendon (which are predominantly type I collagen) exist in two isoforms: unipolar and bipolar (Holmes *et al.* 1994; Figure 3). Unipolar fibrils have all collagen molecules in the fibril oriented in one direction, which gives the fibril a carboxyl and an amino tip. Bipolar fibrils (more precisely, N, N-bipolar fibrils) have two amino-terminal ends and a molecular polarity switch region (or transition zone) in which the orientation of collagen molecules switches, for example, from N-to-C to C-to-N (Holmes *et al.* 1994). The switch in orientation occurs over an 8 *D*-period range in chick tendon collagen fibrils (Holmes *et al.* 1994). Notably, sea cucumbers (*Actinopyga echinites*) lack unipolar fibrils and all their bipolar fibrils have the molecular switch region located precisely midway from each fibril tip; also, the switch varies in extent from 14 to 41 *D*-periods in invertebrate fibrils (Trotter *et al.* 1998, 2000). Earlier studies had shown that collagen fibrils formed by cleavage of procollagen to collagen (explained below) grow from pointed tips (i.e. the pointed ends of fibrils) and the collagen molecules were oriented in one direction along the long axis of the fibril (Kadler *et al.* 1990). Moreover, the C-tip of a unipolar

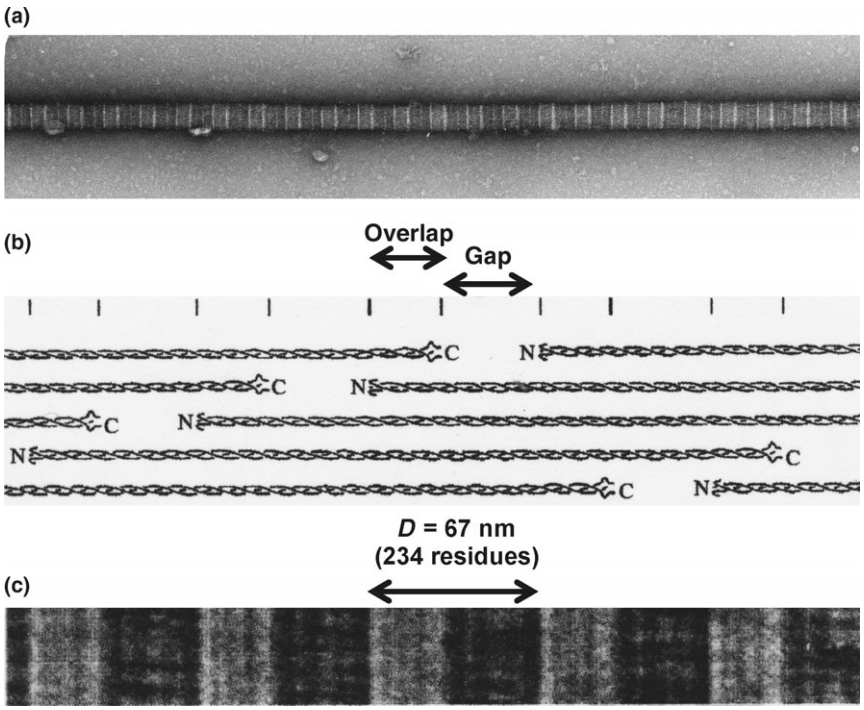

Figure 2 Transmission electron microscopy of individual collagen fibrils. (a) Single collagen fibril from 18-day chick embryonic metatarsal tendon. The fibril is negatively stained with 2% uranyl acetate to show the characteristic light and dark banding pattern. (b) Schematic representation of the axial arrangement of collagen molecules in a collagen fibril. Each collagen molecule is represented with three coiled chains with amino- and carboxy-termini indicated. Each molecule is 4.4 × *D* in length, where *D* approximately 67 nm. The *D*-stagger of molecules that are 4.4 *D* long leads to the formation of a gap zone in the axial structure. (c) The characteristic negative staining pattern of collagen fibrils, as shown by 1% sodium phosphotungstate staining at neutral pH. Images recorded by D. Holmes.

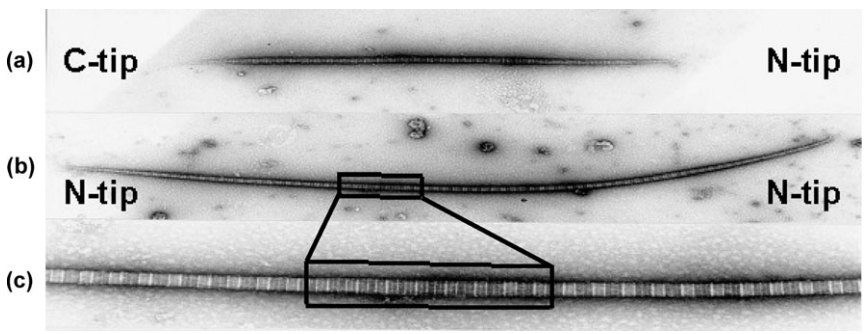

Figure 3 Unipolar and bipolar collagen fibrils. (a) Negatively stained unipolar collagen fibril isolated from embryonic chick metatarsal tendon. Analysis of the staining pattern shows that the collagen molecules are oriented with their amino-terminal to the right-hand side (as shown) and the carboxy-termini to the left. (b) Negatively stained N, N-bipolar collagen fibril from embryonic chick tendon showing the molecular polarity switch region (box). (c) Enlargement of the box in b. Images recorded by D. Holmes.

fibril is required for end-to-end fusion of either two unipolar fibrils (to generate a new N, N-bipolar fibril) or to one end of an N, N-bipolar to generate a longer N, N-bipolar fibril (Graham *et al.* 2000; Kadler *et al.* 2000). Notably, C, C-bipolar collagen fibrils have not been described. Presumably, the structure of C-tips exposes binding sites to promote carboxyl-to-amino fusion of fibril tips.

Two schools of thought developed about how collagen molecules assemble into fibrils: (i) precipitation from a solution of 'bulk' collagen by liquid crystalline ordering of molecules (e.g. see Martin *et al.* 2000) or (ii) 'nucleation and propagation' in which a finite number of collagen molecules form a nucleus that then grows in length and diameter to become the mature fibril (Gross *et al.* 1954). This latter mechanism is analogous to the formation of inorganic crystals. The existence of fibrils of different lengths supports the notion that the fibrils grow in size (which supports the nucleation and propagation model) but collagen and procollagen molecules can form a liquidlike structure when packed in high concentration (which supports the liquid crystalline model). In reality, these two hypotheses might not be mutually exclusive; work by Hulmes and Bruns showed that procollagen molecules can align in zero-*D* register in secretory vacuoles of fibroblasts (analogous to liquid crystalline packing) (Bruns *et al.* 1979; Hulmes *et al.* 1983), which might increase the rate of conversion of procollagen to collagen to nucleate fibrils. Therefore, it is possible that elements of both assembly mechanisms exist *in vivo*.

A system for generating collagen fibrils in vitro starting with procollagen

In 1984, I joined Darwin Prockop's laboratory at UMDNJ, Piscataway, NJ, USA, to develop a system of studying collagen fibril formation by cleavage of procollagen with its physiological convertases, the procollagen N- and C-proteinases (Figure 4). Procollagen had previously been shown to be the biosynthetic precursor of collagen (Bellamy & Bornstein 1971) and there had been initial success in purifying the N- and C-proteinases that convert procollagen to collagen (Njieha *et al.* 1982; Tuderman & Prockop 1982). With the collaboration of Yoshio Hojima who purified the procollagen N- and C-proteinases from chick tendon (Hojima *et al.* 1985, 1989), we developed a method of purifying type I procollagen and cleaving it with N-proteinase to generate pCcollagen, and then cleaving the repurified pCcollagen with the C-proteinase in a bicarbonate buffer. pCcollagen is a cleavage intermediate of procollagen that retains the C-propeptide but lacks the N-propeptide. This system allowed us to study collagen fibril formation in the absence of lysyl oxidase and cross-link precursors (Eyre *et al.* 2008). The presence of cross-link precursors in extracted collagen can affect collagen fibril formation *in vitro* (Herchenhan *et al.* 2015). Using this new system of forming fibrils by cleavage of procollagen, we defined the thermodynamic parameters of the assembly process (Kadler *et al.* 1987), the temperature dependence of collagen fibril assembly (Kadler *et al.* 1988), and showed that the fibrils form as a nucleus that grows at its pointed tips (Kadler *et al.* 1990). These observations indicated that collagen fibrils (in the absence of lysyl oxidase-derived cross-links) exhibit a critical concentration of assembly, analogous to the self-formation of inorganic crystals. Our ability to purify procollagen from cells paved the way to study how mutations in collagen genes that cause osteogenesis imperfecta affect procollagen structure and fibril assembly. These studies showed that mutations in type I collagen genes can produce procollagen molecules that are 'kinked' (Vogel *et al.* 1988), slow the rate of conversion of procollagen to collagen (Lightfoot *et al.* 1992), lead to the formation of abnormal collagen fibrils (Kadler *et al.* 1991) and impair the ability of collagen fibrils to be mineralized during the formation of bone (Culbert *et al.* 1995). These studies led to a better understanding of how mutations in collagen genes can change the structure and processing of collagen molecules and how the resultant collagen fibrils are poorer scaffolds for mineralization, as occurs in osteogenesis imperfecta *in vivo* (Culbert *et al.* 1996). In parallel studies, we also showed that the tips are the sites of diameter regulation (Holmes *et al.* 1998), that fibrils formed at low C-proteinase/pCcollagen ratios bore the closest resemblance to fibrils in vivo (Holmes *et al.* 1996) and that the tips of fibrils are paraboloidal in shape (Holmes *et al.* 1992).

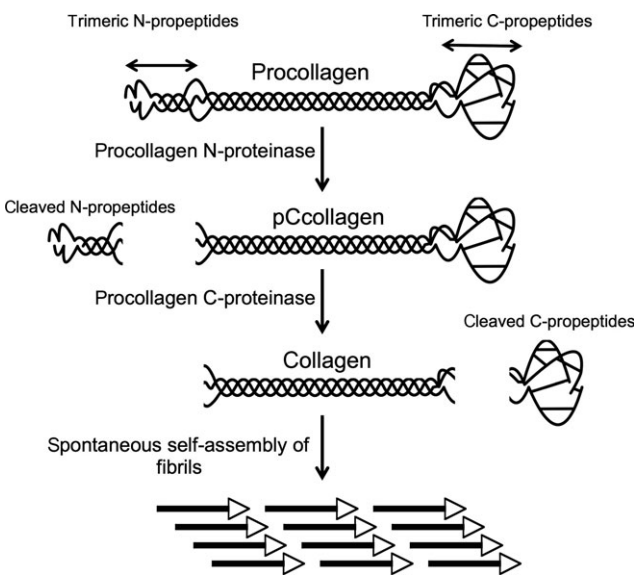

Figure 4 Schematic representation of collagen fibril formation by cleavage of procollagen. Sequential cleavage of the N-propeptides (by procollagen N-proteinase, which are ADAM 2, 3, 14) and the C-propeptides (by procollagen C-proteinase, which are the BMP-1/tolloid family) of procollagen generates collagen that self-assembles into unipolar collagen fibrils (Kadler *et al.* 1987).

Collagen fibril formation in vivo

Although collagen molecules can spontaneously self-assemble into fibrils *in vitro*, additional factors must exist *in vivo* to explain the exquisite three-dimensional supramolecular organization of fibrils, as well as the regulation of diameter, length and composition, that depend on tissue, stage of development, state of tissue ageing and repair, and which vary in disease. The *in vivo* regulation of collagen fibril formation has been studied for over a century, and although enormous progress has been made, the cellular mechanisms of fibril assembly and organization *in vivo* remain elusive.

Some of the earliest reports on the existence of collagen fibrils date back to the end of the 19th century and beginning of the 20th century. For example, Mallory described a 'fibrillar substance' produced by connective tissue cells (i.e. fibroblasts) (Mallory 1903). Studies of collagen fibrils continued during the 1920s and 1930s during which time several groups attempted to develop methods to observe the assembly of the fibrils *in vivo*. A breakthrough came in 1940 when Mary Stearns published her first observations of fibroblasts secreting and assembling collagen fibres (Stearns 1940). Her paper is a 'must-read' for students of collagen fibril formation; the 46 hand-drawn plates are exquisite. Stearns used the camera lucida to visualize and draw details of cytoplasmic connections between cells, striations within cells, 'vacuoles de secretion' and fibres growing at the cell surface. In so doing, she produced the first evidence that fibroblasts are instrumental in assembling collagen fibrils in tissues. Almost 40 years later, Trelstad and Hayashi used transmission electron microscopy (TEM) to show that collagen fibrils occurred in invaginations of the plasma

membrane of embryonic fibroblasts (Trelstad & Hayashi 1979). A decade later this observation was extended using high-voltage TEM to study collagen fibrillogenesis in cornea as well as embryonic chick tendon (Birk & Trelstad 1984, 1985, 1986; Trelstad & Birk 1985). In 2006, we used serial section TEM and immunoEM of embryonic tendon to describe a variety of structures at the plasma membrane that contained collagen fibrils, and which we collectively called 'fibripositors' (Canty *et al.* 2004). Collectively, these studies demonstrate the exquisite control the cell exerts over the self-assembly of collagen fibrils to generate tissues with highly organized collagen matrices.

Fibripositors

In 1989, I returned to the UK as a Wellcome Trust Senior Research Fellow in Basic Biomedical Science and joined Michael Grant's Department of Medical Biochemistry. During the next 10 years, we extended our knowledge of how mutations in collagen genes affect procollagen structure and fibril formation. In collaboration with Peter Byers and Gillian Wallis, these studies focussed on the Ehlers–Danlos syndrome (type VII) that is caused by mutations in COL1A1 and COL1A2 genes that encode the chains of type I procollagen. PhD students Rod Watson, Samantha Lightfoot and Ainsley Culbert, and a postdoc David Holmes, joined my laboratory, and together we showed how mutations in COL1A1 and COL1A2 that cause EDS VII disrupt the structure of procollagen, slow the cleavage of procollagen by N-proteinase and lead to the 'cauliflower' appearance of collagen fibrils in affected individuals (Wallis *et al.* 1992; Watson *et al.* 1992, 1998; Holmes *et al.* 1993; Culbert *et al.* 1996). We also studied the function of the CUB domains in bone morphogenetic protein-1, which is a potent procollagen C-proteinase (Garrigue-Antar *et al.* 2001, 2002, 2004; Hartigan *et al.* 2003; Petropoulou *et al.* 2005; Canty *et al.* 2006a,b). Here, CUB is an evolutionary conserved protein domain named after its discovery in complement components (C1r/C1s), the sea urchin protein Uegf and BMP-1 [for a review, see Bork and Beckmann (1993)].

However, it was during a staff meeting in 2002 that I heard good advice that scientists should change their experimental approach every 10 years. Up until this time, I had used cells as a factory for procollagen production and had overlooked the importance of the cell in fibril assembly. A new postdoc in the laboratory, Elizabeth Canty, took up the challenge of taking our laboratory into new, *in vivo*, directions. We were inspired by the work of Hayashi, Trelstad and Birk and decided to ask questions about how cells regulate fibril assembly and fibril number. With the assistance of David Holmes, Roger Meadows, Tobias Starborg and Yinhui Lu in the laboratory, Liz Canty embarked on studying collagen fibril formation in embryonic chick tendon using serial section electron microscopy and 3D reconstruction. Our first paper, in 2004, showed 3D reconstructions from 50 × 100 nm serial sections of embryonic chick tendon, cut perpendicular to the tissue long axis. These were the deepest and most detailed 3D reconstructions at the time and showed fingerlike projections of the plasma membrane containing thin collagen fibrils (Figure 5). The 3D reconstructions showed that the projections were part of an invagination of the plasma membrane and that the fibril within the invagination and the projection were co-aligned to the long axis of the tendon (Canty *et al.* 2004) (Figure 6). We called these structures 'fibripositors' (a portmanteau of 'fibril' and 'depositors'). We also showed that fibripositors are actin-dependent structures (Canty *et al.* 2006a,b) that projected into intercellular channels stabilized by cadherin-11 containing junctions (Richardson *et al.* 2007).

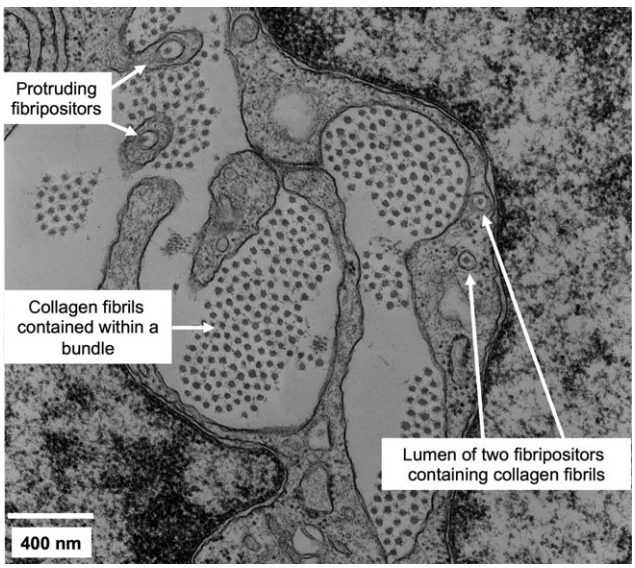

Figure 5 Transmission electron microscopy of embryonic tendon. Embryonic tendon contains bundles of collagen fibrils between adjacent fibroblasts. The image shows profiles of fibripositors. Image obtained by Yinhui Lu.

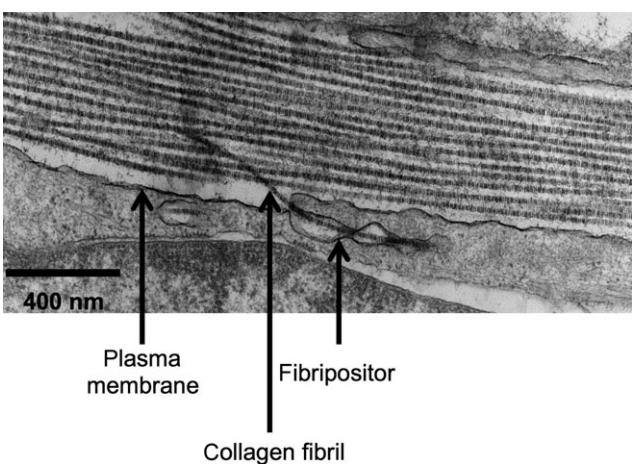

Figure 6 A fibripositor at the plasma membrane of an embryonic fibroblast. Transmission electron microscope image of a collagen fibril contained within a fibripositor at the surface of an embryonic mouse tail-tendon fibroblast. Image obtained by Yinhui Lu.

Serial block face-scanning electron microscopy

The fact that fibripositors are too thin to be seen by light microscopy and that no marker has been identified that can aid in their visualization by fluorescence light microscopy has been a severe hurdle to studies of fibripositor structure, function and formation. Also, the effort and time involved in producing serial sections for electron microscopy is a significant hurdle to further progress; sections can be lost or distorted during processing, and the process requires exceptional skills in ultrathin sectioning and handling. A major breakthrough came with the commercialization of serial block face-scanning electron microscopy (SBF-SEM) (Denk & Horstmann 2004). Here, images of a block face are recorded using a scanning electron microscope prior to the removal of a section by an in-microscope ultramicrotome. The ability to produce serial images without manual sectioning opened up new opportunities to explore fibripositor function. After optimization of sample preparation and staining, image acquisition and data analysis, Toby Starborg, Nick Kalson and Yinhui Lu showed that we could use SBF-SEM as a semi-high-throughput system to examine fibripositor structure and function at the cell–matrix interface (Starborg *et al.* 2013) (Figure 7 and Movie S1). With this new approach, we were able to show that fibripositors are the site of fibril assembly in tendon and that non-muscle myosin II is required for fibril transport and formation (Kalson *et al.* 2013). We also showed that fibripositor-like structures called keratopodia exist in corneal keratocytes (Young *et al.* 2014). SBF-SEM also gave us the opportunity to explore how collagen fibril formation contributes to tendon development. In a *tour de force* of SBF-SEM, Nick Kalson, Yinhui Lu and Susan Taylor outlined a new hypothesis for tendon development in which the number of collagen fibrils is determined by embryonic tendon fibroblasts, and that the growth in lateral size of the tendon is driven by matrix expansion caused by the increase in girth and length of collagen fibrils (Kalson *et al.* 2015). SBF-SEM studies have also revealed a new function for membrane type I-matrix metalloproteinase (MT1-MMP or MMP14) in being essential for tendon development (Taylor *et al.* 2015). Taylor *et al.* showed that release of collagen fibrils from fibripositors at birth requires MT1-MMP and that the process does not rely on the cleavage of collagen at the ¾-¼ vertebrate collagenase cleavage site in the molecule.

Negative regulation of collagen fibril formation during intracellular protein trafficking

Canty *et al.* also made the observation that procollagen can be cleaved to collagen prior to secretion by tendon fibroblasts *in vivo*. Evidence that procollagen can be cleaved to collagen within the cell without forming fibrils demonstrates active negative control of the self-assembly properties of collagen fibrillogenesis *in vivo* (Humphries *et al.* 2008). These observations are in contrast to what happens in conventional cell culture, in which procollagen is readily purified

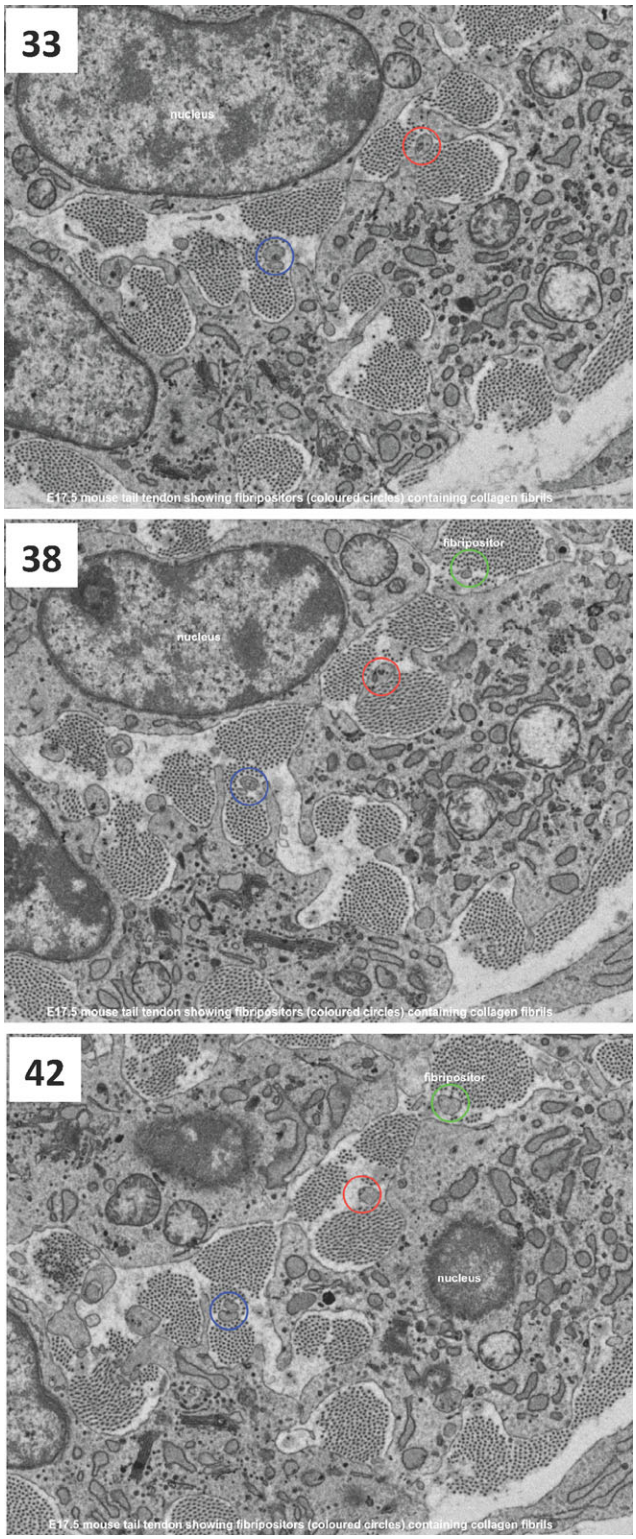

Figure 7 Serial block face-scanning electron microscopy for studies of the cell–matrix interface. Three images from the downloadable Movie S1 generated by serial block face-scanning electron microscopy. The coloured circles show fibripositors. Numbers refer to the image sequence. Images recorded by Tobias Starborg and Nicholas Kalson.

from the cell culture medium. Presumably the environment of the cell and matrix influence the trafficking of procollagen. A halfway house between *in vivo* and *in vitro* is the use of 3D cell culture systems; Kapacee *et al.* showed that fibroblasts incubated in fibrin gels under linear tension replace the fibrin with collagen fibrils that are aligned parallel to the lines of stress and exhibit features of embryonic fibroblasts *in vivo*, including fibripositors (Kapacee *et al.* 2008, 2010; Bayer *et al.* 2010; Kalson *et al.* 2010, 2011). This approach facilitates studies of the role of cells, in a near-physiological environment with tissue-derived mechanical forces, in assembling collagen fibrils.

Regulators of collagen fibril assembly in vivo

The fact that collagen fibrils are comprised of different collagens, that they occur in different numbers and with different diameters and packing densities in different tissues, that the supramolecular organization of fibrils is different in different tissues and that collagen molecules provide interaction sites for receptors and a wide range of extracellular matrix molecules suggests that there are multiple steps in the assembly and organization of fibrils and that each step can be error prone. Defective collagen fibrillogenesis can arise from mutations in genes encoding fibrillar collagens (Table 2), fibril-associated collagens with interrupted triple helices that bind to the surfaces of collagens fibrils, for example type XII and type XIV collagen (Young *et al.* 2002); proteoglycans that interact with fibrils, for example decorin (Danielson *et al.* 1997), lumican (Chakravarti *et al.* 1998) fibromodulin (Hedlund *et al.* 1994; Svensson *et al.* 1999), osteoglycin (Tasheva *et al.* 2002), keratocan (Liu *et al.* 2003) and biglycan (Heegaard *et al.* 2007) [for a review, see Kalamajski and Oldberg (2010)]; enzymes required for post-translational modification of collagen α-chains, for example prolyl 4-hydroxylase (Mussini *et al.* 1967), lysyl hydroxylases (Takaluoma *et al.* 2007) and lysyl oxidases (Maki *et al.* 2002); proteins involved in transporting collagens through the secretory pathway, for example HSP47 (Satoh *et al.* 1996), sedlin (Venditti *et al.* 2012) and TANGO1 (Saito *et al.* 2009; Wilson *et al.* 2011); and proteinases involved in collagen turnover, for example MMP14 (Taylor *et al.* 2015). Loss of the collagen network in cartilage occurs in end-stage osteoarthritis (Ehrlich *et al.* 1977). Conversely ectopic or excessive accumulation of collagen occurs in fibrosis, which can be stimulated by TGF-β (Roberts *et al.* 1986), and can affect any organ often resulting in death. Thus, collagen fibrillogenesis is a precisely regulated process in which the mechanisms that maintain the appropriate number, size and organization of collagen fibrils in adult tissues appear to be sensitive to a wide range of genetic mutations and environmental stimuli.

A personal perspective on some of the most important unanswered questions in the field of collagen fibril homoeostasis

We do not have clear line of sight of how the three-dimensional organization of collagen fibrils is established in tissues. Collagen fibrils first appear part-way through vertebrate embryonic development when the mass of matrix begins to exceed the mass of cells. At this pivotal stage of development, the patterning of tissue progenitor cells has, to a close approximation, been established and might be expected to dictate

Table 2 Diseases caused by mutations in genes encoding fibrillar collagens

Collagen type	Gene	OMIM	Disease	Mouse models
I	COL1A1	120150	Osteogenesis imperfecta (OI); Ehlers–Danlos syndrome type VII	Mov13 (Bonadio *et al.* 1990); Col1a1(Jrt/+) OI/EDS mouse (Chen *et al.* 2014)
	COL1A2	120160	Osteogenesis imperfecta (OI); Ehlers–Danlos syndrome type VII	OIM (Chipman *et al.* 1993); Col1a2(+/G610C) OI (Amish) mouse (Daley *et al.* 2010)
II	COL2A1	120140	Stickler syndrome; achondrogenesis; familial avascular necrosis of the femoral head; Legg–Calve–Perthes disease Kniest dysplasia; spondyloepiphyseal dysplasia congenita (SEDC); Czech dysplasia; myopia 2; Marshall syndrome; epiphyseal dysplasia; platyspondylic lethal skeletal dysplasia	Garofalo *et al.* (1991), Vandenberg *et al.* (1991), Li *et al.* (1995a,b), Gaiser *et al.* (2002), Donahue *et al.* (2003)
III	COL3A1	120180	Ehlers–Danlos syndrome type IV; intracranial berry aneurysm	Liu *et al.* (1997); Tsk2 mouse (Long *et al.* 2015)
V	COL5A1	120215	Nail patella syndrome; Ehlers–Danlos syndrome classic type	Wenstrup *et al.* (2004)
	COL5A2	120190	Ehlers–Danlos syndrome type I or type II	Andrikopoulos *et al.* (1995)
	COL5A3	120216		Huang *et al.* (2011)
XI	COL11A1	120280	Stickler syndrome; otospondylomegaepiphyseal dysplasia (OSMED); Marshall syndrome	Cho/cho mouse (Li *et al.* 1995a,b)
	COL11A2	120290	Stickler syndrome; otospondylomegaepiphyseal dysplasia	McGuirt *et al.* (1999), Li *et al.* (2001)
XXIV	COL24A1	610025		
XXVII	COL27A1	608461	Steel syndrome (Gonzaga-Jauregui *et al.* 2015)	Plumb *et al.* (2011)

the patterning of the matrix. Perhaps novel insights into organogenesis will come from a better understanding of the interplay between cell positioning, cell–cell communication, cell–matrix interactions, cell polarity, the role of the secretory pathway in directing matrix assembly, and mechanical forces.

A further exciting area of research is matrix homoeostasis; it will be fascinating to learn how changes in this process lead to diseases such as osteoarthritis, tendinopathies, fibrosis and cell migration through the matrix. The realization that the bulk of the collagen in tendon and cartilage is synthesized during adolescence and remains unchanged during the lifetime of a person raises intriguing questions about how the collagen network is maintained during life despite countless cycles of mechanical loading. Advances in genome editing and super-resolution light microscopy are all likely to be brought to bear on this question. These approaches are expected to lead to a better understanding of how matrix homoeostasis goes wrong in diseases such as fibrosis, where ectopic and excessive deposition of collagen fibrils can cause death. Recent discoveries show that matrix-rich tissues are peripheral circadian clock tissues and that defects in the rhythm in these tissues lead to pathologies such as calcific tendinopathy (Yeung *et al.* 2014), osteoarthritis (Guo *et al.* 2015; Dudek *et al.* 2016a,b) and intervertebral disc disease (Dudek *et al.* 2016a,b). Thus, the mechanical environment of the cell, the role of the matrix in modulating cell behaviour and peripheral circadian clocks are all likely to contribute to matrix homoeostasis.

Acknowledgements

I am indebted to The Wellcome Trust for continuous support of my research for over 30 years. Numerous students, postdoctoral fellows and technical staff have made valuable contributions to this research. However, a few people need a special mention: Laure Garrigue-Antar, Samantha Lightfoot and Rod Watson performed the studies on osteogenesis imperfecta, the Ehlers–Danlos syndrome and bone morphogenetic protein-1 that established my laboratory at the University of Manchester. In recent years, my gratitude goes out to Nick Kalson, Zoher Kapacee, Elizabeth Laird (nee Canty), Yinhui Lu, Susan Taylor (nee Richardson), Toby Starborg and Chloé Yeung who led the work on 3D electron microscopy, fibripositors and tendonlike constructs. Finally, very special thanks go to David Holmes, who I have known since my PhD days, and who joined my group in the 1990s. David is an enormously talented electron microscopist who has made outstanding contributions to our research.

Conflict of Interest

There is no conflict of interest to declare.

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**, 31–36.

Bard J.B. & Chapman J.A. (1968) Polymorphism in collagen fibrils precipitated at low pH. *Nature* **219**, 1279–1280.

Bayer M.L., Yeung C.Y., Kadler K.E. *et al.* (2010) The initiation of embryonic-like collagen fibrillogenesis by adult human tendon fibroblasts when cultured under tension. *Biomaterials* **31**, 4889–4897.

Bella J. (2016) Collagen structure: new tricks from a very old dog. *Biochem. J.* **473**, 1001–1025.

Bella J., Eaton M., Brodsky B. & Berman H.M. (1994) Crystal and molecular structure of a collagen-like peptide at 1.9 A resolution. *Science* **266**, 75–81.

Bellamy G. & Bornstein P. (1971) Evidence for procollagen, a biosynthetic precursors of collagen. *Proc. Natl. Acad. Sci. USA* **68**, 1138–1142.

Berisio R. & Vitagliano L. (2012) Polyproline and triple helix motifs in host-pathogen recognition. *Curr. Protein Pept. Sci.* **13**, 855–865.

Birk D.E. & Trelstad R.L. (1984) Extracellular compartments in matrix morphogenesis: collagen fibril, bundle, and lamellar formation by corneal fibroblasts. *J. Cell Biol.* **99**, 2024–2033.

Birk D.E. & Trelstad R.L. (1985) Fibroblasts create compartments in the extracellular space where collagen polymerizes into fibrils and fibrils associate into bundles. *Ann. N. Y. Acad. Sci.* **460**, 258–266.

Birk D.E. & Trelstad R.L. (1986) Extracellular compartments in tendon morphogenesis: collagen fibril, bundle, and macroaggregate formation. *J. Cell Biol.* **103**, 231–240.

Bonadio J., Saunders T.L., Tsai E. *et al.* (1990) Transgenic mouse model of the mild dominant form of osteogenesis imperfecta. *Proc. Natl. Acad. Sci. USA* **87**, 7145–7149.

Bork P. & Beckmann G. (1993) The CUB domain. A widespread module in developmentally regulated proteins. *J. Mol. Biol.* **231**, 539–545.

Bourhis J.M., Mariano N., Zhao Y. *et al.* (2012) Structural basis of fibrillar collagen trimerization and related genetic disorders. *Nat. Struct. Mol. Biol.* **19**, 1031–1036.

Brodsky B. & Persikov A.V. (2005) Molecular structure of the collagen triple helix. *Adv. Protein Chem.* **70**, 301–339.

Brodsky B. & Ramshaw J.A. (1997) The collagen triple-helix structure. *Matrix Biol.* **15**, 545–554.

Brodsky B. & Shah N.K. (1995) Protein motifs. 8. The triple-helix motif in proteins. *FASEB J.* **9**, 1537–1546.

Bruns R.R., Hulmes D.J., Therrien S.F. & Gross J. (1979) Procollagen segment-long-spacing crystallites: their role in collagen fibrillogenesis. *Proc. Natl. Acad. Sci. USA* **76**, 313–317.

Canty E.G., Lu Y., Meadows R.S., Shaw M.K., Holmes D.F. & Kadler K.E. (2004) Coalignment of plasma membrane channels and protrusions (fibripositors) specifies the parallelism of tendon. *J. Cell Biol.* **165**, 553–563.

Canty E.G., Garrigue-Antar L. & Kadler K.E. (2006a) A complete domain structure of *Drosophila* tolloid is required for cleavage of short gastrulation. *J. Biol. Chem.* **281**, 13258–13267.

Canty E.G., Starborg T., Lu Y. *et al.* (2006b) Actin filaments are required for fibripositor-mediated collagen fibril alignment in tendon. *J. Biol. Chem.* **281**, 38592–38598.

Chakravarti S., Magnuson T., Lass J.H., Jepsen K.J., LaMantia C. & Carroll H. (1998) Lumican regulates collagen fibril assembly: skin fragility and corneal opacity in the absence of lumican. *J. Cell Biol.* **141**, 1277–1286.

Chapman J.A. (1989) The regulation of size and form in the assembly of collagen fibrils in vivo. *Biopolymers* **28**, 1367–1382.

Chen F., Guo R., Itoh S. *et al.* (2014) First mouse model for combined osteogenesis imperfecta and Ehlers-Danlos syndrome. *J. Bone Miner. Res.* **29**, 1412–1423.

Chipman S.D., Sweet H.O., McBride D.J. Jr *et al.* (1993) Defective pro alpha 2(I) collagen synthesis in a recessive mutation in mice: a model of human osteogenesis imperfecta. *Proc. Natl. Acad. Sci. USA* **90**, 1701–1705.

Craig A.S., Birtles M.J., Conway J.F. & Parry D.A. (1989) An estimate of the mean length of collagen fibrils in rat tail-tendon as a function of age. *Connect. Tissue Res.* **19**, 51–62.

Culbert A.A., Lowe M.P., Atkinson M., Byers P.H., Wallis G.A. & Kadler K.E. (1995) Substitutions of aspartic acid for glycine-220 and of arginine for glycine-664 in the triple helix of the pro alpha 1(I) chain of type I procollagen produce lethal osteogenesis imperfecta and disrupt the ability of collagen fibrils to incorporate crystalline hydroxyapatite. *Biochem. J.* **311**(Pt 3), 815–820.

Culbert A.A., Wallis G.A. & Kadler K.E. (1996) Tracing the pathway between mutation and phenotype in osteogenesis imperfecta: isolation of mineralization-specific genes. *Am. J. Med. Genet.* **63**, 167–174.

Daley E., Streeten E.A., Sorkin J.D. *et al.* (2010) Variable bone fragility associated with an Amish COL1A2 variant and a knock-in mouse model. *J. Bone Miner. Res.* **25**, 247–261.

Danielson K.G., Baribault H., Holmes D.F., Graham H., Kadler K.E. & Iozzo R.V. (1997) Targeted disruption of decorin leads to abnormal collagen fibril morphology and skin fragility. *J. Cell Biol.* **136**, 729–743.

Denk W. & Horstmann H. (2004) Serial block-face scanning electron microscopy to reconstruct three-dimensional tissue nanostructure. *PLoS Biol.* **2**, e329.

Di Lullo G.A., Sweeney S.M., Korkko J., Ala-Kokko L. & San Antonio J.D. (2002) Mapping the ligand-binding sites and disease-associated mutations on the most abundant protein in the human, type I collagen. *J. Biol. Chem.* **277**, 4223–4231.

Donahue L.R., Chang B., Mohan S. *et al.* (2003) A missense mutation in the mouse Col2a1 gene causes spondyloepiphyseal dysplasia congenita, hearing loss, and retinoschisis. *J. Bone Miner. Res.* **18**, 1612–1621.

Dudek M., Gossan N., Yang N. *et al.* (2016a) The chondrocyte clock gene Bmal1 controls cartilage homeostasis and integrity. *J. Clin. Invest.* **126**, 365–376.

Dudek M., Yang N., Ruckshanthi J.P. *et al.* (2016b) The intervertebral disc contains intrinsic circadian clocks that are regulated by age and cytokines and linked to degeneration. *Ann. Rheum. Dis.* doi: 10.1136/annrheumdis-2016-209428. [Epub ahead of print]

Ehrlich M.G., Mankin H.J., Jones H., Wright R., Crispen C. & Vigliani G. (1977) Collagenase and collagenase inhibitors in osteoarthritic and normal cartilage. *J. Clin. Invest.* **59**, 226–233.

Exposito J.Y., Valcourt U., Cluzel C. & Lethias C. (2010) The fibrillar collagen family. *Int. J. Mol. Sci.* **11**, 407–426.

Eyre D.R., Paz M.A. & Gallop P.M. (1984) Cross-linking in collagen and elastin. *Annu. Rev. Biochem.* **53**, 717–748.

Eyre D.R., Weis M.A. & Wu J.J. (2008) Advances in collagen cross-link analysis. *Methods* **45**, 65–74.

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

Fertala A., Holmes D.F., Kadler K.E., Sieron A.L. & Prockop D.J. (1996) Assembly in vitro of thin and thick fibrils of collagen II from recombinant procollagen II. The monomers in the tips of thick fibrils have the opposite orientation from monomers in the growing tips of collagen I fibrils. *J. Biol. Chem.* **271**, 14864–14869.

Gaiser K.G., Maddox B.K., Bann J.G. *et al.* (2002) Y-position collagen II mutation disrupts cartilage formation and skeletal development in a transgenic mouse model of spondyloepiphyseal dysplasia. *J. Bone Miner. Res.* **17**, 39–47.

Garofalo S., Vuorio E., Metsaranta M. *et al.* (1991) Reduced amounts of cartilage collagen fibrils and growth plate anomalies in transgenic mice harboring a glycine-to-cysteine mutation in the mouse type II procollagen alpha 1-chain gene. *Proc. Natl. Acad. Sci. USA* **88**, 9648–9652.

Garred P., Genster N., Pilely K. *et al.* (2016) A journey through the lectin pathway of complement-MBL and beyond. *Immunol. Rev.* **274**, 74–97.

Garrigue-Antar L., Barker C. & Kadler K.E. (2001) Identification of amino acid residues in bone morphogenetic protein-1 important for procollagen C-proteinase activity. *J. Biol. Chem.* **276**, 26237–26242.

Garrigue-Antar L., Hartigan N. & Kadler K.E. (2002) Post-translational modification of bone morphogenetic protein-1 is required for secretion and stability of the protein. *J. Biol. Chem.* **277**, 43327–43334.

Garrigue-Antar L., Francois V. & Kadler K.E. (2004) Deletion of epidermal growth factor-like domains converts mammalian tolloid into a chordinase and effective procollagen C-proteinase. *J. Biol. Chem.* **279**, 49835–49841.

Ghosh N., McKillop T.J., Jowitt T.A. *et al.* (2012) Collagen-like proteins in pathogenic *E. coli* strains. *PLoS One* **7**, e37872.

Gistelinck C., Witten P.E., Huysseune A. *et al.* (2016) Loss of type I collagen telopeptide lysyl hydroxylation causes musculoskeletal abnormalities in a zebrafish model of Bruck syndrome. *J. Bone Miner. Res.* **31**, 1930–1942.

Gonzaga-Jauregui C., Gamble C.N., Yuan B. *et al.* (2015) Mutations in COL27A1 cause Steel syndrome and suggest a founder mutation effect in the Puerto Rican population. *Eur. J. Hum. Genet.* **23**, 342–346.

Graham H.K., Holmes D.F., Watson R.B. & Kadler K.E. (2000) Identification of collagen fibril fusion during vertebrate tendon morphogenesis. The process relies on unipolar fibrils and is regulated by collagen-proteoglycan interaction. *J. Mol. Biol.* **295**, 891–902.

Gross J. & Kirk D. (1958) The heat precipitation of collagen from neutral salt solutions: some rate-regulating factors. *J. Biol. Chem.* **233**, 355–360.

Gross J., Highberger J.H. & Schmitt F.O. (1954) Collagen structures considered as states of aggregation of a kinetic unit. The tropocollagen particle. *Proc. Natl. Acad. Sci. USA* **40**, 679–688.

Guo B., Yang N., Borysiewicz E. *et al.* (2015) Catabolic cytokines disrupt the circadian clock and the expression of clock-controlled genes in cartilage via an NFκB-dependent pathway. *Osteoarthritis Cartilage* **23**, 1981–1988.

Hansen U. & Bruckner P. (2003) Macromolecular specificity of collagen fibrillogenesis: fibrils of collagens I and XI contain a heterotypic alloyed core and a collagen I sheath. *J. Biol. Chem.* **278**, 37352–37359.

Hartigan N., Garrigue-Antar L. & Kadler K.E. (2003) Bone morphogenetic protein-1 (BMP-1). Identification of the minimal domain structure for procollagen C-proteinase activity. *J. Biol. Chem.* **278**, 18045–18049.

Hedlund H., Mengarelli-Widholm S., Heinegard D., Reinholt F.P. & Svensson O. (1994) Fibromodulin distribution and association with collagen. *Matrix Biol.* **14**, 227–232.

Heegaard A.M., Corsi A., Danielsen C.C. *et al.* (2007) Biglycan deficiency causes spontaneous aortic dissection and rupture in mice. *Circulation* **115**, 2731–2738.

Heinemeier K.M., Schjerling P., Heinemeier J., Magnusson S.P. & Kjaer M. (2013) Lack of tissue renewal in human adult Achilles tendon is revealed by nuclear bomb (14)C. *FASEB J.* **27**, 2074–2079.

Heinemeier K.M., Schjerling P., Heinemeier J. *et al.* (2016) Radiocarbon dating reveals minimal collagen turnover in both healthy and osteoarthritic human cartilage. *Sci. Transl. Med.* **8**, 346ra390.

Herchenhan A., Uhlenbrock F., Eliasson P. *et al.* (2015) Lysyl oxidase activity is required for ordered collagen fibrillogenesis by tendon cells. *J. Biol. Chem.* **290**, 16440–16450.

Hodge A.J. (1989) Molecular models illustrating the possible distributions of 'holes' in simple systematically staggered arrays of type I collagen molecules in native-type fibrils. *Connect. Tissue Res.* **21**, 137–147.

Hojima Y., van der Rest M. & Prockop D.J. (1985) Type I procollagen carboxyl-terminal proteinase from chick embryo tendons. Purification and characterization. *J. Biol. Chem.* **260**, 15996–16003.

Hojima Y., McKenzie J.A., van der Rest M. & Prockop D.J. (1989) Type I procollagen N-proteinase from chick embryo tendons. Purification of a new 500-kDa form of the enzyme and identification of the catalytically active polypeptides. *J. Biol. Chem.* **264**, 11336–11345.

Holmes D.F. & Chapman J.A. (1979) Axial mass distributions of collagen fibrils grown in vitro: results for the end regions of early fibrils. *Biochem. Biophys. Res. Commun.* **87**, 993–999.

Holmes D.F., Chapman J.A., Prockop D.J. & Kadler K.E. (1992) Growing tips of type I collagen fibrils formed in vitro are near-paraboloidal in shape, implying a reciprocal relationship between accretion and diameter. *Proc. Natl. Acad. Sci. USA* **89**, 9855–9859.

Holmes D.F., Watson R.B., Steinmann B. & Kadler K.E. (1993) Ehlers-Danlos syndrome type VIIB. Morphology of type I collagen fibrils formed in vivo and in vitro is determined by the conformation of the retained N-propeptide. *J. Biol. Chem.* **268**, 15758–15765.

Holmes D.F., Lowe M.P. & Chapman J.A. (1994) Vertebrate (chick) collagen fibrils formed in vivo can exhibit a reversal in molecular polarity. *J. Mol. Biol.* **235**, 80–83.

Holmes D.F., Watson R.B., Chapman J.A. & Kadler K.E. (1996) Enzymic control of collagen fibril shape. *J. Mol. Biol.* **261**, 93–97.

Holmes D.F., Graham H.K. & Kadler K.E. (1998) Collagen fibrils forming in developing tendon show an early and abrupt limitation in diameter at the growing tips. *J. Mol. Biol.* **283**, 1049–1058.

Huang G., Ge G., Wang D. *et al.* (2011) alpha3(V) collagen is critical for glucose homeostasis in mice due to effects in pancreatic islets and peripheral tissues. *J. Clin. Invest.* **121**, 769–783.

Hudson B.G., Tryggvason K., Sundaramoorthy M. & Neilson E.G. (2003) Alport's syndrome, Goodpasture's syndrome, and type IV collagen. *N. Engl. J. Med.* **348**, 2543–2556.

Hulmes D.J., Miller A., Parry D.A., Piez K.A. & Woodhead-Galloway J. (1973) Analysis of the primary structure of collagen for the origins of molecular packing. *J. Mol. Biol.* **79**, 137–148.

Hulmes D.J., Bruns R.R. & Gross J. (1983) On the state of aggregation of newly secreted procollagen. *Proc. Natl. Acad. Sci. USA* **80**, 388–392.

Humphries S.M., Lu Y., Canty E.G. & Kadler K.E. (2008) Active negative control of collagen fibrillogenesis in vivo. Intracellular cleavage of the type I procollagen propeptides in tendon fibroblasts without intracellular fibrils. *J. Biol. Chem.* **283**, 12129–12135.

Huxley-Jones J., Robertson D.L. & Boot-Handford R.P. (2007) On the origins of the extracellular matrix in vertebrates. *Matrix Biol.* **26**, 2–11.

Johnson C.D., Smith S.P. & Russell R.L. (1977) Electrophorus electricus acetylcholinesterases; separation and selective modification by collagenase. *J. Neurochem.* **28**, 617–624.

Johnstone I.L. (2000) Cuticle collagen genes. Expression in *Caenorhabditis elegans*. *Trends Genet.* **16**, 21–27.

Jokinen J., Dadu E., Nykvist P. *et al.* (2004) Integrin-mediated cell adhesion to type I collagen fibrils. *J. Biol. Chem.* **279**, 31956–31963.

Kadler K.E., Hojima Y. & Prockop D.J. (1987) Assembly of collagen fibrils de novo by cleavage of the type I pC-collagen with procollagen C-proteinase. Assay of critical concentration demonstrates that collagen self-assembly is a classical example of an entropy-driven process. *J. Biol. Chem.* **262**, 15696–15701.

Kadler K.E., Hojima Y. & Prockop D.J. (1988) Assembly of type I collagen fibrils de novo. Between 37 and 41 degrees C the process is limited by micro-unfolding of monomers. *J. Biol. Chem.* **263**, 10517–10523.

Kadler K.E., Hojima Y. & Prockop D.J. (1990) Collagen fibrils in vitro grow from pointed tips in the C- to N-terminal direction. *Biochem. J.* **268**, 339–343.

Kadler K.E., Torre-Blanco A., Adachi E., Vogel B.E., Hojima Y. & Prockop D.J. (1991) A type I collagen with substitution of a cysteine for glycine-748 in the alpha 1(I) chain copolymerizes with normal type I collagen and can generate fractallike structures. *Biochemistry* **30**, 5081–5088.

Kadler K.E., Holmes D.F., Graham H. & Starborg T. (2000) Tip-mediated fusion involving unipolar collagen fibrils accounts for rapid fibril elongation, the occurrence of fibrillar branched networks in skin and the paucity of collagen fibril ends in vertebrates. *Matrix Biol.* **19**, 359–365.

Kadler K.E., Baldock C., Bella J. & Boot-Handford R.P. (2007) Collagens at a glance. *J. Cell Sci.* **120**, 1955–1958.

Kalamajski S. & Oldberg A. (2010) The role of small leucine-rich proteoglycans in collagen fibrillogenesis. *Matrix Biol.* **29**, 248–253.

Kalson N.S., Holmes D.F., Kapacee Z. *et al.* (2010) An experimental model for studying the biomechanics of embryonic tendon: evidence that the development of mechanical properties depends on the actinomyosin machinery. *Matrix Biol.* **29**, 678–689.

Kalson N.S., Holmes D.F., Herchenhan A., Lu Y., Starborg T. & Kadler K.E. (2011) Slow stretching that mimics embryonic growth rate stimulates structural and mechanical development of tendon-like tissue in vitro. *Dev. Dyn.* **240**, 2520–2528.

Kalson N.S., Starborg T., Lu Y. *et al.* (2013) Nonmuscle myosin II powered transport of newly formed collagen fibrils at the plasma membrane. *Proc. Natl. Acad. Sci. USA* **110**, E4743–E4752.

Kalson N.S., Lu Y., Taylor S.H., Starborg T., Holmes D.F. & Kadler K.E. (2015) A structure-based extracellular matrix expansion mechanism of fibrous tissue growth. *Elife* **4**, e05958.

Kapacee Z., Richardson S.H., Lu Y. *et al.* (2008) Tension is required for fibripositor formation. *Matrix Biol.* **27**, 371–375.

Kapacee Z., Yeung C.Y., Lu Y., Crabtree D., Holmes D.F. & Kadler K.E. (2010) Synthesis of embryonic tendon-like tissue by human marrow stromal/mesenchymal stem cells requires a three-dimensional environment and transforming growth factor beta3. *Matrix Biol.* **29**, 668–677.

Koch M., Laub F., Zhou P. *et al.* (2003) Collagen XXIV, a vertebrate fibrillar collagen with structural features of invertebrate collagens: selective expression in developing cornea and bone. *J. Biol. Chem.* **278**, 43236–43244.

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

Lees J.F., Tasab M. & Bulleid N.J. (1997) Identification of the molecular recognition sequence which determines the type-specific assembly of procollagen. *EMBO J.* **16**, 908–916.

Legendre M., Santini S., Rico A., Abergel C. & Claverie J.M. (2011) Breaking the 1000-gene barrier for Mimivirus using ultra-deep genome and transcriptome sequencing. *Virol. J.* **8**, 99.

Li S.W., Prockop D.J., Helminen H. *et al.* (1995a) Transgenic mice with targeted inactivation of the Col2 alpha 1 gene for collagen II develop a skeleton with membranous and periosteal bone but no endochondral bone. *Genes Dev.* **9**, 2821–2830.

Li Y., Lacerda D.A., Warman M.L. *et al.* (1995b) A fibrillar collagen gene, Col11a1, is essential for skeletal morphogenesis. *Cell* **80**, 423–430.

Li S.W., Takanosu M., Arita M. *et al.* (2001) Targeted disruption of Col11a2 produces a mild cartilage phenotype in transgenic mice: comparison with the human disorder otospondylomegaepiphyseal dysplasia (OSMED). *Dev. Dyn.* **222**, 141–152.

Lightfoot S.J., Holmes D.F., Brass A., Grant M.E., Byers P.H. & Kadler K.E. (1992) Type I procollagens containing substitutions of aspartate, arginine, and cysteine for glycine in the pro alpha 1 (I) chain are cleaved slowly by N-proteinase, but only the cysteine substitution introduces a kink in the molecule. *J. Biol. Chem.* **267**, 25521–25528.

Liu X., Wu H., Byrne M., Krane S. & Jaenisch R. (1997) Type III collagen is crucial for collagen I fibrillogenesis and for normal cardiovascular development. *Proc. Natl. Acad. Sci. USA* **94**, 1852–1856.

Liu C.Y., Birk D.E., Hassell J.R., Kane B. & Kao W.W. (2003) Keratocan-deficient mice display alterations in corneal structure. *J. Biol. Chem.* **278**, 21672–21677.

Long K.B., Li Z., Burgwin C.M. *et al.* (2015) The Tsk2/+ mouse fibrotic phenotype is due to a gain-of-function mutation in the PIIINP segment of the Col3a1 gene. *J. Invest. Dermatol.* **135**, 718–727.

Maass T., Bayley C.P., Morgelin M. *et al.* (2016) Heterogeneity of Collagen VI Microfibrils: structural analysis of non-collagenous regions. *J. Biol. Chem.* **291**, 5247–5258.

Maki J.M., Rasanen J., Tikkanen H. *et al.* (2002) Inactivation of the lysyl oxidase gene Lox leads to aortic aneurysms, cardiovascular dysfunction, and perinatal death in mice. *Circulation* **106**, 2503–2509.

Mallory F.B. (1903) A hitherto undescribed fibrillar substance produced by connective-tissue cells. *J. Med. Res.* **10**, 334–341.

Martin R., Farjanel J., Eichenberger D. *et al.* (2000) Liquid crystalline ordering of procollagen as a determinant of three-dimensional extracellular matrix architecture. *J. Mol. Biol.* **301**, 11–17.

McGuirt W.T., Prasad S.D., Griffith A.J. *et al.* (1999) Mutations in COL11A2 cause non-syndromic hearing loss (DFNA13). *Nat. Genet.* **23**, 413–419.

Mienaltowski M.J. & Birk D.E. (2014) Structure, physiology, and biochemistry of collagens. *Adv. Exp. Med. Biol.* **802**, 5–29.

Mussini E., Hutton J.J. Jr & Udenfriend S. (1967) Collagen proline hydroxylase in wound healing, granuloma formation, scurvy, and growth. *Science* **157**, 927–929.

Myllyharju J. & Kivirikko K.I. (2004) Collagens, modifying enzymes and their mutations in humans, flies and worms. *Trends Genet.* **20**, 33–43.

Njieha F.K., Morikawa T., Tuderman L. & Prockop D.J. (1982) Partial purification of a procollagen C-proteinase. Inhibition by synthetic peptides and sequential cleavage of type I procollagen. *Biochemistry* **21**, 757–764.

Orgel J.P., San Antonio J.D. & Antipova O. (2011) Molecular and structural mapping of collagen fibril interactions. *Connect. Tissue Res.* **52**, 2–17.

Parry D.A., Barnes G.R. & Craig A.S. (1978) A comparison of the size distribution of collagen fibrils in connective tissues as a function of age and a possible relation between fibril size distribution and mechanical properties. *Proc. R. Soc. Lond. B Biol. Sci.* **203**, 305–321.

Petropoulou V., Garrigue-Antar L. & Kadler K.E. (2005) Identification of the minimal domain structure of bone morphogenetic protein-1 (BMP-1) for chordinase activity: chordinase activity is not enhanced by procollagen C-proteinase enhancer-1 (PCPE-1). *J. Biol. Chem.* **280**, 22616–22623.

Pingel J., Lu Y., Starborg T. *et al.* (2014) 3-D ultrastructure and collagen composition of healthy and overloaded human tendon: evidence of tenocyte and matrix buckling. *J. Anat.* **224**, 548–555.

Plumb D.A., Dhir V., Mironov A. *et al.* (2007) Collagen XXVII is developmentally regulated and forms thin fibrillar structures distinct from those of classical vertebrate fibrillar collagens. *J. Biol. Chem.* **282**, 12791–12795.

Plumb D.A., Ferrara L., Torbica T. *et al.* (2011) Collagen XXVII organises the pericellular matrix in the growth plate. *PLoS One* **6**, e29422.

Rasmussen M., Jacobsson M. & Bjorck L. (2003) Genome-based identification and analysis of collagen-related structural motifs in bacterial and viral proteins. *J. Biol. Chem.* **278**, 32313–32316.

Reid K.B. & Day A.J. (1990) Ig-binding domains of C1q. *Immunol. Today* **11**, 387–388.

Richardson S.H., Starborg T., Lu Y., Humphries S.M., Meadows R.S. & Kadler K.E. (2007) Tendon development requires regulation of cell condensation and cell shape via cadherin-11-mediated cell-cell junctions. *Mol. Cell. Biol.* **27**, 6218–6228.

Roberts A.B., Sporn M.B., Assoian R.K. *et al.* (1986) Transforming growth factor type beta: rapid induction of fibrosis and angiogenesis in vivo and stimulation of collagen formation in vitro. *Proc. Natl. Acad. Sci. USA* **83**, 4167–4171.

Saito K., Chen M., Bard F. *et al.* (2009) TANGO1 facilitates cargo loading at endoplasmic reticulum exit sites. *Cell* **136**, 891–902.

Satoh M., Hirayoshi K., Yokota S., Hosokawa N. & Nagata K. (1996) Intracellular interaction of collagen-specific stress protein HSP47 with newly synthesized procollagen. *J. Cell Biol.* **133**, 469–483.

Silver F.H. & Trelstad R.L. (1980) Type I collagen in solution. Structure and properties of fibril fragments. *J. Biol. Chem.* **255**, 9427–9433.

Starborg T., Kalson N.S., Lu Y. *et al.* (2013) Using transmission electron microscopy and 3View to determine collagen fibril size and three-dimensional organization. *Nat. Protoc.* **8**, 1433–1448.

Stearns M.L. (1940) Studies on the development of connective tissue in transparent chambers in the rabbit's ear. II. *Am. J. Anat.* **67**, 55–97.

Svensson L., Aszodi A., Reinholt F.P., Fassler R., Heinegard D. & Oldberg A. (1999) Fibromodulin-null mice have abnormal collagen fibrils, tissue organization, and altered lumican deposition in tendon. *J. Biol. Chem.* **274**, 9636–9647.

Sweeney S.M., Orgel J.P., Fertala A. *et al.* (2008) Candidate cell and matrix interaction domains on the collagen fibril, the predominant protein of vertebrates. *J. Biol. Chem.* **283**, 21187–21197.

Takaluoma K., Hyry M., Lantto J. *et al.* (2007) Tissue-specific changes in the hydroxylysine content and cross-links of collagens and alterations in fibril morphology in lysyl hydroxylase 1 knock-out mice. *J. Biol. Chem.* **282**, 6588–6596.

Tasheva E.S., Koester A., Paulsen A.Q. *et al.* (2002) Mimecan/osteoglycin-deficient mice have collagen fibril abnormalities. *Mol. Vis.* **8**, 407–415.

Taylor S.H., Yeung C.Y., Kalson N.S. *et al.* (2015) Matrix metalloproteinase 14 is required for fibrous tissue expansion. *Elife* **4**, e09345.

Thorpe C.T., Streeter I., Pinchbeck G.L., Goodship A.E., Clegg P.D. & Birch H.L. (2010) Aspartic acid racemization and collagen degradation markers reveal an accumulation of damage in tendon collagen that is enhanced with aging. *J. Biol. Chem.* **285**, 15674–15681.

Thurmond F.A. & Trotter J.A. (1994) Native collagen fibrils from echinoderms are molecularly bipolar. *J. Mol. Biol.* **235**, 73–79.

Trelstad R.L. & Birk D.E. (1985) The fibroblast in morphogenesis and fibrosis: cell topography and surface-related functions. *Ciba Found. Symp.* **114**, 4–19.

Trelstad R.L. & Hayashi K. (1979) Tendon collagen fibrillogenesis: intracellular subassemblies and cell surface changes associated with fibril growth. *Dev. Biol.* **71**, 228–242.

Trotter J.A. & Koob T.J. (1989) Collagen and proteoglycan in a sea urchin ligament with mutable mechanical properties. *Cell Tissue Res.* **258**, 527–539.

Trotter J.A., Chapman J.A., Kadler K.E. & Holmes D.F. (1998) Growth of sea cucumber collagen fibrils occurs at the tips and centers in a coordinated manner. *J. Mol. Biol.* **284**, 1417–1424.

Trotter J.A., Kadler K.E. & Holmes D.F. (2000) Echinoderm collagen fibrils grow by surface-nucleation-and-propagation from both centers and ends. *J. Mol. Biol.* **300**, 531–540.

Tuderman L. & Prockop D.J. (1982) Procollagen N-proteinase. Properties of the enzyme purified from chick embryo tendons. *Eur. J. Biochem.* **125**, 545–549.

Vandenberg P., Khillan J.S., Prockop D.J., Helminen H., Kontusaari S. & Ala-Kokko L. (1991) Expression of a partially deleted gene of human type II procollagen (COL2A1) in transgenic mice produces a chondrodysplasia. *Proc. Natl. Acad. Sci. USA* **88**, 7640–7644.

Venditti R., Scanu T., Santoro M. *et al.* (2012) Sedlin controls the ER export of procollagen by regulating the Sar1 cycle. *Science* **337**, 1668–1672.

Verzijl N., DeGroot J., Thorpe S.R. *et al.* (2000) Effect of collagen turnover on the accumulation of advanced glycation end products. *J. Biol. Chem.* **275**, 39027–39031.

Vogel B.E., Doelz R., Kadler K.E., Hojima Y., Engel J. & Prockop D.J. (1988) A substitution of cysteine for glycine 748 of the alpha 1 chain produces a kink at this site in the procollagen I molecule and an altered N-proteinase cleavage site over 225 nm away. *J. Biol. Chem.* **263**, 19249–19255.

Wallis G.A., Kadler K.E., Starman B.J. & Byers P.H. (1992) A tripeptide deletion in the triple-helical domain of the pro alpha 1 (I) chain of type I procollagen in a patient with lethal osteogenesis imperfecta does not alter cleavage of the molecule by N-proteinase. *J. Biol. Chem.* **267**, 25529–25534.

Watson R.B., Wallis G.A., Holmes D.F., Viljoen D., Byers P.H. & Kadler K.E. (1992) Ehlers Danlos syndrome type VIIB. Incomplete cleavage of abnormal type I procollagen by N-proteinase in vitro results in the formation of copolymers of collagen and partially cleaved pNcollagen that are near circular in cross-section. *J. Biol. Chem.* **267**, 9093–9100.

Watson R.B., Holmes D.F., Graham H.K., Nusgens B.V. & Kadler K.E. (1998) Surface located procollagen N-propeptides on dermatosparactic collagen fibrils are not cleaved by procollagen N-proteinase and do not inhibit binding of decorin to the fibril surface. *J. Mol. Biol.* **278**, 195–204.

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**, 53331–53337.

Wilson D.G., Phamluong K., Li L. *et al.* (2011) Global defects in collagen secretion in a Mia3/TANGO1 knockout mouse. *J. Cell Biol.* **193**, 935–951.

Wood G.C. & Keech M.K. (1960) The formation of fibrils from collagen solutions. 1. The effect of experimental conditions: kinetic and electron-microscope studies. *Biochem. J.* **75**, 588–598.

Yeung C.Y., Gossan N., Lu Y. *et al.* (2014) Gremlin-2 is a BMP antagonist that is regulated by the circadian clock. *Sci. Rep.* **4**, 5183.

Young B.B., Zhang G., Koch M. & Birk D.E. (2002) The roles of types XII and XIV collagen in fibrillogenesis and matrix assembly in the developing cornea. *J. Cell. Biochem.* **87**, 208–220.

Young R.D., Knupp C., Pinali C. *et al.* (2014) Three-dimensional aspects of matrix assembly by cells in the developing cornea. *Proc. Natl. Acad. Sci. USA* **111**, 687–692.

Supporting information

Additional Supporting Information may be found online in the supporting information tab for this article:

Movie S1. Step-through movie of consecutive images of embryonic mouse tendon generated by serial block face-scanning electron microscopy.