

Published in final edited form as:

*Essays Biochem.* 2012 ; 52: 113–133. doi:10.1042/bse0520113.

## Lysine post-translational modifications of collagen

Mitsuo Yamauchi<sup>1</sup> and Marnisa Sricholpech

NC Oral Health Institute, University of North Carolina at Chapel Hill, NC 27599, U.S.A.

### Abstract

Type I collagen is the most abundant structural protein in vertebrates. It is a heterotrimeric molecule composed of two  $\alpha 1$  chains and one  $\alpha 2$  chain, forming a long uninterrupted triple helical structure with short non-triple helical telopeptides at both the N- and C-termini. During biosynthesis, collagen acquires a number of post-translational modifications, including lysine modifications, that are critical to the structure and biological functions of this protein. Lysine modifications of collagen are highly complicated sequential processes catalysed by several groups of enzymes leading to the final step of biosynthesis, covalent intermolecular cross-linking. In the cell, specific lysine residues are hydroxylated to form hydroxylysine. Then specific hydroxylysine residues located in the helical domain of the molecule are glycosylated by the addition of galactose or glucose-galactose. Outside the cell, lysine and hydroxylysine residues in the N- and C-telopeptides can be oxidatively deaminated to produce reactive aldehydes that undergo a series of non-enzymatic condensation reactions to form covalent intra- and inter-molecular cross-links. Owing to the recent advances in molecular and cellular biology, and analytical technologies, the biological significance and molecular mechanisms of these modifications have been gradually elucidated. This chapter provides an overview on these enzymatic lysine modifications and subsequent cross-linking.

### Introduction

Collagens comprise a large family of triple helical proteins and are the most abundant protein in vertebrates, representing ~30% of the total proteins. There are now at least 29 genetically distinct types of collagen identified that are encoded by at least 44 genes [1]. Depending on the molecular structure and assembly mode, they can be divided into several subgroups [2]. Fibril-forming collagens comprise the largest subgroup, including types I, II, III, V, XI, XXIV and XXVII. Of all types of collagens, fibrillar type I collagen is the most abundant type providing most tissues and organs with form, stability and connectivity. In addition to the structural functions, type I and other collagens also function as ligands for specific cell receptors, such as integrins, discoidin domain receptors, glycoprotein VI and the mannose receptor family etc., to control cellular activities and extracellular matrix remodelling [3]. Studies have indicated that glycosylation of hydroxylysine residues of collagen plays a key role in the interaction with some of these receptors [4–6].

One of the critical factors for the structural and biomechanical functions of type I collagen fibrils are the PTMs (post-translational modifications) of peptidyl lysine residues. Lysine modifications of collagen are highly regulated sequential processes that take place inside and outside the cell. In the cell, specific peptidyl lysine residues both in the helical and non-helical (telopeptide) domains of the molecule (for the structure of type I collagen, see below) can be hydroxylated forming 5-hydroxylysine. Specific hydroxylysine residues in the helical

domain can then be glycosylated with the addition of galactose, some of which can be further glycosylated with the addition of glucose. Specific enzymes catalyse each of these sequential and domain-specific lysine modifications in the cell. Outside the cell, an enzymatic oxidative deamination occurs to the telopeptidyl lysine and hydroxylysine residues producing the reactive aldehydic residues. The aldehydes produced then initiate a series of non-enzymatic condensation reactions to form extensive covalent intra- and intermolecular cross-links which are critical for the biomechanical functions of the collagen fibrils [7]. Since type I collagen is the most well investigated collagen for its lysine modifications and they are probably shared by other fibrillar collagens, this chapter focuses on those modifications of type I collagen. Owing to the limited space, non-enzymatic glycosylation, known as glycation, of collagen is not dealt with in this chapter (for a detailed review of this, see [8]).

## Collagen biosynthesis

The biosynthesis of collagen is a long complicated process involving a number of PTMs, chain association and folding, secretion, procollagen processing, self-assembly and progressive cross-linking (Figure 1). Type I collagen is a long (~300 nm long, ~1.5 nm thick) heterotrimeric molecule composed of two  $\alpha 1$  chains and one  $\alpha 2$  chain. An  $\alpha 1$  homotrimeric form exists as a minor form. The molecule consists of three domains: the N-terminal non-triple helical domain (N-telopeptide), the central triple helical domain and the C-terminal non-triple helical domain (C-telopeptide). The single (uninterrupted) triple helical domain represents more than 95% of the molecule (for the molecular structure, see Figure 2).

After synthesis in the ER (endoplasmic reticulum), three pro $\alpha$  chains (two pro $\alpha 1$  and one pro $\alpha 2$ ) associate with one another near the C-terminus of the chains, and the triple helix is propagated towards the N-terminus in a zipper-like fashion to form a procollagen molecule. Several specific enzymes and chaperones are involved in chain association, stabilization and correct folding [2,9]. Before the formation of a triple helix procollagen molecule, however, the nascent pro $\alpha$  chains acquire several PTMs. They include the hydroxylation of specific proline residues, which are catalysed by prolyl-4-hydroxylases and prolyl-3-hydroxylase. The former reacts on proline with the minimum sequence X-Pro-Gly and the latter appears to require a Pro-4-Hyp-Gly (Hyp is hydroxyproline) sequence. The proline hydroxylation is critical for the stabilization of the triple helix conformation. The hydroxylation of specific lysine residues and glycosylation of hydroxylysine residues also occur enzymatically (see below) before the formation of a triple helical procollagen molecule. The procollagen molecules are then transported from the ER to the Golgi stacks, and are secreted to the extracellular space. The N- and C-propeptides of procollagen are then cleaved off by a group of metalloproteinases belonging to the ADAMTS (a disintegrin and metalloproteinase with thrombospondin motifs) and BMP1 (bone morphogenetic protein 1)/Tolloid-like proteinases respectively, releasing a mature type I collagen molecule.

After both N- and C-propeptides are cleaved, the collagen molecules are then spontaneously packed in parallel and longitudinally staggered with respect to one another by some multiple of axial repeat distance, called a D period (~67 nm). This packing arrangement creates repeated zones of high and low packing density, i.e. overlap and hole zones respectively, showing a characteristic banding pattern of a collagen fibril seen at the ultrastructural level (Figure 1). The molecules in the fibril are then stabilized by extensive covalent intermolecular cross-linking, which is critical for the formation of mechanically functional collagen fibrils. For the details of fibrillar collagen biosynthesis, see [2].

## Intracellular lysine modifications of collagen

### Hydroxylation of lysine

In human type I collagen, there are 38 lysine residues in an  $\alpha 1$  chain (36 in the helical domain, one in the C-telopeptide domain and one in the N-telopeptide domain) and 31 in an  $\alpha 2$  chain (30 in the helical domain, one in the N-telopeptide domain, and none in the C-telopeptide domain) (UniProt numbers P02452, P02464 and P08123) (Figure 3A). The numbers and sites of lysine residues in the  $\alpha 1$  chain are conserved among different species, e.g. human, mouse, rat and bovine (UniProt numbers P02452, P11087, P02454 and P02453 respectively); however, there are slight differences in those in the  $\alpha 2$  chain. There is one extra lysine residue in the helical region in mouse and rat (total lysine residues 32, 31 in the helical domain and one in N-telopeptide domain) and there are three sites that vary among these four species (UniProt numbers P08123, Q01149, P02466 and P02465 respectively). The lysine residues in both  $\alpha 1$  and  $\alpha 2$  chains that are involved in the formation of covalent cross-linking (telopeptide and helical domains, see below) are all conserved among species. Biochemical analysis has indicated that the extent of lysine hydroxylation of collagen is highly variable in comparison with proline hydroxylation. Approximately 50% of proline residues in collagen are hydroxylated in various genetic types of collagen in different tissues, and it does not significantly change under physiological conditions. The extent of lysine hydroxylation, however, can vary from 15 to 90% depending on the collagen types [10] and, even within type I collagen, it varies significantly from tissue to tissue and under the physiological/pathological condition of the tissue [11]. Furthermore, a difference in the extent of lysine hydroxylation exists between the helical and telopeptide domains of a type I collagen molecule [11–14]. These observations led investigators to speculate that there is more than one mechanism for lysine hydroxylation of collagen [12,15].

LH (lysyl hydroxylase; EC1.14.11.4) catalyses the formation of hydroxylysine in -X-Lys-Gly- sequences in a reaction that requires  $\text{Fe}^{2+}$ , 2-oxoglutarate,  $\text{O}_2$  and ascorbate [16]. In addition to the -X-Lys-Gly- sequence, -X-Lys-Ala- and -X-Lys-Ser- sequences present in the telopeptides (both N- and C-termini) of fibrillar collagen can be hydroxylated. However, for a long time, it was not clear whether one LH hydroxylates lysine residues in both helical and telopeptide domains of the molecule. In the late 1990s, three genes encoding isoforms of LH (procollagen-lysine, 2-oxoglutarate and 5-dioxygenase, *PLOD1–PLOD3*/LH1–3) were identified and partially characterized [17,18]. The substrate specificities of these isoforms *in vivo* are still not clearly established; however, evidence indicates that LH1 primarily hydroxylates lysine residues in the helical domains of fibrillar/non-fibrillar collagens [9]. For LH2, two alternatively spliced isoforms were identified, termed LH2a or LH2(short) and LH2b or LH2(long) respectively. The latter includes an additional 21 amino acids encoded by the 63 nucleotide exon 13A and the former does not. The LH2b appears to be ubiquitously expressed and is the major form of LH2 in most tissues [19]. Our initial observation that LH2 gene expression was closely associated with lysine hydroxylation in the telopeptide domains of a type I collagen molecule led us to propose LH2 (LH2b) as a telopeptidyl LH [14]. Several subsequent reports confirmed that LH2b is indeed the telopeptidyl LH and that forced overexpression of the LH2b gene in cells results in an increase in the  $\text{Hyl}^{\text{ald}}$  ( $\delta$ -hydroxy,  $\alpha$ -amino adipic acid- $\delta$ -semialdehyde)-derived collagen cross-links [20–23] (for collagen cross-linking, see below). LH3 is a multifunctional enzyme possessing LH, GT (hydroxylysyl galactosyltransferase) and GGT (galactosylhydroxylysyl glucosyltransferase) activities [24]. However, its LH activity may be significant for collagen types IV and V, but not type I collagen [25].

The biological significance of lysine hydroxylation/subsequent modifications cannot be overemphasized. In humans, mutations in the gene encoding LH1 cause Ehlers–Danlos syndrome type VIA and the clinically similar Nevo syndrome, characterized by hypermobile

joints, hyperextensible skin and kyphoscoliosis etc. [26,27]. Mutations in LH2 are known to cause Bruck syndrome type 2, a rare autosomal-recessive disorder characterized by bone fragility and congenital joint contractures [21,28]. Most recently, mutations in the LH3 gene and reduced LH3 protein have also been shown to cause several congenital connective tissue malformations and epidermolysis bullosa simplex [29,30]. In an osteoblastic cell-culture system, altered levels of LH2b cause defective cross-linking patterns, collagen fibrillogenesis and matrix mineralization [31], which also indicates the critical importance of telopeptidyl lysine hydroxylation of collagen in determining bone quality [32].

### Glycosylation of hydroxylysine

In collagens and collagen-like proteins, some of the hydroxylysine residues in the helical domain are further modified by the sequential steps of O-linked glycosylation producing G-Hyl (galactosylhydroxylysine) and GG-Hyl (glucosylgalactosylhydroxylysine). These reactions are catalysed by GT (EC 2.4.1.50) and GGT (EC 2.4.1.66) respectively (Figure 3B) [10,33]. Structural analysis of GG-Hyl shows that galactose is attached to the hydroxy group of hydroxy lysine by a  $\beta$ -glycosidic bond, whereas glucose is linked by an  $\alpha$ -glycosidic bond to C-2 of the galactose unit forming the complex structure of 2-O- $\alpha$ -D-glucopyranosyl-O- $\beta$ -D-galactopyranosylhydroxylysine (Figure 3B). Both reactions require the presence of a free  $\epsilon$ -amino group of the hydroxylysine residue, UDP-galactose or UDP-glucose as carbohydrate donors, and the most preferred bivalent cation cofactor  $Mn^{2+}$  [9]. The complete triple helical structure of a collagen molecule inhibits further glycosylation, indicating that the reactions take place on single  $\alpha$  chains prior to triple helix formation of procollagen ([34] and reviewed in [9]). The enzymatic activities of GT and GGT have previously been extensively studied and characterized from various tissue extracts (reviewed in [9]), but it was only approximately a decade ago, that it was found that they were modulated by the multifunctional enzyme LH3 [35,36]. By analysing LH3-deficient mice it was shown that impairment in type IV and VI collagen biosynthesis and secretion resulted in embryonic lethal phenotypes, in which the severity of the phenotype showed a significant correlation only to the level of GGT activity, not LH or GT activity [25,37,38]. It has also been reported that decreased GGT activity from mutations discovered in the human LH3 gene results in congenital disorders affecting connective tissues of various organs [29,30]. For bone type I collagen, most recently, we have shown that the major function of LH3 is GGT activity, but not LH or GT activity [34]. This is consistent with previous studies demonstrating that the GT activity of LH3, when compared with GGT, was significantly lower [35,36] or undetectable [37,39]. As for the collagen GT enzyme, a novel family of collagen galactosyltransferase enzymes, GLT25D1 and GLT25D2 (glycosyltransferase 25 domain 1 and glycosyltransferase 25 domain 2) which are encoded by the *GLT25D1* and *GLT25D2* genes, has been identified. Their tissue distribution and enzymatic activity, as GT, were reported, but the substrate specificity of each isoform is not clearly defined. Of the two isoforms, *GLT25D1* exhibited broad expression in several fetal and adult human tissues, whereas *GLT25D2* expression was detected only in a few cell types, suggesting the former to be the main GT enzyme [39]. We have also shown that *Glt25d1* is the major isoform expressed in the mouse osteoblast cell line [34]. Taken together, these recent studies indicate that, in type I collagen, galactosylation of hydroxylysine is catalysed by GLT25D1 and subsequent glucosylation by LH3 (Figure 3).

The type (mono- or di-) and extent of collagen glycosylation varies among different collagen types, showing a higher abundance in the less-organized collagen type IV of the basement membrane and lower levels in the highly organized fibrillar collagens, types I, II or III [10,33]. Furthermore, even those in type I collagen vary depending on the functional regions within the tissue [40], maturation [41,42] and pathological conditions [29,30,43–46]. In bone, for instance, alterations in the levels of collagen glycosylation have been reported in

several bone/skeletal disorders, such as osteogenesis imperfecta [44, 47–49], post-menopausal osteoporosis [43,50,51], osteosarcoma, osteofibrous dysplasia [45] and Kashin–Beck disease [52], suggesting a role of collagen glycosylation in bone mineralization. Considering the bulky structure of the sugars and the glycosylation sites, structural as well as biological functions have been proposed for collagen glycosylation, i.e. control of collagen fibrillogenesis [34,53–57], crosslinking [58–62], remodelling [55], collagen–cell interaction [4,6] and induction of vessel-like structures [63]. Despite all of the findings with regard to collagen glycosylation reported to date, the specific biological functions of the glycosylated hydroxylysine residues in collagen, in relation to the extent and type of glycosylation at their molecular loci, are still not clearly defined.

## Extracellular lysine modifications

### Oxidative deamination of lysine and hydroxylysine by LOX (lysyl oxidase)

After procollagen molecules were secreted into the extracellular space and both N- and C-propeptide extensions are cleaved off, collagen molecules then self-assemble into fibrils. Recent studies have suggested that procollagen processing and fibril assembly may begin within the confines of the cell membrane [64]. These cleavages, especially that of C-propeptide by BMP1/Tolloid-like proteinases, is necessary for the initiation of fibrillogenesis [65]. During fibrillogenesis, specific lysine and hydroxylysine residues in the N- and C-telopeptides are oxidatively deaminated by LOX to form the respective aldehydic forms, Lys<sup>ald</sup> ( $\alpha$ -amino adipic acid- $\delta$ -semialdehyde) and Hy<sup>lald</sup> (Figure 4), which initiates a series of condensation reactions to form covalent intra- and inter-molecular cross-links (see below). The reaction catalysed by LOX is  $RCH_2NH_2 + H_2O + O_2 \rightarrow RCH=O + NH_3 + H_2O_2$ , where the amine group is derived from telopeptidyl lysine and hydroxylysine. In type I collagen, there are two residues of lysine or hydroxylysine in the C-telopeptides (both from  $\alpha 1$  chains, i.e. there is no lysine residue in the  $\alpha 2$  C-telopeptide) and three in the N-telopeptides (a total of five residues/molecule) that can be oxidized by LOX (Figure 4). LOX contains two cofactors essential to its catalytic function. This enzyme contains one tightly bound copper ion ( $Cu^{2+}$ ) and a unique covalently integrated organic cofactor, LTQ (lysine tyrosylquinone). LTQ is probably the product of an autocatalytic reaction involving  $Cu^{2+}$ -mediated/assisted oxidation of a specific peptidyl tyrosine residue (residue 345 in rat LOX) followed by covalent cross-linking with the  $\epsilon$ -amino group of a conserved peptidyl lysine residue (residue 314 in rat LOX) [66]. It is interesting that this cross-linking enzyme requires its own covalent intramolecular cross-link, LTQ, to function. Apparently LOX cannot catalyse this reaction on lysine and hydroxylysine residues located in the rigid helical domain of the molecule.

LOX is a copper-dependent amine oxidase (EC 1.4.3.13, protein-lysine 6-oxidases) and it is synthesized as a catalytically quiescent N-glycosylated  $\sim 50$  kDa proLOX [67] by various cell types. ProLOX is processed extracellularly by BMP1/Tolloid-like proteinases, the same proteinases that cleave the C-propeptide of type I procollagen, to produce an active mature  $\sim 30$  kDa LOX. Recently, an additional processing of proLOX has been reported, resulting in a truncated form of active LOX, although its biological function is not clear at present [68]. The N-terminal propeptide of LOX has been implicated in regulating the localization of the enzyme, suppressing tumours and inhibiting cell proliferation [69–71]. It has been reported that type I collagen molecules need to be pre-assembled into quarter-staggered fibrillar arrays prior to oxidation of the telopeptidyl lysine and hydroxylysine by LOX [66,72]. Thus it is interesting to note that the same proteinases produce an active LOX enzyme and, at the same time, prepare for its collagen substrate. They also process other types of collagen, a number of extracellular matrix proteins and some of the growth factors, demonstrating their broad and key role in regulating and orchestrating the formation of the extracellular matrix [73].



The critical importance of LOX-mediated cross-linking in tissue stability is well exemplified by 'lathyrism'. This is the condition where deleterious effects in connective tissues are caused by lathyrogens such as BAPN ( $\beta$ -aminopropionitrile) [74], a potent and irreversible inhibitor of LOX, thus preventing collagen and elastin cross-linking. It has been proposed that the inhibition occurs by the formation of a 'dead-end' complex between BAPN and the carbonyl cofactor of LOX, LTQ (Figure 4) [75]. In addition to this classic example of specific inhibition of a biochemical pathway, gene-deletion studies have also provided significant insights into the critical role of LOX in connective tissue formation and maintenance. Mice lacking the *lox* gene die at the end of gestation or as neonates exhibiting the rupture of the aorta and diaphragm, and other connective tissues defects [76,77]. However, collagen cross-links, although diminished, were still formed in these mice, indicating some compensatory mechanisms by other LOX family members.

During the last two decades, four LOXL (LOX-like) proteins (LOXL1–4) have been identified, thus establishing a gene family of five members [78]. Structurally, LOX and LOXL1 form one subfamily, and LOXL2–4 another family. Although the N-terminal regions of the LOX family members are highly variable, all possess high similarity in the C-terminal catalytic domain, i.e. a putative  $\text{Cu}^{2+}$ -binding site, a CRL (cytokine receptor-like) domain and amino acid residues forming the LTQ site [79]. Previous studies have shown that LOX family members possess a wide range of biological functions other than collagen/elastin cross-linking, and they include tumour suppression or metastases promotion, control of growth-factor activity, actin filament formation and chemotaxis for various cell types etc. [80–85]. Although the substrate specificity of each LOXL isoform is not clear at present, studies have indicated their distinct functions. The direct involvement of these LOXL proteins in collagen cross-linking is also still not clearly demonstrated.

The enzymes that catalyse the reactions of lysine modifications of collagen described above, their characteristics and related human diseases/disorders are summarized in Table 1.

### Covalent intra- and inter-molecular cross-linking

The final step of type I collagen biosynthesis is the formation of covalent crosslinks, which is critical for the biomechanical functions of collagen. In this process, again, lysine modifications play a crucial role.

The telopeptidyl aldehydes generated by LOX react with another aldehyde or with the  $\epsilon$ -amine groups derived from the unmodified lysine and hydroxylysine residues to form the initial intra- and inter-molecular cross-links. They can then further mature into multivalent cross-links by involving additional amino acid residues including peptidyl histidine. The presence of such complex intermolecular cross-links demonstrates the highly specific molecular packing arrangement in the fibril, as the amino acid residues involved in cross-linking are relatively sparse (especially histidine) in collagen and they need to be in the correct proximity. Depending on the combination of amino acid residues, several cross-linking pathways have evolved (for details of LOX-mediated and non-enzymatic cross-linking in various connective tissues, see [8,32,86,87]). Figure 5 illustrates cross-linking pathways that are derived from  $\text{Lys}^{\text{ald}}$  or  $\text{Hyl}^{\text{ald}}$ .

One of the most well-investigated pathways is the one leading to a Pyr (pyridinoline) cross-link, which was first identified by Fujimoto [88]. It starts with the  $\text{Hyl}^{\text{ald}}$  residue in the telopeptides, i.e. the 16th residue in the  $\alpha 1$  C-telopeptide ( $\alpha 1-16^{\text{C}}$ ), and the 9th and 5th residue in the  $\alpha 1$  and  $\alpha 2$  N-telopeptide ( $\alpha 1-9^{\text{N}}$  and  $\alpha 2-5^{\text{N}}$ ) (Figure 5). Then the  $\text{Hyl}^{\text{ald}}$  reacts with the  $\epsilon$ -amino group of the helical hydroxylysine residue on a neighbouring molecule producing an intermolecular aldimine (Schiff base) cross-link, deH-DHLNL (dehydro-dihydroxylysinonorleucine). This can be formed between  $16^{\text{C}}\text{-Hyl}^{\text{ald}}$  ( $\alpha 1$ ) and the

juxtaposed helical hydroxylysine (Hyl<sup>87</sup>) of  $\alpha 1$  and  $\alpha 2$  chains, and between the 9<sup>N</sup> ( $\alpha 1$ )- or 5<sup>N</sup> ( $\alpha 2$ )-Hyl<sup>ald</sup> and the juxtaposed helical hydroxylysine ( $\alpha 1$ -930 or  $\alpha 2$ -933) (Figure 5). The aldimine can spontaneously rearrange to form a ketoamine form by Amadori rearrangement. However, some may remain as an aldimine form *in situ*. Both forms are reducible with reducing agents, such as sodium borohydride, under mild conditions. Then, two residues of deH-DHLNL/ketoamine can condense with each other forming a tri-functional cross-link, Pyr, by releasing a helical hydroxylysine residue. Thus Pyr is composed of two residues of telopeptidyl Hyl<sup>ald</sup> and one helical hydroxylysine residue. Although several mechanisms were proposed for Pyr formation [87,89–91] and the details are still not elucidated, Pyr is clearly a maturational product of deH-DHLNL tying together three [86] or two collagen molecules [91]. In mineralized tissues, the major locus of deH-DHLNL and Pyr involves the 16<sup>C</sup>-Hyl<sup>ald</sup> ( $\alpha 1$ ) and 87-Hyl ( $\alpha 1$ ) [87,92–94]. Apparently not all deH-DHLNL mature into Pyr as the rate of the decrease of deH-DHLNL/ketoamine and an increase of Pyr is often disproportional *in vivo* as well as during an *in vitro* incubation ([95], and M. Sricholpech and M. Yamauchi, unpublished work). The fate of all deH-DHLNL is not fully understood but, recently, an alternative mechanism to Pyr formation through its further oxidation and free arginine addition has been suggested [95].

A minor Pyr analogue form, d-Pyr (deoxyipyridinoline) [96], is formed when a helical lysine residue (instead of hydroxylysine) is involved, thus the analogue form is composed of two telopeptidyl Hyl<sup>ald</sup> and one helical lysine residue. This analogue is relatively abundant in mineralized tissues.

Prl (pyrrole) cross-links, another type of trivalent cross-link found in skeletal tissues, are probably formed by a similar condensation reaction composed of telopeptidyl Hyl<sup>ald</sup>, telopeptidyl Lys<sup>ald</sup> and helical hydroxylysine or lysine [59,97]. However, due to its inherent lability, its quantitative analysis and characterization is limited.

Although Pyr is present in most skeletal and connective tissues, such as cartilage, bone, dentin, tendon, aorta and ligament, it is almost absent from skin and cornea [86]. Interestingly, these two connective tissues cover the external surface of the human body and are subject to UV light, to which Pyr is susceptible [98]. Nature clearly chose a different cross-linking pathway in these tissues, producing a more UV-resistant stable tri-functional cross-link, HHL (histidinohydroxylysinonorleucine) [99,100].

The Lys<sup>ald</sup> derived from the C-telopeptide in skin and cornea collagen reacts with the  $\epsilon$ -amine group of the helical Hyl<sup>87</sup> on a neighbouring molecule to form an aldimine intermolecular cross-link, deH-HLNL (dehydro-hydroxylysinonorleucine). deH-HLNL can then form a stable non-reducible cross-link, HHL, by involving the helical residue His<sup>92</sup> on an  $\alpha 2$  chain [99,101]. Quantitative analysis demonstrated that this cross-link rapidly increased from birth through maturational stage, with a gradual, but steady, increase with aging in both cow and human skin collagen [102]. The structure and experimental evidence indicate that the imidazole C-2 of histidine is linked to C-6 of norleucine ( $\epsilon$ -deaminated lysine residue) which in turn is linked to the C-6 of the amino group of hydroxylysine [99]. Thus two of the three amino acid residues, i.e. hydroxylysine and histidine, in HHL are derived from the helical domain of the molecule. HHL, like Pyr, is stable in acid hydrolysis but, in contrast with Pyr, is stable against UV light [103], and thus suitable for skin and cornea.

Another major cross-link in soft connective tissues is derived from an ACP (aldol condensation product;  $\alpha, \beta$ -unsaturated aldol) formed between two residues of Lys<sup>ald</sup>, which occurs as an intramolecular cross-link located in the N-telopeptides [104,105]. Michael addition of N3 imidazole of histidine to the ACP produces deH-AH (dehydro-aldolhistidine)

which can further condense with the  $\epsilon$ -amino group of hydroxylysine forming an aldimine bond to produce deH-HHMD (dehydro-histidinohydroxymersodesmosine) [105] (Figure 5), one of the major cross-links in soft connective tissue collagen [62,106]. This crosslink, which involves four amino acids, has the most complex chemical structure in all of the known collagen cross-links.

An aldimine cross-link, dehydro-lysinonorleucine, can also occur when Lys<sup>ald</sup> condenses with helical lysine (Figure 5). However, this cross-link is minor in collagen.

The LOX-mediated collagen cross-linking is controlled at three levels, i.e. quantity, quality and maturation. (i) Quantity of cross-links: since all of the crosslinks described are initiated by the LOX-catalysed aldehyde formation in the telopeptides, the levels of aldehyde primarily determine the quantity of cross-links formed. Thus the expression and activity of LOX (possibly LOXLs) are key to controlling the quantity of cross-links. (ii) Quality of cross-links: for quality, i.e. type of cross-links, the extent of lysine hydroxylation in the N- or C-telopeptides and those in the juxtaposed helical domain of a neighbouring molecule is the critical determinant. Thus the levels and activities of LH isoforms are key to control the quality of cross-links. (iii) Maturation of the cross-links: as described above, the formation of the initial aldimine cross-links or ACP and their maturation are spontaneous, and thus not controlled by specific enzymes. It is likely that the microenvironment around the cross-linking regions controls these processes. One such potential factor is the state (extent and type) of glycosylation of hydroxylysine involved in cross-linking due to their bulky structures and hydrophilic nature.

## Conclusions and perspectives

Lysine modifications of collagen have been investigated for the last five decades, and their biological significance, molecular mechanisms and detailed chemistries have been gradually elucidated. The modifications are tissue-specific and probably associated with their functions. Identification of isoforms of LH, LOX and GLT25D has provided significant new insights into the tissue-specific control of such modifications. However, there are still a number of unanswered questions including: (i) specificity and redundancy of these isoforms, (ii) regulatory mechanisms of gene expression and enzymatic activity of these isoforms in different cells and tissues, (iii) function(s) of hydroxylysine glycosylation, and (iv) controlling factor(s) of cross-link maturation. Some of these issues could be addressed by studies employing gain- and loss-of-function approaches (e.g. single- and double-gene knockdown, and overexpression) in well-defined cell-culture systems, and tissue- and developmental stage-specific conditional gene-deletion experiments *in vivo*.

## Acknowledgments

Work in the authors laboratory is supported by the National Institute of Health (NIH) [grant numbers R21 DE019569, DE020909 and AR060978].

## References

1. Carter EM, Raggio CL. Genetic and orthopedic aspects of collagen disorders. *Curr. Opin. Pediatr.* 2009; 21:46–54. [PubMed: 19253462]
2. Hulmes, DJS. Collagen diversity, synthesis and assembly. In: Fratzl, P., editor. *Collagen*. London: Springer; 2008. p. 15–47.
3. Leitinger B. Transmembrane collagen receptors. *Annu. Rev. Cell Dev. Biol.* 2011; 27:265–290. [PubMed: 21568710]
4. Vogel W, Gish GD, Alves F, Pawson T. The discoidin domain receptor tyrosine kinases are activated by collagen. *Mol. Cell.* 1997; 1:13–23. [PubMed: 9659899]



5. Bhadriraju K, Chung KH, Spurlin TA, Haynes RJ, Elliott JT, Plant AL. The relative roles of collagen adhesive receptor DDR2 activation and matrix stiffness on the downregulation of focal adhesion kinase in vascular smooth muscle cells. *Biomaterials*. 2009; 30:6687–6694. [PubMed: 19762078]
6. Jurgensen HJ, Madsen DH, Ingvarsen S, Melander MC, Gardsvoll H, Patthy L, Engelholm LH, Behrendt N. A novel functional role of collagen glycosylation: interaction with the endocytic collagen receptor uparap/ENDO180. *J. Biol. Chem*. 2011; 286:32736–32748. [PubMed: 21768090]
7. Yamauchi M, Shiiba M. Lysine hydroxylation and cross-linking of collagen. *Methods Mol. Biol*. 2008; 446:95–108. [PubMed: 18373252]
8. Avery, NC.; Bailey, AJ. Restraining cross-links responsible for the mechanical properties of collagen fibers: natural and artificial. In: Fratzl, P., editor. *Collagen*. Springer; 2008. p. 81-110.
9. Kivirikko KI, Myllyla R. Posttranslational enzymes in the biosynthesis of collagen: intracellular enzymes. *Methods Enzymol*. 1982; 82:245–304. [PubMed: 6210830]
10. Miller, EJ. Collagen chemistry. In: Piez, KA.; Reddi, AH., editors. *Extracellular Matrix Biochemistry*. Elsevier; 1984. p. 41-78.
11. Uzawa K, Yeowell HN, Yamamoto K, Mochida Y, Tanzawa H, Yamauchi M. Lysine hydroxylation of collagen in a fibroblast cell culture system. *Biochem. Biophys. Res. Commun*. 2003; 305:484–487. [PubMed: 12763018]
12. Gerriets JE, Curwin SL, Last JA. Tendon hypertrophy is associated with increased hydroxylation of nonhelical lysine residues at two specific cross-linking sites in type I collagen. *J. Biol. Chem*. 1993; 268:25553–25560. [PubMed: 8244992]
13. Bank RA, Robins SP, Wijmenga C, Breslau-Siderius LJ, Bardoel AF, van der Sluijs HA, Pruijs HE, TeKoppele JM. Defective collagen crosslinking in bone, but not in ligament or cartilage, in Bruck syndrome: indications for a bone-specific telopeptide lysyl hydroxylase on chromosome 17. *Proc. Natl. Acad. Sci. U.S.A.* 1999; 96:1054–1058. [PubMed: 9927692]
14. Uzawa K, Grzesik WJ, Nishiura T, Kuznetsov SA, Robey PG, Brenner DA, Yamauchi M. Differential expression of human lysyl hydroxylase genes, lysine hydroxylation, and cross-linking of type I collagen during osteoblastic differentiation *in vitro*. *J. Bone Miner. Res*. 1999; 14:1272–1280. [PubMed: 10457259]
15. Royce PM, Barnes MJ. Failure of highly purified lysyl hydroxylase to hydroxylate lysyl residues in the non-helical regions of collagen. *Biochem. J*. 1985; 230:475–480. [PubMed: 3931636]
16. Kivirikko KI, Pihlajaniemi T. Collagen hydroxylases and the protein disulfide isomerase subunit of prolyl 4-hydroxylases. *Adv. Enzymol. Relat. Areas Mol. Biol*. 1998; 72:325–398. [PubMed: 9559057]
17. Valtavaara M, Papponen H, Pirttila AM, Hiltunen K, Helander H, Myllyla R. Cloning and characterization of a novel human lysyl hydroxylase isoform highly expressed in pancreas and muscle. *J. Biol. Chem*. 1997; 272:6831–6834. [PubMed: 9054364]
18. Valtavaara M, Szpirer C, Szpirer J, Myllyla R. Primary structure, tissue distribution, and chromosomal localization of a novel isoform of lysyl hydroxylase (lysyl hydroxylase 3). *J. Biol. Chem*. 1998; 273:12881–12886. [PubMed: 9582318]
19. Yeowell HN, Walker LC. Tissue specificity of a new splice form of the human lysyl hydroxylase 2 gene. *Matrix Biol*. 1999; 18:179–187. [PubMed: 10372558]
20. Mercer DK, Nicol PF, Kimbembe C, Robins SP. Identification, expression, and tissue distribution of the three rat lysyl hydroxylase isoforms. *Biochem. Biophys. Res. Commun*. 2003; 307:803–809. [PubMed: 12878181]
21. van der Slot AJ, Zuurmond AM, Bardoel AF, Wijmenga C, Pruijs HE, Sillence DO, Brinckmann J, Abraham DJ, Black CM, Verzijl N, et al. Identification of PLOD2 as telopeptide lysyl hydroxylase, an important enzyme in fibrosis. *J. Biol. Chem*. 2003; 278:40967–40972. [PubMed: 12881513]
22. van der Slot AJ, Zuurmond AM, van den Bogaerd AJ, Ulrich MM, Middelkoop E, Boers W, Karel Runday H, DeGroot J, Huizinga TW, Bank RA. Increased formation of pyridinoline cross-links due to higher telopeptide lysyl hydroxylase levels is a general fibrotic phenomenon. *Matrix Biol*. 2004; 23:251–257. [PubMed: 15296939]

23. Pornprasertsuk S, Duarte WR, Mochida Y, Yamauchi M. Lysyl hydroxylase-2b directs collagen cross-linking pathways in MC3T3-E1 cells. *J. Bone Miner. Res.* 2004; 19:1349–1355. [PubMed: 15231023]
24. Myllyla R, Wang C, Heikkinen J, Juffer A, Lampela O, Risteli M, Ruotsalainen H, Salo A, Sipila L. Expanding the lysyl hydroxylase toolbox: new insights into the localization and activities of lysyl hydroxylase 3 (LH3). *J. Cell Physiol.* 2007; 212:323–329. [PubMed: 17516569]
25. Ruotsalainen H, Sipila L, Vapola M, Sormunen R, Salo AM, Uitto L, Mercer DK, Robins SP, Risteli M, Aszodi A, et al. Glycosylation catalyzed by lysyl hydroxylase 3 is essential for basement membranes. *J. Cell Sci.* 2006; 119:625–635. [PubMed: 16467571]
26. Yeowell HN, Walker LC. Mutations in the lysyl hydroxylase 1 gene that result in enzyme deficiency and the clinical phenotype of Ehlers–Danlos syndrome type VI. *Mol. Genet. Metab.* 2000; 71:212–224. [PubMed: 11001813]
27. Giunta C, Randolph A, Al-Gazali LI, Brunner HG, Kraenzlin ME, Steinmann B. Nevo syndrome is allelic to the kyphoscoliotic type of the Ehlers–Danlos syndrome (EDS VIA). *Am. J. Med. Genet. A.* 2005; 133A:158–164. [PubMed: 15666309]
28. Ha-Vinh R, Alanay Y, Bank RA, Campos-Xavier AB, Zankl A, Superti-Furga A, Bonafe L. Phenotypic and molecular characterization of Bruck syndrome (osteogenesis imperfecta with contractures of the large joints) caused by a recessive mutation in PLOD2. *Am. J. Med. Genet. A.* 2004; 131:115–120. [PubMed: 15523624]
29. Salo AM, Cox H, Farndon P, Moss C, Grindulis H, Risteli M, Robins SP, Myllyla R. A connective tissue disorder caused by mutations of the lysyl hydroxylase 3 gene. *Am. J. Hum. Genet.* 2008; 83:495–503. [PubMed: 18834968]
30. Risteli M, Ruotsalainen H, Salo AM, Sormunen R, Sipila L, Baker NL, Lamande SR, Vimpari-Kauppinen L, Myllyla R. Reduction of lysyl hydroxylase 3 causes deleterious changes in the deposition and organization of extracellular matrix. *J. Biol. Chem.* 2009; 284:28204–28211. [PubMed: 19696018]
31. Pornprasertsuk S, Duarte WR, Mochida Y, Yamauchi M. Overexpression of lysyl hydroxylase-2b leads to defective collagen fibrillogenesis and matrix mineralization. *J. Bone Miner. Res.* 2005; 20:81–87. [PubMed: 15619673]
32. Saito M, Marumo K. Collagen cross-links as a determinant of bone quality: a possible explanation for bone fragility in aging, osteoporosis, and diabetes mellitus. *Osteoporosis Int.* 2010; 21:195–214.
33. Shinkai H, Yonemasu K. Hydroxylysine-linked glycosides of human complement subcomponent C1q and various collagens. *Biochem. J.* 1979; 177:847–852. [PubMed: 444207]
34. Sricholpech M, Perdivara I, Nagaoka H, Yokoyama M, Tomer KB, Yamauchi M. Lysyl hydroxylase 3 glucosylates galactosylhydroxylysine residues in type I collagen in osteoblast culture. *J. Biol. Chem.* 2011; 286:8846–8856. [PubMed: 21220425]
35. Heikkinen J, Risteli M, Wang C, Latvala J, Rossi M, Valtavaara M, Myllyla R. Lysyl hydroxylase 3 is a multifunctional protein possessing collagen glucosyltransferase activity. *J. Biol. Chem.* 2000; 275:36158–36163. [PubMed: 10934207]
36. Wang C, Luosujarvi H, Heikkinen J, Risteli M, Uitto L, Myllyla R. The third activity for lysyl hydroxylase 3: galactosylation of hydroxylysyl residues in collagens *in vitro*. *Matrix Biol.* 2002; 21:559–566. [PubMed: 12475640]
37. Rautavuoma K, Takaluoma K, Sormunen R, Myllyharju J, Kivirikko KI, Soininen R. Premature aggregation of type IV collagen and early lethality in lysyl hydroxylase 3 null mice. *Proc. Natl. Acad. Sci. U.S.A.* 2004; 101:14120–14125. [PubMed: 15377789]
38. Sipila L, Ruotsalainen H, Sormunen R, Baker NL, Lamande SR, Vapola M, Wang C, Sado Y, Aszodi A, Myllyla R. Secretion and assembly of type IV and VI collagens depend on glycosylation of hydroxylysines. *J. Biol. Chem.* 2007; 282:33381–33388. [PubMed: 17873278]
39. Schegg B, Hulsmeier AJ, Rutschmann C, Maag C, Hennet T. Core glycosylation of collagen is initiated by two  $\beta(1-O)$ galactosyltransferases. *Mol. Cell. Biol.* 2009; 29:943–952. [PubMed: 19075007]

40. Moro L, Romanello M, Favia A, Lamanna MP, Lozupone E. Posttranslational modifications of bone collagen type I are related to the function of rat femoral regions. *Calcif. Tissue Int.* 2000; 66:151–156. [PubMed: 10652964]
41. Schofield JD, Freeman IL, Jackson DS. The isolation, and amino acid and carbohydrate composition, of polymeric collagens prepared from various human tissues. *Biochem. J.* 1971; 124:467–473. [PubMed: 5002672]
42. Toole BP, Kang AH, Trelstad RL, Gross J. Collagen heterogeneity within different growth regions of long bones of rachitic and non-rachitic chicks. *Biochem. J.* 1972; 127:715–720. [PubMed: 4651137]
43. Michalsky M, Norris-Suarez K, Bettica P, Pecile A, Moro L. Rat cortical and trabecular bone collagen glycosylation are differently influenced by ovariectomy. *Biochem. Biophys. Res. Commun.* 1993; 192:1281–1288. [PubMed: 8507198]
44. Tenni R, Valli M, Rossi A, Cetta G. Possible role of overglycosylation in the type I collagen triple helical domain in the molecular pathogenesis of osteogenesis imperfecta. *Am. J. Med. Genet.* 1993; 45:252–256. [PubMed: 8456811]
45. Lehmann HW, Wolf E, Roser K, Bodo M, Delling G, Muller PK. Composition and posttranslational modification of individual collagen chains from osteosarcomas and osteofibrous dysplasias. *J. Cancer Res. Clin. Oncol.* 1995; 121:413–418. [PubMed: 7635871]
46. Brinckmann J, Notbohm H, Tronnier M, Acil Y, Fietzek PP, Schmeller W, Muller PK, Batge B. Overhydroxylation of lysyl residues is the initial step for altered collagen cross-links and fibril architecture in fibrotic skin. *J. Invest. Dermatol.* 1999; 113:617–621. [PubMed: 10504450]
47. Brenner RE, Vetter U, Nerlich A, Worsdorfer O, Teller WM, Muller PK. Altered collagen metabolism in osteogenesis imperfecta fibroblasts: a study on 33 patients with diverse forms. *Eur J. Clin. Invest.* 1990; 20:8–14. [PubMed: 2108039]
48. Bateman JF, Mascara T, Chan D, Cole WG. Abnormal type I collagen metabolism by cultured fibroblasts in lethal perinatal osteogenesis imperfecta. *Biochem. J.* 1984; 217:103–115. [PubMed: 6421277]
49. Cetta G, De Luca G, Tenni R, Zanaboni G, Lenzi L, Castellani AA. Biochemical investigations of different forms of osteogenesis imperfecta. Evaluation of 44 cases. *Connect. Tissue Res.* 1983; 11:103–111. [PubMed: 6224635]
50. Moro L, Bettica P, Romanello M, Suarez KN. 17  $\beta$ -Estradiol and tamoxifen prevent the overglycosylation of rat trabecular bone collagen induced by ovariectomy. *Eur J. Clin. Chem. Clin. Biochem.* 1997; 35:29–33. [PubMed: 9156563]
51. Moro L, Suarez KN, Romanello M. The influence of orchidectomy on collagen glycosylation of trabecular bone in rat. *Eur. J. Clin. Chem. Clin. Biochem.* 1997; 35:269–273. [PubMed: 9166968]
52. Yang C, Niu C, Bodo M, Gabriel E, Notbohm H, Wolf E, Muller PK. Fulvic acid supplementation and selenium deficiency disturb the structural integrity of mouse skeletal tissue. An animal model to study the molecular defects of Kashin–Beck disease. *Biochem. J.* 1993; 289:829–835. [PubMed: 8435081]
53. Amudeswari S, Liang JN, Chakrabarti B. Polar-apolar characteristics and fibrillogenesis of glycosylated collagen. *Collagen Relat. Res.* 1987; 7:215–223.
54. Torre-Blanco A, Adachi E, Hojima Y, Wootton JA, Minor RR, Prockop DJ. Temperature-induced post-translational over-modification of type I procollagen. Effects of over-modification of the protein on the rate of cleavage by procollagen N-proteinase and on self-assembly of collagen into fibrils. *J. Biol. Chem.* 1992; 267:2650–2655. [PubMed: 1733961]
55. Yang CL, Rui H, Mosler S, Notbohm H, Sawaryn A, Muller PK. Collagen II from articular cartilage and annulus fibrosus. Structural and functional implication of tissue specific posttranslational modifications of collagen molecules. *Eur. J. Biochem.* 1993; 213:1297–1302. [PubMed: 8504821]
56. Notbohm H, Nokelainen M, Myllyharju J, Fietzek PP, Muller PK, Kivirikko KI. Recombinant human type II collagens with low and high levels of hydroxylysine and its glycosylated forms show marked differences in fibrillogenesis *in vitro*. *J. Biol. Chem.* 1999; 274:8988–8992. [PubMed: 10085145]

57. Batge B, Winter C, Notbohm H, Acil Y, Brinckmann J, Muller PK. Glycosylation of human bone collagen I in relation to lysylhydroxylation and fibril diameter. *J. Biochem.* 1997; 122:109–115. [PubMed: 9276678]
58. Eyre DR, Glimcher MJ. Analysis of a crosslinked peptide from calf bone collagen: evidence that hydroxylysyl glycoside participates in the crosslink. *Biochem. Biophys. Res. Commun.* 1973; 52:663–671. [PubMed: 4351144]
59. Hanson DA, Eyre DR. Molecular site specificity of pyridinoline and pyrrole cross-links in type I collagen of human bone. *J. Biol. Chem.* 1996; 271:26508–26516. [PubMed: 8900119]
60. Robins SP. Cross-linking of collagen. Isolation, structural characterization and glycosylation of pyridinoline. *Biochem. J.* 1983; 215:167–173. [PubMed: 6626172]
61. Yamauchi M, Noyes C, Kuboki Y, Mechanic GL. Collagen structural microheterogeneity and a possible role for glycosylated hydroxylysine in type I collagen. *Proc. Natl. Acad. Sci. U.S.A.* 1982; 79:7684–7688. [PubMed: 6961443]
62. Yamauchi M, Katz EP, Mechanic GL. Intermolecular cross-linking and stereospecific molecular packing in type I collagen fibrils of the periodontal ligament. *Biochemistry.* 1986; 25:4907–4913. [PubMed: 3768322]
63. Palmieri D, Valli M, Viglio S, Ferrari N, Ledda B, Volta C, Manduca P. Osteoblasts extracellular matrix induces vessel like structures through glycosylated collagen I. *Exp. Cell Res.* 2010; 316:789–799. [PubMed: 20006603]
64. Canty EG, Lu Y, Meadows RS, Shaw MK, Holmes DF, Kadler KE. Coalignment of plasma membrane channels and protrusions (fibripositors) specifies the parallelism of tendon. *J. Cell Biol.* 2004; 165:553–563. [PubMed: 15159420]
65. Kadler KE, Hojima Y, Prockop DJ. 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.* 1987; 262:15696–15701. [PubMed: 3316206]
66. Kagan HM, Li W. Lysyl oxidase: properties, specificity, and biological roles inside and outside of the cell. *J. Cell. Biochem.* 2003; 88:660–672. [PubMed: 12577300]
67. Trackman PC, Bedell-Hogan D, Tang J, Kagan HM. Post-translational glycosylation and proteolytic processing of a lysyl oxidase precursor. *J. Biol. Chem.* 1992; 267:8666–8671. [PubMed: 1349020]
68. Atsawasuwan P, Mochida Y, Katafuchi M, Tokutomi K, Mocanu V, Parker CE, Yamauchi M. A novel proteolytic processing of prollysyl oxidase. *Connect. Tissue Res.* 2011; 52:479–486. [PubMed: 21591931]
69. Thomassin L, Werneck CC, Broekelmann TJ, Gleyzal C, Hornstra IK, Mecham RP, Sommer P. The Pro-regions of lysyl oxidase and lysyl oxidase-like 1 are required for deposition onto elastic fibers. *J. Biol. Chem.* 2005; 280:42848–42855. [PubMed: 16251195]
70. Palamakumbura AH, Jeay S, Guo Y, Pischon N, Sommer P, Sonenshein GE, Trackman PC. The propeptide domain of lysyl oxidase induces phenotypic reversion of rastransformed cells. *J. Biol. Chem.* 2004; 279:40593–40600. [PubMed: 15277520]
71. Vora SR, Guo Y, Stephens DN, Salih E, Vu ED, Kirsch KH, Sonenshein GE, Trackman PC. Characterization of recombinant lysyl oxidase propeptide. *Biochemistry.* 2010; 49:2962–2972. [PubMed: 20192271]
72. Siegel RC, Fu JC. Collagen cross-linking. Purification and substrate specificity of lysyl oxidase. *J. Biol. Chem.* 1976; 251:5779–5785. [PubMed: 9401]
73. Hopkins DR, Keles S, Greenspan DS. The bone morphogenetic protein 1/Tolloid-like metalloproteinases. *Matrix Biol.* 2007; 26:508–523. [PubMed: 17560775]
74. Levene CI, Gross J. Alterations in state of molecular aggregation of collagen induced in chick embryos by  $\beta$ -aminopropionitrile (lathyrus factor). *J. Exp. Med.* 1959; 110:771–790. [PubMed: 14416144]
75. Tang SS, Trackman PC, Kagan HM. Reaction of aortic lysyl oxidase with  $\beta$ -aminopropionitrile. *J. Biol. Chem.* 1983; 258:4331–4338. [PubMed: 6131892]

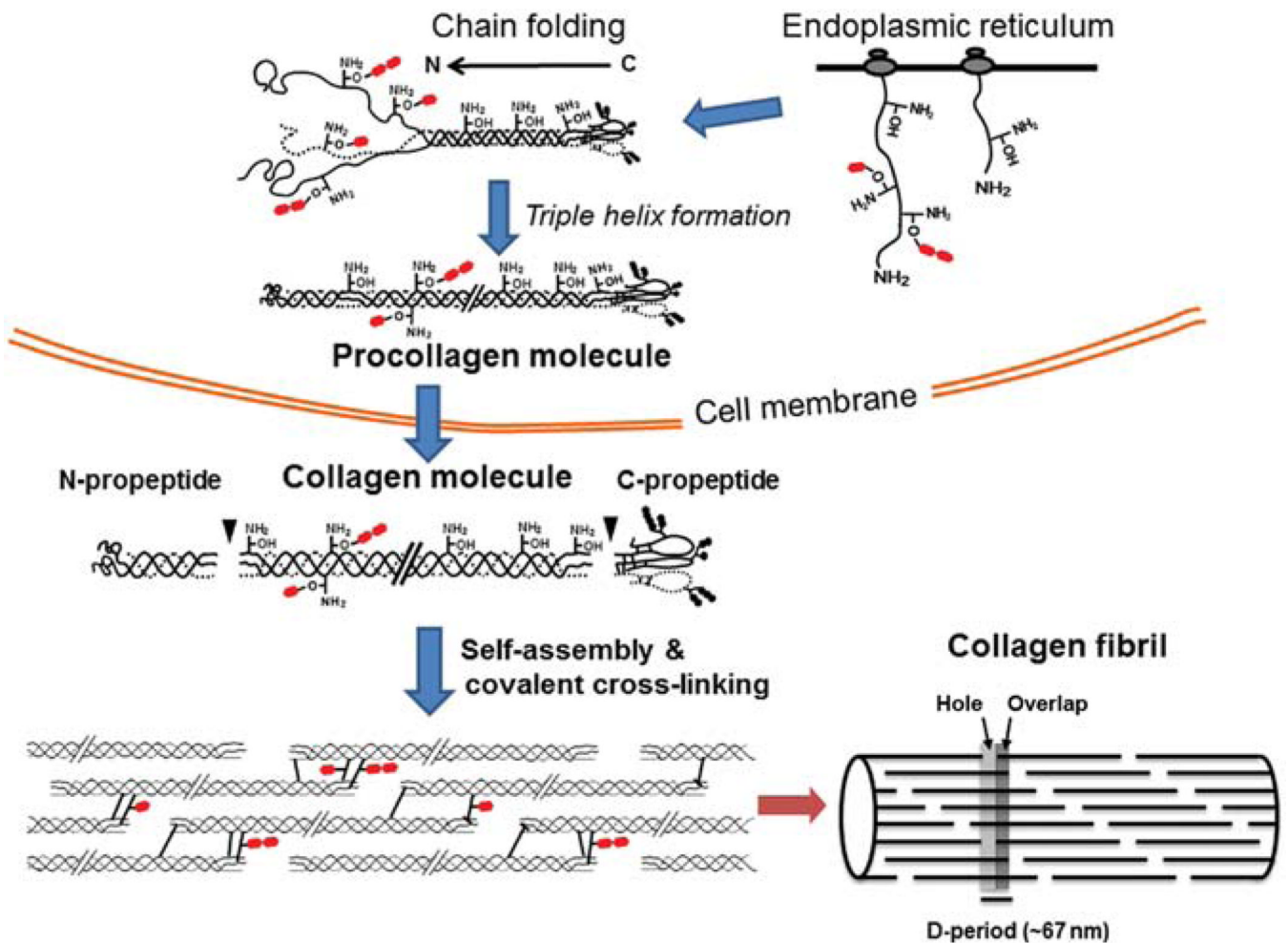
76. Maki JM, Rasanen J, Tikkanen H, Sormunen R, Makikallio K, Kivirikko KI, Soininen R. Inactivation of the lysyl oxidase gene *Lox* leads to aortic aneurysms, cardiovascular dysfunction, and perinatal death in mice. *Circulation*. 2002; 106:2503–2509. [PubMed: 12417550]
77. Hornstra IK, Birge S, Starcher B, Bailey AJ, Mecham RP, Shapiro SD. Lysyl oxidase is required for vascular and diaphragmatic development in mice. *J. Biol. Chem*. 2003; 278:14387–14393. [PubMed: 12473682]
78. Csiszar K. Lysyl oxidases: a novel multifunctional amine oxidase family. *Prog. Nucleic Acid Res. Mol. Biol*. 2001; 70:1–32. [PubMed: 11642359]
79. Maki JM. Lysyl oxidases in mammalian development and certain pathological conditions. *Histol. Histopathol*. 2009; 24:651–660. [PubMed: 19283672]
80. Lucero HA, Kagan HM. Lysyl oxidase: an oxidative enzyme and effector of cell function. *Cell. Mol. Life Sci*. 2006; 63:2304–2316. [PubMed: 16909208]
81. Payne SL, Hendrix MJ, Kirschmann DA. Paradoxical roles for lysyl oxidases in cancer: a prospect. *J. Cell. Biochem*. 2007; 101:1338–1354. [PubMed: 17471532]
82. Atsawasuwan P, Mochida Y, Katafuchi M, Kaku M, Fong KS, Csiszar K, Yamauchi M. Lysyl oxidase binds transforming growth factor- $\beta$  and regulates its signaling via amine oxidase activity. *J. Biol. Chem*. 2008; 283:34229–34240. [PubMed: 18835815]
83. Barry-Hamilton V, Spangler R, Marshall D, McCauley S, Rodriguez HM, Oyasu M, Mikels A, Vaysberg M, Ghermazien H, Wai C, et al. Allosteric inhibition of lysyl oxidase-like-2 impedes the development of a pathological microenvironment. *Nat. Med*. 2010; 16:1009–1017. [PubMed: 20818376]
84. Siddikuzzaman, Grace VM, Guruvayoorappan C. Lysyl oxidase: a potential target for cancer therapy. *Inflammopharmacology*. 2011; 19:117–129. [PubMed: 21107914]
85. Schlotzer-Schrehardt U. Genetics and genomics of pseudoexfoliation syndrome/glaucoma. *Middle East Afr. J. Ophthalmol*. 2011; 18:30–36. [PubMed: 21572731]
86. Eyre DR, Paz MA, Gallop PM. Cross-linking in collagen and elastin. *Annu. Rev. Biochem*. 1984; 53:717–748. [PubMed: 6148038]
87. Yamauchi, M.; Mechanic, G. Cross-linking of collagen. In: Nimni, ME., editor. *Collagen*. Boca Raton: CRC Press; 1988. p. 157-172.
88. Fujimoto D. Isolation and characterization of a fluorescent material in bovine achilles tendon collagen. *Biochem. Biophys. Res. Commun*. 1977; 76:1124–1129. [PubMed: 901463]
89. Fujimoto D, Moriguchi T, Ishida T, Hayashi H. The structure of pyridinoline, a collagen crosslink. *Biochem. Biophys. Res. Commun*. 1978; 84:52–57. [PubMed: 728134]
90. Eyre DR, Oguchi H. The hydroxypyridinium crosslinks of skeletal collagens: their measurement, properties and a proposed pathway of formation. *Biochem. Biophys. Res. Commun*. 1980; 92:403–410. [PubMed: 7356471]
91. Robins SP, Duncan A. Cross-linking of collagen. Location of pyridinoline in bovine articular cartilage at two sites of the molecule. *Biochem. J*. 1983; 215:175–182. [PubMed: 6626173]
92. Yamauchi M, Katz EP, Otsubo K, Teraoka K, Mechanic GL. Cross-linking and stereospecific structure of collagen in mineralized and nonmineralized skeletal tissues. *Connect. Tissue Res*. 1989; 21:159–167. [PubMed: 2605940]
93. Yamauchi M, Katz EP. The post-translational chemistry and molecular packing of mineralizing tendon collagens. *Connect. Tissue Res*. 1993; 29:81–98. [PubMed: 8403898]
94. Otsubo K, Katz EP, Mechanic GL, Yamauchi M. Cross-linking connectivity in bone collagen fibrils: the COOH-terminal locus of free aldehyde. *Biochemistry*. 1992; 31:396–402. [PubMed: 1731897]
95. Eyre DR, Weis MA, Wu JJ. Maturation of collagen ketoimine cross-links by an alternative mechanism to pyridinoline formation in cartilage. *J. Biol. Chem*. 2010; 285:16675–16682. [PubMed: 20363745]
96. Ogawa T, Ono T, Tsuda M, Kawanishi Y. A novel fluor in insoluble collagen: a crosslinking moiety in collagen molecule. *Biochem. Biophys. Res. Commun*. 1982; 107:1252–1257. [PubMed: 7138535]



97. Kuypers R, Tyler M, Kurth LB, Jenkins ID, Horgan DJ. Identification of the loci of the collagen-associated Ehrlich chromogen in type I collagen confirms its role as a trivalent cross-link. *Biochem. J.* 1992; 283:129–136. [PubMed: 1567360]
98. Sakura S, Fujimoto D, Sakamoto K, Mizuno A, Motegi K. Photolysis of pyridinoline, a cross-linking amino acid of collagen, by ultraviolet light. *Can. J. Biochem.* 1982; 60:525–529. [PubMed: 7104830]
99. Yamauchi M, London RE, Guenat C, Hashimoto F, Mechanic GL. Structure and formation of a stable histidine-based trifunctional cross-link in skin collagen. *J. Biol. Chem.* 1987; 262:11428–11434. [PubMed: 3624221]
100. Yamauchi M, Chandler GS, Tanzawa H, Katz EP. Cross-linking and the molecular packing of corneal collagen. *Biochem. Biophys. Res. Commun.* 1996; 219:311–315. [PubMed: 8604983]
101. Mechanic GL, Katz EP, Henmi M, Noyes C, Yamauchi M. Locus of a histidinebased, stable trifunctional, helix to helix collagen cross-link: stereospecific collagen structure of type I skin fibrils. *Biochemistry.* 1987; 26:3500–3509. [PubMed: 3651393]
102. Yamauchi M, Woodley DT, Mechanic GL. Aging and cross-linking of skin collagen. *Biochem. Biophys. Res. Commun.* 1988; 152:898–903. [PubMed: 3130057]
103. Yamauchi M, Prisanh P, Haque Z, Woodley DT. Collagen cross-linking in sun-exposed and unexposed sites of aged human skin. *J. Invest. Dermatol.* 1991; 97:938–941. [PubMed: 1919057]
104. Kang AH, Faris B, Franzblau C. Intramolecular cross-link of chick skin collagen. *Biochem. Biophys. Res. Commun.* 1969; 36:345–349. [PubMed: 5822395]
105. Tanzer ML, Housley T, Berube L, Fairweather R, Franzblau C, Gallop PM. Structure of two histidine-containing crosslinks from collagen. *J. Biol. Chem.* 1973; 248:393–402. [PubMed: 4684687]
106. Tanzer, ML. Cross-linking. In: Ramachandran, GN.; Reddi, AH., editors. *Biochemistry of Collagen*. New York: Plenum; 1976. p. 137-16220.

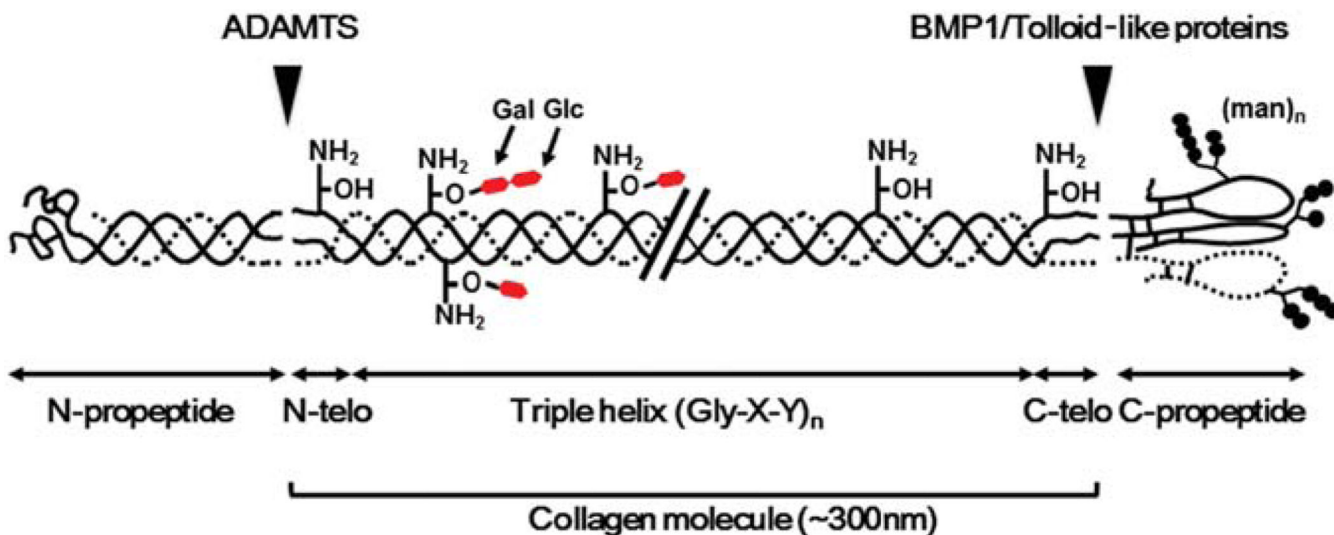
### Summary

- Lysine modifications of fibrillar type I collagen occur both intra- and extracellularly in a tissue-specific manner.
- Inside cells, specific lysine residues in the helical domain are hydroxylated primarily by LH1, and those in the telopeptide domains probably by LH2(b).
- Specific hydroxylysine residues in the helical domain can be sequentially glycosylated by the additions of galactose and then glucose. These reactions may be catalysed by GLT25D1 and LH3 respectively.
- Outside the cells, lysine and hydroxylysine residues in the telopeptides can be oxidatively deaminated by LOX, and possibly by its family members, generating the respective aldehydes, i.e. Lys<sup>ald</sup> and Hyl<sup>ald</sup>.
- The aldehydes then condense with the juxtaposed lysine, hydroxylysine, Lys<sup>ald</sup>, Hyl<sup>ald</sup> and histidine to spontaneously form various covalent crosslinks, which are tissue-specific.



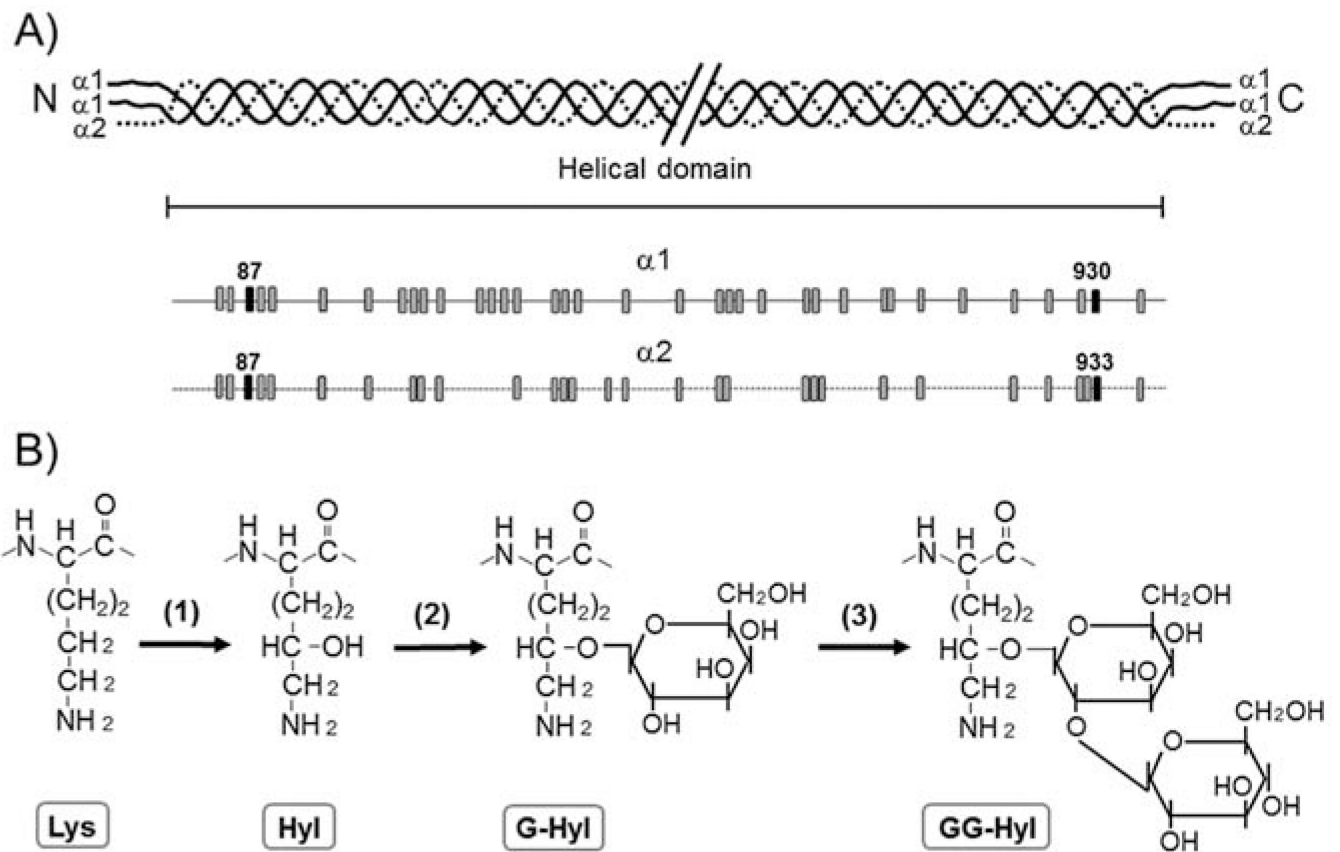
**Figure 1. Type I collagen biosynthesis and lysine modifications**

The top part of the Figure (above the cell membrane) illustrates the intracellular events and the bottom part of the Figure (below the cell membrane) illustrates the extracellular events. During the synthesis of pro $\alpha$  chains in the ER, specific peptidyl lysine residues are hydroxylated to form hydroxylysine (-OH-NH<sub>2</sub>) and, subsequently, specific glycosylated hydroxylysine residues (O-linked glycosylation). For the latter, either single galactose (a red hexagon) or glucose-galactose (two red hexagons) is attached. After these and other modifications (e.g. hydroxylation of proline, asparagine-linked glycosylation shown as closed circles in the C-propeptide), two pro $\alpha$ 1 chains (solid line) and one pro $\alpha$ 2 chain (dotted line) associate with one another and fold into a triple helical molecule from the C- to the N-terminus to form a procollagen molecule, packaged and secreted into the extracellular space. Then both N- and C-propeptides are cleaved to release a collagen molecule. The collagen molecules are then spontaneously self-assembled into a fibril and stabilized by covalent intra- and inter-molecular covalent cross-linking. During fibrillogenesis, molecules are packed in parallel and are longitudinally staggered by an axial repeat distance, D period (~67 nm) creating two repeated regions, i.e. overlap and hole regions, in the fibril.



**Figure 2. The structure of type I procollagen**

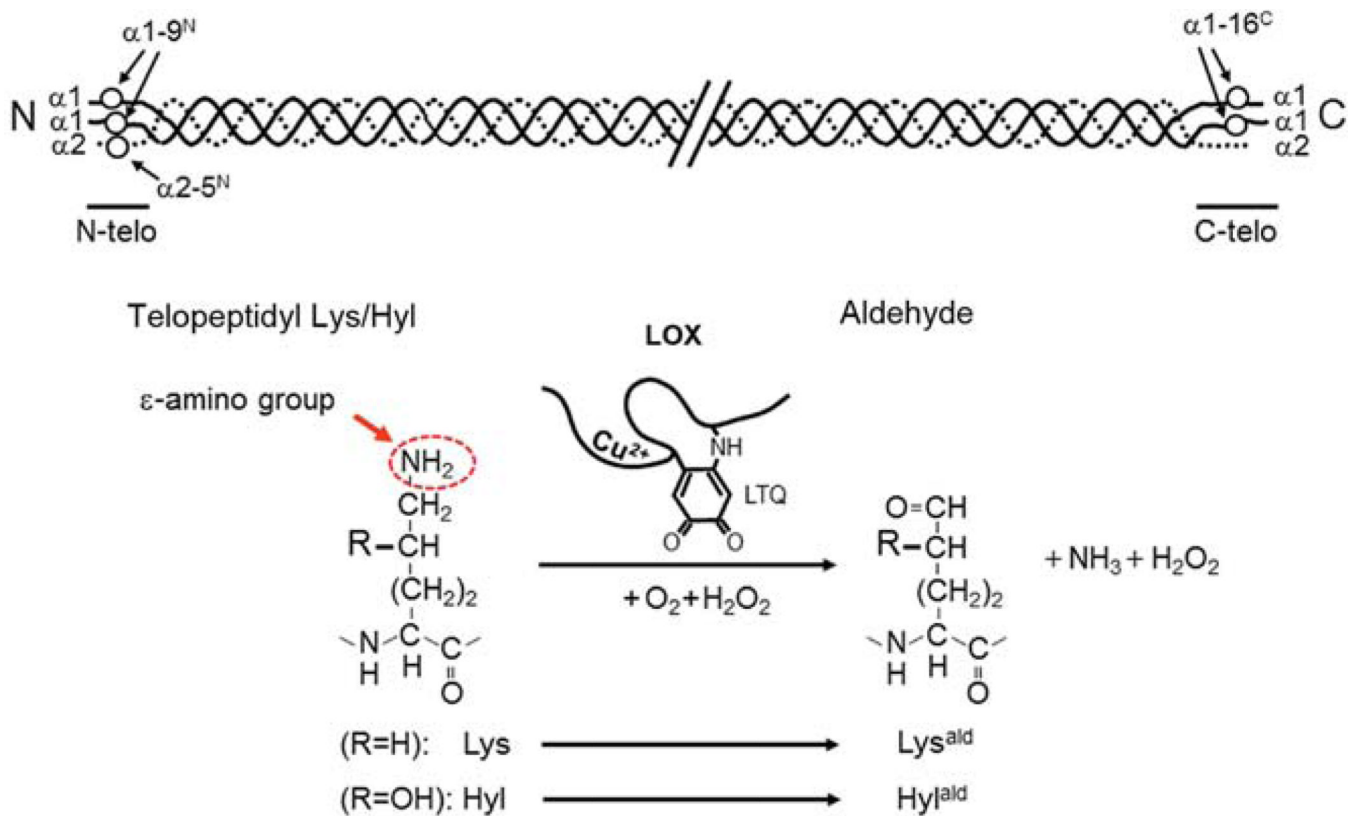
Type I procollagen is composed of two pro $\alpha$ 1(I) chains (solid line) and one pro $\alpha$ 2(I) chain (dotted line) forming a unique triple helix. The N-propeptide contains a short triple helical domain and intrachain-disulfide bonds. Each chain of the globular C-propeptide is glycosylated by high-mannose asparagine-linked oligosaccharide (man)<sub>n</sub> and this domain is stabilized by interchain-disulfide bonds (lines). The procollagen molecule is cleaved at specific sites at both N- and C-terminal regions (indicated by arrows) by ADAMTS and BMP1/Tolloid-like proteinases respectively. For the illustrations of hydroxylysine and its glycosylation, see Figure 1. Type I collagen is a ~300 nm long and ~1.5 nm thick molecule containing three domains: N-telopeptide (N-telo), uninterrupted triple helix and C-telopeptide (C-telo). Each  $\alpha$  chain in the molecule is coiled into an extended left-handed helix and, then, three  $\alpha$  chains are folded into a right-handed triple helix structure. Gal, galactose unit; Glc, glucose unit.



**Figure 3. Lysine modifications in the helical domain of type I collagen**

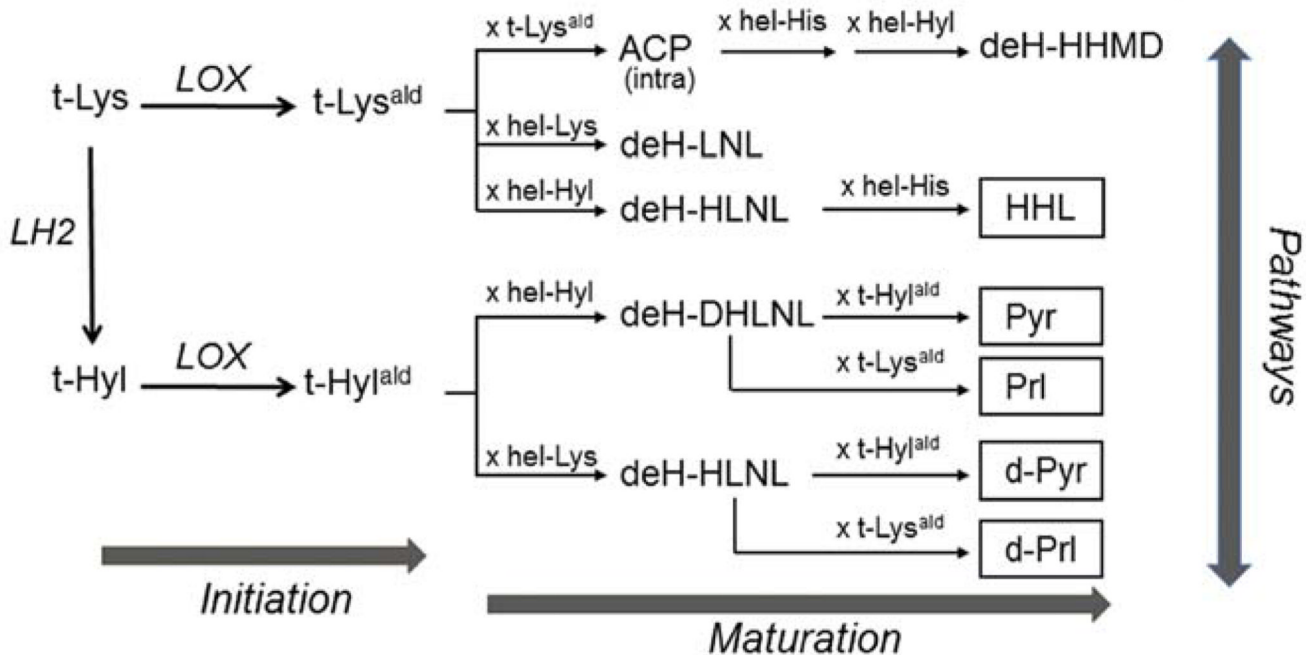
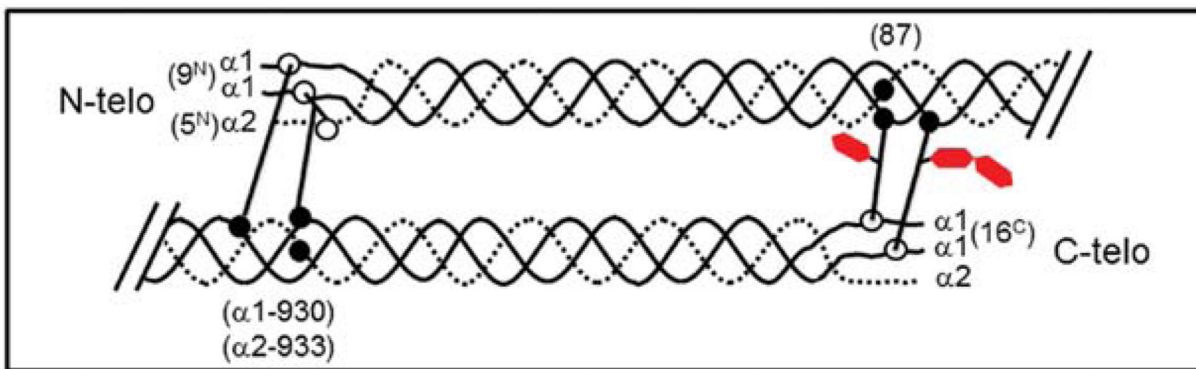
(A) Molecular distribution of lysine residues in the helical domain of  $\alpha 1$  (solid line) and  $\alpha 2$  (broken line) chains based on human type I collagen. There are 36 lysine residues in an  $\alpha 1$  chain and 30 lysine residues in an  $\alpha 2$  chain. Grey squares indicate the relative positions of lysine residues in each chain and black squares indicate the intermolecular cross-linking sites ( $\alpha 1$ , residues 87 and 930;  $\alpha 2$ , residues 87 and 933). (B) Enzyme-mediated sequential modifications of helical lysine residues: (1) lysine hydroxylation forming hydroxylysine; (2) galactosylation of hydroxylysine forming G-Hyl; and (3) glucosylation of G-Hyl forming GG-Hyl. For details of the enzymes involved in each reaction, please see the text.





**Figure 4. Lysine modifications in the telopeptide domains of type I collagen**

There are five lysine or hydroxylysine residues in the telopeptide domains of type I collagen, two in the C- (two  $\alpha 1-16^C$ ) and three in the N- (two  $\alpha 1-9^N$  and one  $\alpha 2-5^N$ ) telopeptide, indicated by open circles. These five lysine/hydroxylysine residues can be oxidatively deaminated by a copper ( $\text{Cu}^{2+}$ )-dependent amine oxidase, LOX, to form the respective aldehydes, i.e. Lys<sup>ald</sup> and Hyl<sup>ald</sup>. Solid line,  $\alpha 1$  chain; broken line,  $\alpha 2$  chain. For details, please see the text.



**Figure 5. LOX-mediated collagen cross-linking**

The top panel (boxed) shows the cross-linking sites of type I collagen. Black lines within and between the molecules indicate examples of the intra- and inter-molecular cross-linkages. Numbers in parentheses indicate the residue numbers of the telopeptidyl aldehydes (open circles) and the helical lysine or hydroxylysine residues (closed circles) involved in cross-linking. Red hexagon, galactose or glucose attached to the helical hydroxylysine involved in cross-linking; telo, telopeptide. Solid line,  $\alpha 1$  chain; broken line,  $\alpha 2$  chain. The bottom panel summarizes initiation, maturation and various pathways of cross-linking. The boxed cross-link compounds are non-reducible cross-links. t, telopeptidyl; hel, helical; ald, aldehyde; d-, deoxyde-; H, dehydro.

**Table 1**

Characteristics of the enzymes for lysine residue modification in collagens and collagen-like proteins, and their association to human disorders EDS VIA, Ehlers-Danlos syndrome type VIA.

Enzyme (gene)	Human chromosomes	Amino acids	Enzymatic functions	Diseases caused by gene mutation or enzyme inhibition [reference(s)]
LH1 ( <i>PLOD1</i> )	1p36.2-36.3	709	Hydroxylates lysine at the Y position of Gly-X-Y in the helical domain	EDS VIA [26] Nevo syndrome [27]
LH2a ( <i>PLOD2a</i> )	3q23-q24	737		
LH2b ( <i>PLOD2b</i> )	+ 63bpE13A	758	Hydroxylates lysine in the telopeptides	Bruck syndrome type 2 [21,28]
LH3 ( <i>PLOD3</i> )	7q22	738	Hydroxylates lysine (type IV and V collagen)  Galactosylates hydroxylysine (controversial) Glucosylates galactosyl-hydroxylysine	A rare connective tissue disorder [29]  Epidermolysis Bullosa Simplex [30]
GLT25D1	19p13.1 1	622	Galactosylates hydroxylysine at the Y position of Gly-X-Y	Unknown
GLT25D2	1q25	626	Galactosylates hydroxylysine at the Y position of Gly-X-Y	Unknown
LOX	5q23.2	417	Oxidatively deaminates lysine/hydroxylysine in the telopeptides	Lathyrism [74]
LOXL1	15q22	574	Oxidatively deaminates lysine of tropoelastin	Susceptible to pseudoexfoliation syndrome/glaucoma [85]
LOXL2	8p21.3-p21.2	774	Oxidatively deaminates lysine/hydroxylysine  Create and maintain microenvironment of cancer and fibrotic disease	Up-regulated in human tumour and fibrotic tissue [83]
LOXL3	2p13	753	Oxidatively deaminates lysine/hydroxylysine	Unknown
LOXL4	10q24	756	Oxidatively deaminates lysine/hydroxylysine	Unknown