

Review

Glycosylation Modulates the Structure and Functions of Collagen: A Review

Igor Tvaroška 

Institute of Chemistry, Slovak Academy of Sciences, 845 38 Bratislava, Slovakia; chemitsa@savba.sk

Abstract: Collagens are fundamental constituents of the extracellular matrix and are the most abundant proteins in mammals. Collagens belong to the family of fibrous or fiber-forming proteins that self-assemble into fibrils that define their mechanical properties and biological functions. Up to now, 28 members of the collagen superfamily have been recognized. Collagen biosynthesis occurs in the endoplasmic reticulum, where specific post-translational modification—glycosylation—is also carried out. The glycosylation of collagens is very specific and adds β -D-galactopyranose and β -D-Glcp-(1 \rightarrow 2)-D-Galp disaccharide through β -O-linkage to hydroxylysine. Several glycosyltransferases, namely COLGALT1, COLGALT2, LH3, and PGGHG glucosidase, were associated with glycosylation of collagens, and recently, the crystal structure of LH3 has been solved. Although not fully understood, it is clear that the glycosylation of collagens influences collagen secretion and the alignment of collagen fibrils. A growing body of evidence also associates the glycosylation of collagen with its functions and various human diseases. Recent progress in understanding collagen glycosylation allows for the exploitation of its therapeutic potential and the discovery of new agents. This review will discuss the relevant contributions to understanding the glycosylation of collagens. Then, glycosyltransferases involved in collagen glycosylation, their structure, and catalytic mechanism will be surveyed. Furthermore, the involvement of glycosylation in collagen functions and collagen glycosylation-related diseases will be discussed.

Keywords: collagen; structure; collagen hydroxylases; collagen glycosyltransferases; catalytic mechanism; glycosylation functions; collagen glycosylation-related diseases



Citation: Tvaroška, I. Glycosylation Modulates the Structure and Functions of Collagen: A Review. *Molecules* **2024**, *29*, 1417. <https://doi.org/10.3390/molecules29071417>

Academic Editor: Juan M. Benito

Received: 26 February 2024

Revised: 19 March 2024

Accepted: 20 March 2024

Published: 22 March 2024



Copyright: © 2024 by the author. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

The extracellular matrix (ECM) is a non-cellular component of the body existing within all tissues and organs and is responsible for their maintenance and integrity. ECM consists of diverse macromolecules. Proteins or glycoproteins, such as collagens, elastin, laminins and tenascins, proteoglycans, glycan, and polysaccharides form a dynamic 3-dimensional network providing structural support for the cells and tissues [1,2]. These macromolecules modulate many cellular processes, such as cell signaling functions, properties, and phenotype. ECM coordinates interactions between cells in organs and tissues using inside-out and outside-in signals [3]. Interactions of ECM components with their cell receptors play significant roles in health and diseases [4–6].

Collagens, with over 30 % of the total proteins, are the most dominant protein in the human body and comprise 80–90 % of ECM as the principal constituent [6–8]. The characteristic structural feature of collagens is the presence of a triple helix consisting of repeated amino acid triplets (XaaYaaGly)_n. In vertebrates, the collagen family comprises 28 diverse members [9–11] divided into several subfamilies based on their supramolecular architecture and function. Collagens are present in soft and hard connective tissues such as the skin, tendons, cartilage, bone, teeth, and corneas [1]. They provide the required tissue integrity and are involved in the other cell processes. Defects in collagen structures, either autoimmune or genetic origin, are associated with most 40 collagen diseases [12,13] of

various severity, such as bone and cartilage abnormalities, skin alternation, visual defects, hearing loss, vessel abnormalities, kidney diseases, lung diseases, and cancer [14–18].

Before forming a triple helix, a collagen single chain undergoes post-translational modifications (PTMs) in the endoplasmic reticulum (ER). The PTM starts with hydroxylation of the proline and lysine residues, which thermally stabilizes the triple helix [19]. Then, some of the formed hydroxylysine (Hyl) residues are glycosylated. The unique glycosylation of collagen involves linking β -D-galactopyranose via β -linkage to Hyl and adding glucopyranose to the C2 position of galactopyranose by α -linkage. The resulting O-glycan on collagen is the simple β -D-galactopyranose or α -D-glucopyranosyl-(1 \rightarrow 2)- β -D-galactopyranose disaccharide [20]. Although collagen has been the subject of numerous studies for about 100 years, progress in understanding the glycosylation of collagens commenced in the last two decades, and many aspects remain unclear [21]. Moreover, collagen glycosylation was recently associated with autoimmune diseases like rheumatoid arthritis [22], systemic sclerosis [23], and cancer [24]. Recently, various crystal structures of human collagen lysyl hydroxylase 3 (LH3, decoded by *PLOD3* gene) that are responsible for the transfer of glucopyranose to the galactosylated Hyl have been solved [25–28].

Several reviews and books have been published on various collagen structures, biosynthesis, biological functions, and application aspects [10–12,29–37]. Therefore, this is not intended to be an exhaustive review of all collagen studies, which are only briefly discussed. Instead, the review will focus on collagens' less understood aspects, namely glycosylation. The review starts with a short overview of collagen structure and biosynthesis. Then, current knowledge about collagen glycosylation, enzymes involved in glycosylation, and their catalytic mechanisms will be discussed. Finally, the question of how alternation in glycosylation is associated with structural changes and various biological and pathological processes will be addressed.

2. Collagen Structure

Collagens interact with various adhesion molecules associated with multiple biological processes. A systematic understanding of the structure of collagens is crucial for understanding their biological functions. In the last 40 years, the knowledge about collagen's structure has considerably progressed and has been described in several reviews [7,9–11,29–31,33,38–42]. The characteristic feature of collagens is triple helix formed by three parallel polypeptide chains.

In vertebrates, the collagen superfamily consists of 28 members with numbers denoted by Roman numerals [9,10,43]. Collagens are classified based on their domain structure and supramolecular organization. They can be grouped into fibril-forming collagens (types I, II, III, V, XI, XXIV, and XXVII); FACIT—fibril-associated collagens (types IX, XII, XIV, and XX); network-forming collagens (types IV, VI, VII, VIII, and X); FACIT-like collagens (types XVI, XIX, XXI, and XXII); transmembrane collagens (types XIII, XVII, XXIII, and XXV); multiplexin collagens (types XV and XVIII); and other molecules with collagenous domains (types XXVI and XXVIII) [43]. Collagens can be homotrimers or heterotrimers. Homotrimeric collagens are formed from three identical α -chains, e.g., $[\alpha 1(\text{III})]_3$ for type III collagen. Heterotrimeric collagens contain different α -chains of the same collagen type, like $[\alpha 1(\text{I})][\alpha 2(\text{I})]_2$ for type I collagen. Some collagens can also occur with various α -chain compositions. An example is type V collagen formed by combinations of three different α -chains, namely $[\alpha 1(\text{V})][\alpha 2(\text{V})]_2$, $[\alpha 1(\text{V})][\alpha 2(\text{V})][\alpha 3(\text{V})]$, or $[\alpha 1(\text{V})]_3$ [43]. Type I collagen is the most abundant and best described of all collagens. Type I collagen provides a basis for the mechanical properties of many biological materials and makes the most significant part of collagen fibrils of skin, tendons and ligaments, veins and arteries, organs, and bone [41,43].

2.1. The Primary Structure

Collagens consist of three α -polypeptide chains (Figure 1). The primary structure of polypeptide chains, termed α -chains, affects collagen structural and physiological prop-

erties at all hierarchical levels. The defining motif of polypeptide chains is a repeating Xaa-Yaa-Gly sequence presented in all collagen types (Figure 1a). The folding of the three chains requires that Gly is every third residue in triplets. The remaining two amino acids, Xaa and Yaa, are often (2*S*)-proline (Pro, 28%) and (2*S*,4*R*)-4-hydroxyproline (Hyp, 38%) but can also be any other amino acid, except glycine [30,44] (Figure 1a). As a consequence, the Pro-Hyp-Gly moiety is the most frequent of all possible triplets in collagen (10.8%). There is also a high amount of lysine (Lys) residues in position Yaa (10.5%), reflected in the occurrence of the tripeptide Pro-Lys-Gly of 2.7% [44]. The presence of lysine residue is critical for the structure and behavior of collagen due to its role in post-translational modifications. Each polypeptide chain contains 662 to 3152 amino acids, depending on the collagen type [9].

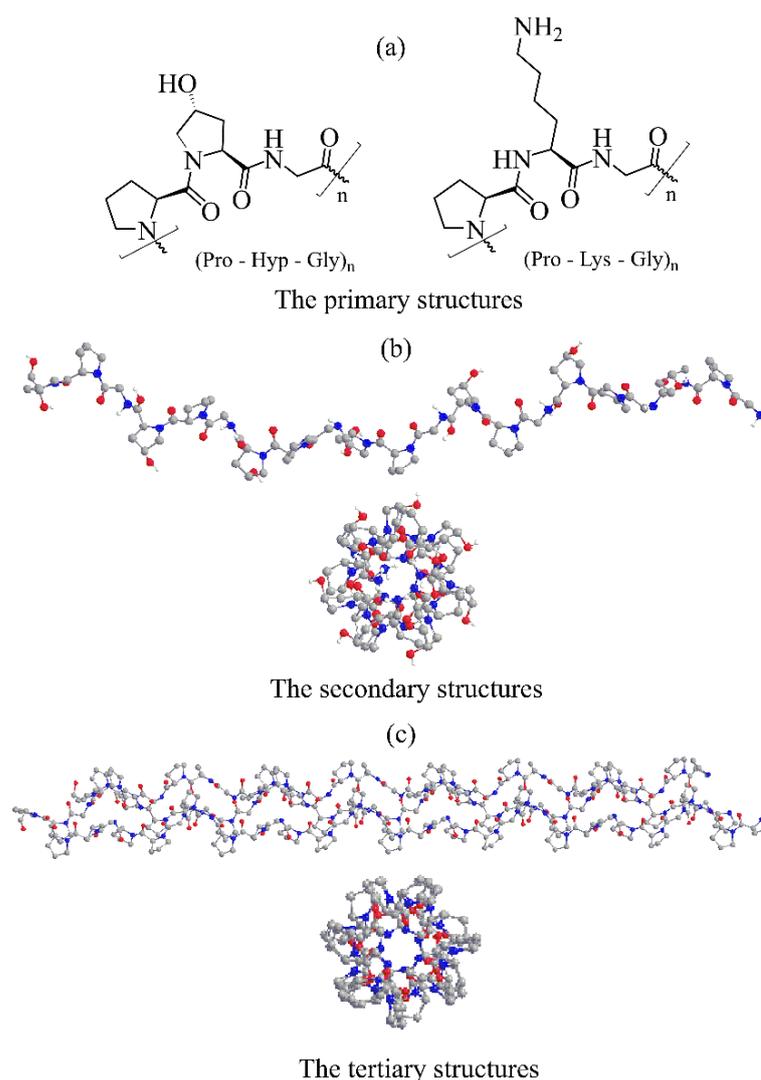


Figure 1. Illustration of collagens molecular structures. (a) Schematic representation of the most abundant collagen Pro-Hyp-Gly and Pro-Lys-Gly triplets. (b) Crystal structure of a collagen model (Gly-Pro-Hyp)₁₀ single chain helix (PDB entry 4OY5, ref. [45]). (c) The collagen triple helix of the (Pro-Pro-Gly)₁₀ model (PDB entry 1K6F, ref. [46]). Crystal structures are depicted in colored ball-and-stick representations.

2.2. The Secondary Structure

In all collagens, the polypeptide chain adopts a polyproline II-type left-handed helical conformation [47] with all peptide bonds in the trans conformation [30]. The high-resolution crystal structure of the collagen peptide model (Pro-Hyp-Gly)₁₀ [45] revealed that

(Figure 1b) the proline rings could adopt two ring conformers, while the hydroxyproline rings prefer only one ring conformation. Hydrogen bonds between backbone amine groups and carbonyl groups of amino acids are essential factors stabilizing the left-handed helix. In collagens, α -chains contain at least one collagenous domain and non-collagenous domains. Both ends of the α -chain, N-terminus and C-terminus are non-collagenous domains.

2.3. The Tertiary Structure

Three parallel α -helical chains supercoil around the central axis to form a right-handed triple helix (Figure 1c). Due to difficulties in investigating collagens at the atomic level, detailed information about the triple helix structure was obtained using various collagen peptide models [40], showing that the triple helix conformation is close to the 7/2 helix. The recently solved high-resolution crystal structure of the (Pro-Hyp-Gly)₁₀ collagen model [45] at 0.98 Å resolution confirmed the 7/2 helix with the helical pitch corresponding to the length of three amino acids. It provided structural details about the α -chain at the atomic level. This tertiary structure, called tropocollagen, is stabilized by inter-chain hydrogen bonds, and three chains are staggered by one residue relative to each other. The tropocollagen can develop a straight or kinked rope-like rod shape. However, this rod can have a kink or flexible interruption depending on (Xaa-Yaa-Gly)_n sequence. Crucial for the helix stability is the presence of glycine that occupies the position in the center of the triple helix. A single replacement of Gly with other amino acids can destabilize the triple helix [30]. The more bulky side chains of the other two amino acids point outward and almost perpendicular to the helix axis. The average molecular weight of the tropocollagen is 300 kDa and <300 nm long with a diameter of 1–2 nm [30]. The collagen triple helix binds to various cell surface receptors such as integrins, discoidin domain receptors, glycoprotein VI, leukocyte-associated IG-like receptor-1, and members of the mannose receptor family [48]. These interactions regulate cell behavior, and collagen gene mutations are associated with various connective tissue diseases [32].

2.4. The Quarternary Structure

Tropocollagens self-associate into intermediate-size fibrils called microfibrils [49]. Each microfibril contains five tropocollagen molecules. In microfibrils, tropocollagens are linked end-to-end in a line and parallel but staggered alignment [50]. There is a space between adjacent molecules about 65–67 nm and a distance of 40 nm between two 300 nm long consecutive molecules [51]. The covalent intramolecular and intermolecular cross-linking between lysine and histidine of the C-terminus or N-terminus of tropocollagen molecules is responsible for the stability of insoluble microfibrils [52].

2.5. Supramolecular Assemblies

Various classes of collagen structures, such as fibrils, networks, filaments, and trans-membrane collagenous domains, are present in the ECM. The following will discuss fibril type I collagen as the principal component of all fibrous tissues except cartilage. After the formation of microfibrils, collagen fibrillogenesis continues by further self-assembling several microfibrils into collagen fibrils with a rope-like structure [53]. The length of collagen fibrils is up to 1 μ m with a diameter of up to 500 nm. Thus, the elongation from the collagen triple helix (<2 nm diameter, 300 nm length) to the fibril (500 nm, 1 mm) during fibrillogenesis is remarkable on its scale. In the next step of the hierarchical building process, several thousand (~4000) collagen fibrils can twist or pack together to form a collagen fiber. Finally, bunches of collagen fibers interact with one another to create the basic organizational structure of the organism with a three-dimensional network [41,49,52].

In vivo, collagen protein materials are organized into various supramolecular structures hierarchically, forming high-order assemblies [49,52]. Figure 2 shows an overview of the hierarchy in the structure with its length scales [41]. Collagens from different groups interact with collagen from other groups and with non-collagenous proteins and, together with their abundance, modulate the higher order of tissue structures. They play vital roles

in tissue development, growth, repair, and remodeling. On the right hand, Figure 2 illustrates steps in forming non-mineralized and mineralized collagen fibers at these different levels. The formation of nanoscale hydroxyapatite (HAP) crystals is a multistep process (biomineralization) in which water within the collagen matrix is replaced by calcium and phosphate complexes. These complexes are precursors of HAP, which fills the gap regions in the collagen fibrils in hard tissues such as bone and teeth [41,54].

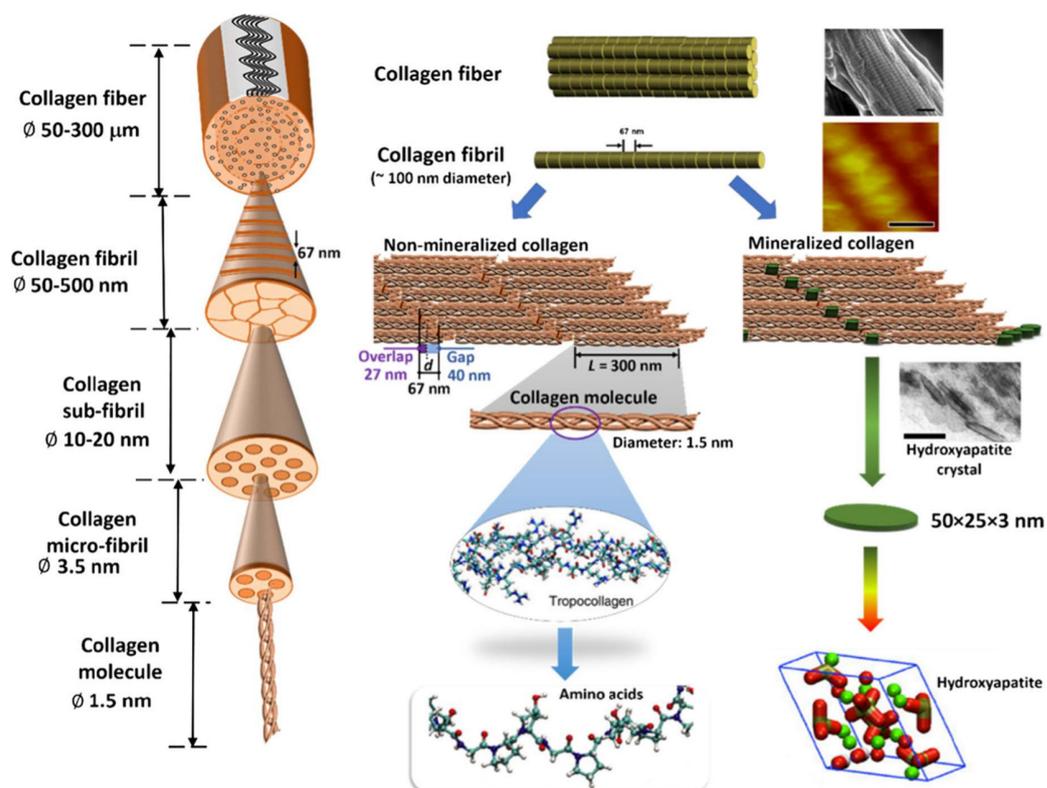


Figure 2. Hierarchical structure of a collagen fiber. The left-hand side shows the hierarchy of the collagen fiber with length-scales. A more detailed illustration of each hierarchy level is explained for non-mineralized and mineralized collagen fibers on the right-hand side. Each collagen molecule comprises three peptide chains forming the ≈ 300 nm long triple helical collagen molecule. Collections of collagen molecules aggregate both in lateral and longitudinal directions to form fibrils. Reprinted with permission from [41]. Copyright 2019 with permission from Elsevier.

3. Post-Translation Modifications of Collagen

Collagen biosynthesis, schematically described in Figure 3, assembly and maturation are multistep processes that occur intracellularly and extracellularly and have been discussed in several papers recently [11,38,43,55,56]. Therefore, this review will focus on post-translational modifications critical to collagen's structure and biological functions. Collagen biosynthesis commences within the nucleus of cells by the synthesis of the single polypeptide pro- α -chain called the pre-procollagen chain, which contains a collagenous α -chain domain and the N- and C-terminal domains. After transcription, the pre-procollagen chain is translocated to the rear endoplasmic reticulum (RER). In the RER, the signaling sequence at its N-terminal domain is cleaved, and post-translational modifications (PTM) are carried out.

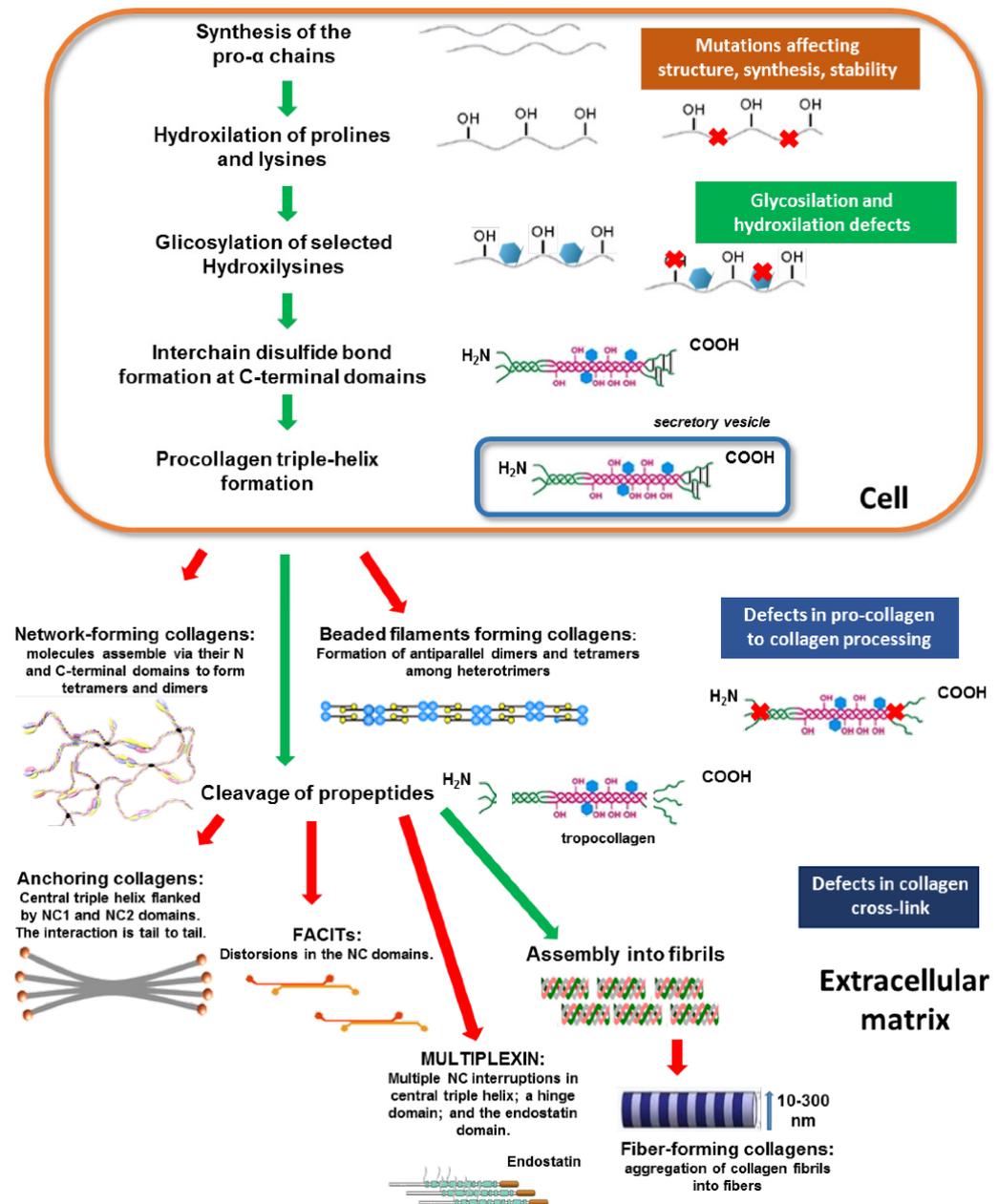


Figure 3. Schematic representation of collagen biosynthesis steps. The main biosynthesis steps of various collagen types are indicated. Green arrows highlight the contiguous processing steps, whereas red arrows indicate the final phase of collagen assembly into the different structural conformations. Colored boxes on the right indicate potential alterations occurring during collagen processing at various steps, thus causing abnormalities in the structure and assembly. Abbreviations: NC, non-collagenous domain. Reprinted with permission from [14]. Copyright 2018 with permission from MDPI.

Intracellular and extracellular PTMs represent critical steps of collagen biosynthesis that influence collagen's structure and functionality [32,55,57–61]. Before triple helix formation, the α -chain is stabilized by intra-chain disulfide bonds, and an *N*-glycan is transferred to an asparagine residue in the C-terminal propeptide domain. The principal intracellular PTMs of polypeptide α -chains include the hydroxylation of selected proline and lysine residues followed by the *O*-glycosylation of specific hydroxylysine residues required for proper folding and assemble into a right-handed triple helix to form pro-collagens. Enzymes involved in these modifications are galactosyltransferases and glucosyltrans-

ferases [62]. In extracellular space, lysine and hydroxylysine are oxidatively deaminated, leading to reactive aldehydic residues that subsequently form intra and intermolecular covalent cross-linkages influencing the functions of fibrils [58].

In the following events, three pro- α -chains begin to associate at the C-domain and propagate toward the N-terminus, becoming a pro-collagen molecule. The triple helix formation prevents further PTM, and pro-collagen molecules are secreted into extracellular space, where N- and C-domains are cleaved, leading to tropocollagens. Finally, tropocollagens self-assemble into various supramolecular structures for their function depending on the collagen types (Figure 3).

4. Hydroxylation of Proline and Lysine Residues

In the procollagen triple helix, repetitive Xaa-Yaa-Gly sequences contain numerous hydroxyproline and hydroxylysine residues in the Xaa and Yaa positions. Their presence is a consequence of enzymatic hydroxylation by three collagen-specific enzymes, namely prolyl 4-hydroxylases (P4H), prolyl 3-hydroxylases (P3H), and lysyl hydroxylases (LH) that hydroxylate numerous proline and lysine residues into hydroxyproline (Hyp) and hydroxylysine (Hyl) throughout PTMs [32,34,57,59,61,63]. These enzymes are a large and diverse superfamily of 2-oxoglutarate-dependent oxygenases (2OGXs) that catalyze various oxidative reactions [64–66]. The catalytic mechanism of these hydroxylases is similar and requires Fe^{2+} as a cofactor and molecular oxygen (O_2), 2-oxoglutarate (2OG), and ascorbate as co-substrates (Figure 4).

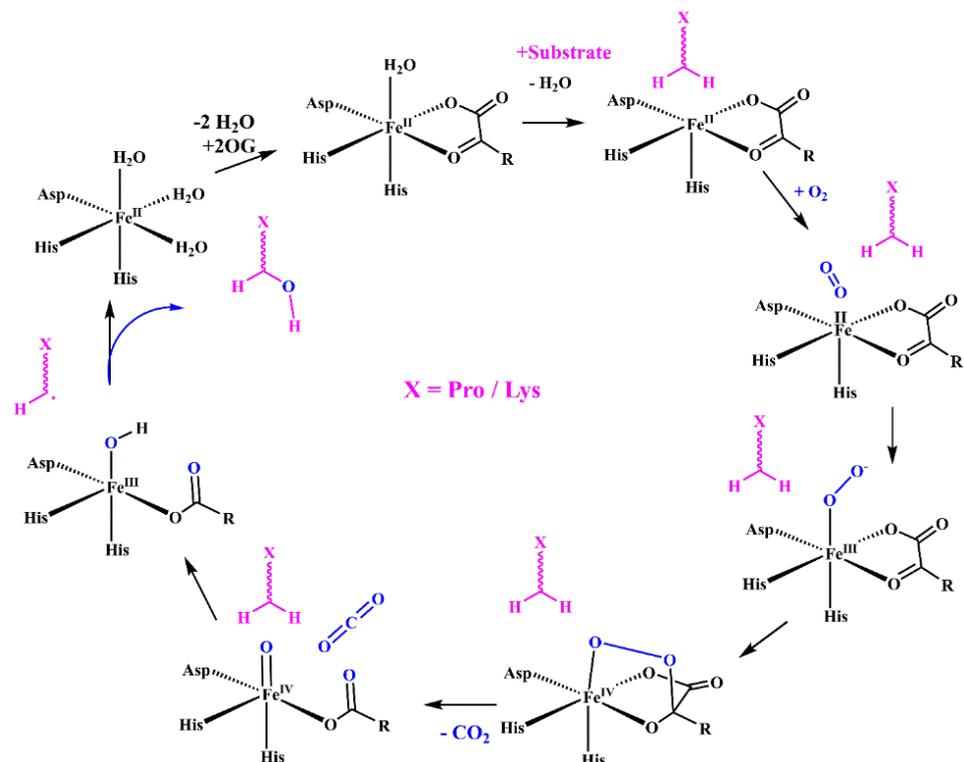


Figure 4. Mechanism of the Fe(II)- and 2-OG-dependent dioxygenases including P4H, P3H, and LH hydroxylases. The enzyme active site consists of a ferrous iron ion coordinated by two His residues and one Asp residue. The catalytic reaction consists of several consecutive steps: binding of 2-OG displacing two water molecules; the substrate binding replaces a third water molecule; binding oxygen and forming an anionic intermediate; the formation of a cyclic peroxide molecule by attacks the ketone of 2-OG; the decomposition of intermediate and the formation of the Fe(IV)oxo species; the abstractions of hydrogen from the substrate; and the reaction of substrate radical with the Fe(III)–OH complex to form the hydroxylated substrate.

4.1. Prolyl 4-Hydroxylase

Collagen prolyl 4-hydroxylase (P4H, EC 1.14.11.2) is the enzyme crucial for collagen biosynthesis. P4H exists as the heterotetramer $\alpha_2\beta_2$ formed from α and β subunits. The tetrameric assembly is required for the stability and total activity of the enzyme. P4H catalyzes proline 4-hydroxylation at the Yaa position in various Xaa-Pro-Gly sequences of the α -chains (Figure 5a) with different efficiency [67–69]. The 4-proline hydroxylation is critical for the thermal stability of collagen at physiological temperatures. Proline 4-hydroxylation is essential for assembling procollagen chains into triple helices through hydrogen bonding and also influences interactions between integrins and collagens [70]. For collagen P4H, several isoenzymes have been identified. Three isoforms of the α subunit, α (I), α (II), and α (III), were identified with α (I) prevailing. The α subunit forms the catalytic domain responsible for substrate recognition and enzymatic activity. The β subunit functions as protein disulfide isomerase (PDI) [59,69,71]. Following decades of unsuccessful attempts to solve the high-resolution structure of P4H, recently, the crystal structure of the catalytic domain of human collagen P4H-II of the CAT domain complexed with the intact β /PDI subunit has been solved at a 3.8 Å resolution [72].

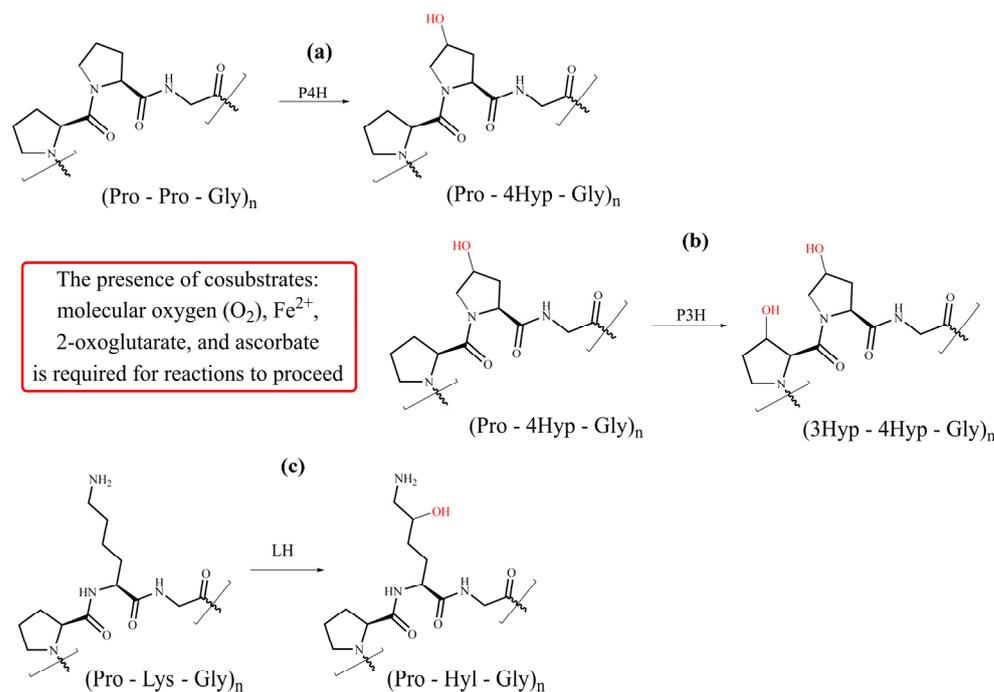


Figure 5. Schematic representation of collagen-specific PTMs—hydroxylation of proline and lysine residues with (a) P4H (prolyl 4-hydroxylases), (b) P3H (prolyl 3-hydroxylases), and (c) LH (lysyl hydroxylases).

Also, structures of several other P4H enzymes and functional domains of the subunit and PDI have been determined [68,72–78]. The crystal structure of P4H has been complemented by a quantum mechanics/molecular mechanics molecular (QM/MM) modeling study on how Glu127 and Arg161 mutations affect substrate binding [79]. P4H is crucial in collagen synthesis and has been associated with fibrotic diseases, hypoxia, and cancer [80]. Therefore, these data have provided valuable information for the rational design of potential therapeutic for fibrosis and cancer [57].

4.2. Prolyl 3-Hydroxylase

Collagen prolyl 3-hydroxylase (P3H, EC 1.14.11.7) catalyzes the hydroxylation of proline residue in position 3 (Figure 5b). The substrate for a P3H action is the (Pro-4Hyp-Gly) sequence [57,59,63,81]. Not all proline residues in such moieties are hydroxylated; the extent of hydroxylation depends on the collagen type and is tissue-specific. Three isoforms

exist in humans, namely P3H1, P3H2, and P3H3. All have 2OGX domain homologous with P4H and proceed via a similar reaction mechanism [81]. An active P3H1 forms a complex with the cartilage-associated protein (CRTAP) and the peptidyl-prolyl *cis-trans* isomerase cyclophilin B (CYPB) in 1:1:1 stoichiometry [81]. Such a complex is not needed for the functioning of P3H2 and P3H3. The P3H1 isoform occurs in tissues full of fibrillar collagens, whereas P3H2 is mainly expressed in tissues rich in basement membranes, such as the kidney [82,83]. The function of P3H is not fully understood, although inactivating P3H1 causes a severe, recessive form of osteogenesis imperfecta characterized by skeletal deformities and bone fragility [84,85]. It was suggested that 3-proline hydroxylation plays an essential biological function by influencing the ordered self-assembly of collagen supramolecular structures [86,87].

4.3. Lysyl Hydroxylases

Collagen lysyl hydroxylase (LH, EC.1.14.11.4) catalyzes the hydroxylation of lysine residues (Figure 5c) in ER and leads to 5-hydroxylysine (Hyl). There are three isoforms of LH in humans: LH1, LH2, and LH3 encoded by *PLOD1*, *PLOD2*, and *PLOD3* genes, respectively. A catalytic mechanism of LH is similar to that described for prolyl hydroxylases (Figure 4). A formation of LH dimers is required for their hydroxylation activity [88,89]. Hydroxylation precedes two fundamental modifications of collagen: glycosylation and covalent cross-linking [57–60]. The hydroxylation of lysine residues in the Xaa-Lys-Gly sequence dominates, although Xaa-Lys-Al and Xaa-Lys-Ser moieties are also hydroxylated in the telopeptide domain. Notably, a lysine residue in the Xaa position is not hydroxylated. Compared with proline hydroxylation, the extent of lysine hydroxylation is more variable, depending on the tissues and physiological conditions, and telopeptide and helical domains are hydrolyzed differently [60]. While the LH1 hydroxylates lysines are located in the helical domain, the LH2 is solely responsible for hydroxylation in the telopeptide domain [25,90,91].

Recent biochemical and homology modeling studies revealed that LH1 and LH2 exhibit, before unobserved, galactosylhydroxylysyl glucosyltransferase (GlcGalHylT) activity, thus being bifunctional enzymes [92]. The LH3 is a multifunctional enzyme, and it was assumed that it bears three different catalytic activities, namely lysine hydroxylase (LH), hydroxylysyl galactosyltransferase (GalHylT), and galactosylhydroxylysyl glucosyltransferase [88,93–96]. However, recent experimental findings indicate that the primary function of LH3 is the final step of glycosylation, that is, the transfer of the glucopyranose residue to galactohydroxylated lysine [97–99].

5. Oxidation of Lysyl Residues

After secretion into extracellular space, some telopeptide hydroxylated lysine residues are catalytically oxidized by enzymes lysyl oxidases. The oxidative deamination of lysine residues is a prerequisite for the spontaneous formation of covalent cross-links that stabilize developing fibrils. The lysyl oxidases family comprises five highly homologous members: lysyl oxidase (LOX, EC 1.4.3.13), lysyl oxidase like-1 (LOXL1), lysyl oxidase like-2 (LOXL2), lysyl oxidase like-3 (LOXL3), and lysyl oxidase like-4 (LOXL4) [100–102]. Lysyl oxidases are copper-dependent enzymes that catalyze the oxidative deamination of Lys and Hyl residues in telopeptide domains of collagen (Figure 6), leading to their aldehydes [58]. The catalytic domain of all LOX family members is highly conserved; therefore, they are assumed to have very similar catalytic properties [103]. This reaction is the first step in covalent collagen crosslinking in ECM [34,42,60,104]. The catalytic site of LOX contains copper ion (Cu^{2+}), and the reaction requires the covalently integrated cofactor lysyl tyrosylquinone (LTQ). Recently, the crystal structure of human lysyl oxidase-like 2 (hLOXL2) at a 2.4-Å resolution [105], and this structure has been used as the template to generate a homology 3D model of LOX [106].

The resulting telopeptidyl aldehydes spontaneously react with other aldehydes or free lysine and hydroxylysine in a series of condensation reactions, generating various intra- and intermolecular cross-links that are crucial for biomechanical properties of collagens.

Collagens have three major cross-linking pathways (Figure 7) [60]. Lysine-based cross-linking occurs mainly in adult skin, cornea, and sclera. In contrast, hydroxylysine-based cross-links prevail in bone, cartilage, ligament, most tendons, embryonic skin, and internal connective tissues of the body. In humans, cross-links decreases with age [34].

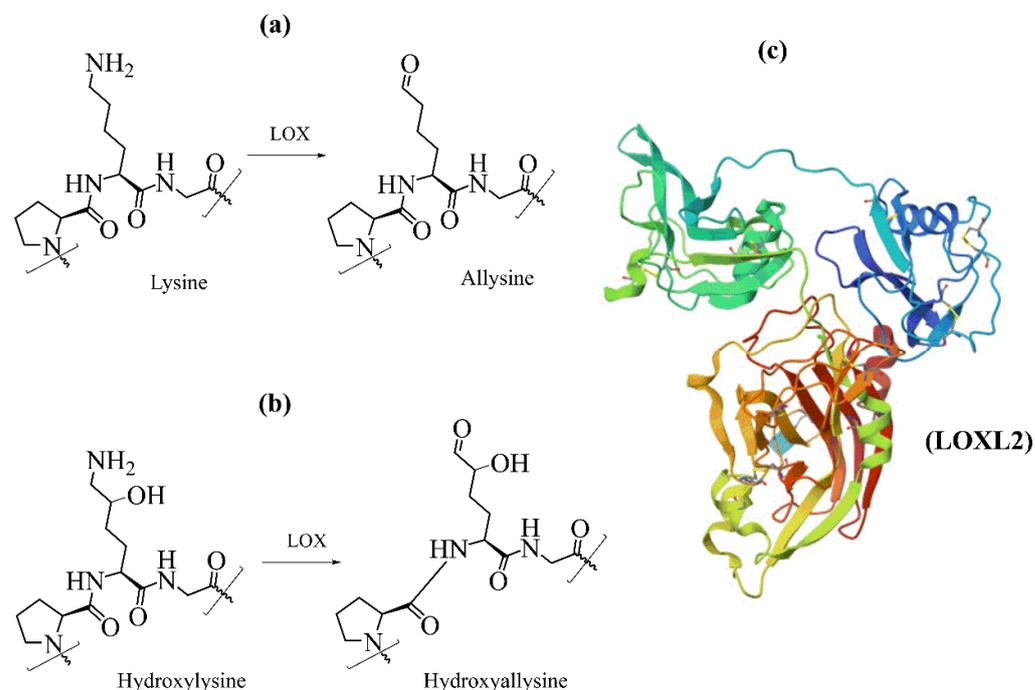


Figure 6. Schematic representation of collagen-specific oxidative deamination oxidation by lysyl oxidase (LOX) of (a) lysine and (b) hydroxylysine and (c) the 3D structure of LOXL2, PDB ID code 5ZE3 [105].

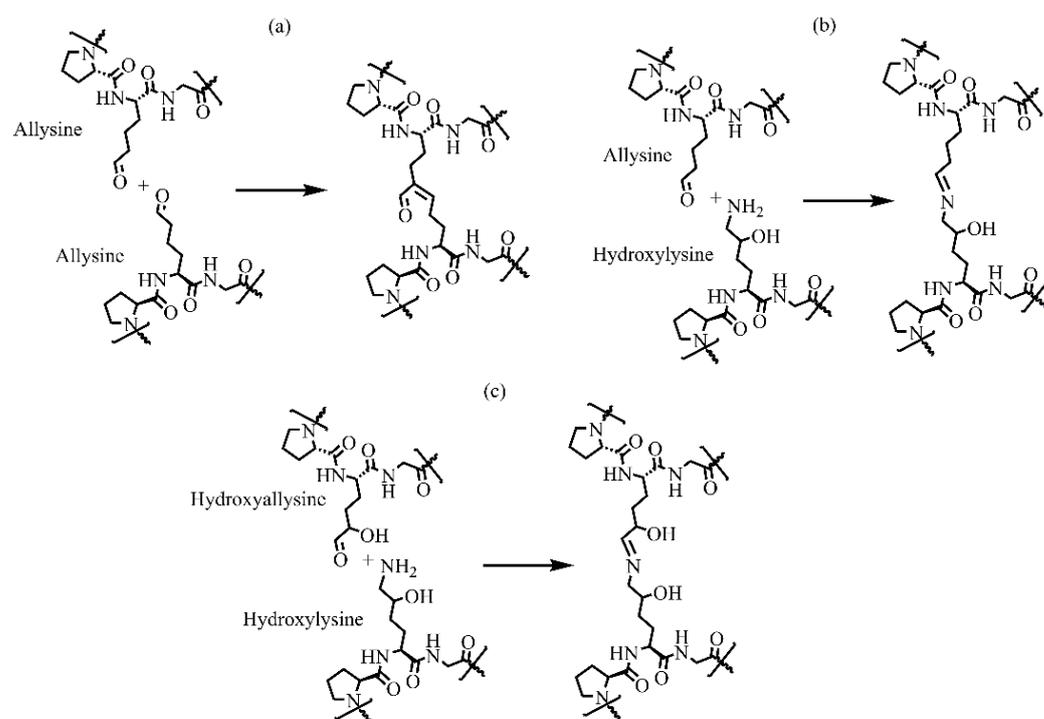


Figure 7. Schematic representation of three major physiological intramolecular cross-links by condensation between (a) two allysines forming an aldol condensation product, (b) allysine and hydroxylysine forming dehydro-hydroxylysinonorleucine, and (c) hydroxyallysine and hydroxylysine forming dehydro-hydroxylysinohydroxynorleucine.

6. Glycosylation of Collagens

Collagens belong to the group of human secretory proteins that are modified by adding carbohydrate molecules using two different mechanisms: enzymatic (glycosylation) and non-enzymatic (glycation). Due to its association with aging and diabetes, glycation has become one of the most studied processes. Many papers have been published, and several reviews discussed the progress in understanding glycation's mechanism and biological consequences [107–112]. It is beyond the scope of this review to describe glycation in detail and reference available literature. Therefore, only a very brief overview of glycation will be given. Less attention has been devoted to the second mechanism, glycosylation, than glycation [21,113]. The aim is to provide the reader with an overview of the recent progress in this area.

6.1. Glycation

The reaction of glucose in the aldehyde form (and several other monosaccharides presented in the blood) with lysine and arginine residues in proteins is called glycation. The formed glucosyl-lysine intermediates are stabilized via further reactions, including Schiff bases and Amadori products. This leads to complex advanced glycation end products (AGEs), including intra and intermolecular cross-links [107,110,114]. Glycation is a slow process since the initial reaction of glucose with lysine requires the conformational change in the sugar from the dominant cyclic form to the open-chain form. Collagens are very long-lived proteins and, therefore, are susceptible to glycation.

The physiological cross-linking of collagen is carried out enzymatically by lysyl oxidases (LOX), located at the ends of tropocollagens, modulating optimal properties and stabilizing the different supramolecular structures of collagen [104]. On the contrary, glycation triggers pathological cross-linking located randomly in the helical regions of fibrous collagens that increases with age. It has been shown that glycation might significantly impair LOX-controlled cross-linking [115]. AGEs cross-linking alters collagen physical properties (fibril stiffness) and influences intermolecular interactions with other ECM components, greatly influencing collagen's biological function. Collagens are relevant ligands of integrins [116], and their altered fibril structure likely affects the binding to integrins and thus impairs integrin outside-in and inside-out signaling. Although the glycation of collagen plays an essential role in aging and diabetes [107,108,110], it has also been implicated in various severe diseases [107,109], among others, in vascular diseases [109,117,118], and cancer [119–122].

6.2. Enzymatic Glycosylation

Glycosylation is the most abundant and perhaps the most structurally diverse post-translational modification of proteins consisting of several sequential steps in the endoplasmic reticulum and Golgi apparatus that involve numerous glycosyltransferases and glycosidases [123–125]. Glycosylation proceeds through several different pathways that reflect the presence of glycosyltransferases (GTs) and their substrates in the cell. Glycosyltransferases synthesize glycan structures by catalyzing the transfer of monosaccharide residue from an activated donor nucleotide sugar to the hydroxyl group of an acceptor. Based on the similarity of their primary sequences, GTs were clustered into 117 families [126,127] and registered in the Carbohydrate Active Enzyme Database (CAZy, available at <http://www.cazy.org>).

An analysis of known crystal structures revealed the existence of three general folds for GTs: GT-A, GT-B, and GT-C [128–131]. From a mechanical point of view, the catalytic reaction is quite complex. It requires the presence of a donor nucleotide sugar in the adequate micro-concentration and an acting glycosyltransferase in a Golgi compartment. Also, the working glycosyltransferase must be sufficiently expressed in the productive conformation that accommodates the donor and acceptor in the reaction Michaelis complex [124]. In sugar-nucleotides, the sugar residue is linked to UDP by the α -glycosidic linkage (α -anomer), and an attachment of the transferred monosaccharide residue to an

acceptor can be achieved by the α - or β -glycosidic bond. Based on the resulting linkage, glycosyltransferases are classified into retaining or inverting enzymes [130].

The produced glycans show enormous structural diversity [132], allowing for adjusting their 3D structure to biological function [131,133]. In the case of collagens, two different glycans have been observed: the highly conserved *N*-glycan in the C-terminal propeptide domain and *O*-glycans on hydroxylated lysine residues. Glycosylation on collagens occurs in the Golgi apparatus and ER before triple helix formation.

6.2.1. *N*-Glycan

The presence of *N*-glycan on collagen has been observed in the C-terminal propeptide domain [134–142]. These studies also revealed that glycan is a high mannose oligosaccharide containing two *N*-acetylglucosamine (GlcNAc) and six to nine mannose residues. *N*-glycosylation is a common PTM occurring in most secretory proteins [143–146]. The glycosylation of proteins starts in ER with the biosynthesis of oligosaccharide GlcNAc₂Man₉Glc₃ linked to dolichol phosphate. It continues by transferring oligosaccharide en bloc to protein by forming a β -1-*N*-linkage between GlcNAc and the Asn side chain. Then, the repertoire of glycosidases and glycosyltransferases in the ER and Golgi apparatus synthesize a massive variety of *N*-glycans that differ in composition.

The structure of collagen *N*-glycan was determined by 500 MHz ¹H NMR spectroscopy [147]. The isolated and purified glycans were a heterogeneous mixture of high-mannose type glycans [147], as shown in Figure 8. The one carbohydrate unit contains two *N*-acetylglucosamine residues and between six and nine mannose residues (GlcNAc₂Man₆₋₉). The percentages on terminal residues (shown in red) in Figure 8 refer to the relative occurrence of these residues in the mixture. This implies that after the transfer of oligosaccharide GlcNAc₂Man₉Glc₃ from dolichol phosphate to the Asn of collagen C-terminal propeptide, only several glycosidases are involved in the final shaping of the glycan structure. In the ER, the first three glucose residues and one mannose residue are trimmed by α -glucosidases α -Glc-I and α -Glc-II and by the ER α -mannosidase (α -Man), respectively [143,145]. Then, the properly folded collagen with linked glycan Glc₂Man₈ is transported to the Golgi apparatus for further processing by Golgi α -mannosidase (α -Man-I), leading to high mannose *N*-glycan GlcNAc₂Man₆ [148]. The NMR structure suggests processing of the *N*-glycan at the C-terminal propeptide finishes here. The observed heterogeneity likely reflects the influence of cellular and micro-environmental factors affecting glycosylation efficiency. In contrast to collagens, *N*-glycans maturation in other glycoproteins continues with the concerned action of glycosyltransferases and glycosidases forming branching, elongation, and capping [144,145].

Glycans are integral to protein function and regulate their diverse biological functions in physiological and pathological events by interacting with lectins. The *N*-glycosylation motif is highly conserved within the C-terminal domain of fibrillar pro-collagens [134,135,137,140], which suggests that the glycan has an important structural or biological function. However, the C- and N-terminal pro-collagen domains are trimmed as soon as fibrillar pro-collagen is secreted into extracellular space. This suggests that the *N*-glycans have a somewhat more relevant function inside the cell than an extracellular function. Although the function of the *N*-glycan inside the cell is well described (Helenius and Aebi 2001, Helenius and Aebi 2004), the function of the collagen *N*-glycan remains incomplete.

6.2.2. *O*-Glycans

The first mention of the carbohydrate molecules in collagens dates almost 90 years ago, when glucopyranose and galactopyranose in a 1:1 ratio were observed in the collagen of animal skin [149]. In the next two decades, further characterizations of saccharides in collagen were determined, and it was shown that these saccharides are bound to the peptide by a glycosidic linkage [150,151]. Studies on carbohydrate contents in various collagens have revealed two different saccharides on hydroxylysine, namely a disaccharide α -D-glucopyranosyl-(1 \rightarrow 2)- β -D-galactopyranose or a monosaccharide β -D-

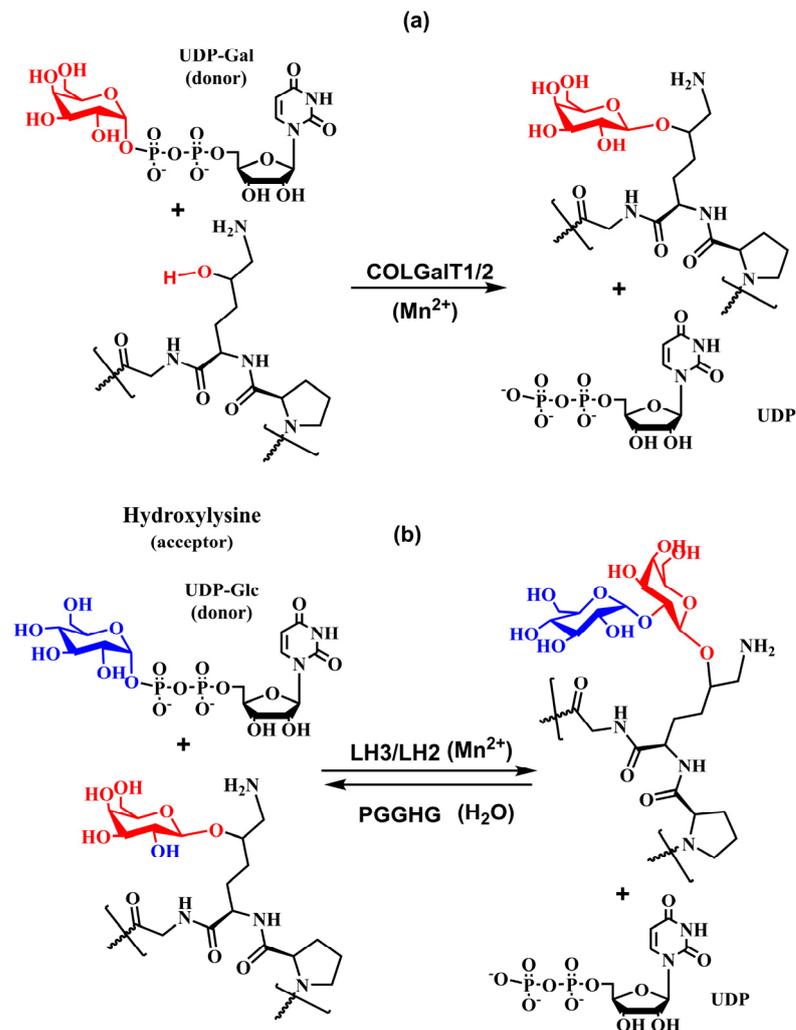


Figure 9. Schematic diagram of the enzymatic reaction catalyzed by (a) COLGALT1/2 and (b) LH2 and LH3.

Site-directed mutagenesis has been used to examine the presence of the DxD motif in COLGALT1 [162], and three DxD motifs were localized in these enzymes, one in the N-domain and two in the C-domain. In COLGALT1, only one DxD in the C-domain was not required for the galactotransferase activity. Also, it was suggested that the UDP-Gal donor and the metal cofactor bind to the DxD motif in the N-domain [162]. The variable DxD motif is the characteristic feature of enzymes with GT-A fold [124,129,131,163]. Two $\alpha/\beta/\alpha$ Rossmann-like folds create the active site containing a DxD motif [128]. The enzymes of the GT-A superfamily have two distinct donor and acceptor binding sites with a crucial role played by a disordered loop that changes conformation after binding an acceptor [164]. The kinetics measurements of apparent K_m of COLGALT1 and COLGALT2 provided values of 13.6 g/liter and 9.8 g/liter for the bovine type I collagen acceptor and 18.77 μM and 33.53 μM for the UDP-Gal donor, respectively [97]. The sequence alignment of both human COLGALT revealed a 63% identity [97], suggesting that the same DxD sequence is responsible for their activity. From a chemical point of view, COLGAT enzymes are the metal-ion-dependent (Mn^{2+}) inverting glycosyltransferases, in which the formation of a new glycosidic linkage between anomeric carbon (C1) of galactose from the donor UDP-Gal and 5-oxygen from the acceptor collagen Hyl proceeds with the inversion of stereochemistry at the anomeric carbon C1. Experimental and theoretical calculations support an $\text{S}_{\text{N}}2$ -like direct-displacement mechanism for inverting glycosyltransferases [124,128].

The analogy with inverting enzymes having GT-A fold leads to the tentative mechanism shown in Figure 10, where the DxD motif binds the bivalent ion (Mn^{2+}). The chelating bivalent metal serves as the donor's active site anchor, which is stabilized by a hydrogen bond between COLGAT amino acids and hydroxyls from ribose. The enzyme also provides a catalytic base that deprotonates attacking hydroxyl from the acceptor for the reaction to proceed [124,130].

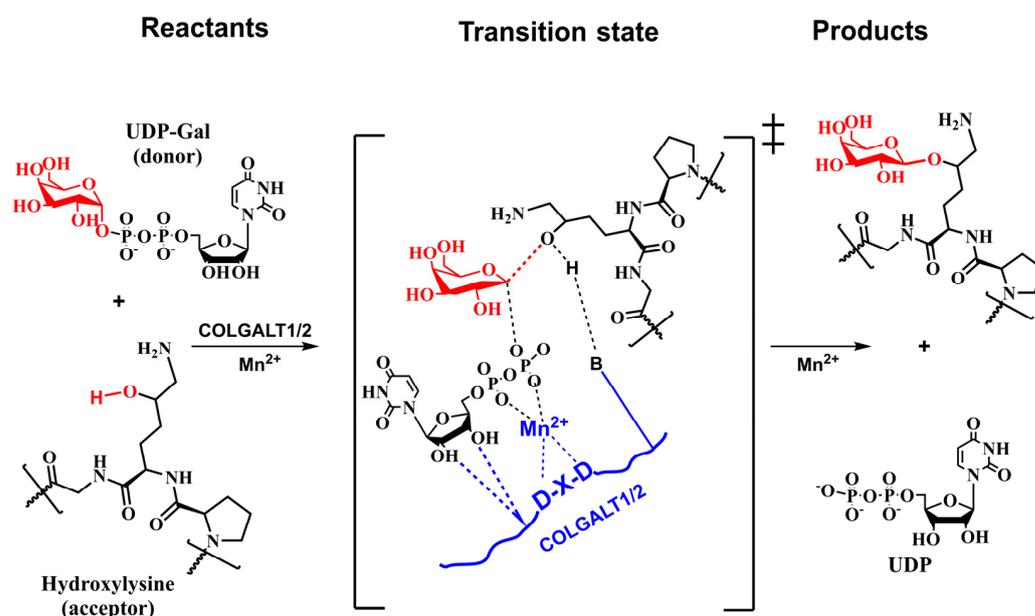


Figure 10. Schematic representation of the tentative SN_2 -like direct-replacement catalytic mechanism of inverting COLGALT1/2 transferases based on the analogy with inverting glycosyltransferases [124,165].

The structure of COLGAT enzymes has not been solved yet. However, the recently developed deep-learning method AlphaFold [166,167] generated, among others, the homology models of COLGALT1 and COLGALT2, as shown in Figure 11a,b. Both AlphaFold models have a high (>89%) per-residue model confidence score (pLDDT).

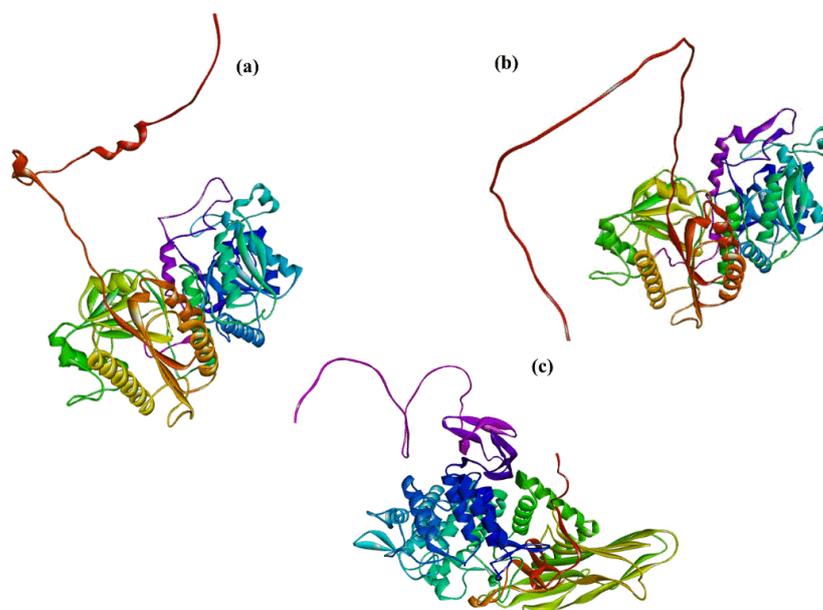


Figure 11. The AlphaFold homology models with per-residue model confidence score pLDDT > 89% of (a) COLGALT1 (AF-Q8NBJ-F1), (b) COLGALT2 (AF-Q8IYK4-F1), and (c) PGGHG (AF-Q32M88-F1).

Surprisingly, no glucosylation activity has been detected for two galactosyltransferases [97,158]. Conversely, the glucosylation enzymatic activity has been assigned to lysyl hydroxylase LH3 [95]. It was also shown that the sub-cellular locations of LH3 and COLGALT1 in ER are comparable [158].

6.4. Lysyl Hydroxylases

For a long time, it was assumed that of three isoforms of LH in humans, LH3 is the only one exhibiting glycosyltransferase activity. However, a recent study [92] discovered glycosyltransferase activity in LH1 and LH2 isoforms, although this activity was higher in the LH2b isoform. The authors used a luciferase assay to follow a release of the reaction product, UDP, from the reaction of the donor UDP-Glc with synthetic amino acids or deglycosylated type IV collagen. In the assay, all LH showed detectable GlcGalHyIT activities that were abolished by a mutation of amino acid residues chelating Mn^{2+} [92]. To support these findings, the authors measured the binding affinity of UDP-Glc to LH2 isoforms microscale thermophoresis, which led to the value of $K_d = 10 \mu M$ for LH2b. A comparison of domains bearing glycosyltransferase activity has shown a 60% identity between LH3 and LH1 and a 57% sequence identity between LH3 and LH2. The LH1, LH2, and LH3 active sites possess D-X-X-D, E-X-X-D, and D-X-X-D motifs, respectively. They also generated a 3D homology model of LH2b using the LH3 crystal structure (PDB 6FXT) as the template [25]. They suggested that a recognized loop may enhance the GlcGalHyIT activity by serving as the active site cap in LH2b.

It was suggested that the N-terminal of LH3 shows glycosyltransferase activity, whereas the C-terminal exhibits lysyl hydroxylase activity [88,94–96]. The *PLOD3* gene encodes the multifunctional enzyme LH3. The loss of glycosylation activity does not impair the lysyl hydroxylase activity. The lysyl hydroxylase activity depends on the dimerization of LH, whereas its dimerization is not required for the glycosylation activities [89,168]. Although LH3 catalyzes transfers of both the galactose and glucose residues to procollagen hydroxylated lysine, the transfer of glucose is more efficient [25,169]. The catalytic mechanism of the LH3 is fascinating. While in the galactosylation reaction, LH3 functions as the inverting GT (Figure 10); in glucosylation, the catalytic reaction proceeds via the retaining mechanism with the retention of stereochemistry at the anomeric carbon C1 of glucose.

On the other hand, both reactions are metal-ion-dependent. A kinetics experiment of glucosylation reaction (Myllyla, 1976) revealed an ordered mechanism with Mn^{2+} binding first, then binding UDP-Glc, and the acceptor is the last entering the active site. The product release proceeds the opposite way; the first leaves glucosylated procollagen, then UDP and Mn^{2+} , although Mn^{2+} may not leave during each catalytic cycle. Although the primary functional location of LH3 is in the ER, LH3 was also found secreted and active in extracellular space [170].

Recently, in a series of papers, the crystal structures of full-length human LH3 in complex with metal bivalent ion (Mn^{2+} and Fe^{2+}), donors (UDP-Glc and UDP-Gal), and various sugar-nucleotide analogs (PDB entries: 6FXK, 6FXM, 6FXR, 6FXT, 6FXX, 6FXY [25], 6TE3, 6TEC, 6TES, 6TEU, 6TEX, 6TEZ, [28] 6WFV [26], and 8ONE [27]) were solved. Crystal and solution data revealed for LH3 an elongated tail-to-tail homodimer structure of 200 kD and 20 nm long [25]. Two LH3 monomers are connected by two C-terminal domains and stabilized by hydrophobic contacts and electrostatic interactions between amino acids at the interface (Figure 12). It was observed that while the LH activity is associated with the dimer, GT activities were undisturbed in the LH3 monomer, thus supporting previously published data [89,168]. The solved crystal structure [25] revealed that LH3 comprises three domains. The first two N-terminal domains show Rossmann-fold resembling GT-A fold architecture of GTs [128–130]. The third C-terminal domain exhibits a double-stranded β -helix fold observed in dioxygenases [171]. The glycosyltransferase and LH activities were located in two 80 Å separated catalytic sites in the N-terminal and C-domain.

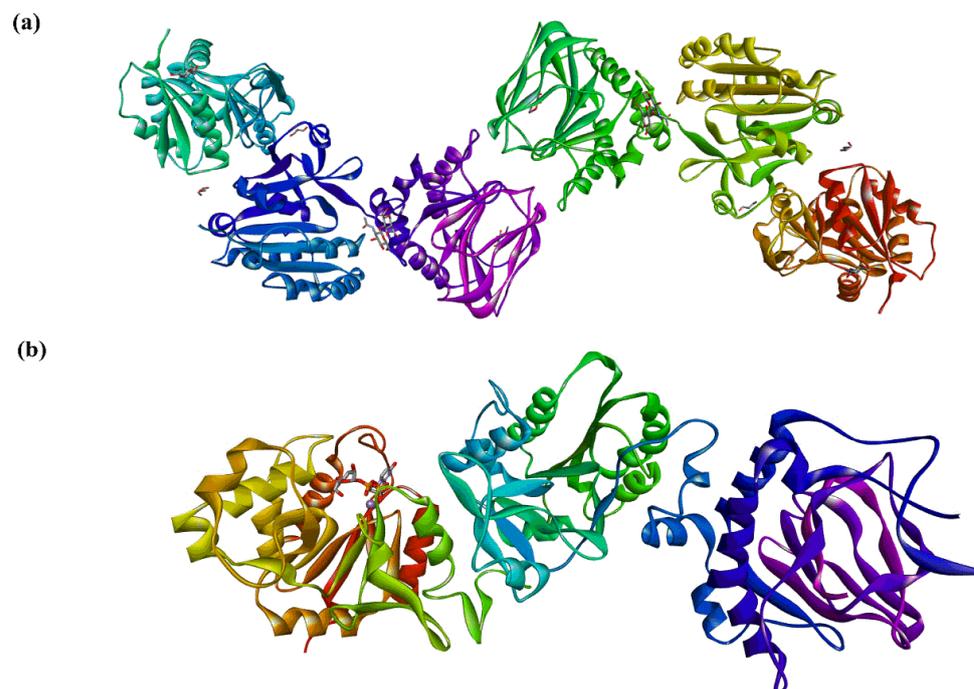


Figure 12. The 3D structure of human LH3 [25]: (a) tail-to-tail dimer observed in the crystal structure (PDB ID code 6FXK), and (b) three domains of the crystal structure of human LH3, the left is the catalytic glucosyltransferase domain; only the bound donor UDP-Glc and metal cofactor Mn^{2+} are shown (PDB ID code 8ONE).

The solved LH3 structures permitted the localization of the glycosylation active site and identified crucial amino acids involved in binding the bivalent metal ion and stabilizing the donor. In the complex of LH3 with UDP-Glc and Mn^{2+} , two aspartate residues, 112 and 115; histidine 253; and two water molecules chelate metal. Co-crystallization with donor either UDP-Gal or UDP-Glc led to the recognition of the LH3-specific D-X-X-D (Asp112-Ser113-Tyr114-Asp115), where Ser113 and Tyr 114 together with Val44 interact with ribose hydroxyls. Tyr 114 and Trp75 residues sandwich the uracil residue and stabilize its position by π - π interactions along with hydrogen bonds between Thr46 and uracil residue [25]. In the case of LH3, similarly to other GTs [128], the electron density of monosaccharide residue was weak, likely due to several Glc (or Gal) conformations, and it was insufficient to model these residues. Interestingly, the crystal structures of LH3 in complexes with UDP-Glc and UDP-Gal showed no difference.

The question is how the structurally identical active site can orchestrate two different catalytic reactions (inverting vs. retaining, Glc vs. Gal, and different acceptors) requiring different micro-environments. However, it was recently claimed that LH3 is not responsible for the transfer of Gal to Hyl, and the primary function of LH3 is glucose transfer [27,98,172]. Another hypothetical explanation might be that binding different acceptors will reshape the active site differently for each glycosylation reaction. The donor binding structuralizes the flexible loop Hyl72-Gly87 and changes the conformation of Trp145. The mutation of Trp145 into alanine completely abolished glycosylation [27,28], supporting a “gating” role in the active site suggested for Trp145 [25]. Glu141 was indicated as a potential catalytic nucleophile from amino acids in the neighborhood of a nucleotide sugar in the active site since the Glu141Ala mutation completely blocked the sugar transfer [27,28]. However, the distance between Glu141 oxygen and the anomeric oxygen in the LH3 crystal structure (PDB entry 8ONE) [28] ($>7 \text{ \AA}$) seems to be too long for this function.

Compared to the inverting [173] GTs, the catalytic mechanism of retaining GTs is more complex, and three different mechanisms were discussed [174]. A classical double-displacement mechanism involving a covalent glycosyl-enzyme intermediate was proposed for α -1,3-

galactosyltransferase [175,176] and for A and B α -1,3-galactosyltransferases [177,178]. The one-step front-face reaction (internal return-like or S_Ni) mechanism was proposed for the retaining GTs LgtC [179–181] and UDP-GalNAc polypeptide: GalNAc transferase-T2 [173] and α -1,3-galactosyltransferase [182,183]. Also, a slightly modified S_Ni mechanism by including a short-life oxocarbenium ion intermediate, the step-wise front-face reaction mechanism (step-wise S_Ni or ion-pair), was proposed [130] and used for the trehalose-6-phosphate synthase [184], α -1,2-mannosyltransferase [185], and the polypeptide UDPGalNAc transferase [186–188]. GTs utilize the S_Ni -like mechanism without a properly located nucleophile in the neighborhood of the anomeric carbon at the transferred sugar donor.

Experimental data characterized the mechanism as an ordered [159] and concerned mechanism [28] and, together with lacking a properly located nucleophile [27,28], indicated that LH3 might utilize the S_Ni -like mechanism for glucosylation of galactosylated-Hyl. The analogy with the proposed LgtC mechanism [179,180] is used to design a tentative mechanism for LH3, as shown in Figure 13.

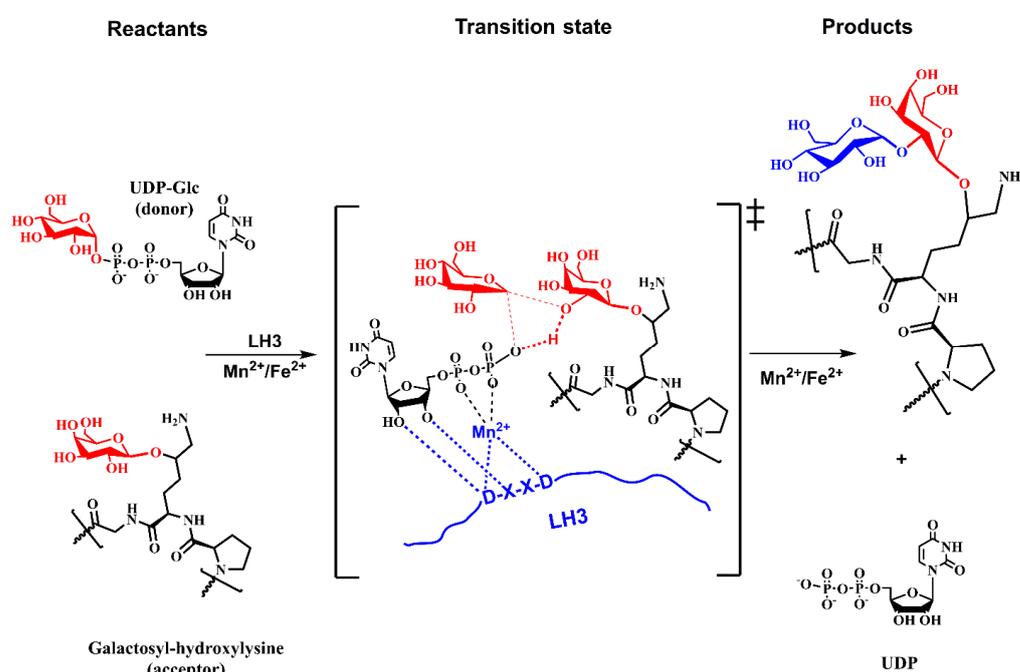


Figure 13. Schematic representation of the tentative S_Ni -like mechanism of the retaining glycosyltransferase LH3 based on the analogy with the LgtC mechanism [179,180].

Figures 10 and 13 show schematic representations of transition state models (TS) for tentative mechanisms for COLGAT and LH3 enzymes. In both catalytic reactions, the most pronounced TS characteristics represent a deformation of the transferred sugar that adopts half-chair/envelope conformation with shortened ring C-O bond; the anomeric carbon has sp^2 hybridization and an oxo-carbenium character; the breaking glycosidic bond is longer than the standard C-O bond; and the forming glycosidic C-O linkage is considerably shorter than this distance in the reaction starting complex. How realistic the tentative mechanisms proposed for COLGALT and LH3 enzymes (Figures 10 and 13) are remains to be determined.

6.5. Collagen Glucosidase

Protein-glycosylgalactosylhydroxylysine glucosidase (PGGHG; EC3.2.1.107) decoded by ATHL1 gene cleaves glucose from disaccharide unit α -D-glucopyranosyl-(1 \rightarrow 2)- β -D-galactopyranose linked to hydroxylysine residues of collagen [189–191]. The PGGHG enzyme belongs to the glycoside hydrolase family 65 (GH65), which cleaves α -glucosidic linkages in oligosaccharides [127]. Although PGGHG is specific for disaccharide α -D-

glucopyranosyl-(1→2)-β-D-galactopyranose linked to hydroxylysine, it can hydrolyze kojibiose (α-D-glucopyranosyl-(1→2)-β-D-galactopyranose) but with the five-times lower activity [189], implying a role played by collagen in the catalytic reaction. The site-directed mutagenesis of human PGGHG indicated three carboxyl residues, Asp301, Glu430, and Glu574, that are assumed to constitute the catalytic site of PGGHG [192]. The first crystal structure of a glycosidase from the GH65 family, the enzyme from *Flavobacterium johnsoniae* (FjGH65A), was recently solved [193]. The FjGH65A enzyme catalyzes the hydrolysis of kojibiose and other α-(1→2)-glucosides. The kinetics investigation of the catalytic mechanism revealed that FjGH65A is the α-(1→2)-glucosidase and utilizes an inverting mechanism. Although the 3D structure of PGGHG has not been solved, homology models have been generated by the developed deep-learning method AlphaFold [166,167] with a high (>89%) per-residue model confidence score and is shown in Figure 11c.

7. Functional Role of Collagen Glycosylation

Collagen PTMs influence their structure, deposition, and amounts and affect cellular processes [6,14,194]. Alternations of collagen structures caused by mutation in genes encoding collagens can influence several pathological conditions. The following will discuss only the effects associated with enzymatic glycosylation. Various biological functions of collagens critically depend on the collagen interactions with numerous ligands. Collagens bind integrins and other extracellular receptors. Collagen glycan components, glucose and galactose residues, have two sides: hydrophilic and hydrophobic. The top surface containing the hydroxyl group and the ring oxygen is more polar (hydrophilic), whereas the bottom surface, which consists of hydrogens, is less polar (hydrophobic). The bottom surface is involved in interactions with hydrophobic amino acids of the protein. Hydroxyl groups usually point to the solvent and interact with surrounding molecules. Molecular modeling studies provided insights into how glycans affect the collagen structure and interactions with surrounding molecules. Molecular modeling has been performed to investigate the role of Hyl glycosylation on interactions of α2β1 and α3β1 integrins with a triple helical model of type IV collagen [195]. The results show that glycan on Hyl significantly inhibits collagen binding to integrin, most likely due to the proximity of galactose to the key interacting residues. Other molecular dynamics simulations were used to investigate how Hyl glycosylation affects the structure of type I collagen [196]. Simulations of the collagen molecular segment in an aqueous solution have shown that (a) the Hyl glycosylation does not impair triple helix structure, (b) glycosylation modulates the peptide backbone conformation and (c) hydrophilic faces of monosaccharide residues are exposed to solvent and hydrophobic surfaces point to a hydrophobic part of collagen.

Understanding the functional role played by collagen glycosylation in many cellular processes requires determining both the sites of glycosylation and the glycan structures associated with each site. This task is quite challenging, and standard structural methods, such as NMR and X-ray crystallography, are very limited for collagen. Due to recent progress, mass spectrometry (MS) has emerged as the primary tool for glycan analysis [197]. Using liquid chromatography–tandem mass spectrometry (LC-MS-MS) provides structural information with high sensitivity [198]. LC-MS-MS application offers unique semiquantitative information on collagen glycosylation that can be correlated with tissue and pathophysiological information.

7.1. N-Glycan Functions

The N-glycosylation site at the C-propeptide domain is conserved in collagens, which, together with its bulky structure, indicates that it should have essential biological functions. Although the crucial biological roles of N-glycans on proteins have been well described [133], it is surprising that the biological functions of the collagen N-glycan are still inconclusive. Docking a model N-glycan into the solved crystal structure of the human C-terminal propeptide domain homotrimer [199] has shown that the N-glycan is located on the outer face of the β-sheet and exposed to the solvent [137]. This supports the assumption

that the *N*-glycan is functional. However, the C- and N-terminal pro-collagen domains are trimmed as soon as fibrillar pro-collagen is secreted into extracellular space. This suggests that the *N*-glycans have a somewhat more relevant function inside the cell than within the extracellular matrix. Surprisingly, it was found [140] that the non-glycosylated variant of C-terminal propeptide folds appropriately, similarly to the wild-type, suggesting that *N*-glycan might play a role as the signal for the clearance of trimmed C-terminal propeptides. Also, it was proposed that receptors recognize *N*-glycan for signaling to regulate collagen synthesis [200,201] or degradation [202].

Furthermore, it has been observed that the misfolding of the C-terminal pro-collagen domain is aggravated when the *N*-glycan is missing [137]. Recently, another functional proposal for *N*-glycan appeared. It was found that proper folding and secretion of procollagen in misfolding mutation requires the presence of *N*-glycosylation, and it was suggested that the critical intracellular function of *N*-glycan is to enable appropriate folding under proteostatic challenges. An absence of the *N*-glycan disturbs procollagen quality control, leading to intracellular aggregation and reduced secretion [137]. Despite the evident progress, further studies are needed to understand the functional role of collagen *N*-glycan fully. Still, available experimental data shows that *N*-glycans modulate various processes during collagen biosynthesis.

7.2. *O*-Glycan Functions

The functional role of collagen *O*-glycosylation has not been explained satisfactorily up to now. However, the progress during the last three decades indicates that *O*-glycosylation is involved in various critical structural, biological, and pathological processes. The extent of *O*-glycosylation modulates the stability of collagens against proteolytic degradation, collagen fibrillogenesis, cross-linking, secretion, assembly, and the distribution of collagens. The *O*-glycosylation also mediates interactions with non-collagenous molecules, and as the ligand for ECM components regulates specific signaling pathways, the *O*-glycosylation modulates fibril assembly and affects the morphology of fibrils [58]. It has been shown that thin fibrils were observed in collagens containing high contents of hydroxylysine. In contrast, thick fibrils have less glycosylated hydroxylysine, likely due to lesser steric interactions. The results indicate that regulation of the extent of lysine hydroxylation and hydroxylysine glycosylation may regulate the formation and morphology of the collagen fibrils. Also, fewer glycosylated collagens self-associate faster [98,203]. Since the quality and quantity of *O*-glycosylation affect the behavior of collagen in the ECM, it is evident that an appropriate folding rate is vital. It has been found that the fibrils formed from collagen with a greater extent of glycosylation had smaller diameters than fibrils formed from standard type I collagen, supporting suggestions that glycosylation alters the morphology of the formed fibrils [204]. However, another study did not observe any relationship between the level of glycosylation and the diameter of formed fibrils [205].

The *O*-glycosylation has been shown to play an essential role in forming basement membranes [172]. In that study, a loss of *O*-glycosylation in mice impaired the localization of type IV collagen and thus disrupted the formation of the basement membrane during mouse embryogenesis and led to lethality. Also, the lack of *O*-glycosylation interrupts the tetramerization and secretion of type VI collagen and the secretion of type IV collagen, implying that *O*-glycosylation is required for supramolecular assembly and collagen interactions with collagens and other proteins [206]. Together, results show that the *O*-glycosylation is fundamental for the ECM and basement membrane structure, suggesting that LH gene mutation may trigger ECM disorders and muscular diseases [207].

Different findings were observed on the role of *O*-glycosylation in melanoma cell invasion. In the study of melanoma cell adhesion [24], it has been observed that collagen glycosylated in the $\alpha 1(\text{IV})1263\text{--}1277$ region inhibits interactions with cell-surface glycoprotein (CD44), leading to a significant decrease in their adhesion. *O*-glycosylation in the $\alpha 1(\text{IV})531\text{--}543$ region also inhibits melanoma cell binding to $\alpha 3\beta 1$ integrin [195]. On the other hand, the binding of the CTLD-2 lectin domain of Endo180 (recycling transmembrane

glycoprotein, collagen binding receptor) to O-glycosylated collagen increases the endocytic efficiency of Endo-180 [208]. Since Endo-180 participates in collagen turnover during the malignant process, it enhances tumor cell invasion. Recently, an investigation of the role of collagen O-glycosylation in regulating carcinogenesis and metastasis by regulating cancer stem cell (CSC) phenotypes in lung cancer was carried out [209]. The results suggested that O-glycosylation creates a favorable microenvironment for the generation and survival of lung CSC.

Though the extracellular location is not clarified yet, some data implicate that LH3 is associated with cell growth and viability throughout cell–matrix interactions or receptor activation [210].

Experiments with the collagen-induced arthritis (CIA) model indicate that PTMs may play a vital role in autoimmune diseases by developing the T cell autoimmune response. In the CIA model, mice were immunized with recombinant type II collagens. The comparison shows that glycosylated collagen induces a higher level of arthritis than non-glycosylated collagen. The results showed that the extent of glycosylation affects the immune response to collagen [211].

8. Glycosylation Diseases

Collagens interact with many ECM components, influencing many cellular processes. Alterations in collagen structures can, therefore, affect various pathological processes, including genetic disorders, fibrosis, skin disorders, and cancer. More than 1300 collagen mutations associated with human diseases have been characterized [32]. Many excellent reviews on collagen diseases have been published where more details and further references can be found. See, for example, references [14,15,31,194,212–216]. Here, only collagen glycosylation-related diseases will be discussed. Recent findings that all three LH enzymes have collagen GlcGalHylT activity indicate that mutations in *PLOD1* and *PLOD2* genes may contribute to impaired hydroxylation and aberrant collagen glycosylation. Therefore, these are also mentioned here.

Undoubtedly, glycosylation is a crucial PTM of collagen and affects collagens' structural and functional behavior. However, understanding the role of glycosylation in collagen diseases is still in its infancy and lags behind other collagen-related diseases. Several aspects of deciphering glycosylation-related diseases—the development of GTs bioassays with proper precision, accuracy, reproducibility, and operational simplicity [217]; the overlapping locations of three LHs and two GTs often complicate their characterization [113]; and the quantitative characterization of hydroxylation vs. glycosylation and their mutations [198]—have made this task challenging. However, recent advances in glycomics and MS of glycans are progressively disclosing the critical roles of glycosylation in diseases and have been reported only recently [21,58,59].

Various anomalies related to small blood vessels in the brain are called small-vessel diseases (SVDs) and are associated with several pathological processes [218]. Among clinical signs are ischemic strokes, a more significant proportion of intracerebral hemorrhages and causes of death, and long-term severe disability. Mutations in *Col4a1* [219,220] and *Col4a2* [221] genes are associated with SVDs. These mutations lead to the impaired secretion and intracellular accumulation of collagens COL4A1 or COL4A2 [220,221], affecting the quality of the basement membrane. Recently, in two unrelated patients with SVDs, biallelic variants in the *COLGALT1* gene encoding ColGalT1 have been discovered [222]. Both patients showed clinical symptoms similar to those in *Col4A1/Col4A2* disease [219]. The study showed that impaired GalHylT activity decreases COL4A1 production. As a result, the secretion of type IV collagen into extracellular space was disrupted. This might contribute to basement membrane fragility, causing cerebral small vessel aberration similar to the pathological characteristics of *Col4A1/Col4A2* mutants.

Connective tissue disorder (CTD) due to LH3 deficiency (bone fragility–contractures–arterial rupture–deafness syndrome) is a rare genetic disease caused by a lack of LH3 activity. The disease involves multiple tissue and organ involvement, including skeletal

abnormalities; ocular involvement; and hair, nail, and skin anomalies. Patients also present intrauterine growth retardation, facial dimorphism, and joint flexion contractures. Growth and developmental delay, bilateral sensorineural deafness, friable diaphragm, and later-onset spontaneous vascular ruptures are additional reported features [223]. In two patients with CTD clinical symptoms, a significantly reduced GlcGalHylT activity was observed in serum [224–226]. Analyses revealed two heritable *PLOD3* mutations, one from the father and the second from the mother. The locations of mutations were found in regions with glycosyltransferase and lysyl hydroxylase activities. The observed data show mainly type II collagen disorders, suggesting the molecular basis of the CTD syndrome. The consequences of small changes in nucleotides of the *PLOD3* gene are dramatic, with all activities of LH3 considerably reduced [224].

Mutations in the genes encoding LH cause several severe or lethal diseases. The Ehlers–Danlos syndrome (EDS) and Nevo syndrome [227] are a group of heritable connective tissue disorders characterized by skin hyperextensibility, articular hypermobility, and tissue fragility. The significant characteristics of kyphoscoliotic-type EDS are severe muscle hypotonia, generalized joint laxity, scoliosis, and scleral fragility and the rupture of the ocular globe [228,229]. At least 20 different mutations have been identified in *PLOD1* [230] and *PLOD3* [224], causing a deficiency in lysyl hydroxylase.

Bruck syndrome is an autosomal recessive syndrome consisting of bone fragility and congenital joint contractures. According to the genotype, it has been classified into types 1 and 2 involving mutations of the *FKBP10* and *PLOD2* collagen genes, respectively [231–233]. For example, data from two siblings, a six-year-old male and a five-year-old female, with Bruck syndrome type 2 of non-affected parents, were recently published [234]. The male has experienced more than 45 fractures, developed severe scoliosis, and has debilitating flexion contractures. The female has minimal flexion contractures, a history of 15 fractures, and severe scoliosis. *PLOD2* mutations result in a dramatic underhydroxylation of collagen telopeptide Lys, causing a decrease in the pyridinoline cross-linking of type I collagen.

Epidermolysis bullosa (EB) simplex [207,224,235] is a group of rare heritable diseases that cause the skin to be fragile and to blister easily. Tears, sores, and blisters happen when something rubs or bumps the skin. Mutations of around ten genes among *PLOD3* [207,224,226,236] inherited from parents cause most forms of epidermolysis bullosa. In patients, the pathological mutation in *PLOD3* led to a decreased expression of LH3 [236] and type VII collagen in the skin, resulting in reduced glucosylation [226]. It was suggested that the decrease in GlcGalHylT activity correlates with the severity of phenotype and the resulting abnormalities in the organization of the extracellular matrix [207].

Stickler syndrome is a group of hereditary conditions characterized by a distinctive facial appearance, eye abnormalities, hearing loss, and joint problems. Stickler syndrome is a genetic condition affecting connective tissues in the face and joints, leading to problems with vision, hearing, and movement problems. This hereditary condition is usually diagnosed in babies and children, and early treatment leads to a positive outcome [237]. Stickler syndrome is caused by mutations in the *COL2A1* and *COL1A1* genes that are associated with the formation of types II, IX, and XI collagens and affect their structure and properties [238,239], and also by mutation of the *PLOD3* gene [224,225]. Affected persons show a decreased glycosyltransferase activity. Although clinical symptoms best describe Stickler syndrome, they also exhibit variable features of EDS and EB [225].

In addition to heritable connective tissue diseases, impairments of glycosylation are associated with various common diseases, and some will be briefly mentioned below. Covalent intermolecular cross-links are the final step in collagen biosynthesis and affect many biomechanical functions of collagen [58]. This process starts after the secretion of collagen into extracellular space by the action of LOX and is followed by spontaneous interactions between neighboring chains (see paragraph on LOX). This process is critical for the properties of fibrils, influences fibrinogenesis and matrix mineralization, and is affected by the glycosylation of collagen [98,99].

An increase in the glycosylation of the collagen triple helix has been shown in degenerative joint disease (osteoarthritis cartilage) [240]. A comparison of lysyl hydroxylation and pyridinoline cross-linking of normal and degenerative knee showed that degenerative cartilage contained significantly more hydroxylysine residues per collagen molecule than healthy cartilage from the same donor.

Inflammation or irritation of a tendon (tendinitis) occurs in the elbow, knee, shoulder, and hip. In a normal tendon, type I collagen accounts for almost 95% of the total collagen [241]. In tendinitis, the previously functional collagen is replaced by aberrant collagen with a lower ratio of pentosidine and higher ratios of hydroxylated lysine residues in collagen cross-links [242,243]. More hydroxylated lysine residues were also present in Achilles tendinosis [244].

Losing bone mineral density weakens bones and can lead to osteoporosis (osteopenia). It was found that the degree of the lysyl hydroxylation of both α -chains of type I collagen showed a significant inverse correlation with the trabecular bone volume. Also, the extent of the glycosylation of type I collagen correlated with both the level of lysylhydroxylation and the degree of osteopenia [205].

The extent of glycosylation has been found to influence metastatic cancer. Collagen cross-link accumulation increases stromal stiffness and stimulates tumor cells' invasive properties. For example, tumor stroma contains higher levels of hydroxylysine aldehyde-derived collagen cross-links and lower levels of lysine aldehyde-derived cross-links compared to normal lung tissues due to the higher expression of LH2, which indicates that LH2 enhances the collagen metastatic properties of tumor cells and functions by controlling the relative abundance of hydroxylysine aldehyde-derived collagen cross-links in the tumor stroma [245]. The role of LH2 in metastasis is supported by pharmacologic inhibition of the *PLOD2* enzymatic activity that suppresses metastases. These data indicate that *PLOD2*-dependent collagen modifications influence metastasis [246]. Also, the higher expression of the *PLOD1* and *PLOD2* genes was found in hypoxic breast cancer cells and is associated with an increased risk of mortality. It was suggested that *PLOD2* is critical for fibrillar collagen formation by breast cancer cells, increases tumor stiffness, and is required for metastasis to lymph nodes and lungs [247].

9. Concluding Remarks

Post-translational modifications are the essential characteristics of collagen, including proline and lysine hydroxylation, *N*-glycosylation, the *O*-glycosylation of Hyl, the oxidative deamination of Lys and Hyl residues in the telopeptide domains, and following intra and intermolecular covalent cross-linking. The hydroxylation of proline in positions 3 and 4 is crucial for the stability of the triple-helical structure, whereas the hydroxylation of lysine residues is a prerequisite for subsequent glycosylation. The glycosylation of collagen takes place in the ER before forming a triple-helical assembly. It is simple but unique since the attached *O*-glycan comprises either monosaccharide galactose or disaccharide α -D-glucopyranosyl-(1 \rightarrow 2)- β -D-galactopyranose. Collagen *O*-glycosylation commences by transferring the galactose residue from the donor UDP-Gal to the 5-hydroxyl group of Hyl by the β -glycosidic linkage. Two inverting galactosyltransferases, COLGALT1 and COLGALT2, catalyze this transfer. The second step is the transfer of the glucose residue from the donor UDP-Glc to the C2 hydroxyl of glucose by the α -glycosidic linkage through the catalytic action of a multifunctional enzyme lysylhydroxylase LH3 utilizing the retaining mechanism. The extent of glycosylation, glycosites, and the heterogeneity of glycan distribution (mono versus disaccharide) depend on the collagen types, tissue, cell type, maturation stage, and pathological condition and seem to modulate collagen structural organization and function in the ECM.

Although the functional role of collagen glycosylation and its involvement in diseases is not fully understood, it is clear that collagen *O*-glycosylation is fundamental for its correct assembly and ECM deposition. Many studies showed that collagen aberrant *O*-glycosylation is associated with connective tissue disorders, cancer development, and

metastasis. Collagens also play a crucial role in the properties of basement membranes, where their impaired deposition can cause pathological conditions. There are indications that O-glycosylations modulate the physiological cross-linking catalyzed by the enzyme lysyl oxidase. Primarily, four carbohydrate processing enzymes, namely COLGALT1, COLGALT2, LH3, and PGGHG, are involved together with lysyl hydroxylases in the biosynthesis of O-glycan, where a delicate balance between their actions defines the O-glycosylation pattern. Recent progress in the understanding of the catalytic mechanism of glycosyltransferases and glycosidases, together with progress in collagen glycosylation analysis by mass spectrometry and biochemical characterization, offer tools to decipher the details of collagen O-glycosylation and its micro- and macro-heterogeneity. Such a multidisciplinary approach can contribute to fully unraveling glycosylation's role in collagen's structure and function. This knowledge will provide a rational guide to treating O-glycosylation-related diseases and to the development of new therapeutic agents for cancer, fibrosis, and regenerative medicine.

Funding: This research received no external funding.

Conflicts of Interest: The author declares no conflicts of interest.

References

1. Karamanos, N.K.; Theocharis, A.D.; Piperigkou, Z.; Manou, D.; Passi, A.; Skandalis, S.S.; Vynios, D.H.; Orian-Rousseau, V.; Ricard-Blum, S.; Schmelzer, C.E.H.; et al. A guide to the composition and functions of the extracellular matrix. *FEBS J.* **2021**, *288*, 6850–6912. [[CrossRef](#)]
2. Frantz, C.; Stewart, K.M.; Weaver, V.M. The extracellular matrix at a glance. *J. Cell Sci.* **2010**, *123*, 4195–4200. [[CrossRef](#)] [[PubMed](#)]
3. Theocharis, A.D.; Skandalis, S.S.; Gialeli, C.; Karamanos, N.K. Extracellular matrix structure. *Adv. Drug Deliv. Rev.* **2016**, *97*, 4–27. [[CrossRef](#)]
4. Yue, B. Biology of the extracellular matrix: An overview. *J. Glaucoma* **2014**, *23*, S20–S23. [[CrossRef](#)]
5. Kular, J.K.; Basu, S.; Sharma, R.I. The extracellular matrix: Structure, composition, age-related differences, tools for analysis and applications for tissue engineering. *J. Tissue Eng.* **2014**, *5*, 2041731414557112. [[CrossRef](#)] [[PubMed](#)]
6. Dzobo, K.; Dandara, C. The Extracellular Matrix: Its Composition, Function, Remodeling, and Role in Tumorigenesis. *Biomimetics* **2023**, *8*, 146. [[CrossRef](#)]
7. Van der Rest, M.; Garrone, R. Collagen Family of Proteins. *FASEB J.* **1991**, *5*, 2814. [[CrossRef](#)] [[PubMed](#)]
8. Beachley, V.Z.; Wolf, M.T.; Sadtler, K.; Manda, S.S.; Jacobs, H.; Blatchley, M.R.; Bader, J.S.; Pandey, A.; Pardoll, D.; Elisseeff, J.H. Tissue matrix arrays for high-throughput screening and systems analysis of cell function. *Nat. Methods* **2015**, *12*, 1197–1204. [[CrossRef](#)]
9. Ricard-Blum, S. The collagen family. *Cold Spring Harb. Perspect. Biol.* **2011**, *3*, a004978. [[CrossRef](#)]
10. Mienaltowski, M.J.; Birk, D.E. Structure, physiology, and biochemistry of collagens. *Adv. Exp. Med. Biol.* **2014**, *802*, 5–29. [[CrossRef](#)]
11. Gelse, K.; Poschl, E.; Aigner, T. Collagens—structure, function, and biosynthesis. *Adv. Drug Deliv. Rev.* **2003**, *55*, 1531–1546. [[CrossRef](#)]
12. Wolff, D. Collagen Diseases. In *xPharm: The Comprehensive Pharmacology Reference*; Enna, S.J., Bylund, D.B., Eds.; Elsevier Inc.: Amsterdam, The Netherlands, 2008; pp. 1–9. [[CrossRef](#)]
13. Fidler, A.L.; Boudko, S.P.; Rokas, A.; Hudson, B.G. The triple helix of collagens—An ancient protein structure that enabled animal multicellularity and tissue evolution. *J. Cell Sci.* **2018**, *131*, jcs203950. [[CrossRef](#)]
14. Arseni, L.; Lombardi, A.; Orioli, D. From Structure to Phenotype: Impact of Collagen Alterations on Human Health. *Int. J. Mol. Sci.* **2018**, *19*, 1407. [[CrossRef](#)] [[PubMed](#)]
15. Romero-Ortuno, R.; Kenny, R.A.; McManus, R. Collagens and elastin genetic variations and their potential role in aging-related diseases and longevity in humans. *Exp. Gerontol.* **2020**, *129*, 110781. [[CrossRef](#)] [[PubMed](#)]
16. Jobling, R.; D'Souza, R.; Baker, N.; Lara-Corrales, I.; Mendoza-Londono, R.; Dupuis, L.; Savarirayan, R.; Ala-Kokko, L.; Kannu, P. The collagenopathies: Review of clinical phenotypes and molecular correlations. *Curr. Rheumatol. Rep.* **2014**, *16*, 394. [[CrossRef](#)] [[PubMed](#)]
17. Necula, L.; Matei, L.; Dragu, D.; Pitica, I.; Neagu, A.; Bleotu, C.; Diaconu, C.C.; Chivu-Economescu, M. Collagen Family as Promising Biomarkers and Therapeutic Targets in Cancer. *Int. J. Mol. Sci.* **2022**, *23*, 12415. [[CrossRef](#)] [[PubMed](#)]
18. Romer, A.M.A.; Thorseth, M.L.; Madsen, D.H. Immune Modulatory Properties of Collagen in Cancer. *Front. Immunol.* **2021**, *12*, 791453. [[CrossRef](#)] [[PubMed](#)]
19. Leikina, E.; Merts, M.V.; Kuznetsova, N.; Leikin, S. Type I collagen is thermally unstable at body temperature. *Proc. Natl. Acad. Sci. USA* **2002**, *99*, 1314–1318. [[CrossRef](#)] [[PubMed](#)]

20. Spiro, R.G. Characterization and quantitative determination of the hydroxylysine-linked carbohydrate units of several collagens. *J. Biol. Chem.* **1969**, *244*, 602–612. [[CrossRef](#)] [[PubMed](#)]
21. Hennes, T. Collagen glycosylation. *Curr. Opin. Struct. Biol.* **2019**, *56*, 131–138. [[CrossRef](#)]
22. Smolen, J.S.; Aletaha, D.; McInnes, I.B. Rheumatoid arthritis. *Lancet* **2016**, *388*, 2023–2038. [[CrossRef](#)]
23. Gabrielli, A.; Avvedimento, E.V.; Krieg, T. Scleroderma. *N. Engl. J. Med.* **2009**, *360*, 1989–2003. [[CrossRef](#)] [[PubMed](#)]
24. Lauer-Fields, J.L.; Malkar, N.B.; Richet, G.; Drauz, K.; Fields, G.B. Melanoma cell CD44 interaction with the alpha 1(IV)1263-1277 region from basement membrane collagen is modulated by ligand glycosylation. *J. Biol. Chem.* **2003**, *278*, 14321–14330. [[CrossRef](#)] [[PubMed](#)]
25. Scietti, L.; Chiapparino, A.; De Giorgi, F.; Fumagalli, M.; Khoriauli, L.; Nergadze, S.; Basu, S.; Olieric, V.; Cucca, L.; Banushi, B.; et al. Molecular architecture of the multifunctional collagen lysyl hydroxylase and glycosyltransferase LH3. *Nat. Commun.* **2018**, *9*, 3163. [[CrossRef](#)]
26. Guo, H.-F.; Tsai, C.-L.; Miller, M.D.; Philips, G.; Tainer, J.; Kurie, J. The Crystal Structure of a Collagen Galactosylhydroxylysyl Glucosyltransferase from Human. 2021. Available online: https://www.wwpdb.org/pdb?id=pdb_00006wfv (accessed on 25 February 2024).
27. Mattoteia, D.; Chiapparino, A.; Fumagalli, M.; De Marco, M.; De Giorgi, F.; Negro, L.; Pinnola, A.; Faravelli, S.; Roscioli, T.; Scietti, L.; et al. Identification of Regulatory Molecular “Hot Spots” for LH/PLOD Collagen Glycosyltransferase Activity. *Int. J. Mol. Sci.* **2023**, *24*, 11213. [[CrossRef](#)] [[PubMed](#)]
28. Chiapparino, F.; De Giorgi, F.; Scietti, L.; Faravelli, S.; Roscioli, T.; Forneris, F. A cooperative network of molecular “hot spots” highlights the complexity of LH3 collagen glycosyltransferase activities. *bioRxiv* **2020**. [[CrossRef](#)]
29. Bella, J. Collagen structure: New tricks from a very old dog. *Biochem. J.* **2016**, *473*, 1001–1025. [[CrossRef](#)]
30. Shoulders, M.D.; Raines, R.T. Collagen structure and stability. *Annu. Rev. Biochem.* **2009**, *78*, 929–958. [[CrossRef](#)]
31. Prockop, D.J.; Kivirikko, K.I. Collagens: Molecular biology, diseases, and potentials for therapy. *Annu. Rev. Biochem.* **1995**, *64*, 403–434. [[CrossRef](#)]
32. Myllyharju, J.; Kivirikko, K.I. Collagens, modifying enzymes and their mutations in humans, flies and worms. *Trends Genet.* **2004**, *20*, 33–43. [[CrossRef](#)]
33. Fratzl, P. (Ed.) *Collagen, Structure and Mechanism*; Springer Science+Business Media: New York, NY, USA, 2008. [[CrossRef](#)]
34. Kielty, C.M.; Grant, M.E. The Collagen Family: Structure, Assembly, and Organization in the Extracellular Matrix. In *Connective Tissue and Its Heritable Disorders: Molecular, Genetic, and Medical Aspects*; Royce, P.M., Steinmann, B., Eds.; Wiley-Liss, Inc.: Hoboken, NJ, USA, 2002; pp. 159–221. [[CrossRef](#)]
35. Sagi, I.; Afratis, N.A. (Eds.) *Collagen, Methods and Protocols*; Humana: New York, NY, USA, 2019. [[CrossRef](#)]
36. Wang, H. A Review of the Effects of Collagen Treatment in Clinical Studies. *Polymers* **2021**, *13*, 3868. [[CrossRef](#)] [[PubMed](#)]
37. Abou Neel, E.A.; Bozec, L.; Knowles, J.C.; Syed, O.; Mudera, V.; Day, R.; Hyun, J.K. Collagen-emerging collagen based therapies hit the patient. *Adv. Drug Deliv. Rev.* **2013**, *65*, 429–456. [[CrossRef](#)]
38. Onursal, C.; Dick, E.; Angelidis, I.; Schiller, H.B.; Staab-Weijnitz, C.A. Collagen Biosynthesis, Processing, and Maturation in Lung Ageing. *Front. Med.* **2021**, *8*, 593874. [[CrossRef](#)]
39. Bornstein, P.; Sage, H. Structurally distinct collagen types. *Annu. Rev. Biochem.* **1980**, *49*, 957–1003. [[CrossRef](#)] [[PubMed](#)]
40. Brodsky, B.; Persikov, A.V. Fibrous Proteins: Coiled-Coils, Collagen and Elastomers. *Adv. Protein Chem.* **2005**, *70*, 301–339. [[CrossRef](#)]
41. Yang, W.; Meyers, M.A.; Ritchie, R.O. Structural architectures with toughening mechanisms in nature: A review of the materials science of type-I collagenous materials. *Prog. Mat. Sci.* **2019**, *103*, 425–483. [[CrossRef](#)]
42. Bächinger, H.P.; Mizuno, K.; Vranka, J.A.; Boudko, S.P. Collagen Formation and Structure. *Compr. Nat. Prod. II* **2010**, *5*, 469–539. [[CrossRef](#)]
43. Mienaltowski, M.J.; Gonzales, N.L.; Beall, J.M.; Pechanec, M.Y. Basic Structure, Physiology, and Biochemistry of Connective Tissues and Extracellular Matrix Collagens. In *Progress in Heritable Soft Connective Tissue Diseases*; Springer: Berlin/Heidelberg, Germany, 2021; pp. 5–43. [[CrossRef](#)]
44. Ramshaw, J.A.; Shah, N.K.; Brodsky, B. Gly-X-Y tripeptide frequencies in collagen: A context for host-guest triple-helical peptides. *J. Struct. Biol.* **1998**, *122*, 86–91. [[CrossRef](#)]
45. Suzuki, H.; Mahapatra, D.; Board, A.J.; Steel, P.J.; Dyer, J.M.; Gerrard, J.A.; Dobson, R.C.J.; Valery, C. Sub-Angstrom structure of collagen model peptide (GPO)(10) shows a hydrated triple helix with pitch variation and two proline ring conformations. *Food Chem.* **2020**, *319*, 126598. [[CrossRef](#)]
46. Berisio, R.; Vitagliano, L.; Mazzarella, L.; Zagari, A. Crystal structure of the collagen triple helix model [(Pro-Pro-Gly)(10)](3). *Protein Sci.* **2002**, *11*, 262–270. [[CrossRef](#)]
47. Adzhubei, A.A.; Sternberg, M.J.; Makarov, A.A. Polyproline-II helix in proteins: Structure and function. *J. Mol. Biol.* **2013**, *425*, 2100–2132. [[CrossRef](#)]
48. Leitinger, B.; Hohenester, E. Mammalian collagen receptors. *Matrix Biol.* **2007**, *26*, 146–155. [[CrossRef](#)] [[PubMed](#)]
49. Gautieri, A.; Vesentini, S.; Redaelli, A.; Buehler, M.J. Hierarchical structure and nanomechanics of collagen microfibrils from the atomistic scale up. *Nano Lett.* **2011**, *11*, 757–766. [[CrossRef](#)]
50. Petruska, J.A.; Hodge, A.J. A Subunit Model for the Tropocollagen Macromolecule. *Proc. Natl. Acad. Sci. USA* **1964**, *51*, 871–876. [[CrossRef](#)] [[PubMed](#)]

51. Smith, J.W. Molecular pattern in native collagen. *Nature* **1968**, *219*, 157–158. [[CrossRef](#)] [[PubMed](#)]
52. Liu, X.; Zheng, C.; Luo, X.; Wang, X.; Jiang, H. Recent advances of collagen-based biomaterials: Multi-hierarchical structure, modification and biomedical applications. *Mater. Sci. Eng. C Mater. Biol. Appl.* **2019**, *99*, 1509–1522. [[CrossRef](#)]
53. Bozec, L.; van der Heijden, G.; Horton, M. Collagen fibrils: Nanoscale ropes. *Biophys. J.* **2007**, *92*, 70–75. [[CrossRef](#)]
54. Sorokina, L.V.; Shahbazian-Yassar, R.; Shokuhfar, T. Collagen biomineralization: Pathways, mechanisms, and thermodynamics. *Emergent Mater.* **2021**, *4*, 1205–1224. [[CrossRef](#)]
55. Amirrah, I.N.; Lokanathan, Y.; Zulkiflee, I.; Wee, M.; Motta, A.; Fauzi, M.B. A Comprehensive Review on Collagen Type I Development of Biomaterials for Tissue Engineering: From Biosynthesis to Bioscaffold. *Biomedicines* **2022**, *10*, 2307. [[CrossRef](#)]
56. Koide, T.; Nagata, K. Collagen Biosynthesis. *Top. Curr. Chem.* **2005**, *247*, 85–114. [[CrossRef](#)]
57. Salo, A.M.; Myllyharju, J. Prolyl and lysyl hydroxylases in collagen synthesis. *Exp. Dermatol.* **2021**, *30*, 38–49. [[CrossRef](#)]
58. Yamauchi, M.; Sricholpech, M. Lysine post-translational modifications of collagen. *Essays Biochem.* **2012**, *52*, 113–133. [[CrossRef](#)]
59. Gjaltema, R.A.; Bank, R.A. Molecular insights into prolyl and lysyl hydroxylation of fibrillar collagens in health and disease. *Crit. Rev. Biochem. Mol. Biol.* **2017**, *52*, 74–95. [[CrossRef](#)]
60. Yamauchi, M.; Terajima, M.; Shiiba, M. Lysine Hydroxylation and Cross-Linking of Collagen. *Methods Mol. Biol.* **2019**, *1934*, 309–324. [[CrossRef](#)]
61. Kivirikko, K.I. Posttranslational processing of collagens. In *Principles of Medical Biology*; Bittar, E.E., Bittar, N., Eds.; Elsevier: Amsterdam, The Netherlands, 1996; Volume 3, pp. 233–254.
62. Kivirikko, K.I.; Myllyla, R. Collagen glycosyltransferases. *Int. Rev. Connect. Tissue Res.* **1979**, *8*, 23–72. [[CrossRef](#)]
63. Rappu, P.; Salo, A.M.; Myllyharju, J.; Heino, J. Role of prolyl hydroxylation in the molecular interactions of collagens. *Essays Biochem.* **2019**, *63*, 325–335. [[CrossRef](#)]
64. Islam, M.S.; Leissing, T.M.; Chowdhury, R.; Hopkinson, R.J.; Schofield, C.J. 2-Oxoglutarate-Dependent Oxygenases. *Annu. Rev. Biochem.* **2018**, *87*, 585–620. [[CrossRef](#)]
65. Myllyharju, J. Collagen Hydroxylases. In *2-Oxoglutarate-Dependent Oxygenases*; Royal Society of Chemistry: London, UK, 2015; pp. 149–168. [[CrossRef](#)]
66. Markolovic, S.; Wilkins, S.E.; Schofield, C.J. Protein Hydroxylation Catalyzed by 2-Oxoglutarate-dependent Oxygenases. *J. Biol. Chem.* **2015**, *290*, 20712–20722. [[CrossRef](#)]
67. Salo, A.M.; Rappu, P.; Koski, M.K.; Karjalainen, E.; Izzi, V.; Drushinin, K.; Miinalainen, I.; Käpylä, J.; Heino, J.; Myllyharju, J. Collagen prolyl 4-hydroxylases have sequence specificity towards different X-Pro-Gly triplets. *bioRxiv* **2023**. [[CrossRef](#)]
68. Murthy, A.V.; Sulu, R.; Koski, M.K.; Tu, H.; Anantharajan, J.; Sah-Teli, S.K.; Myllyharju, J.; Wierenga, R.K. Structural enzymology binding studies of the peptide-substrate-binding domain of human collagen prolyl 4-hydroxylase (type-II): High affinity peptides have a PxGP sequence motif. *Protein Sci.* **2018**, *27*, 1692–1703. [[CrossRef](#)]
69. Myllyharju, J. Prolyl 4-hydroxylases, the key enzymes of collagen biosynthesis. *Matrix Biol.* **2003**, *22*, 15–24. [[CrossRef](#)]
70. Sipila, K.H.; Drushinin, K.; Rappu, P.; Jokinen, J.; Salminen, T.A.; Salo, A.M.; Käpylä, J.; Myllyharju, J.; Heino, J. Proline hydroxylation in collagen supports integrin binding by two distinct mechanisms. *J. Biol. Chem.* **2018**, *293*, 7645–7658. [[CrossRef](#)] [[PubMed](#)]
71. Gorres, K.L.; Raines, R.T. Prolyl 4-hydroxylase. *Crit. Rev. Biochem. Mol. Biol.* **2010**, *45*, 106–124. [[CrossRef](#)] [[PubMed](#)]
72. Murthy, A.V.; Sulu, R.; Lebedev, A.; Salo, A.M.; Korhonen, K.; Venkatesan, R.; Tu, H.; Bergmann, U.; Janis, J.; Laitaaja, M.; et al. Crystal structure of the collagen prolyl 4-hydroxylase (C-P4H) catalytic domain complexed with PDI: Toward a model of the C-P4H alpha(2)beta(2) tetramer. *J. Biol. Chem.* **2022**, *298*, 102614. [[CrossRef](#)]
73. Wang, C.; Li, W.; Ren, J.; Fang, J.; Ke, H.; Gong, W.; Feng, W.; Wang, C.C. Structural insights into the redox-regulated dynamic conformations of human protein disulfide isomerase. *Antioxid. Redox Signal.* **2013**, *19*, 36–45. [[CrossRef](#)]
74. Longbotham, J.E.; Levy, C.; Johannissen, L.O.; Tarhonskaya, H.; Jiang, S.; Loenarz, C.; Flashman, E.; Hay, S.; Schofield, C.J.; Scrutton, N.S. Structure and Mechanism of a Viral Collagen Prolyl Hydroxylase. *Biochemistry* **2015**, *54*, 6093–6105. [[CrossRef](#)] [[PubMed](#)]
75. Koski, M.K.; Hieta, R.; Hirsila, M.; Ronka, A.; Myllyharju, J.; Wierenga, R.K. The crystal structure of an algal prolyl 4-hydroxylase complexed with a proline-rich peptide reveals a novel buried tripeptide binding motif. *J. Biol. Chem.* **2009**, *284*, 25290–25301. [[CrossRef](#)]
76. Anantharajan, J.; Koski, M.K.; Kursula, P.; Hieta, R.; Bergmann, U.; Myllyharju, J.; Wierenga, R.K. The structural motifs for substrate binding and dimerization of the alpha subunit of collagen prolyl 4-hydroxylase. *Structure* **2013**, *21*, 2107–2118. [[CrossRef](#)]
77. Koski, M.K.; Anantharajan, J.; Kursula, P.; Dhavala, P.; Murthy, A.V.; Bergmann, U.; Myllyharju, J.; Wierenga, R.K. Assembly of the elongated collagen prolyl 4-hydroxylase alpha(2)beta(2) heterotetramer around a central alpha(2) dimer. *Biochem. J.* **2017**, *474*, 751–769. [[CrossRef](#)]
78. Chowdhury, R.; McDonough, M.A.; Mecinovic, J.; Loenarz, C.; Flashman, E.; Hewitson, K.S.; Domene, C.; Schofield, C.J. Structural basis for binding of hypoxia-inducible factor to the oxygen-sensing prolyl hydroxylases. *Structure* **2009**, *17*, 981–989. [[CrossRef](#)]
79. Timmins, A.; de Visser, S.P. How Are Substrate Binding and Catalysis Affected by Mutating Glu(127) and Arg(161) in Prolyl-4-hydroxylase? A QM/MM and MD Study. *Front. Chem.* **2017**, *5*, 94. [[CrossRef](#)]
80. Stegen, S.; Laperre, K.; Eelen, G.; Rinaldi, G.; Fraisl, P.; Torreken, S.; Van Loooveren, R.; Loopmans, S.; Bultynck, G.; Vinckier, S.; et al. HIF-1alpha metabolically controls collagen synthesis and modification in chondrocytes. *Nature* **2019**, *565*, 511–515. [[CrossRef](#)]

81. Vranka, J.A.; Sakai, L.Y.; Bachinger, H.P. Prolyl 3-hydroxylase 1, enzyme characterization and identification of a novel family of enzymes. *J. Biol. Chem.* **2004**, *279*, 23615–23621. [[CrossRef](#)]
82. Tiainen, P.; Pasanen, A.; Sormunen, R.; Myllyharju, J. Characterization of recombinant human prolyl 3-hydroxylase isoenzyme 2, an enzyme modifying the basement membrane collagen IV. *J. Biol. Chem.* **2008**, *283*, 19432–19439. [[CrossRef](#)] [[PubMed](#)]
83. Vranka, J.; Stadler, H.S.; Bachinger, H.P. Expression of prolyl 3-hydroxylase genes in embryonic and adult mouse tissues. *Cell Struct. Funct.* **2009**, *34*, 97–104. [[CrossRef](#)] [[PubMed](#)]
84. Forlino, A.; Marini, J.C. Osteogenesis imperfecta. *Lancet* **2016**, *387*, 1657–1671. [[CrossRef](#)] [[PubMed](#)]
85. Marini, J.C.; Forlino, A.; Bachinger, H.P.; Bishop, N.J.; Byers, P.H.; Paepe, A.; Fassier, F.; Fratzl-Zelman, N.; Kozloff, K.M.; Krakow, D.; et al. Osteogenesis imperfecta. *Nat. Rev. Dis. Primers* **2017**, *3*, 17052. [[CrossRef](#)] [[PubMed](#)]
86. Weis, M.A.; Hudson, D.M.; Kim, L.; Scott, M.; Wu, J.J.; Eyre, D.R. Location of 3-hydroxyproline residues in collagen types I, II, III, and V/XI implies a role in fibril supramolecular assembly. *J. Biol. Chem.* **2010**, *285*, 2580–2590. [[CrossRef](#)] [[PubMed](#)]
87. Hudson, D.M.; Eyre, D.R. Collagen prolyl 3-hydroxylation: A major role for a minor post-translational modification? *Connect. Tissue Res.* **2013**, *54*, 245–251. [[CrossRef](#)] [[PubMed](#)]
88. Rautavuoma, K.; Takaluoma, K.; Passoja, K.; Pirskanen, A.; Kvist, A.P.; Kivirikko, K.I.; Myllyharju, J. Characterization of three fragments that constitute the monomers of the human lysyl hydroxylase isoenzymes 1–3. The 30-kDa N-terminal fragment is not required for lysyl hydroxylase activity. *J. Biol. Chem.* **2002**, *277*, 23084–23091. [[CrossRef](#)] [[PubMed](#)]
89. Heikkinen, J.; Risteli, M.; Lampela, O.; Alavesä, P.; Karppinen, M.; Juffer, A.H.; Myllyla, R. Dimerization of human lysyl hydroxylase 3 (LH3) is mediated by the amino acids 541–547. *Matrix Biol.* **2011**, *30*, 27–33. [[CrossRef](#)]
90. Takaluoma, K.; Lantto, J.; Myllyharju, J. Lysyl hydroxylase 2 is a specific telopeptide hydroxylase, while all three isoenzymes hydroxylate collagenous sequences. *Matrix Biol.* **2007**, *26*, 396–403. [[CrossRef](#)]
91. Terajima, M.; Taga, Y.; Nakamura, T.; Guo, H.F.; Kayashima, Y.; Maeda-Smithies, N.; Parag-Sharma, K.; Kim, J.S.; Amelio, A.L.; Mizuno, K.; et al. Lysyl hydroxylase 2 mediated collagen post-translational modifications and functional outcomes. *Sci. Rep.* **2022**, *12*, 14256. [[CrossRef](#)]
92. Guo, H.F.; Bota-Rabassedas, N.; Terajima, M.; Leticia Rodriguez, B.; Gibbons, D.L.; Chen, Y.; Banerjee, P.; Tsai, C.L.; Tan, X.; Liu, X.; et al. A collagen glucosyltransferase drives lung adenocarcinoma progression in mice. *Commun. Biol.* **2021**, *4*, 482. [[CrossRef](#)] [[PubMed](#)]
93. 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. [[CrossRef](#)]
94. Wang, C.; Risteli, M.; Heikkinen, J.; Hussa, A.K.; Uitto, L.; Myllyla, R. Identification of amino acids important for the catalytic activity of the collagen glucosyltransferase associated with the multifunctional lysyl hydroxylase 3 (LH3). *J. Biol. Chem.* **2002**, *277*, 18568–18573. [[CrossRef](#)] [[PubMed](#)]
95. 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. [[CrossRef](#)]
96. Wang, C.; Luosujärvi, H.; Heikkinen, J.; Risteli, M.; Uitto, L.; Myllyla, R. The third activity for lysyl hydroxylase 3: Galactosylation of hydroxyllysyl residues in collagens in vitro. *Matrix Biol.* **2002**, *21*, 559–566. [[CrossRef](#)]
97. Schegg, B.; Hulsmeier, A.J.; 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. [[CrossRef](#)]
98. Sricholpech, M.; Perdivara, I.; Nagaoka, H.; Yokoyama, M.; Tomer, K.B.; Yamauchi, M. Lysyl hydroxylase 3 glucosylates galactosylhydroxylysine residues in type I collagen in osteoblast culture. *J. Biol. Chem.* **2011**, *286*, 8846–8856. [[CrossRef](#)] [[PubMed](#)]
99. Sricholpech, M.; Perdivara, I.; Yokoyama, M.; Nagaoka, H.; Terajima, M.; Tomer, K.B.; Yamauchi, M. Lysyl hydroxylase 3-mediated glucosylation in type I collagen: Molecular loci and biological significance. *J. Biol. Chem.* **2012**, *287*, 22998–23009. [[CrossRef](#)]
100. Moon, H.J.; Finney, J.; Ronnebaum, T.; Mure, M. Human lysyl oxidase-like 2. *Bioorg. Chem.* **2014**, *57*, 231–241. [[CrossRef](#)]
101. Rodriguez-Pascual, F.; Rosell-Garcia, T. Lysyl Oxidases: Functions and Disorders. *J. Glaucoma* **2018**, *27* (Suppl. S1), S15–S19. [[CrossRef](#)]
102. Vallet, S.D.; Ricard-Blum, S. Lysyl oxidases: From enzyme activity to extracellular matrix cross-links. *Essays Biochem.* **2019**, *63*, 349–364. [[CrossRef](#)]
103. Csizsar, K. Lysyl oxidases: A novel multifunctional amine oxidase family. *Prog. Nucleic Acid Res. Mol. Biol.* **2001**, *70*, 1–32. [[CrossRef](#)]
104. Avery, N.C.; Bailey, A.J. Restraining Cross-Links Responsible for the Mechanical Properties of Collagen Fibers: Natural and Artificial. In *Collagen: Structure and Mechanics*; Springer: Berlin/Heidelberg, Germany, 2008; pp. 81–110. [[CrossRef](#)]
105. Zhang, X.; Wang, Q.; Wu, J.; Wang, J.; Shi, Y.; Liu, M. Crystal structure of human lysyl oxidase-like 2 (hLOXL2) in a precursor state. *Proc. Natl. Acad. Sci. USA* **2018**, *115*, 3828–3833. [[CrossRef](#)]
106. Vallet, S.D.; Gueroult, M.; Belloy, N.; Dauchez, M.; Ricard-Blum, S. A Three-Dimensional Model of Human Lysyl Oxidase, a Cross-Linking Enzyme. *ACS Omega* **2020**, *4*, 8495–8505. [[CrossRef](#)]
107. Steiner, D.; Macron, C.R.; Cohen Sabban, E.N. Diabetes, Non-enzymatic Glycation, and Aging. In *Dermatology and Diabetes*; Springer: Berlin/Heidelberg, Germany, 2018; pp. 243–279. [[CrossRef](#)]
108. Bailey, A.J. Molecular mechanisms of ageing in connective tissues. *Mech. Ageing Dev.* **2001**, *122*, 735–755. [[CrossRef](#)]

109. Singh, R.; Barden, A.; Mori, T.; Beilin, L. Advanced glycation end-products: A review. *Diabetologia* **2001**, *44*, 129–146. [[CrossRef](#)] [[PubMed](#)]
110. Paul, R.G.; Bailey, A.J. Glycation of Collagen: The Basis of its Central Role in the Late Complications of Ageing and Diabetes. *Int. J. Biochem. Cell Biol.* **1996**, *28*, 1297–1310. [[CrossRef](#)] [[PubMed](#)]
111. Bailey, A.J.; Paul, R.G.; Knott, L. Mechanisms of maturation and ageing of collagen. *Mech. Ageing Dev.* **1998**, *106*, 1–56. [[CrossRef](#)]
112. Zheng, W.; Li, H.; Go, Y.; Chan, X.H.F.; Huang, Q.; Wu, J. Research Advances on the Damage Mechanism of Skin Glycation and Related Inhibitors. *Nutrients* **2022**, *14*, 4588. [[CrossRef](#)] [[PubMed](#)]
113. De Giorgi, F.; Fumagalli, M.; Scietti, L.; Forneris, F. Collagen hydroxylysine glycosylation: Non-conventional substrates for atypical glycosyltransferase enzymes. *Biochem. Soc. Trans.* **2021**, *49*, 855–866. [[CrossRef](#)]
114. Monnier, V.M.; Mustata, G.T.; Biemel, K.L.; Reihl, O.; Lederer, M.O.; Zhenyu, D.; Sell, D.R. Cross-linking of the extracellular matrix by the maillard reaction in aging and diabetes: An update on “a puzzle nearing resolution”. *Ann. N. Y. Acad. Sci.* **2005**, *1043*, 533–544. [[CrossRef](#)]
115. Hudson, D.M.; Archer, M.; King, K.B.; Eyre, D.R. Glycation of type I collagen selectively targets the same helical domain lysine sites as lysyl oxidase-mediated cross-linking. *J. Biol. Chem.* **2018**, *293*, 15620–15627. [[CrossRef](#)]
116. Tvaroska, I.; Kozmon, S.; Kona, J. Molecular Modeling Insights into the Structure and Behavior of Integrins: A Review. *Cells* **2023**, *12*, 324. [[CrossRef](#)] [[PubMed](#)]
117. Lee, J.; Yun, J.S.; Ko, S.H. Advanced Glycation End Products and Their Effect on Vascular Complications in Type 2 Diabetes Mellitus. *Nutrients* **2022**, *14*, 3086. [[CrossRef](#)] [[PubMed](#)]
118. Basta, G.; Schmidt, A.M.; De Caterina, R. Advanced glycation end products and vascular inflammation: Implications for accelerated atherosclerosis in diabetes. *Cardiovasc. Res.* **2004**, *63*, 582–592. [[CrossRef](#)] [[PubMed](#)]
119. Turner, D.P. The Role of Advanced Glycation End-Products in Cancer Disparity. *Adv. Cancer Res.* **2017**, *133*, 1–22. [[CrossRef](#)] [[PubMed](#)]
120. Schroter, D.; Hohn, A. Role of Advanced Glycation End Products in Carcinogenesis and their Therapeutic Implications. *Curr. Pharm. Des.* **2018**, *24*, 5245–5251. [[CrossRef](#)] [[PubMed](#)]
121. Muthyalaiyah, Y.S.; Jonnalagadda, B.; John, C.M.; Arockiasamy, S. Impact of Advanced Glycation End products (AGEs) and its receptor (RAGE) on cancer metabolic signaling pathways and its progression. *Glycoconj. J.* **2021**, *38*, 717–734. [[CrossRef](#)] [[PubMed](#)]
122. Dariya, B.; Nagaraju, G.P. Advanced glycation end products in diabetes, cancer and phytochemical therapy. *Drug Discov. Today* **2020**, *25*, 1614–1623. [[CrossRef](#)] [[PubMed](#)]
123. Schjoldager, K.T.; Narimatsu, Y.; Joshi, H.J.; Clausen, H. Global view of human protein glycosylation pathways and functions. *Nat. Rev. Mol. Cell Biol.* **2020**, *21*, 729–749. [[CrossRef](#)] [[PubMed](#)]
124. Tvaroska, I. Glycosyltransferases as targets for therapeutic intervention in cancer and inflammation: Molecular modeling insights. *Chem. Pap.* **2022**, *76*, 1953–1988. [[CrossRef](#)]
125. Lowe, J.B.; Marth, J.D. A genetic approach to Mammalian glycan function. *Annu. Rev. Biochem.* **2003**, *72*, 643–691. [[CrossRef](#)]
126. Campbell, J.A.; Davies, G.J.; Bulone, V.; Henrissat, B. A classification of nucleotide-diphospho-sugar glycosyltransferases based on amino acid sequence similarities. *Biochem. J.* **1997**, *326 Pt 3*, 929–939. [[CrossRef](#)]
127. Lombard, V.; Golaconda Ramulu, H.; Drula, E.; Coutinho, P.M.; Henrissat, B. The carbohydrate-active enzymes database (CAZy) in 2013. *Nucleic Acids Res.* **2014**, *42*, D490–D495. [[CrossRef](#)]
128. Unligil, U.M.; Rini, J.M. Glycosyltransferase structure and mechanism. *Curr. Opin. Struct. Biol.* **2000**, *10*, 510–517. [[CrossRef](#)]
129. Breton, C.; Snajdrova, L.; Jeanneau, C.; Koca, J.; Imberty, A. Structures and mechanisms of glycosyltransferases. *Glycobiology* **2006**, *16*, 29r–37r. [[CrossRef](#)] [[PubMed](#)]
130. Lairson, L.L.; Henrissat, B.; Davies, G.J.; Withers, S.G. Glycosyltransferases: Structures, functions, and mechanisms. *Annu. Rev. Biochem.* **2008**, *77*, 521–555. [[CrossRef](#)] [[PubMed](#)]
131. Perez, S.; Tvaroska, I. Carbohydrate-protein interactions: Molecular modeling insights. *Adv. Carbohydr. Chem. Biochem.* **2014**, *71*, 9–136. [[CrossRef](#)] [[PubMed](#)]
132. Laine, R.A. A calculation of all possible oligosaccharide isomers both branched and linear yields 1.05×10^{12} structures for a reducing hexasaccharide: The *Isomer Barrier* to development of single-method saccharide sequencing or synthesis systems. *Glycobiology* **1994**, *4*, 759–767. [[CrossRef](#)] [[PubMed](#)]
133. Varki, A. Biological roles of glycans. *Glycobiology* **2017**, *27*, 3–49. [[CrossRef](#)]
134. Olsen, B.R.; Guzman, N.A.; Engel, J.; Condit, C.; Aase, S. Purification and characterization of a peptide from the carboxy-terminal region of chick tendon procollagen type I. *Biochemistry* **1977**, *16*, 3030–3036. [[CrossRef](#)]
135. Clark, C.C. The distribution and initial characterization of oligosaccharide units on the COOH-terminal propeptide extensions of the pro-alpha 1 and pro-alpha 2 chains of type I procollagen. *J. Biol. Chem.* **1979**, *254*, 10798–10802. [[CrossRef](#)] [[PubMed](#)]
136. Clark, C.C.; Kefalides, N.A. Carbohydrate moieties of procollagen: Incorporation of isotopically labeled mannose and glucosamine into propeptides of procollagen secreted by matrix-free chick embryo tendon cells. *Proc. Natl. Acad. Sci. USA* **1976**, *73*, 34–38. [[CrossRef](#)]
137. Li, R.C.; Wong, M.Y.; DiChiara, A.S.; Hosseini, A.S.; Shoulders, M.D. Collagen’s enigmatic, highly conserved N-glycan has an essential proteostatic function. *Proc. Natl. Acad. Sci. USA* **2021**, *118*, e2026608118. [[CrossRef](#)]
138. Anttinen, H.; Oikarinen, A.; Ryhanen, L.; Kivirikko, K.I. Evidence for the transfer of mannose to the extension peptides of procollagen within the cisternae of the rough endoplasmic reticulum. *FEBS Lett.* **1978**, *87*, 222–226. [[CrossRef](#)]

139. Duksin, D.; Bornstein, P. Impaired conversion of procollagen to collagen by fibroblasts and bone treated with tunicamycin, an inhibitor of protein glycosylation. *J. Biol. Chem.* **1977**, *252*, 955–962. [[CrossRef](#)]
140. Lamande, S.R.; Bateman, J.F. The type I collagen pro alpha 1(I) COOH-terminal propeptide N-linked oligosaccharide. Functional analysis by site-directed mutagenesis. *J. Biol. Chem.* **1995**, *270*, 17858–17865. [[CrossRef](#)]
141. Clark, C.C. Asparagine-linked glycosides. *Methods Enzymol.* **1982**, *82 Pt A*, 346–360. [[CrossRef](#)]
142. Clark, C.C.; Kefalides, N.A. Localization and partial composition of the oligosaccharide units on the propeptide extensions of type I procollagen. *J. Biol. Chem.* **1978**, *253*, 47–51. [[CrossRef](#)] [[PubMed](#)]
143. Nagae, M.; Yamaguchi, Y.; Taniguchi, N.; Kizuka, Y. 3D Structure and Function of Glycosyltransferases Involved in N-glycan Maturation. *Int. J. Mol. Sci.* **2020**, *21*, 437. [[CrossRef](#)] [[PubMed](#)]
144. Stanley, P.; Taniguchi, N.; Aebi, M. N-Glycans. In *Essentials of Glycobiology*; Cold Spring Harbor Laboratory Press: Harbor, NY, USA, 2015; pp. 99–111. [[CrossRef](#)]
145. Vasconcelos-Dos-Santos, A.; Oliveira, I.A.; Lucena, M.C.; Mantuano, N.R.; Whelan, S.A.; Dias, W.B.; Todeschini, A.R. Biosynthetic Machinery Involved in Aberrant Glycosylation: Promising Targets for Developing of Drugs Against Cancer. *Front. Oncol.* **2015**, *5*, 138. [[CrossRef](#)]
146. Aebi, M. N-linked protein glycosylation in the ER. *Biochim. Biophys. Acta* **2013**, *1833*, 2430–2437. [[CrossRef](#)]
147. Marti, T.; Kosen, P.A.; Honory, K.; Franzblau, C.; Schmid, K.; Van Halbeek, H.; Gerwig, G.J.; Vliegthart, J.F. The structure of the carbohydrate units of the carboxyl-terminal peptide of procollagen as elucidated by 500 MHz ¹H-NMR spectroscopy. *Biochim. Biophys. Acta* **1984**, *799*, 305–312. [[CrossRef](#)]
148. Kornfeld, R.; Kornfeld, S. Assembly of asparagine-linked oligosaccharides. *Annu. Rev. Biochem.* **1985**, *54*, 631–664. [[CrossRef](#)] [[PubMed](#)]
149. Grassmann, W.; Schleich, H. Über den Kohlenhydratgehalt des Kollagens. 2. Mitteilung zur Kenntnis des Kollagens. *Biochem. Z.* **1935**, *277*, 320–328.
150. Graßmann, W.; Hörmann, H.; Hafter, R. Eine quantitative Bestimmung von Kohlenhydraten als Osazone. *Hoppe-Seyler's Z. Physiol. Chem.* **1957**, *307*, 87–96. [[CrossRef](#)]
151. Blumenfeld, O.O.; Paz, M.A.; Gallop, P.M.; Seifter, S. The Nature, Quantity, and Mode of Attachment of Hexoses in Ichthyocol. *J. Biol. Chem.* **1963**, *238*, 3835–3839. [[CrossRef](#)]
152. Spiro, R.G.; Spiro, M.J. Studies on the biosynthesis of the hydroxylysine-linked disaccharide unit of basement membranes and collagens. 1. Tissue and subcellular distribution of glycosyltransferases and the effect of various conditions on the enzyme levels. *J. Biol. Chem.* **1971**, *246*, 4899–4910. [[CrossRef](#)] [[PubMed](#)]
153. Spiro, R.G.; Spiro, M.J. Studies on the biosynthesis of the hydroxylysine-linked disaccharide unit of basement membranes and collagens. 2. Tissue and subcellular distribution of glycosyltransferases and the effect of various conditions on the enzyme levels. *J. Biol. Chem.* **1971**, *246*, 4910–4918. [[CrossRef](#)] [[PubMed](#)]
154. Spiro, R.G.; Spiro, M.J. Studies on the biosynthesis of the hydroxylysine-linked disaccharide unit of basement membranes and collagens. 3. Tissue and subcellular distribution of glycosyltransferases and the effect of various conditions on the enzyme levels. *J. Biol. Chem.* **1971**, *246*, 4919–4925. [[CrossRef](#)] [[PubMed](#)]
155. Bosmann, H.B.; Eylar, E.H. Glycoprotein biosynthesis: The biosynthesis of the hydroxylysine-galactose linkage in collagen. *Biochem. Biophys. Res. Commun.* **1968**, *33*, 340–346. [[CrossRef](#)] [[PubMed](#)]
156. Bosmann, H.B. Collagen-Glucosyl Transferase in Fibroblasts transformed by Oncogenic Viruses. *Nature* **1968**, *218*, 582–583. [[CrossRef](#)] [[PubMed](#)]
157. Kivirikko, K.I.; Myllyla, R. Posttranslational enzymes in the biosynthesis of collagen: Intracellular enzymes. *Methods Enzymol.* **1982**, *82*, 245–304. [[CrossRef](#)]
158. Liefhebber, J.M.; Punt, S.; Spaan, W.J.; van Leeuwen, H.C. The human collagen beta(1-O)galactosyltransferase, GLT25D1, is a soluble endoplasmic reticulum localized protein. *BMC Cell Biol.* **2010**, *11*, 33. [[CrossRef](#)] [[PubMed](#)]
159. Myllyla, R.; Risteli, L.; Kivirikko, K.I. Collagen glucosyltransferase. Partial purification and characterization of the enzyme from whole chick embryos and chick-embryo cartilage. *Eur. J. Biochem.* **1976**, *61*, 59–67. [[CrossRef](#)]
160. Risteli, A.L.; Myllyla, R.; Kivirikko, K.I. Partial purification and characterization of collagen galactosyltransferase from chick embryos. *Biochem. J.* **1976**, *155*, 145–153. [[CrossRef](#)]
161. Risteli, B.L.; Myllyla, R.; Kivirikko, K.I. Affinity chromatography of collagen glycosyltransferases on collagen linked to agarose. *Eur. J. Biochem.* **1976**, *67*, 197–202. [[CrossRef](#)]
162. Perrin-Tricaud, C.; Rutschmann, C.; Hennet, T. Identification of domains and amino acids essential to the collagen galactosyltransferase activity of GLT25D1. *PLoS ONE* **2011**, *6*, e29390. [[CrossRef](#)] [[PubMed](#)]
163. Breton, C.; Imberty, A. Structure/function studies of glycosyltransferases. *Curr. Opin. Struct. Biol.* **1999**, *9*, 563–571. [[CrossRef](#)]
164. Qasba, P.K.; Ramakrishnan, B.; Boeggeman, E. Substrate-induced conformational changes in glycosyltransferases. *Trends Biochem. Sci.* **2005**, *30*, 53–62. [[CrossRef](#)] [[PubMed](#)]
165. Unligil, U.M.; Zhou, S.; Yuwaraj, S.; Sarkar, M.; Schachter, H.; Rini, J.M. X-ray crystal structure of rabbit N-acetylglucosaminyltransferase I: Catalytic mechanism and a new protein superfamily [In Process Citation]. *EMBO J.* **2000**, *19*, 5269–5280. [[CrossRef](#)] [[PubMed](#)]
166. Jumper, J.; Evans, R.; Pritzel, A.; Green, T.; Figurnov, M.; Ronneberger, O.; Tunyasuvunakool, K.; Bates, R.; Zidek, A.; Potapenko, A.; et al. Highly accurate protein structure prediction with AlphaFold. *Nature* **2021**, *596*, 583–589. [[CrossRef](#)] [[PubMed](#)]

167. Varadi, M.; Anyango, S.; Deshpande, M.; Nair, S.; Natassia, C.; Yordanova, G.; Yuan, D.; Stroe, O.; Wood, G.; Laydon, A.; et al. AlphaFold Protein Structure Database: Massively expanding the structural coverage of protein-sequence space with high-accuracy models. *Nucleic Acids Res.* **2022**, *50*, D439–D444. [[CrossRef](#)] [[PubMed](#)]
168. Guo, H.F.; Tsai, C.L.; Terajima, M.; Tan, X.; Banerjee, P.; Miller, M.D.; Liu, X.; Yu, J.; Byemerwa, J.; Alvarado, S.; et al. Pro-metastatic collagen lysyl hydroxylase dimer assemblies stabilized by Fe(2+)-binding. *Nat. Commun.* **2018**, *9*, 512. [[CrossRef](#)]
169. Rautavuoma, K.; Takaluoma, K.; Sormunen, R.; Myllyharju, J.; Kivirikko, K.I.; Soininen, R. Premature aggregation of type IV collagen and early lethality in lysyl hydroxylase 3 null mice. *Proc. Natl. Acad. Sci. USA* **2004**, *101*, 14120–14125. [[CrossRef](#)]
170. Salo, A.M.; Wang, C.; Sipila, L.; Sormunen, R.; Vapola, M.; Kervinen, P.; Ruotsalainen, H.; Heikkinen, J.; Myllyla, R. Lysyl hydroxylase 3 (LH3) modifies proteins in the extracellular space, a novel mechanism for matrix remodeling. *J. Cell. Physiol.* **2006**, *207*, 644–653. [[CrossRef](#)]
171. Aik, W.; McDonough, M.A.; Thalhammer, A.; Chowdhury, R.; Schofield, C.J. Role of the jelly-roll fold in substrate binding by 2-oxoglutarate oxygenases. *Curr. Opin. Struct. Biol.* **2012**, *22*, 691–700. [[CrossRef](#)] [[PubMed](#)]
172. Ruotsalainen, H.; Sipila, L.; Vapola, M.; Sormunen, R.; Salo, A.M.; Uitto, L.; Mercer, D.K.; Robins, S.P.; Risteli, M.; Aszodi, A.; et al. Glycosylation catalyzed by lysyl hydroxylase 3 is essential for basement membranes. *J. Cell Sci.* **2006**, *119*, 625–635. [[CrossRef](#)]
173. Gómez, H.; Rojas, R.; Patel, D.; Tabak, L.A.; Lluch, J.M.; Masgrau, L. A computational and experimental study of O-glycosylation. Catalysis by human UDP-GalNAc polypeptide: GalNAc transferase-T2. *Org. Biomol. Chem.* **2014**, *12*, 2645–2655. [[CrossRef](#)]
174. Tvaroska, I. Molecular modeling of retaining glycosyltransferases. *ACS Symp. Ser.* **2006**, *930*, 285–301. [[CrossRef](#)]
175. Andre, I.; Tvaroska, I.; Carver, J.P. On the reaction pathways and determination of transition-state structures for retaining alpha-galactosyltransferases. *Carbohydr. Res.* **2003**, *338*, 865–877. [[CrossRef](#)]
176. Monegal, A.; Planas, A. Chemical rescue of alpha 3-galactosyltransferase. Implications in the mechanism of retaining glycosyltransferases. *J. Am. Chem. Soc.* **2006**, *128*, 16030–16031. [[CrossRef](#)] [[PubMed](#)]
177. Rojas-Cervellera, V.; Ardevol, A.; Boero, M.; Planas, A.; Rovira, C. Formation of a covalent glycosyl-enzyme species in a retaining glycosyltransferase. *Chemistry* **2013**, *19*, 14018–14023. [[CrossRef](#)] [[PubMed](#)]
178. Soya, N.; Fang, Y.; Palcic, M.M.; Klassen, J.S. Trapping and characterization of covalent intermediates of mutant retaining glycosyltransferases. *Glycobiology* **2011**, *21*, 547–552. [[CrossRef](#)] [[PubMed](#)]
179. Persson, K.; Ly, H.D.; Dieckelmann, M.; Wakarchuk, W.W.; Withers, S.G.; Strynadka, N.C.J. Crystal structure of the retaining galactosyltransferase LgtC from *Neisseria meningitidis* in complex with donor and acceptor sugar analogs: Towards a mechanism. *Nat. Struct. Biol.* **2001**, *8*, 166–175. [[CrossRef](#)] [[PubMed](#)]
180. Tvaroška, I. Molecular modeling insights into the catalytic mechanism of the retaining galactosyltransferase LgtC. *Carbohydr. Res.* **2004**, *339*, 1007–1014. [[CrossRef](#)] [[PubMed](#)]
181. Gomez, H.; Polyak, I.; Thiel, W.; Lluch, J.M.; Masgrau, L. Retaining Glycosyltransferase Mechanism Studied by QM/MM Methods: Lipopolysaccharyl-alpha-1,4-galactosyltransferase C Transfers alpha-Galactose via an Oxocarbenium Ion-like Transition State. *J. Am. Chem. Soc.* **2012**, *134*, 4743–4752. [[CrossRef](#)] [[PubMed](#)]
182. Gómez, H.; Lluch, J.M.; Masgrau, L. Substrate-Assisted and Nucleophilically Assisted Catalysis in Bovine alpha 1,3-Galactosyltransferase. Mechanistic Implications for Retaining Glycosyltransferases. *J. Am. Chem. Soc.* **2013**, *135*, 7053–7063. [[CrossRef](#)] [[PubMed](#)]
183. Albesa-Jove, D.; Sainz-Polo, M.A.; Marina, A.; Guerin, M.E. Structural Snapshots of alpha-1,3-Galactosyltransferase with Native Substrates: Insight into the Catalytic Mechanism of Retaining Glycosyltransferases. *Angew. Chem. Int. Ed. Engl.* **2017**, *56*, 14853–14857. [[CrossRef](#)] [[PubMed](#)]
184. Ardevol, A.; Rovira, C. The Molecular Mechanism of Enzymatic Glycosyl Transfer with Retention of Configuration: Evidence for a Short-Lived Oxocarbenium-Like Species. *Angew. Chem.-Int. Ed.* **2011**, *50*, 10897–10901. [[CrossRef](#)] [[PubMed](#)]
185. Bobovska, A.; Tvaroska, I.; Kona, J. A theoretical study on the catalytic mechanism of the retaining alpha-1,2-mannosyltransferase Kre2p/Mnt1p: The impact of different metal ions on catalysis. *Org. Biomol. Chem.* **2014**, *12*, 4201–4210. [[CrossRef](#)] [[PubMed](#)]
186. Janos, P.; Trnka, T.; Kozmon, S.; Tvaroska, I.; Koca, J. Different QM/MM Approaches To Elucidate Enzymatic Reactions: Case Study on ppGalNAcT2. *J. Chem. Theory. Comput.* **2016**, *12*, 6062–6076. [[CrossRef](#)]
187. Lira-Navarrete, E.; Iglesias-Fernandez, J.; Zandberg, W.F.; Companon, I.; Kong, Y.; Corzana, F.; Pinto, B.M.; Clausen, H.; Peregrina, J.M.; Vocadlo, D.J.; et al. Substrate-guided front-face reaction revealed by combined structural snapshots and metadynamics for the polypeptide N-acetylgalactosaminyltransferase 2. *Angew. Chem. Int. Ed. Engl.* **2014**, *53*, 8206–8210. [[CrossRef](#)]
188. Trnka, T.; Kozmon, S.; Tvaroska, I.; Koca, J. Stepwise catalytic mechanism via short-lived intermediate inferred from combined QM/MM MERP and PES calculations on retaining glycosyltransferase ppGalNAcT2. *PLoS Comput. Biol.* **2015**, *11*, e1004061. [[CrossRef](#)]
189. Sternberg, M.; Spiro, R.G. Studies on the catabolism of the hydroxylysine-linked disaccharide units of basement membranes and collagens. Isolation and characterization of a rat kidney alpha-glucosidase of high specificity. *J. Biol. Chem.* **1979**, *254*, 10329–10336. [[CrossRef](#)]
190. Sternberg, M.; Spiro, R.G. Studies on the Catabolism of the Hydroxylysine-Linked Disaccharide Units of Basement Membranes and Collagens: Isolation and Characterization of a New Rat-Kidney α -Glucosidase of High Specificity. *Ren. Physiol.* **1980**, *3*, 1–3. [[CrossRef](#)]
191. Hamazaki, H.; Hotta, K. Purification and characterization of an alpha-glucosidase specific for hydroxylysine-linked disaccharide of collagen. *J. Biol. Chem.* **1979**, *254*, 9682–9687. [[CrossRef](#)] [[PubMed](#)]

192. Hamazaki, H.; Hamazaki, M.H. Catalytic site of human protein-glucosylgalactosylhydroxylysine glucosidase: Three crucial carboxyl residues were determined by cloning and site-directed mutagenesis. *Biochem. Biophys. Res. Commun.* **2016**, *469*, 357–362. [[CrossRef](#)] [[PubMed](#)]
193. Nakamura, S.; Nihira, T.; Kurata, R.; Nakai, H.; Funane, K.; Park, E.Y.; Miyazaki, T. Structure of a bacterial alpha-1,2-glucosidase defines mechanisms of hydrolysis and substrate specificity in GH65 family hydrolases. *J. Biol. Chem.* **2021**, *297*, 101366. [[CrossRef](#)]
194. Fang, M.; Yuan, J.; Peng, C.; Li, Y. Collagen as a double-edged sword in tumor progression. *Tumour Biol.* **2014**, *35*, 2871–2882. [[CrossRef](#)]
195. Stawikowski, M.J.; Aukszi, B.; Stawikowska, R.; Cudic, M.; Fields, G.B. Glycosylation modulates melanoma cell alpha2beta1 and alpha3beta1 integrin interactions with type IV collagen. *J. Biol. Chem.* **2014**, *289*, 21591–21604. [[CrossRef](#)]
196. Tang, M.; Wang, X.; Gandhi, N.S.; Foley, B.L.; Burrage, K.; Woods, R.J.; Gu, Y. Effect of hydroxylysine-O-glycosylation on the structure of type I collagen molecule: A computational study. *Glycobiology* **2020**, *30*, 830–843. [[CrossRef](#)]
197. An, H.J.; Froehlich, J.W.; Lebrilla, C.B. Determination of glycosylation sites and site-specific heterogeneity in glycoproteins. *Curr. Opin. Chem. Biol.* **2009**, *13*, 421–426. [[CrossRef](#)]
198. Perdivara, I.; Yamauchi, M.; Tomer, K.B. Molecular Characterization of Collagen Hydroxylysine O-Glycosylation by Mass Spectrometry: Current Status. *Aust. J. Chem.* **2013**, *66*, 760–769. [[CrossRef](#)]
199. Sharma, U.; Carrique, L.; Vadon-Le Goff, S.; Mariano, N.; Georges, R.N.; Delolme, F.; Koivunen, P.; Myllyharju, J.; Moali, C.; Aghajari, N.; et al. Structural basis of homo- and heterotrimerization of collagen I. *Nat. Commun.* **2017**, *8*, 14671. [[CrossRef](#)]
200. Davies, D.; Tuckwell, D.S.; Calderwood, D.A.; Weston, S.A.; Takigawa, M.; Humphries, M.J. Molecular characterisation of integrin-procollagen C-propeptide interactions. *Eur. J. Biochem.* **1997**, *246*, 274–282. [[CrossRef](#)]
201. Katayama, K.; Armendariz-Borunda, J.; Raghov, R.; Kang, A.H.; Seyer, J.M. A pentapeptide from type I procollagen promotes extracellular matrix production. *J. Biol. Chem.* **1993**, *268*, 9941–9944. [[CrossRef](#)] [[PubMed](#)]
202. Smedsrod, B.; Melkko, J.; Risteli, L.; Risteli, J. Circulating C-terminal propeptide of type I procollagen is cleared mainly via the mannose receptor in liver endothelial cells. *Biochem. J.* **1990**, *271*, 345–350. [[CrossRef](#)]
203. Notbohm, H.; Nokelainen, M.; Myllyharju, J.; Fietzek, P.P.; Muller, P.K.; Kivirikko, K.I. 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. [[CrossRef](#)]
204. Torre-Blanco, A.; Adachi, E.; Hojima, Y.; Wootton, J.A.; Minor, R.R.; Prockop, D.J. 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. [[CrossRef](#)]
205. Batge, B.; Winter, C.; Notbohm, H.; Acil, Y.; Brinckmann, J.; Muller, P.K. Glycosylation of human bone collagen I in relation to lysylhydroxylation and fibril diameter. *J. Biochem.* **1997**, *122*, 109–115. [[CrossRef](#)]
206. Sipila, L.; Ruotsalainen, H.; Sormunen, R.; Baker, N.L.; Lamande, S.R.; 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. [[CrossRef](#)]
207. Risteli, M.; Ruotsalainen, H.; Salo, A.M.; Sormunen, R.; Sipila, L.; Baker, N.L.; Lamande, S.R.; Vimpri-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. [[CrossRef](#)]
208. Jurgensen, H.J.; Madsen, D.H.; Ingvarsen, S.; Melander, M.C.; Gardsvoll, H.; Patthy, L.; Engelholm, L.H.; Behrendt, N. A novel functional role of collagen glycosylation: Interaction with the endocytic collagen receptor uparap/ENDO180. *J. Biol. Chem.* **2011**, *286*, 32736–32748. [[CrossRef](#)]
209. Gardelli, C.; Russo, L.; Cipolla, L.; Moro, M.; Andriani, F.; Rondinone, O.; Nicotra, F.; Sozzi, G.; Bertolini, G.; Roz, L. Differential glycosylation of collagen modulates lung cancer stem cell subsets through beta1 integrin-mediated interactions. *Cancer Sci.* **2021**, *112*, 217–230. [[CrossRef](#)]
210. Wang, C.; Kovanen, V.; Raudasoja, P.; Eskelinen, S.; Pospiech, H.; Myllyla, R. The glycosyltransferase activities of lysyl hydroxylase 3 (LH3) in the extracellular space are important for cell growth and viability. *J. Cell. Mol. Med.* **2009**, *13*, 508–521. [[CrossRef](#)]
211. Myers, L.K.; Myllyharju, J.; Nokelainen, M.; Brand, D.D.; Cremer, M.A.; Stuart, J.M.; Bodo, M.; Kivirikko, K.I.; Kang, A.H. Relevance of posttranslational modifications for the arthritogenicity of type II collagen. *J. Immunol.* **2004**, *172*, 2970–2975. [[CrossRef](#)]
212. Myllyharju, J.; Kivirikko, K.I. Collagens and collagen-related diseases. *Ann. Med.* **2001**, *33*, 7–21. [[CrossRef](#)]
213. Ricard-Blum, S.; Baffet, G.; Theret, N. Molecular and tissue alterations of collagens in fibrosis. *Matrix Biol.* **2018**, *68–69*, 122–149. [[CrossRef](#)]
214. Williams, L.; Layton, T.; Yang, N.; Feldmann, M.; Nanchahal, J. Collagen VI as a driver and disease biomarker in human fibrosis. *FEBS J.* **2022**, *289*, 3603–3629. [[CrossRef](#)]
215. Muschler, J.; Streuli, C.H. Cell-matrix interactions in mammary gland development and breast cancer. *Cold Spring Harb. Perspect. Biol.* **2010**, *2*, a003202. [[CrossRef](#)]
216. Cox, T.R.; Erler, J.T. Molecular pathways: Connecting fibrosis and solid tumor metastasis. *Clin. Cancer. Res.* **2014**, *20*, 3637–3643. [[CrossRef](#)]
217. Wagner, G.K.; Pesnot, T. Glycosyltransferases and their Assays. *Chembiochem* **2010**, *11*, 1939–1949. [[CrossRef](#)]

218. Li, Q.; Yang, Y.; Reis, C.; Tao, T.; Li, W.; Li, X.; Zhang, J.H. Cerebral Small Vessel Disease. *Cell Transplant.* **2018**, *27*, 1711–1722. [[CrossRef](#)]
219. Gould, D.B.; Phalan, F.C.; van Mil, S.E.; Sundberg, J.P.; Vahedi, K.; Massin, P.; Bousser, M.G.; Heutink, P.; Miner, J.H.; Tournier-Lasserre, E.; et al. Role of COL4A1 in small-vessel disease and hemorrhagic stroke. *N. Engl. J. Med.* **2006**, *354*, 1489–1496. [[CrossRef](#)]
220. Weng, Y.C.; Sonni, A.; Labelle-Dumais, C.; de Leau, M.; Kauffman, W.B.; Jeanne, M.; Biffi, A.; Greenberg, S.M.; Rosand, J.; Gould, D.B. COL4A1 mutations in patients with sporadic late-onset intracerebral hemorrhage. *Ann. Neurol.* **2012**, *71*, 470–477. [[CrossRef](#)]
221. Jeanne, M.; Labelle-Dumais, C.; Jorgensen, J.; Kauffman, W.B.; Mancini, G.M.; Favor, J.; Valant, V.; Greenberg, S.M.; Rosand, J.; Gould, D.B. COL4A2 mutations impair COL4A1 and COL4A2 secretion and cause hemorrhagic stroke. *Am. J. Hum. Genet.* **2012**, *90*, 91–101. [[CrossRef](#)]
222. Miyatake, S.; Schneeberger, S.; Koyama, N.; Yokochi, K.; Ohmura, K.; Shiina, M.; Mori, H.; Koshimizu, E.; Imagawa, E.; Uchiyama, Y.; et al. Biallelic COLGALT1 variants are associated with cerebral small vessel disease. *Ann. Neurol.* **2018**, *84*, 843–853. [[CrossRef](#)]
223. National Academies of Sciences, Engineering, and Medicine. *Selected Heritable Disorders of Connective Tissue and Disability*; The National Academies Press: Washington, DC, USA, 2022. [[CrossRef](#)]
224. Salo, A.M.; Cox, H.; Farndon, P.; Moss, C.; Grindulis, H.; Risteli, M.; Robins, S.P.; Myllyla, R. A connective tissue disorder caused by mutations of the lysyl hydroxylase 3 gene. *Am. J. Hum. Genet.* **2008**, *83*, 495–503. [[CrossRef](#)]
225. Ewans, L.J.; Colley, A.; Gaston-Massuet, C.; Gualtieri, A.; Cowley, M.J.; McCabe, M.J.; Anand, D.; Lachke, S.A.; Scietti, L.; Forneris, F.; et al. Pathogenic variants in PLOD3 result in a Stickler syndrome-like connective tissue disorder with vascular complications. *J. Med. Genet.* **2019**, *56*, 629–638. [[CrossRef](#)]
226. Vahidnezhad, H.; Youssefian, L.; Saeidian, A.H.; Touati, A.; Pajouhanfar, S.; Baghdadi, T.; Shadmehri, A.A.; Giunta, C.; Kraenzlin, M.; Syx, D.; et al. Mutations in PLOD3, encoding lysyl hydroxylase 3, cause a complex connective tissue disorder including recessive dystrophic epidermolysis bullosa-like blistering phenotype with abnormal anchoring fibrils and type VII collagen deficiency. *Matrix Biol.* **2019**, *81*, 91–106. [[CrossRef](#)]
227. Giunta, C.; Randolph, A.; Al-Gazali, L.I.; Brunner, H.G.; Kraenzlin, M.E.; Steinmann, B. Nevo syndrome is allelic to the kyphoscoliotic type of the Ehlers-Danlos syndrome (EDS VIA). *Am. J. Med. Genet. A* **2005**, *133*, 158–164. [[CrossRef](#)]
228. Beighton, P.; De Paepe, A.; Steinmann, B.; Tsipouras, P.; Wenstrup, R.J. Ehlers-Danlos syndromes: Revised nosology, Villefranche, 1997. Ehlers-Danlos National Foundation (USA) and Ehlers-Danlos Support Group (UK). *Am. J. Med. Genet.* **1998**, *77*, 31–37. [[CrossRef](#)]
229. Steinmann, B.; Royce, P.M.; Superti-Furga, A. *Connective Tissue and Its Heritable Disorders: Molecular, Genetic, and Medical Aspects*; Wiley-Liss, Inc.: Hoboken, NJ, USA, 2002. [[CrossRef](#)]
230. Yeowell, H.N.; Walker, L.C. 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. [[CrossRef](#)]
231. Ha-Vinh, R.; Alanay, Y.; Bank, R.A.; Campos-Xavier, A.B.; 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. [[CrossRef](#)]
232. van der Slot, A.J.; Zuurmond, A.M.; Bardoeel, A.F.; Wijmenga, C.; Pruijs, H.E.; Sillence, D.O.; Brinckmann, J.; Abraham, D.J.; Black, C.M.; Verzijl, N.; et al. Identification of PLOD2 as telopeptide lysyl hydroxylase, an important enzyme in fibrosis. *J. Biol. Chem.* **2003**, *278*, 40967–40972. [[CrossRef](#)]
233. Puig-Hervas, M.T.; Temtamy, S.; Aglan, M.; Valencia, M.; Martinez-Glez, V.; Ballesta-Martinez, M.J.; Lopez-Gonzalez, V.; Ashour, A.M.; Amr, K.; Pulido, V.; et al. Mutations in PLOD2 cause autosomal-recessive connective tissue disorders within the Bruck syndrome–osteogenesis imperfecta phenotypic spectrum. *Hum. Mutat.* **2012**, *33*, 1444–1449. [[CrossRef](#)]
234. Luce, L.; Casale, M.; Waldron, S. A Rare Case of Bruck Syndrome Type 2 in Siblings with Broad Phenotypic Variability. *Ochsner J.* **2020**, *20*, 204–208. [[CrossRef](#)]
235. Uitto, J.; Richard, G. Progress in epidermolysis bullosa: From eponyms to molecular genetic classification. *Clin. Dermatol.* **2005**, *23*, 33–40. [[CrossRef](#)]
236. Savolainen, E.R.; Kero, M.; Pihlajaniemi, T.; Kivirikko, K.I. Deficiency of galactosylhydroxyllysyl glucosyltransferase, an enzyme of collagen synthesis, in a family with dominant epidermolysis bullosa simplex. *N. Engl. J. Med.* **1981**, *304*, 197–204. [[CrossRef](#)]
237. Snead, M.P.; Yates, J.R. Clinical and Molecular genetics of Stickler syndrome. *J. Med. Genet.* **1999**, *36*, 353–359. [[CrossRef](#)]
238. Soh, Z.; Richards, A.J.; McNinch, A.; Alexander, P.; Martin, H.; Snead, M.P. Dominant Stickler Syndrome. *Genes* **2022**, *13*, 1089. [[CrossRef](#)]
239. Boothe, M.; Morris, R.; Robin, N. Stickler Syndrome: A Review of Clinical Manifestations and the Genetics Evaluation. *J. Pers. Med.* **2020**, *10*, 105. [[CrossRef](#)]
240. Bank, R.A.; Verzijl, N.; Lafeber, F.P.; Tekoppele, J.M. Putative role of lysyl hydroxylation and pyridinoline cross-linking during adolescence in the occurrence of osteoarthritis at old age. *Osteoarthr. Cartil.* **2002**, *10*, 127–134. [[CrossRef](#)]
241. Kannus, P. Structure of the tendon connective tissue. *Scand. J. Med. Sci. Sports* **2000**, *10*, 312–320. [[CrossRef](#)]
242. Bank, R.A.; Tekoppele, J.M.; Oostingh, G.; Hazleman, B.L.; Riley, G.P. Lysylhydroxylation and non-reducible crosslinking of human supraspinatus tendon collagen: Changes with age and in chronic rotator cuff tendinitis. *Ann. Rheum. Dis.* **1999**, *58*, 35–41. [[CrossRef](#)]
243. Xu, Y.; Murrell, G.A. The basic science of tendinopathy. *Clin. Orthop. Relat. Res.* **2008**, *466*, 1528–1538. [[CrossRef](#)]

244. de Mos, M.; van El, B.; DeGroot, J.; Jahr, H.; van Schie, H.T.; van Arkel, E.R.; Tol, H.; Heijboer, R.; van Osch, G.J.; Verhaar, J.A. Achilles tendinosis: Changes in biochemical composition and collagen turnover rate. *Am. J. Sports Med.* **2007**, *35*, 1549–1556. [[CrossRef](#)] [[PubMed](#)]
245. Chen, Y.; Terajima, M.; Yang, Y.; Sun, L.; Ahn, Y.H.; Pankova, D.; Puperi, D.S.; Watanabe, T.; Kim, M.P.; Blackmon, S.H.; et al. Lysyl hydroxylase 2 induces a collagen cross-link switch in tumor stroma. *J. Clin. Investig.* **2015**, *125*, 1147–1162. [[CrossRef](#)] [[PubMed](#)]
246. Eisinger-Mathason, T.S.; Zhang, M.; Qiu, Q.; Skuli, N.; Nakazawa, M.S.; Karakasheva, T.; Mucaj, V.; Shay, J.E.; Stangenberg, L.; Sadri, N.; et al. Hypoxia-dependent modification of collagen networks promotes sarcoma metastasis. *Cancer Discov.* **2013**, *3*, 1190–1205. [[CrossRef](#)] [[PubMed](#)]
247. Gilkes, D.M.; Bajpai, S.; Wong, C.C.; Chaturvedi, P.; Hubbi, M.E.; Wirtz, D.; Semenza, G.L. Procollagen lysyl hydroxylase 2 is essential for hypoxia-induced breast cancer metastasis. *Mol. Cancer Res.* **2013**, *11*, 456–466. [[CrossRef](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.