nature portfolio

Peer Review File

The structural basis for the human procollagen lysine hydroxylation and dual-glycosylation

Corresponding Author: Professor Yu Cao

This file contains all reviewer reports in order by version, followed by all author rebuttals in order by version.

Version 0:

Reviewer comments:

Reviewer #1

(Remarks to the Author)
The manuscript from Peng et al reports on the structural characterization of a complex involving two human enzymes responsible for collagen lysine post-translational modifications, namely LH3/PLOD3 and GLT25D1/ColGalT1. This is a very interesting topic highly suitable for consideration in a high profile journal like Nature Communications, considering the complexity of a fundamental process such as collagen biosynthesis and the implications that malfunctions in its post-translational modification machineries have in causing a large variety of pathologies.

Nevertheless, I cannot endorse publication of the current version of this manuscript. In my view, this work requires extensive revision and additional supporting experimental evidence prior to become considerable for publication, as several statements currently provided by the authors are either conflicting with the presented data or lack solid experimental evidence.

In particular:

- The most remarkable observation emerging from this work is the interaction interface between LH3/PLOD3 and GLT25D1/ColGalT1. This interface, presented at page 5 in the paragraph *"The interfaces mediating the KOGG complex formation"* is solely described based on CryoEM data obtained from samples clearly showing heterogeneous oligomeric ensembles not necessarily physiologically relevant, as highlighted by the 2D classes shown in Supplementary Figures 2-4. There is absolutely no biochemical validation to support the proposed contacts yielding the extremely unusual heterotetrameric assembly. The authors should make an effort generating LH3/PLOD3 and/or GLT25D1/ColGalT1 variants around the interaction site to demonstrate that the proposed contact interactions are not artifacts of CryoEM sample preparation.

- Related to my previous point, the description of possible LH3/PLOD3:GLT25D1/ColGalT1 fiber-like polymers is extremely weak and may entirely be associated with CryoEM sample preparation artifacts. In absence of supporting experimental evidence, the section at page 7 dedicated to *"The polymerization of KOGG quaternary complex"*, together with figure 5 and all associated comments in the manuscript abstract, in the discussion, as well as in the supplementary movie should either be corroborated by solid experimental evidence or be removed integrally from the manuscript.

- The PDB files and their associated validation reports provided in the reviewing package allowed additional evaluation of the proposed macromolecular assemblies, further highlighting the weakness exposed in my previous comments. Besides the "core" LH3/PLOD3 dimer and the GLT25D1/ColGalT1 dimer, the 3D reconstruction is fuzzy and large segments of the proposed protein ensembles, even in the proposed 2:2:2:2 complex assemblies provided in PDB files 8ZGC, 8ZGG, and 8ZGH are not covered by experimental density. This is further highlighted by the models provided, which do not contain side chains for the outmost LH3/PLOD3 and GLT25D1/ColGalT1 dimers. Based on the current data, I suggest the authors to focus (and validate) the LH3/PLOD3:GLT25D1/ColGalT1 complex interfaces, prior to speculating about large oligomeric assemblies currently not supported by experimental evidence.

- Additionally, I recommend the authors to increase their modelling efforts towards correctness for protein and ligand geometries, as the PDB validation reports highlight an extensive number of geometry outliers in nearly all structures accompanying this manuscript.

- The authors do not provide a solid rationale for the presence of UDP-sugar donor substrates in the GalT-N domain of GLT25D1/COLGALT1 in the structures obtained without supplementing UDP-Gal to the sample. Are the authors stating that GLT25D1/COLGALT1 bears two functional catalytic sites in two neighboring domains within each monomer? The analysis of the catalytic sites of this enzyme provided at page 5 is very poor for such unexpected and unusual observation, and the results obtained from the mutagenesis experiments and biochemical assays reported at page 6 are not conclusive, rendering the final statement of the associated paragraph *"Nevertheless, the robust enzymatic capacity demonstrated by the UDP-only local structure of GalT-C implies its primary role in galactosylation, warranting further investigation"* extremely weak.

- Regarding biochemical characterization of LH3/PLOD3 and GLT25D1/ColGalT1 mutants, evidence is missing regarding the proper folding and complex assembly for the various mutants that have been presented. SDS-PAGE and size exclusion chromatography analysis of all KOGG complexes containing enzyme mutants, should be included in the supplementary materials of a revised version.

- The authors also report as particularly noteworthy *"the identification of a long blob of density near the C1 carboxylate group of 2OG and Fe ion in the LH domain of KOGG supplemented with either UDP-Glc or UDP-Gal, a position characteristic of the hydroxyl-acceptor substrate in 2OG-dependent oxygenases, suggesting the presence of potential substrate peptides"*. Considering the large macromolecular nature of the procollagen substrate (assuming that such multi-enzyme:substrate complex would survive the cell extraction and chromatography steps used to purify the KOGG complexes), and considering that the CryoEM grids described in the present work were not carried out using any synthetic peptide substrate, the authors should have been able to identify possible procollagen molecules in their SDS-PAGE analysis. Furthermore, it should be noted that previous reports on LH3/PLOD3 indicate that these enzymes can adopt self-inhibited conformations (originally described in doi: 10.1038/s41467-018-05631-5 and analyzed in depth in doi: 10.3389/fmolb.2022.876352), with flexible loops occupying the exact same positions proximate to the 2OG cofactor described in this manuscript. I suggest the authors to carefully revise the entire contents of this section.

- Likewise, the comments at Page 7, lines 5-6, associated with the loose and fragmented (i.e., spurious and discontinuous) densities located near the β-phosphate group of UDP in the GlcT domain of LH3/PLOD3 should be removed, as they unlikely support the possible presence of galactosyl or glucosyl moieties in low occupancy given the low resolution and poor definition of the density in that region.

- The discussion section analyzing LH3/PLOD3 and GLT25D1/ColGalT1 pathogenic mutations is presently addressing LH3/PLOD3 pathogenic mutations which were already mapped on the enzyme's 3D structure (doi: 10.1038/s41467-018-05631-5), and provides very limited analysis of the GLT25D1/ColGalT1 pathogenic mutations. The authors should either expand this section with a more in-depth analysis of the significance of the molecular mapping of GLT25D1/ColGalT1 pathogenic mutations on the enzyme structure, or remove this section and the associated Figure 6 from the manuscript.

Additional comments:

- The enzymes responsible for collagen lysine hydroxylation and glucosylation are mostly described in the literature as "multifunctional lysyl hydroxylases", with the "LH" acronym, whereas the "PLOD" name is used to describe the genes encoding for these proteins. I recommend using the LH nomenclature, or possibly combining both LH and PLOD nomenclature when discussing these enzymes throughout the manuscript text, graphics, and supplementary materials.

- Likewise, the enzymes responsible for hydroxylysine galactosylation are named GLT25D or COLGALT or ColGalT. The authors should either use only one of these names or use them both as suggested for LH/PLOD enzymes.

- Both LH3/PLOD3 and GLT25D1/ColGalT1 bear N-linked glycosylations, some of which were already reported for LH3/PLOD3 in previous crystal structures. The authors should comment on the N-linked glycosylations observed and modeled in their EM structures, and properly describe how they were represented in graphical elements such as in Figure 2, where there is no description of such features in the current figure legend.

- Page 5, line 7: RMSD values of 14.468 Å cannot be considered *"high similarity"*.

- For biochemical assays, the authors describe usage of a synthetic peptide (LPGTAGLPGMKGHRGFSGLDG) mimicking the collagen substrate region. Given the extensive presence of lysine residues on the numerous collagen polypeptide sequences, a more precise description of the reasoning associated with selection of the peptide sequence used for the biochemical assays is needed.

- Page 2, Line 29, recent publications highlight that also LH1/PLOD1 and LH2/PLOD2 possess glucosyltransferase activity. See doi: 10.1016/j.trsl.2021.08.002 and doi: 10.1038/s42003-021-01982-w.

- Page 4, Line 37, *"RMSD (root mean square deviations) around XXX-XXX Å"* requires attention.

- Page 6, line 14, *"and surrounding residues (Supplementary Figure XXX)"* requires attention.

- Page 6, line 35, Mn should be Mn^{2+}

Reviewer #2

(Remarks to the Author)
This work resolves the structure of the KOGG complex, which is composed of PLOD3 & ColGalT1. which forms an extended and probably quite flexible structure. I am reviewing it as an expert in single-particle cryo-EM rather than an expert in the biochemistry of this system. The paper is well-written and interesting for a subject-non-expert to read. While I have a few points I think the cryo-EM is in good shape and was clearly carefully carried out and reported.

Major points.
- The flexible nature of the system means that even with extensive classification the authors are working at modest resolution ('3.4 Å and a local resolution ranging from 2.5 Å to 6.5 Å at the core region of the complex') and the authors are clear about this limitation. They should have a supplementary figure showing density from each domain, including areas at 2.5, 3.5, and 6.5 A resolution.

Minor points
- The authors are careful to not overinterpret the oligomeric state observed when sugar-UDPs are added to the samples and I am aware that the 2D classes are in the supplementary but they are a nice way for nonexperts to get their head around the data so this may suit being a panel in fig 4.
- The authors use the term dose rather than fluence. While this is very common, technically fluence (particles or photons/ unit area) is correct rather than does (a measure of photons or particles being absorbed per unit mass). The authors should feel free to do what they think is best here.
- SupFig10- unsolved should be unidentified or unmodeled (as it is in other figures actually)

Reviewer #3

(Remarks to the Author)
The manuscript by Peng et al explores the structural basis of Human Procollagen Lysine hydroxylation and glycosylation. In collagen, Lysine can undergo hydroxylation (Hyl) followed by O-linked glycosylation to form Hyl-O-glyco which helps form mature collagen triple helices and cross-linking processes. Procollagen-lysine-2-oxoglutarate-5-dioxygenase (PLOD) family of enzymes (PLOD1-3) and procollagen galactosyltransferase 1 (COLGALT1) are responsible for the enzymatic conversion of Lysine to Lys-O-glyco. The manuscript provides the first Cryo-EM structure for the human PLOD3 and COLGALT1 enzyme complex, known as the KOGG complex which is a significant discovery in the field. This study showed that PLOD3 and COLGALT1 form a procollagen lysine hydroxylation and dual-glycosylation complex (KOGG complex), suggesting a coupling mechanism for lysyl hydroxylation-galactosylation-glucosylation reactions in collagen processing. The structures revealed a KOGG quaternary complex with a stoichiometry of 2:2 between the two enzymes in the PLOD3/COLGALT1 and the existence of an enzyme matrix formed by interactions between quaternary complexes. It also provided insights into the spatial configuration of catalytic domains of the hydroxylation-galactosylation-glucosylation cascade of procollagens and the details of substrate/product binding sites. However, the authors do not address multiple critical questions that are mandatory for the conclusions made in this study. Authors must address the following comments for further consideration of this manuscript for publication.
Major comments:
1. The author talks about the fiber-like arrangement of the KOGG complex without any biological context. The authors should provide some more evidence of a fiber-like arrangement through some additional biophysical characterization. It may also be an artefact of the protein purification process. The gel filtration profile of the purified complex shows a peak in void volume overlapping with the peak of the complex. There is the possibility of soluble aggregate coming in complex peak fraction. Authors should do a cryo-EM analysis for void volume fraction only. There is the possibility of getting a higher population of fiber-like structures. Does the fiber-like assembly have any impact on the enzymatic characteristic of the complex?
2. The RMSD values between GalT-N and AC (overall RMSD ~ 3.958 Å) and between GalT-C and GlcT (overall RMSD ~ 14.468 Å) are quite high. The structural alignment for GalT-N and AC shows the presence of a similar fold. However, GalT-C and GlcT structural alignments do not seem to converse also reflected in high RMSD values.
3. For the density of the ligand in both PLOD3 and COLGALT1, authors must make new figures only showing the density for the ligands with the contour level of the map. It would be also advisable to put the metal ion densities with the distances to the coordinated residues.
4. Fig 1d and 1e should be represented in a simple domain coloring scheme unless the authors want to emphasize on secondary structure contents of the domain. Fig 1e can be put into supplementary material as it does not add any valuable information to the main text or main figure.
5. Authors should discuss how this PLOD3-COLGALT1 complex could affect fibrillr collagens and basement membrane collagens?
Minor comments:
1. Line 35: Post-translational
2. Line 60: It mentioned only COLGALT1, not COLGALT2—no mention of COLGALT2 in this study being involved in the same process. COLGALT1 and COLGALT2 act as galactosyltransferase and come under the COLGALT/GLT25D family of enzymes. Why does the author choose only COLGALT1 galactosyltransferase in association with PLOD3 and not COLGALT2 in this study?
3. Line 132: What does XXX-XXX Å represent? No RMSD value on page 4 line number 132.
4. Line 188-189: "In ColGalT1, two DXD motifs have been identified as residues 166-168 in its GalT-N and residues 461-

463 in its GalT-C domain" (mentioned in this study).
No mention of the third DXD motif in this study.
5. Line 264: "The biological significance of KOGG's linear polymerization remains unclear, and we speculate that this long fiber-like arrangement might endow KOGG with the capacity to catalyze hydroxylation/glycosylation simultaneously on multiple lysyl sites on a single polypeptide chain, such as COL1A1, COL1A2, and COL4A1." PLOD3 is reported to hydroxylate lysine residues in COL4 but not in COL1. As per previous studies, this KOGG complex's lysine hydroxylation and dual-glycosylation features should be present in COL4 and only dual-glycosylation features in COL1.
6. In both, Figure 3.d and Figure 4.c glucosyltransferase activity is less than galactosyltransferase activity. In Figure 3.d, it is mentioned that the mutations are at the catalytic sites of GalT-N and GalT-C of KOGG complex. GalT-N and GalT-C are present in COLGALT1 and are associated with galactosyltransferase activity. But why in the case of PLOD3 the galactosyltransferase activity is more than that of glucosyltransferase activity?
7. Page 6, line 197, Supplementary Figure XXX (figure number not denoted).

Version 1:

Reviewer comments:

Reviewer #1

(Remarks to the Author)
The revised manuscript from Peng et al incorporates additional experiments that addressed my previous remarks in a proper and comprehensive way. The revised statements and new display items appropriately complete the proposed work. The biochemical validation of the proposed KOGG complex strengthens the insights provided by the Cryo-EM data. I also appreciated how the authors decided to tone down their speculative statements about polymerization of the complex.

I have no other remarks and congratulate the authors for this nice work.

Reviewer #2

(Remarks to the Author)
The authors have thoughtfully responded to my comments, which were cryo-EM focused.

I note that the two other reviewers, presumably who have much more subject-specific knowledge than me, also raised concerns that seem to have been addressed.

Jamie Blaza
University of York

Reviewer #3

(Remarks to the Author)
The authors have responded very satisfactorily.

Reviewer #1 (Remarks to the Author)

The manuscript from Peng et al reports on the structural characterization of a complex involving two human enzymes responsible for collagen lysine post-translational modifications, namely LH3/PLOD3 and GLT25D1/ColGalT1. This is a very interesting topic highly suitable for consideration in a high profile journal like Nature Communications, considering the complexity of a fundamental process such as collagen biosynthesis and the implications that malfunctions in its post-translational modification machineries have in causing a large variety of pathologies.

Nevertheless, I cannot endorse publication of the current version of this manuscript. In my view, this work requires extensive revision and additional supporting experimental evidence prior to become considerable for publication, as several statements currently provided by the authors are either conflicting with the presented data or lack solid experimental evidence.

In particular:

- The most remarkable observation emerging from this work is the interaction interface between LH3/PLOD3 and GLT25D1/ColGalT1. This interface, presented at page 5 in the paragraph *"The interfaces mediating the KOGG complex formation"* is solely described based on CryoEM data obtained from samples clearly showing heterogeneous oligomeric ensembles not necessarily physiologically relevant, as highlighted by the 2D classes shown in Supplementary Figures 2-4. There is absolutely no biochemical validation to support the proposed contacts yielding the extremely unusual heterotetrameric assembly. The authors should make an effort generating LH3/PLOD3 and/or GLT25D1/ColGalT1 variants around the interaction site to demonstrate that the proposed contact interactions are not artifacts of CryoEM sample preparation.

Our reply: Thank you for the questions and suggestions regarding the interactions between the enzymatic components of the KOGG complex. We recognize that structural data from CryoEM should be complemented by functional assays to validate the proposed contact interactions. In response, we introduced a series of mutations within the interface between PLOD3 and ColGalT1 and examined their effects on complex assembly and catalytic activities. The Co-IP assay revealed that mutations in the N-loop of ColGalT1, including single alanine substitutions at residues 37-42, 44, and 46, disrupted the formation of the KOGG complex, while leaving the expression level of ColGalT1 largely unaffected (Fig. R1a). Similarly, alanine substitutions at interacting residues on PLOD3, such as 229-231, 241, 264, 265, 446, and 452, also resulted in a decrease or complete loss of complex formation (Fig. R1b). These results suggest that the PLOD3-ColGalT1 heterotetrameric complex is not an artifact of cryo-EM sample preparation. Instead, the interactions between the N-loop of ColGalT1 and the LH & AC domains of PLOD3 are crucial for complex formation in the cell. We greatly appreciate the reviewer's emphasis on validating our structural discoveries, and we have incorporated these functional results into **Figure 2f &2g**, along with an expanded discussion in the revised manuscript.

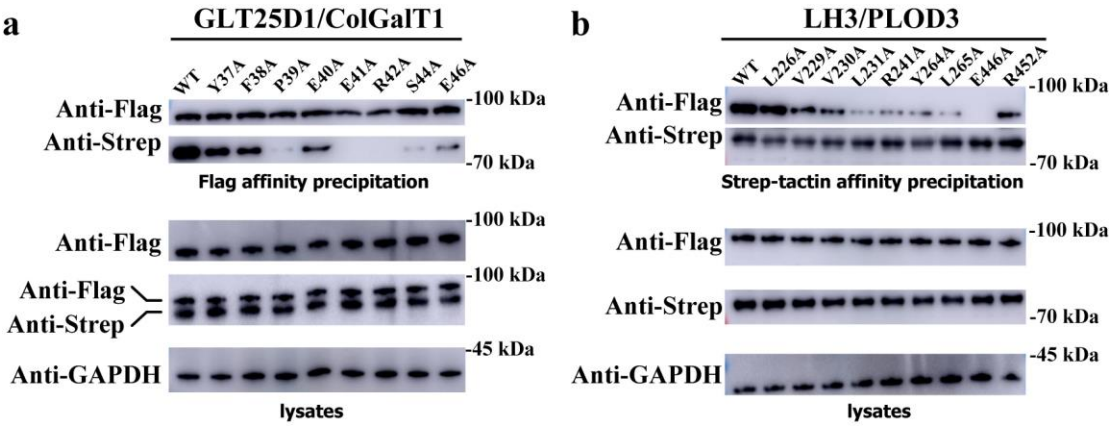

Fig. R1: The mutations in PLOD3-ColGalT1 interface impair formation of KOGG complex. a, The 293T cells were co-transfected with wild type PLOD3 with flag tag and ColGalT1 variants with Strep tag carrying mutations at its N-loop. **b,** The 293T cells were co-transfected with wild type ColGalT1 with Strep tag and PLOD3 variants with Flag tag carrying mutations at PLOD3-ColGalT1 interface.

- Related to my previous point, the description of possible LH3/PLOD3:GLT25D1/ColGalT1 fiber-like polymers is extremely weak and may entirely be associated with CryoEM sample preparation artifacts. In absence of supporting experimental evidence, the section at page 7 dedicated to *"The polymerization of KOGG quaternary complex"*, together with figure 5 and all associated comments in the manuscript abstract, in the discussion, as well as in the supplementary movie should either be corroborated by solid experimental evidence or be removed integrally from the manuscript.

Our reply: Thank you for your insightful comments regarding the potential polymerization of the PLOD3-ColGalT1complex. We share this concerns about the possibility of artifacts arising from the sample preparation process. In response, we have conducted additional assays to test the presence of higher oligomeric forms of the PLOD3-ColGalT1 complex *in vivo*.

To investigate the molecular behavior of PLOD3-ColGalT1 polymerization, we first performed a crosslinking assay with the purified PLOD3-ColGalT1complex prior to cryo-EM sample preparation. The SDS-PAGEs showed that upon incubation with the crosslinking agent $BS(PEG)_5$, the migration rates of both PLOD3 and ColGalT1 shifted, with appearance of bands corresponding to tetrameric PLOD3-ColGalT1 complex (300-400 kD), as well as several higher molecular weight bands, exceeding 400 kD (the maximum size available for molecular weight standards) indicating the presence of even higher oligomeric states (Fig. R2a).

To further explore the oligomeric state of endogenous PLOD3-ColGalT1 complexes, we performed native gel electrophoresis using cell lysates from untreated MC3T3-E1, SaoS-2, HEK293T, and Expi293F cell lines. Immunoblotting of the native gel detected KOGG bands corresponding to both tetrameric (300-400 kD) and higher molecular weight states, suggesting the existence of higher-order oligomers for the PLOD3-ColGalT1 complex. These additional experimental findings provide evidence for the existence of higher oligomeric forms of the

KOGG complex, which support the notion of polymerization. We have now incorporated these results into supplement materials to better substantiate the polymerization process description. However, we recognize that additional in-situ visualization and functional assays may be necessary to fully elucidate the biological role of the fiber-like form of the KOGG complex. In response, we tuned down the statement about the KOGG polymer and moved Figure 5 into the supplement materials as **Supplementary Figure 11**.

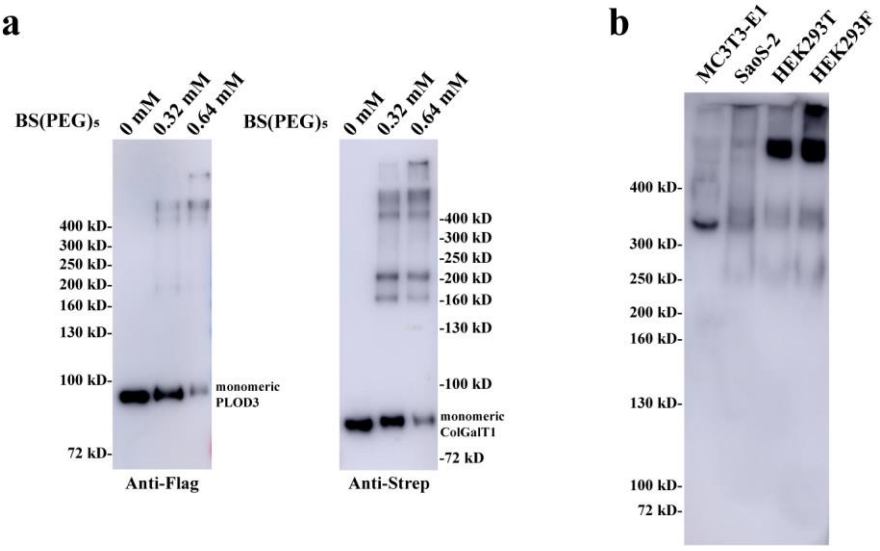

Fig. R2: The polymerization of PLOD3-ColGalT1 complex. a, The purified KOGG complex protein sample was incubated with crosslinking agent, $BS(PEG)_5$, and subjected to SDS-PAGE. **b,** The cell lysates from MC3T3-E1, SaoS-2, HEK293T, and Expi293F cell culture was subjected to native-gel electrophoresis and visulized with anti-ColGalT1 antibody.

- The PDB files and their associated validation reports provided in the reviewing package allowed additional evaluation of the proposed macromolecular assemblies, further highlighting the weakness exposed in my previous comments. Besides the "core" LH3/PLOD3 dimer and the GLT25D1/ColGalT1 dimer, the 3D reconstruction is fuzzy and large segments of the proposed protein ensembles, even in the proposed 2:2:2:2 complex assemblies provided in PDB files 8ZGC, 8ZGG, and 8ZGH are not covered by experimental density. This is further highlighted by the models provided, which do not contain side chains for the outmost LH3/PLOD3 and GLT25D1/ColGalT1 dimers. Based on the current data, I suggest the authors to focus (and validate) the LH3/PLOD3:GLT25D1/ColGalT1 complex interfaces, prior to speculating about large oligomeric assemblies currently not supported by experimental evidence.

Our reply: Thank you for raising important points and providing suggestions regarding the polymerization of the KOGG complex. We agree with the reviewer's observation that the map quality in the elongating regions of the cryo-EM reconstruction for the KOGG complex in its oligomeric state is suboptimal. The resolution and continuity in areas outside the core complex are below average, making it challenging to construct reliable molecular models for those regions However, when visualized at a lower contour level, we observe clear elongating

polymerization with repeated 2+2 components, as demonstrated in Figure 5 of the original manuscript. In addition, our native gel electrophoresis and crosslinking assays further support the existence of such large oligomeric assemblies.
After carefully considering the reviewer's comments and concerns, we have decided to remove Figure 5 from the main text. These results, along with newly obtained data from native gel electrophoresis and crosslinking assays, have been moved to the supplementary materials as **Supplementary Figure 11**. Furthermore, we have revised the manuscript to soften statements about the KOGG polymer, explicitly noting the limitations of our structural model and the need for further validation. We hope these revisions improve the clarity and robustness of our manuscript and address the reviewer's concerns.

- Additionally, I recommend the authors to increase their modelling efforts towards correctness for protein and ligand geometries, as the PDB validation reports highlight an extensive number of geometry outliers in nearly all structures accompanying this manuscript.

Our reply: Thank you for the questions and suggestions regarding the geometry statistics of the validation reports. We have carefully reviewed the structural models and made necessary modifications during the revision process to improve the protein and ligand geometries. After further refinement, we observed slightly improved geometry in three of the four structural models. As a result, we have updated these three structural coordinate files and have provided the corresponding updated validation reports for them.

- The authors do not provide a solid rationale for the presence of UDP-sugar donor substrates in the GalT-N domain of GLT25D1/COLGALT1 in the structures obtained without supplementing UDP-Gal to the sample. Are the authors stating that GLT25D1/COLGALT1 bears two functional catalytic sites in two neighboring domains within each monomer? The analysis of the catalytic sites of this enzyme provided at page 5 is very poor for such unexpected and unusual observation, and the results obtained from the mutagenesis experiments and biochemical assays reported at page 6 are not conclusive, rendering the final statement of the associated paragraph *"Nevertheless, the robust enzymatic capacity demonstrated by the UDP-only local structure of GalT-C implies its primary role in galactosylation, warranting further investigation"* extremely weak.

Our reply: Thank you for the thoughtful questions and suggestions regarding the function of the GalT domains in ColGalT1. The biological roles of the two GalT domains in ColGalT1 have long been a subject of debate, as mutations in both domains can disrupt enzymatic activity. Previous studies involving chimera proteins, where ColGalT1 segments were fused with CEECAM1 (a non-glycosylating homolog of ColGalT1), showed that the chimera with the N-terminal domain of CEECAM1 and the C-terminal segment of ColGalT1 remained active, whereas the chimera with the N-terminal domain of ColGalT1 and the C-terminal segment of CEECAM1 lost glycosyltransferase activity [1]. This observation suggested that while the N-terminal structural motif, GalT-N, is essential for the functional integrity of ColGalT1, it might not play a direct catalytic role but rather contributes to maintaining the enzyme's structural stability.

Our structural findings align well with these previous observations. Cryo-EM analysis revealed that both GT domains in ColGalT1 adopt a GT-A folding, with substrate density observed in the cavities of both the GalT-N and GalT-C domains. However, in the GalT-N domain, UDP-glucose remains intact, whereas in the GalT-C domain, the sugar moiety is lost, indicating uncoupled enzymatic activity. A closer examination of the local structures around the substrate-interacting residues in GalT-N and GalT-C reveals a hydrophobic environment around the β-face of the phosphate-sugar bond in GalT-N, with no side-chain carboxylate available to serve as the base catalyst required for an inverting glycosyltransferase reaction [2]. The only carboxylate-containing residue near the β-phosphate in GalT-N, D265, is positioned on the distal side of the sugar anomeric carbon, making it unlikely to participate directly in the glycosylation reaction. In contrast, in the GalT-C domain, two carboxylates - D522 and E572 - are located near the β-phosphate of UDP (Fig. R3). Structural comparisons with other functional GT-A enzymes, such as Bacillus subtilis SpsA (PDB ID: 1QGQ) and rabbit N-acetylglucosaminyltransferase GNT-I (PDB ID: 1FOA), reveal that D522 in ColGalT1 is structurally related to the catalytic residues D191 in SpsA and D291 in GNT-I (Fig. R3), positioning it as a likely candidate for the base catalyst in the inverting mechanism typical of Clan I GT-A enzymes [2,3].

In summary, our structural analysis suggests only one functional catalytic site in GLT25D1/COLGALT1, located in the GalT-C domain. Although the GalT-N domain also adopts a GT-A folding similar to GalT-C, the arrangement of local residues in its substrate-binding pocket indicates that it likely lacks catalytic activity, as evidenced by the intact UDP-sugar bound in this domain. We appreciate the reviewer's suggestion to further elaborate on the differences between the GT domains of ColGalT1. In response, we have revised the discussion on the structure and function of ColGalT1 in the resubmitted manuscript. The structural comparison shown in Fig. R3 has also been included as a **Supplementary Fig. 9c** for further clarification.

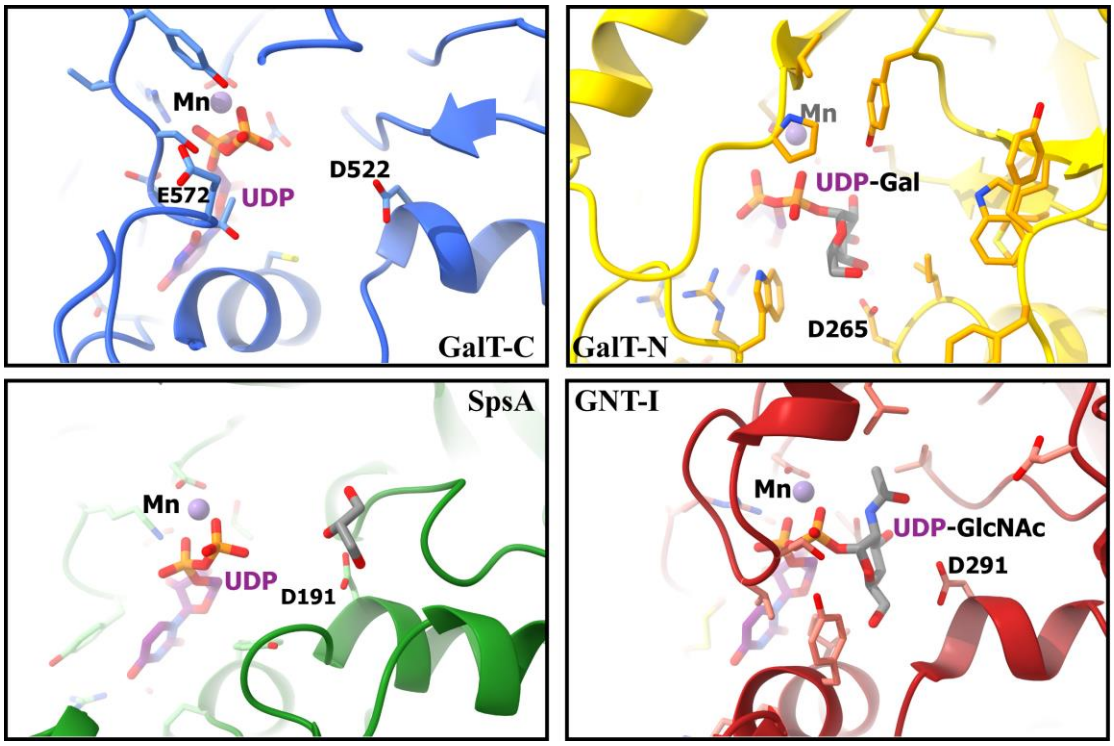

Fig. R3: The structural comparison among the GalT domains of ColGalT1 and previously reported GT-A folding enzymes. The substrate-binding cavities of GalT-C, GalT-N, SpsA, and GNT-I were enlarged and shown as cartoon-stick mixed models. The molecular models of SpsA and GNT-I were generated using the coordinate files with PDB IDs 1qgq and 1foa, respectively.

- Regarding biochemical characterization of LH3/PLOD3 and GLT25D1/ColGalT1 mutants, evidence is missing regarding the proper folding and complex assembly for the various mutants that have been presented. SDS-PAGE and size exclusion chromatography analysis of all KOGG complexes containing enzyme mutants, should be included in the supplementary materials of a revised version.

Our reply: Thank you for the questions and suggestions regarding the quality control of folding and complex assembly for the KOGG complex mutants. We agree that it is essential to verify the proper folding and complex formation for the mutants to draw reliable conclusions on their functional effects. In response, we purified all mutant PLOD3 and ColGalT1 proteins used in the biochemical assays and performed size exclusion chromatography (SEC) analysis, as summarized in Fig. R4. The results showed that most mutants of PLOD3 and ColGalT1 were expressed and purified at comparable levels, with the exceptions of PLOD3^{H719A} and PLOD3^{G256A}, which had lower yields. Moreover, SDS-PAGE and SEC-FPLC analyses confirmed that most mutants retained their ability to form the PLOD3-ColGalT1 complex. We appreciate the reviewer's suggestion to include these quality control data, and we have now incorporated the results into the supplementary materials as **Supplementary Figure 1e** for the revised manuscript.

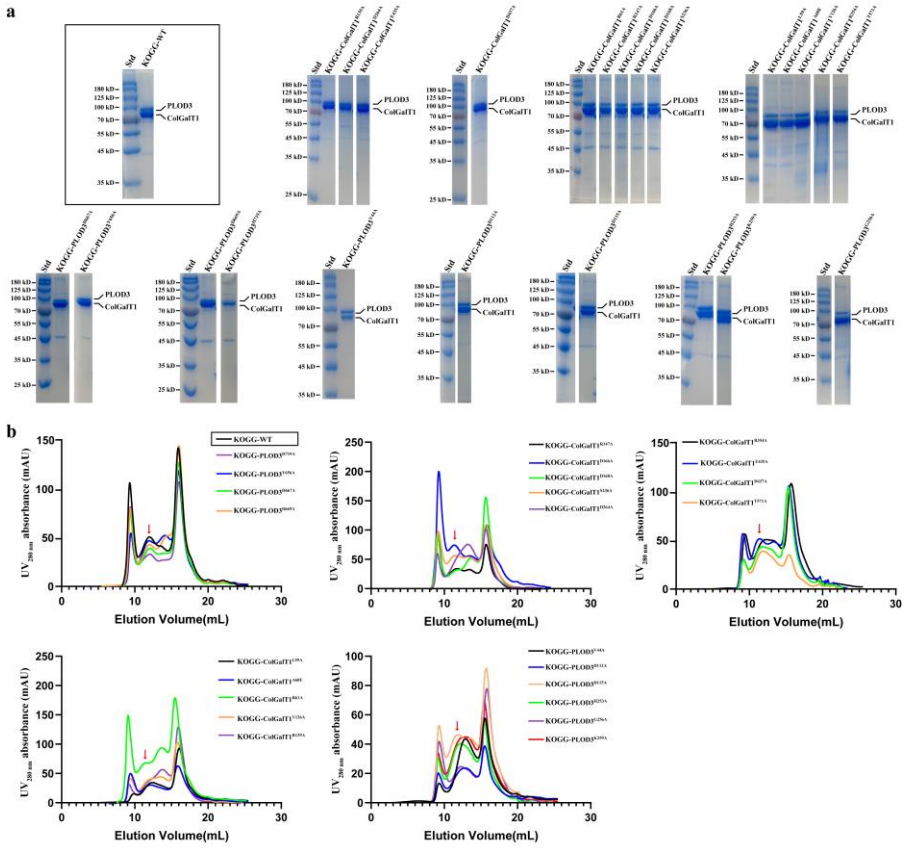

Fig. R4: The purification of KOGG complexes with functional mutations. a, SDS-PAGE analysis of the elutates from the Strep-tag affinity chromatography of KOGG complex with the mutations as indicated above the gels. **b,** size-exclusion chromatography profiles of the KOGG complexes with functional mutations. The elution fractions corresponding to the heterotetrameric PLOD3-ColGalT1 complex were indicated with red arrows.

- The authors also report as particularly noteworthy *"the identification of a long blob of density near the C1 carboxylate group of 2OG and Fe ion in the LH domain of KOGG supplemented with either UDP-Glc or UDP-Gal, a position characteristic of the hydroxyl-acceptor substrate in 2OG-dependent oxygenases, suggesting the presence of potential substrate peptides"*. Considering the large macromolecular nature of the procollagen substrate (assuming that such multi-enzyme:substrate complex would survive the cell extraction and chromatography steps used to purify the KOGG complexes), and considering that the CryoEM grids described in the present work were not carried out using any synthetic peptide substrate, the authors should have been able to identify possible procollagen molecules in their SDS-PAGE analysis. Furthermore, it should be noted that previous reports on LH3/PLOD3 indicate that these enzymes can adopt self-inhibited conformations (originally described in doi: 10.1038/s41467-018-05631-5 and analyzed in depth in doi: 10.3389/fmolb.2022.876352), with flexible loops occupying the exact same positions proximate to the 2OG cofactor described in this manuscript. I suggest the authors to carefully revise the entire contents of this section.

Our reply: Thank you for your insightful questions and suggestions regarding the unresolved density near the C1 carboxylate group of 2OG and Fe2+ in the LH domain of PLOD3. We understand the reviewer's concern and share the curiosity about the identity of this density. As you correctly pointed out, the presence of procollagen in our KOGG sample is plausible, and we did indeed detect COL1A1 in the purified KOGG complex using Western blot analysis (Fig. R5a). However, due to the discontinuous and low-resolution nature of the non-proteinaceous density, it was challenging to model the substrate peptide confidently into the map. Consequently, we refrained from incorporating the collagen sequence into this region. Additionally, we recognize the previous studies by the Forneris group, which describe a flexible capping loop (residues 590-610) near the catalytic center of the LH domain playing a gating role in response to Fe concentration. In response, we closely examined the local map density near the unresolved region. Our analysis revealed that two potentially flexible segments -- the loop$^{588\text{-}612}$ (capping loop) and the C-terminal segment of PLOD3 (residues 729-738) -- were well-resolved with clear and continuous density (Fig. R5b). While this observation suggests that the unresolved density may not correspond to the capping loop, we found that structural superposition between our cryo-EM structure and the previously published crystal structures [4,5] showed a significant displacement of the capping loop (Fig. R5c). In the crystal structure, the second Fe ion prevents the capping loop from moving, maintaining a closed conformation that limits space around the 2OG, thus decreasing accessibility to the substrate peptide. In contrast, in our cryo-EM structure, the capping loop moves outward in the absence of the second Fe ion, thereby opening up space in the reaction

center and potentially allowing the substrate to enter. Given the current cryo-EM data, it remains challenging to definitively determine whether the unresolved density corresponds to the alternative conformation of the capping loop or to the substrate procollagen, which could have a low occupancy. In response to your suggestion, we have revised the manuscript to include a more thorough structural analysis and comparison with the crystal structures, along with the Figure R5 as part of **Supplementary Figure 10**. This revision addresses both our new findings and the uncertainties surrounding the unresolved density. We deeply thank the reviewer for bringing our attention to the capping loop and auto-inhibition mechanism of PLODs, as it has provided valuable insight into the conformational differences of LH3 in low and high Fe conditions. This has certainly improved the quality and depth of our manuscript.

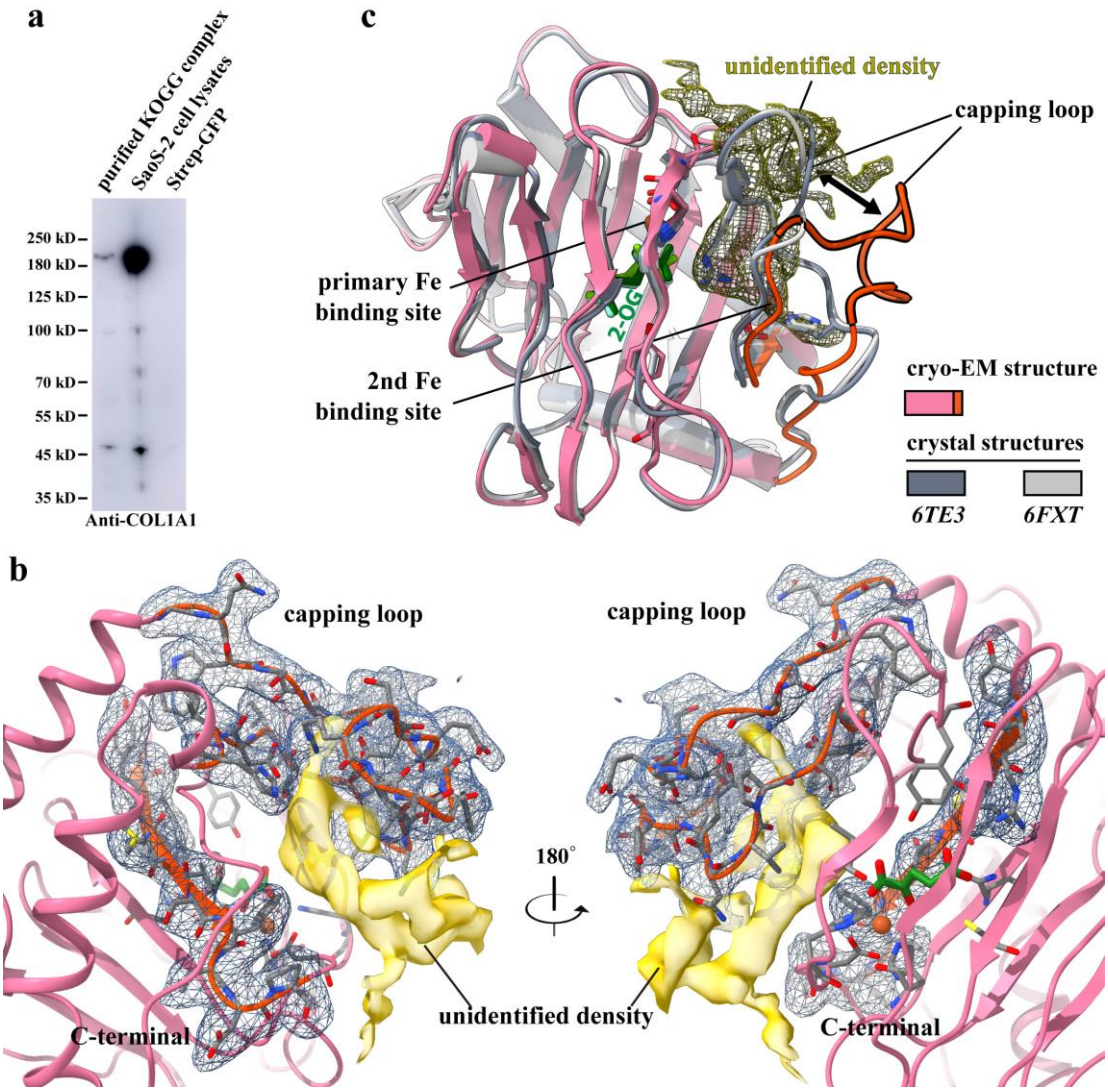

Fig. R5: The analysis on the unidentified density near the lysyl hydroxylation site. a, the presence of collagen proteins in the cryo-EM samples. **b,** the local cryo-EM map density near the lysyl hydroxylation site of LH3/PLOD3. The LH domain was depicted as cartoon mode in pink, with the residue 588-612 (capping loop) and 729-738 (C-terminal segment) highlighted in orange. The map densities enclosing the capping loop and the C-terminal segment were shown as blue mesh, with the unidentified density as a yellow surface model. **c,** The structural superposition among the LH domains from KOGG/UDP-Gal structure

(pink) and crystal structures of LH3/PLOD3, generated using the coordinate files with PDB IDs 6TE3 (gray-blue) and 6FXT (gray). The unidentified density was shown as a yellow mesh.

- Likewise, the comments at Page 7, lines 5-6, associated with the loose and fragmented (i.e., spurious and discontinuous) densities located near the β-phosphate group of UDP in the GlcT domain of LH3/PLOD3 should be removed, as they unlikely support the possible presence of galactosyl or glucosyl moieties in low occupancy given the low resolution and poor definition of the density in that region.

Our reply: Thank you for the suggestions regarding the unresolved density near the UDP bound in the GlcT domain. After a careful re-examination of the non-proteinaceous density in this region, we acknowledge the concerns about the density quality. As a result, given the poor resolution and lack of definition in this area, we have removed the sentences that previously speculated on the presence of galactosyl or glucosyl moieties at low occupancy. Instead, we now describe these as sparse density blobs, with the molecular attribution still to be determined. We appreciate the reviewer's input, which has helped improve the precision and clarity of our manuscript.

- The discussion section analyzing LH3/PLOD3 and GLT25D1/ColGalT1 pathogenic mutations is presently addressing LH3/PLOD3 pathogenic mutations which were already mapped on the enzyme's 3D structure (doi: 10.1038/s41467-018-05631-5), and provides very limited analysis of the GLT25D1/ColGalT1 pathogenic mutations. The authors should either expand this section with a more in-depth analysis of the significance of the molecular mapping of GLT25D1/ColGalT1 pathogenic mutations on the enzyme structure, or remove this section and the associated Figure 6 from the manuscript.

Our reply: Thank you for your suggestion to expand the analysis of pathogenic mutations in the KOGG complex. In the revised manuscript, we have included a more detailed discussion of the disease-related mutations found in both LH3 and ColGalT1. This includes several mutations identified in mice that result in severe developmental defects affecting the craniofacial, ocular, and digital structures. We hope this expanded analysis will provide insights for researchers and clinicians, furthering the understanding of the molecular pathology underlying collagen-related disorders and diseases.

Additional comments:

- The enzymes responsible for collagen lysine hydroxylation and glucosylation are mostly described in the literature as "multifunctional lysyl hydroxylases", with the "LH" acronym, whereas the "PLOD" name is used to describe the genes encoding for these proteins. I recommend using the LH nomenclature, or possibly combining both LH and PLOD nomenclature when discussing these enzymes throughout the manuscript text, graphics, and supplementary materials.

Our reply: We thank the reviewer for the helpful suggestion regarding the nomenclature in this manuscript. We notice that "LH3" is widely used in the literature to refer to the protein. To improve the clarity of our manuscript, we have revised the text to introduce the full name

"multifunctional procollagen lysine hydroxylase and glycosyltransferase LH3" in the ***Introduction*** session, and we now use "LH3" or "LH3/PLOD3" consistently throughout the text and figures.

- Likewise, the enzymes responsible for hydroxylysine galactosylation are named GLT25D or COLGALT or ColGalT. The authors should either use only one of these names or use them both as suggested for LH/PLOD enzymes.

Our reply: Following the same nomenclature approach, we have revised the manuscript to consistently use the "GLT25D/ColGalT1" designation when referring to the enzyme encoded by the COLGALT1 gene. For conciseness, we retained the use of "ColGalT1" in instances such as the dimeric interface or "LH3-ColGalT1 interface," as it is also a recommended name in the UniProt database.

- Both LH3/PLOD3 and GLT25D1/ColGalT1 bear N-linked glycosylations, some of which were already reported for LH3/PLOD3 in previous crystal structures. The authors should comment on the N-linked glycosylations observed and modeled in their EM structures, and properly describe how they were represented in graphical elements such as in Figure 2, where there is no description of such features in the current figure legend.

Our reply: Thank you for your suggestions regarding the N-linked glycosylations. In the revised manuscript, we have now included a description of the N-linked glycosylation sites identified in the cryo-EM structure within the ***result*** session. Additionally, we have updated the legend of **Figure 2** to clarify the graphical representation of the N-glycans shown in the illustration.

- Page 5, line 7: RMSD values of 14.468 Å cannot be considered *"high similarity"*.

Our reply: We appreciate the reviewer's insight and have revised the manuscript accordingly. The updated description now reads:
"*Interestingly, structural comparisons reveal a high similarity between GalT-N and AC (overall RMSD ~ 3.958 Å) and a moderate similarity between GalT-C and GlcT (overall RMSD ~ 14.468 Å)…….*"

- For biochemical assays, the authors describe usage of a synthetic peptide (LPGTAGLPGMKGHRGFSGLDG) mimicking the collagen substrate region. Given the extensive presence of lysine residues on the numerous collagen polypeptide sequences, a more precise description of the reasoning associated with selection of the peptide sequence used for the biochemical assays is needed.

Our reply: Thank you for your comments and suggestions. The synthetic peptide used, *NH_2*-LPGTAGLPGMKGHRGFSGLDG-*COOH*, is derived from the human collagen alpha-1(I) chain (Uniprot ID P02452, residues 255-275), a segment that is conserved between humans and mice. Specifically, K265 corresponds to α1 Lys-87 in the triple-helical region of type I collagen and represents a major substrate site for the lysyl hydroxylation and glycosylation by LH3 [6,7]. In prior studies, a shorter version of this peptide, *NH_2*-GTAGLPGMKGHRGFS-*COOH*, was also used in functional assays of LH3/PLOD3 [8]. For clarity, we have now included this explanation in the revised manuscript.

- Page 2, Line 29, recent publications highlight that also LH1/PLOD1 and LH2/PLOD2 possess glucosyltransferase activity. See doi: 10.1016/j.trsl.2021.08.002 and doi: 10.1038/s42003-021-01982-w.

Our reply: Thank you for highlighting the recent publications regarding glucosyltransferase activity in LH1/PLOD1 and LH2/PLOD2. While LH3 has been traditionally considered the primary glucosyltransferase in the PLOD family, we acknowledge the new findings that suggest LH1 and LH2 also exhibit glucosyltransferase activity *in vitro* [9,10], albeit at lower levels compared to LH3 [11,12]. These recent reports expand our understanding of the PLOD family's biological role, and we have updated the introduction to reflect these recent advancements in collagen biology.

- Page 4, Line 37, *"RMSD (root mean square deviations) around **XXX-XXX** Å"* requires attention.

Our reply: Thank you for pointing out the incomplete information in our manuscript. We have revised it to incorporate the RMSD values.

- Page 6, line 14, *"and surrounding residues (Supplementary Figure **XXX**)"* requires attention.

Our reply: Thank you for pointing out the incomplete information in our manuscript. We have revised it to provide the figure number (**Supplementary Figure 9a**).

- Page 6, line 35, **Mn should be Mn^{2+}**

Our reply: Thank you. We have made revisions to correct the valence information accordingly.

Reviewer #2 (Remarks to the Author)

This work resolves the structure of the KOGG complex, which is composed of PLOD3 & ColGalT1. which forms an extended and probably quite flexible structure. I am reviewing it as an expert in single-particle cryo-EM rather than an expert in the biochemistry of this system. The paper is well-written and interesting for a subject-non-expert to read. While I have a few points I think the cryo-EM is in good shape and was clearly carefully carried out and reported.

Major points.

- The flexible nature of the system means that even with extensive classification the authors are working at modest resolution ('3.4 Å and a local resolution ranging from 2.5 Å to 6.5 Å at the core region of the complex') and the authors are clear about this limitation. They should have a supplementary figure showing density from each domain, including areas at 2.5, 3.5, and 6.5 A resolution.

Our reply: Thank you for the suggestions regarding the local resolution map for individual domains of KOGG. We have revised **Supplementary Fig. 2** to add a panel showing local resolution density for GalT-N, GalT-C, GlcT, AC, and LH domains.

Minor points

- The authors are careful to not overinterpret the oligomeric state observed when sugar-UDPs are added to the samples and I am aware that the 2D classes are in the

supplementary but they are a nice way for nonexperts to get their head around the data so this may suit being a panel in fig 4.

Our reply: Thank you for suggesting adding 2D class graphs to compare the high and low oligomeric states of KOGG. We have revised the figure to add representative 2D graphs accordingly.

- The authors use the term dose rather than fluence. While this is very common, technically fluence (particles or photons/ unit area) is correct rather than does (a measure of photons or particles being absorbed per unit mass). The authors should feel free to do what they think is best here.

Our reply: Thank you for pointing that out. We have revised the manuscript to ensure the correct terminology is used, changing "dose" to "fluence" where appropriate.

- SupFig10- unsolved should be unidentified or unmodeled (as it is in other figures actually)

Our reply: Thank you for pointing that out. We have revised the manuscript to replace "unsolved" density with "unidentified" density in **Supplementary Fig. 10** and **Figure 4b**.

Reviewer #3 (Remarks to the Author):

The manuscript by Peng et al explores the structural basis of Human Procollagen Lysine hydroxylation and glycosylation. In collagen, Lysine can undergo hydroxylation (Hyl) followed by O-linked glycosylation to form Hyl-O-glyco which helps form mature collagen triple helices and cross-linking processes. Procollagen-lysine-2-oxoglutarate-5-dioxygenase (PLOD) family of enzymes (PLOD1-3) and procollagen galactosyltransferase 1 (COLGALT1) are responsible for the enzymatic conversion of Lysine to Lys-O-glyco. The manuscript provides the first Cryo-EM structure for the human PLOD3 and COLGALT1 enzyme complex, known as the KOGG complex which is a significant discovery in the field. This study showed that PLOD3 and COLGALT1 form a procollagen lysine hydroxylation and dual-glycosylation complex (KOGG complex), suggesting a coupling mechanism for lysyl hydroxylation-galactosylation-glucosylation reactions in collagen processing. The structures revealed a KOGG quaternary complex with a stoichiometry of 2:2 between the two enzymes in the PLOD3/COLGALT1 and the existence of an enzyme matrix formed by interactions between quaternary complexes. It also provided insights into the spatial configuration of catalytic domains of the hydroxylation-galactosylation-glucosylation cascade of procollagens and the details of substrate/product binding sites. However, the authors do not address multiple critical questions that are mandatory for the conclusions made in this study. Authors must address the following comments for further consideration of this manuscript for publication.

Major comments:

1. The author talks about the fiber-like arrangement of the KOGG complex without any biological context. The authors should provide some more evidence of a fiber-like arrangement through some additional biophysical characterization. It may also be an artefact of the protein purification process. The gel filtration profile of the purified complex shows a peak in void volume overlapping with the peak of the complex. There

is the possibility of soluble aggregate coming in complex peak fraction. Authors should do a cryo-EM analysis for void volume fraction only. There is the possibility of getting a higher population of fiber-like structures. Does the fiber-like assembly have any impact on the enzymatic characteristic of the complex?

Our reply: Thank you for the questions and suggestions regarding the polymerization of the KOGG quaternary complex. We understand the reviewer's concerns about the fiber-like arrangement's biological relevance and acknowledge the necessity of excluding the artifact of the protein purification process. Based on the reviewer's suggestion, we first analyzed the void volume fraction with SDS-PAGE and electronic microscopy. No PLOD3 or ColGalT1 could be identified in the void volume fractions by SDS-PAGE (Fig R6a). Additionally, no fiber or other high-order polymer form of protein could be observed via EM methods (Fig R6b). For the enzymatic activities of the fiber-like assembly, we are unable to conduct a biochemical assay due to the highly dynamic among the multi-oligomeric states of KOGG and the lack of a separation approach to obtain fiber-like species.

We propose the polymerization of KOGG based on the observation of the cryo-EM map density corresponding to the elongation of KOGG with additional PLOD3-ColGalT1 tetramer units. Although the continuity and resolution of the elongating density prevent reliable molecular modeling, the overall shape and size align well with the higher polymerization of KOGG. Nevertheless, we acknowledge that more structural evidence and functional data are required to establish *in vivo* presence of KOGG polymer and elucidate its biological role. To avoid overinterpreting the oligomeric state of KOGG, we decide to tune down the statement about the KOGG polymer and move Figure 5 into the supplement materials as **Supplementary Fig. 11**. We hope this revision would improve the precision and clarity of our manuscript.

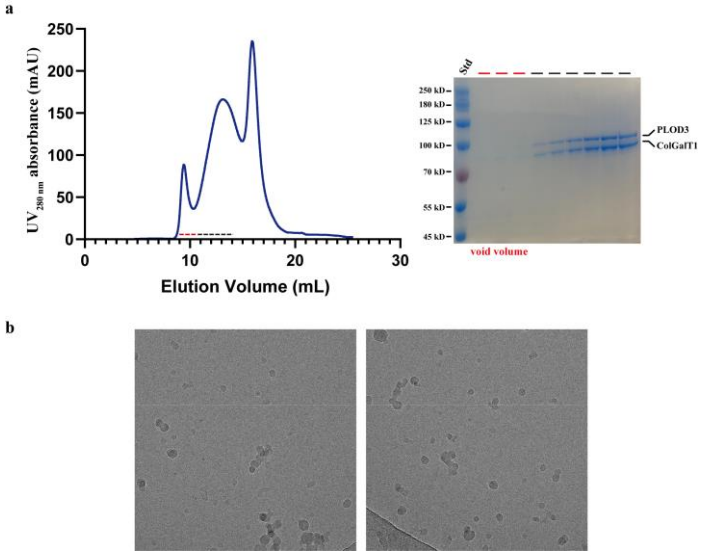

Fig. R6: The analysis of polymerization of KOGG complex. a, The SEC-FPLC profile of purified KOGG complex (left) and the SDS-PAGE analysis of the void volume. b, the cryo-EM graphs of the samples prepared from the fractions corresponding to the void volume.

2. The RMSD values between GalT-N and AC (overall RMSD ~ 3.958 Å) and between GalT-C and GlcT (overall RMSD ~ 14.468 Å) are quite high. The structural alignment

for GalT-N and AC shows the presence of a similar fold. However, GalT-C and GlcT structural alignments do not seem to converse also reflected in high RMSD values.

Our reply: Thank you for the comments regarding the structural similarity between GalT-C and GlcT domains. We agree that an RMSD value over 10 Å should not be classified as "high similarity" as suggested in the original manuscript. However, despite the higher RMSD value, GalT-C and GlcT domains do adopt similar Rossmann folds, with conserved spatial arrangements of key secondary structural motifs. To clarify this similarity, we focused the structural superposition on the core β-strands and helices of both domains, which yielded a more reliable comparison with a local RMSD of 5.385 Å (Fig. R7). To improve the precision of the manuscript, we have revised the ***results*** section to describe the similarity between GalT-C and GlcT as "moderate."

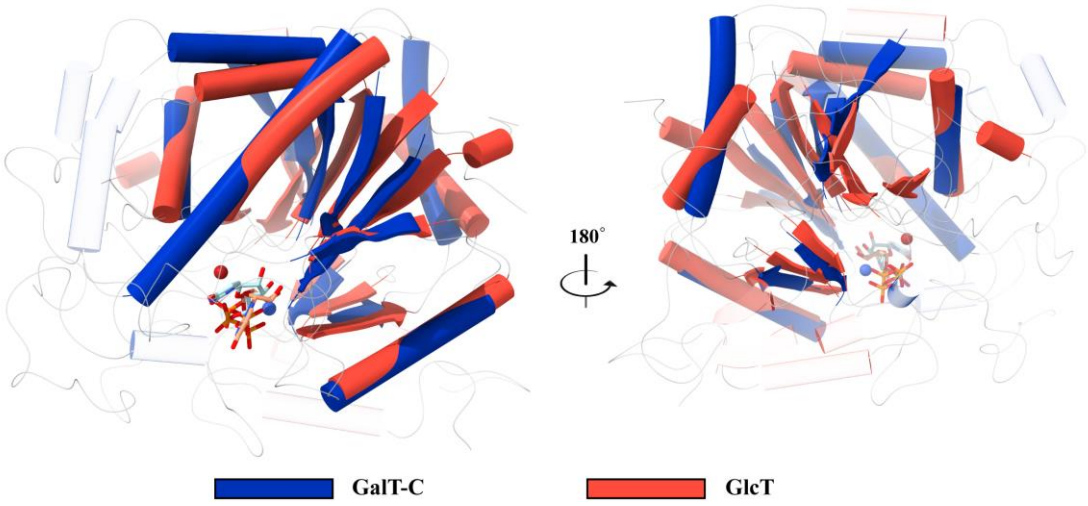

Fig. R7: The structural comparison between GalT-C and GlcT domains. The GalT-C domain from GLT25D1/ColGalT1 and the GlcT domain from LH3/PLOD3 were aligned and shown as cartoon models. The core helices and strands used for the superposition are highlighted in blue for GalT-C and red for GlcT, respectively. Outlying secondary structures, which were excluded from the superposition, are displayed with semi-transparency.

3. For the density of the ligand in both PLOD3 and COLGALT1, authors must make new figures only showing the density for the ligands with the contour level of the map. It would be also advisable to put the metal ion densities with the distances to the coordinated residues.

Our reply: Thank you for suggesting presenting the density for the key ligands bound in KOGG complex. We acknowledge the importance of showing details for those critical substrates and moieties in the cryo-EM map and have generated the map and marked the distances between metal ions and coordinating residues, as shown in Fig. R8 below. In the revised manuscript, this figure was added in to supplementary materials as **Supplementary Fig. 12**.

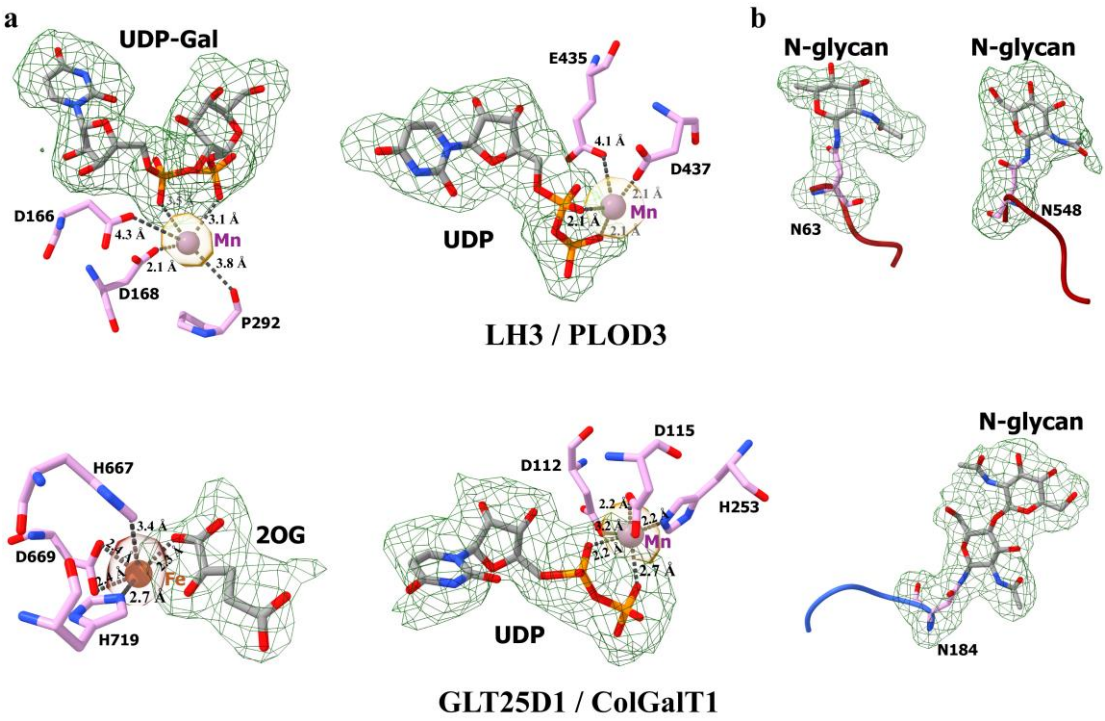

Fig. R8: The cryo-EM map density and the model-fitting for the key ligands and moieties in the KOGG complex. a, The 2-OG, UDP, and UDP-Gal molecules identified in KOGG/UDP-Gal complex were shown as stick models enclosed by the local densities as green mesh models. The metal ions identified in KOGG/UDP-Gal complex were shown as spheres enclosed by the local densities as surface models. The metal ion-coordinating residues were shown as stick models, with the distances between ions and residues indicated with dotted lines. **b,** The N-glycosylation moieties on residues N63 and N548 of LH3/PLOD3 and N184 of GLT25D1/ColGalT1 were shown as stick models enclosed by the local densities as green mesh models. All the local maps were contoured to the same level (0.162).

4. Fig 1d and 1e should be represented in a simple domain coloring scheme unless the authors want to emphasize on secondary structure contents of the domain. Fig 1e can be put into supplementary material as it does not add any valuable information to the main text or main figure.

Our reply: Thank you for the questions and suggestions regarding Figure 1. For Figure 1d, we initially used purple to highlight the β-strands to emphasize the secondary structure content of the domain. However, we acknowledge the reviewer's concern that the contrast between the purple β-strands and red helices may be difficult for readers to distinguish. To improve clarity, we have revised the color scheme in Figure 1d to provide stronger contrast between the β-strands and helices. As suggested, we have also moved Figure 1e to the Supplementary Materials and reorganized the relevant content across **Supplementary Figs. 5 and 6**.

5. Authors should discuss how this PLOD3-COLGALT1 complex could affect fibrillr collagens and basement membrane collagens?

Our reply: Thank you for the questions and suggestions regarding the biological role of KOGG complex in fibrillar collagens and basement membrane collagens biosynthesis. Fibrillar collagen, such as type I collagen, carries fewer lysine modification sites compared to basement membrane collagens. PLOD1 functions as the primary lysyl hydroxylase, and the PLOD3-COLGALT1 complex carries out glycosylation modifications after the hydroxylation by PLOD1. These hydroxylation and glycosylation modifications are essential for properly forming the triple helix structure in pro-collagen. Studies have shown that PLOD3 mutant mice exhibit disorganized collagen fibers in the skin and lungs [13], while disruption of GLT25D1/ColGalT1 results in the intracellular accumulation of type I collagen and growth arrest in osteosarcoma cells [14], confirming the crucial role of the PLOD3-COLGALT1 complex in the maturation process of fibrillar collagen. Compared with fibrillar collagens, numerous lysine modification sites were identified in basement membrane collagen, such as type IV collagen. Research by Ruotsalainen, H. et al. has demonstrated that type IV collagen accumulates intracellularly in PLOD3 knockout mice embryos, and the basement membrane structure is fragmented [13]. ColGalT1 with mutations impairing the enzymatic activity could also lead to cerebral small vessel disease, a disorder characterized by basement membrane fragility and insufficient COL4A1 production [15]. This correlation between the pathological effects of LH3/PLOD3 and GLT25D1/ColGalT1, along with their spatial arrangement in the KOGG complex, suggests that the post-translational modifications carried out by the PLOD3-COLGALT1 complex are crucial for the proper localization and functional assembly of type IV collagen in the basement membrane.
We appreciate the reviewer's suggestion to expand on the discussion of how the KOGG complex affects fibrillar and basement membrane collagens. We have revised the manuscript to incorporate this information.

Minor comments:

1. Line 35: Post-translational

Our reply: Thank you for the comments, and we have made revisions accordingly.

2. Line 60: It mentioned only COLGALT1, not COLGALT2—no mention of COLGALT2 in this study being involved in the same process. COLGALT1 and COLGALT2 act as galactosyltransferase and come under the COLGALT/GLT25D family of enzymes. Why does the author choose only COLGALT1 galactosyltransferase in association with PLOD3 and not COLGALT2 in this study?

Our reply: Thank you for the insightful question regarding GLT25D2/ColGalT2. Both GLT25D1/ColGalT1 and GLT25D2/ColGalT2, members of the GLT25D family, play roles in collagen glycosylation. However, GLT25D1/ColGalT1 is more widely expressed in many tissues and organs, while GLT25D2/ColGalT2 is primarily expressed in the nervous system, suggesting a more specialized role. Recent studies have shown that when GLT25D1/ColGalT1 is disrupted, there is a compensatory upregulation of GLT25D2/ColGalT2, which may indicate its potential involvement in collagen processing[14]. Nevertheless, due to its lower expression levels and thermostability observed in our preliminary experiments, we decided to focus our study on GLT25D1/ColGalT1 at this stage.

3. Line 132: What does XXX-XXX Å represent? No RMSD value on page 4 line number 132.

Our reply: We apologize for the missing information in our original manual. It should be 0.31-0.36 Å for the overall RMSD among KOGG in different cryo-EM structure models. We have revised the manuscript to add this information.

4. Line 188-189: “In ColGalT1, two DXD motifs have been identified as residues 166-168 in its GalT-N and residues 461-463 in its GalT-C domain” (mentioned in this study). No mention of the third DXD motif in this study.

Our reply: Thank you for pointing out the omission regarding the third DXD motif in ColGalT1. In a prior biochemical study by Perrin-Tricaud et al., three DXD motifs were proposed based on sequence analysis, specifically residues 166-168, 461-463, and 585-587 [1]. However, in their follow-up biochemical assays, they found that mutations in the third DXD motif (585-587) did not affect the glycosyltransferase activity of GLT25D1/ColGalT1, leading them to conclude that the third DXD motif is dispensable for its enzymatic function. Given that DXD motifs are generally crucial for glycosyltransferase activity, we excluded the third DXD motif from our description. To enhance clarity, we have revised the manuscript to fully incorporate the findings from this prior work.

Here is the updated description in the manuscript:

"*In ColGalT1, a previous study identified three DXD motifs based on sequence analysis, including residues 166-168 in the GalT-N domain, and residues 461-463 and 585-587 in the GalT-C domain. However, the third DXD motif (585-587) was found not to impact glycosyltransferase activity.*"

5. Line 264: “The biological significance of KOGG's linear polymerization remains unclear, and we speculate that this long fiber-like arrangement might endow KOGG with the capacity to catalyze hydroxylation/glycosylation simultaneously on multiple lysyl sites on a single polypeptide chain, such as COL1A1, COL1A2, and COL4A1.” PLOD3 is reported to hydroxylate lysine residues in COL4 but not in COL1. As per previous studies, this KOGG complex’s lysine hydroxylation and dual-glycosylation features should be present in COL4 and only dual-glycosylation features in COL1.

Our reply: Thank you for the questions and suggestions the collagen-specificity of PLOD3. It is a good point that the hydroxylation-glycosylation coordinating activities suggested by the KOGG polymerization might only apply to the COL4A1 for the substrate-specificity of PLOD3. In light of reviewer’s comments, we have revised the manuscript to make correction accordingly.

6. In both, Figure 3.d and Figure 4.c glucosyltransferase activity is less than galactosyltransferase activity. In Figure 3.d, it is mentioned that the mutations are at the catalytic sites of GalT-N and GalT-C of KOGG complex. GalT-N and GalT-C are present in COLGALT1 and are associated with galactosyltransferase activity. But why in the case of PLOD3 the galactosyltransferase activity is more than that of glucosyltransferase activity?

Our reply: Thank you for the questions and suggestions regarding the glucosyltransferase (GlcT) and galactosyltransferase (GalT) activities. In our biochemical assays, all the mutant data were normalized against their corresponding wild-type counterparts for each enzyme activity (GlcT, GalT, or LH) before plotting. This means that the galactosyltransferase and glucosyltransferase data were obtained from two separate groups of experiments and,

therefore, are not directly comparable. Although we observed a higher readout for glucosyltransferase activity in some experiments, we did not compare the two activities directly, as they were not part of parallel experiments.

7. Page 6, line 197, Supplementary Figure XXX (figure number not denoted).

Our reply: Thank you for pointing out the missed information. It should be **Supplementary Figure 9a**, and we have updated the information accordingly in our revised manuscript.

References

1 Perrin-Tricaud, C., Rutschmann, C. & Hennet, T. Identification of domains and amino acids essential to the collagen galactosyltransferase activity of GLT25D1. *PLoS One* **6**, e29390 (2011). https://doi.org/10.1371/journal.pone.0029390

2 Lairson, L. L., Henrissat, B., Davies, G. J. & Withers, S. G. Glycosyltransferases: structures, functions, and mechanisms. *Annu Rev Biochem* **77**, 521-555 (2008). https://doi.org/10.1146/annurev.biochem.76.061005.092322

3 Ardevol, A., Iglesias-Fernandez, J., Rojas-Cervellera, V. & Rovira, C. The reaction mechanism of retaining glycosyltransferases. *Biochem Soc Trans* **44**, 51-60 (2016). https://doi.org/10.1042/BST20150177

4 Scietti, L. *et al.* Molecular architecture of the multifunctional collagen lysyl hydroxylase and glycosyltransferase LH3 (vol 9, 3163, 2018). *Nature Communications* **9** (2018). https://doi.org/10.1038/s41467-018-06481-x

5 Mattoteia, D. *et al.* Identification of Regulatory Molecular "Hot Spots" for LH/PLOD Collagen Glycosyltransferase Activity. *International journal of molecular sciences* **24** (2023). https://doi.org/10.3390/ijms241311213

6 Terajima, M. *et al.* Cyclophilin B control of lysine post-translational modifications of skin type I collagen. *PLoS Genet* **15**, e1008196 (2019). https://doi.org/10.1371/journal.pgen.1008196

7 Sricholpech, M. *et al.* Lysyl hydroxylase 3-mediated glucosylation in type I collagen: molecular loci and biological significance. *J Biol Chem* **287**, 22998-23009 (2012). https://doi.org/10.1074/jbc.M112.343954

8 Mattoteia, D. *et al.* Identification of Regulatory Molecular "Hot Spots" for LH/PLOD Collagen Glycosyltransferase Activity. *Int J Mol Sci* **24** (2023). https://doi.org/10.3390/ijms241311213

9 Guo, H. F. *et al.* A collagen glucosyltransferase drives lung adenocarcinoma progression in mice. *Commun Biol* **4**, 482 (2021). https://doi.org/10.1038/s42003-021-01982-w

10 Koenig, S. N. *et al.* New mechanistic insights to PLOD1-mediated human vascular disease. *Transl Res* **239**, 1-17 (2022). https://doi.org/10.1016/j.trsl.2021.08.002

11 Heikkinen, J. *et al.* Lysyl hydroxylase 3 is a multifunctional protein possessing collagen glucosyltransferase activity. *J Biol Chem* **275**, 36158-36163 (2000). https://doi.org/10.1074/jbc.M006203200

12 Rautavuoma, K. *et al.* Premature aggregation of type IV collagen and early lethality in lysyl hydroxylase 3 null mice. *Proceedings of the National Academy of Sciences of the United States of America* **101**, 14120-14125 (2004). https://doi.org/10.1073/pnas.0404966101

13 Ruotsalainen, H. *et al.* Glycosylation catalyzed by lysyl hydroxylase 3 is essential for basement membranes. *J Cell Sci* **119**, 625-635 (2006). https://doi.org/10.1242/jcs.02780

14 Baumann, S. & Hennet, T. Collagen Accumulation in Osteosarcoma Cells lacking GLT25D1 Collagen Galactosyltransferase. *J Biol Chem* **291**, 18514-18524 (2016). https://doi.org/10.1074/jbc.M116.723379

15 Miyatake, S. *et al.* Biallelic COLGALT1 variants are associated with cerebral small vessel disease. *Ann Neurol* **84**, 843-853 (2018). https://doi.org/10.1002/ana.25367