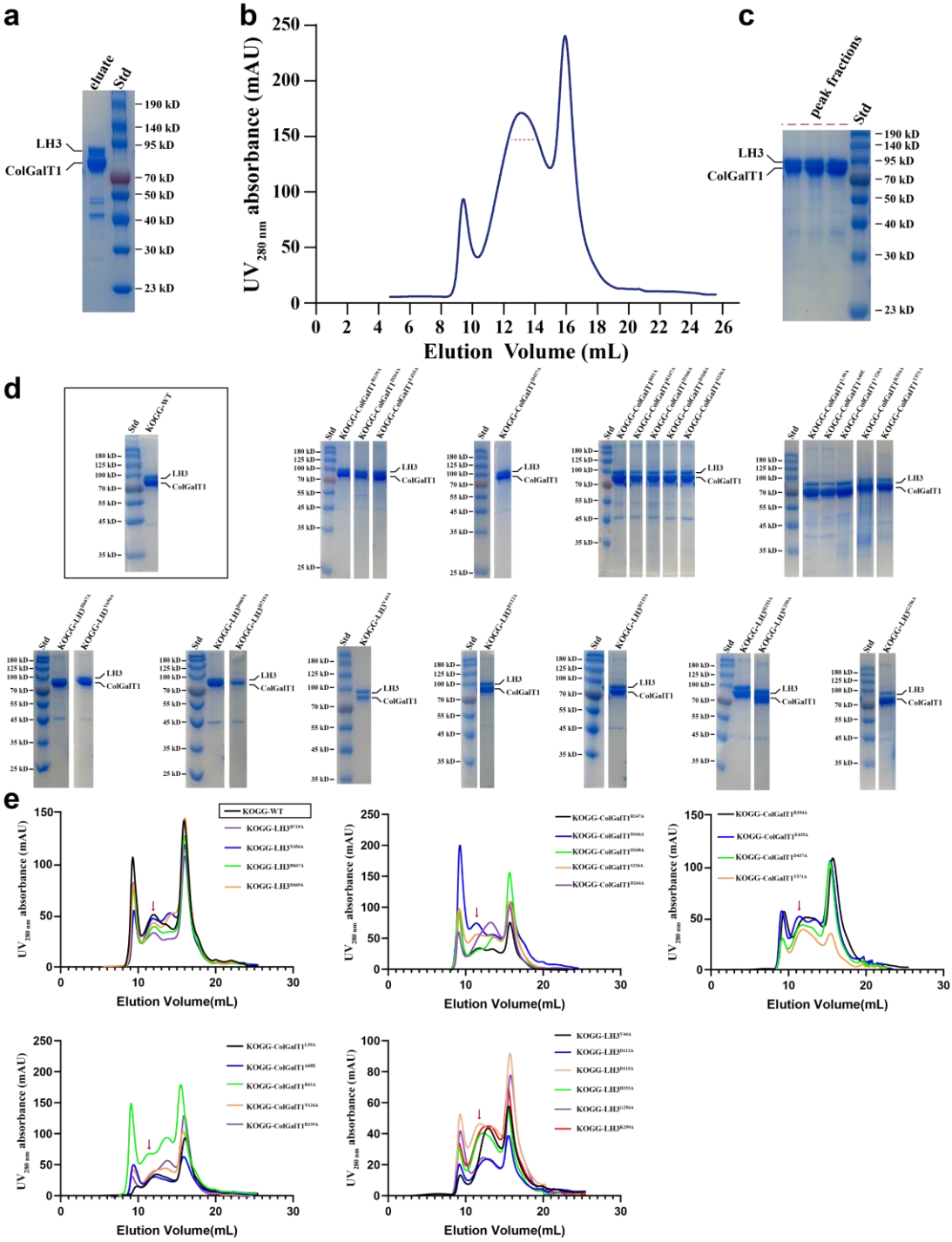
Supplementary information for

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

Junjiang Peng[1,2†], Wenguo Li[1,2†], Deqiang Yao[3†], Ying Xia[2], Qian Wang[2], Yan Cai[2], Shaobai Li[2], Mi Cao[2], Yafeng Shen[2], Peixiang Ma[1,4], Rijing Liao[2], Jie Zhao[1,4], An Qin[1,4*], Yu Cao[1,2*]

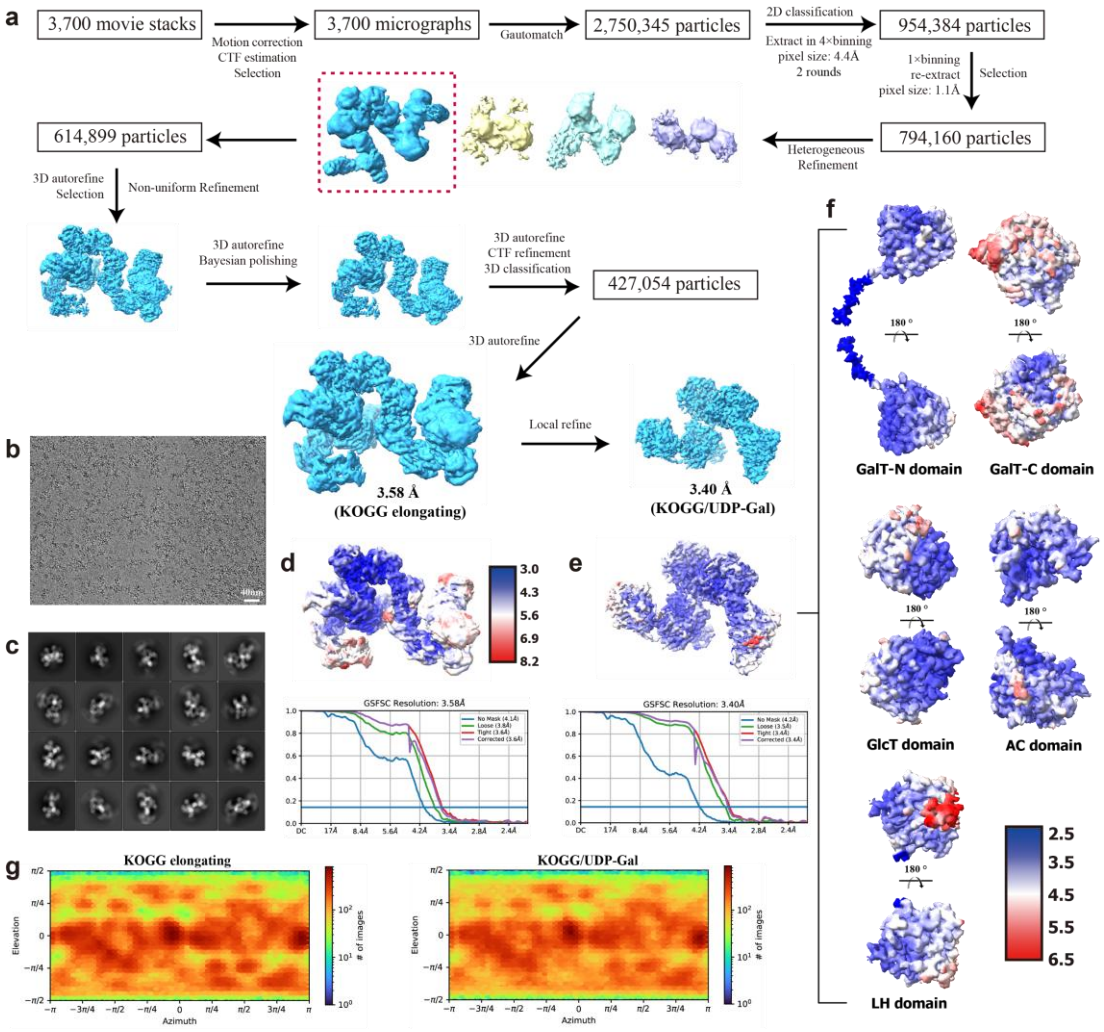
Corresponding authors: Email: yu.cao@shsmu.edu.cn (Y. C.); dr_qinan@163.com (A. Q.)

Supplementary items:
Supplementary Figs. 1-13
Supplementary Tables 1-3



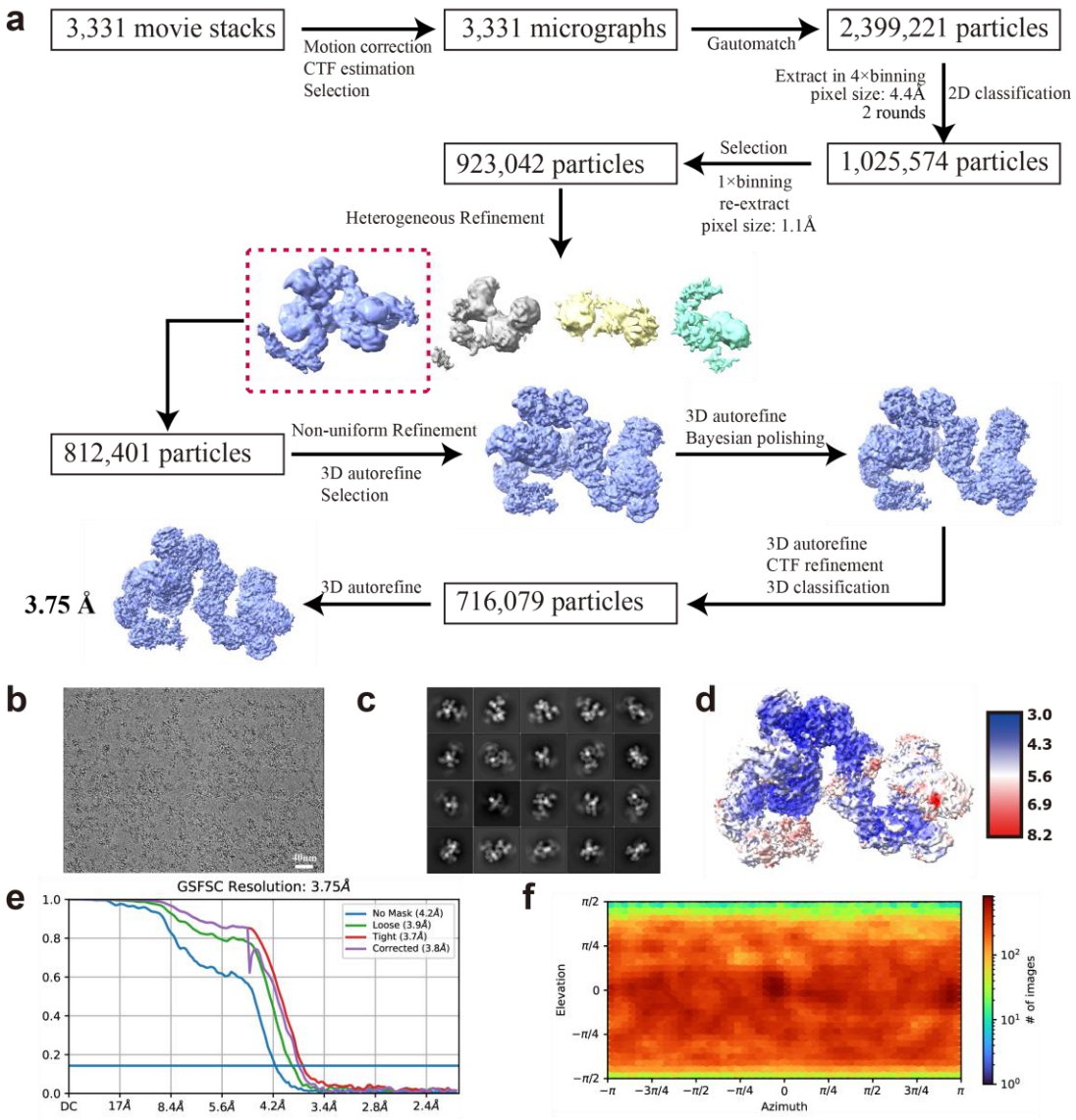
Supplementary Fig. 1. The purification of the LH3 and ColGalT1 complex for cryo-EM and enzymatic assays.

a, The purification of LH3 and ColGalT1 protein complex by affinity chromatography. **b&c,** The SEC FPLC purification of LH3 and ColGalT1 protein complex (b) and the SDS-PAGE analysis of the peak fractions (c). **d,** SDS-PAGE analysis of the eluates from the Strep-tag affinity chromatography of the KOGG complex with the mutations indicated above the gels. **e,** SEC FPLC profiles of the KOGG complexes with functional mutations. The elution fractions corresponding to the heterotetrameric LH3-ColGalT1 complex were indicated with red arrows.



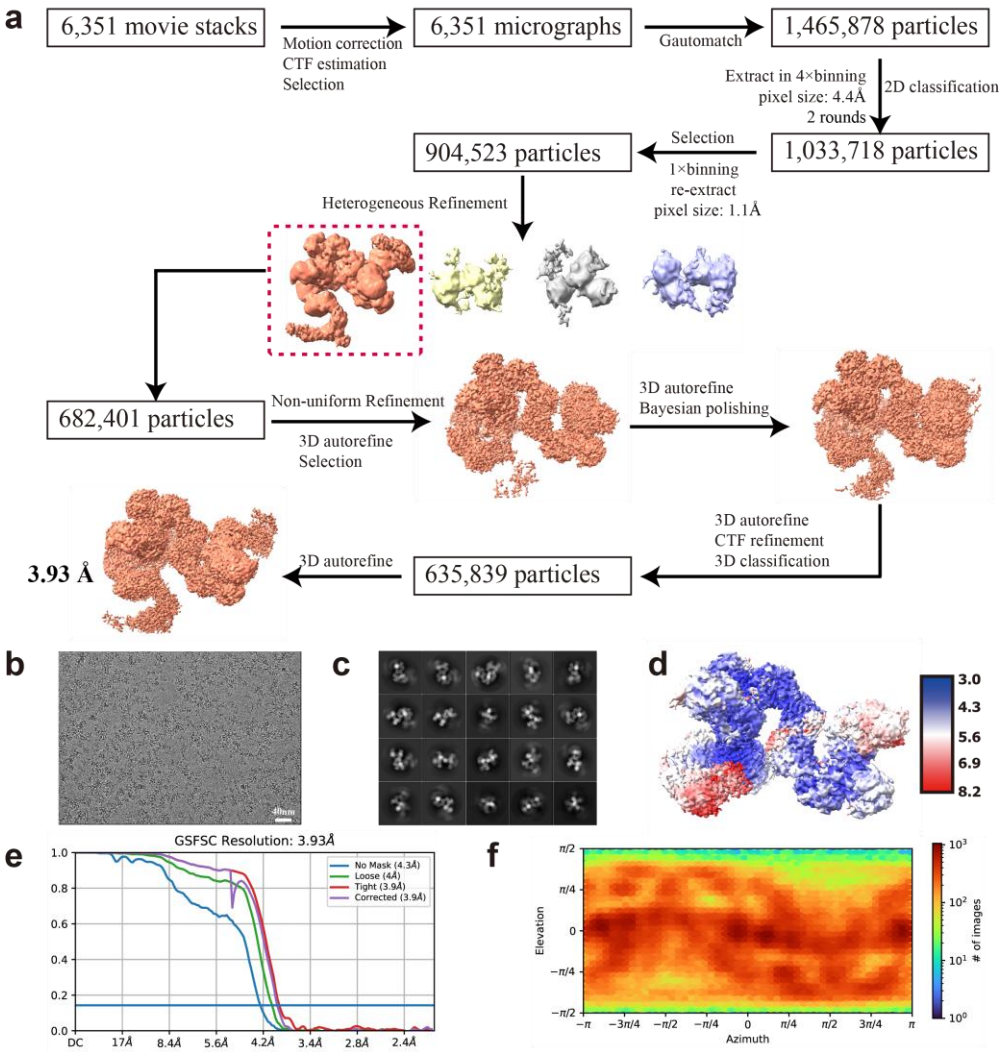
Supplementary Fig. 2. The cryo-EM analysis of KOGG/UDP-Gal.

a, The flow chart of cryo-EM data processing on LH3-ColGalT1 complex supplemented with UDP-galactose, resulting in two complex structures: KOGG elongating upon global refinement and KOGG/UDP-Gal after local refinement. **b,** A representative micrograph of LH3-ColGalT1 complex supplemented with UDP-galactose. Most of the micrographs were similarly of high quality. **c,** A representative 2D class average. **d,** The resolution map for KOGG elongating (upper) and the gold-standard Fourier shell correlation (FSC) curves for the final cryo-EM map, generated by cryoSPARC with non-uniform refinement (lower). **e,** The resolution map for KOGG/UDP-Gal generated by cryoSPARC with non-uniform refinement (upper) and the gold-standard Fourier shell correlation (FSC) curves for the final cryo-EM map, generated by cryoSPARC with non-uniform refinement (lower). **f,** Local resolution maps for the individual domains of KOGG/UDP-Gal. **g,** Orientation distributions of particles for the final map reconstruction of KOGG elongating (left) and KOGG/UDP-Gal (right).



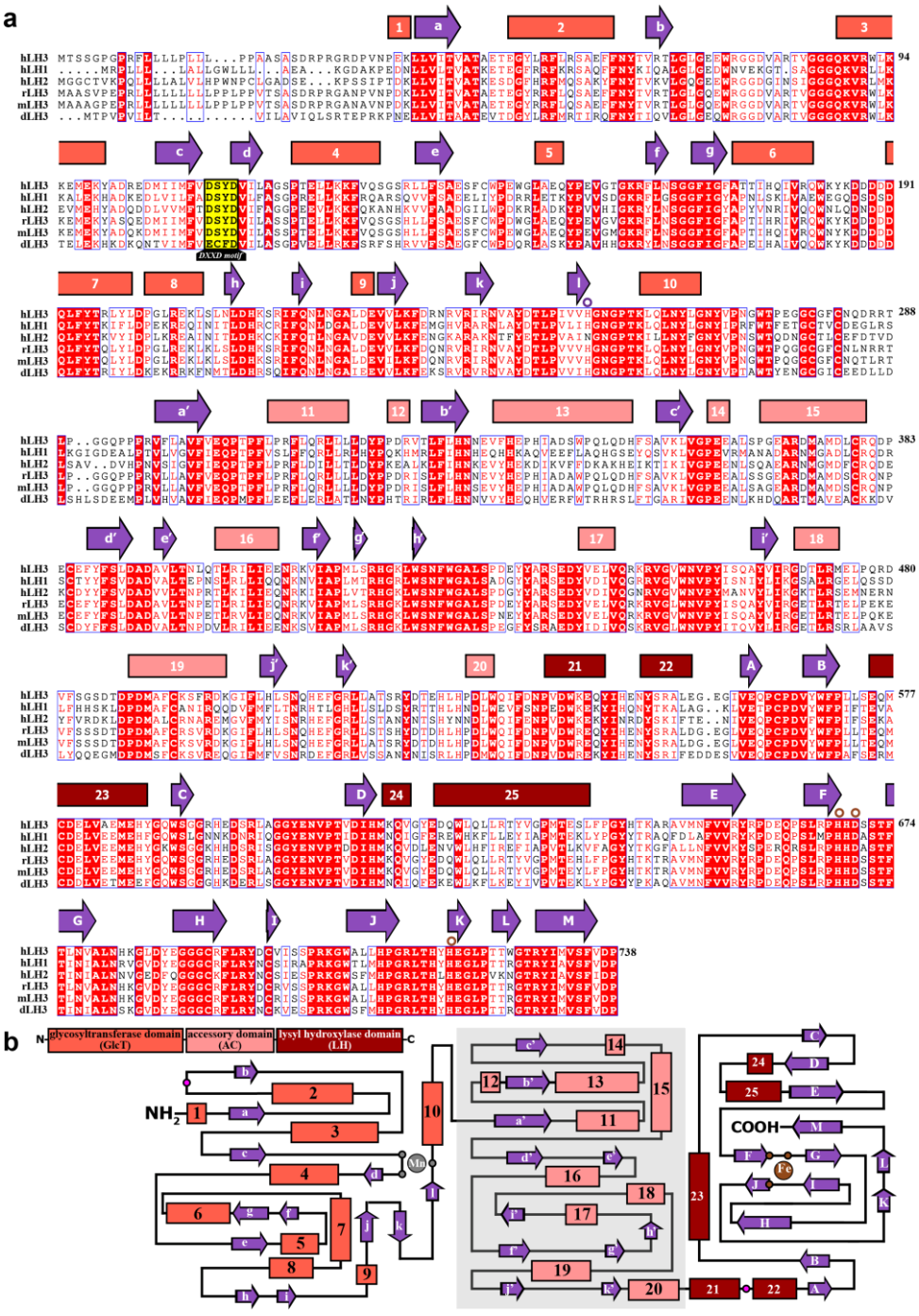
Supplementary Fig. 3. The cryo-EM analysis of KOGG/UDP-Glc.

a, The flow chart of cryo-EM data processing on LH3-ColGalT1 complex supplemented with UDP-glucose. **b,** A representative micrograph of LH3-ColGalT1 complex supplemented with UDP-glucose. Most of the micrographs were similarly of high quality. **c,** A representative 2D class average. **d,** Local-resolution map of KOGG/UDP-Glc. **e,** The gold-standard Fourier shell correlation (FSC) curves for the final cryo-EM map of KOGG/UDP-Glc, generated by cryoSPARC with non-uniform refinement. **f,** Orientation distributions of particles for the final map reconstruction of KOGG/UDP-Glc.



Supplementary Fig. 4. The cryo-EM analysis of KOGG apo.

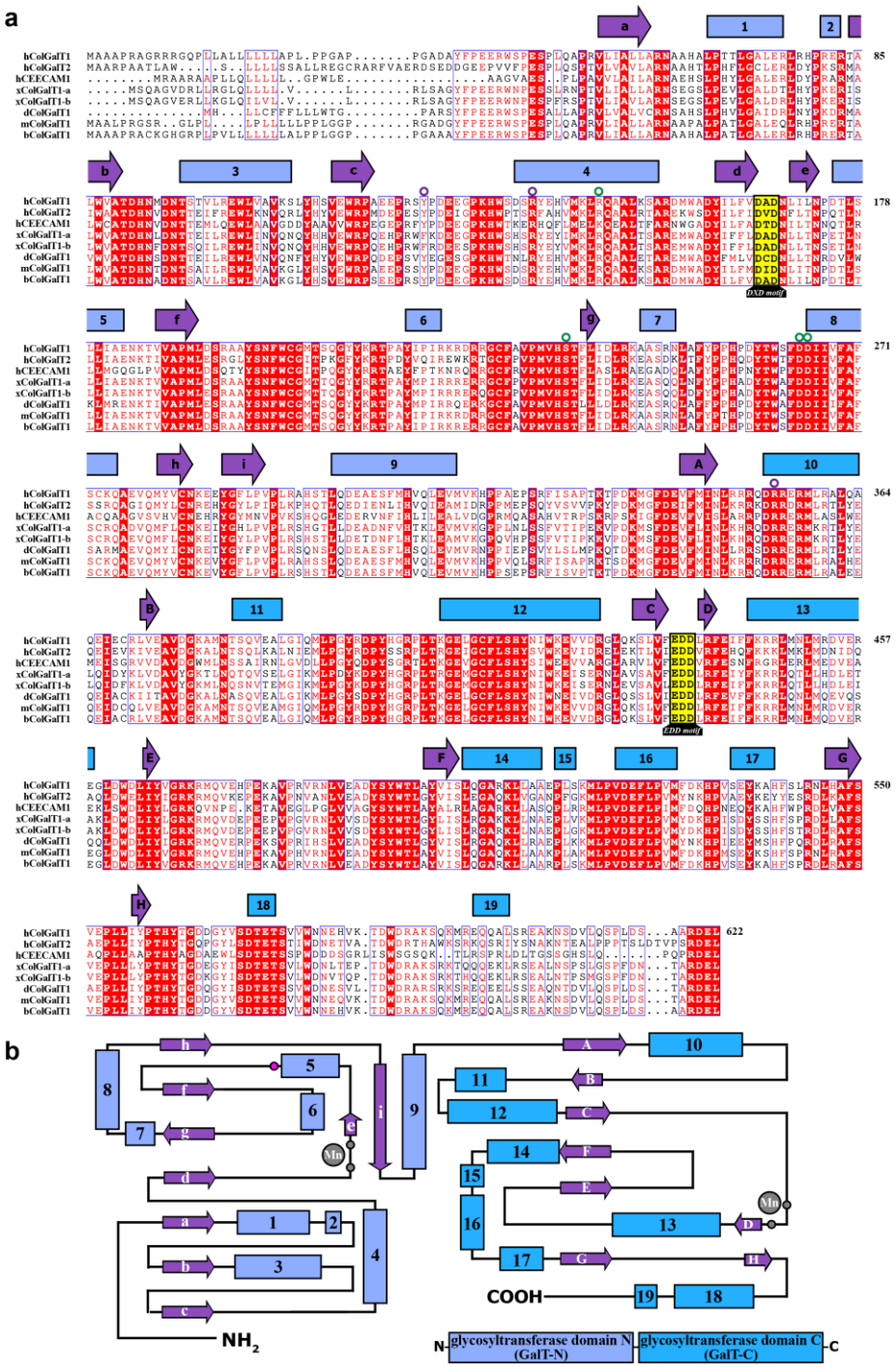
a, The flow chart of cryo-EM data processing on LH3-ColGalT1 complex non-supplemented with substrates. **b,** A representative micrograph of LH3-ColGalT1 complex non-supplemented. Most of the micrographs were similarly of high quality. **c,** A representative 2D class average. **d,** Local-resolution map of KOGG apo. **e,** The gold-standard Fourier shell correlation (FSC) curves for the final cryo-EM map of KOGG apo, generated by cryoSPARC with non-uniform refinement. **f,** Orientation distributions of particles for the final map reconstruction of KOGG apo.



Supplementary Fig. 5. The amino acid sequence analysis on LH3/PLOD3.

a, The protein sequences of LH1-3 from *Homo sapiens* (hLH1, UniProtKB-Q02809, [https://www.uniprot.org/uniprotkb/Q02809]; hLH2, UniProtKB-O00469, [https://www.uniprot.org/uniprotkb/O00469]; hLH3, UniProtKB-O60568, [https://www.uniprot.org/uniprotkb/O60568]), LH3 from *Danio rerio* (dLH3, UniProtKB-A3KNH4, [https://www.uniprot.org/uniprotkb/A3KNH4]), LH3 from *Rattus norvegicus* (rLH3, UniProtKB-Q5U367, [https://www.uniprot.org/uniprotkb/Q5U367]), LH3 from *Mus musculus* (mLH3, UniProtKB-Q9R0E1, [https://www.uniprot.org/uniprotkb/Q9R0E1]) were aligned with the secondary structural elements of hLH3 marked above the alignment. Residues are colored based on their

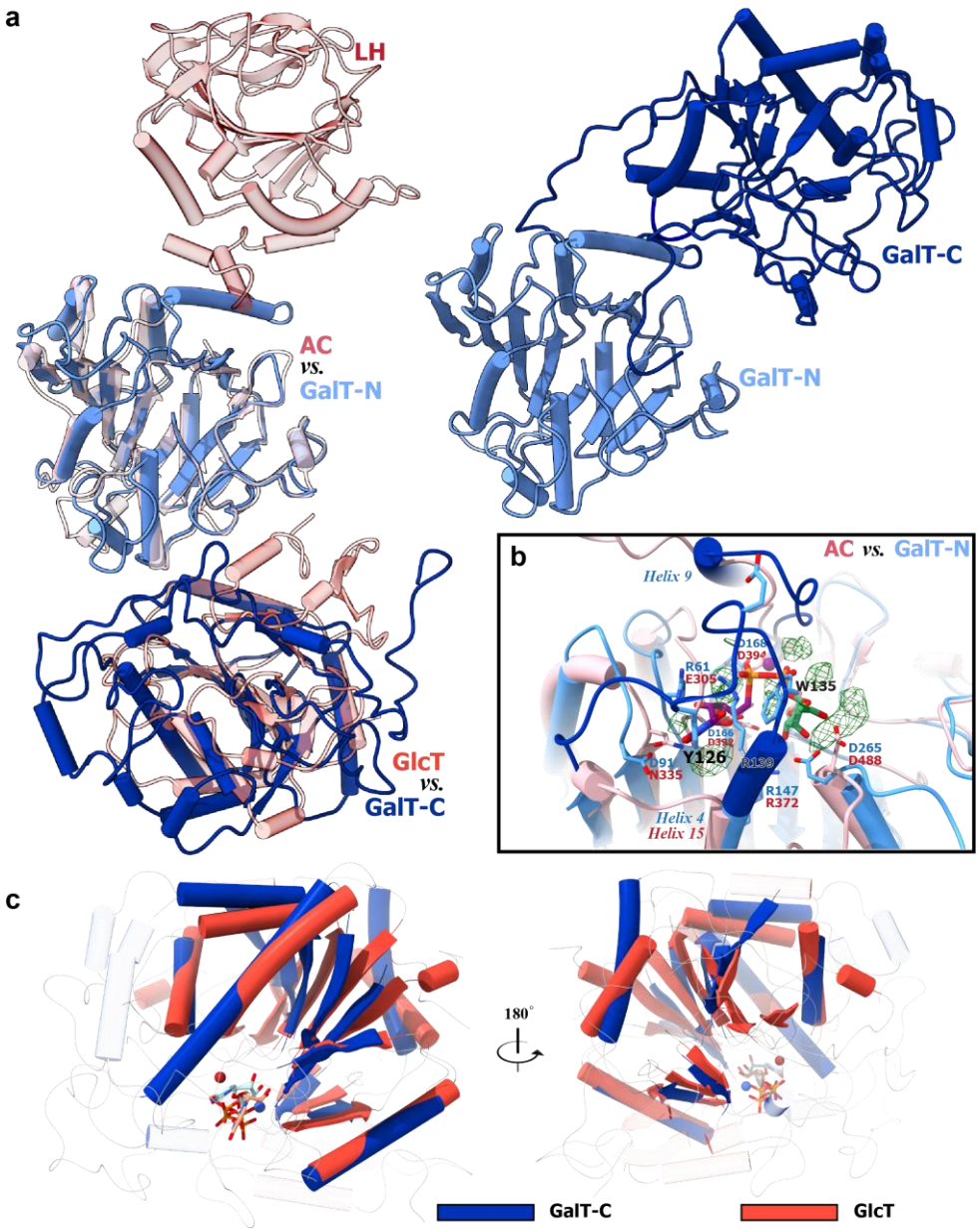
conservation using the ESPript server. **b,** Schematic diagram illustrating the organization of secondary structures in LH3/PLOD3.



Supplementary Fig. 6. The amino acid sequence analysis on GLT25D1/ColGalT1.

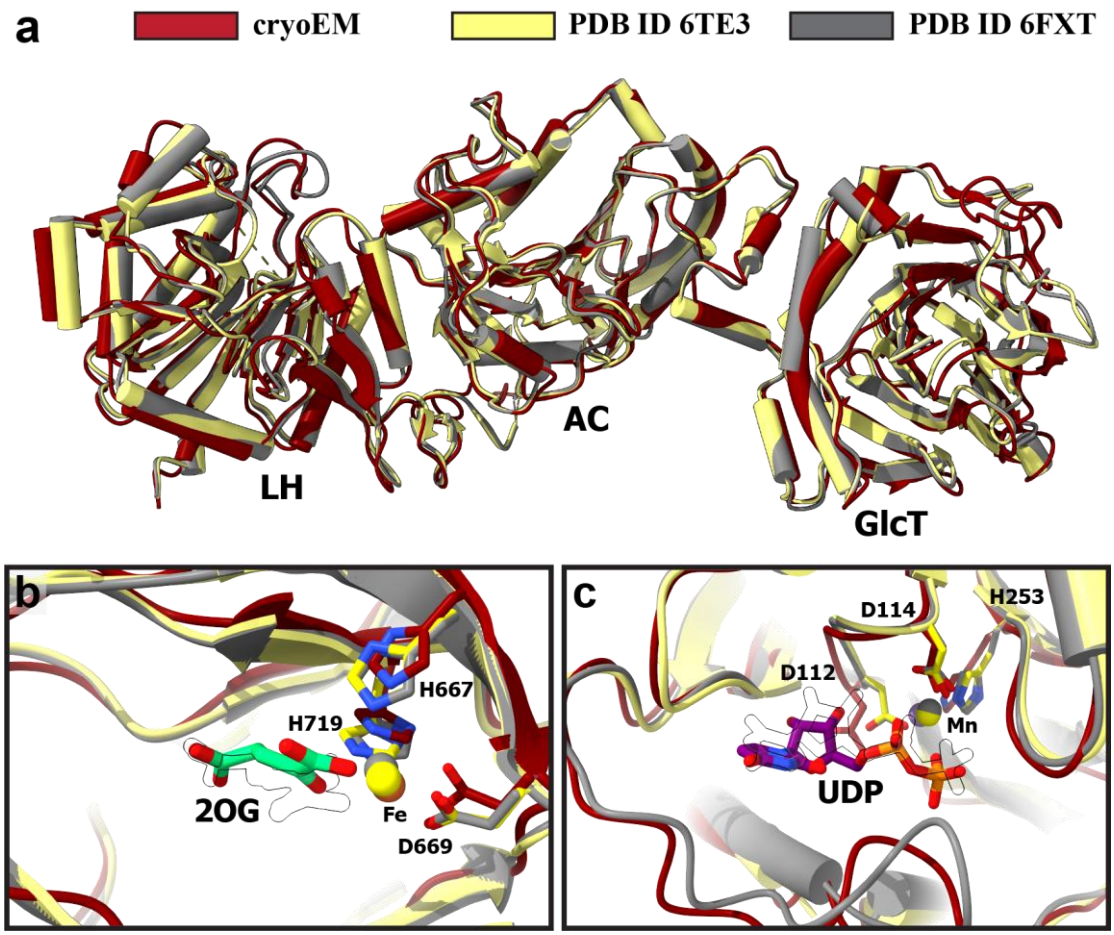
a, The protein sequences of Procollagen galactosyltransferase 1 from *Homo sapiens* (hColGalT1, UniProtKB-Q8NBJ5, [https://www.uniprot.org/uniprotkb/Q8NBJ5]), Procollagen galactosyltransferase 2 from *Homo sapiens* (hColGalT2, UniProtKB-Q8IYK4, [https://www.uniprot.org/uniprotkb/Q8IYK4]), Inactive glycosyltransferase 25 family member 3 from *Homo sapiens* (hCEECAM1, UniProtKB-Q5T4B2, [https://www.uniprot.org/uniprotkb/Q5T4B2]), Procollagen galactosyltransferase 1-a from *Xenopus laevis* (xColGalT1-a, UniProtKB-A0JPH3, [https://www.uniprot.org/uniprotkb/A0JPH3]), Procollagen galactosyltransferase 1-b from *Xenopus laevis* (xColGalT1-b, UniProtKB-Q5U483,

[https://www.uniprot.org/uniprotkb/Q5U483]), Procollagen galactosyltransferase 1 from *Danio rerio* (dColGalT1, UniProtKB-A5PMF6, [https://www.uniprot.org/uniprotkb/A5PMF6]), Procollagen galactosyltransferase 1 from *Mus musculus* (mColGalT1, UniProtKB-Q8K297, [https://www.uniprot.org/uniprotkb/Q8K297]), Procollagen galactosyltransferase 1 from *Bos taurus* (bColGalT1, UniProtKB-A5PK45, [https://www.uniprot.org/uniprotkb/A5PK45]) were aligned with the secondary structural elements of hColGalT1 marked above the alignment. Residues are colored based on their conservation using the ESPript server. **b,** Schematic diagram illustrating the organization of secondary structures in GLT25D1/ColGalT1.



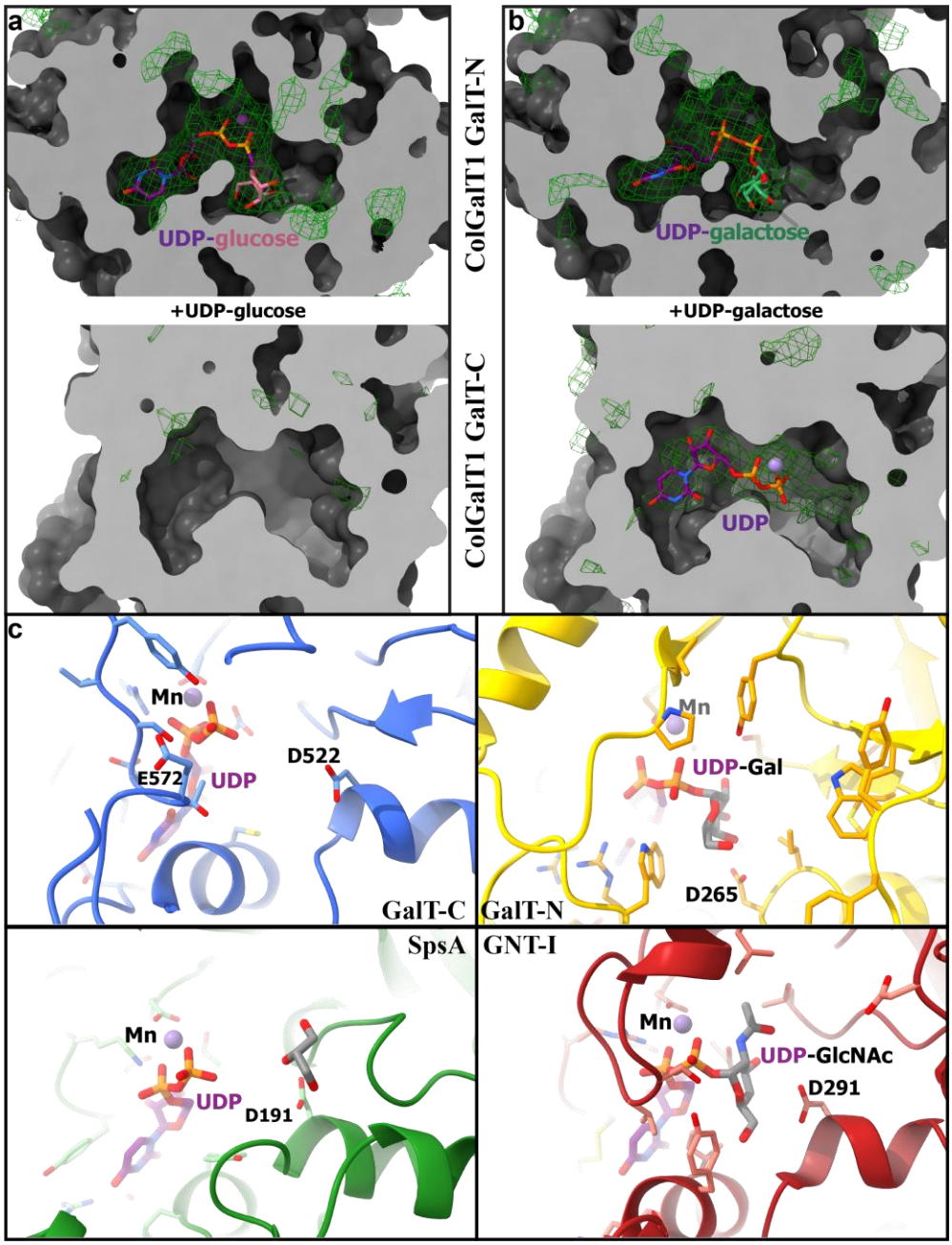
Supplementary Fig. 7. The structural comparison among the functional domains in the KOGG complex.

a, Left: The structural superpositions among the GlcT and AC domains from LH3 and GalT-N and GalT-C domains from ColGalT1. The LH3 structure from KOGG/UDP-Gal was shown as a cartoon model, and the GalT-N and GalT-C domains from ColGalT1 were superposed individually to LH3. Right: The ColGalT1 structure was shown as a reference with the same orientation as the GalT-N domain in the superposition. **b,** Enlarged view of the inner cavity in the AC vs GalT-N domain superposition. The AC domain was shown as a pink cartoon model, and the GalT-N domain was shown as a light blue cartoon model, except for the local structure in dark blue, which had a critically different conformation from the AC domain. **c,** 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.



Supplementary Fig. 8. The structural comparison of LH3/PLOD3 cryo-EM structure with crystal structures

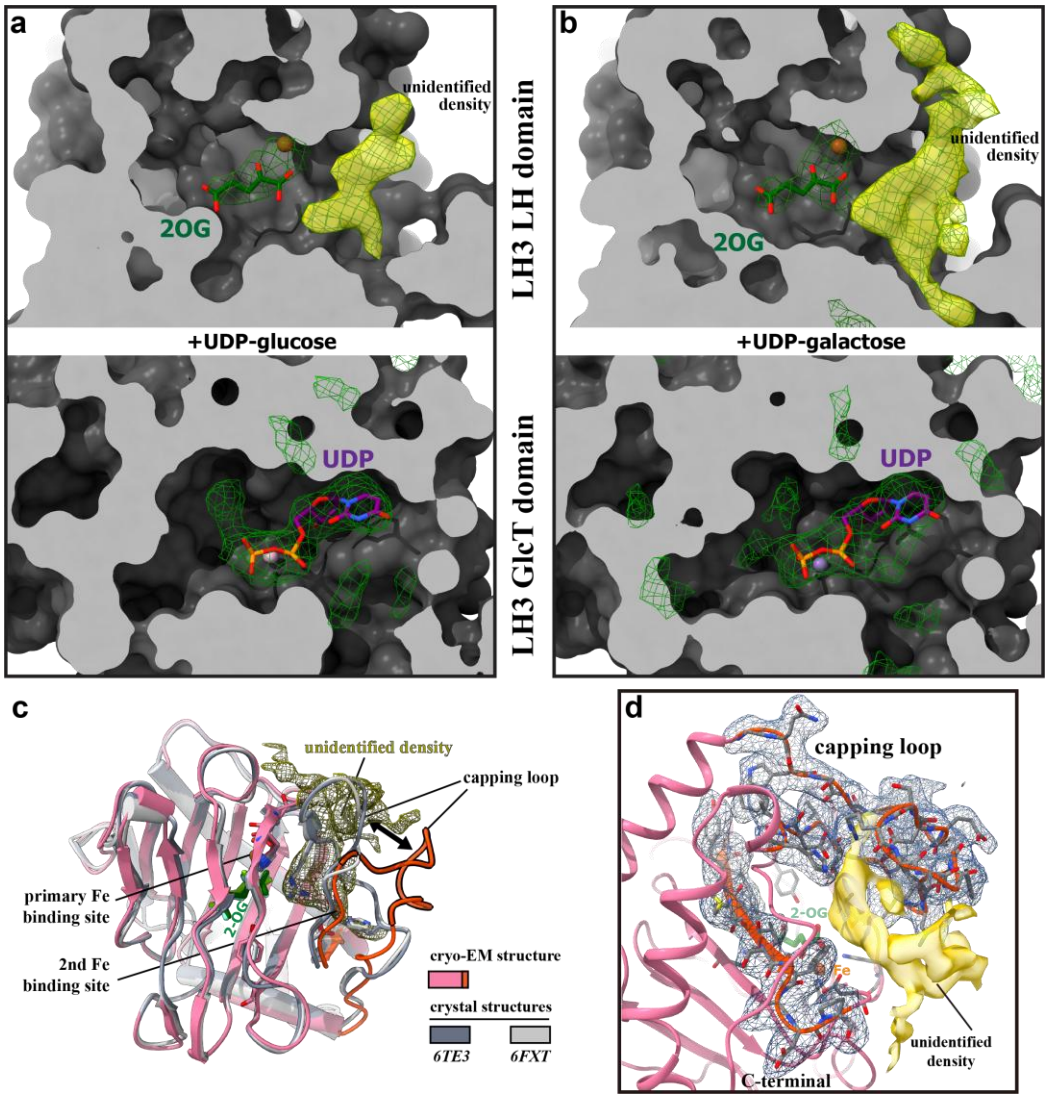
a, The structural superpositions among LH3^{B} structure from KOGG/UDP-Glc complex and crystal structures. The LH3^{B} structure from KOGG/UDP-Glc and two crystal structures generated using coordinate files with PDB IDs 6TE3 and 6FXT were shown as cartoon models. **b,** Enlarged view of the active site in LH domain. The 2OG molecules in three structures were shown as stick model, with the 2OG in LH3^{B} structure colored in green and red and other 2OG molecules set to transparent. **c,** Enlarged view of the active site in GlcT domain. The UDP molecules in three structures were shown as stick model, with the UDP in LH3^{B} structure colored in purple-red-orange and other UDP molecules set to transparent.



Supplementary Fig. 9. The cryo-EM density identification and modeling in ColGalT1.

a, The non-proteinous densities in the inner cavity of ColGalT1 from the KOGG/UDP-Glc complex. The GalT-N (upper) and GalT-C (lower) domains were shown as surface models and sliced to show the inner cavities for substrate binding. The non-proteinous densities in the cryo-EM map were shown as green mesh. **b,** The non-proteinous densities in the inner cavity of ColGalT1 from the KOGG/UDP-Gal complex. The GalT-N (upper) and GalT-C (lower) domains were shown as surface models and sliced to show the inner cavities for substrate binding. The non-proteinous densities in the cryo-EM map were shown as green mesh. **c,** 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

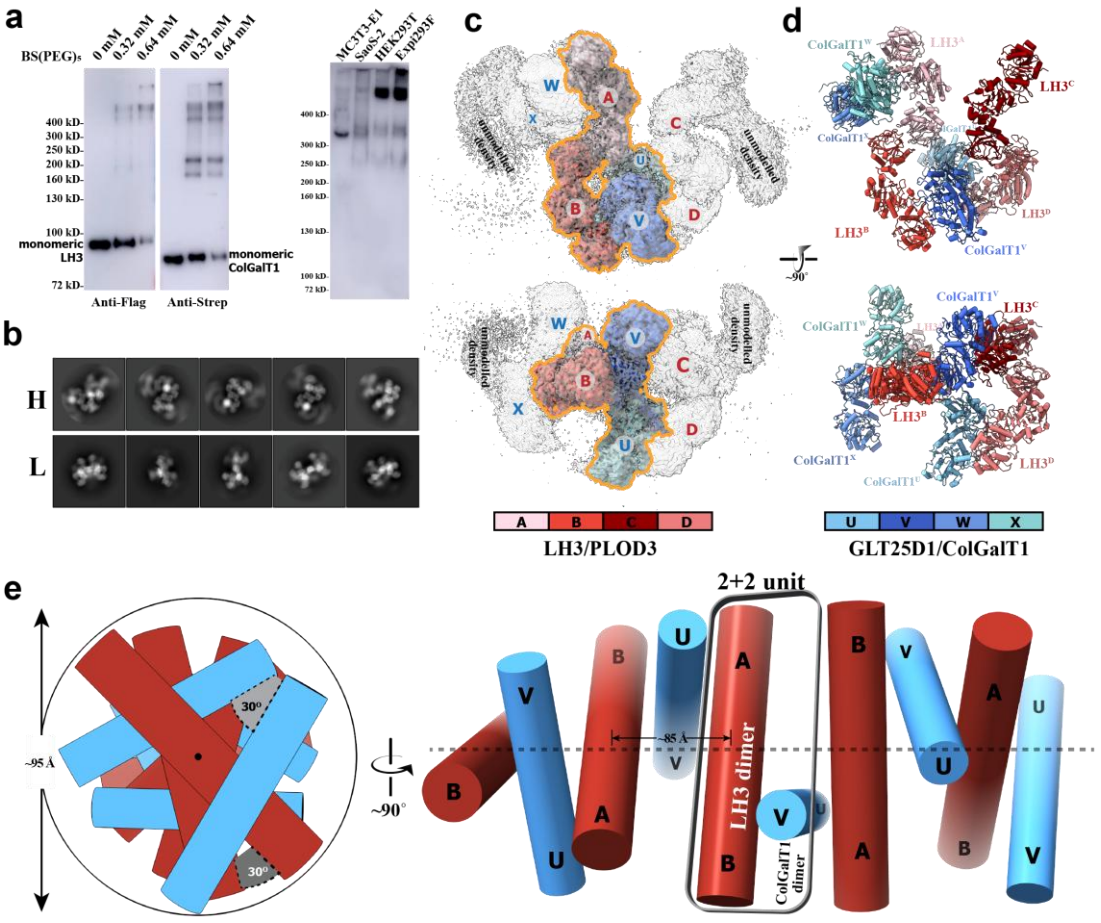
PDB IDs 1QGQ and 1FOA, respectively.



Supplementary Fig. 10. The cryo-EM density identification and modeling in LH3/PLOD3.

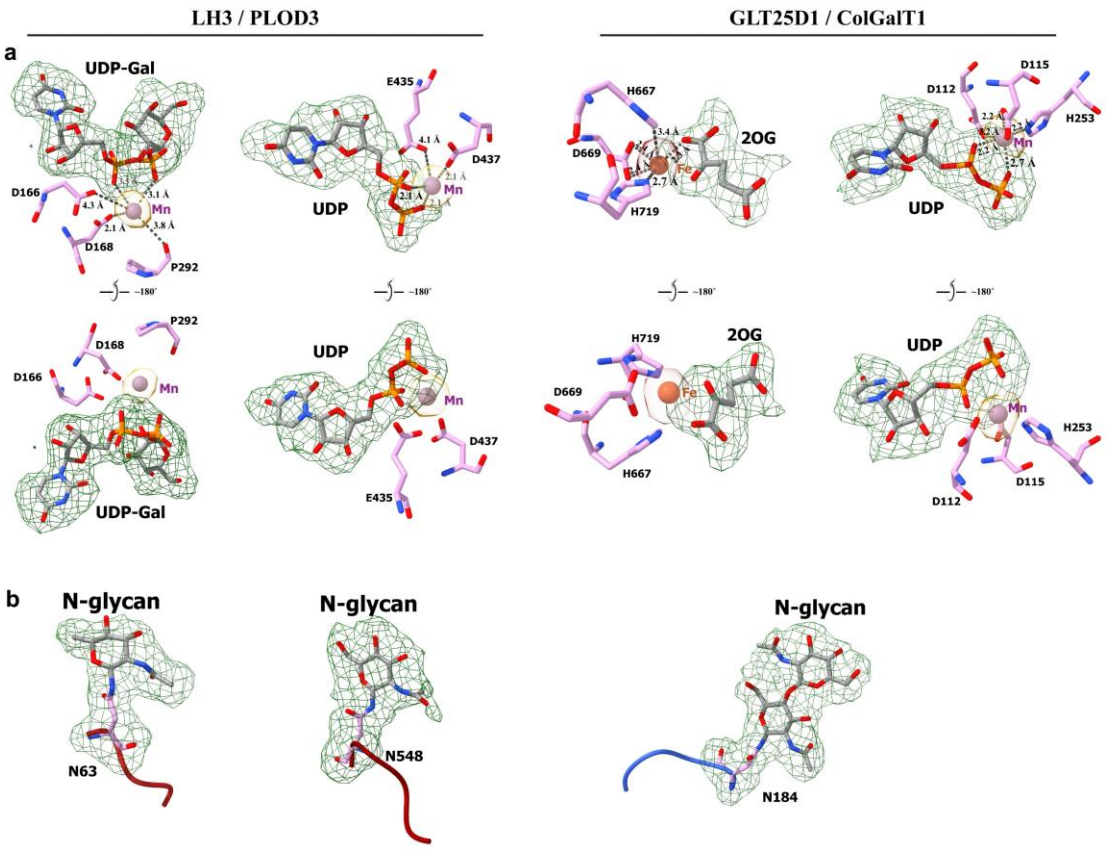
a, The non-proteinous densities in the inner cavity of LH3 from the KOGG/UDP-Glc complex. The LH (upper) and GlcT (lower) domains were shown as surface models and sliced to show the inner cavities for substrate binding. The non-proteinous densities in the cryo-EM map were shown as green mesh. The large continuous density identified near the 2OG was highlighted in the yellow surface model. **b,** The non-proteinous densities in the inner cavity of LH3 from the KOGG/UDP-Gal complex. The LH (upper) and GlcT (lower) domains were shown as surface models and sliced to show the inner cavities for substrate binding. The non-proteinous densities in the cryo-EM map were shown as green mesh. The large continuous density identified near the 2OG was highlighted in the yellow surface model. **c,** The local cryo-EM map density near the lysyl hydroxylation site of LH3. 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. **d,** The structural superposition among the LH domains from KOGG/UDP-Gal structure (pink) and crystal structures of LH3, generated using the coordinate files with PDB IDs 6TE3 (gray-blue) and 6FXT (gray). The unidentified density was shown as a yellow

mesh.



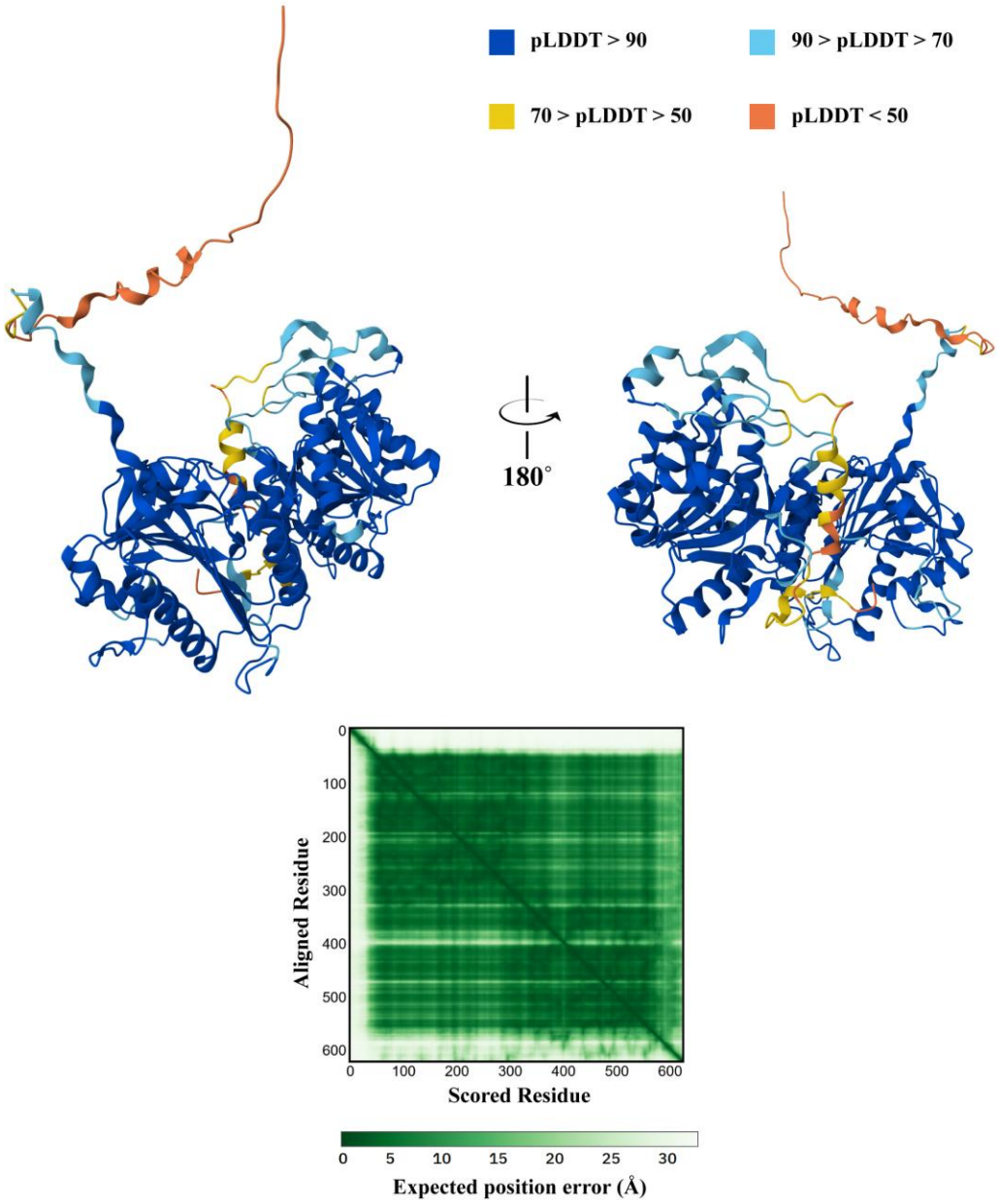
Supplementary Fig. 11. Fiber-like polymerizing pattern of the KOGG complex.

a, The high oligomeric states of the KOGG complex. **Left,** The purified KOGG complex protein sample was incubated with crosslinking agent BS(PEG)$_5$ and subjected to SDS-PAGE. **Right,** The cell lysates from MC3T3-E1, SaoS-2, HEK293T, and Expi293F cell cultures were subjected to native-gel electrophoresis and visualized with anti-ColGalT1 antibody. **b,** The 2D classes representing the KOGG complex in the high oligomeric (H) and low oligomeric states (L). **c,** Cryo-EM map of KOGG/UDP-Gal without local refinement. The cryo-EM map was set to transparent, and the density corresponding to the KOGG quaternary complex was shown within the map. The map portions for the additional LH3 (C and D) and ColGalT1 (W and X) were marked. **d,** Heterooctameric KOGG elongated complex. The four LH3 protomers (A-D) and ColGalT1 protomers (U-X) were shown as cartoon models and colored to the scheme below. **e,** Schematic diagram illustrating the fiber-like organization of KOGG polymers. The polymerization pattern of KOGG polymer was simulated by repeatedly superposing the protomers C-D of KOGG elongated complex with the protomers A-B from an additional KOGG elongated complex. The LH3 dimer and ColGalT1 dimer in the KOGG polymer were shown as red and blue cylinders, respectively. The KOGG polymer was exemplified as five 2+2 units assembled with the U-A interunit contacting and viewed along the elongation axis (left) and perpendicularly to the elongation axis (right).



Supplementary Fig. 12. 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).



Supplementary Fig. 13. The AlphaFold2-predicted structural model of human ColGalT1.

The predicted model was shown with the pLDDT score overlaid from two viewing angles (upper panel), with the predicted aligned error (PAE) shown below the model.

Supplementary Table 1. Cryo-EM data information

Data	Particles for final map reconstitution	Unique protein component 1	Unique protein component 2	Other components	Resolution (Å)
KOGG apo complex	635,839	LH3/PLOD3 × 4	GLT25D1/ColGalT1 × 4	none	3.93
KOGG/UDP-Glc complex	716,079	LH3/PLOD3 × 4	GLT25D1/ColGalT1 × 4	UDP-glucose	3.75
KOGG elongating complex	427,054	LH3/PLOD3 × 4	GLT25D1/ColGalT1 × 4	UDP-galactose	3.58
KOGG/UDP-Gal complex	427,054	LH3/PLOD3 × 2	GLT25D1/ColGalT1 × 2	UDP-galactose	3.40

Supplementary Table 2. Cryo-EM data collection, model refinement, and validation statistics.

	KOGG apo complex (EMDB-60079) (PDB 8ZGH)	KOGG/UDP-Glc complex (EMDB-60078) (PDB 8ZGG)	KOGG elongating complex (EMDB-60075) (PDB 8ZGC)	KOGG/UDP-Gal complex (EMDB-60076) (PDB 8ZGE)
Data collection & processing				
Magnification	81000			
Voltage (kV)	300			
Electron exposure ($e^-/Å^2$)	50			
Defocus range (μm)	1.5-2.6			
Pixel size (Å/pixel)	1.1			
Symmetry imposed	C1			
Number of movies	6351	3331	3700	3700
Final particle images (no.)	635,839	716,079	427,054	427,054
Map resolution (Å)	3.93	3.75	3.58	3.40
FSC threshold	0.143			
Refinement				
Initial model used (PDB code)	AF-Q8NBJ5-F1 6FXK			
Model resolution (Å)	4.1	3.9	3.7	3.5
FSC threshold	0.143	0.143	0.143	0.143
Model composition				
Non-hydrogen atoms	34096	34148	34214	21428
Protein residues	5172	5172	5172	2586
Ligands	GDU:2; FE2:2; NAG:7	AKG:2;AKG:2; 660:2; MN:2;MN:4; NAG:7	FE2:2;AKG:2; FE2:2; MN:6; NAG:8	GDU:2;AKG:2; GDU:2; FE2:2; MN:6; NAG:8
B factors ($Å^2$)				
Protein	69.18	64.77	61.88	32.24
Ligand	50.89	57.27	49.55	39.40
R.m.s. deviations				
Bond length (Å)	0.004	0.005	0.005	0.006
Bond angles (°)	0.703	1.101	1.116	1.092
Validation				
MolProbity score	1.94	1.89	1.86	1.92
Clashscore	9.34	9.47	7.63	8.16
Rotamer outliers (%)	0.35	0.62	0.66	0.75
Ramachandran plot				
Outlier (%)	0.08	0.12	0.29	0.23
Allowed (%)	6.94	5.72	6.59	7.53
Favored (%)	92.98	94.16	93.11	92.24

Supplementary Table 3. Primer List.

LH3-F	CCGGCGCGCCATGACCTCCTCGGGGC
LH3-R	CCGCGGCCGCGGGGTCGACAAAGGACA
ColGalT1-F	CCGGCGCGCCGCCCCGCCGGGCGCC
ColGalT1-R	CCGCGGCCGCGAGTTCATCCCGGGCAGCACT
ColGalT1-L59A-F	TCGCGCTGGCGGCGCGAAACGCGGCCCACGCGTTG
ColGalT1-L59A-R	TTTCGCGCCGCCAGCGCGATGAGCACGCGCGGCG
ColGalT1-A60E-F	CGCTGTTGGAGCGAAACGCGGCCCACGCGTTGCC
ColGalT1-A60E-R	GCGTTTCGCTCCAACAGCGCGATGAGCACGCGCG
ColGalT1-R61A-F	CTGTTGGCGGCAAACGCGGCCCACGCGTTGCCCAC
ColGalT1-R61A-R	CCGCGTTTGCCGCCAACAGCGCGATGAGCACGCGC
ColGalT1-Y126A-F	CCCAGGTCCGCCCCGGACGAGGAAGGCCCGAAAC
ColGalT1-Y126A-R	CGTCCGGGGCGGACCTGGGCTCCTCTGCTGGCCG
ColGalT1-R139A-F	GGTCTGACTCAGCCTACGAGCATGTCATGAAGTTG
ColGalT1-R139A-R	ATGCTCGTAGGCTGAGTCAGACCAGTGTTTCGGGC
ColGalT1-R147A-F	GTCATGAAGTTGGCCCAGGCAGCCCTGAAATCAGC
ColGalT1-R147A-R	GCTGCCTGGGCCAACTTCATGACATGCTCGTAGCG
ColGalT1-D166A-F	CCTGTTTGTAGCTGCGGACAACCTGATCCTCAACC
ColGalT1-D166A-R	TTGTCCGCAGCTACAAACAGGATGTAATCAGCCCAC
ColGalT1-D168A-F	GTAGATGCGGCCAACCTGATCCTCAACCCTGACAC
ColGalT1-D168A-R	GATCAGGTTGGCCGCATCTACAAACAGGATGTAATCAG
ColGalT1-S236A-F	ATGGTGCACGCGACCTTCCTGATCGACCTGCGGAA
ColGalT1-S236A-R	GGAAGGTCGCGTGCACCATGGGAACTGCAAAGCAG
ColGalT1-D264A-F	TGGTCCTTTGCCGACATCATCGTCTTTGCCTTCTC
ColGalT1-D264A-R	GATGATGTCGGCAAAGGACCAGGTGTAGTCAGGGT
ColGalT1-R354A-F	CGGCAGGACGCGCGGGAGCGCATGCTGCGGGCG
ColGalT1-R354A-R	GCTCCCGCGCGTCCTGCCGCCGCCTCAGGTTGAT
ColGalT1-E435A-F	GCTTGTGTTTGCGGATGACCTGCGTTTTGAGATCTTC
ColGalT1-E435A-R	CAGGTCATCCGCAAACACAAGCGATTTCTGCAGCCC
ColGalT1-D437A-F	GTTTGAGGATGCCCTGCGTTTTGAGATCTTCTTCAAG
ColGalT1-D437A-R	AAACGCAGGGCATCCTCAAACACAAGCGATTTCTGC
ColGalT1-T571A-F	TGTGAGTGACGCCGAGACCTCAGTCGTATGGAAC
ColGalT1-T571A-R	AGGTCTCGGCGTCACTCACATAGCCATCGTCTCCT
ColGalT1-Y37A-F	CCGACGCCGCCTTCCCCGAGGAGCGCTGGAGCCCG
ColGalT1-Y37A-R	TCGGGGAAGGCGGCGTCGGCGCCCGGCGGGGC
ColGalT1-F38A-F	GACGCCTACGCCCCCGAGGAGCGCTGGAGCCCGGA
ColGalT1-F38A-R	CCTCGGGGGCGTAGGCGTCGGCGCCCGGCGGGGC
ColGalT1-P39A-F	CGCCTACTTCGCCGAGGAGCGCTGGAGCCCGGAG
ColGalT1-P39A-R	GCTCCTCGGCGAAGTAGGCGTCGGCGCCCGGCGG
ColGalT1-E40A-F	CTTCCCCGCGGAGCGCTGGAGCCCGGAGTCGCC
ColGalT1-E40A-R	AGCGCTCCGCGGGGAAGTAGGCGTCGGCGCCCG
ColGalT1-E41A-F	CCCCGAGGCGCGCTGGAGCCCGGAGTCGCCCCT
ColGalT1-E41A-R	TCCAGCGCGCCTCGGGGAAGTAGGCGTCGGCGC

ColGalT1-R42A-F	CCCGAGGAGGCCTGGAGCCCGGAGTCGCCCCTG
ColGalT1-R42A-R	GGGCTCCAGGCCTCCTCGGGGAAGTAGGCGTCG
ColGalT1-S44A-F	AGCGCTGGGCCCCGGAGTCGCCCCTGCAGGCGCC
ColGalT1-S44A-R	GACTCCGGGGCCCAGCGCTCCTCGGGGAAGTAGG
ColGalT1-E46A-F	GAGCCCGGCGTCGCCCCTGCAGGCGCCGCGCGT
ColGalT1-E46A-R	GGGGCGACGCCGGGCTCCAGCGCTCCTCGGGGA
LH3-V44A-F	GTGATCACTGCGGCCACAGCTGAAACCGAGGGGTA
LH3-V44A-R	CTGTGGCCGCAGTGATCACCAGCAGCTTCTCTGGG
LH3-D112A-F	CATGTTTGTGGCTAGCTACGACGTGATTCTGGCCG
LH3-D112A-R	GTCGTAGCTAGCCACAAACATGATGATCATATCCTCCC
LH3-D115A-F	GGATAGCTACGCCGTGATTCTGGCCGGCAGCCCC
LH3-D115A-R	CAGAATCACGGCGTAGCTATCCACAAACATGATGATC
LH3-H253A-F	CATTGTGGTCGCTGGAAACGGTCCCACTAAGCTG
LH3-H253A-R	ACCGTTTCCAGCGACCACAATGGGGAGCGTGTCGT
LH3-G256A-F	CATGGAAACGCTCCCACTAAGCTGCAGCTCAACTA
LH3-G256A-R	GCTTAGTGGGAGCGTTTCCATGGACCACAATGGGG
LH3-K259A-F	GGTCCCACTGCGCTGCAGCTCAACTACCTGGGAAA
LH3-K259A-R	GCTGCAGCGCAGTGGGACCGTTTCCATGGACCACA
LH3-Y656A-F	GTGGTTCGCGCCCGGCCAGACGAGCAGCCGTCTCT
LH3-Y656A-R	TCTGGCCGGGCGCGAACCACAAAGTTCATCACCGC
LH3-H667A-F	TGCGGCCAGCCCACGACTCATCCACCTTCACCCTCA
LH3-H667A-R	GAGTCGTGGGCTGGCCGCAGAGACGGCTGCTCGTC
LH3-D669A-F	CACACCACGCCTCATCCACCTTCACCCTCAACGTT
LH3-D669A-R	GGTGGATGAGGCGTGGTGTGGCCGCAGAGACGGCT
LH3-H719A-F	CACCCACTACGCCGAGGGGCTGCCAACGACCTGGG
LH3-H719A-R	GCCCCTCGGCGTAGTGGGTGAGGCGGCCGGGGTGC
LH3-L226A-F	ACGGGGCTGCAGATGAAGTGGTTTTAAAGTTTGATCGG
LH3-L226A-R	CCACTTCATCTGCAGCCCCGTTGAGGTTCTGAAAGA
LH3-V229A-F	GCTTTAGATGAAGCGGTTTTAAAGTTTGATCGGAACCG
LH3-V229A-R	AAACTTTAAAACCGCTTCATCTAAAGCCCCGTTGAGG
LH3-V230A-F	AGATGAAGTGGCTTTAAAGTTTGATCGGAACCGTGTGC
LH3-V230A-R	GATCAAACTTTAAAGCCACTTCATCTAAAGCCCCGTTGAG
LH3-L231A-F	ATGAAGTGGTTGCAAAGTTTGATCGGAACCGTGTGCG
LH3-L231A-R	CGATCAAACTTTGCAACCACTTCATCTAAAGCCCCGTTG
LH3-R241A-F	TGTGCGTATCGCGAACGTGGCCTACGACACGCTC
LH3-R241A-R	GCCACGTTCGCGATACGCACACGGTTCCGATCA
LH3-Y264A-F	GCAGCTCAACGCCCTGGGAAACTACGTCCCCAATGG
LH3-Y264A-R	GTTTCCCAGGGCGTTGAGCTGCAGCTTAGTGGGAC
LH3-L265A-F	AGCTCAACTACGCGGGAAACTACGTCCCCAATGGCT
LH3-L265A-R	GTAGTTTCCCGCGTAGTTGAGCTGCAGCTTAGTGG
LH3-E446A-F	GACTACGTGGCGCTGGTGCAGCGGAAGCGAGTG
LH3-E446A-R	TGCACCAGCGCCACGTAGTCCTCGGAGCGGGCG
LH3-R452A-F	CAGCGGAAGGCAGTGGGTGTGTGGAATGTACCA

LH3-R452A-R	ACACCCACTGCCTTCCGCTGCACCAGCTCCACG

Supplementary Movie: The assembling of KOGG tetramer and its polymerization.