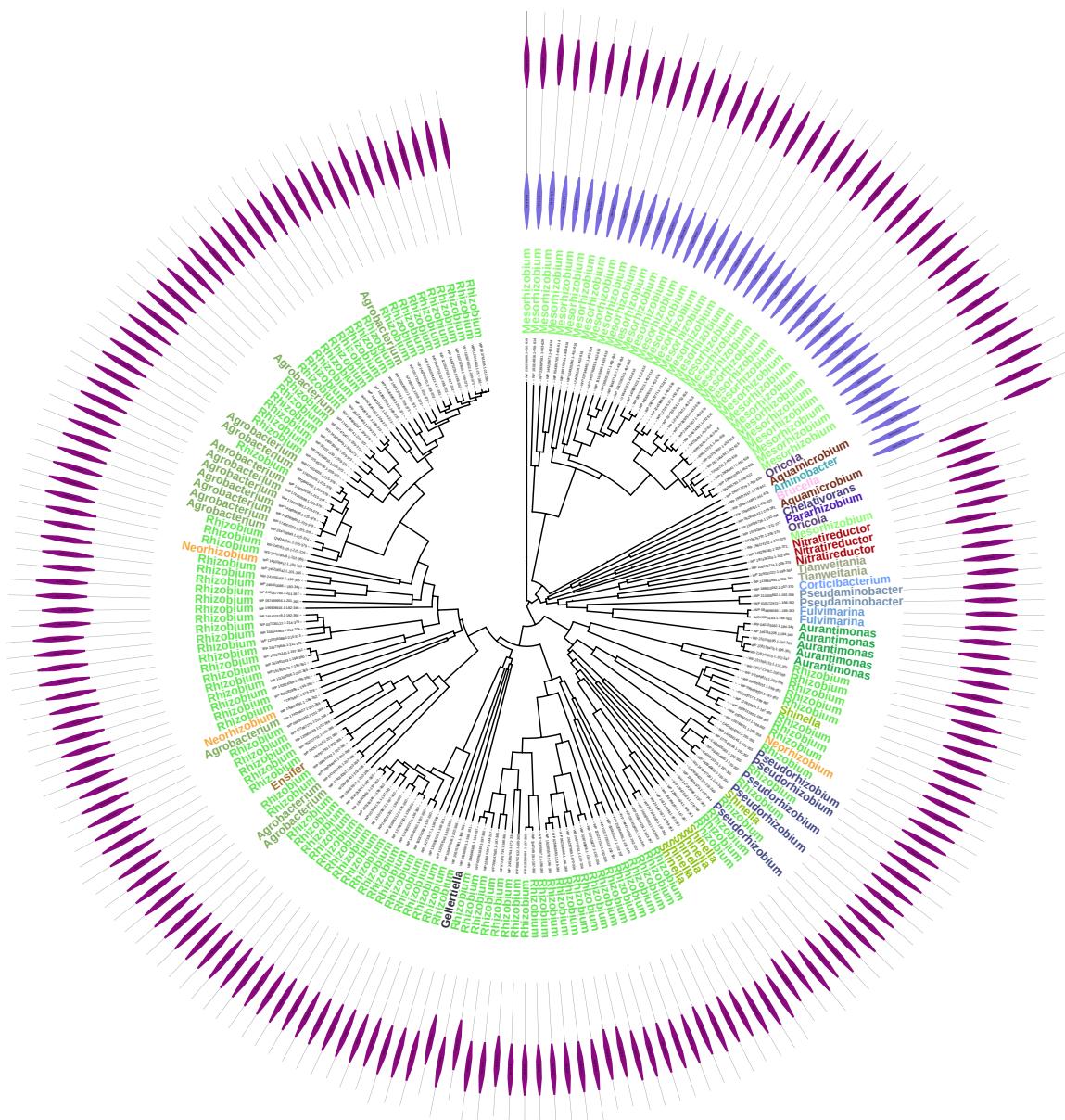


Supplementary Fig. 1. Aclust tree of representative O-Lig sequences. The four subfamilies are shown in red, green, turquoise, and yellow. Seed sequences are marked with a red bar.



Supplementary Fig. 2. Phylogenetic tree of representative WadA homologs from the GTxx3 family, showing the taxonomic origin of the sequences and the domain modularity. N-terminal GT25 domains are shown in blue and GTxx3 domain in purple. The two-domain sequences segregate from proteins with a single domain GTxx3, suggesting co-evolution of the two domains.

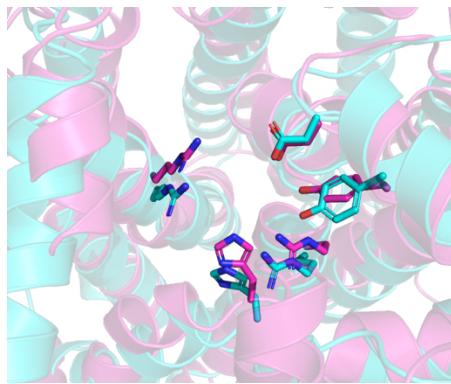
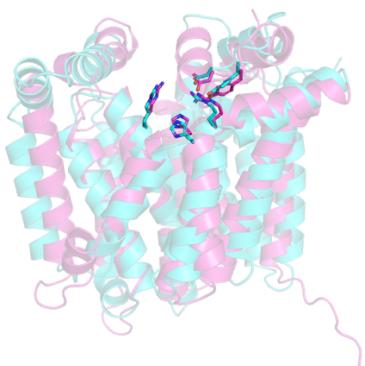
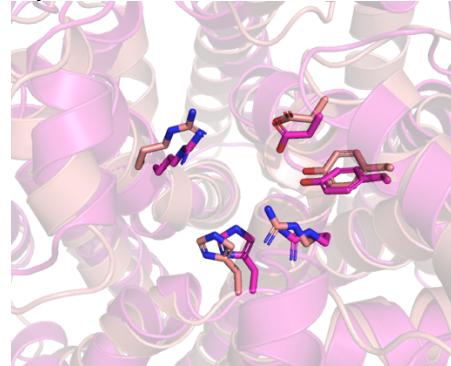
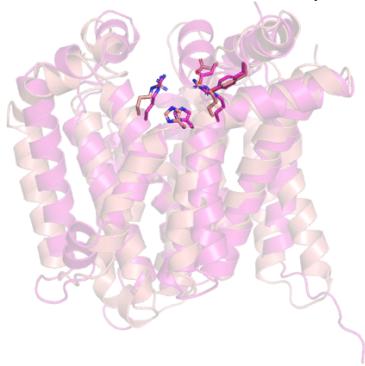
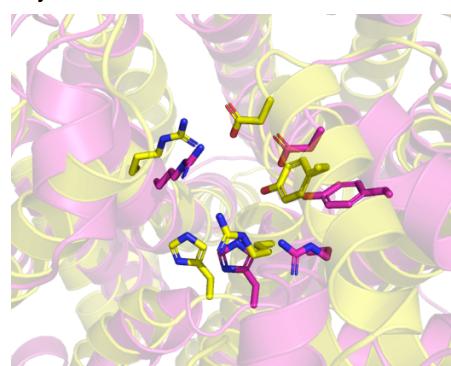
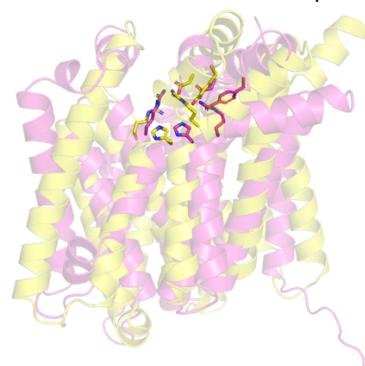


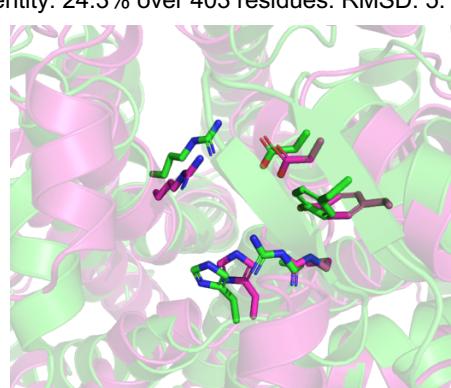
ABB29906.1 and AHB32334.1. Sequence identity: 21.4% over 421 residues. RMSD: 4.3 over 304 residues.



ACA24821.1 and AHB32334.1. Sequence identity: 21.7% over 415 residues. RMSD: 3.9 over 312 residues.



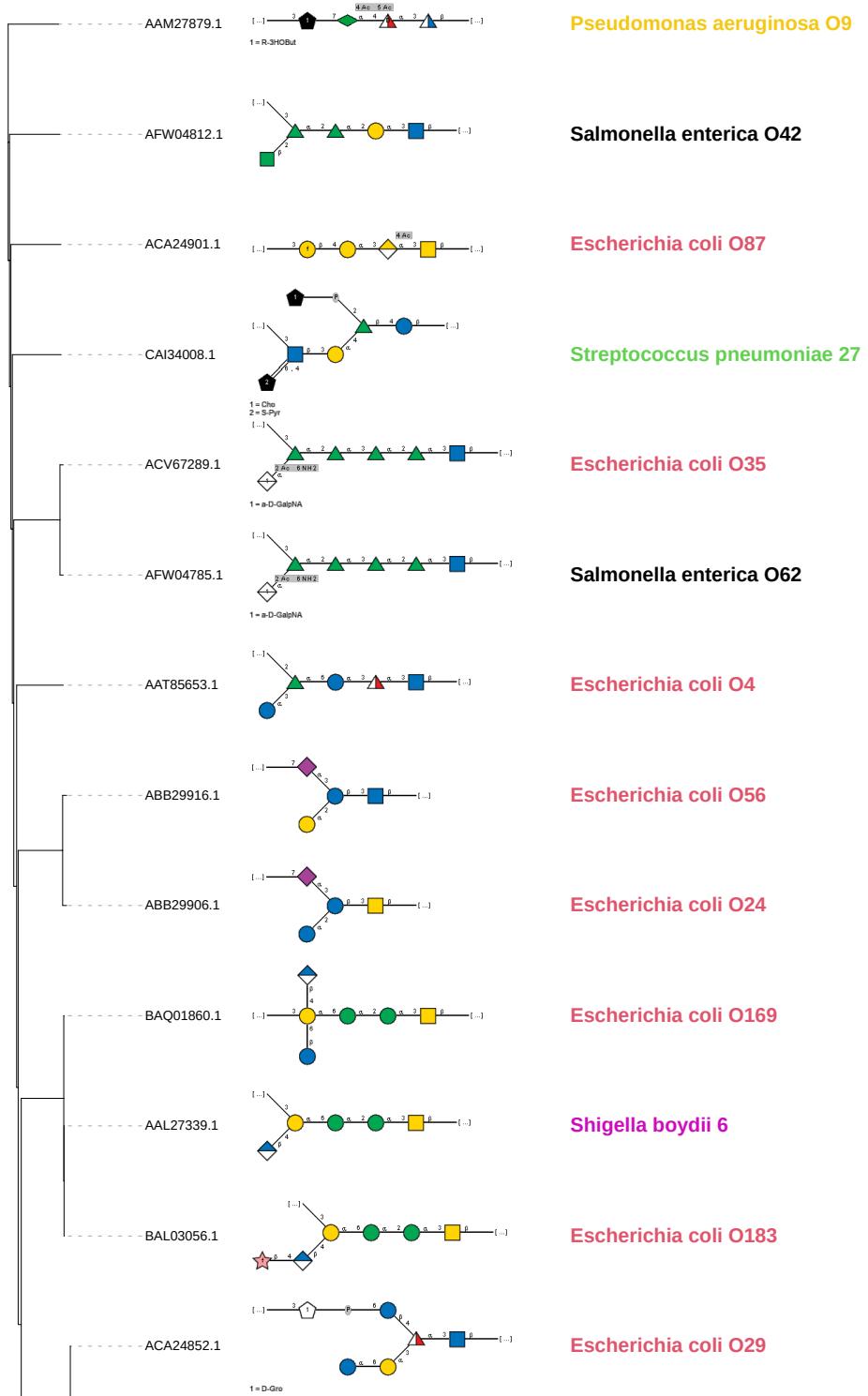
AHB32215.1 and AHB32334.1. Sequence identity: 24.3% over 403 residues. RMSD: 5.1 over 248 residues.

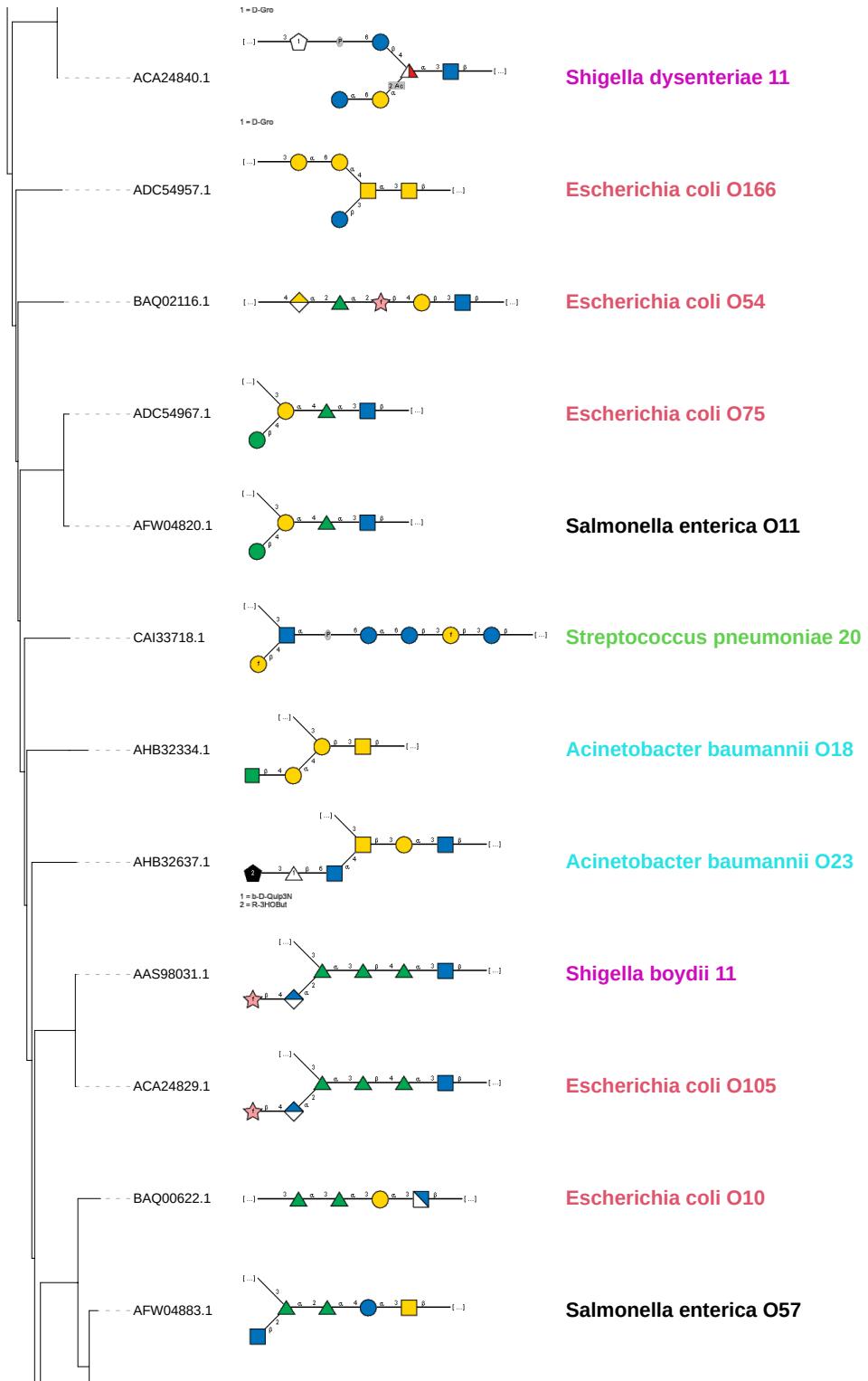


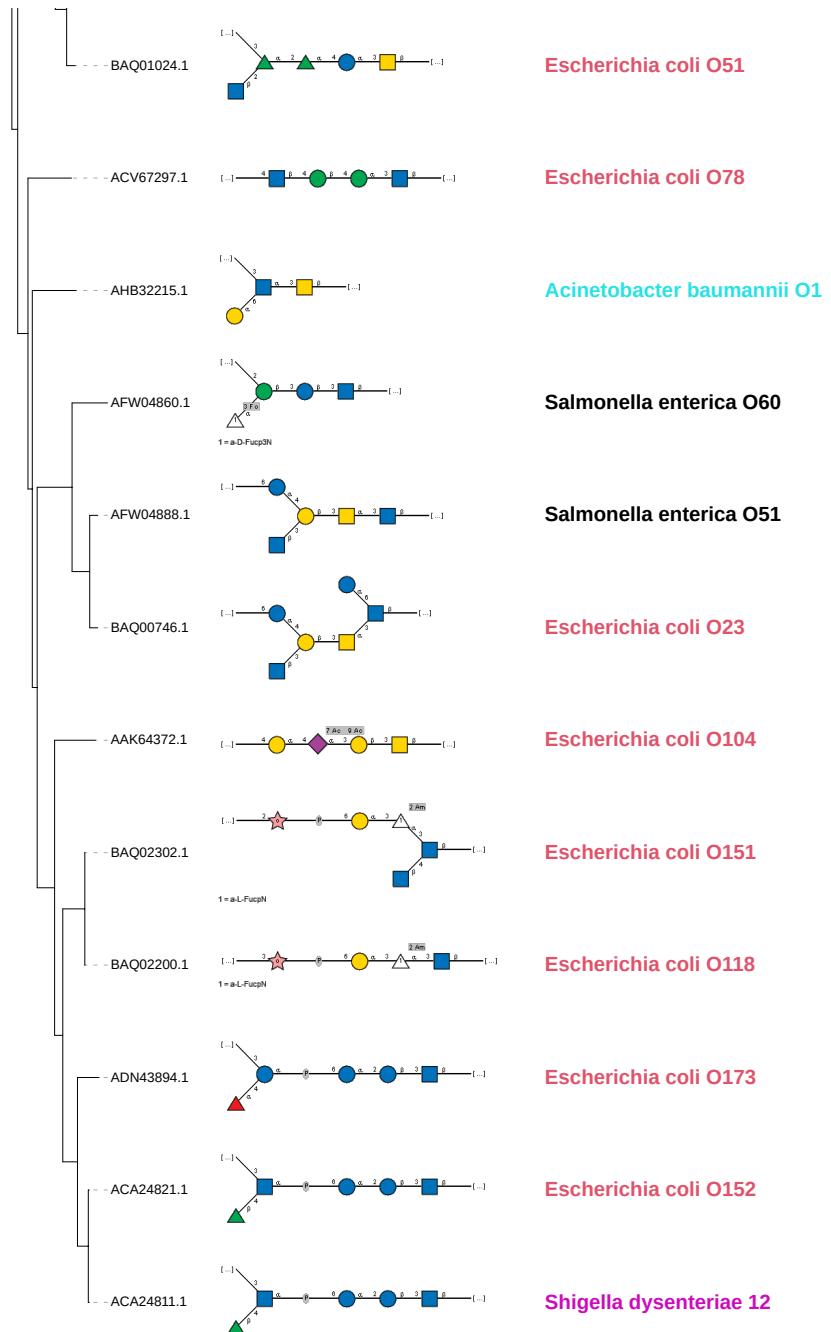
CAI34008.1 and AHB32334.1. Sequence identity: 22.4% over 416 residues. RMSD: 4.2 over 320 residues.

Supplementary Fig. 3. Structural superimpositions of AlphaFold models of distantly related BP-Pols from family GTxx4 aligned to a reference, AHB32334.1. AHB32334.1 is shown in pink in all of the figures.

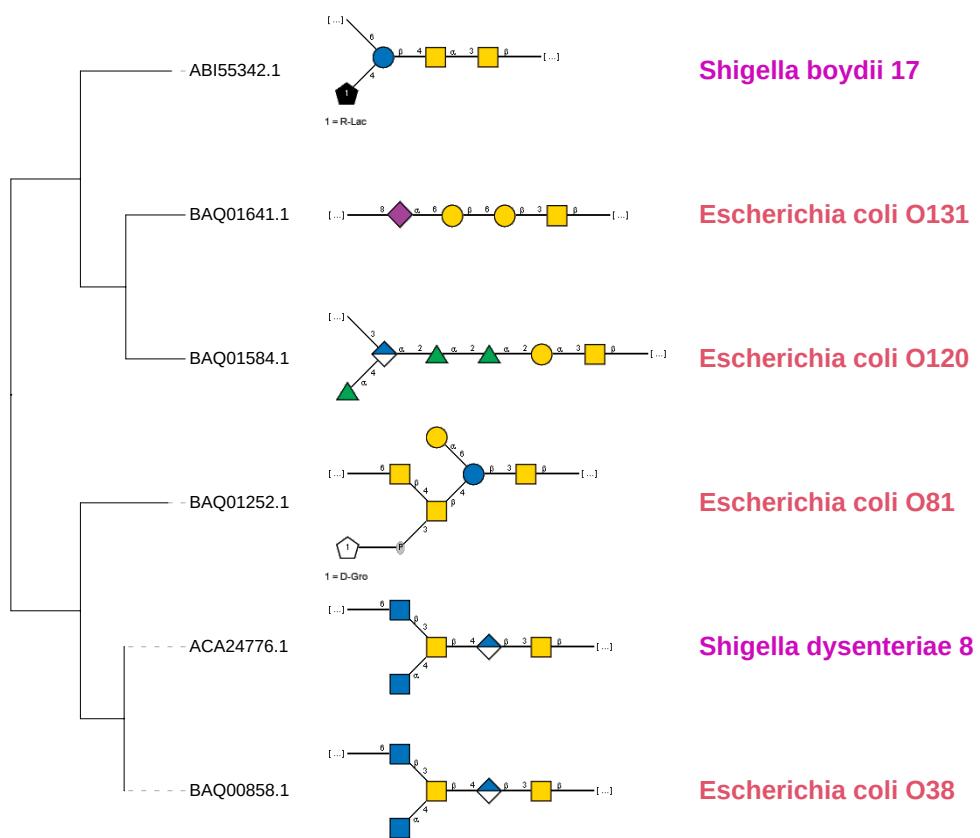
GTxx4:



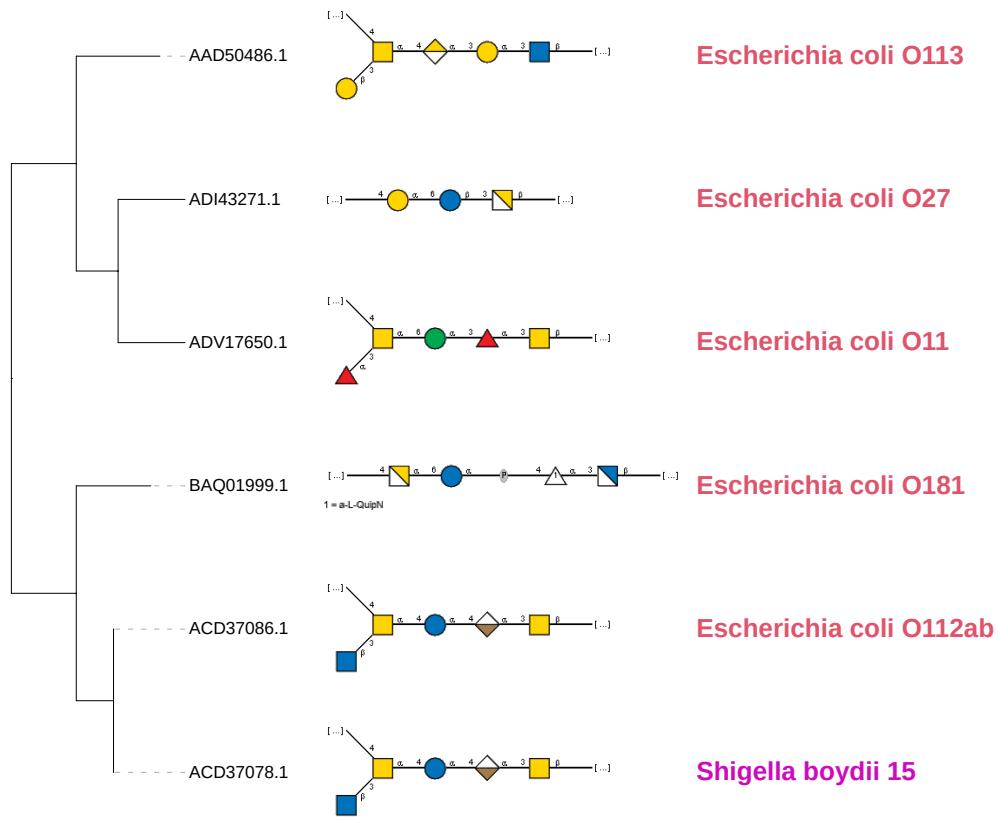




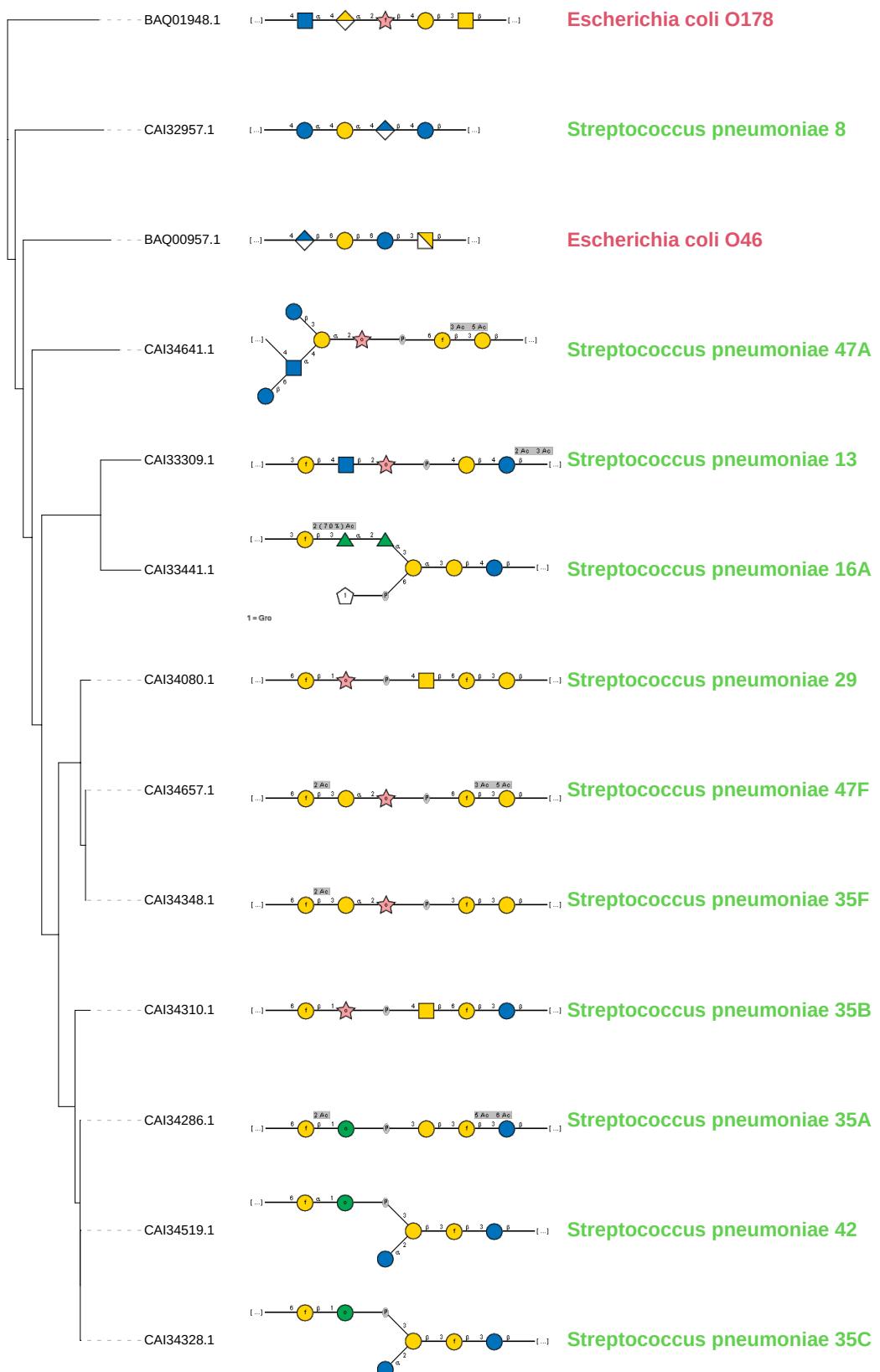
GTxx5:



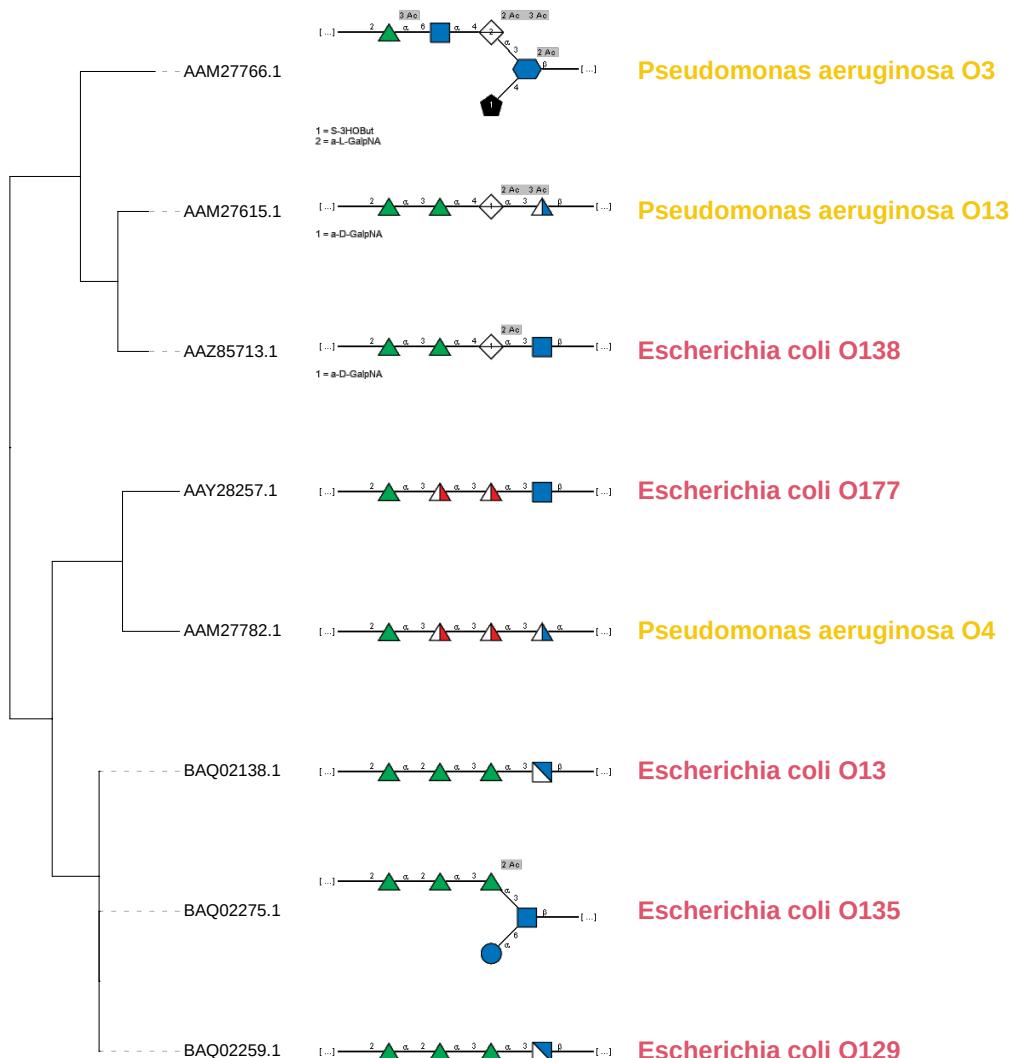
GTxx6:



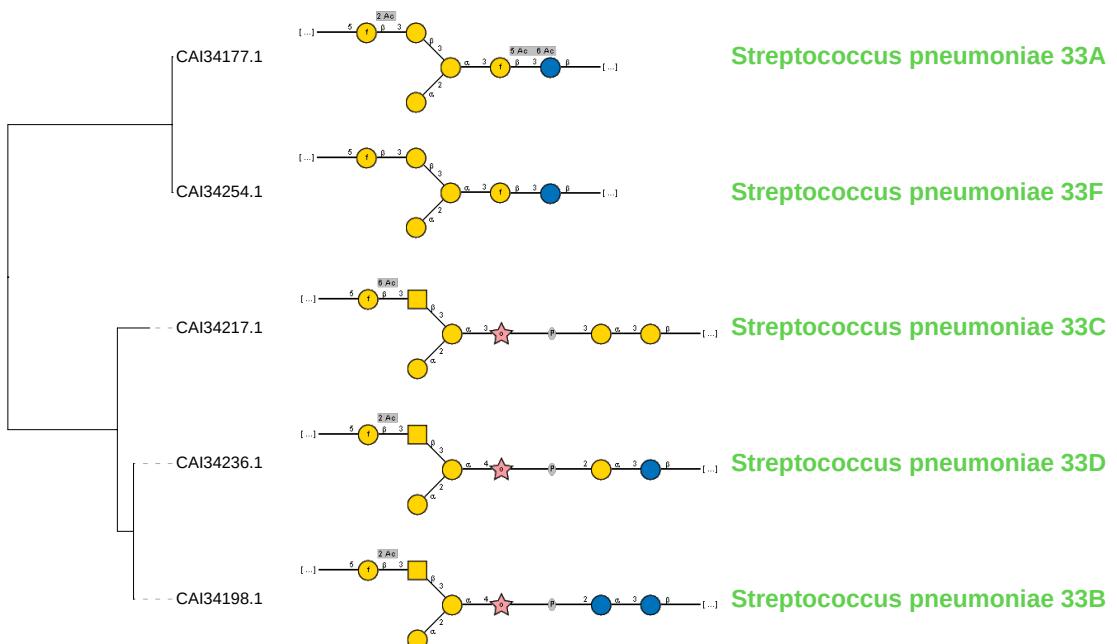
GTxx7:



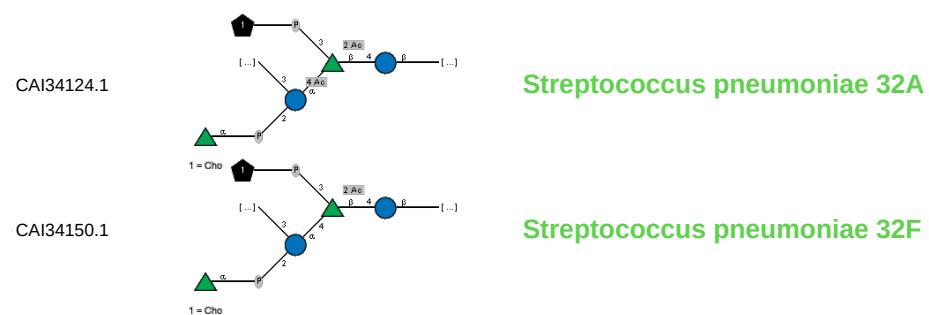
GTxx8:



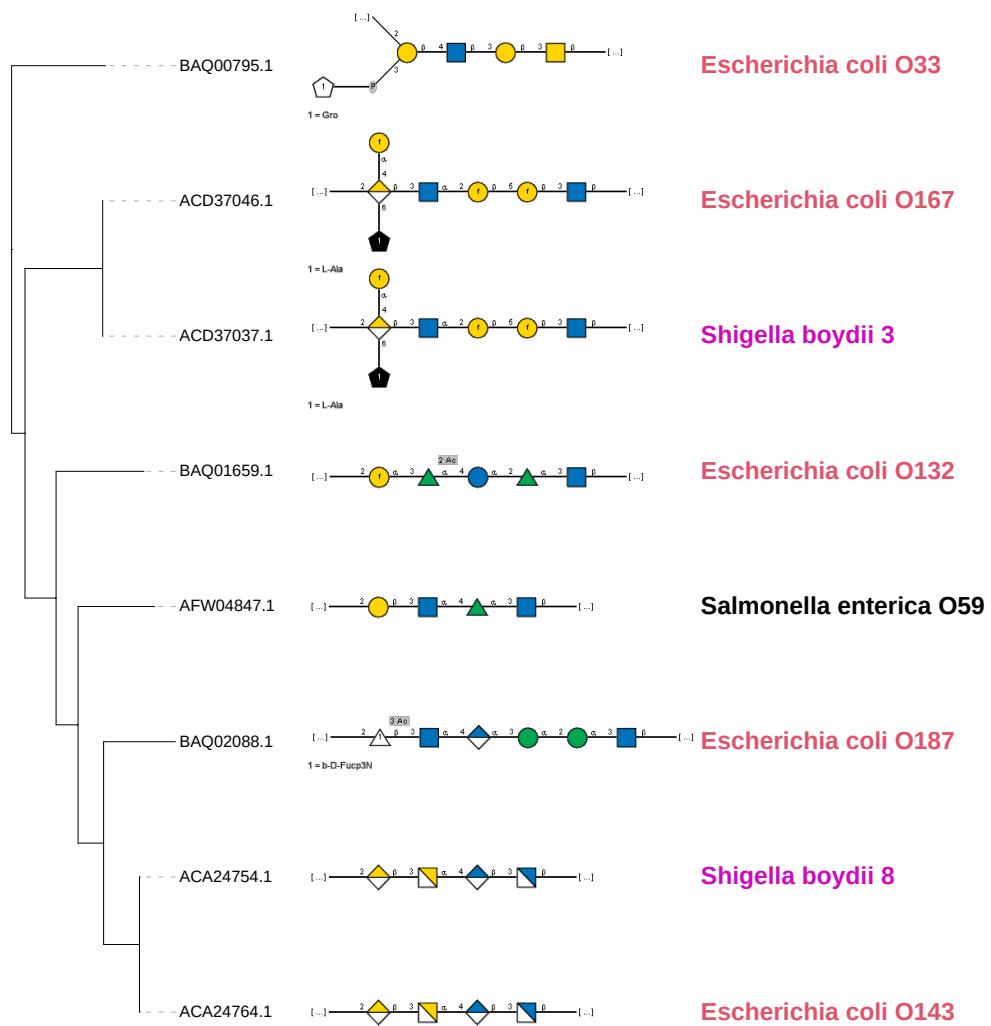
GTxx9:



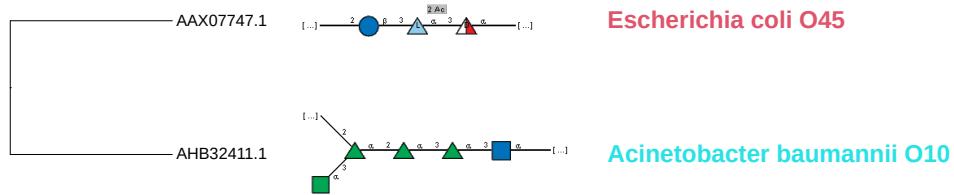
GTx10:



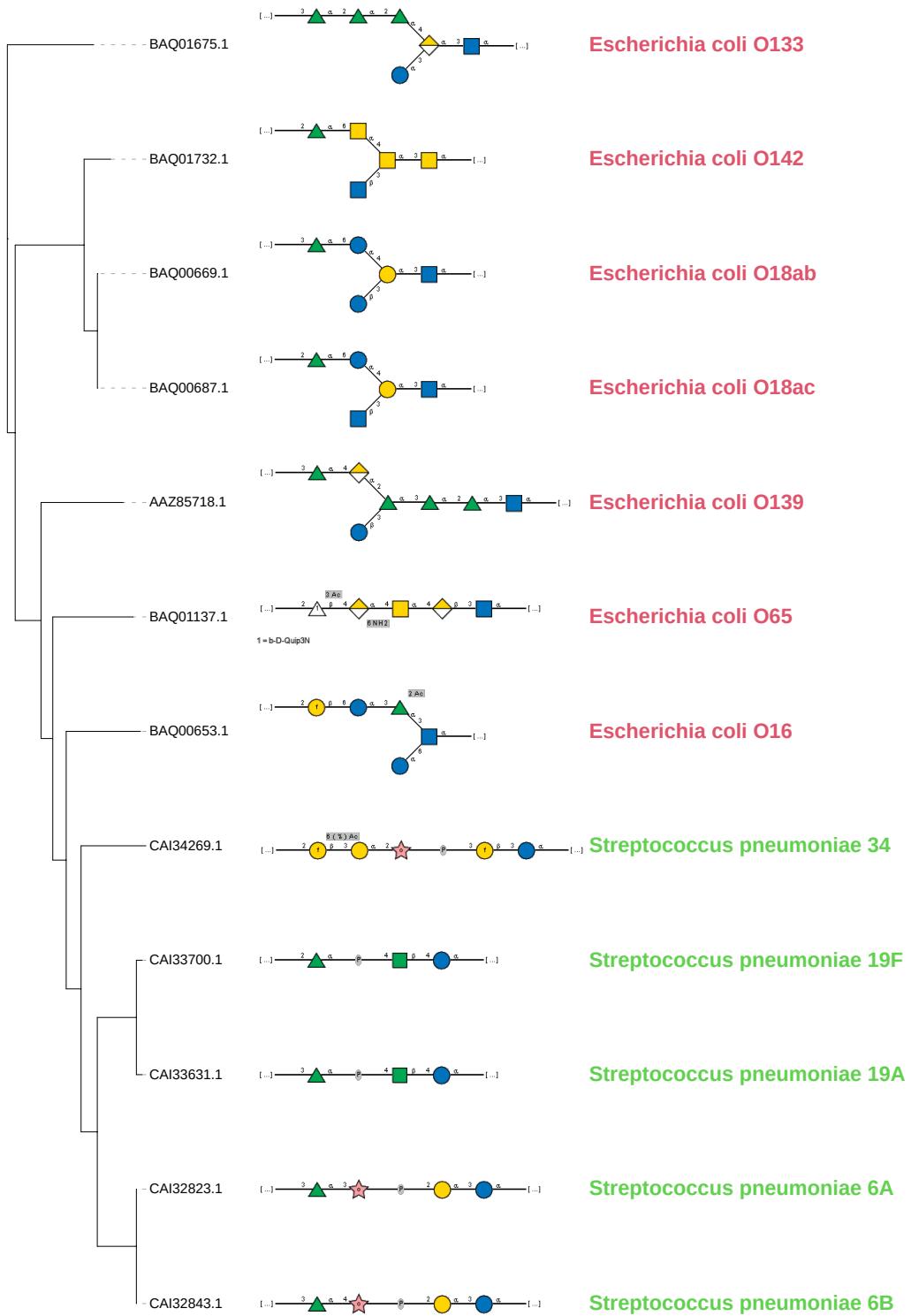
GTx11:



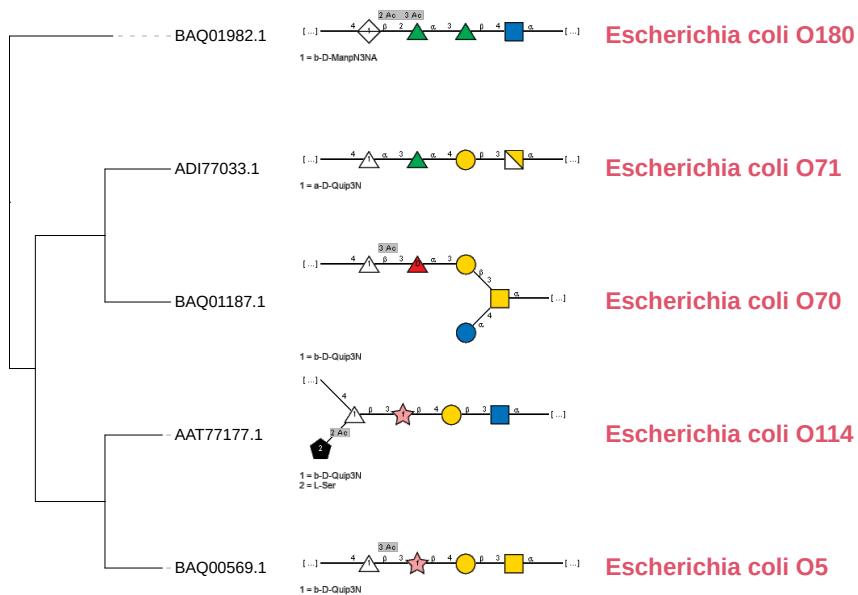
GTx12:



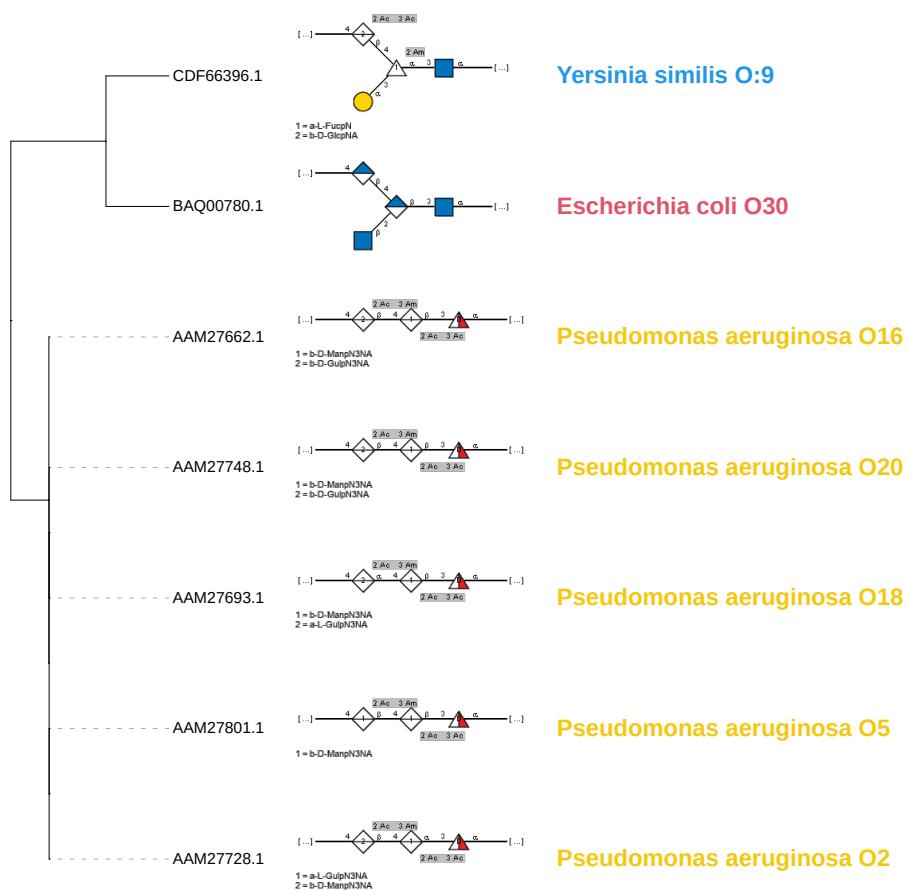
GTx13:



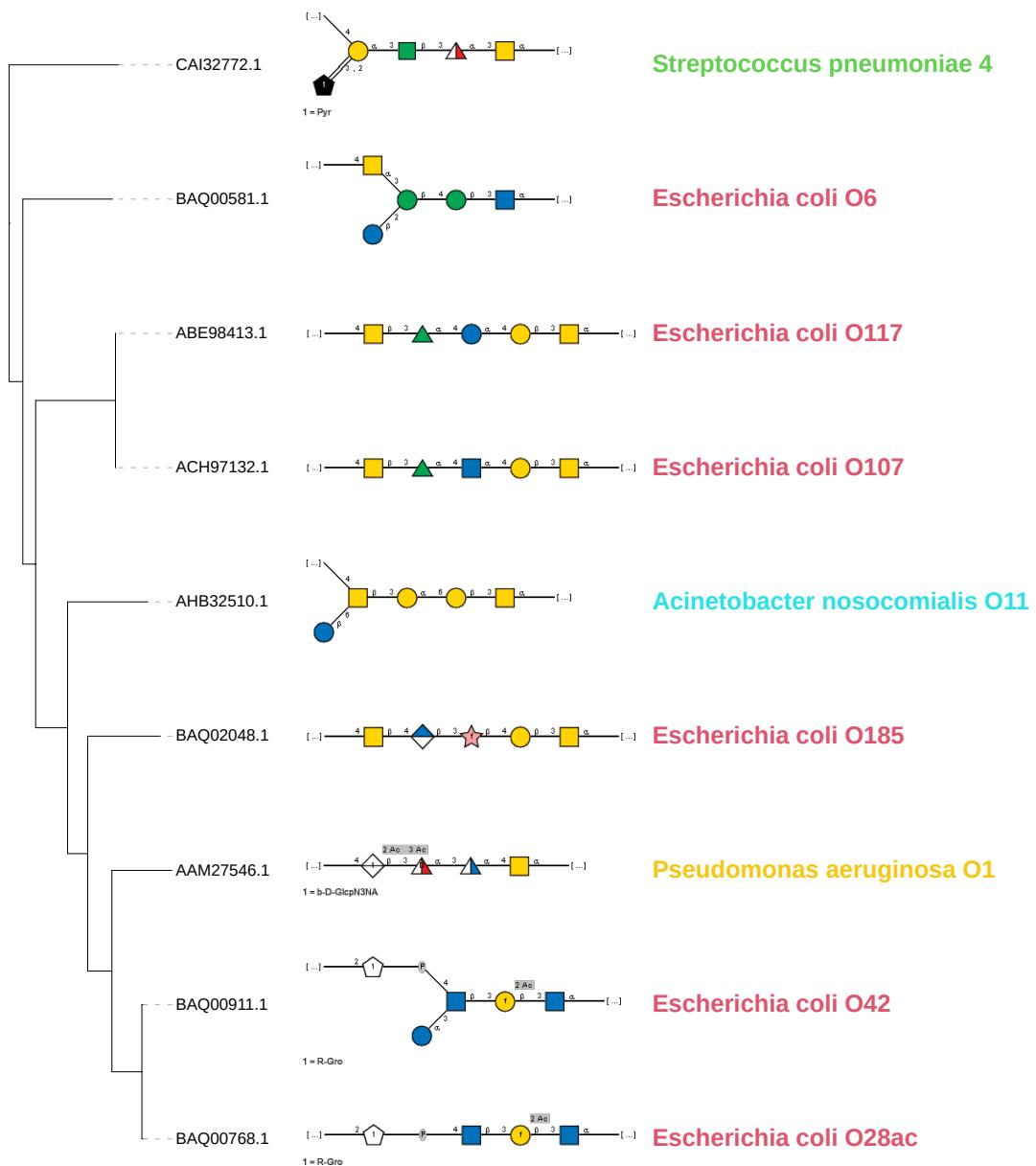
GTx14:



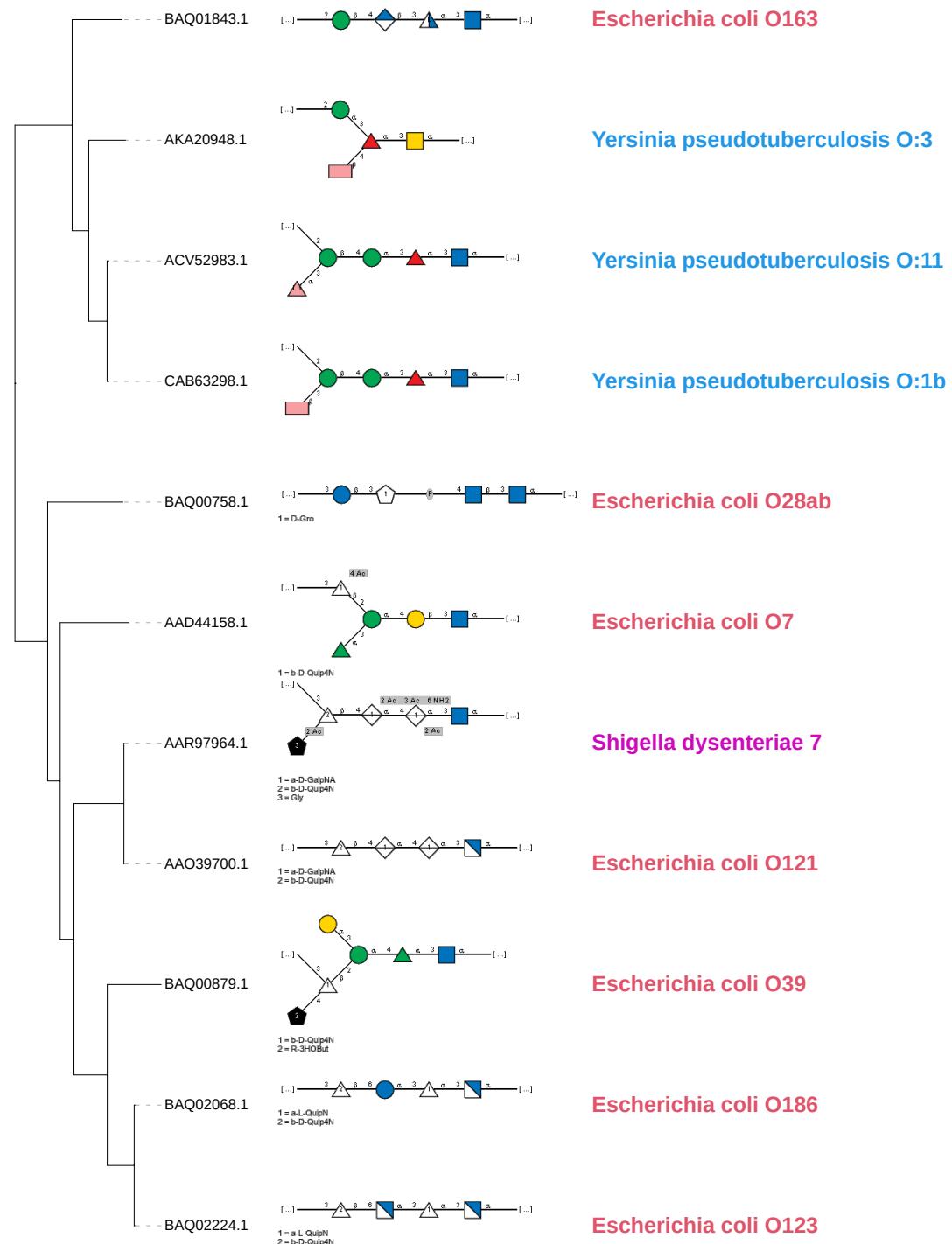
GTx15:



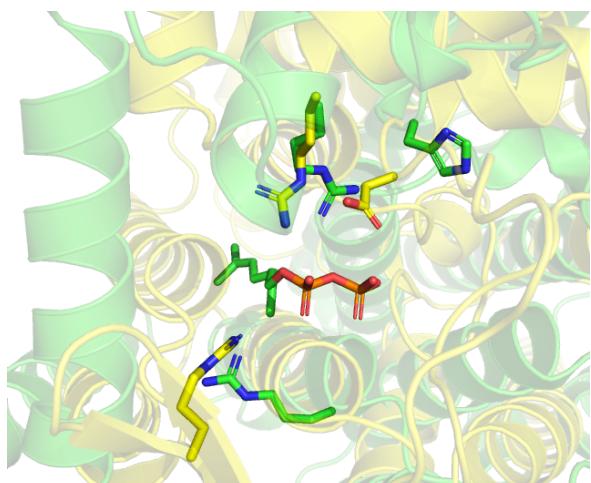
GTx16:



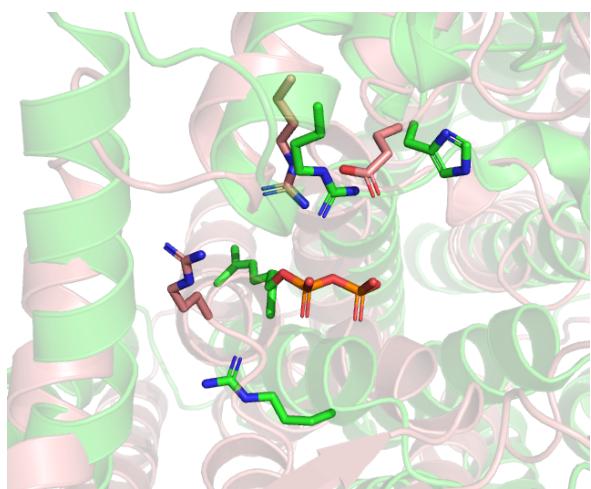
GTx17:



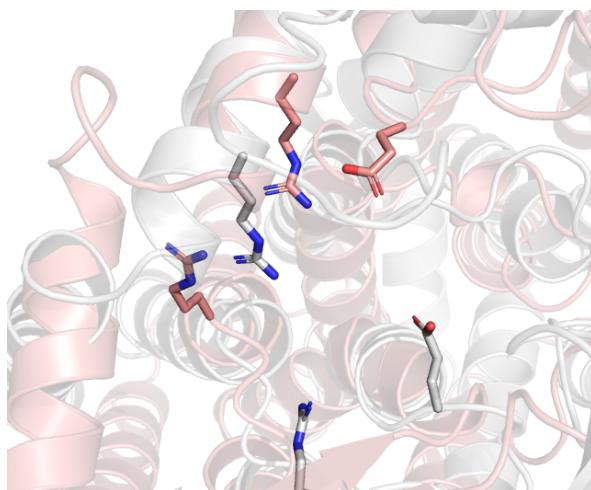
Supplementary Fig. 4. Phylogenetic trees of the BP-Pols in each family with structures of the corresponding sugar repeat units in SNFG format visualized using iTOL. The tree for GTxx4 is split in three pages due to its size. No tree is shown for GTx10, because the two BP-Pol members with known repeat units have identical sequences.



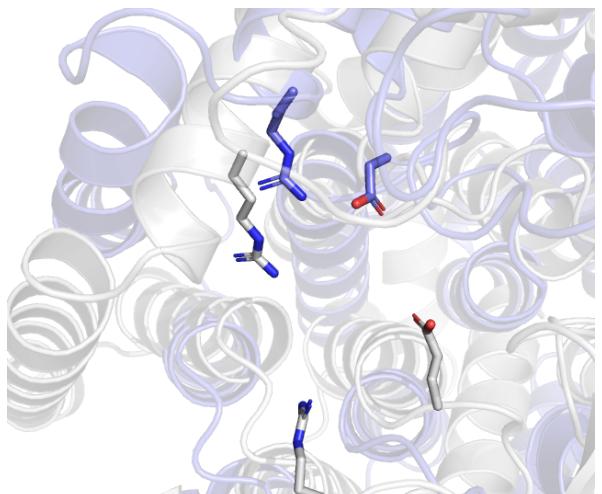
CAI34369 from GTxx7 (yellow) and 7TPG from GTxx3 (green). RMSD: 5.8 over 160 residues.



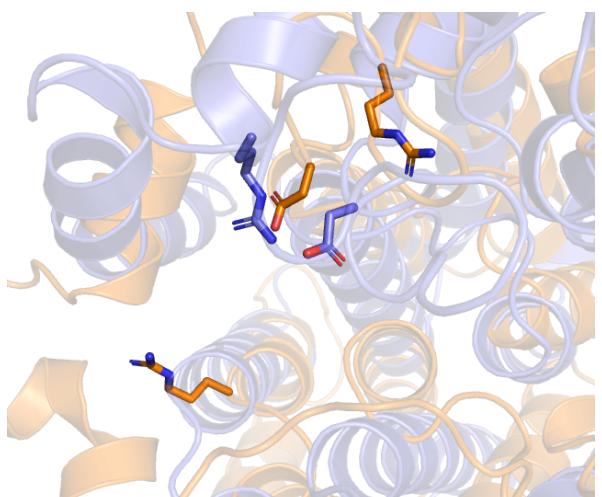
7TPG from GTxx3 (green) and AAM27615.1 from GTxx8 (pink). RMSD: 5.3 over 192 residues



AAM27615.1 from GTxx8 (pink) and CAI34254.1 from GTxx9 (grey). RMSD: 5.4 over 232 residues.

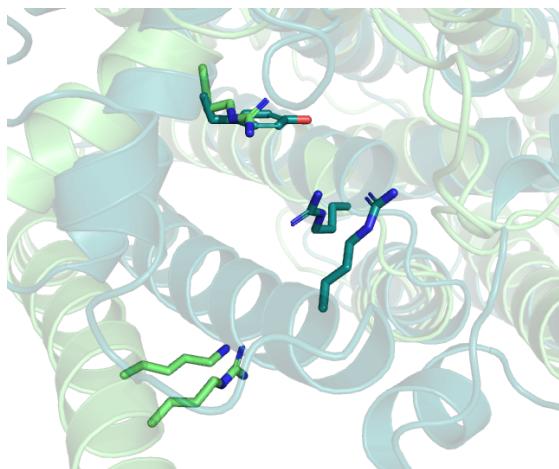


CAI34254.1 from GTx9 (grey) and CAI34124.1 from GTx10 (blue). RMSD: 5.7 over 208 residues.

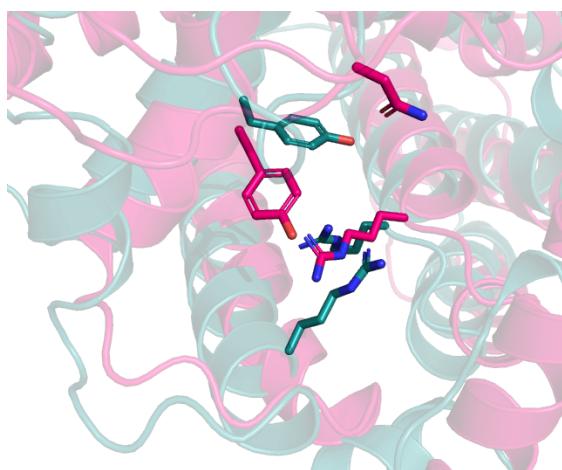


CAI34124.1 from GTx10 (blue) and BAQ02088.1 from GTx11 (orange). RMSD: 5.9 over 144 residues.

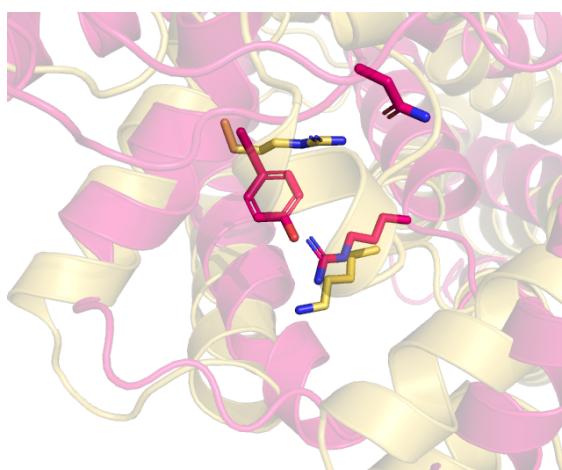
Supplementary Fig. 5. Structural superimpositions of representative members from each family in clan GT-C_{B1}. All structures are AlphaFold models except for 7TPG.



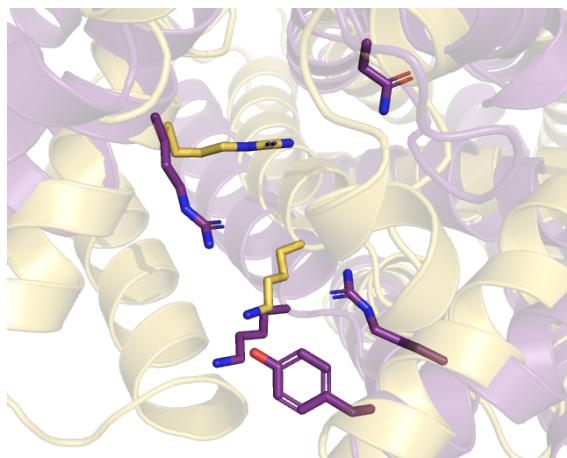
BAQ02224.1 from GTx17 (green) and AHB32411.1 from GTx12 (teal). RMSD: 4.9 over 168 residues.



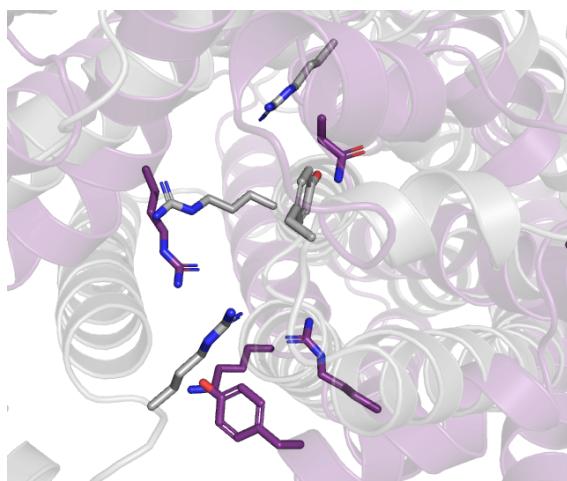
AHB32411.1 from GTx12 (teal) and CAI32823.1 from GTx13 (red). RMSD: 5.2 over 328 residues.



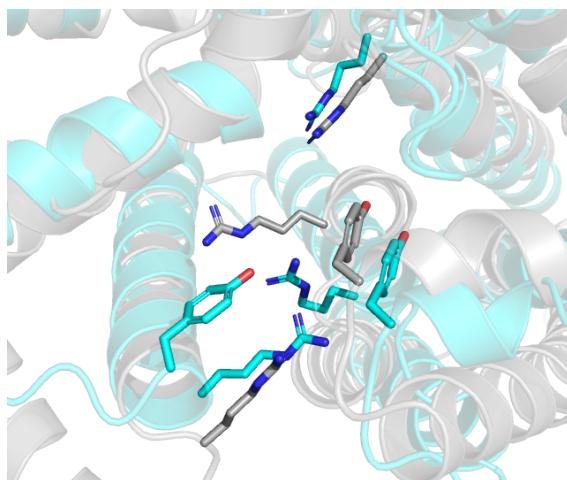
CAI32823.1 from GTx13 (red) and AAT77177.1 from GTx14 (yellow). RMSD: 4.9 over 232 residues.



AAT77177.1 from GTx14 (yellow) and AAM27801.1 from GTx15 (purple). RMSD: 4.7 over 200 residues.

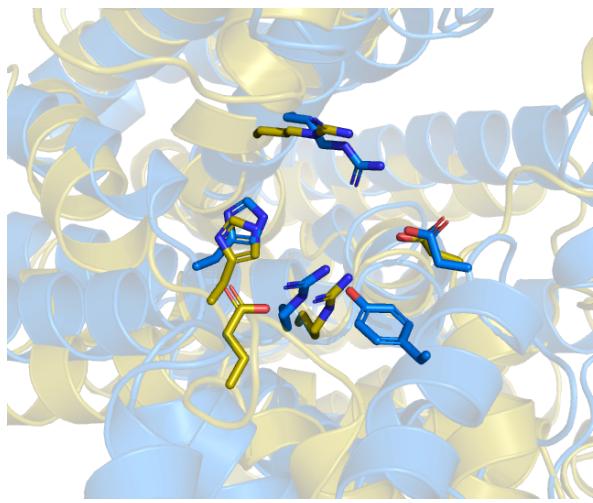


AAM27801.1 from GTx15 (purple) and CAI32772.1 from GTx16 (grey). RMSD: 5.8 over 184 residues.

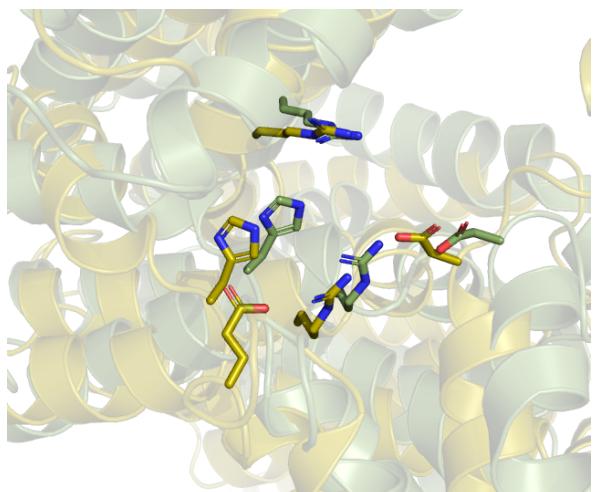


CAI32772.1 from GTx16 (grey) and ACH50550.1 from GTxx2 (turquoise). RMSD: 5.5 over 360 residues.

Supplementary Fig. 6. Structural superimpositions of representative members from each family in clan GT-C_{B2}. All structures are AlphaFold models.



AHB32861.1 from GTxx4 (blue) and BAQ01641.1 from GTxx5 (gold). RMSD: 4.4 over 200 residues.



BAQ01641.1 from GTxx5 (gold) and ADI43271.1 from GTxx6 (green). RMSD 5.7 over 352 residues.

Supplementary Fig. 7. Structural superimpositions of representative members from each family in clan GT-C_{B3}. All structures are AlphaFold models.