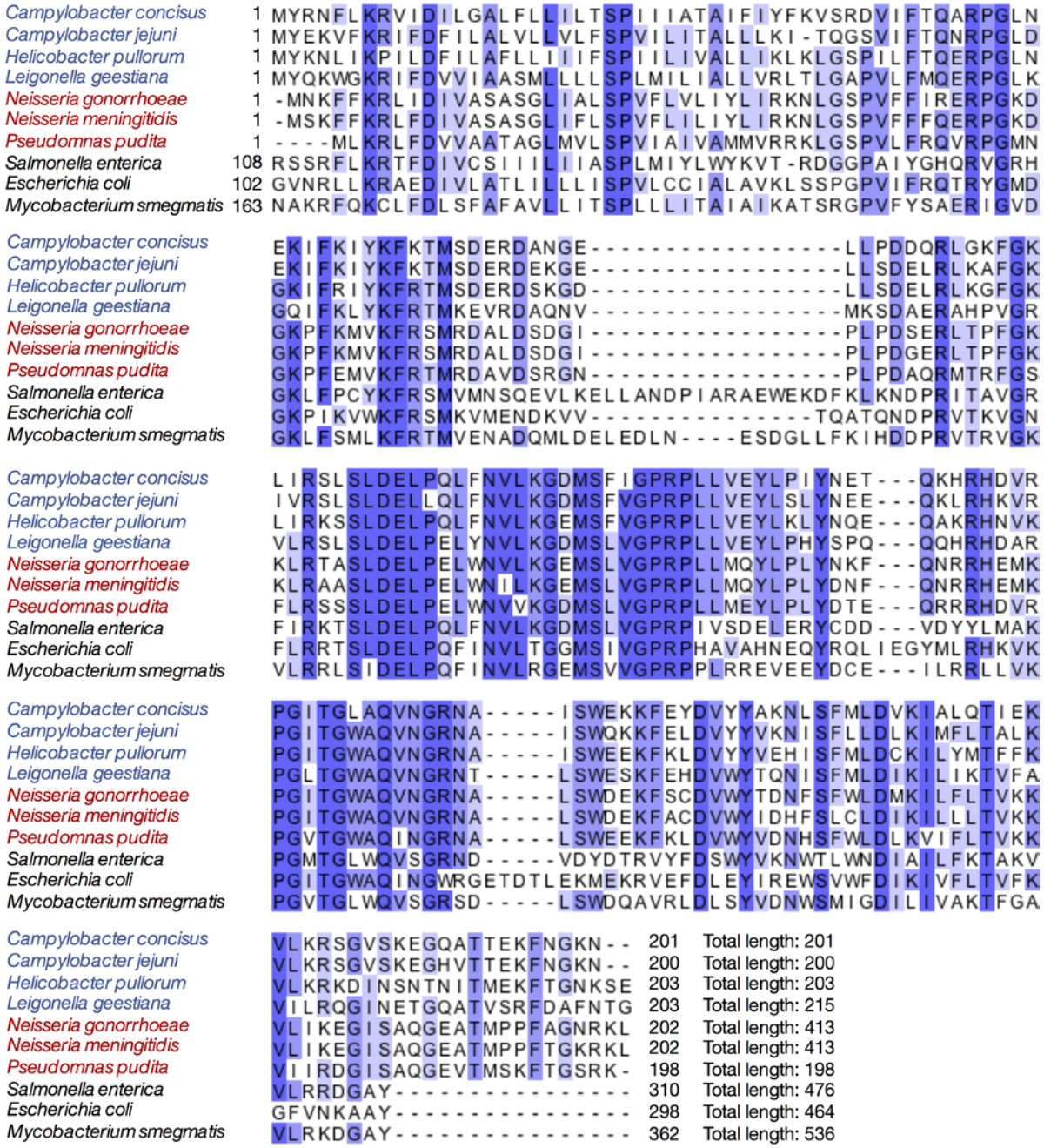
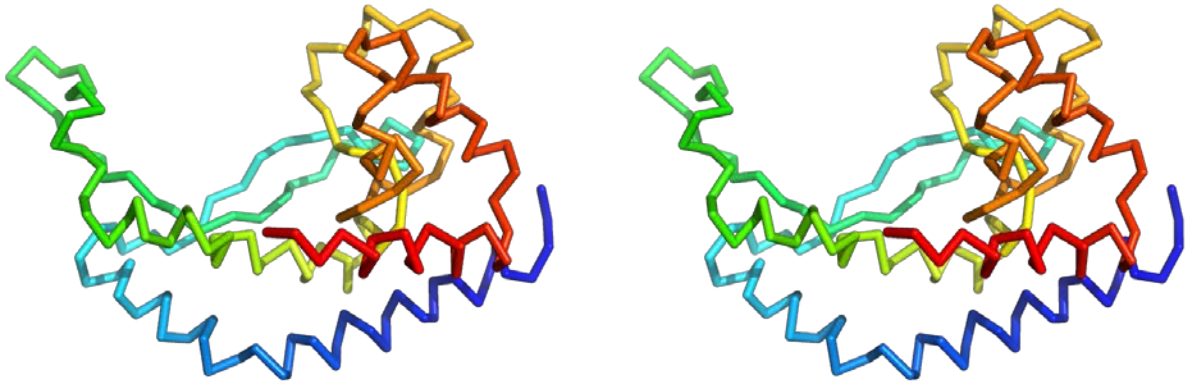
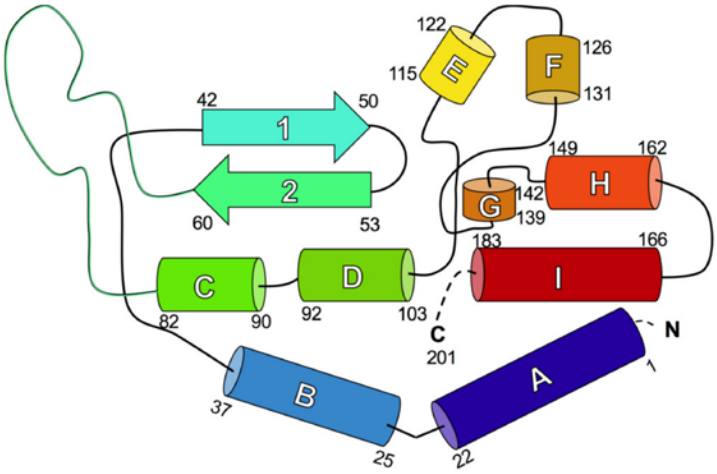


Supplementary Figure 1 | Sequence alignment of PglC-like minimal catalytic domain across the three monotopic-PGT families. Representative sequences from the PglC subfamily (blue): *C. concisus* PglC (A0A0M4SI81)/ *C. jejuni* PglC (O86156)/ *H. pullorum* PglC (E1B268)/ *L. geestiana* PglC (A0A0W0UA41). Representative sequences from the PglB-bifunctional subfamily (red): *N. gonorrhoeae* PglB (A0A1D3HQ90)/ *N. meningitidis* PglB (Q9RR58)/ *P. putida* Sugar transferase (A0A0P7CW64). Representative sequences from the WbaP subfamily (black): *S. enterica* WbaP (S4IKQ0)/ *E. coli* WcaJ (P71241)/ *M. smegmatis* WcaJ (A0A0D6J209).

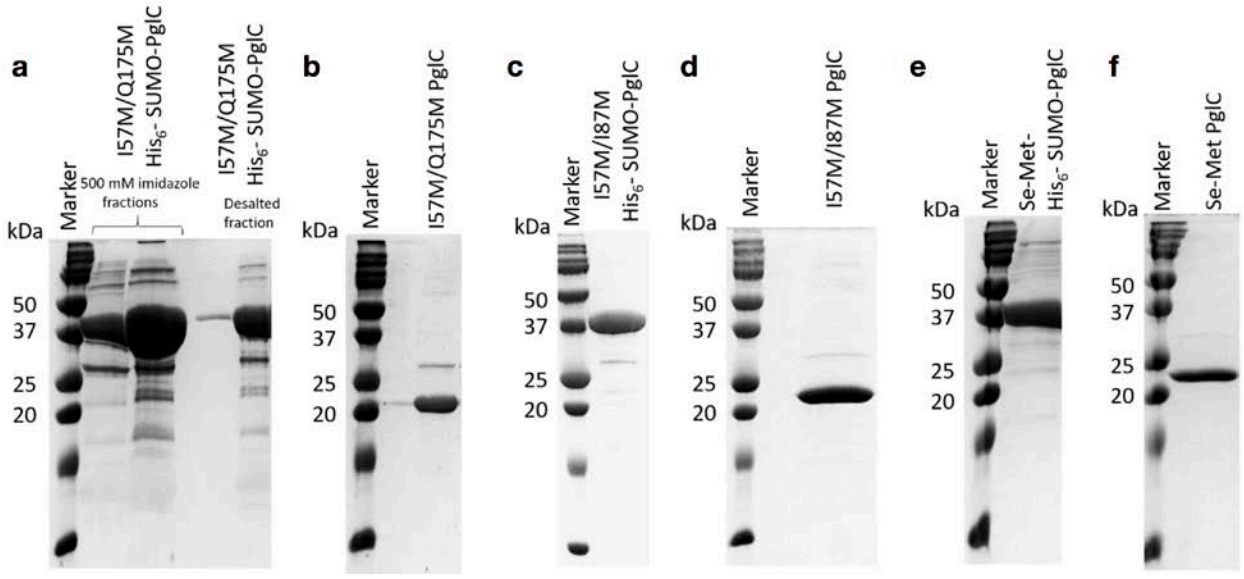
a



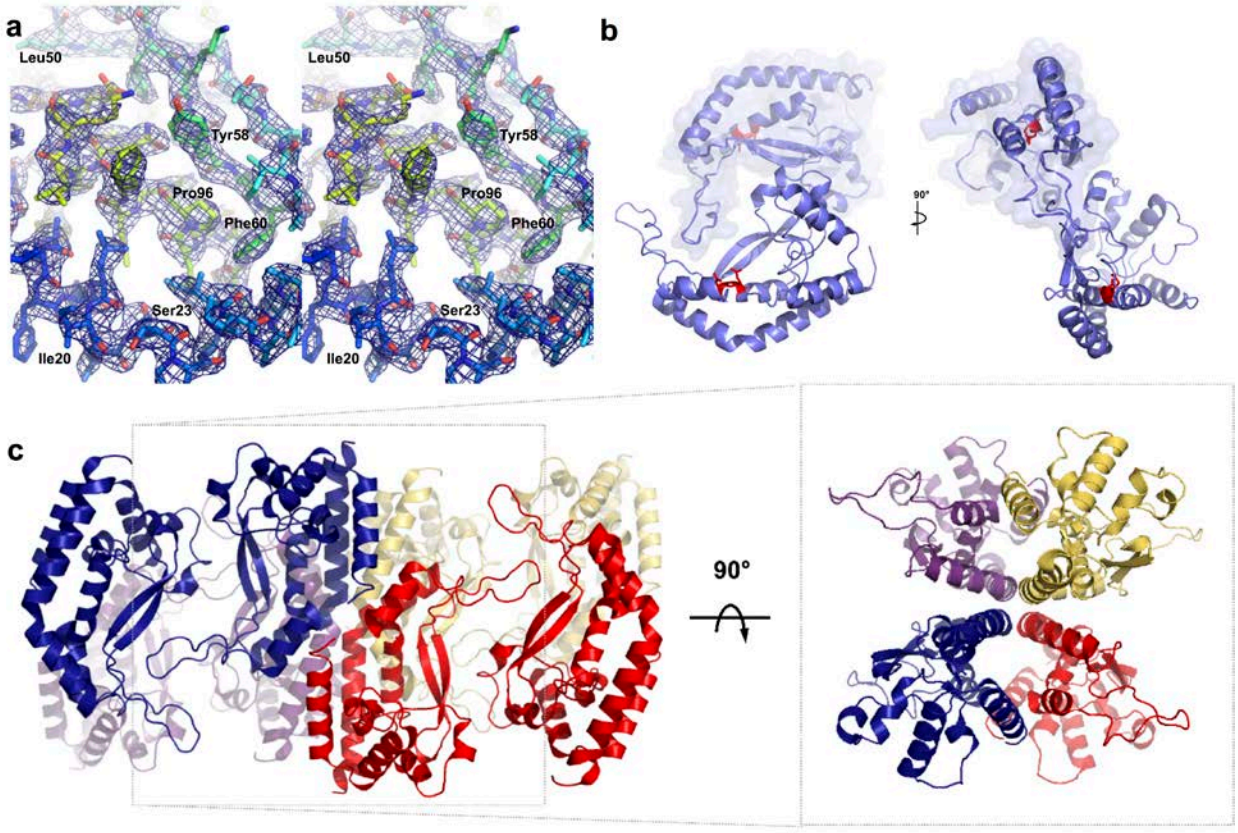
b



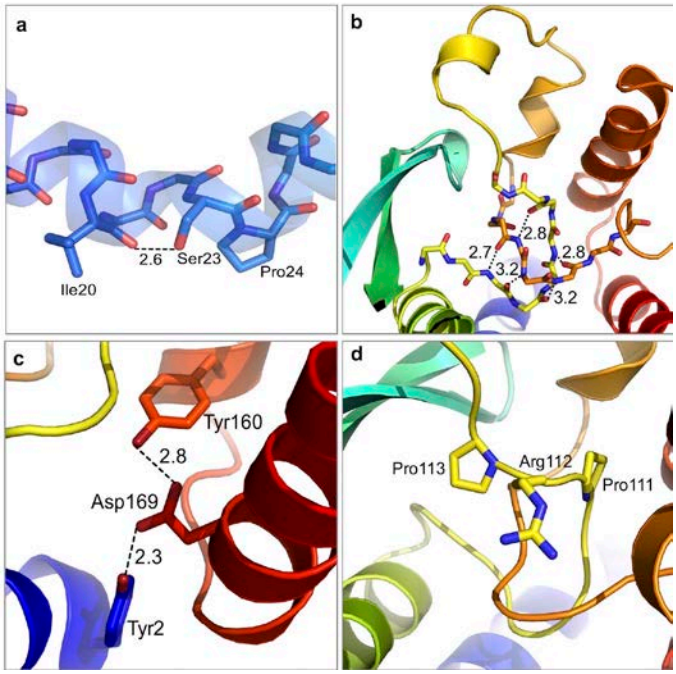
Supplementary Figure 2 | The novel fold of PglC. a, Stereoview of the Cα trace of PglC colored from N-terminus (blue) to C-terminus (red). **b,** Topology diagram for PglC. The reentrant membrane helix (RMH) is formed by the helix-break-helix motif of helices A and B. Helices C, D, and I are amphipathic.



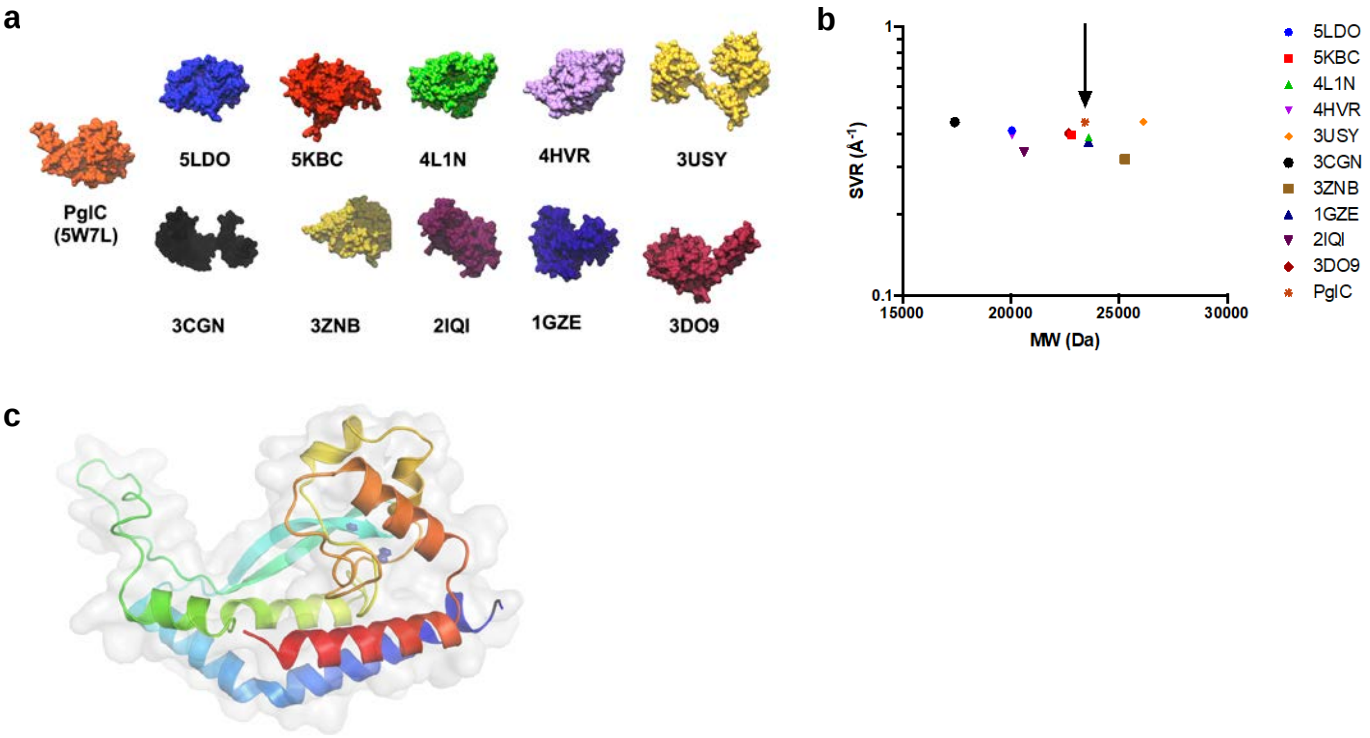
Supplementary Figure 3 | SDS-PAGE analysis of purified PglCs. a, I57M/Q175M His_6-SUMO-PglC was obtained employing Ni-NTA affinity chromatography. The expected molecular weight of the protein is 35.8 kDa. **b,** Tag-less I57M/Q175M PglC was obtained after treating His_6-SUMO-tagged PglC with SUMO-protease followed by purification by Ni-NTA affinity chromatography. The expected molecular weight of the protein is 23.4 kDa. **c,** I57M/I87M His_6-SUMO-PglC was obtained employing Ni-NTA affinity chromatography. The expected molecular weight of the protein is 35.8 kDa. **d,** Tag-less I57M/I87M PglC was obtained after treating His_6-SUMO-tagged PglC with SUMO-protease followed by purification by Ni-NTA affinity chromatography. The expected molecular weight of the protein is 23.4 kDa. **e,** Purification of Se-Met His_6-SUMO-PglC was carried out employing Ni-NTA affinity chromatography. The expected molecular weight of the protein is 36.2 kDa. **f,** Tag-less pure Se-Met PglC was obtained by treating Se-Met His_6-SUMO-PglC with SUMO-protease followed by purification by Ni-NTA affinity chromatography. The expected molecular weight of the protein is 23.6 kDa.



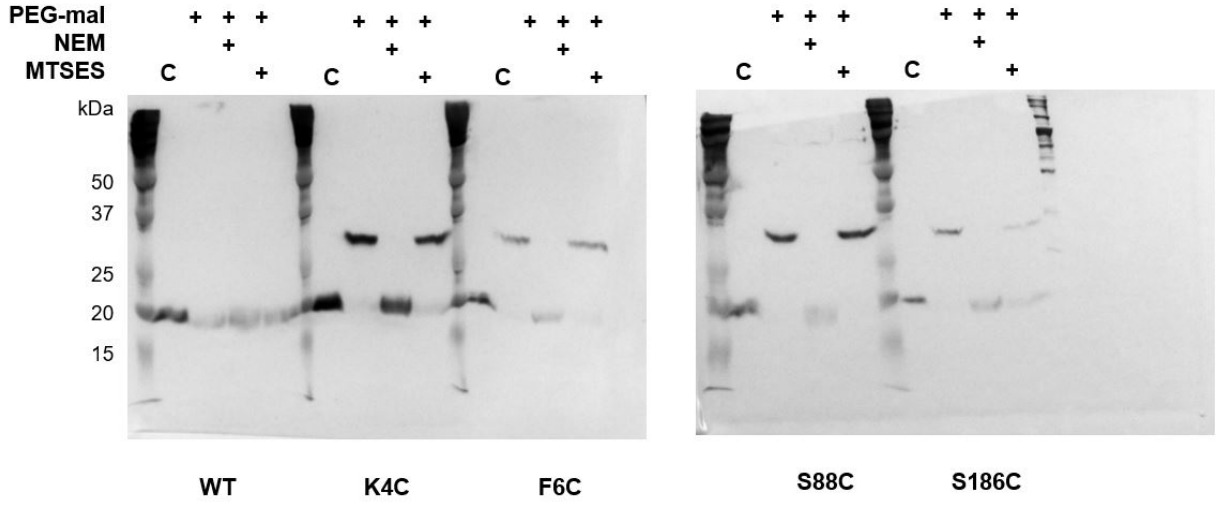
Supplementary Figure 4 | Electron density and crystal packing observed for *C. concisus* I57M/Q175M PglC. a, Stereo-view of the final $2F_O$-F_C electron density map contoured to 2σ in the region of the RMH (blue) and AHABh-motif (cyan-green). **b,** Crystal contacts near the active site do not preclude analysis of the PglC active site. Crystallographic dimer of PglC in the asymmetric unit (ASU) shown in cartoon representation. Molecular surface of Chain B displayed to illustrate solvent excluded volume of the monomer. Catalytic Asp-Glu dyad shown in red sticks. **c,** Crystal contacts are mediated through the RMH neighboring symmetry mates. Crystallographic ASU and symmetry mates are depicted in ribbon representations with each asymmetric unit depicted in one color. Each crystallographic ASU contacts a second crystallographic ASU (right) through contacts of the RMH. The Pro24 residue of each RMH interdigitates with the corresponding Pro24 of the proximal protomer.



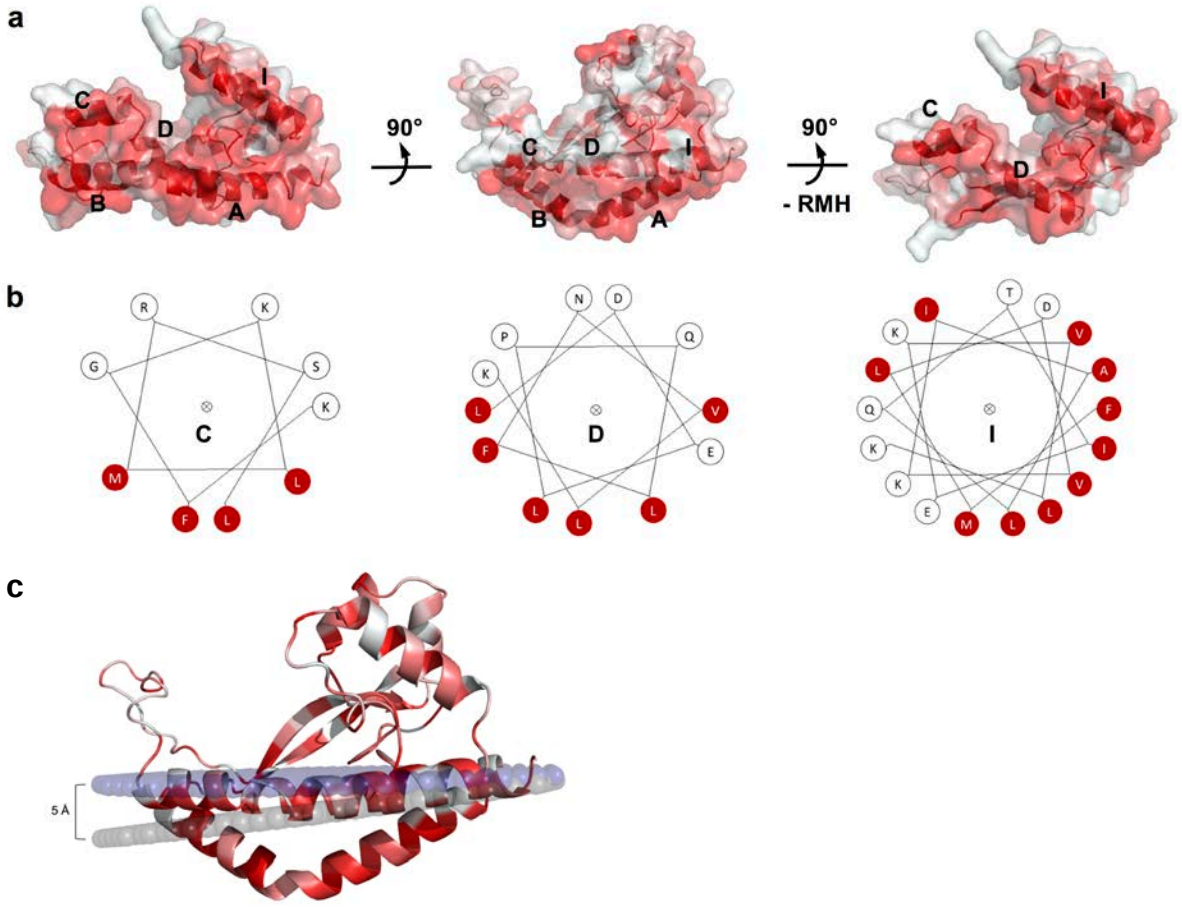
Supplementary Figure 5 | Proline-kinks and hydrogen bonds establish critical interactions in PglC. a, A proline-kink from Pro24 in tandem with a 2.6 Å intra-molecular hydrogen bond between Ser23 and the backbone carbonyl of Ile20 establishes the helix-break-helix motif of the RMH. **b**, An extensive hydrogen-bonding network between backbone amide and carbonyl groups stabilizes the double-twisted loop motif. **c**, A triad of polar residues (Tyr2/Asp169/Tyr160) forms a hydrogen-bonding network that establishes intra-molecular interactions between helices A, F and G. **d**, A strictly-conserved Pro-Arg-Pro (111-113) orients Arg112 towards the active site Asp-Glu dyad, which would potentially position the Arg112 side chain for interaction with the uracil nucleobase of the UDP-sugar.



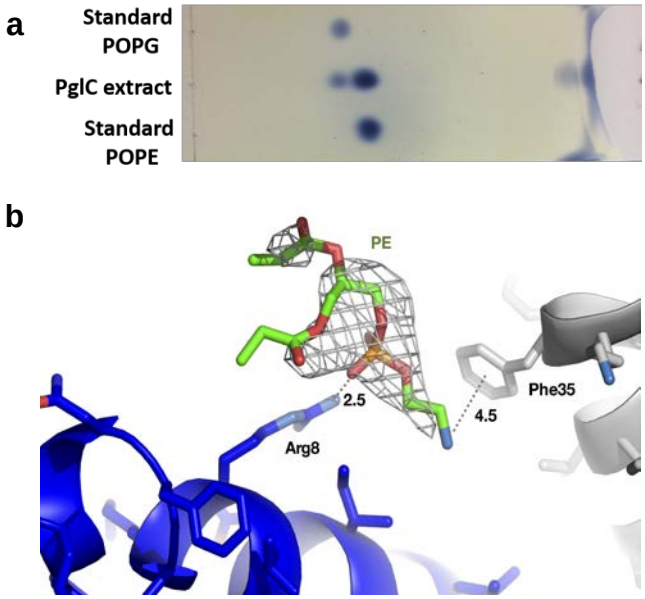
Supplementary Figure 6 | Molecular packing of PglC. a, The surface-to-volume ratio (SVR) (SVR = SAS/SEV) of PglC is similar to those of a test set of proteins of comparable molecular weight from the PDB. Test set of similar structures from the PDB shown in surface representation. **b,** SVR for all test-set structures, including PglC, remain similar across the MW range of the test set. **c,** Solvent-excluded cavities detected in the double-twisted loop domain of PglC. Cavities detected by Voidoo shown in dark blue with volumes of 0.554 ± 0.232 $Å^3$ and 0.754 ± 0.344 $Å^3$, respectively. Solvent-accessible surface of PglC displayed as rendered by PyMol in transparent gray.



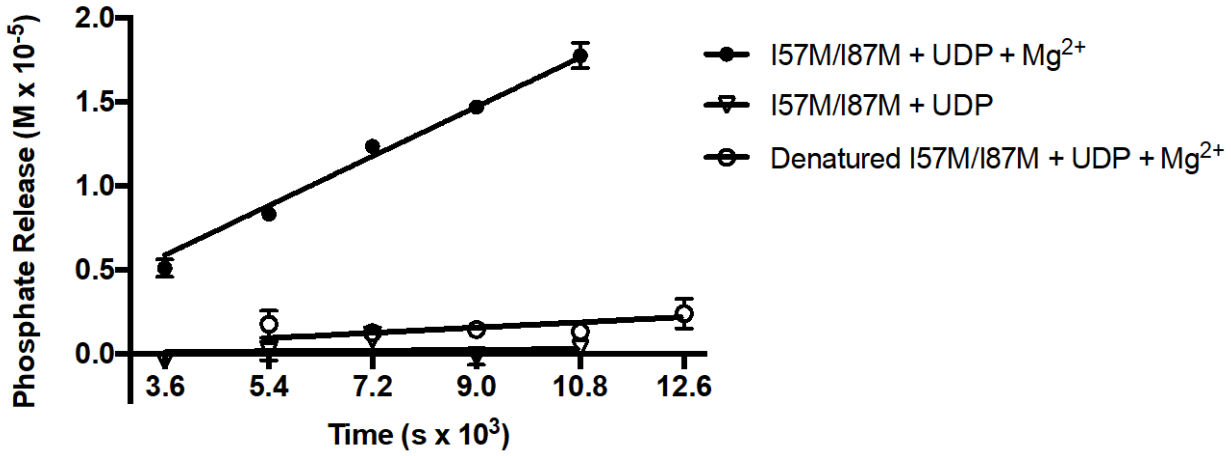
Supplementary Figure 7 | Uncropped images of Western blots shown in Figure 2. *In vivo* SCAM analysis indicates that the N-and C-termini of the RMH are localized at the cytoplasmic face (C = control, no PEG-mal labeling). PEG-maleimide (PEG-mal); N-ethylmaleimide (NEM); 2-sulfonatoethyl methanethiosulfonate (MTSES).



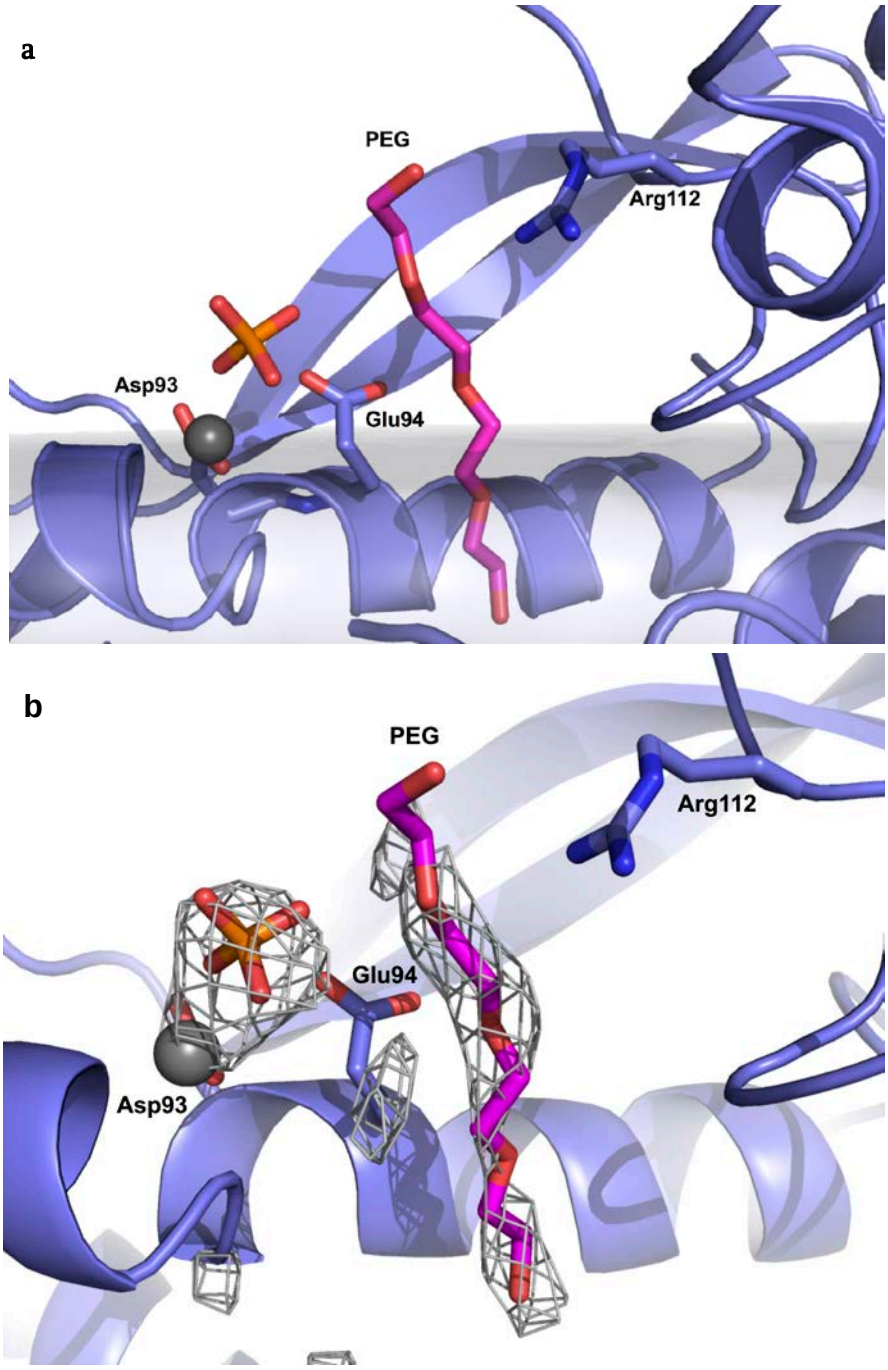
Supplementary Figure 8 | Hydropathy analyses of PglC. a, Hydrophobic surface representations reveal a highly hydrophobic surface of PglC comprising the RMH (helices A, B) and three amphipathic helices (helices C, D, I). Visualization after removal of the RMH confirms the hydrophobic plane formed by the amphipathic helices. **b,** Helical wheel representations for helices C, D, and I illustrate the clear amphipathic nature of the co-planar helices. **c,** Comparison of PglC membrane plane location calculated by the PPM server (gray spheres) and manually-positioned plane (blue spheres) using hydropathy analyses of the RMH and coplanar α-helices. Ribbon diagram of PglC rendered with hydrophobic coloring.



Supplementary Figure 9 | TLC analysis of chloroform extract of PglC and identification of bound lipid head group. a, The TLC analysis demonstrates the presence of both phosphatidylethanolamine and phosphatidylglycerol phospholipids in the purified PglC. A mixture of chloroform:methanol:water (65:25:4) was used as the eluent. Cerium ammonium molybdate (CAM) was used for staining. The CAM stain was prepared by adding 12 g of ammonium molybdate, 0.5 g of ceric ammonium molybdate and 15 mL of concentrated sulfuric acid to 235 mL of distilled water. After staining, the TLC plates were developed by heating. **b,** F_o-F_c simulated annealing omit map contoured to 3$\boldsymbol{\sigma}$ shown as gray wire mesh. PE molecule 303 (chain A) shown in green sticks. RMH of PglC chain A shown in blue. Phe35 from a symmetry mate shown in gray. Interactions between the PE phosphoryl moiety and Arg8 (hydrogen bond) and Phe35 (cation-$\boldsymbol{\pi}$ [Gallivan, J. P. & Dougherty, D. A. Cation-pi interactions in structural biology. *Proc. Natl. Sci. USA*, *96*, 9459-9464 (1999)]) shown with distances in Å.



Supplementary Figure 10 | Time course of phosphate release in the reaction of 20 μM PglC with 200 μM UDP in the presence and absence of 5 mM $MgCl_2$. Assays were carried out in triplicate. Error bars represent mean ± standard deviation. Reaction with I57M/I87M PglC at 20 μM with 200 μM UDP and 5 mM $MgCl_2$ gave a rate of phosphate release from UDP of 1.76 nM s^{-1}. Control phosphate analysis experiments using solutions of PglC, UDP, buffer, water, DDM (0.03%) and $MgCl_2$ (5 mM) alone confirmed that none of these introduced background phosphate and that all of the measured phosphate came from hydrolysis of UDP.



Supplementary Figure 11 | Polyethylene glycol (PEG) position in the PglC binding site identifies the putative Pren-P binding site. a, PEG molecule 305 (chain A; pink sticks), Arg112, Glu94, Asp93 (blue sticks), and phosphate molecule 302 (chain A; orange sticks) are shown with the cofactor Mg^{2+} depicted as a grey sphere. The hypothetical membrane surface is shown as a transparent grey surface. **b,** F_o-F_c simulated annealing omit map contoured to 3$\boldsymbol{\sigma}$ (gray wire mesh) calculated with models of PEG 305 and phosphate 302 removed.

Supplementary Table 1 | Data collection and refinement statistics

	I57M/Q175M Variant	I57M/I87M Variant	WT Se-Met
Data collection			
Space group	P 3_2 2 1	P 3_2 2 1	P 3_1 2 1
Cell dimensions			
a, *b*, *c* (Å)	70.802, 70.802, 188.442	71.61, 71.61, 189.442	143.375, 143.375, 194.004
α, β, γ (°)	90, 90, 120	90, 90, 120	90, 90, 120
Resolution (Å)	62. 81 - 2.74 (2.84-2.74)	94.93 - 2.59 (2.71 - 2.59)	124.17- 3.11 (3.22-3.11)
R_{merge}	0.0987 (1.1)	0.084 (1.708)	0.118 (1.414)
$I / \sigma I$	21.4 (2.5)	11.6 (0.7)	10.5 (0.9)
Completeness (%)	0.99 (1.0)	0.98 (0.97)	0.97 (0.97)
Redundancy	18.0 (11.8)	3.8 (3.8)	5.0 (3.4)
Refinement			
Resolution (Å)	62. 81 - 2.74 (2.84-2.74)		
No. reflections	27735 (1941)		
R_{work} / R_{free}	0.2587/0.2815		
No. atoms			
Protein	3043		
Ligand/ion	82		
Water	20		
B-factors			
Protein	79.49		
Ligand/ion	93.44		
Water	64.34		
R.m.s. deviations			
Bond lengths (Å)	0.003		
Bond angles (°)	0.66		

*1 crystal was used for each structure. *Values in parentheses are for highest-resolution shell.

Supplementary Table 2 | Top 10 matches from DALI search with PglC as query.

Rank	PDB - Chain ID	Z-score	RMSD (Å)	# of aligned residues	# of aa in matched structure	% identity	Molecule Description
1	5j0i-A	4.3	3.7	61	73	8	DESIGNED PROTEIN 2L6HC3_12
2	3csx-B	4.0	3.3	61	71	7	PUTATIVE UNCHARACTERIZED PROT.
3	5j2l-A	3.9	4.4	60	76	2	PROTEIN DESIGN 2L4HC2_11
4	4knh-B	3.8	8.8	90	857	11	NUP192P
5	4knh-A	3.8	8.8	90	914	11	NUP192P
6	3hr0-B	3.8	8.2	74	249	7	COG4
7	4i0x-E	3.8	5.1	61	67	3	ESAT-6-LIKE PROTEIN MAB_3112
8	1jq0-A	3.7	4.4	57	71	2	GP41 ENVELOPE PROTEIN
9	2vs0-B	3.7	6.5	68	84	9	VIRULENCE FACTOR ESXA
10	4i0x-K	3.7	5.6	65	77	3	ESAT-6-LIKE PROTEIN MAB_3112

Supplementary Table 3 | Results of helical geometry analysis for helix D from HELANAL-plus[a].

Partition	Helix Length	Twist (°)	Residues/Turn	Unit Height (Å)
Full-length	13	107.2 ± 6.35	3.36 ± 0.20	1.75 ± 0.28
N-terminal	5	112.8 ± 8.08	3.19 ± 0.23	2.06 ± 0.40
C-terminal	9	104.3 ± 5.14	3.45 ± 0.17	1.58 ± 0.14

[a]Kumar, P. & Bansal, M. HELANAL-Plus: a web server for analysis of helix geometry in protein structures. *J. Biomol. Struct. Dyn.* **30**, 773-783, doi:10.1080/07391102.2012.689705 (2012).

Supplementary Table 4 | Primers used for site-directed mutagenesis of PglC for crystallography.

Primer ID	Primer sequence (altered codon underlined)
I57M_F	5’ CCAGGGCTTAATGAGAAAATTTTTAAAATGTATAAATTTAAGACGATGAGCGATGAGCG 3’
I57M_R	5’ CGCTCATCGCTCATCGTCTTAAATTTATACATTTTAAAAATTTTCTCATTAAGCCCTGG 3’
I87M_F	5’ GGCAAATTTGGCAAACTTATGCGCTCACTTAGCCTCG 3’
I87M_R	5’ CGAGGCTAAGTGAGCGCATAAGTTTGCCAAATTTGCC 3’
Q175M_F	5’ GCTTGATGTAAAGATCGCCTTAATGACAATAGAAAAGGTGCTAAAACG 3’
Q175M_R	5’ CGTTTTAGCACCTTTTCTATTGTCATTAAGGCGATCTTTACATCAAGC 3’

Supplementary Table 5 | Primers used for site-directed mutagenesis of PglC for SCAM analysis.

Primer ID	Primer sequence (altered codon underlined)
PglC_F_NdeI	5’-AAAAAACATATGTATGAAAAAGTTTTTAAAAGAATTTTTG-3’
PglC_R_XhoI	5’-AAAAAACTCGAGGTTCTTGCCATTAAATTTCTCTG-3’
K4C_F	5’-GGAGATATACATATGTATGAA<u>TGC</u>GTTTTTAAAAGAATTTTTG-3’
K4C_R	5’-CAAAAATTCTTTTAAAAAC<u>GCA</u>TTCATACATATGTATATCTCC-3’
F6C_F	5’-CATATGTATGAAAAAGTT<u>TGC</u>AAAAGAATTTTTGATTTTATTTTAGC-3’
F6C_R	5’-GCTAAAATAAAATCAAAAATTCTTTT<u>GCA</u>AACTTTTTCATACATATG-3’
S88C_F	5’-GGAAAAATCGTTAGA<u>TGC</u>TTAAGTTTGGATGAGCTTTTGC-3’
S88C_R	5’-GCAAAAGCTCATCCAAACTTAA<u>GCA</u>TCTAACGATTTTTCC-3’
S186_F	5’-GGTTTTAAAACGAAGTGGGGTA<u>TGC</u>AAAGAAGGCCATGTTAC-3’
S186_R	5’-GTAACATGGCCTTCTTT<u>GCA</u>TACCCCACTTCGTTTTAAAACC-3’