

Molecular basis of lipid-linked oligosaccharide recognition and processing by bacterial oligosaccharyltransferase

Journal Article**Author(s):**

Napiórkowska, Maja; Boilevin, Jérémy; Sovdat, Tina; Darbre, Tamis; Reymond, Jean-Louis; Aebi, Markus; Locher, Kaspar P.

Publication date:

2017-12

Permanent link:

<https://doi.org/10.3929/ethz-b-000230849>

Rights / license:

[In Copyright - Non-Commercial Use Permitted](#)

Originally published in:

Nature Structural & Molecular Biology 24, <https://doi.org/10.1038/nsmb.3491>

Funding acknowledgement:

166672 - Structural and mechanistic studies of components of bacterial protein N-glycosylation pathway and of vitamin B12 transport (SNF)

Molecular basis of lipid-linked oligosaccharide recognition and processing by bacterial oligosaccharyltransferase

Maja Napiórkowska¹, Jérémy Boilevin², Tina Sovdat^{2,†}, Tamis Darbre², Jean-Louis Reymond², Markus Aebi³, Kaspar P. Locher^{1*}

¹ Institute of Molecular Biology and Biophysics, ETH Zurich, Switzerland

² Department of Chemistry and Biochemistry, University of Berne, Switzerland

³ Institute of Microbiology, ETH Zurich, Switzerland

[†] Present address, Rolic Technologies Ltd., Allschwil, Switzerland

*Correspondence and requests should be addressed to K.P.L (email: locher@mol.biol.ethz.ch)

Oligosaccharyltransferase (OST) is a membrane-integral enzyme that catalyzes the transfer of glycans from lipid-linked oligosaccharides (LLOs) onto asparagine side chains, the first step in protein *N*-glycosylation. We report the X-ray structure of a single-subunit OST, the PglB protein from *Campylobacter lari*, trapped in an intermediate state bound to acceptor peptide and a synthetic LLO analog. The structure reveals the role of the external loop EL5, present in all OST enzymes, for substrate recognition. Whereas the N-terminal half binds LLO, the C-terminal half interacts with the peptide. The glycan moiety of LLO must thread under EL5 to access the active site. Reducing EL5 mobility decreases the catalytic rate of OST when full-size heptasaccharide LLO is provided, but not for a monosaccharide-containing LLO analog. Our results define the chemistry of a ternary complex state, assign functional roles to conserved OST motifs, and provide opportunities for glyco-engineering by rational design of PglB.

Introduction

N-linked protein glycosylation is one of most abundant post-translational modifications of proteins that is present in all domains of life¹⁻⁴. It is essential in eukaryotes, facilitating many biological processes such as protein folding, quality control or host-pathogen interactions⁵⁻⁷, and influences virulence in bacteria^{8,9}. The molecular mechanism of the glycan transfer from lipid-linked oligosaccharide (LLO) to the acceptor asparagine located in glycosylation sequons N-X-(S/T) is conserved^{10,11} and catalyzed by the membrane-integral oligosaccharyltransferase (OST). OST is a member of the C family of glycosyltransferases (GT-C), which consists of membrane proteins that are located in the endoplasmic reticulum (ER) or the plasma membrane and that contain 8-13 predicted transmembrane helices^{12,13}. OST of higher eukaryotes is a multi-protein complex with STT3 as a catalytic subunit. In contrast, archaea, kinetoplastids and some proteobacteria contain single-subunit OST enzymes (ssOST) that are homologous to STT3^{7,14-16}. Eukaryotic and bacterial LLOs differ in the length and configuration of the polyprenyl tail and in the nature of the attached glycans. However, all contain a pyrophosphate or phosphate moieties that serve as the leaving groups of the nucleophilic substitution reaction^{1,3,10}.

The ssOST from the Gram-negative, pathogenic bacterium *Campylobacter jejuni* is the PglB protein. It is a well-studied, *bona fide* model system that allows the reaction mechanism of the OST-catalyzed process to be investigated without the complexity of additional subunits present in multi-subunit OSTs. The previously reported structures of PglB from *Campylobacter lari*¹⁷ revealed the fold of STT3 proteins. The same fold was subsequently found in the related AglB protein from *Archaeoglobus fulgidus*^{18,19}. Although structural and biochemical studies have provided insight into the mechanism of sequon recognition^{17,20}, the interactions of OST with LLO and the transfer mechanism remain poorly understood due to the absence of structural insight. Here we describe the crystal structure of PglB bound to synthetic substrate analogs, an acceptor peptide and a non-hydrolyzable LLO. This allowed a ternary complex to be trapped and its structure to be determined at 2.7 Å resolution. By

combining the structural insight with chemo-enzymatic approaches and functional studies, we can define a key intermediate in the reaction mechanism of OST.

Results

In vitro functional assays and inhibitory LLO analogs

The main challenge for trapping a ternary complex intermediate of OST was to ensure that the glycan transfer did not proceed. This could in principle be achieved by generating non-functional PglB mutants or by using substrate analogs that could bind PglB but did not allow glycan transfer. We did not use PglB mutants because for all active site mutations that were investigated in the past, binding of the peptide substrate was also severely reduced. We next considered the use of non-reactive peptide substrates. Arguably, the least-reactive substitution of the acceptor asparagine residue in the sequon is 2,4-diaminobutanoic acid (Dab). Dab-containing peptides were previously found to act as competitive inhibitors both of prokaryotic²¹ and eukaryotic OST^{22–24}. While we had earlier shown that a Dab-containing sequon also inhibited PglB²¹, the absence of the carboxamide group reduced the affinity of the corresponding peptide to PglB approximately 10-fold relative to an Asn-containing sequon. Also, the activation mechanism of the carboxamide group cannot be studied using a Dab-containing peptide. We therefore synthesized various LLO analogs (Fig. 1a) with the goal of identifying a useful compound for structural studies. It was previously shown that water-soluble, monosaccharide-containing LLOs can serve as substrates for PglB²⁵. We followed a similar synthetic approach and generated LLO analogs containing one *N*-acetyl-D-glucosamine (GlcNAc) moiety, a pyrophosphate group, and polyprenyl tails of different lengths in either *cis* or *trans* configuration.

To determine the rates and Michaelis-Menten constants of the synthetic LLOs, we used a fluorescently labeled peptide carrying the glycosylation sequon to monitor the glycosylation reaction by in-gel fluorescence²⁰. Although it was previously suggested that at least 4 isoprenyl units are required for PglB activity²⁵, we found that compounds with as few as two isoprenyl units allowed the glycosylation reaction to proceed, albeit with lower rates (Supplementary Fig. 1). We tested two compounds that contained 3 or 4 isoprenyl units in *trans* configuration, (ω EE)-PP-GlcNAc (or farnesyl-PP-GlcNAc) and (ω EEE)-PP-GlcNAc (or geranylgeranyl-PP-GlcNAc), and one compound containing 4 isoprenyl units in *cis* configuration, (ω ZZZ)-PP-GlcNAc (or nerylneryl-PP-GlcNAc). The latter recapitulates the stereochemistry of the wild type LLO lipid moiety (Supplementary Fig. 2a). Although all synthetic LLO analogs were recognized as substrates by PglB, their turnover rates (Fig. 1b, Table 1) were considerably lower than the 1.55 \pm 0.32 peptide/s (error denotes s.d. of six independent cell cultures, n=6) we observed with full-length, *C. jejuni* heptasaccharide LLO (wild type LLO). Intriguingly, whereas the K_M of the (ω EE)-PP-GlcNAc compound was 6-fold higher than that of the (ω EEE)-PP-GlcNAc compound, the turnover rate was only 2-fold lower (Table 1, Supplementary Fig. 3). This suggests that while longer polyprenyl tails indeed increase the affinity of LLO analogs for PglB, the steps following LLO binding also determine the rate of the reaction. While the stereochemistry of the double bonds of

polyprenyl tail had no impact on LLO affinity (unaltered K_M values), it affected the turnover rates (Fig. 1b, Supplementary Fig. 3, Table 1), indicating an impact on the transfer reaction. Similar specificity was observed for the O-OST PglL from *Neisseria meningitidis*, where the stereochemistry of the lipid carrier strongly influenced the glycosylation reaction²⁶. We identified (ω ZZZ)-PP-GlcNAc as a PglB substrate with a high turnover rate (Fig. 1b) while having sufficiently high solubility, which is required for co-crystallization experiments.

This compound then served as a template for the design of non-reactive LLOs. It has been shown that non-hydrolyzable substrate analogs act as inhibitors rather than donors for glycosyltransferases that use nucleotide-activated²⁷ or lipid-linked glycans²⁸. In part following earlier procedures²⁸⁻³⁰ three non-hydrolyzable LLOs were synthesized, (ω ZZZ)-PPS-GlcNAc, (ω ZZZ)-PPC-GlcNAc and (ω ZZZ)-PPC(OH)-GlcNAc, which contain unreactive thiophosphate, phosphonate, or hydroxy-phosphonate groups, respectively, replacing the first phosphate group connected to the anomeric C1 atom of the reducing-end GlcNAc. The GlcNAc phosphonate²⁹ (Fig. 1c, **1**), GlcNAc hydroxy phosphonate²⁹ (Fig. 1c, **2**) and GlcNAc thiophosphate^{31,32} (Fig. 1c, **3**) were individually coupled to nerylneryl phosphate²⁵ (Fig. 1c, **4**) using 1,1'-carbonyldiimidazole (CDI), followed by deacetylation with NH_4OH and purification on basified silica gel²⁸. Apparent IC_{50} of these compounds were determined in PglB-catalyzed glycosylation containing (ω ZZZ)-PP-GlcNAc as the functional LLO (Fig. 1d). The results showed similar inhibitory potencies for the phosphonate- and hydroxy-phosphonate-containing LLO analogs, but a lower potency (higher IC_{50} value) for the thiophosphate-containing analog (Table 2). All IC_{50} values were considerably higher than the K_M value of the corresponding, functionally active (ω ZZZ)-PP-GlcNAc compound, suggesting that the oxygen atom of the natural pyrophosphate moiety directly or indirectly (e.g. due to a different bond angle or bond length³³) contributes to LLO binding.

Structure of the ternary complex and peptide binding pocket

We co-crystallized a sequence-optimized variant of *C. lari* PglB with the hexapeptide DQNATF and the synthetic, non-hydrolyzable LLO analog (ω ZZZ)-PPC-GlcNAc (Supplementary Fig. 2b) and determined its structure at 2.7Å resolution (Table 3, Fig. 2a). The PglB construct contained several amino acid changes introduced with the aim of improving lattice contacts. The construct exhibited an increased thermostability (data not shown), was less prone to aggregation, but its *in vitro* glycosylation activity was indistinguishable from the 1.50 +/- 0.04 peptide/s determined previously for wild type PglB²⁰.

The architecture of PglB is similar to the previously determined, peptide-bound PglB structure¹⁷, where the C-terminal fragment of EL5 (C-EL5, located between transmembrane helices TM9 and TM10a) pins the bound acceptor peptide against the periplasmic domain of PglB. However, we found that the previously disordered, N-terminal segment of the external loop EL5 (N-EL5) now adopts a defined three-dimensional conformation, folding over and significantly contributing to the binding of LLO (Fig. 2a, Supplementary Fig. 4, Supplementary Fig. 5). Notably, N-EL5 forms an α -helix located at the membrane boundary, with the helix axis forming an angle of $\sim 30^\circ$ relative to the membrane plane. A similar helix

was observed in one of the apo-structures of AglB¹⁸, suggesting that N-EL5 has a propensity for forming this secondary structure element. However, this is functionally relevant only upon LLO binding.

High quality electron density was found for both substrates (Fig. 2b, 2c), which are bound on opposite sides of the tunnel connecting the active site of PglB, indicated by the catalytic Mn²⁺ ion, to the peptide-binding pocket. The location of bound peptide in the ternary complex structure is unaltered relative to the peptide-only structure¹⁷, with the side chain of the acceptor asparagine located in a tunnel formed between the C-terminal segment of EL5 (C-EL5) and the core of PglB (Fig. 2b). The carboxamide group of the acceptor Asn is in close proximity to the catalytically essential residues D56 and E319 as well as the catalytic Mn²⁺ ion (Supplementary Fig. 6a). While the exact rotamer of the carboxamide group of the acceptor Asn side chain cannot be defined, we find continuous electron density to the nearby D56 residue (but not to E319), suggesting a strong contact in the form of a hydrogen bond. No bound water could be identified near the carboxamide group of the acceptor Asn, but three water molecules could be placed near the Mn²⁺ ion. Combined with the contacts to the conserved residues D56, E319 and D154, this leads to a distorted octahedral coordination of the metal (Supplementary Fig. 6b). Importantly, the nitrogen atom of the acceptor Asn side chain is approximately 6 Å apart from the C1 atom of the reducing-end GlcNAc of the LLO analog (Fig. 2b), suggesting that in order to reach the transition state of the reaction, further closing of the distance is required.

Our structure reveals that a short loop containing the DGGK [⁴⁸¹DGGK] motif that is conserved in many bacteria, is present near the active site and interacts both with bound peptide and with the LLO molecule (Fig 2b, 2c). The backbone carbonyl of G482 forms a hydrogen bond with the amide nitrogen of the acceptor asparagine (2.9Å distance), but not with the GlcNAc moiety of the LLO. The closest contact between the DGGK motif and the GlcNAc moiety of the LLO is formed by two backbone atoms (N and C α) of G483, with a distance of 3.8 Å, representing weak van der Waals interactions. The side chains of D481 and K484 point away from the substrates and are engaged in salt bridges with residues from the periplasmic domain and therefore probably have structural roles. The DGGK loop had previously been investigated and found to be essential for the function of *C. jejuni* PglB^{34,35}. Given the proximity of this loop to the reducing-end GlcNAc of the bound LLO analog, it is likely to impose steric constraints during the transfer reaction.

LLO binding site and interactions with PglB

The density for the synthetic (ω ZZZ)-PPC-GlcNAc molecule was of excellent quality, allowing unambiguous building of its structure (Fig. 2c). The polyprenyl tail reaches approximately halfway across the membrane (Fig. 3a) and has largely defined conformation. It is tightly embedded in a hydrophobic groove of the transmembrane domain of PglB but also forms van der Waals contacts with hydrophobic residues on N-EL5 located ~4 Å away. Our finding that electron density of the entire nerylanyl tail was resolved suggests that at least 4 isoprenyl units contribute to the binding of wild type LLO to PglB.

The pyrophosphate and GlcNAc moieties are bound in the active site, near the Mn^{2+} ion. However, there is no direct contact between the pyrophosphate oxygens and the metal ion, as the shortest distance is $\sim 5\text{\AA}$. Instead, the phosphate connected to GlcNAc forms a salt bridge with R375, a conserved residue whose essential role was previously reported²⁰. The second phosphate group (connected to the polyprenyl tail) forms multiple hydrogen bonds, one of which with Y293, a residue from N-EL5 that was previously identified as essential. Intriguingly, the contacts between PglB and the pyrophosphate and GlcNAc moieties of the LLO are all of polar or electrostatic nature. The above interactions can explain why the previously identified R375 and Y293 residues are essential for the glycosylation activity of PglB, but not for sequon binding. A mutation of R375 to alanine had earlier been reported to abolish turnover of PglB, without affecting acceptor peptide binding²⁰. In contrast, a R375K mutant had revealed only a moderate reduction in activity, in line with the functional importance of the salt bridge provided by this residue. Based on the functional analysis of Y293 mutants, we had earlier proposed that it might form a stacking interaction with the reducing-end sugar³⁶. Our structure disproves this interpretation, clearly showing a hydrogen bond with a pyrophosphate oxygen. While a Y293A mutation had been found to abolish PglB function, a Y293F mutation, despite being unable to form a hydrogen bond with the pyrophosphate moiety, retained some activity. This suggests that in addition to a direct contact with the LLO, the aromatic ring of Y293 has a structural role in stabilizing the conformation of EL5 when LLO is bound.

There are several additional residues interacting with the LLO. These include Y196, S198 and Q289, all of which form hydrogen bonds with the pyrophosphate moiety of LLO (Fig. 3b). In addition, the hydroxyl group of Y468 forms a hydrogen bond to the *N*-acetyl substituent of the reducing-end GlcNAc moiety. Y468 is part of the extended WWD motif and is strictly conserved in OSTs that accept HexNAc reducing-end sugars and had therefore been proposed to provide interactions with the *N*-acetyl group¹⁷. To analyze the relevance of the interactions observed in the structure, we mutated the involved residues and determined turnover rates of PglB both using full-length, wild type LLO and the synthetic (ω ZZZ)-PP-GlcNAc analog (Fig. 3c, 3d). Mutations of Y468 and Y196 cause a strong reduction or complete loss of PglB activity, whereas mutations of Y79, S198 and Q289 did not affect the activity equally strongly. The activity of a Y468A mutant was reduced by 1000-fold using wild type LLO and 600-fold using synthetic LLO. In contrast, the mutation Y468F showed only 50-fold and 2.3-fold reduction in turnover rate when using synthetic or wild type LLO, respectively. This suggests that in addition to the hydrogen bond provided by the hydroxyl group, the aromatic ring of Y468 has a structural role in the active site. There was an intriguing difference in the impact of the mutations on PglB function with full-length, wild type LLO versus the shortened, synthetic analog. The rates observed with the synthetic LLO analog were generally more strongly reduced, indicating that additional interactions exist between PglB and the glycan or undecaprenol moieties in the wild type LLO.

Functional studies with disaccharide-containing LLO analogs

Our structure reveals that both the C3 and C4 atoms of the GlcNAc moiety are exposed to the solvent, and that non-reducing extensions with both α -1,3 and β -1,4 glycosidic linkages would be compatible without steric clashes. To further explore the impact of additional saccharides on LLO binding and specificity, we used synthetic farnesyl-PP-chitobiose (containing a GlcNAc- β -1,4-GlcNAc glycan)²⁸ and chemo-enzymatically generated farnesyl-PP-GlcNAc- α -1,3-GalNAc, which we produced from synthetic farnesyl-PP-GlcNAc using purified PglA protein^{37,38} (Fig. 4a). We found that both disaccharide LLOs could serve as glycan donors for PglB, and the resulting glycopeptides containing two saccharides could indeed be observed by tricine SDS-PAGE (Fig. 4b). Compared to farnesyl-PP-GlcNAc, the chemo-enzymatically extended farnesyl-PP-GlcNAc- α -1,3-GalNAc revealed a 2-fold increase in turnover and a reduction of K_M value by a factor of two (Table 1). Intriguingly, whereas the K_M of the chitobiose-containing analog was only twice that of farnesyl-PP-GlcNAc- α -1,3-GalNAc, the observed turnover rate was reduced \sim 300-fold (Fig. 4c and d, Table 1). This suggests that whereas the α -1,3-linked second saccharide unit contributes little to the initial binding of LLO, it is compatible with the subsequent transfer reaction, whereas the chitobiose-containing LLO does not allow PglB to reach the transition state effectively, which could be due to steric clashes or missing interactions with the enzyme.

Immobilization of the External loop 5 by disulfide cross-linking

The structure reveals that in order to access the active site, the glycan of LLO diffusing to PglB in the lipid bilayer will have to “dive” under the external loop 5, which subsequently becomes ordered. We hypothesized that for efficient LLO binding, full disordering at least of the N-EL5 segment would have to occur. We therefore generated a mutant of PglB that contained two cysteines, one in N-EL5 (V294C) and another in the transmembrane domain of PglB (R371C). Upon addition of an oxidizing agent, a disulfide cross-link could be formed, tethering N-EL5 to the transmembrane domain and restricting the opening for a glycan to thread under EL5 (Fig. 5a). The cross-linking yield was determined to be 80 \pm 2.9% (Supplementary Fig. 7). The *in vitro* activity of this PglB mutant was unaltered compared to the wild type protein both for full-length, wild type LLO and for the synthetic LLO analog neryleryl-PP-GlcNAc (1.42 \pm 0.12 peptide/s and 2.8 \pm 0.7 peptide/min, respectively, error denotes s.d. of three independent cell cultures and protein preparations, n=3). However, upon disulfide cross-linking, a difference was evident. Whereas the cross-link did not affect the activity with the monosaccharide-containing LLO analog, the turnover observed with full-length, wild type LLO was reduced 3.8-fold (Fig. 5b). This suggests that LLO binding, including glycan passage through the loop EL5, was at least partially rate-limiting.

Discussion

Our results provide the first structural insight into a STT3 protein (and to the best of our knowledge any GT-C family member) with both the acceptor peptide and the lipid-linked donor substrates bound. We have captured PglB in an intermediate state, prior to glycan transfer from the LLO to the peptide. To reach its observed binding site, LLO has to be pulled out of the membrane. PglB facilitates this step by providing a hydrophobic groove that binds the polyprenyl chain and by providing several specific contacts to the hydrophilic moiety of the LLO, including a salt bridge from R375 to the pyrophosphate group.

The distance between the atoms that would ultimately form a *N*-glycosidic bond (carboxamide nitrogen of the acceptor Asn and C1 atom of the reducing-end sugar) is $\sim 6\text{\AA}$ in our structure, a distance that is too long for the reaction to proceed. It is possible that the use of a phosphonate LLO analog for our structural studies might have contributed to the observed distance to the acceptor peptide, because phosphonates have a slightly different geometry than pyrophosphates (longer P–C than P–O bonds and slightly different bond angle), and the methylene group cannot form a strong contact with the Mn^{2+} ion. However, these differences are small and cannot explain the gap that needs to be bridged to reach the transition state of the reaction. This requires a conformational change, during which the LLO most likely will have to shift towards the peptide, because the peptide is more firmly embedded in its binding site. LLO shifting would involve a repositioning of the pyrophosphate moiety, whereby the phosphate group currently in contact with the conserved R375 would replace water molecules and interact with the Mn^{2+} ion, in agreement with the essential role of the metal ion for reactivity^{12,13,20,34,39}. The second phosphate would then take the place of the first in forming a salt bridge with R375. Importantly, in our structure, steric clashes would prevent such a repositioning. In particular, the loop containing the DGGK motif blocks a further approaching of the LLO to the acceptor asparagine. The importance of this motif for PglB function was previously recognized: In *C. jejuni* PglB, mutations of the aspartate side chain of the DGGK motif to alanine or glutamate significantly decreased glycosylation activity, although the effect was partially reversed when higher LLO concentrations were used³⁴. Molecular dynamics simulations implicated the motif in correctly orienting the carboxamide group of the acceptor Asn residue⁴⁰. However, beyond the requirement for a conformational change to allow the acceptor Asn and the LLO to approach, the exact role of the motif cannot be defined at present. Our results also cannot clarify the exact mechanism of asparagine carboxamide group activation (the twisted amide hypothesis) further. Nevertheless, our results provide a promising starting point for computational approaches aimed at exploring the molecular events of the transfer step.

Our results also define the role of EL5, a loop that is present in all STT3 proteins and that has a dual function in binding peptide (C-EL5) and LLO (N-EL5). Because enzymatic cleavage of EL5 has been observed to result in only a slight reduction in activity³⁶, the two functions of EL5 are independent of each other and thus modular. In contrast to recent molecular dynamics simulations that proposed that EL5 becomes disordered upon LLO binding⁴¹, our structural data demonstrate that the opposite is correct. Our disulfide cross-linking experiment also demonstrates that the larger the glycan moiety, the slower the diffusion of the donor LLO to

the active site given that the glycan must dive or thread under / into EL5. This has consequences for the mechanism of PglB, as outlined in Fig. 6. In the absence of substrates (state 1), EL5 is flexible, disordered, and disengaged from the OST core. This matches one of the apo-AglB structures¹⁸. The binding of the substrates can in principle occur in either order. If peptide binds first (state 2'), C-EL5 becomes ordered, as demonstrated by the peptide-only structure of PglB¹⁷. If LLO binds first (state 2), N-EL5 might be engaged, but no structure is currently available of this state. Given that eukaryotic OST does not have a longer EL5 than bacterial PglB, but eukaryotic LLO has a larger, more branched glycan, it is tempting to speculate that binding of LLO prior to the binding of peptide could be a more efficient mechanistic route for OST, because when both N-EL5 and C-EL5 are disengaged, the large glycan moiety can more easily access the active site.

Not only does the ternary complex structure of PglB provide a mechanistic understanding of bacterial OSTs, but also gives insight into structural changes that may be required for the function of all OSTs, as the glycosylation mechanism is thought to be broadly conserved in all domains of life. It also provides opportunities for engineering the substrate specificity with respect to the LLO. This may be applied in the development and production of glycoconjugate therapeutics by inducing relaxed substrate specificity of PglB in order to synthesize various new bacterial and eukaryotic peptidoglycans to use as novel vaccines.

Acknowledgements

This research was supported by the Swiss National Science Foundation (SNF 310030B_166672 to K.P.L. and Transglyco Sinergia grant to M.A., J.-L.R. and K.P.L.). We thank the beamline staff at the Swiss Light Source for assistance with data collection, C. Lizak for preparing a cysteine-less construct of PglB and A. Ramírez for assistance with the chemo-enzymatic synthesis of farnesyl-PP-GlcNAc-1,3 α -GalNAc.

Author contributions

M.N. performed the over-expression, purification, disulfide cross-linking, functional characterization, and crystallization of PglB. K.P.L. and M.N. performed X-ray data collection, structure determination, and model building and refinement. J.B., T.S., T.D. and J.-L.R. synthesized LLO analogs. M.N., M.A., and K.P.L. devised experiments and analyzed the data. K.P.L. and M.N. wrote the manuscript with input from all authors.

Competing Financial Interests Statement

The authors declare no competing financial interests.

References

1. Schwarz, F. & Aebi, M. Mechanisms and principles of N-linked protein glycosylation. *Curr. Opin. Struct. Biol.* **21**, 576–582 (2011).
2. Nothaft, H. & Szymanski, C. M. Protein glycosylation in bacteria: sweeter than ever. *Nat. Rev. Microbiol.* **8**, 765–778 (2010).
3. Calo, D., Kaminski, L. & Eichler, J. Protein glycosylation in Archaea: Sweet and extreme. *Glycobiology* **20**, 1065–1076 (2010).
4. Szymanski, C. M., Yao, R., Ewing, C. P., Trust, T. J. & Guerry, P. Evidence for a system of general protein glycosylation in *Campylobacter jejuni*. *Mol. Microbiol.* **32**, 1022–1030 (1999).
5. Helenius, A. & Aebi, M. Roles of N-linked glycans in the endoplasmic reticulum. *Annu. Rev. Biochem.* **73**, 1019–49 (2004).
6. Kornfeld, R. & Kornfeld, S. Assembly of Asparagine-Linked Oligosaccharides. *Annu. Rev. Biochem.* **54**, 631–664 (1985).
7. Cherepanova, N., Shrimal, S. & Gilmore, R. N-linked glycosylation and homeostasis of the endoplasmic reticulum. *Curr. Opin. Cell Biol.* **41**, 57–65 (2016).
8. Nothaft, H. & Szymanski, C. M. Bacterial protein n-glycosylation: New perspectives and applications. *J. Biol. Chem.* **288**, 6912–6920 (2013).
9. Valguarnera, E., Kinsella, R. L. & Feldman, M. F. Sugar and Spice Make Bacteria Not Nice: Protein Glycosylation and Its Influence in Pathogenesis. *J. Mol. Biol.* **428**, 3206–3220 (2016).
10. Wacker, M. *et al.* Substrate specificity of bacterial oligosaccharyltransferase suggests a common transfer mechanism for the bacterial and eukaryotic systems. *Proc. Natl. Acad. Sci. U. S. A.* **103**, 7088–93 (2006).
11. Marshall, R. D. Glycoproteins. *Annu. Rev. Biochem.* **41**, 673–702 (1972).
12. Lairson, L. L., Henrissat, B., Davies, G. J. & Withers, S. G. Glycosyltransferases: Structures, Functions, and Mechanisms. *Annu. Rev. Biochem.* **77**, 521–555 (2008).
13. Liu, J. & Mushegian, A. Three monophyletic superfamilies account for the majority of the known glycosyltransferases. *Protein Sci.* **12**, 1418–31 (2003).
14. Kelleher, D. J. & Gilmore, R. An evolving view of the eukaryotic oligosaccharyltransferase. *Glycobiology* **16**, 47R–62R (2005).
15. Yan, Q. & Lennarz, W. J. Studies on the function of oligosaccharyl transferase subunits: a glycosylatable photoprobe binds to the luminal domain of Ost1p. *Proc. Natl. Acad. Sci. U. S. A.* **99**, 15994–9 (2002).
16. Nasab, F. P., Schulz, B. L., Gamarro, F., Parodi, A. J. & Aebi, M. All in one: *Leishmania* major STT3 proteins substitute for the whole oligosaccharyltransferase complex in *Saccharomyces cerevisiae*. *Mol. Biol. Cell* **19**, 3758–68 (2008).
17. Lizak, C., Gerber, S., Numao, S., Aebi, M. & Locher, K. P. X-ray structure of a

- bacterial oligosaccharyltransferase. *Nature* **474**, 350–355 (2011).
18. Matsumoto, S. *et al.* Crystal structures of an archaeal oligosaccharyltransferase provide insights into the catalytic cycle of N-linked protein glycosylation. *Proc. Natl. Acad. Sci. U. S. A.* **110**, 17868–73 (2013).
 19. Matsumoto, S., Taguchi, Y., Shimada, A., Igura, M. & Kohda, D. Tethering an N-Glycosylation Sequon-Containing Peptide Creates a Catalytically Competent Oligosaccharyltransferase Complex. *Biochemistry* **56**, 602–611 (2017).
 20. Gerber, S. *et al.* Mechanism of bacterial oligosaccharyltransferase: In vitro quantification of sequon binding and catalysis. *J. Biol. Chem.* **288**, 8849–8861 (2013).
 21. Lizak, C. *et al.* Unexpected reactivity and mechanism of carboxamide activation in bacterial N-linked protein glycosylation. *Nat. Commun.* **4**, 2627 (2013).
 22. Bause, E., Breuer, W. & Peters, S. Investigation of the active site of oligosaccharyltransferase from pig liver using synthetic tripeptides as tools. *Biochem. J.* **312**, 979–985 (1995).
 23. Imperiali, B., Shannon, K. L. & Rickert, K. W. Role of peptide conformation in asparagine-linked glycosylation. *J. Am. Chem. Soc.* **114**, 7942–7944 (1992).
 24. Imperiali, B. & Tai, V. W. Chemistry and Biochemistry of Asparagine-Linked Protein Glycosylation. 281–303 (2003).
 25. Liu, F. *et al.* Rationally Designed Short Polyisoprenol-Linked PglB Substrates for Engineered Polypeptide and Protein N-Glycosylation. *J. Am. Chem. Soc.* **136**, 566–569 (2014).
 26. Musumeci, M. A. *et al.* In vitro activity of *Neisseria meningitidis* PglL O-oligosaccharyltransferase with diverse synthetic lipid donors and a UDP-activated sugar. *J. Biol. Chem.* **288**, 10578–87 (2013).
 27. Compain, P. & Martin, O. R. Carbohydrate mimetics-based glycosyltransferase inhibitors. *Bioorg. Med. Chem.* **9**, 3077–3092 (2001).
 28. Ramírez, A. S. *et al.* Characterization of the single-subunit oligosaccharyltransferase STT3A from *Trypanosoma brucei* using synthetic peptides and lipid-linked oligosaccharide analogs. *Glycobiology* **27**, 525–535 (2017).
 29. Hajduch, J. *et al.* A convenient synthesis of the C-1-phosphonate analogue of UDP-GlcNAc and its evaluation as an inhibitor of O-linked GlcNAc transferase (OGT). *Carbohydr. Res.* **343**, 189–95 (2008).
 30. Knapp, S. & Myers, D. S. α -GlcNAc Thioconjugates. *J. Org. Chem.* **66**, 3636–3638 (2001).
 31. Knapp, S. & Ajayi, K. The anomeric Pudovik rearrangement. *Tetrahedron Lett.* **48**, 1945–1949 (2007).
 32. Knapp, S. *et al.* Shortcut to Mycothiol Analogues. *Org. Lett.* **4**, 4337 (2002).
 33. Engel, R. Phosphonates as Analogues of Natural Phosphates. *Chem. Rev.* **77**, 349–367 (1977).

34. Jaffee, M. B. & Imperiali, B. Exploiting topological constraints to reveal buried sequence motifs in the membrane-bound N-linked oligosaccharyl transferases. *Biochemistry* **50**, 7557–67 (2011).
35. Ihssen, J. *et al.* Structural insights from random mutagenesis of *Campylobacter jejuni* oligosaccharyltransferase PglB. *BMC Biotechnol.* **12**, (2012).
36. Lizak, C. *et al.* A catalytically essential motif in external loop 5 of the bacterial oligosaccharyltransferase PglB. *J. Biol. Chem.* **289**, 735–746 (2014).
37. Weerapana, E., Glover, K. J., Chen, M. M. & Imperiali, B. Investigating bacterial N-linked glycosylation: synthesis and glycosyl acceptor activity of the undecaprenyl pyrophosphate-linked bacillosamine. *J. Am. Chem. Soc.* **127**, 13766–7 (2005).
38. Glover, K. J., Weerapana, E. & Imperiali, B. In vitro assembly of the undecaprenylpyrophosphate-linked heptasaccharide for prokaryotic N-linked glycosylation. *Proc. Natl. Acad. Sci. U. S. A.* **102**, 14255–9 (2005).
39. Igura, M. *et al.* Structure-guided identification of a new catalytic motif of oligosaccharyltransferase. *EMBO J.* **27**, 234–43 (2008).
40. Pedebos, C., Arantes, P. R., Giesel, G. M. & Verli, H. *In silico* Investigation of the PglB Active Site Reveals Transient Catalytic States and Octahedral Metal Ion Coordination. *Glycobiology* **25**, 1183–1195 (2015).
41. Sun Lee, H. & Im, W. Transmembrane Motions of PglB Induced by LLO are Coupled with EL5 Loop Conformational Changes Necessary for OST Activity. *Glycobiology* **27**, 734–742 (2017).

Figure 1 | *In vitro* inhibition of PglB with synthetic LLO analogs.

(a) Structure of synthetic lipid-linked monosaccharides. Red “R” denotes the polyprenyl tails used: R₁: farnesyl or (EE ω), R₂: geranylgeranyl or (EEE ω), R₃: nerylneryl or (ZZZ ω). (b) Maximal turnover rates of PglB-catalyzed glycosylation reaction using fluorescently labeled peptide (DQNATF) and synthetic donor LLOs shown in a. The data represent three independent cell cultures (error bars indicate s.d. n \geq 6) (c) Synthesis of inhibitory LLO analogs 5-7 by a coupling reaction between saccharide precursors 1-3 and lipid phosphate 4. (d) Kinetic analysis of inhibitory LLO analogs shown in c. PglB-catalyzed reaction using 17 μ M ω ZZZ-PP-GlcNAc with various concentrations of inhibitory LLO analogs. Deduced IC₅₀ values are given in Table 2. Each point represents the average initial turnover rate from three independent protein preparations. Error bars indicate s.d (n \geq 4). Source data for b and c are available online.

Figure 2 | Structure of the ternary complex.

(a) Ribbon diagram of PglB with transmembrane domain in light blue and periplasmic domain in light green. TM helices flanking EL5 are indicated. N-terminal part of EL5 (N-EL5) is colored in turquoise, C-terminal part of EL5 (C-EL5) in dark purple. Mn²⁺ is shown as a pink sphere. Bound acceptor peptide and inhibitory LLO are shown as orange and yellow sticks, respectively. (b) Close-up view of the active site, with PglB shown as a ribbon diagram and colored as in a. The green mesh depicts an omit Fo-Fc map contoured at $\sigma=3.0$ and only shown around bound acceptor peptide, which is shown as orange sticks, with numbers indicating sequon positions relative to the acceptor Asn (zero). Relevant residues of the WWD and DGGK motifs or contributing to the active site are shown as sticks and labeled. Bound LLO is shown as yellow sticks. EL5 was removed for clarity. The dashed line depicts the distance between the carboxamide nitrogen of the acceptor Asn and the C1 atom of the reducing-end GlcNAc. (c) Close-up view of the active site similar to a, but depicting N-EL5 (colored turquoise) and C-EL5 (colored purple). The green mesh depicts the same omit map as in b but shown around bound LLO. Black spheres represent the C α atoms of the DGGK motif.

Figure 3 | LLO binding site.

(a) Electrostatic surface potential of LLO binding site of PglB, color-coded ranging from blue (most positive) to red (most negative). Residues interacting with LLO are indicated. (b) Close-up view of the LLO binding site, with PglB colored as in Fig. 2a. Acceptor peptide is shown as orange sticks, LLO as yellow sticks, and bound Mn²⁺ as a pink sphere. Residues interacting with bound LLO are shown as sticks and labeled. Black dashed lines indicate hydrogen bond or salt bridge interactions. (c) Normalized initial turnover rates of selected PglB mutants using 50 μ M wild type LLO as a donor substrate. (d) Normalized initial turnover rates of selected PglB mutants using 500 μ M synthetic (ω ZZZ)-PP-GlcNAc as a donor substrate. Data in c and d represents three independent cell cultures (error bars indicate s.d., n \geq 3). Source data for c and d are available online.

Figure 4 | Effect of disaccharide linkage on PglB activity.

(a) Structure of the synthetic or chemo-enzymatically generated disaccharide LLO analogs. (b) Tricine SDS-PAGE analysis of the *in vitro* glycosylation activity using fluorescently labeled peptide and different LLOs. The transferred glycans are shown schematically for each lane: Yellow squares, N-acetyl-galactosamine (GalNAc); blue squares, N-acetyl-glucosamine (uncropped gel image shown in Supplementary Data Set 4). (c) and (d) Kinetic analysis of PglB activity using LLO donors (ω EE)-PP-GlcNAc-1,3 α -GalNAc and (ω EE)-PP-GlcNAc-1,4 β -GlcNAc, respectively. 1 μ M PglB and varying concentrations of disaccharide LLOs were used. The deduced K_M and k_{cat} values are indicated in Table 1. Data in c and d represent technical triplicate (error bars indicate s.d., $n \geq 3$). Source data for c and d are available online.

Figure 5 | Disulfide cross-linking of EL5 to the TM domain of PglB.

(a) Schematic representation of PglB (grey surface) showing the localization of the engineered disulfide cross-link. EL5 is depicted as a dark purple (C-EL5) and turquoise (N-EL5) line and labeled. Helices are shown as cylinders, TM helices are labeled. Residues V294 (EL5) and R371 (TM11) were mutated to cysteines and are depicted as yellow asterisks. (b) Glycosylation activity of the PglB V194C/R371C double mutant using 50 μ M wild type LLO (left panel) and 500 μ M synthetic (ω ZZZ)-PP-GlcNAc (right panel). Bars labeled “ox” (oxidized) depict turnover rates of disulfide cross-linked samples, whereas bars labeled “red” (reduced) depict turnover rates of samples containing 10 mM DTT, breaking the disulfide bond. Turnover rates of the disulfide cross-linked samples were corrected taking cross-linking efficiency into account. Data in b represent three independent cell cultures (error bars indicate s.d., $n=3$). Source data for b are available online.

Figure 6 | Glycosylation mechanism.

PglB is shown as a grey-colored molecular surface, with N-EL5 and C-EL5 colored turquoise and purple, respectively. States are indicated with circled numbers. In state 1 no substrates are bound and EL5 is disordered, as indicated by a dashed line. Depending on whether LLO or peptide bind first, states 2 or 2' are reached. Bound LLO is represented by a yellow line for isoprenoid moieties, yellow squares for N-acetyl-galactosamine (GalNAc), and blue circles for glucose (Glc). The pyrophosphate moiety is hidden behind N-EL5. State 2' is shown in a yellow-shaded box surrounded by dashed lines to indicate that no structure exists of this state. The ordering of N-EL5 in state 2' is putative. The observed crystal structure of the ternary complex reflects state 3, with acceptor peptide and LLO bound and the entire EL5 ordered.

Tables

LLO	K_M (μM)	k_{cat} (peptide/min)
(ω EE)-PP-GlcNAc	46.3 +/- 5.8	0.94 +/- 0.22
(ω EEE)-PP-GlcNAc	7.2 +/- 0.7	1.60 +/- 0.12
(ω ZZZ)-PP-GlcNAc	8.5 +/- 0.8	6.63 +/- 1.70
(ω EE)-PP-GlcNAc-1,3 α -GalNAc	20.8 +/- 2.0	1.64 +/- 0.04
(ω EE)-PP-GlcNAc-1,4 β -GlcNAc	46.0 +/- 3.0	0.0053 +/- 0.0001

Table 1 K_M and k_{cat} values of synthetic LLO analogs used for PglB-catalyzed glycosylation, calculated from data shown in Supplementary Fig. 2 and Figs. 4c and 4d. For Michaelis-Menten kinetics of the monosaccharide LLO analogs and disaccharide LLO analogs, 0.1 μM and 1 μM PglB was used respectively. For monosaccharide-containing LLO analogs, errors denote s.d of three independent cell cultures ($n \geq 6$), whereas for disaccharide-containing LLO analogs, errors denote s.d. of three technical replicates ($n \geq 3$).

Inhibitory LLO	IC_{50} (μM)
(ω ZZZ)-PPS-GlcNAc	283 +/- 35
(ω ZZZ)-PPC(OH)-GlcNAc	77.2 +/- 5.8
(ω ZZZ)-PPC-GlcNAc	48.5 +/- 2.8

Table 2 Apparent IC_{50} values of different inhibitory LLO analogs calculated from data shown in Fig. 1d. Errors denote s.d of four independent protein preparations ($n \geq 4$).

Ternary complex¹	
Data collection	
Space group	P2 ₁ 2 ₁ 2 ₁
Cell dimensions	
<i>a</i> , <i>b</i> , <i>c</i> (Å)	83.94, 116.18, 172.02
(°)	90.0, 90.0, 90.0
Resolution (Å)	47.77-2.70 (2.80-2.70)*
R _{merge}	0.087 (1.36)*
I / σ	21.42 (2.18)*
Completeness (%)	100 (100)*
Redundancy	12.8 (13.5)*
Refinement	
Resolution (Å)	47.7-2.7
No. reflections	601398 (62190)*
R _{work} / R _{free}	0.217/0.244
No. non hydrogen atoms	5941
Protein	5816
Peptide	59
LLO	43
Mn ²⁺ ion (I)	1
Mn ²⁺ ion (II)	1
Na ⁺ ion	1
Water	20
B-factors (Å ²)	
Protein	79.2
Peptide	58.3
LLO	77.9
Mn ²⁺ ion (I)	48.8
Mn ²⁺ ion (II)	70.0
Na ⁺ ion	55.2
Water	57.6
R.m.s. deviations	
Bond lengths (Å)	0.009
Bond angles (°)	1.12

Table 3 Data collection and refinement statistics (molecular replacement)

¹ data collected from one crystal

*Values in parentheses are for highest-resolution shell.

METHODS

No statistical methods were used to predetermine sample size. The experiments were not randomized. The investigators were not blinded to allocation during experiments and outcome assessment.

Construction of plasmids

A construct of PglB used in all experiments contained several amino acid changes generated to improve protein crystallization, i.e. lattice contacts. The optimized construct was fully functional, less prone to aggregation and exhibited an increased thermostability compared to wild type PglB (data not shown). Mutations in the *pglB* gene of *C. lari* were generated by the QuikChange method on the pSF2 plasmid¹⁷. The resulting cysteine-less construct with removed glycosylation sites and carrying mutations: K2E, C17A, C30A, A108T, C360L, N535Q, Q536K, K549P, D550N, F553I, N556P, A600P, A602D, T606K, T607Q, V610I, M611T, I619S, F622Y, A624S, V627I, A630N, F663Y and F670Y was referred to as PglB. All subsequent mutations in the LLO binding site were based on this construct.

Over-expression and purification of PglB

Overexpression and purification of PglB and PglB mutants was performed as previously described¹⁷. Shortly, PglB was overexpressed in *E. coli* BL21-Gold (DE3) (Stratagene) cells at 37 °C in 5 l flasks using Terrific Broth medium supplemented with 1% glycerol (w/v). The cells were induced at 37 °C for four hours at A₆₀₀ of 3.0 by adding 0.1% arabinose (w/v). Cells were harvested by centrifugation and cell pellets were stored at -80 °C. All following steps were carried out at 4 °C. Cells were resuspended in 25mM Tris-HCl, pH 8.0, 250mM NaCl and disrupted in a M-110-L microfluidizer (Microfluidics) at 15,000 p.s.i. chamber pressure. Membranes were pelleted by centrifugation at 100,000g for 30 min and solubilized in 25 mM Tris-HCl, pH 8.0, 250 mM NaCl, 10% glycerol (v/v) and 1% *N*-dodecyl- β -D-maltopyranoside (w/v) (DDM, Anatrace) for 1.5h. All subsequent purification buffers contained 0.016% DDM. PglB was purified on a nickel-nitrilotriacetic acid (Ni-NTA) Superflow affinity column (Qiagen) and desalted into 10 mM MES pH 6.5, 100 mM NaCl, 3% glycerol (v/v), 0.016% DDM. For crystallization experiments desalted protein was concentrated to 6 mg mL⁻¹ in an Amicon Ultra-15 concentrator (Milipore) with a molecular mass cutoff of 100 kDa and purified by size exclusion chromatography (Superdex 200 10/300, GE Healthcare) in desalting buffer. Collected fractions were pooled and concentrated to 12 mg mL⁻¹.

Purification of the PglB V294C/R371C mutant

The PglB V294C/R371C mutant was purified as described above with addition of 10mM β -mercaptoethanol to resuspension, solubilization, washing and elution buffers. The mutant was firstly desalted into 25mM Tris-HCl, pH 8.0, 250mM NaCl, 10% glycerol (v/v), 0.016% DDM supplemented with 1mM CuCl₂ and incubated for 30 min at 4 °C, then desalted into

10mM MES pH 6.5, 100mM NaCl, 3% glycerol (v/v), 0.016% DDM. A small fraction of noncross-linked V294C/R371C PglB was used as a control for quantification of the cross-linking efficiency and glycosylation activity. For glycosylation turnover measurements of the reduced V294C/R371C PglB mutant, cross-linked samples were treated with 10mM dithiothreitol (DTT).

Quantification of cross-linking efficiency

The estimation of the cross-linking efficiency was performed by labeling free cysteines with fluoresceine 5-maleimide. The cross-link and control sample were incubated with a 10-fold molar excess of fluoresceine 5-maleimide over cysteines for 10 min at 4 °C. The reaction was quenched with 2mM *N*-ethylmaleimide. The same amount (0.5 μ g) of labeled samples was resolved by SDS-PAGE and fluorescent bands were visualized using the Typhoon Trio Plus imager (GE Healthcare). Band intensities were calculated in ImageJ.

Protein crystallization

Ac-DQNATF{p-NO₂}-NH₂, neryleryl-PPC-GlcNAc and MnCl₂ were added to the concentrated protein to the final concentration of 0.75 mM, 2 mM, 1 mM, respectively and incubated for 15 min at 4 °C. PglB was crystallized by vapor diffusion in hanging drops at 20 °C against a reservoir containing 100 mM glycine pH 9.4, 120 mM Li₂SO₄, 29-31% PEG400 at the protein to reservoir ratio 2:1. Crystals usually appeared after 10-12 days and grew to full size within 3-4 weeks. Crystals were cryoprotected by stepwise addition of cryoprotectant PEG400 up to 30% final concentration and directly flash frozen in liquid nitrogen before data collection.

Data collection and structural determination

The X-ray data was collected at the microfocus X06SA beamline at the Swiss Light Source (Villigen). The wavelength for data collection was 0.9998 Å. All data were processed using XDS⁴². There was no indication of any severe diffraction anisotropy. The ternary complex crystallized in P2₁2₁2₁ space group, containing one PglB molecule per asymmetric unit. The structure was solved by molecular replacement using peptide-only PglB (PDB: 3RCE) as a model. The iterative model building and refinement was performed in Coot⁴³ and Phenix⁴⁴. Final refinement statistics and data collection are summarized in Table 3. The average B factors were calculated with Baverage from the CCP4 program suite. Ramachandran analysis: Ramachandran favored 93.23%, Ramachandran allowed 5.78%, Ramachandran outliers 0.99%. Structural figures were drawn using PyMOL⁴⁵.

***C. jejuni* wild type LLO extraction**

Extraction of the *C. jejuni* heptasaccharide wild type LLO (WT LLO) was done as described before⁴⁶. Shortly, LLO was extracted from SCM6 *E.coli* cells carrying a *C. jejuni* *pglB*_{mut}

cluster, containing an inactivated *pglB* gene. LLO was extracted with a mixture of chloroform: methanol: water (10:10:3 v:v:v) followed by rotary evaporation. The LLO was reconstituted in 10mM MES pH 6.5, 100mM NaCl, 1% Triton X-100 (w/v) buffer and its concentration was determined by titrating different amounts of LLO against constant amounts of fluorescently labeled peptide in an *in vitro* glycosylation assay (see below).

***In vitro* glycosylation assay**

In vitro glycosylation assay were performed as described previously²⁰ using a labeled acceptor peptide 5-carboxyfluorescein-GS-DQNATF-NH₂ (10 μ M) as an acceptor substrate and wild type LLO (50 μ M) or neryleryl-PP-GlcNAc (500 μ M) as a donor substrate. Reaction mixtures with wild type LLO and neryleryl-PP-GlcNAc contained 1nM or 100nM PglB, respectively. 1% Triton X-100 (w/v) was added only to wild type LLO reaction mixtures. For turnover rate determination, a total of 6 samples were taken at different time intervals so that the reaction was in the linear range. Data was fitted by linear regression in GraphPad PRISM7.

Determination of the Michaelis- Menten constants

Michaelis-Menten kinetics of the monosaccharide and disaccharide LLO analogs were determined using the *in vitro* glycosylation assay described above. 100 nM PglB was used for assays containing monosaccharide LLO analogs, whereas 1 μ M PglB was used for assays containing disaccharide LLO analogs. The concentrations of LLO analogs varied from 1-1000 μ M. Samples were taken at different time points and initial turnover rates were calculated for each concentration of LLO analog using linear regression. The resulting rates were plotted against LLO concentration and the data was fitted by nonlinear regression using the Michaelis-Menten analysis tool in GraphPad PRISM7.

Determination of the apparent IC₅₀ values

Apparent IC₅₀ values of non-hydrolyzable LLO analogs were determined using the *in vitro* glycosylation assay described above. For the kinetics of inhibitory LLO analogs, reaction mixtures contained 17 μ M neryleryl-PP-GlcNAc and 10-5000 μ M of competitive LLO inhibitors. PglB was added at a final concentration of 100 nM to start the reaction. Samples were taken at different time points and initial turnover rates were calculated using linear regression. The resulting rates were plotted against the concentration of the inhibitory LLO and the data was fitted by nonlinear regression in GraphPad PRISM7.

Synthesis of LLO analogs

Six LLO substrate analogs were synthesized according to previously reported procedures with slight modifications^{25,28,47}. Five of them were composed of a *N*-acetyl-D-glucosamine (GlcNAc) residue linked to polyprenyl pyrophosphates of different chain length

and stereochemistry (*E*-geraniol, *Z*-nerol, *E,E*-farnesol, *E,E,E*-geranylgeraniol and *Z,Z,Z*-nerylnerol) while the last one had a chitobiose moiety coupled to *E,E*-farnesol. The synthetic strategy included three steps: the synthesis of GlcNAc or chitobiose α -phosphate, the preparation of the lipid phosphate precursors and the coupling of both monophosphates to the desired final pyrophosphates. Nerylneryl-PP-GlcNAc ((ωZZZ) -PP-GlcNAc), geranylgeranyl-PP-GlcNAc ((ωEEE) -PP-GlcNAc), farnesyl-PP-GlcNAc ((ωEE) -PP-GlcNAc), neryl-PP-GlcNAc ((ωZ) -PP-GlcNAc), geranyl-PP-GlcNAc ((ωE) -PP-GlcNAc) and farnesyl-PP-chitobiose ((ωEE) -PP-GlcNAc-1,4- β -GlcNAc) were isolated as final compounds. Phosphonate LLO analogs nerylneryl-PPC-GlcNAc ((ωZZZ) -PPC-GlcNAc) and nerylneryl-PPC(OH)-GlcNAc ((ωZZZ) -PPC(OH)-GlcNAc) were obtained using previously described procedures with modifications^{28,29}. The synthesis required 10 and 8 steps respectively from GlcNAc, where the hydroxyl-phosphonate analog was an intermediate in the phosphonate synthesis. Nerylneryl-phosphate (Fig. 1c, 4) (1.00 eq) and CDI (5.00 eq) were dissolved in *N,N*-dimethylformamide (DMF) and stirred for three hours at room temperature under an argon atmosphere. MeOH (10.00 eq) was added and the reaction mixture was stirred for one hour to quench the unreacted CDI. MeOH was removed *in vacuo*. A solution of the corresponding saccharide precursor (Fig. 1c, 1-3) (3.00-4.00 eq) in DMF was added to the reaction mixture which was then stirred for six days. The reaction mixture was concentrated *in vacuo* and the residue was dissolved in MeOH and a large excess of NH₄OH was added. The reaction mixture was stirred at room temperature for twelve hours, then concentrated at the rotavapor and freeze-dried. The residue was purified by flash chromatography to yield the corresponding LLOs (Fig. 1c, 5-7) as colorless lyophilisats. A detailed description of the synthesis of the inhibitory LLO analogs can be found in Supplementary Note. All synthetic compounds but nerylneryl-PPC(OH)-GlcNAc were shown to be pure by ¹H, ¹³C, ³¹P NMR and ESI-HRMS. nerylneryl-PPC(OH)-GlcNAc could not be obtained at a purity of more than 90%.

Chemoenzymatic synthesis of farnesyl-PP-GlcNAc-1,3 α -GalNAc

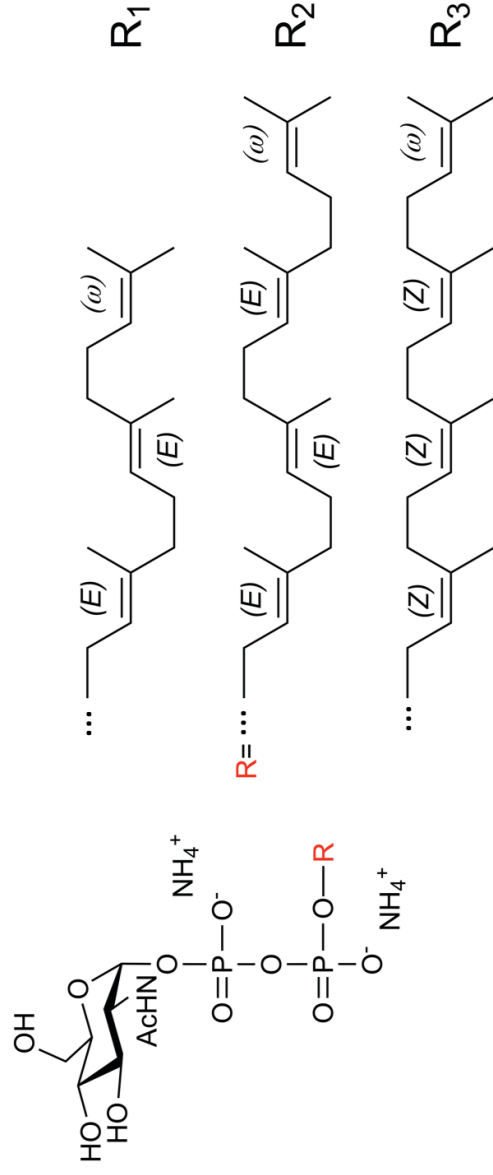
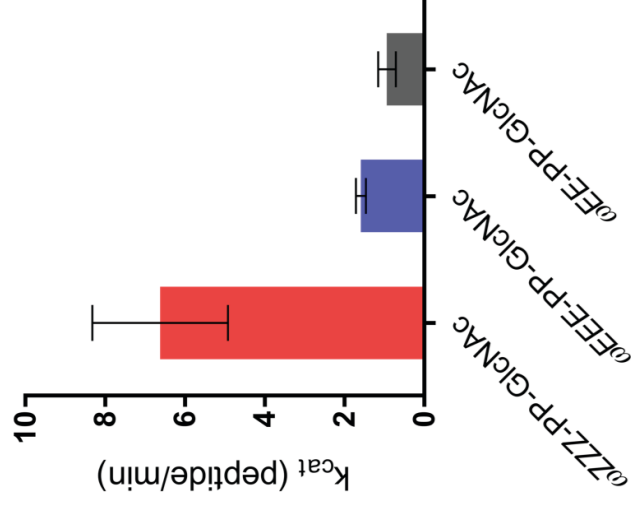
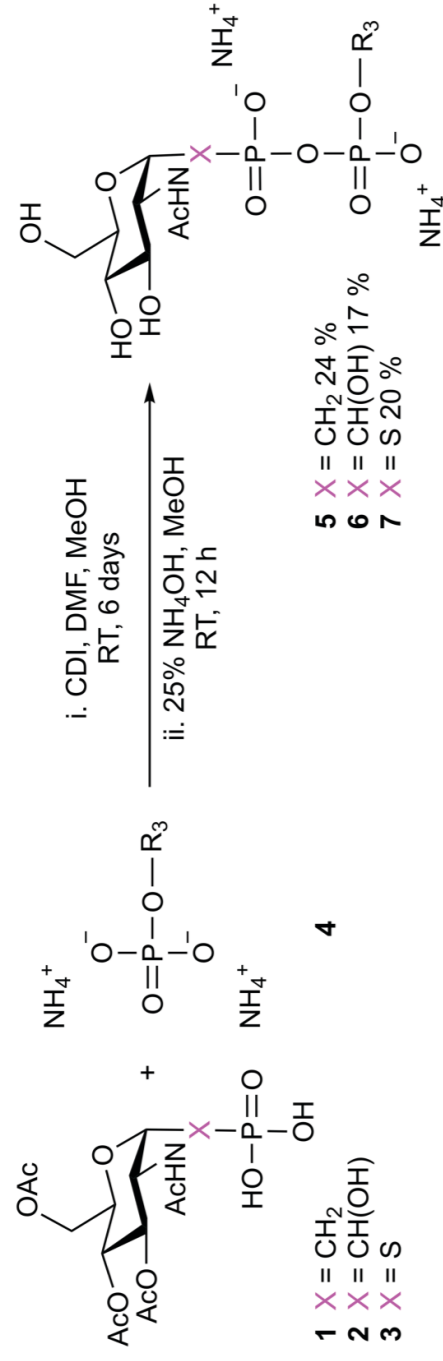
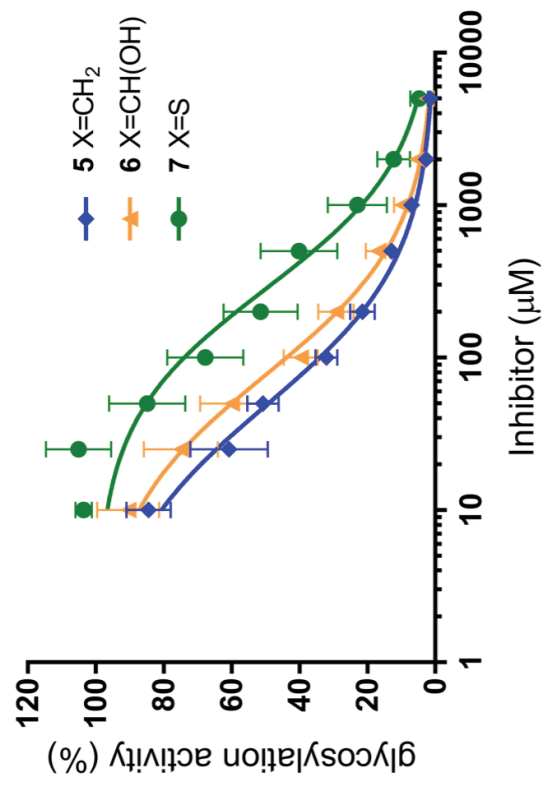
Farnesyl-PP-GlcNAc-1,3 α -GalNAc was obtained by the chemoenzymatic approach previously described, where GalNAc was enzymatically added to synthetic undecaprenyl-PP-GlcNAc by PglA^{37,38}. The obtained product could be verified by Tricine SDS-PAGE analysis.

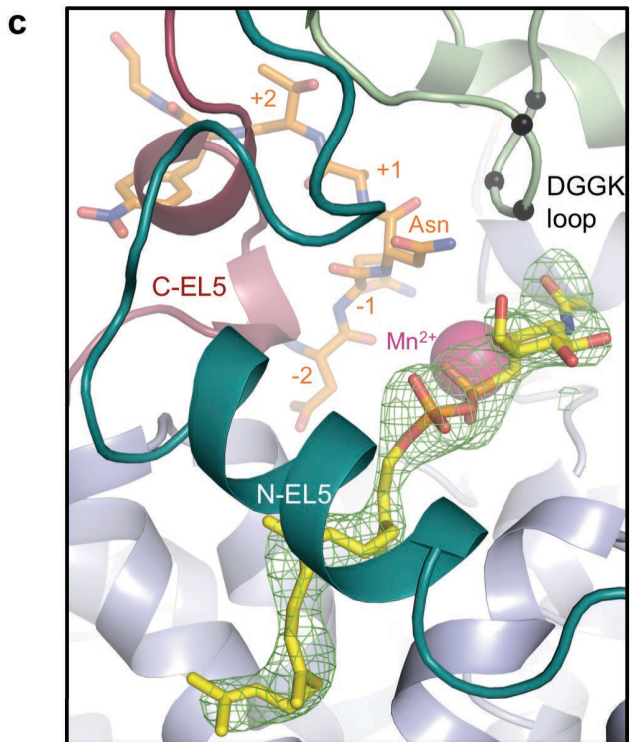
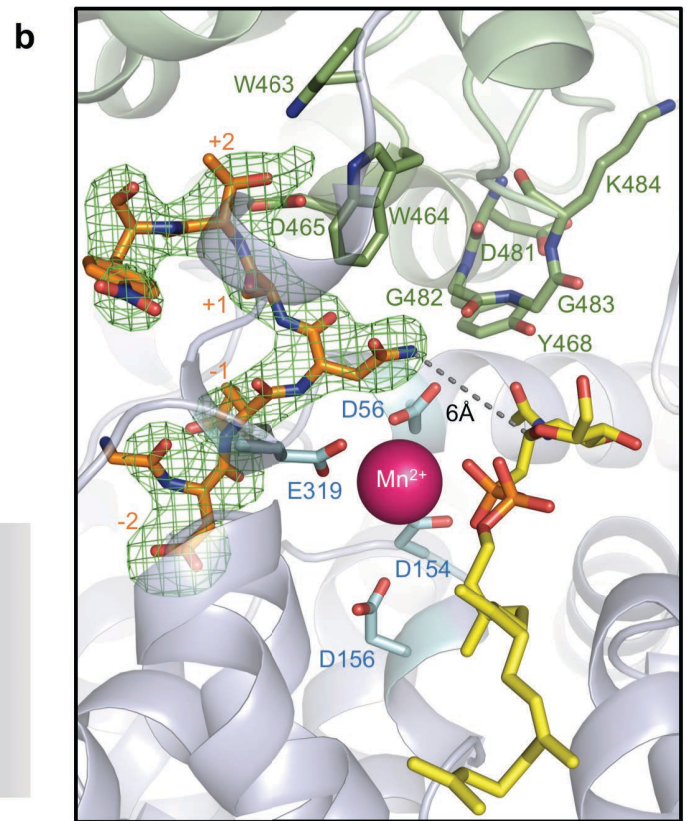
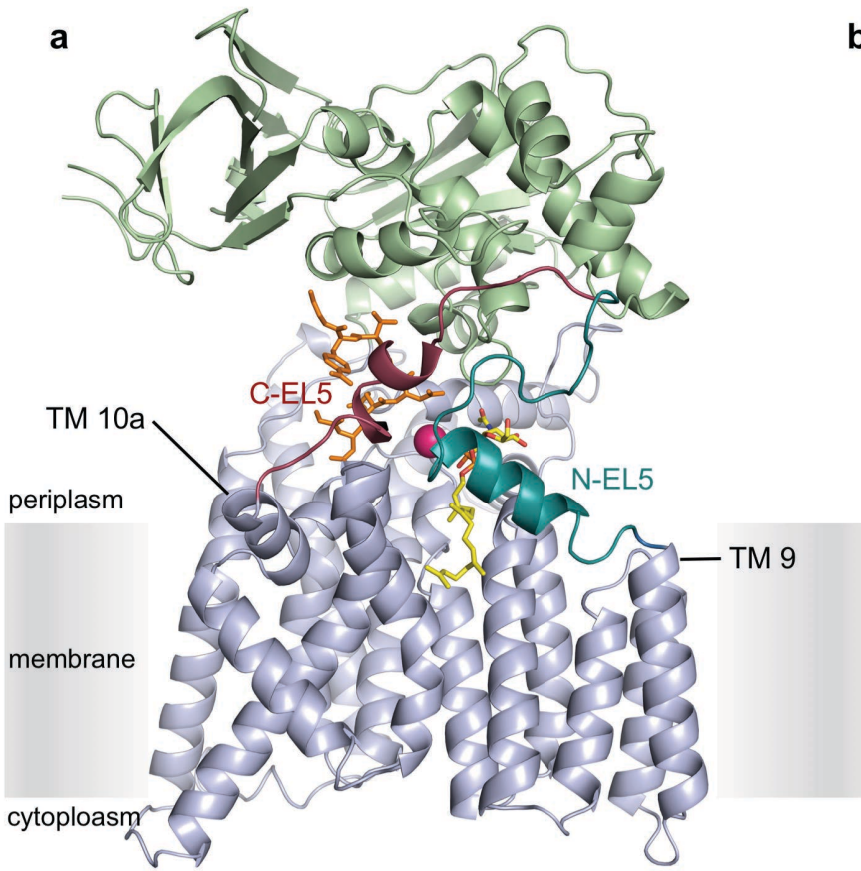
A Life Sciences Reporting Summary for this article is available online.

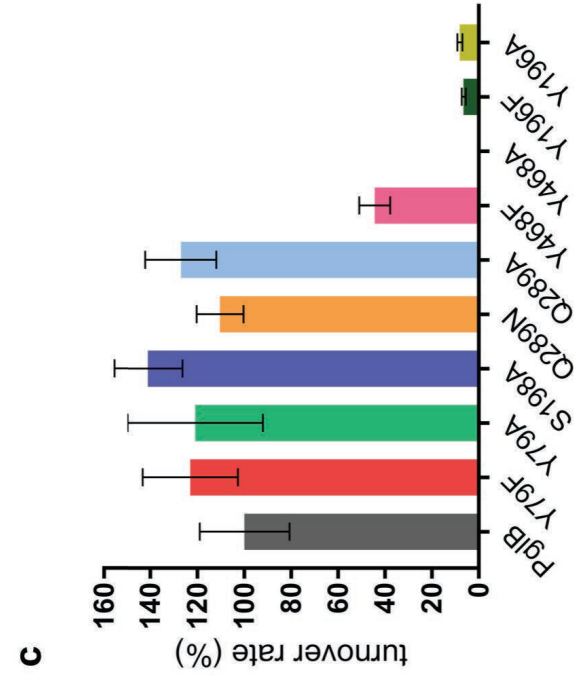
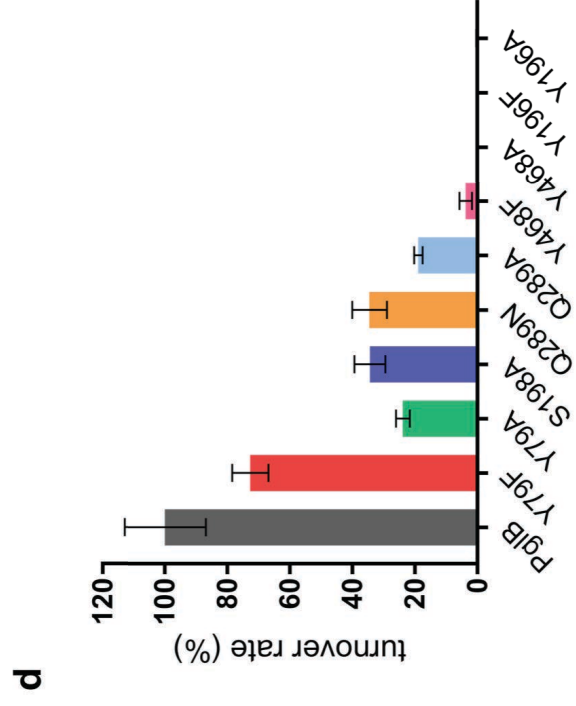
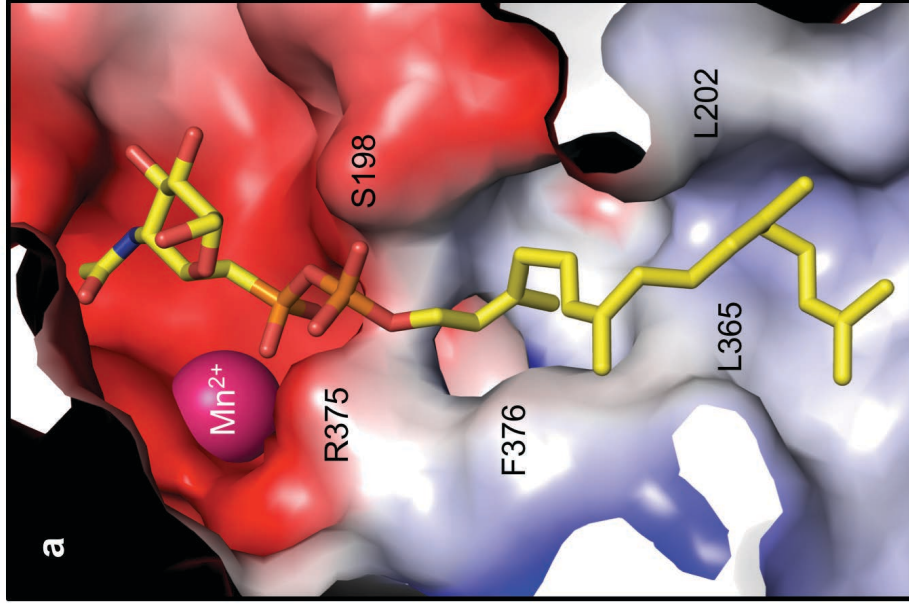
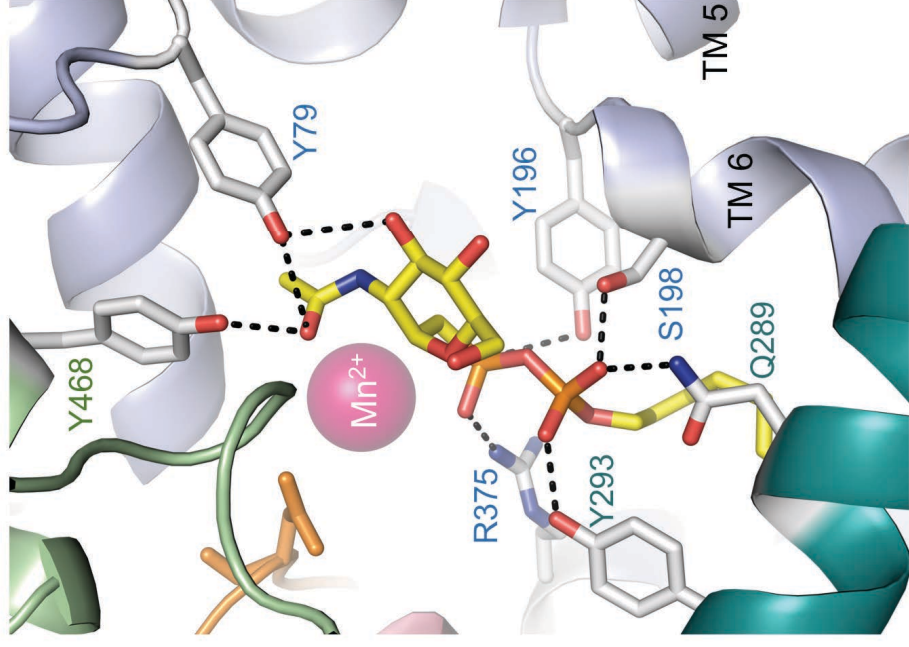
Data availability

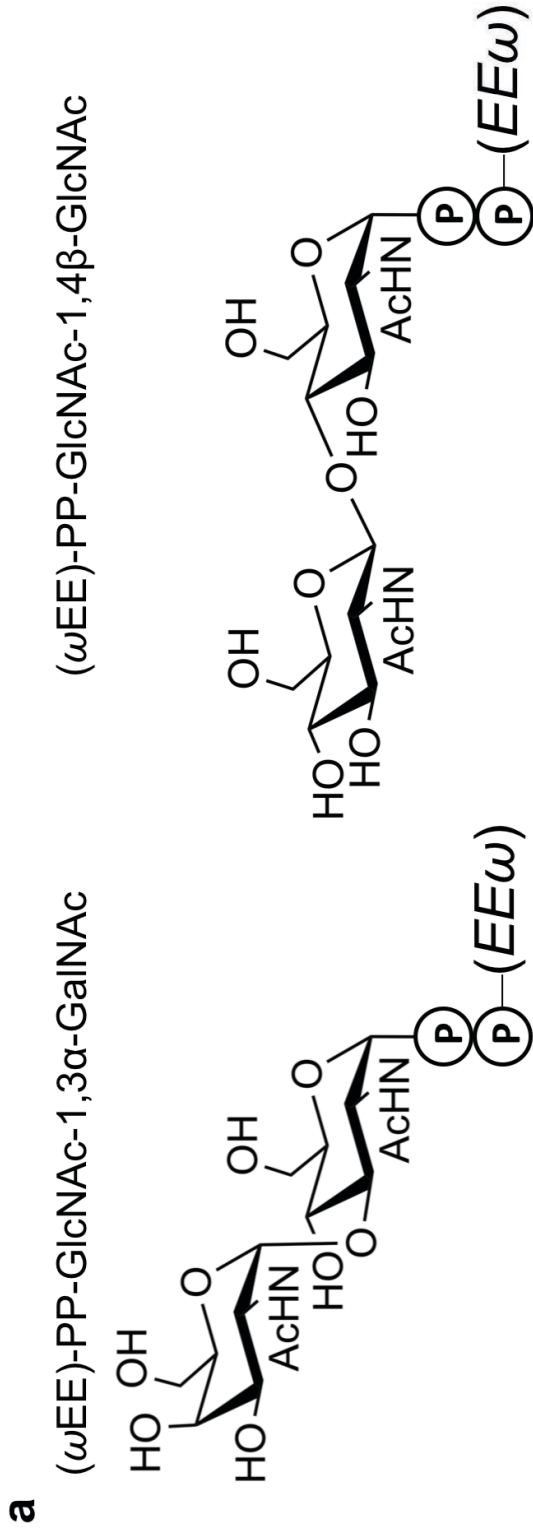
Atomic coordinates of the X-ray structure of PglB with an acceptor peptide and an LLO analog bound were deposited in the Protein Data Bank (PDB) under accession code: 5OGL. All other data are available from the corresponding author upon reasonable request.

42. Kabsch, W. Xds. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **66**, 125–132 (2010).
43. Emsley, P. & Cowtan, K. Coot: model-building tools for molecular graphics. *Acta Cryst* **60**, 2126–2132 (2004).
44. Adams, P. D. *et al.* PHENIX: building new software for automated crystallographic structure determination. *Res. Pap. PHENIX Acta Cryst* (1948).
45. W. L. Delano. The PyMOL Molecular Graphics System. (2002).
46. Kowarik, M. *et al.* N-linked glycosylation of folded proteins by the bacterial oligosaccharyltransferase. *Science* **314**, 1148–50 (2006).
47. Perez, C. *et al.* Structure and mechanism of an active lipid-linked oligosaccharide flippase. *Nature* **524**, 433–438 (2015).

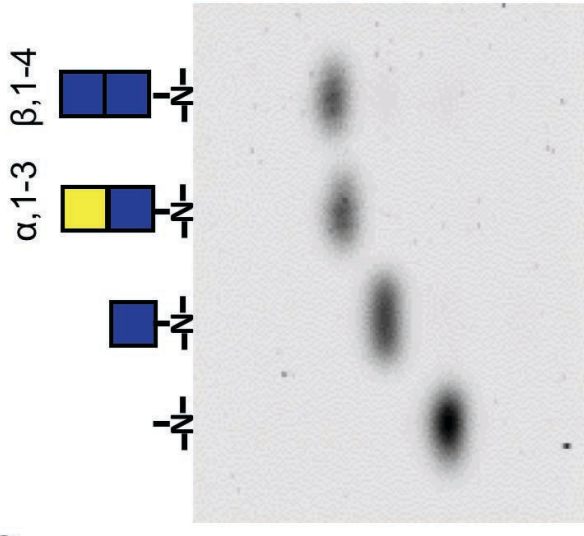
a**b****c****d**



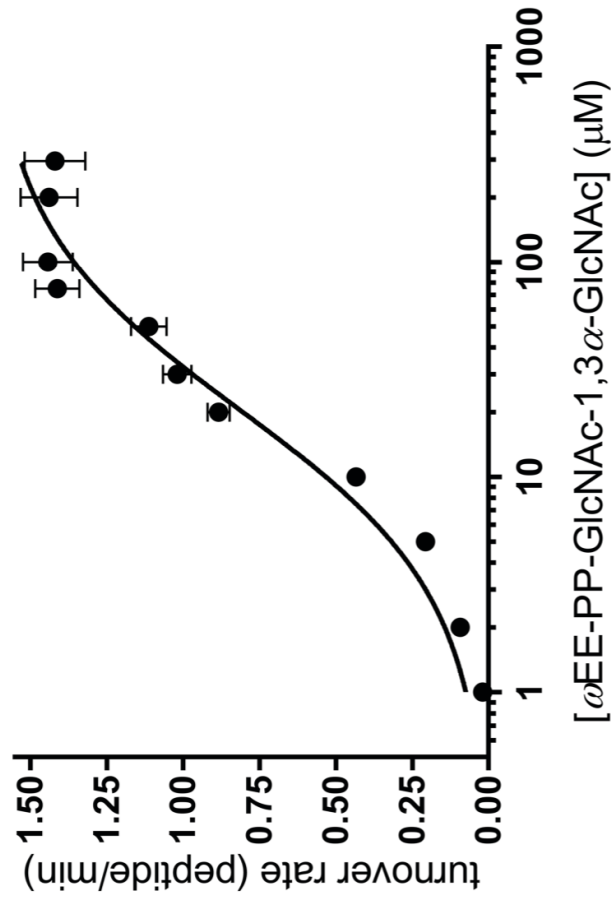




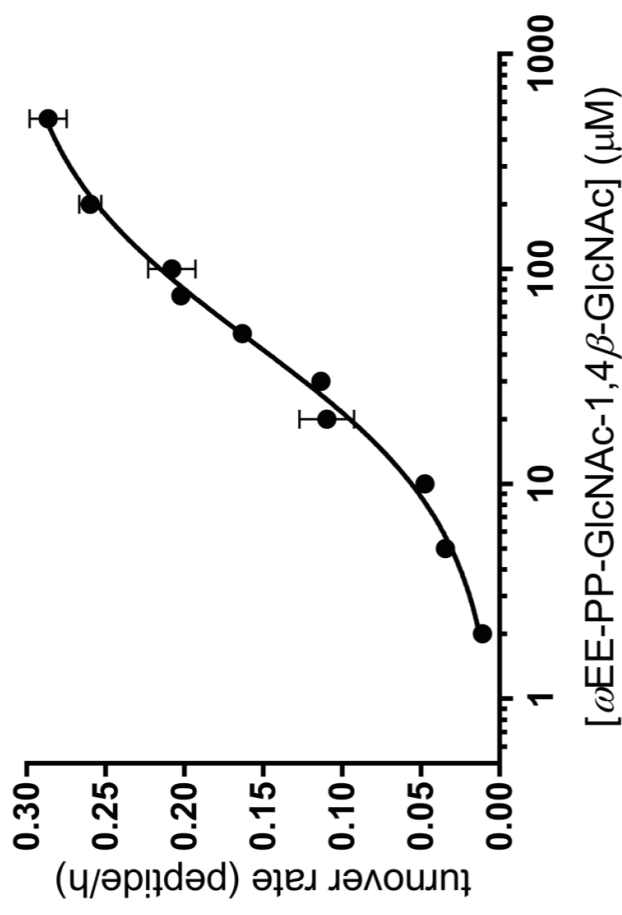
b



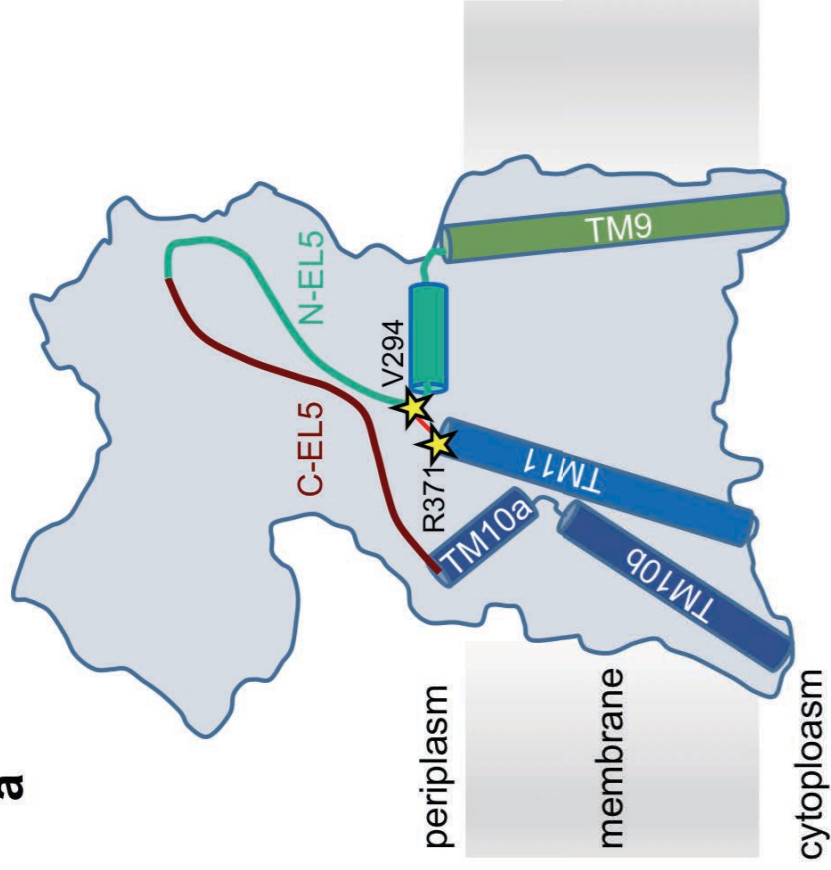
c



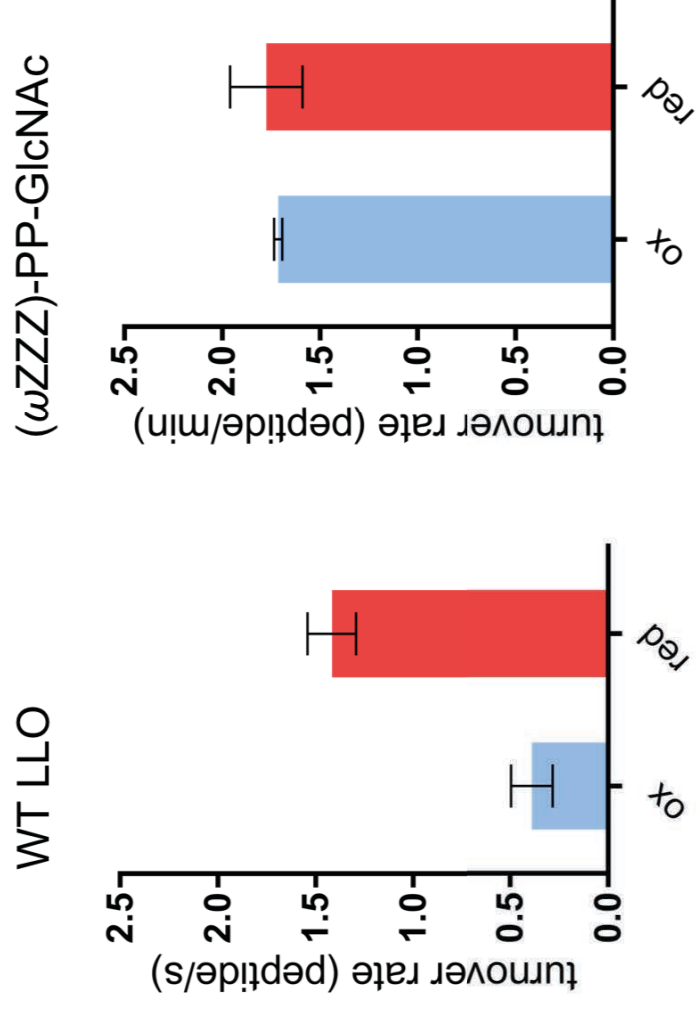
d

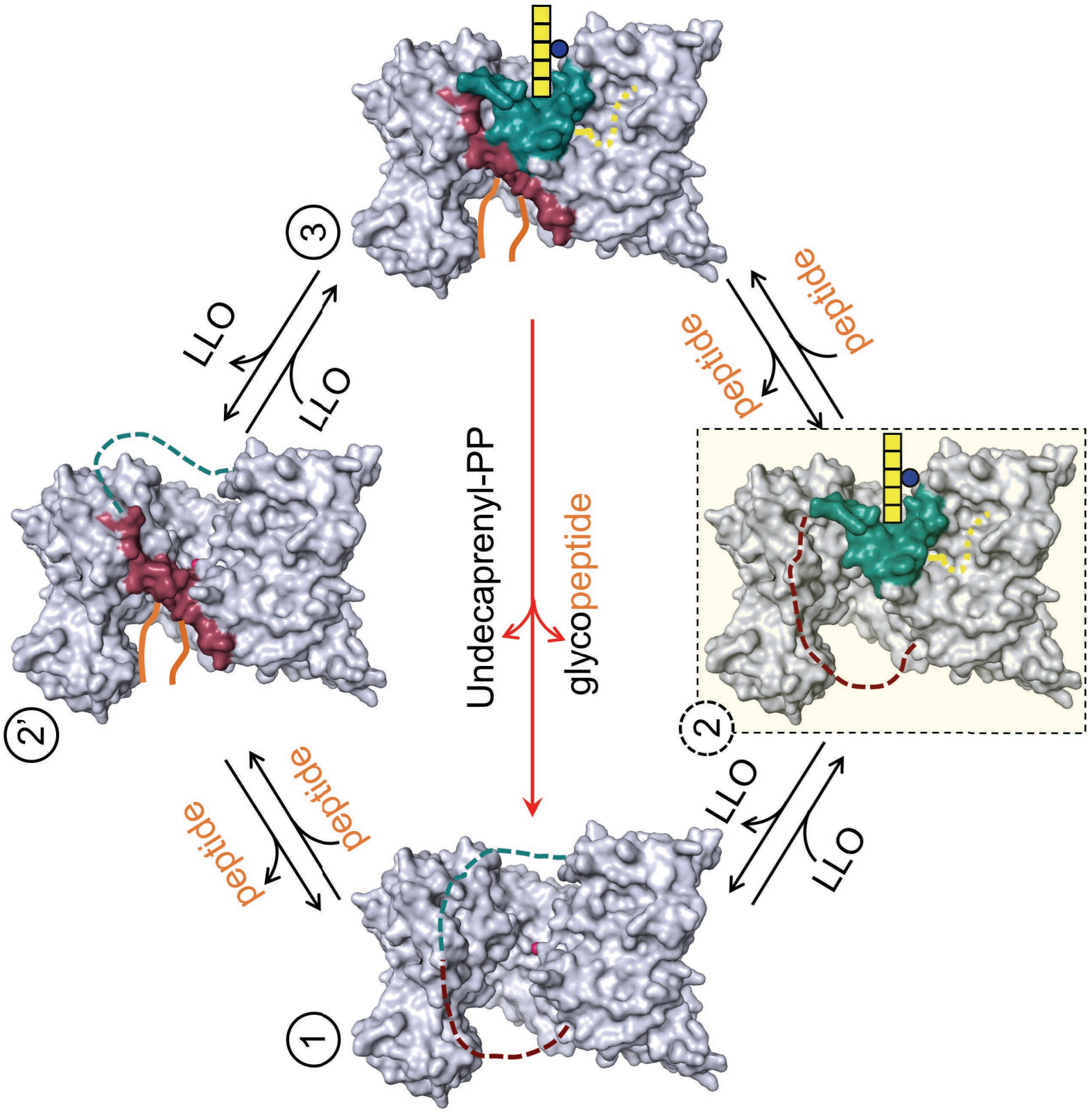


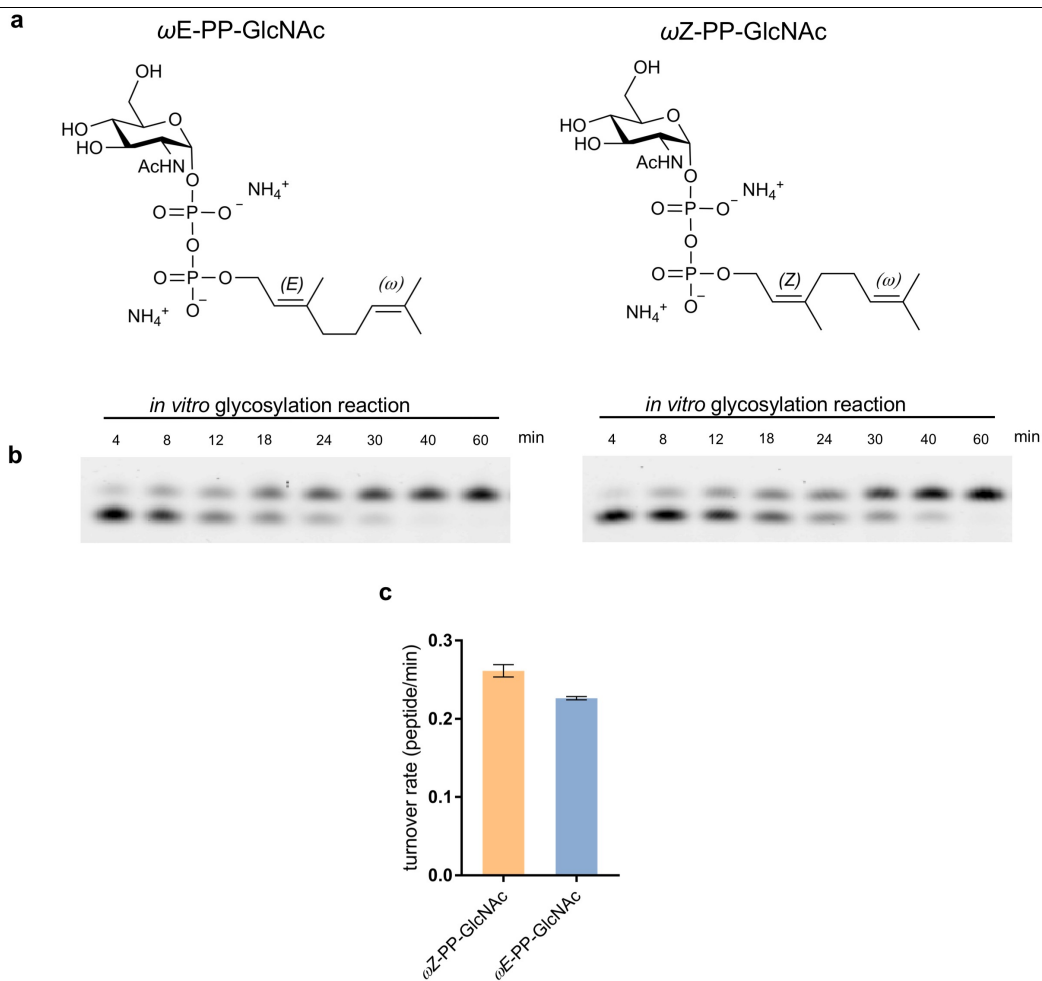
a



b



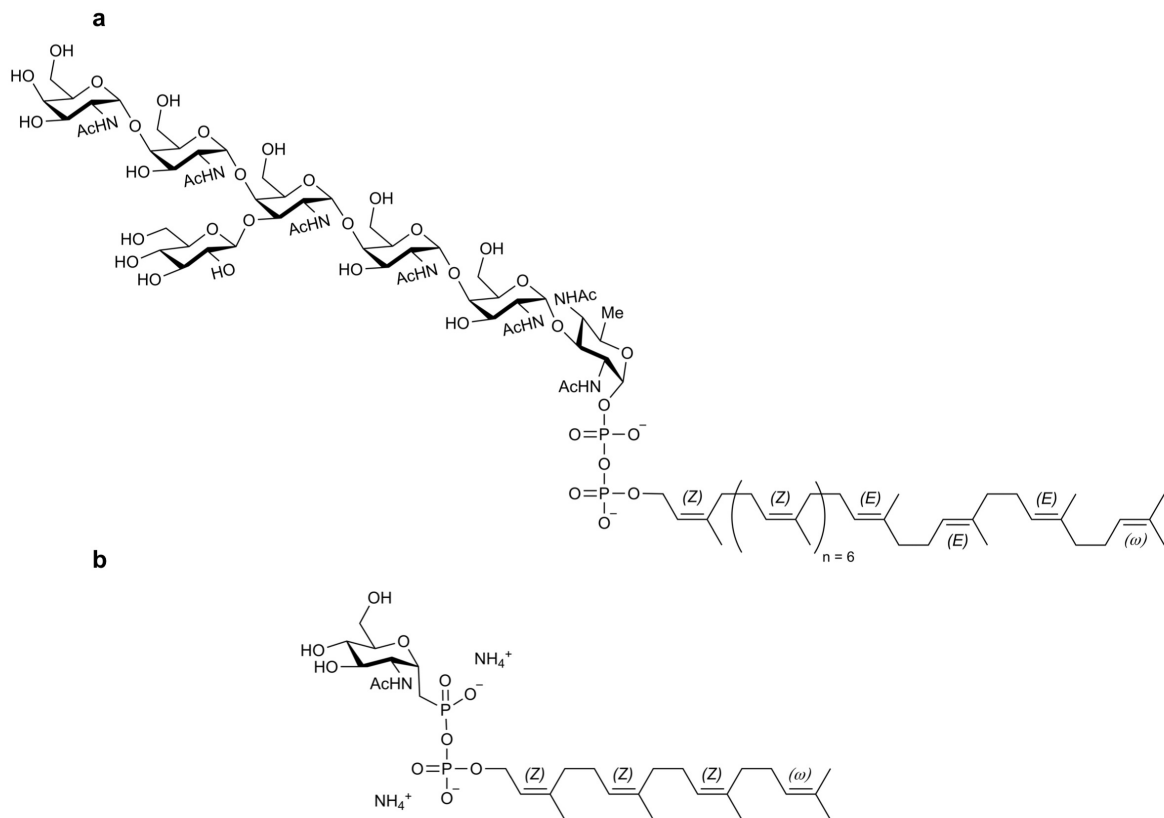




Supplementary Figure 1

In vitro activity of PglB with synthetic LLO analogs.

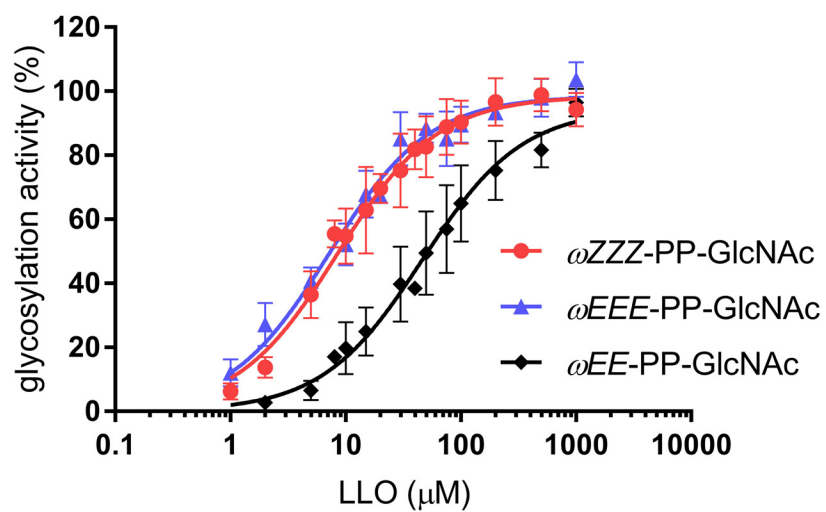
PglB activity was measured *in vitro* using synthetic LLO analogs with distinct polyprenyl tails. **(a)** Structures of synthetic LLO analogs featuring *cis* (ω Z-PP-GlcNAc) or *trans* (ω E-PP-GlcNAc) configuration of double bonds. **(b)** Tricine SDS-PAGE analysis of peptide glycosylation determined by quantification of fluorescently labeled substrate. The assays were performed at 30 °C and contained 1 μ M PglB and 100 μ M synthetic LLO analog. **(c)** Turnover numbers of glycosylation assays shown in **b**. Data represent three independent protein preparations (error bars indicate s.d., n=3).



Supplementary Figure 2

Structures of lipid-linked oligosaccharides.

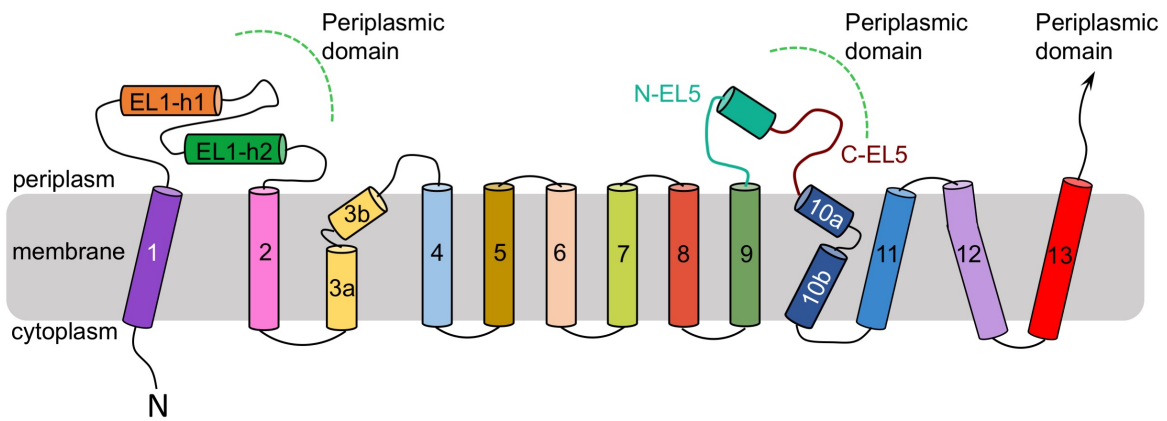
(a) Structure of full-length, wild type *C. jejuni* LLO, abbreviated as undecaprenyl-PP-BacGalNAc₅Glc. (b) Structure of synthetic, inhibitory LLO nerylneryl-PPC-GlcNAc as used for co-crystallization and structure determination of PgIB ternary complex.



Supplementary Figure 3

Kinetic analysis of synthetic LLO substrates.

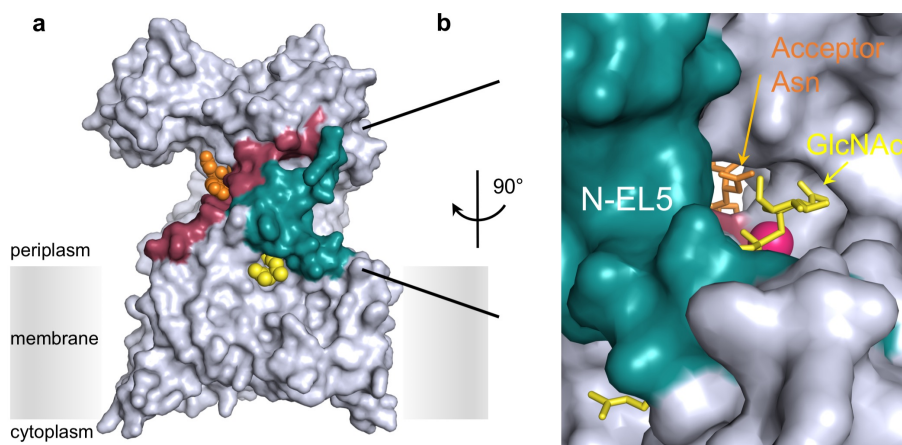
Normalized PgIB activities as a function of varying concentrations of three LLO analogs (structures shown in Fig. 1a). Activities were normalized to the maximal turnover rate for each LLO analog independently to allow the comparison of the Michaelis constants K_M . Each data point represents three independent cell cultures, calculated from the slope of the linear regression fit (error bars indicate s.d., $n \geq 4$)



Supplementary Figure 4

PglB membrane topology.

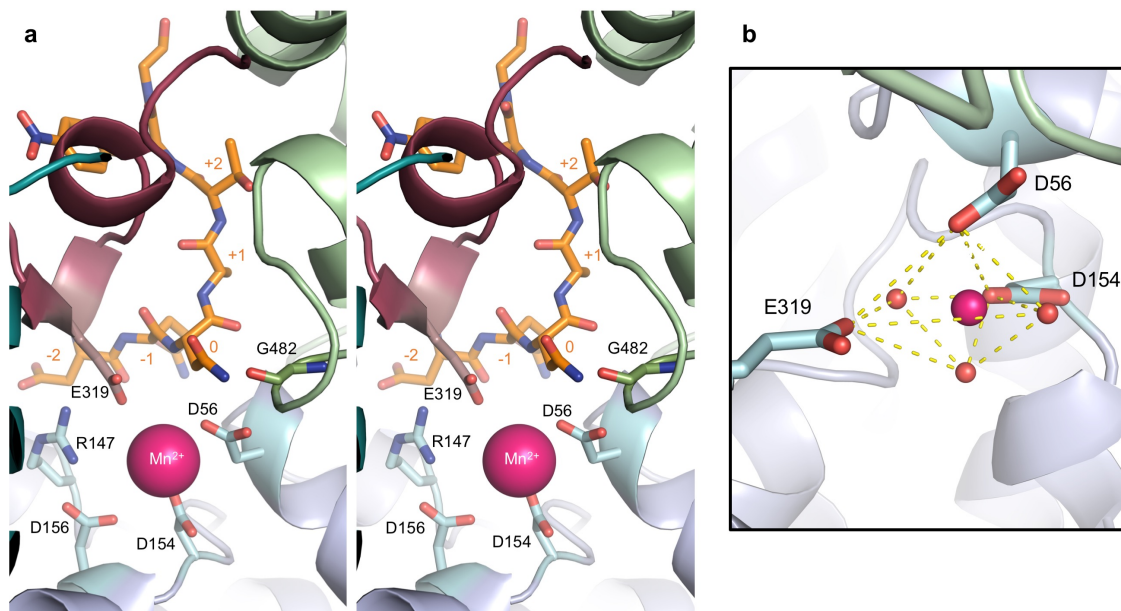
Topological scheme of PglB transmembrane domain, with TM helices depicted as cylinders and numbered. Dashed green lines indicate non-covalent contacts to the periplasmic domain. The N-terminal and C-terminal segments of the external loop EL5 are indicated as N-EL5 and C-EL5, respectively.



Supplementary Figure 5

Surface representation of the ternary complex structure. PglB- acceptor peptide-LLO complex.

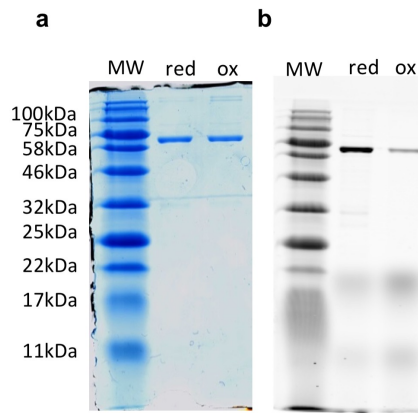
(a) Surface representation of the ternary complex structure of PglB bound to acceptor peptide and synthetic, inhibitory LLO analog. PglB is colored grey, N-EL5 in turquoise, C-EL5 in purple, bound peptide in orange, bound LLO in yellow. (b) Close-up view of a. The acceptor peptide and inhibitory LLO are shown in orange and yellow sticks, respectively. The substrates are bound to cavities at opposite entrances / exits of the PglB tunnel.



Supplementary Figure 6

Catalytic site and acceptor asparagine binding.

(a) Stereo view of the catalytic site and acceptor peptide binding. PglB is colored as in Fig. 2a, selected residues are shown as sticks and labeled. The acceptor peptide is shown as orange sticks, with residues numbered relative to their position in the acceptor sequence (Asn at position zero). The bound Mn^{2+} ion is shown as a pink sphere and labeled. (b) Schematic view of coordination geometry of Mn^{2+} ion, yellow dashed lines depict octahedral shape. The six ligands are the side-chain atoms of D56, D154, E319 and three water molecules, shown as red spheres.



Supplementary Figure 7

Disulfide cross-linking efficiency of the cysteine double mutant.

The ratio of disulfide cross-linking was determined by labeling unreacted, free cysteines with fluorescein maleimide and quantitation of in-gel fluorescence. (a) Coomassie-stained gel; (b) fluorescence-scanned gel. The oxidative disulfide bond formation was determined to be 80 +/- 2.9% from three independent cell cultures (error denotes s.d., n=3). Lanes labeled "ox" (oxidized) indicates a sample that was cross-linked with CuCl_2 during the purification process, whereas "red" (reduced) indicates a control sample that contained 10 mM β -mercaptoethanol during purification, which was removed by desalting. MW denotes marker proteins, with masses indicated on the side.

Click inside this box and **insert a single image for Supplementary Figure 8.**

For best results, use Insert menu to select a saved file; do not paste images. Source images must be JPEGs (no larger than 10 MB) saved in RGB color profile, at a resolution of 150–300 dpi. Optimize panel arrangement to a 2:3 height-to-width ratio; maximum online display is 600h x 900w pixels. Reduce empty space between panels and around image. Keep each image to a single page.

Delete these instructions before inserting the image.

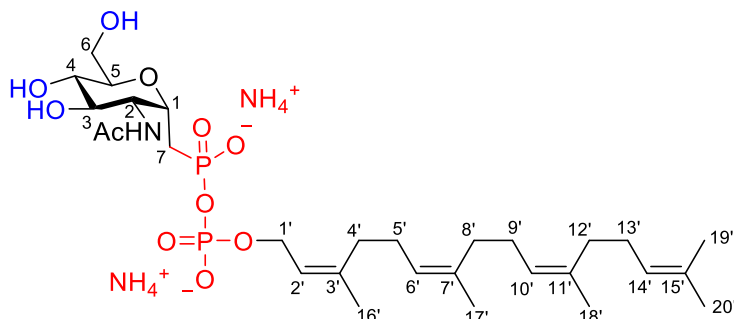
Supplementary Figure 8

Insert figure title here by deleting or overwriting this text; keep title to a single sentence; use Symbol font for symbols and Greek letters.

Insert figure caption here by deleting or overwriting this text; captions may run to a second page if necessary. To ensure accurate appearance in the published version, please use Symbol font for all symbols and Greek letters.

Supplementary Note

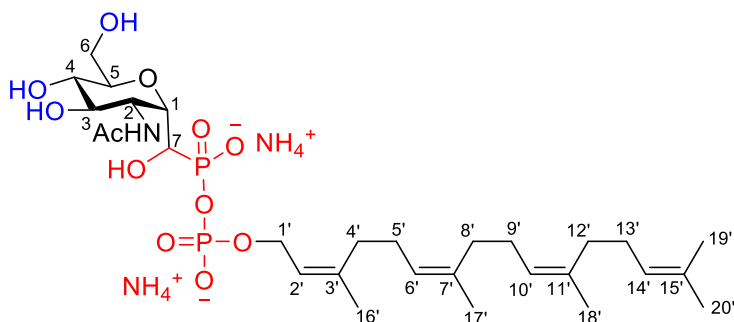
Synthesis of (((2S,3R,4R,5S,6R)-3-acetamido-4,5-dihydroxy-6-(hydroxymethyl)tetrahydro-2H-pyran-2-yl)methyl)phosphonic ((2Z,6Z,10Z)-3,7,11,15-tetramethylhexadeca-2,6,10,14-tetraen-1-yl phosphoric) anhydride (**5**)



4 (25 mg, 0.06 mmol, 1.00 eq) and CDI (50.1 mg, 0.31 mmol, 5.00 eq) were dissolved in 2 ml DMF and stirred for 3 h at RT under an Ar atmosphere. 0.03 ml dry MeOH was added and the reaction mixture was stirred for 1 h to quench the unreacted CDI. MeOH was removed *in vacuo*. **1** (110 mg, 0.26 mmol, 4.18 eq) in 2 ml DMF was

added to the reaction mixture which was then stirred for 6 days. The reaction mixture was concentrated in *vacuo* and the residue was dissolved in 5 ml MeOH and ammonium hydroxide (11.8 ml, 74.20 mmol, 1200 eq) was added. The reaction mixture was stirred at RT overnight, then concentrated at the rotavapor and freeze-dried. The residue was purified by flash column chromatography on basified silica gel with 25% ammoniac solution in water (EtOAc:iPrOH:H₂O / 4:2:1, R_f = 0.15) to yield **5** (10.3 mg, 0.015 mmol, 24 %) as a colorless lyophilisat. ¹H NMR (400 MHz, MeOD) δ = 5.44 (t, *J* = 6.4 Hz, 1H, H-2'), 5.09-5.16 (m, 3H, H-6', H-10', H-14'), 4.50 (t, *J* = 6.8 Hz, 3H, H-1', H-1), 4.00 (dd, *J* = 9.2 Hz, 5.6 Hz, 2.0 Hz, 1H, H-2), 3.92 (d, *J* = 11.6 Hz, 1H, H-6a), 3.78 (t, *J* = 8.4 Hz, 1H, H-5), 3.59-3.65 (m, 2H, H-3, H-6b), 3.19 (t, *J* = 8.8 Hz, 1H, H-4), 2.26-2.34 (m, 1H, H-7a), 2.04-2.13 (m, 12H, H-4', H-5', H-8', H-9', H-12', H-13'), 1.99 (s, 3H, NHAc), 1.94-1.96 (m, 1H, H-7b), 1.74, 3x1.68, 1.61 (5xs, 5x3H, H-16', H-17', H-18', H-19', H-20'). ¹³C NMR (101 MHz, MeOD) δ = 173.6 (s, NHC=OCH₃), 140.5 (s, C-3'), 136.5 (s, C-7'), 136.3 (s, C-11'), 132.4 (s, C-15'), 126.1 (s, C-16'), 125.9 (s, C-10'), 125.4 (s, C-14'), 123.5 (s, C-2'), 75.0 (s, C-5), 73.6 (s, C-4), 72.5 (s, C-3), 71.7 (d, *J* = 3.3 Hz, C-1), 63.6 (s, C-1'), 63.5 (s, C-6), 55.1 (d, *J* = 1.9 Hz, C-2), 33.3 (s, C-4'), 33.3 (s, C-8'), 32.9 (s, C-12'), 27.7 (s, C-5'), 27.7 (s, C-9'), 27.5 (s, C-13'), 26.4 (s, C-7), 26.0 (s, NHC=OCH₃), 23.7 (s, C-19'), 23.7 (s, C-20'), 23.7 (s, C-16'), 22.8 (s, C-17'), 17.8 (s, C-18'). ³¹P NMR (162 MHz, MeOD) δ = 13.3 (d, *J* = 23.9 Hz, P1), -9.9 (d, *J* = 23.9 Hz, P2). ESI-HRMS (-) *m/z* calculated 650.2865 (M-[H⁺]), found 650.2847 for C₂₉H₅₀NO₁₁P₂⁻.

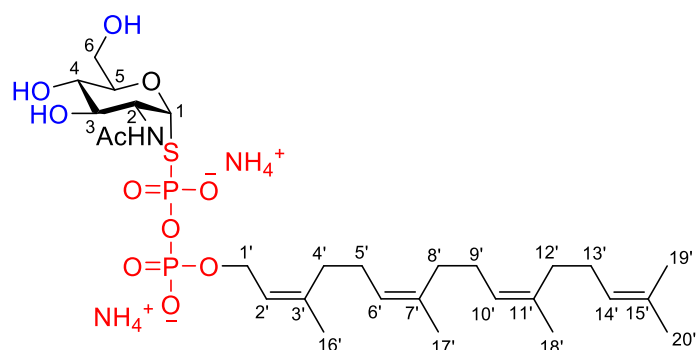
Synthesis of (((2S,3R,4R,5S,6R)-3-acetamido-4,5-dihydroxy-6-(hydroxymethyl)tetrahydro-2H-pyran-2-yl)(hydroxy)methyl)phosphonic ((2Z,6Z,10Z)-3,7,11,15-tetramethylhexadeca-2,6,10,14-tetraen-1-yl phosphoric) anhydride (major diastereomer at C7) (**6**)



4 (30 mg, 0.07 mmol, 1.00 eq) and CDI (60.1 mg, 0.37 mmol, 5.00 eq) were dissolved in 2.4 ml DMF and stirred for 3 h at RT under an Ar atmosphere. 0.04 ml dry MeOH was added and the reaction mixture was stirred for 1 h to quench the unreacted CDI. MeOH was removed *in vacuo*. **2** (major

diastereomer) (98.2 mg, 0.22 mmol, 3.00 eq) in 2.4 ml DMF was added to the reaction mixture which was then stirred for 6 days. The reaction mixture was concentrated in vacuo and the residue was dissolved in 6 ml MeOH and ammonium hydroxyde (7.09 ml, 44.50 mmol, 600 eq) was added. The reaction mixture was stirred at RT overnight, then concentrated at the rotavapor and freeze-dried. The residue was purified by flash column chromatography on basified silica gel with 25% ammoniac solution in water (EtOAc:iPrOH:H₂O / 4:2:1, R_f = 0.15) to yield **6** (major diastereomer) (8.5 mg, 0.012 mmol, 17 %) as a colorless lyophilisat. ¹H NMR (400 MHz, MeOD) δ = 5.45 (t, *J* = 6.8 Hz, 1H, H-2'), 5.11-5.16 (m, 3H, H-6', H-10', H-14'), 4.53 (t, *J* = 6.8 Hz, 2H, H-1'), 4.35-4.40 (m, 1H, H-1), 4.23 (dd, *J* = 8.0 Hz, 6.4 Hz, 1H, H-7), 4.17 (dd, *J* = 8.0 Hz, 4.8 Hz, 1H, H-2), 4.02 (t, *J* = 7.2 Hz, 2H, H-3), 3.94 (dt, *J* = 7.6 Hz, 2.4 Hz, 1H, H-5), 3.86 (dd, *J* = 12.0 Hz, 8.0 Hz, 1H, H-6a), 3.66 (dd, *J* = 12.0 Hz, 2.4 Hz, 1H, H-6b), 3.35 (t, *J* = 6.4 Hz, 1H, H-4), 2.04-2.13 (m, 12H, H-4', H-5', H-8', H-9', H-12', H-13'), 2.02 (s, 3H, NHAc), 1.74, 3x1.68, 1.61 (5xs, 5x3H, H-16', H-17', H-18', H-19', H-20'). ¹³C NMR (101 MHz, MeOD) δ = 173.9 (s, NHC=OCH₃), 140.6 (s, C-3'), 136.4 (s, C-7'), 136.3 (s, C-11'), 132.4 (s, C-15'), 126.1 (s, C-16'), 125.9 (s, C-10'), 125.4 (s, C-14'), 123.3 (s, C-2'), 78.7 (s, C-5), 73.6 (s, C-3), 72.3 (s, C-1), 72.1 (s, C-7), 71.9 (s, C-4), 63.8 (s, C-1'), 62.3 (s, C-6), 53.3 (s, C-2), 33.3 (s, C-4'), 33.3 (s, C-8'), 32.9 (s, C-12'), 27.7 (s, C-5'), 27.6 (s, C-9'), 27.5 (s, C-13'), 25.9 (s, NHC=OCH₃), 23.7 (s, C-19'), 23.7 (s, C-20'), 23.7 (s, C-16'), 23.2 (s, C-17'), 17.8 (s, C-18'). ³¹P NMR (162 MHz, MeOD) δ = 11.3 (d, *J* = 19.6 Hz, P1), -9.0 (d, *J* = 19.6 Hz, P2). ESI-HRMS (-) *m/z* calculated 666.2814 (M-[H⁺]), found 666.2796 for C₂₉H₅₀NO₁₂P₂⁻.

Synthesis of (((2S,3R,4R,5S,6R)-3-acetamido-4,5-dihydroxy-6-(hydroxymethyl)tetrahydro-2H-pyran-2-yl)thiophosphonic ((2Z,6Z,10Z)-3,7,11,15-tetramethylhexadeca-2,6,10,14-tetraen-1-yl phosphoric) anhydride (7**))**



4 (40 mg, 0.10 mmol, 1.00 eq) and CDI (80 mg, 0.49 mmol, 5.00 eq) were dissolved in 3 ml DMF and stirred for 3 h at RT under an Ar atmosphere. 0.05 ml dry MeOH was added and the reaction mixture was stirred for 1 h to quench the unreacted CDI. MeOH was removed *in vacuo* and **3** (174 mg, 0.41 mmol, 4.14 eq) in 3 ml DMF was added to the reaction mixture which

was then stirred for 6 days. The reaction mixture was concentrated to dryness *in vacuo*. The residue was dissolved in 10 ml MeOH and 12.60 ml NH₄OH was added dropwise at 0°C. The reaction mixture was then stirred at RT overnight, concentrated to remove MeOH and freeze-dried. The residue was purified by flash column chromatography on basified silica gel with 25% ammoniac solution in water (EtOAc:iPrOH:H₂O / 4:2:1, R_f = 0.25) to yield **7** (14 mg, 0.02 mmol, 20 %) as a colorless lyophilisat. ¹H NMR (400 MHz, MeOD) δ = 5.89 (dd, *J* = 10.4 Hz, 4.4 Hz, 1H, H-1), 5.44 (t, *J* = 6.4 Hz, 1H, H-2'), 5.13-5.16 (m, 3H, H-6', H-10', H-14'), 4.50 (t, *J* = 6.8 Hz, 2H, H-1'), 4.14 (dd, *J* = 10.8 Hz, 4.8 Hz, 1H, H-2), 4.06 (dt, *J* = 9.2 Hz, 1.6 Hz, 1H, H-5), 3.95 (dd, *J* = 12.0 Hz, 2.0 Hz, 1H, H-6a), 3.65 (dd, *J* = 12.0 Hz, 4.8 Hz, 1H, H-6b), 3.60 (m, 1H, H-3), 3.27 (t, *J* = 9.2 Hz, 1H, H-4), 2.05-2.13 (m, 12H, H-4', H-5', H-8', H-9', H-12', H-13'), 2.02 (s, 3H, NHAc), 1.74, 3x1.68, 1.61 (5xs, 5x3H, H-16', H-17', H-18', H-19', H-20'). ¹³C NMR (101 MHz, MeOD) δ = 172.7 (s, NHC=OCH₃), 139.3, (s, C-3'), 135.1 (s, C-7'), 134.9 (s, C-11'), 131.0 (s, C-15'), 124.7 (s, C-14'), 124.5 (s, C-10'), 124.0 (s, C-6'), 122.6 (s, C-2'), 84.6 (d, *J* = 3.3 Hz, C-1), 74.2 (s, C-5), 71.8 (s, C-3), 71.7 (s, C-4), 62.3 (d, *J* =

5.3 Hz, C-1'), 61.7 (s, C-6), 54.5 (d, $J = 6.2$ Hz, C-2), 31.9 (s, C-4'), 31.9 (s, C-8'), 31.5 (s, C-12'), 26.3 (s, C-5'), 26.2 (s, C-9'), 26.1 (s, C-13'), 24.6 (s, C-19'), 22.4 (s, C-20'), 22.4 (s, C-16'), 22.3 (s, C-17'), 21.5 (s, NHC=OCH₃), 16.4 (s, C-18'). ³¹P NMR (122 MHz, MeOD) $\delta = 4.5$ (d, $J = 26.9$ Hz, P1), -10.9 (d, $J = 26.9$ Hz, P2). ESI-HRMS (-) m/z calculated 668.2429 (M-[H⁺]), found 668.2414 for C₂₈H₄₈NO₁₁P₂S⁻.