ARTICLE

doi:10.1038/nature10151

X-ray structure of a bacterial oligosaccharyltransferase

Christian Lizak[1,2], Sabina Gerber[2], Shin Numao[1]†, Markus Aebi[1] & Kaspar P. Locher[2]

Asparagine-linked glycosylation is a post-translational modification of proteins containing the conserved sequence motif Asn-X-Ser/Thr. The attachment of oligosaccharides is implicated in diverse processes such as protein folding and quality control, organism development or host–pathogen interactions. The reaction is catalysed by oligosaccharyltransferase (OST), a membrane protein complex located in the endoplasmic reticulum. The central, catalytic enzyme of OST is the STT3 subunit, which has homologues in bacteria and archaea. Here we report the X-ray structure of a bacterial OST, the PglB protein of *Campylobacter lari*, in complex with an acceptor peptide. The structure defines the fold of STT3 proteins and provides insight into glycosylation sequon recognition and amide nitrogen activation, both of which are prerequisites for the formation of the N-glycosidic linkage. We also identified and validated catalytically important, acidic amino acid residues. Our results provide the molecular basis for understanding the mechanism of N-linked glycosylation.

It is estimated that more than half of all eukaryotic proteins are glycoproteins; that is, specific amino acid side chains are chemically modified with carbohydrates in a process termed glycosylation[1,2]. The most abundant of these modifications is asparagine-linked (N-linked) glycosylation, which affects a multitude of cellular functions[3–5]. Asparagines are specifically glycosylated in the context of a consensus sequon Asn-X-Ser/Thr when located in the endoplasmic reticulum (ER). The reaction takes place on the luminal surface of the ER membrane and is catalysed by OST, a hetero-oligomeric membrane protein complex in most eukaryotes[6]. A hallmark of N-linked glycosylation is its broad specificity with respect to the polypeptide substrate, which is a direct consequence of the short recognition sequon[7]. This distinguishes OST from *O*-glycosyltransferases that modify serine or threonine residues and exhibit a higher specificity for their protein substrates[8].

The key step of OST-catalysed glycosylation is the formation of an N-glycosidic linkage between the amide nitrogen of the acceptor asparagine and the C1 carbon of the first saccharide moiety of a lipid-linked oligosaccharide (LLO) donor (Fig. 1a). This results in the en bloc transfer of the oligosaccharide onto the acceptor asparagine. Details of the reaction mechanism are poorly understood owing to the absence of structural insight into OST at high resolution, but also to the complex chemical nature of LLO, its low abundance in biological samples, and its insolubility in water. In contrast, crystal structures of various soluble *O*-glycosyltransferases have been published and their reaction mechanisms investigated in detail[9–12]. For OST, the currently accepted model suggests that glycosylation sequons are recognized when in unfolded protein segments[13–15], either during protein translocation into the ER or after translocation is completed[16]. The central, catalytically active component within OST is the STT3 subunit[6,17]. The other subunits are thought to assist and refine the reaction[18,19].

N-linked glycosylation is not restricted to eukaryotes. Homologous processes are found in archaea and in defined taxa of proteobacteria[20,21]. Eukaryotic and bacterial LLOs contain isoprenoid moieties that anchor the oligosaccharides in the membrane, and pyrophosphates as leaving groups of the substitution reactions. However, the attached oligosaccharides are chemically distinct, and unlike their eukaryotic counterparts, the OSTs of prokaryotes (and of eukaryotic kinetoplastids) consist of a single subunit, which is homologous to the STT3 subunits of eukaryotic OST complexes (Supplementary Fig. 1)[22–24]. The best-studied bacterial OST, termed PglB (84 kDa), is encoded in the protein glycosylation locus *pgl* of the Gram-negative bacterium *Campylobacter jejuni*. This gene cluster is sufficient for catalysing protein glycosylation when transferred into *Escherichia coli* cells[25]. The similarity in sequence and membrane topology indicates that PglB and eukaryotic STT3s share a common reaction mechanism[6,25–27]. To understand the molecular basis of N-linked glycosylation, we have determined the X-ray structure of PglB from *C. lari*, which is 56% identical to that of *C. jejuni*[28]. *C. lari* PglB is functional when co-expressed with the *C. jejuni pgl* cluster in *E. coli* cells (Fig. 1b). We co-crystallized *C. lari* PglB with the hexapeptide DQNATF, an optimal acceptor sequence for *C. jejuni* PglB[29]. X-ray diffraction data was anisotropic and extended to 3.4 Å resolution. The structure was refined to R/R_{free} values of 23.8% and 27.1%, respectively (Supplementary Table 1).

Structure of *C. lari* PglB

In agreement with earlier predictions (reviewed in ref. 6), the structure of PglB revealed two domains: a transmembrane domain comprising residues 1–432 and a periplasmic domain comprising residues 433–712 (Fig. 1c). In addition to the covalent linkage, the two domains have extensive non-covalent interactions, provided mainly by the first external loop (EL1) of the transmembrane domain that forms two helices parallel to the membrane plane. The periplasmic domain features a mixed α/β fold that was previously observed in the structures of the homologous domains of *C. jejuni* PglB and of *Pyrococcus furiosus* AglB[30,31]. However, these isolated domains were catalytically inactive and unable to bind acceptor peptide. Our structure of full-length PglB provides a molecular explanation by revealing that the transmembrane domain is indispensible both for peptide binding and catalysis.

In contrast to the periplasmic domain, the transmembrane domain features a novel fold (Fig. 2). Thirteen transmembrane segments are

[1]Institute of Microbiology, Department of Biology, ETH Zurich, CH-8093 Zurich, Switzerland. [2]Institute of Molecular Biology and Biophysics, Department of Biology, ETH Zurich, CH-8093 Zurich, Switzerland. †Present address: Novartis Pharma AG, CH-4002 Basel, Switzerland.

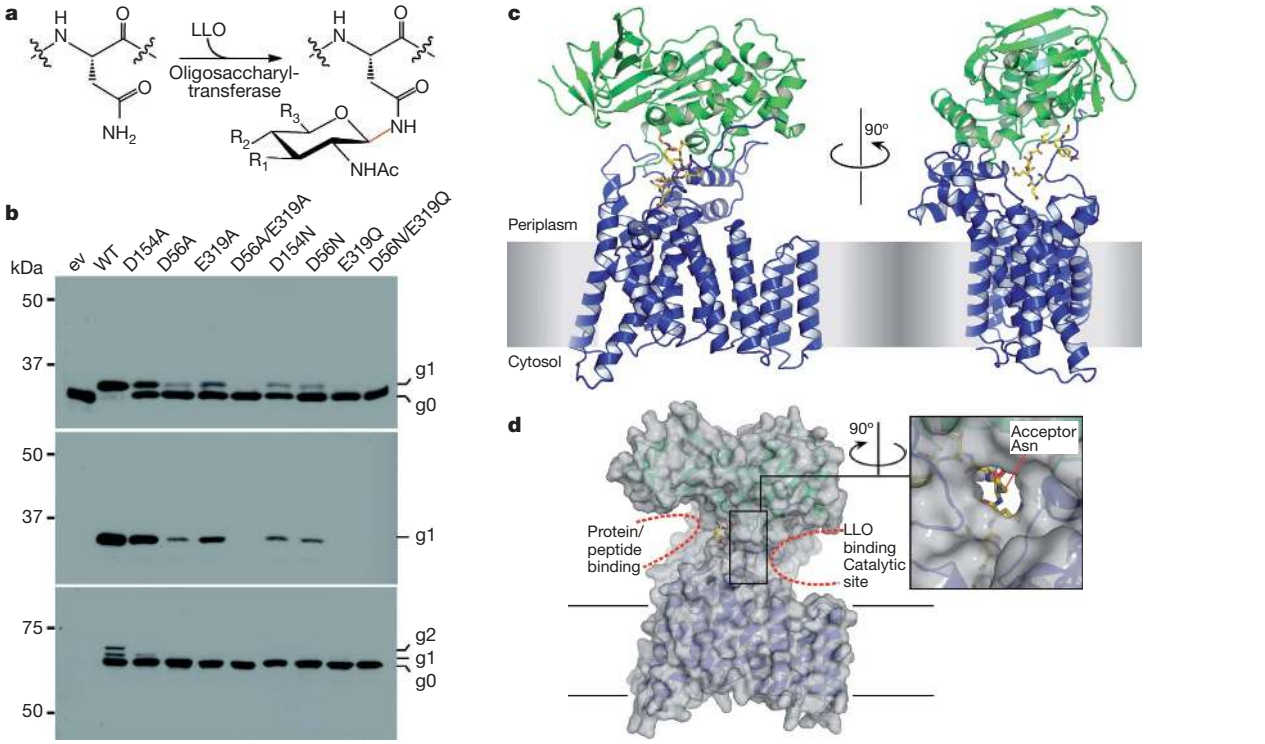

Figure 1 | Activity and structure of *C. lari* PglB. **a**, Reaction scheme of N-linked glycosylation, yielding an N-glycosidic bond (red). In bacteria, R_1 = oligosaccharyl, R_2 = NH-Ac, R_3 = CH_3. In eukaryotes, R_1 = OH, R_2 = oligosaccharyl, R_3 = CH_2OH. **b**, *In vivo* glycosylation assay in *E. coli*. Immunoblots detecting acceptor protein 3D5 (top), glycans (middle) or PglB (bottom). PglB constructs indicated above the lanes include vector control (ev), wild type (WT), or PglB mutants. Glycosylation yields a mobility shift from the unmodified (g0) to the glycosylated form (g1). Functional PglB is partially auto-glycosylated at N535 and N556, resulting in two additional bands (g1 and g2). **c**, Ribbon diagrams of PglB structure, with transmembrane and periplasmic domains in blue and green, respectively, and bound acceptor peptide in stick representation and coloured yellow. The presumed position of the membrane is indicated by a grey rectangle. **d**, Surface representation in semi-transparent grey, with ribbons as in **a**. Two cavities at opposite sides of PglB are indicated by dashed lines, providing access for substrates. The cavities are connected by a tunnel that accommodates the acceptor asparagine.

connected by short cytoplasmic and external loops, with the exception of the long external loops EL1 and EL5. Whereas EL1 is well ordered, EL5 is only partially ordered, with 25 residues disordered in the electron density map. Transmembrane segments TM1–4 and TM10–13 form the sequon-binding and catalytic sites and provide the bulk of the interface with the periplasmic domain. In the peptide-bound state, PglB forms two large cavities above the membrane surface, located at opposite sides of the protein (Fig. 1d). The left-side cavity provides access for acceptor proteins as suggested by the presence of bound peptide in the structure, whereas the right-side cavity harbours the catalytic residues and probably serves as the binding pocket for LLO. The two cavities are connected where the side chain of the acceptor asparagine reaches from the peptide-binding site into the catalytic pocket.

Acceptor sequon binding and recognition

Clear density for the bound hexapeptide DQNATF was observed in a location that placed the acceptor asparagine some 15 Å above the membrane surface (Fig. 1c). Almost 80% of the contact surface of the peptide (calculated by areaimol[32]) is buried at the interface of the transmembrane and periplasmic domains of PglB (Fig. 3a), indicating tight binding and a firmly imposed conformation. The peptide forms a loop that almost completes a 180° turn; accordingly, protein substrates have to present their glycosylation sequons in sufficiently large, flexible and surface-exposed loops[33], because the peptide-binding cavity of PglB does not appear to fit fully folded protein domains. The observed conformation of the peptide would be incompatible with a proline residue at the +1 position, in agreement with the observation that +1 prolines are not allowed in glycosylation sequons[34,35].

A hallmark of N-linked glycosylation is the requirement of a serine or threonine at the +2 position of the acceptor sequon. The PglB structure provides a molecular explanation by revealing that the β-hydroxyl group of the +2 Thr of bound hexapeptide forms three hydrogen bonds, one with each of the side chains of the 'WWD motif', which is strictly conserved in STT3 proteins (Fig. 3b and Supplementary Fig. 3). The motif is located in the periplasmic domain, and the interaction of the two tryptophan and the aspartate side chains saturate the hydrogen-bonding capacity of the β-hydroxyl group. The arrangement physically separates the +2 Thr from the acceptor asparagine, and we conclude that the WWD motif defines the polypeptide substrate specificity, but is not directly involved in catalysis. Notably, the structure can also explain preferences and deviations at the +2 position of glycosylation sequons. The γ-methyl group of the +2 Thr is in van der Waals contact with Ile 572 of PglB (3.6 Å distance to the γ-methyl group of Ile 572; Fig. 3b and Supplementary Fig. 4). This stabilizing interaction is absent if a serine is in the +2 position, which may explain that acceptor sequons containing a +2 Thr are glycosylated 40 times more efficiently than if they contain a +2 Ser (ref. 36). The structure further suggests that the non-natural, *S*-configured threonine would cause a steric clash with Ile 572. *S*-configured threonine is indeed not allowed at the +2 position, with a 15,000-fold reduction in glycosylation efficiency compared to *R*-configured threonine[37]. Ile 572 is conserved in bacteria and has been suggested to be part of a MXXI motif[31]. However, the corresponding residue in the archaeal AglB protein was found to be a lysine[30], and sequence alignments with eukaryotic STT3 homologues reveal no clear conservation of Ile 572, indicating that residues other than isoleucine can provide contacts to the +2 Thr in homologous proteins. The PglB structure can also explain allowed deviations from

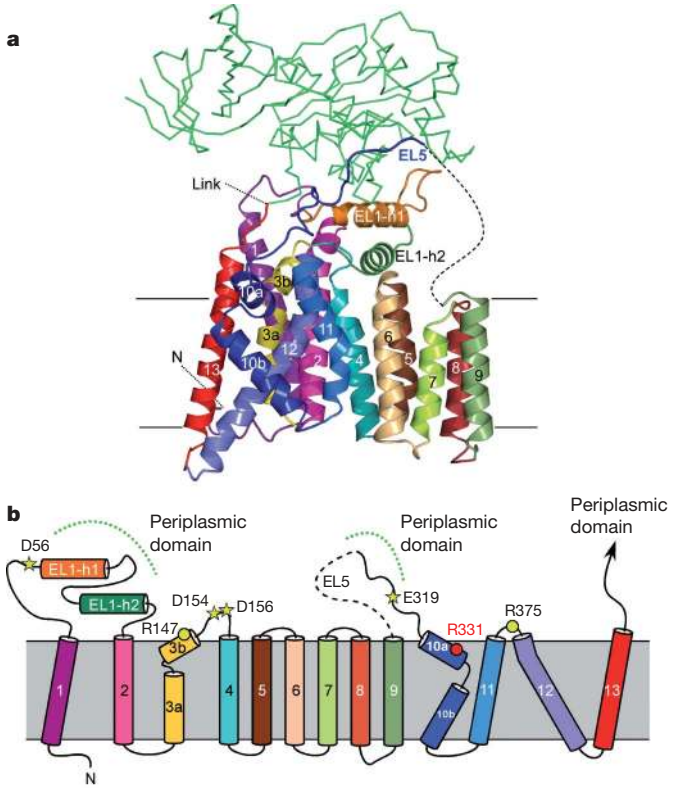

Figure 2 | Topology of transmembrane domain. **a**, Ribbon diagram, with helices numbered and coloured as in **b**. The periplasmic domain is shown as a green backbone trace. **b**, Topological schematic indicating helices and connecting loops. Dashed green lines indicate non-covalent contacts to the periplasmic domain. Conserved residues forming the active site of PglB are indicated by yellow stars and spheres and are labelled. R331 contributing to peptide recognition is indicated in red.

the consensus sequon: the acceptor sequence N-X-C, present in ~2.2% of experimentally determined glycosylation sites of the mouse glycoproteome[2], is probably allowed because the β-sulphhydryl group of cysteine can form similar hydrogen bonds as a β-hydroxyl group. Glycines, alanines and valines have also been reported at the +2 position of glycosylated sequons, albeit only at low abundance[2,28,38]. These residues can, in principle, be accommodated in the binding pocket of PglB because they are equal in size or smaller than threonine. However, glycosylation of sequons such as N-G-X, with X being larger than threonine, or of T/S-X-N (reverse sequons)[2] cannot be explained by the PglB structure.

Compared to the eukaryotic enzymes, bacterial OSTs analysed thus far have an extended acceptor sequon: glycosylation is only efficient if a negatively charged residue (Asp or Glu) is present at the −2 position, resulting in a consensus sequon D/E-X-N-X-S/T[35]. In PglB, R331 provides a salt bridge to the −2 Asp of bound peptide (Fig. 3a and Supplementary Fig. 4), thereby strengthening the PglB–peptide interaction. R331 is conserved in bacteria, but not in eukaryotes, where no requirement for a negative charge at the −2 position is observed. The extended sequon recognition may reflect the need for tighter peptide binding in bacteria, where the local concentration of the acceptor polypeptide is probably lower than in eukaryotes.

Catalytic site

The catalytic pocket is located in the right-side cavity of PglB (Fig. 1d) and is marked by a bound cation, located ~8 Å above the membrane boundary. Owing to the high concentration of magnesium salt in the crystallization solution, it was modelled as Mg^{2+}. PglB, like all OSTs, is only functional with a bound divalent cation (M^{2+})[39,40]. The physiological cation was suggested to be Mn^{2+}, but PglB is also active in Mg^{2+}, a property that was previously observed for other metal-dependent glycosyltransferases[11]. The catalytic pocket of PglB features three acidic side chains (D56, D154, E319) that are part of the transmembrane domain and coordinate M^{2+} (Fig. 4 and Supplementary Fig. 5). At the current resolution, water molecules that might be additional ligands of M^{2+} cannot be modelled. The residues located in the catalytic pocket are generally conserved in STT3 proteins (Supplementary Fig. 3). The aspartates D154 and D156 belong to a previously reported D-X-D motif, and mutation of either aspartate to an alanine abolished the activity of the mannosyltransferase GPI-MT-1, a member of the same glycosyltransferase superfamily as PglB (GT-C)[41,42]. In contrast, D56 and E319 have not been previously identified as catalytically relevant, but their carboxyl groups interact both with M^{2+} and the amido group of the acceptor asparagine. Such interactions would not be possible if a glutamine or an aspartate side chain was present instead of the acceptor asparagine, which may explain why glutamines and aspartates are not glycosylated by OST. To confirm the catalytic involvement of the three M^{2+}-binding residues, we mutated them individually to alanines and tested the activity of the resulting PglB mutants in a complementation assay (Fig. 1b). Even though OST activity is not limiting in our assay, the mutation D154A reduced the observed glycosylation yield by >50%, whereas D56A

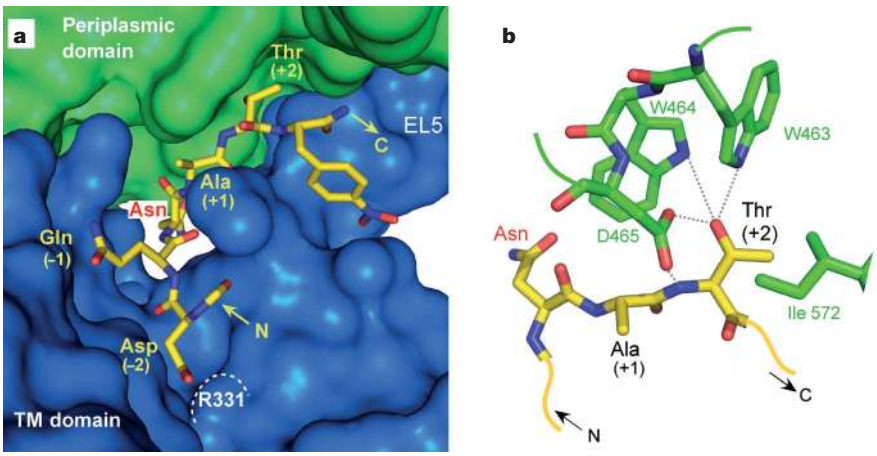

Figure 3 | Sequon binding and recognition. **a**, Transmembrane (TM) and periplasmic domains of PglB are in blue and green, respectively. Acceptor peptide is in ball and stick representation; N and C denote amino and carboxy termini. Peptide residues are labelled yellow, acceptor Asn is labelled red. **b**, PglB residues interacting with the +2 Thr of bound peptide are shown in green and labelled. Hydrogen bonds from the WWD motif to the β-hydroxyl group are indicated by dashed lines.

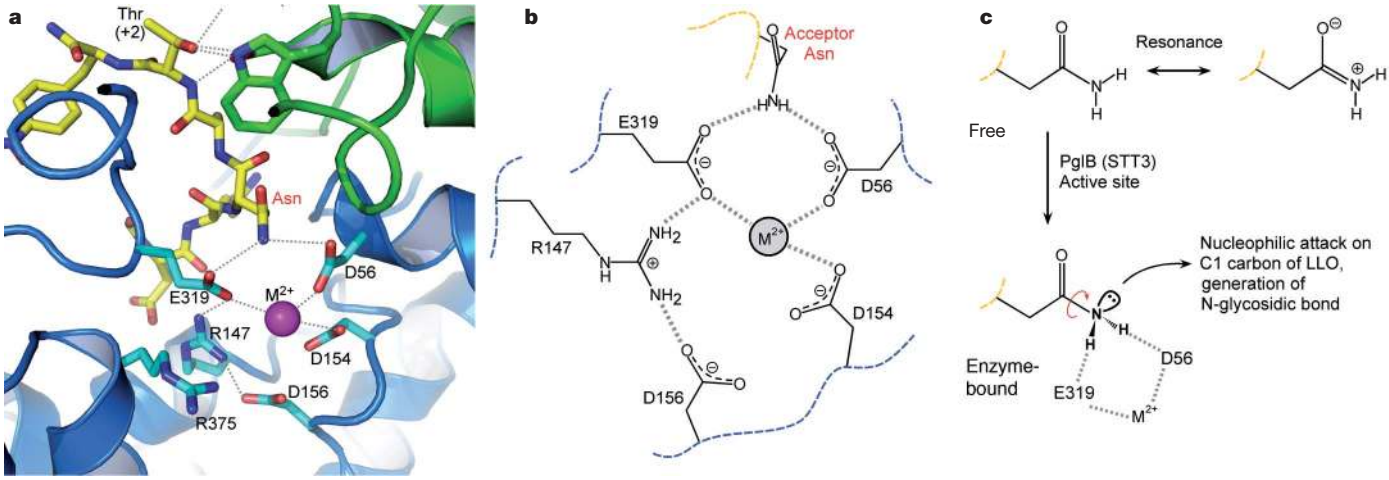

Figure 4 | Catalytic site and amide nitrogen activation. **a**, Transmembrane and periplasmic domains of PglB are coloured blue and green, respectively. Selected side chains are in ball and stick representation, with carbon atoms coloured cyan and green for transmembrane and periplasmic domain residues, respectively, and yellow for acceptor peptide. Grey dashed lines indicate hydrogen bonds or interactions with the divalent cation M^{2+}. **b**, Chemical structure of the catalytic site, indicating interactions as in **a**. Blue and yellow dashed lines indicate transmembrane domain and peptide backbones, respectively. **c**, Presumed mechanism of amide nitrogen activation. Yellow dashed lines indicate peptide backbone. The amido group of free acceptor Asn features electron delocalization, indicated by resonance. When bound to PglB, the amido group may form hydrogen bonds with the catalytically essential D56 and E319, requiring rotation around the C–N bond (red arrow).

and E319A reduced it by >90%. A PglB double mutant D56A/E319A was completely inactive.

There is a controversial discussion on how the amido group of the acceptor asparagine might be activated for a nucleophilic attack on the C1 carbon of LLO substrate. Amides are poor nucleophiles because the free electron pair of the nitrogen is conjugated to the carbonyl group. As a consequence, the N–C bond has double-bond character, and the nucleophilicity of the nitrogen is low. To explain the reactivity of the amido group, specific conformations of the acceptor peptide, such as a 'β-turn' or an 'Asx turn', have been proposed, invoking direct involvement of the +2 Ser/Thr in catalysis[34,43]. Given the firm binding of the +2 Thr to the WWD motif in our PglB structure, we can rule out such an involvement. Instead, the structure of PglB presents a distinct possibility for explaining amide nitrogen activation: the catalytically essential residues D56 and E319 are optimally positioned to form hydrogen bonds with the two amide protons of the acceptor asparagine. Forming such hydrogen bonds would require a rotation of the N–C bond of the amido group, thereby abolishing the conjugation of the nitrogen electrons with the carbonyl group (Fig. 4c). Not only would this increase the electronegative nature of the amide nitrogen (by polarizing the N–H bonds and increasing the electron density on the nitrogen), but it would also generate an sp^3 hybridized nitrogen with a reactive lone pair optimally positioned for a nucleophilic attack. The energy barrier for rotating the N–C bond in most amides is estimated to be 16–20 kcal mol^{-1}, and the 270° amide conformation shown in Fig. 4c has been calculated to have an energy of ~18.6 kcal mol^{-1} relative to the planar conformation[44]. Hence, it would take 1–2 low-barrier hydrogen bonds[45] (each worth ~10 kcal mol^{-1}) to provide sufficient energy to break permanently the conjugation of the carboxamide group of the acceptor asparagine. The carboxyl groups of D56 or E319 might provide such interactions in the transition state of the reaction. However, it will require a higher resolution structure to measure reliably the lengths of the hydrogen bonds. Mutating D56 to asparagine (D56N) has an even more pronounced inhibitory effect than truncation to alanine, and the E319Q mutant is completely inactive (Fig. 1b). This demonstrates that the negative charges provided by the carboxyl groups of D56 and E319 are essential for catalysis and that the acidic side chains cannot be replaced by iso-electronic amides. Steric effects might explain the increased inhibition of D56N relative to D56A and of E319Q relative to E319A.

Glycosylation mechanism

Given that PglB is active even when solubilized in detergent (used for crystallization), our structure has probably captured a functionally competent state. Glycosylation occurs with inversion of configuration at the substituted C1 carbon. We modelled the LLO substrate into the PglB structure such that the di-*N*-acetyl-bacillosamine moiety is properly aligned for a nucleophilic attack by the activated amide nitrogen, while the leaving pyrophosphate group is in contact with M^{2+} and the conserved R375 (Fig. 5a and Supplementary Fig. 3). This arrangement places the additional saccharide moieties in the right-hand cavity of PglB and the C2 substituent of the first saccharide moiety, a *N*-acetyl group present in bacterial and eukaryotic LLO, in the vicinity of a conserved tyrosine residue (Y468, Supplementary Fig. 3), where density consistent with a bound water molecule is observed. When modelled as shown in Fig. 5a, the isoprenoid moieties of LLO are embedded in a hydrophobic groove on the PglB surface, pointing into the lipid bilayer (Supplementary Fig. 6). The function of M^{2+} in PglB appears to be twofold: on the one hand, it orients the acidic residues D56 and E319 that interact with the acceptor asparagine, and on the other, it stabilizes the leaving group (lipid-pyrophosphate) of the substitution. The dual function seems to be distinct from metal-dependent, configuration-inverting glycosyltransferases of the GT-A family, where M^{2+} primarily serves the stabilization of the leaving group[11].

With the acceptor peptide present in the structure and the LLO molecule tentatively modelled, we can propose a basic, three-state catalytic cycle for PglB-catalysed glycosylation (Fig. 5b). A critical element of the proposed mechanism is the engagement and disengagement of EL5, which we expect to be flexible and disordered in the absence of bound acceptor peptide (ground state). Upon sequon binding, the C-terminal half of EL5 is ordered and pins the bound acceptor peptide against the periplasmic domain of PglB, thereby restricting its motion. Because the essential E319 is part of EL5, this simultaneously results in the formation of the catalytic site, where the acceptor asparagine is oriented and the amide nitrogen activated. Binding of LLO will then result in a nucleophilic attack of the activated amide nitrogen, resulting in glycosylation. Once the glycosidic bond is formed, the newly attached saccharides are tightly pressed against PglB, causing steric tension that can be released by disengagement of EL5. This allows the glycopeptide to dissociate from the enzyme. Subsequent cleavage of the lipid-linked pyrophosphate

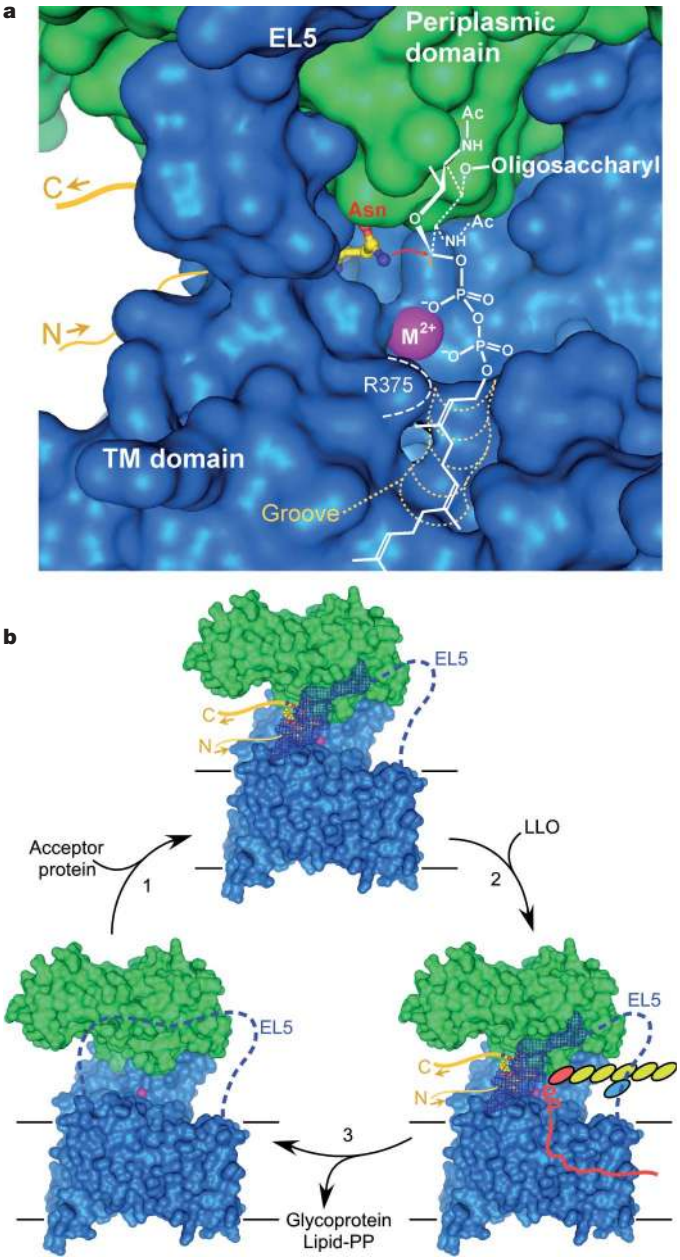

Figure 5 | Proposed glycosylation mechanism of PglB. Transmembrane (TM) and periplasmic domain surfaces are shown in blue and green, respectively. Bound acceptor peptide is in ball and stick representation, and yellow lines indicate the N and C termini. **a**, The chemical structure of bacterial LLO is shown in white, highlighting presumed interactions with M^{2+} and R375 of PglB, and a collinear arrangement of attacking and leaving groups of the substitution. A red arrow indicates the nucleophilic attack. A predominantly hydrophobic groove accommodating the isoprenoid moieties is indicated. **b**, The observed crystal structure reflects the top state, with acceptor peptide bound and the C-terminal half of EL5 ordered and engaged. The bottom-left state reflects the ground state, with no substrates bound and EL5 disordered, as indicated by a dashed line. In the bottom-right state, bound *C. jejuni* LLO is shown (red line for isoprenoid moieties, P for phosphate, ellipsoids for saccharide moieties). The molecular events indicated by arrows are: 1, sequon binding, EL5 engagement, acceptor asparagine activation; 2, LLO binding, glycosylation; 3, disengagement of EL5, release of glycosylated sequon and lipid-pyrophosphate (PP).

anhydride and folding of the newly glycosylated protein probably provide the main contributions to the driving force of the *in vivo* reaction. We should point out that there is no experimental evidence indicating a strict sequence of events. Thus, the binding of LLO to PglB might precede that of acceptor peptide.

Conclusions and outlook

Our results provide the basis for understanding the mechanism of N-linked glycosylation at a molecular level. Future studies will be directed at probing the outlined chemical concepts, which are at the core of both bacterial and eukaryotic OST. In addition, the interaction of the catalytic STT3 with the additional subunits in eukaryotic OST needs to be defined. At present, the highest resolution structure of a eukaryotic OST is a 12 Å electron microscopy map of the yeast enzyme[46]. The PglB structure also provides opportunities for engineering the substrate specificity both with respect to the acceptor sequon and LLO, which may open up new avenues for the production of glycoprotein and glycoconjugate therapeutics.

METHODS SUMMARY

***In vivo* glycosylation assay.** *Escherichia coli* SCM6 cells were transformed with three separate plasmids containing: (1) the *C. jejuni* $pglB_{mut}$ cluster (containing an inactivated *pglB* gene) to generate and flip LLO; (2) the glycosylation acceptor protein 3D5, a single-chain Fv fragment containing a DQNAT acceptor sequon[47]; (3) *C. lari* PglB, wild type or mutants. Expression and glycosylation of 3D5 was monitored by SDS–PAGE of periplasmic cell extracts and visualized by mobility shift due to increased size in an immunoblot using anti-c-Myc antibody, or the reactivity of the glycoprotein in an anti-glycan immunoblot using hR6 antiserum. Expression of *C. lari* PglB was monitored by immunoblots of whole-cell extracts using anti-haemagglutinin (HA) antiserum.

Crystallization and structure determination. *Campylobacter lari* PglB containing a C-terminal His_{10} tag was expressed in *E. coli* cells and purified in β-D-dodecyl maltoside. PglB was co-crystallized with an acceptor peptide Ac-DQNATF $\{4\text{-}NO_2\}\text{-}NH_2$, containing an acetylated N terminus and an amidated C terminus. For experimental phasing, crystals were soaked in ethyl mercury phosphate (EMP) before data collection. Three distinct data sets of EMP derivatives were collected. The phase problem was solved by a combination of MIRAS and molecular replacement, using the periplasmic domain of *C. jejuni* PglB (Protein Data Bank code 3AAG) as a search model.

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 27 January; accepted 26 April 2011.

1. Apweiler, R., Hermjakob, H. & Sharon, N. On the frequency of protein glycosylation, as deduced from analysis of the SWISS-PROT database. *Biochim. Biophys. Acta* **1473,** 4–8 (1999).
2. Zielinska, D. F., Gnad, F., Wisniewski, J. R. & Mann, M. Precision mapping of an *in vivo* N-glycoproteome reveals rigid topological and sequence constraints. *Cell* **141,** 897–907 (2010).
3. Helenius, A. & Aebi, M. Roles of N-linked glycans in the endoplasmic reticulum. *Annu. Rev. Biochem.* **73,** 1019–1049 (2004).
4. Kornfeld, R. & Kornfeld, S. Assembly of asparagine-linked oligosaccharides. *Annu. Rev. Biochem.* **54,** 631–664 (1985).
5. Varki, A. Biological roles of oligosaccharides—all of the theories are correct. *Glycobiology* **3,** 97–130 (1993).
6. Kelleher, D. J. & Gilmore, R. An evolving view of the eukaryotic oligosaccharyltransferase. *Glycobiology* **16,** 47R–62R (2006).
7. Petrescu, A. J., Milac, A. L., Petrescu, S. M., Dwek, R. A. & Wormald, M. R. Statistical analysis of the protein environment of N-glycosylation sites: implications for occupancy, structure, and folding. *Glycobiology* **14,** 103–114 (2004).
8. Breton, C. & Imberty, A. Structure/function studies of glycosyltransferases. *Curr. Opin. Struct. Biol.* **9,** 563–571 (1999).
9. Charnock, S. J. & Davies, G. J. Structure of the nucleotide-diphospho-sugar transferase, SpsA from *Bacillus subtilis*, in native and nucleotide-complexed forms. *Biochemistry* **38,** 6380–6385 (1999).
10. Ünligil, U. M. & Rini, J. M. Glycosyltransferase structure and mechanism. *Curr. Opin. Struct. Biol.* **10,** 510–517 (2000).
11. Lairson, L. L., Henrissat, B., Davies, G. J. & Withers, S. G. Glycosyltransferases: Structures, functions, and mechanisms. *Annu. Rev. Biochem.* **77,** 521–555 (2008).
12. Lazarus, M. B., Nam, Y., Jiang, J., Sliz, P. & Walker, S. Structure of human O-GlcNAc transferase and its complex with a peptide substrate. *Nature* **469,** 564–567 (2011).
13. Chen, W. & Helenius, A. Role of ribosome and translocon complex during folding of influenza hemagglutinin in the endoplasmic reticulum of living cells. *Mol. Biol. Cell* **11,** 765–772 (2000).
14. Silberstein, S. & Gilmore, R. Biochemistry, molecular biology, and genetics of the oligosaccharyltransferase. *FASEB J.* **10,** 849–858 (1996).
15. Whitley, P., Nilsson, I. & von Heijne, G. A nascent secretory protein may traverse the ribosome endoplasmic reticulum translocase complex as an extended chain. *J. Biol. Chem.* **271,** 6241–6244 (1996).

16. Ruiz-Canada, C., Kelleher, D. J. & Gilmore, R. Cotranslational and Posttranslational N-glycosylation of polypeptides by distinct mammalian OST isoforms. *Cell* **136,** 272–283 (2009).
17. Zufferey, R. *et al.* Stt3, a highly conserved protein required for yeast oligosaccharyl transferase activity *in vivo*. *EMBO J.* **14,** 4949–4960 (1995).
18. Wilson, C. M. & High, S. Ribophorin I acts as a substrate-specific facilitator of N-glycosylation. *J. Cell Sci.* **120,** 648–657 (2007).
19. Schulz, B. L. *et al.* Oxidoreductase activity of oligosaccharyltransferase subunits Ost3p and Ost6p defines site-specific glycosylation efficiency. *Proc. Natl Acad. Sci. USA* **106,** 11061–11066 (2009).
20. Mescher, M. F. & Strominger, J. L. Purification and characterization of a prokaryotic glycoprotein from cell-envelope of *Halobacterium salinarium*. *J. Biol. Chem.* **251,** 2005–2014 (1976).
21. Szymanski, C. M., Yao, R. J., 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).
22. Izquierdo, L. *et al.* Distinct donor and acceptor specificities of *Trypanosoma brucei* oligosaccharyltransferases. *EMBO J.* **28,** 2650–2661 (2009).
23. 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,** 3758–3768 (2008).
24. Szymanski, C. M. & Wren, B. W. Protein glycosylation in bacterial mucosal pathogens. *Nature Rev. Microbiol.* **3,** 225–237 (2005).
25. Wacker, M. *et al.* N-linked glycosylation in *Campylobacter jejuni* and its functional transfer into *E. coli*. *Science* **298,** 1790–1793 (2002).
26. Abu-Qarn, M. & Eichler, J. Protein N-glycosylation in Archaea: defining *Haloferax volcanii* genes involved in S-layer glycoprotein glycosylation. *Mol. Microbiol.* **61,** 511–525 (2006).
27. Li, L. *et al*. Overexpression and topology of bacterial oligosaccharyltransferase PglB. *Biochem. Biophys. Res. Commun.* **394,** 1069–1074 (2010).
28. Schwarz, F. *et al.* Relaxed acceptor site specificity of bacterial oligosaccharyltransferase *in vivo*. *Glycobiology* **21,** 45–54 (2011).
29. Chen, M. M., Glover, K. J. & Imperiali, B. From peptide to protein: Comparative analysis of the substrate specificity of N-linked glycosylation in *C. jejuni*. *Biochemistry* **46,** 5579–5585 (2007).
30. Igura, M. *et al.* Structure-guided identification of a new catalytic motif of oligosaccharyltransferase. *EMBO J.* **27,** 234–243 (2008).
31. Maita, N., Nyirenda, J., Igura, M., Kamishikiryo, J. & Kohda, D. Comparative structural biology of eubacterial and archaeal oligosaccharyltransferases. *J. Biol. Chem.* **285,** 4941–4950 (2010).
32. Collaborative Computational Project, Number 4. The CCP4 Suite: programs for protein crystallography. *Acta Crystallogr. D* **50,** 760–763 (1994).
33. Kowarik, M. *et al.* N-linked glycosylation of folded proteins by the bacterial oligosaccharyltransferase. *Science* **314,** 1148–1150 (2006).
34. Bause, E. & Legler, G. The role of the hydroxy amino acid in the triplet sequence Asn-Xaa-Thr(Ser) for the N-glycosylation step during glycoprotein biosynthesis. *Biochem. J.* **195,** 639–644 (1981).
35. Kowarik, M. *et al.* Definition of the bacterial N-glycosylation site consensus sequence. *EMBO J.* **25,** 1957–1966 (2006).
36. Bause, E. Model studies on N-glycosylation of proteins. *Biochem. Soc. Trans.* **12,** 514–517 (1984).
37. Breuer, W., Klein, R. A., Hardt, B., Bartoschek, A. & Bause, E. Oligosaccharyltransferase is highly specific for the hydroxy amino acid in Asn-Xaa-Thr/Ser. *FEBS Lett.* **501,** 106–110 (2001).
38. Valliere-Douglass, J. F. *et al.* Asparagine-linked oligosaccharides present on a non-consensus amino acid sequence in the C_H1 domain of human antibodies. *J. Biol. Chem.* **284,** 32493–32506 (2009).
39. Imperiali, B. & Rickert, K. W. Conformational implications of asparagine-linked glycosylation. *Proc. Natl Acad. Sci. USA* **92,** 97–101 (1995).
40. Sharma, C., Lehle, L. & Tanner, W. N-glycosylation of yeast proteins—characterization of the solubilized oligosaccharyl transferase. *Eur. J. Biochem.* **116,** 101–108 (1981).
41. Liu, J. & Mushegian, A. Three monophyletic superfamilies account for the majority of the known glycosyltransferases. *Protein Sci.* **12,** 1418–1431 (2003).
42. Maeda, Y. *et al.* PIG-M transfers the first mannose to glycosylphosphatidylinositol on the lumenal side of the ER. *EMBO J.* **20,** 250–261 (2001).
43. Imperiali, B., Shannon, K. L. & Rickert, K. W. Role of peptide conformation in asparagine-linked glycosylation. *J. Am. Chem. Soc.* **114,** 7942–7944 (1992).
44. Wiberg, K. B. & Breneman, C. M. Resonance interactions in acyclic systems. 3. Formamide internal rotation revisited. Charge and energy redistribution along the C-N bond rotational pathway. *J. Am. Chem. Soc.* **114,** 831–840 (1992).
45. Cleland, W. W. & Kreevoy, M. M. Low-barrier hydrogen bonds and enzymatic catalysis. *Science* **264,** 1887–1890 (1994).
46. Li, H., Chavan, M., Schindelin, H., Lennarz, W. J. & Li, H. L. Structure of the oligosaccharyl transferase complex at 12 Å resolution. *Structure* **16,** 432–440 (2008).
47. Lizak, C., Fan, Y. Y., Weber, T. C. & Aebi, M. N-linked glycosylation of antibody fragments in *Escherichia coli*. *Bioconjug. Chem.* **22,** 488–496 (2011).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank the beamline staff at the Swiss Light Source for assistance with data collection, S. Fleurkens and M. Bucher for technical assistance, and D. Arigoni for discussions. *E. coli* SCM6 was provided by C. Marolda and M. Valvano. C.L. was affiliated with the Life Science Zurich Graduate School. This research was supported by the NCCR Structural Biology Zurich (grant to K.P.L.) and Swiss National Science Foundation grants to M.A. (SNF 31003A-127098/1) and K.P.L. (SNF 31003A-131075/1).

Author Contributions C.L., M.A. and K.P.L. designed the experiments. C.L., S.G. and S.N. performed the experiments; K.P.L. performed crystallographic calculations and model building; C.L., M.A. and K.P.L. analysed the data and wrote the manuscript.

Author Information Atomic coordinates and structure factors have been deposited with the Protein Data Bank under accession code 3RCE. Reprints and permissions information is available at www.nature.com/reprints. The authors declare competing financial interests: details accompany the full-text HTML version of the paper at www.nature.com/nature. Readers are welcome to comment on the online version of this article at www.nature.com/nature. Correspondence and requests for materials should be addressed to K.P.L. (locher@mol.biol.ethz.ch).

METHODS

***In vivo* complementation.** To analyse the activity of PglB from *C. lari in vivo* the gene encoding PglB was amplified from the *pgl* gene cluster of *Campylobacter lari* isolate (sample provided by H. Hächler) genomic DNA by polymerase chain reaction (PCR) and was cloned into a pMLBAD plasmid[48] with a C-terminal HA tag fused to PglB, resulting in the plasmid pMIK71 (ref. 28). For complementation studies, pMIK71 or pMLBAD empty vector were transformed into *E. coli* SCM6 cells carrying the plasmids pCL21 (ref. 47) and pACYCpgl_{mut} (ref. 25). pCL21 encodes for the expression of the single-chain Fv fragment of 3D5 carrying a DQNAT glycosylation site in the linker region and a C-terminal Myc tag fused to 3D5. pACYCpgl_{mut} codes for the biosynthesis of the *C. jejuni* lipid-linked oligosaccharide (LLO) with an inactivated *C. jejuni pglB* gene (W458A and D459A). A 5 ml pre-culture was inoculated from a single clone and grown over night at 37 °C in LB medium. The main culture was inoculated to an optical density (A_{600}) of 0.05 in 15 ml LB medium and grown at 37 °C to A_{600} of 0.5. The culture was induced by addition of arabinose to 0.1% (w/v) and grown for 4 h at 24 °C. For extraction of periplasmic proteins, an equivalent of 1 ml culture volume with an A_{600} of 3 was harvested by centrifugation, re-suspended in 150 µl extraction buffer, consisting of 30 mM Tris-HCl, pH 8.5; 20% (w/v) sucrose; 1 mM EDTA and 1 mg ml^{-1} lysozyme (Sigma) and incubated for 1 h at 4 °C. A final centrifugation step yielded periplasmic proteins in the supernatant. Glycosylation of 3D5 and expression of PglB were analysed by immunoblot following SDS–PAGE. Immunodetection was performed with anti-c-Myc monoclonal antibody (Calbiochem) and anti-glycan serum hR6 (S. Amber and M. Aebi, personal communication) to observe glycosylated 3D5. Immunodetection of *C. lari* PglB was performed with anti-HA antiserum (Santa Cruz). All experiments were performed at least in triplicate, and representative samples are shown.

Mutagenesis study. Mutant PglB was generated by the QuickChange method, and the resulting plasmids of all constructs were validated by DNA sequencing. The mutant PglB variants were cloned into pMLBAD as above and used in complementation assays.

PglB purification. The gene encoding PglB was cloned into a modified pBAD (Invitrogen) expression plasmid with a C-terminal decahistidine affinity tag fused to PglB, resulting in the plasmid pSF2. Owing to the applied cloning strategy, PglB carried the mutation K2E and the plasmid was confirmed by DNA sequencing (Microsynth).

PglB from *C. lari* was overexpressed from pSF2 in *E. coli* BL21-Gold (DE3) (Stratagene) cells in a 30 l fermentor (Infors). Cells were grown at 37 °C in Terrific Broth medium supplemented with 1% glycerol (w/v) to A_{600} of 10.0 before the culture was induced by the addition of 0.1% arabinose (w/v) for 2 h. All following steps were performed at 4 °C unless specified differently. Cells were harvested by centrifugation, re-suspended in 25 mM Tris-HCl, pH 8.0; 250 mM NaCl and disrupted in a M-110L microfluidizer (Microfluidics) at 15,000 p.s.i. chamber pressure. Membranes were pelleted by ultracentrifugation at 100,000*g* for 0.5 h. PglB was 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 h. All subsequent buffers contained DDM as detergent.

The supernatant was supplemented with 25mM imidazole and loaded onto a NiNTA superflow affinity column (Qiagen), washed with 60 mM imidazole before PglB was eluted with 200 mM imidazole. The protein was desalted into 10 mM MES-NaOH, pH 6.5; 100 mM NaCl; 0.5 mM EDTA; 3% glycerol (v/v); 3% polyethylene glycol 400 (v/v) and concentrated to 7–10 mg ml^{-1} in an Amicon Ultra-15 concentrator (Millipore) with a molecular mass cutoff of 100 kDa.

Native crystals. The peptide Ac-DQNATF{4-NO_2}-NH_2 was added to concentrated PglB to a final concentration of 0.75 mM, incubated for 0.5 h, and crystallized by vapour diffusion in sitting drops at 20 °C against a reservoir of 100 mM glycine, pH 9.4; 50 mM magnesium acetate; 6% dimethyl sulphoxide (DMSO) (v/v) and 23–34% (v/v) polyethylene glycol 400. The protein to reservoir volume ratio in the sitting drop was 2:1. Crystals typically appeared after 3–4 weeks and matured to full size within 6 weeks. Crystals were directly flash frozen by immersion in liquid nitrogen before data collection.

Heavy-metal derivatives. Native crystals were soaked for 30–60 min in 1 mM ethyl mercury phosphate (EMP) before back-soaking and flash-freezing by immersion in liquid nitrogen.

Data collection. Crystals belonged to the space group $P2_12_12_1$, with one PglB–peptide complex in the asymmetric unit. Native data were collected at the microdiffractometer beamline X06SA at the Swiss Light Source (SLS, Villigen) because not all sections of the crystals diffracted equally well. EMP2 and EMP3 derivative data sets (Supplementary Table 1) were collected at the same station, whereas EMP1 was collected at the high-resolution station of the same beam line. Data were processed and merged with XDS[49] or HKL2000 (HKL Research, Inc.) and anisotropic scaling/ellipsoid truncation was performed[50].

Structure determination. The structure was determined using a combination of molecular replacement using the periplasmic domain of *C. jejuni* PglB (Protein Data Bank code 3AAG) as a search model and Phaser[51] on the one hand, and multiple isomorphous replacement with anomalous scattering (MIRAS) using SHARP (Global Phasing Limited) on the other. The process of phase calculation and model building (using O[52]) and refinement (using Phenix[53]) was iterated, starting with the periplasmic domain and extending into the best-ordered regions of the transmembrane domain (TM1–4 and TM10–13), followed by TM5–9. The locations of three cysteines in the transmembrane domain (indicated by Hg anomalous peaks) served as starting points for tracing, until very good density allowed placement of bulky residues, confirming the sequence register. The final structure excludes two disordered loops of PglB (residues 283–306 and residues 605–607) as well as the C-terminal polyhistidine tag. Data collection and refinement statistics are given in Supplementary Table 1.

48. Lefebre, M. D. & Valvano, M. A. Construction and evaluation of plasmid vectors optimized for constitutive and regulated gene expression in *Burkholderia cepacia* complex isolates. *Appl. Environ. Microbiol.* **68,** 5956–5964 (2002).
49. Kabsch, W. XDS. *Acta Crystallogr. D* **66,** 125–132 (2010).
50. Strong, M. *et al.* Toward the structural genomics of complexes: Crystal structure of a PE/PPE protein complex from *Mycobacterium tuberculosis*. *Proc. Natl Acad. Sci. USA* **103,** 8060–8065 (2006).
51. McCoy, A. J. *et al.* Phaser crystallographic software. *J. Appl. Cryst.* **40,** 658–674 (2007).
52. Jones, T. A., Zou, J. Y., Cowan, S. W. & Kjeldgaard, M. Improved methods for building protein models in electron-density maps and the location of errors in these models. *Acta Crystallogr. A* **47,** 110–119 (1991).
53. Adams, P. D. *et al.* PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. D* **66,** 213–221 (2010).