A Transmembrane Crenarchaeal Mannosyltransferase Is Involved in N-Glycan Biosynthesis and Displays an Unexpected Minimal Cellulose-Synthase-like Fold

Rosaria Gandini†, Tom Reichenbach, Oliver Spadiut‡, Tien-Chye Tan§, Dayanand C. Kalyani and Christina Divne

School of Engineering Sciences in Chemistry, Biotechnology, and Health (CBH), KTH Royal Institute of Technology, 10691 Stockholm, Sweden

Correspondence to Christina Divne: divne@kth.se
https://doi.org/10.1016/j.jmb.2020.06.016
Edited by Georg Schulz

Abstract

Protein glycosylation constitutes a critical post-translational modification that supports a vast number of biological functions in living organisms across all domains of life. A seemingly boundless number of enzymes, glycosyltransferases, are involved in the biosynthesis of these protein-linked glycans. Few glycan-biosynthetic glycosyltransferases have been characterized in vitro, mainly due to the majority being integral membrane proteins and the paucity of relevant acceptor substrates. The crenarchaeote Pyrobaculum calidifontis belongs to the TACK superphylum of archaea (Thaumarchaeota, Aigarchaeota, Crenarchaeota, Korarchaeota) that has been proposed as an eukaryotic ancestor. In archaea, N-glycans are mainly found on cell envelope surface-layer proteins, archaeal flagellins and pili. Archaeal N-glycans are distinct from those of eukaryotes, but one noteworthy exception is the high-mannose N-glycan produced by P. calidifontis, which is similar in sugar composition to the eukaryotic counterpart. Here, we present the characterization and crystal structure of the first member of a crenarchaeal membrane glycosyltransferase, PcManGT. We show that the enzyme is a GDP-, dolichylphosphate-, and manganese-dependent mannosyltransferase. The membrane domain of PcManGT includes three transmembrane helices that topologically coincide with "half" of the six-transmembrane helix cellulose-binding tunnel in Rhodobacter spheroides cellulose synthase BcsA. Conceivably, this "half tunnel" would be suitable for binding the dolichylphosphate-linked acceptor substrate. The PcManGT gene (Pcal_0472) is located in a large gene cluster comprising 14 genes of which 6 genes code for glycosyltransferases, and we hypothesize that this cluster may constitute a crenarchaeal N-glycosylation (PNG) gene cluster.

Introduction

Across all three domains of life, protein glycosylation constitutes a critical post-translational modification that supports a vast number of biological functions in living organisms, such as protein stability, sorting, cell-cell recognition and signaling, transport, and dynamic adaptation to changing environments [1–4]. About two-thirds of all eukaryotic proteins are predicted to be glycosylated, and the majority of these carry asparagine-linked (N)-glycans [5]. While all three domains of life perform protein glycosylation, the archaeal glycosylation machineries appear more closely related to the eukaryotic systems [6], albeit highly blended with distinct traits of both bacteria and eukarya [7]. Especially N-glycosylation is more common in eukarya and archaea than in bacteria [6].

All N-glycans are assembled on a membrane-bound lipid carrier. Eukaryotes use dolichylphosphate (Dol-PP) as lipid carrier, archaea use either Dol-PP or Dol-P, and bacteria use undecaprenol monophosphate. In eukaryotes, the Dol-PP-linked heptasaccharide core (GlcNAc)2-(Man)5 of the lipid-linked oligosaccharide (LLO) is assembled at the cytoplasmic face of the ER membrane, and
translocation of the LLO to the luminal side where the glycan is further mannosylated and glucosylated to give the LLO tetradecasaccharide, \((\text{GlcNAc})_{2}-(\text{Man})_{9}-(\text{Glc})_{3}\), to be transferred \("en bloc\"\) by the oligosaccharyltransferase (OST) to the appropriate protein substrate [8]. The enzymes responsible for completing the heptasaccharide core are Alg7, Alg13/14, Alg1, Alg2 and Alg11, where the latter three are mannosyltransferases that transfer a mannosyl residue from GDP-\(\alpha\)-Man to an intermediate built on a chitobiose-charged Dol-PP-carrier lipid (Figure S1). After translocation of the heptasaccharide to the ER lumen, Alg3, Alg9 and Alg12 add four additional mannosyl residues, and Alg6, Alg8 and Alg10 each add one glucosyl units to complete the LLO tetradecasaccharide. The GTs resident on the ER-luminal face use Dol-\(\beta\)-Man to an intermediate built on a chitobiose-charged Dol-PP-carrier lipid and glucose (Dol-\(\beta\)-Glc) as activated glycosyl donors for the transfer reactions instead of nucleotide sugars.

Archaea produce a wider variety of glycans compared with eukaryotes, which is manifested as an equally impressive diversity of the glycan-biosynthetic enzymes [6,7]. As of May 2020, more than 10,000 gene sequences coding for archael glycosyltransferases (GTs), originating from 364 archael genomes, have been classified in the carbohydrate-active enzymes database in 31 distinct GT families (CAZY (www.cazy.org) [9]). In archaea, glycosylation occurs mainly on the cytoplasmic face of the plasma membrane, but can in some cases be further modified after the lipid-linked glycan has been translocated to the extracellular side for subsequent transfer to surface proteins; as for instance for \textit{Haloferax volcanii} AgIS, which transfers the terminal mannose from Dol-\(\beta\)-Man to the lipid-linked tetrasaccharide at the extracellular face of the membrane [10], and the post-transfer modification observed for the \textit{Sulfolobus solfataricus} LLO, which becomes extended by one hexose on the extracellular side [11]. The main substrates for protein \(N\)-glycosylation are typically cell envelope surface-layer proteins (S-layer proteins) and archaeilins (archaeal flagillins) [3,12,13].

Of the few archael \(N\)-glycan structures that have been characterized to date, one stands out as noteworthy from an evolutionary perspective. The biantennary high-mannose \(N\)-glycan (HMG) produced by the crenarchaeon \textit{Pyrobaculum calidifontis} bears, unlike other characterized archael \(N\)-glycans, distinct resemblance with respect to sugar composition to its eukaryotic HMG counterpart [14]. The \textit{P. calidifontis} \(N\)-glycan is an HMG built on a chitobiose-type core consisting of two \(\beta\)-1,4-linked modified GlcNAc

---

**Figure 1.** Overall fold of \(\text{PcManGT}\). The overall fold of \(\text{PcManGT}\) shown in two views. Coloring scheme: GT-A fold in orange, IF helices in green, substratum IF helix in violet, and the TM helices in light blue. The model of the \(\text{PcManGT}\cdot\text{Mn}^{2+}\cdot\text{GDP}\) complex is shown with GDP drawn as a stick object and the \(\text{Mn}^{2+}\) ion as a gray sphere. The \(\beta\)-strands are drawn as arrows and \(\alpha\)-helices as cylinders. The position of the disordered acceptor loop, which is protruding from IFH1, is indicated as red coil.

---

4659 Structure of crenarchaeal mannosyltransferase
Figure 2. ITC for nucleotides. Dissociation constants derived from ITC for the binding of nucleotides by PcManGT in the presence of Mn^{2+}. (a–d) Raw ITC thermograms and fitted binding curves for ADP (a), CDP (b), UDP (c), and GDP (d).
residues that are di-N-acetylated at the C2 and C3 positions (Glc(NAc)2), and with an additional carboxylate group at C6 of the second unit (GlcA(NAc)2) (Figure S1) [14]. The eukarya-lookalike HMG of *P. calidifontis* is particularly interesting since this crenarchaeote belongs to the archaeal TACK superphylum (comprising Thaumarchaeota, Aigarchaeota, Crenarchaeota and Korarchaeota), which has been proposed as an origin of eukaryotes [15,16].

Despite the evolutionary diversity of archaeal glycans, the underlying biosynthetic processes remain poorly understood, and few archaeal glycans and GTs have been characterized biochemically and functionally. Of the known archaeal GT genes, three-dimensional structural information for the gene products exists for only six unique enzyme activities, all of which are of euryarchaeal origin. Furthermore, of these few characterized enzymes, merely two represent integral membrane proteins, namely, the dolichylphosphate mannose synthase from the euryarchaeon *Pyrococcus furiosus* (PfDPMS) belonging to family GT2 [17], and the GT66-family member OST AglB from the euryarchaeon *Archaeoglobus fulgidus* [18]. Specifically, there are no biochemical or structural data available for any GT member belonging to the TACK superphylum.

Here, we present the first crystal structure and biochemical characterization of a glycosyltransferase of TACK origin, namely, the membrane glycosyltransferase PcManGT from *P. calidifontis*. We report the structures of PcManGT in the unliganded state, in complex with GDP·Mn2+ and with GDP-α-Man·Mn2+, as well as structural comparisons with related GTs, and discuss the possible function of PcManGT in N-glycan biosynthetic pathways of *P. calidifontis*.

**Results**

**Structure determination and overall fold**

The diffraction capacity of PcManGT crystals could be improved from 4 to 2.6 Å resolution by crystallization in meso using bicelles (i.e., bilayer micelles), which are open disk-like structures that offer an environment that is more akin to natural, native membranes. Structure determination was initially hampered by insufficient isomorphous or anomalous signal for the large number of screened data sets for heavy atom-exposed crystals, but experimental starting phases could eventually be estimated by the use of sodium-tungstate cluster (Table S1).
The overall fold of PcManGT features a catalytic extramembrane domain with the, for family-2 GTs, canonical GT-A fold (www.cazy.org) (Figures 1 and S2). As for membrane GTs [17], the catalytic domain is connected to two membrane-interface helices, IFH1 and IFH2. The transmembrane (TM) domain is composed of three antiparallel TM helices (TMH1–TMH3). TMH3 is short and is disrupted by a 90°-kink midway across the membrane, after which it continues roughly parallel with, but below, the membrane surface where it forms a substratum for IFH1 and IFH2 (Figures 1 and S2). Hence, we refer to this helix as the substratum interface helix (sub-IFH3).

Donor specificity and donor-crystal complexes

Isothermal titration calorimetry (ITC) was used to test the affinity of PcManGT for a relevant selection of diphosphonucleotides, diphosphonucleotide sugars and divalent metal cations. Of the four diphosphonucleotides (ADP, CDP, UDP and GDP) tested in the presence of the metal cofactor Mn$^{2+}$, dissociation constants could only be obtained for GDP (Figure 2 and Table S2). ITC analysis of GDP-$\alpha$-mannose (GDP-$\alpha$-Man) and GDP-$\alpha$-glucose (GDP-$\alpha$-Glc), both of which are candidate donors for PcManGT, revealed a 2-fold lower $K_d$ value for GDP-$\alpha$-Man in the presence of Mn$^{2+}$ (Figure 3 and Table S2). For GDP and GDP-$\alpha$-Man in the presence of Mn$^{2+}$, Mg$^{2+}$ or Ca$^{2+}$ indicated Mn$^{2+}$ as preferred divalent metal cation. Guided by the ITC results regarding the donor preference, PcManGT·Mn$^{2+}$ crystals were soaked with either GDP, GDP-$\alpha$-Man or GDP-$\alpha$-Glc. Crystals of PcManGT soaked with GDP-$\alpha$-Glc did not produce ordered electron density for the donor substrate; however, crystal complexes with GDP and GDP-$\alpha$-Man were determined at 2.7 Å resolution.

The guanosine ring in both the GDP (Figure 4(a)) and GDP-$\alpha$-Man (Figure 4(b)) complex forms hydrogen bonds via its N1 and N2 groups to Ser82, and the O6 oxygen is within interaction range of Arg112 Ne. The two ribose hydroxyl groups (O2’ and O3’) are involved in hydrogen bonding with the backbone amide of Tyr52, the backbone amide nitrogen of Pro50, and the side-chain of Glu54. Additional stacking occurs between the ring of Pro50 and the GDP ribose ring. The DXD motif in PcManGT is defined by Asp135 and Asp137. Asp137 coordinates the diphosphate group indirectly via the Mn$^{2+}$ ion, while Asp135 stacks with the ribose C4’–C5’ moiety of GDP. The importance of the divalent metal cation is confirmed by the ITC data where the presence of the Mn$^{2+}$ significantly increases the affinity of PcManGT for the substrate (Table S2).

In the GDP complex (Figure 4(a)), the $\beta$-phosphate oxygen atoms are stabilized by Arg261 Nh2 and Ne, and the Trp257 Ne1 nitrogen. Arg260 Nh2 is suitably positioned to interact with the bridging oxygen between the $\alpha$ and $\beta$-phosphate groups, but the Arg260 side-chain has less well-defined electron density making the assignment of these interactions more uncertain. In the GDP-$\alpha$-Man complex (Figure 4(b)), the phosphate oxygen atoms are only interacting with the manganese ion, while Arg261 interacts with the mannosyl group.

The overall electron density for the mannosyl group in the PcManGT·Mn$^{2+}$·GDP-$\alpha$-Man complex is somewhat featureless (Figure S3), probably due to local motion. Nonetheless, possible interactions with the protein can be predicted: Asp135 Od2, which is positioned within hydrogen-bonding distance of the mannosyl O5 oxygen; His264 Nd1 with the O4 hydroxyl group; and the Gly220 backbone carbonyl oxygen that can reach within hydrogen-bonding distance of the mannosyl O3 hydroxyl group (Figure 4(b)). Arg260 is still mainly disordered, but residual side-chain density...
indicates that it no longer interacts with the donor's phosphate groups. Compared with the two liganded PcManGT structures, we observe only minor positional shifts of the active-site side-chains in the unliganded structure.

Thermal stabilization by lipids

The effect of phospholipids and dolichol compounds on the stabilization to thermal unfolding of PcManGT was tested (Figure 5(a) and Table S3). The anionic 55- and 95-carbon dolichylmonophosphates (Dol55-P and Dol95-P) provided stabilization corresponding to an increased melting temperature ($\Delta T_m$) of 30 °C and 22 °C, respectively. Similarly, the anionic phospholipids phosphatidylserine (PS) and phosphatidylglycerol (PG) provided relative thermal stabilization of PcManGT by $\Delta T_m$ values of 21 °C (PS) and 24 °C (PG). While the neutral lipid phosphatidylethanolamine (PC) was marginally stabilizing, phosphatidylethanolamine (PE) did not show a stabilizing effect. PDPMS, which like PcManGT depends on dolichyl lipids for function, shows a similar, but weaker, trend for preferential stabilization by anionic lipids (Figure 5(b)). Dolichylphosphates do not play a functional role for the prokaryotic MFS transporter used as membrane protein control, and while there is a certain amount of stabilization, no distinct preference for the lipid head group or an isoprenoid-type chain is observed for this transporter (Figure 5(c)). As expected for the lipid-independent soluble control protein EcArnA, the addition of lipids did not provide any significant protection against thermal denaturation (Figure 5(d)).

Characterization of enzyme-dependent donor hydrolysis

Both retaining GTs [19,20] and inverting GTs [21] have been reported to be able to catalyze the destructive hydrolysis of its donor substrate in the

![Figure 5. Thermal stabilization assay. The effect of lipids on the thermal stability of (a) PcManGT, (b) PDPMS, (c) a prokaryotic transporter belonging to the MFS superfamily, and (d) EcArnA, a soluble protein from E. coli. The relative stabilization to thermal unfolding by the addition of lipids is expressed as the change in melting temperature ($\Delta T_m$). The lipids used were: anionic dolichylmonophosphates (Dol55-P, Dol95-P), anionic phospholipids (PG, PS); and neutral phospholipids (PC, PE). The Dol55-P and Dol95-P data for PDPMS are from Gandini et al. [17]. All experiments were performed at least twice ($n \geq 2$).]
absence of a natural acceptor. Based on the very strong relative stabilization of \(Pc\)ManGT by dolichylphosphate, we expect the natural acceptor to be an archaeal dolichylphosphate-linked sugar needed for synthesis of an LLO intermediate; however, Dol-\(P\)-linked sugars are not easily available. In the absence of suitable acceptors, we noted a distinct signal in reactions with GDP-\(\alpha\)-Man as sole donor substrate. We initially considered this phenomenon to be attributed to true \(Pc\)ManGT-dependent GDP-\(\alpha\)-Man hydrolysis, or to enzyme-independent spontaneous hydrolysis of GDP-\(\alpha\)-Man, or both. To investigate this further, GDP-\(\alpha\)-Man and GDP-\(\alpha\)-Glc were used as substrates in reactions using the GDP-GLO™ assay that is highly sensitive, and importantly, does not produce false positive results due to contaminating phosphate in the reagents used. GDP-\(\alpha\)-Man hydrolysis was monitored using wild-type \(Pc\)ManGT, and a mutant, D135A/D137A, where the metal-coordinating aspartate residues had been replaced by alanine. We tested the enzymes in the absence and presence of PS, Dol55-\(P\), Dol95-\(P\), and bicelles. The reason for testing both dolichylphosphates was to investigate whether the higher stabilization of \(Pc\)ManGT by Dol55-\(P\) (Figure 5(a)) is also manifested in higher catalytic activity. The reason for testing bicelles was to evaluate whether a bilayer-like environment provides an advantage for activity compared with only adding free phospholipids (PS).

The results show that spontaneous enzyme-independent hydrolysis is distinct, but comparably low for GDP-\(\alpha\)-Man, and close to zero for GDP-\(\alpha\)-Glc (Figure 6, bars 1 and 11), which highlights an inherent instability of GDP-\(\alpha\)-Man compared with GDP-\(\alpha\)-Glc under the reaction conditions used (for details, see Materials and Methods). \(Pc\)ManGT reactions with or without the addition of a natural anionic phospholipids (PS) show a 2-fold enzyme-dependent increase in GDP-\(\alpha\)-Man hydrolysis relative to solvent-mediated hydrolysis (Figure 6, bars 3 and 4). Further increase of enzyme-dependent hydrolysis is observed when Dol55-\(P\), or Dol95-\(P\), is added, irrespectively of the presence of PS (Figure 6, bars 5–6 and 8–9). The precise length of the dolichyl chain does not seem to influence the activity significantly when the enzyme is present only in a detergent solution.

The most noticeable results were obtained when performing the hydrolysis reactions in the presence of Dol-\(P\) and \(Pc\)ManGT incorporated in bicelles (Figure 6, bars 7 and 10). A remarkable increase in hydrolysis of the GDP-\(\alpha\)-Man donor is observed when \(Pc\)ManGT is incorporated into bicelles in the presence of a Dol55-\(P\) (Figure 6, bar 7), which corresponds to the natural length of the lipid carrier. Moreover, the results show that (i) the DXD mutant lacks detectable activity beyond spontaneous solvent-mediated donor hydrolysis (Figure 6, bar 2), which supports the importance of Asp135 and Asp137 for coordinating the donor substrate and metal-ion cofactor, respectively; and (ii) that GDP-\(\alpha\)-Glc is not a substrate for \(Pc\)ManGT (Figure 6, bars 12–13). In addition, we performed a preliminary attempt to identify possible acceptor substrates using soluble intermediates of the eukaryotic LLO-biosynthetic pathway (\(2\alpha\)-mannobiose, \(3\alpha\)-mannobiose, \(6\alpha\)-mannobiose, \(3\alpha,6\alpha\)-mannotriose, and chitobiose; Figure S4). Possible product formation was monitored using HPAEC-PAD analysis (Figure S5); however, no products were generated in the presence of soluble

Figure 6. Influence of different lipids on enzyme-dependent donor hydrolysis. Analysis using the GDP-Glo™ assay of solvent-mediated (1 and 11) and enzyme-dependent hydrolysis of the donor substrate GDP-\(\alpha\)-Man (2–10) or GDP-\(\alpha\)-Glc (12–13) by \(Pc\)ManGT. All reactions were carried out in Hepes buffer (pH 7.5) in the presence of Mn\(^{2+}\) and LAPAO.
sugar intermediates, and only GDP-α-Man hydrolysis was observed.

Comparison with related enzymes

In a search for the experimental 3-D structure most similar to that of full-length PcManGT structure, the DALI server (http://ekhidna2.biocenter.helsinki.fi/dali/) returned the catalytic cellulose synthase domain BcsA of Rhodobacter sphaeroides cellulose synthase (PDB 4HG6) [22] as the closest match with an r.m.s.d. value of 3.2 Å for 325 aligned residues and a resulting sequence identity of 19%. A DALI search against only the catalytic extramembrane domain of PcManGT gave the best match for chondroitin synthase (PDB Z286), resulting in an r.m.s.d. of 2.7 Å for 204 aligned residues corresponding to a sequence identity of 20%. These global measures of similarity highlight that PcManGT is distinct from GTs with known structure.

A structure-based alignment of PcManGT with the few available transmembrane GT2 members with known 3-D structure emphasizes the similarities and differences (Figures S6 and S2). The most similar protein, the synthase subunit RsBcsA (PDB 4HG6) [22], can be divided into three principal parts: the N-terminal region (residues 1–60) folds as two TM helices (TMH1, TMH2) that coil with the C-terminal TM helix of the BcsB subunit to form a bundle; a cellulose synthase region with an extramembrane catalytic GT-A domain connected to a TM domain with six TM helices (TMH3–TMH8) that form a narrow cellulose channel (residues 63–123, 404–467, 522–572); a PilZ domain with a β-barrel fold (residues 583–675) linked to an extensive C-terminal α-helical region that wraps around the GT-A domain like a collar (residues 676–759). Since PcManGT does not require either a BcsB or PilZ domain, only the cellulose synthase region of BcsA (residues 64–572) is relevant for a structural comparison.

The PcManGT structure maps to the region in BcsA defined by residues 96–496 (Figures S6 and S2). PcManGT and BcsA shares the same triangular arrangement of IFH1, IFH2 and the substratum IF helix (sub-IFH3, residues 474–496). With respect to the arrangement of TM helices, the three TM helices in the TM domain of PcManGT can be regarded as one half of the six-TMH domain in BcsA that forms the cellulose-binding channel (i.e., TMH3–TMH8) (Figure S7). By this comparison, TMH1, TMH2 and TMH3 in PcManGT correspond to TMH4, TMH5 and TMH6 in BcsA, respectively. However, a positional shift of PcManGT–TMH3, and the shorter loop preceding it, causes PcManGT–TMH3 to superimpose with BcsA–TMH6 rather than BcsA–TMH8. Furthermore, PcManGT lacks the sequence pattern [D,D,D,Q(Q/R)XRW] that is typical for processive GTs [23], emphasizing that PcManGT does not act as a sugar polymerase. The presumed lack of processivity and the topologically simpler “half-channel”-TM-domain motif in PcManGT (Figure S7A) would be in good agreement with an enzyme that catalyzes a single glycosyl-transfer reaction to a Dol-PP-linked acceptor sugar during N-glycan biosynthesis in P. calidifontis.

In PDPMs, the A-loop is defined as an extensive connection between the last β-strand of the GT-A domain (β7 and IFH2 (Figure S2C), which displays a particularly high sensitivity to the absence or presence of acceptor substrate [17]. Typically, this loop is ordered in the absence of acceptor, and once the glycolipid product is bound, the IF helices are separated and pull the A-loop into an open and disordered state. The A-loop in PDPMs is directly tethered to the metal center via Arg202, which provides a rationale for the release of the A-loop in the post-catalysis state when the metal has departed.

In PcManGT, the β7/IFH2 loop appears well ordered, but instead, the loop β5/IFH1 (residues 168–177) shows disorder. PcManGT lacks a counterpart to Arg202 in DPMS and thus does not seem to couple IFH2 to the presence of metal ion, acceptor substrate or product. Interestingly, BcsA has a corresponding loop (residues 278–297) that closes as a lid over the triangular opening between IFH1, IFH2 and the sub-IFH3. This opening leads directly to the position of a presumed third glucosyl-binding subsite at Phe301 in IFH1 and Trp417 in TMH5 (corresponding in PcManGT to Tyr181 in IFH1 and His293 in TMH2) (Figure S8A). A similar opening exists in PcManGT (Figure S8B), and we anticipate that the A-loop, once ordered, presumably in the presence of acceptor, would fold over this opening site, shielding the transfer site in a way similar to that seen in BcsA. Regardless of whether the A-loop is located at β7/IFH2 as in PDPMs and GtrB, or at β5/IFH1 as in PcManGT and RsBcsA, the common theme is that the loop connects to either of the two IF helices, which emphasizes the functional role of IF helices to regulate lipid-linked acceptor association/dissociation in transmembrane GTs.

Hypothetical N-glycosylation gene cluster in P. calidifontis

Biosynthesis of the eukaryotic LLO heptasaccharide core requires three mannosyltransferases Alg1 (GT33), Alg2 (GT4) and Alg11 (GT4), whereas the complete LLO requires three additional non-Leloir mannosyltransferases (Alg3, Alg9 and Alg12). We analyzed the P. calidifontis genome for Leloir-mannosyltransferase genes, specifically GT counterparts for alg1, alg2, and alg11. The P. calidifontis genome contains 21 GT genes: 9 GT2, 6 GT4, 1 GT5, 1 GT20, 1 GT35, 1 GT66 and 2 GTNC, where all but GT2 and GT66 are retaining. Most of the GT genes are scattered throughout the genome, with one exception: a 14-gene cluster of 13,404 bp containing the genes Pcal_0470 to Pcal_0483.
This gene cluster is the largest predicted cluster in the *P. calidifontis* genome (Table S4) and contains six GT genes: three GT2 (Pcal_0472, Pcal_0476, Pcal_0481) and three GT4 (Pcal_0473, Pcal_0475, Pcal_0483). Thus, half of all GT4 genes are present in this cluster. Considering that several of the mannosyltransferases in the euryarchaeal HMG-biosynthetic pathway are GT4 enzymes, the presence of three GT4 genes in this cluster is interesting and raises the possibility that it may constitute a protein N-glycosylation gene (PNG) cluster.

Alg1 adds the first mannosyl unit to the chitobiose core of the LLO and belongs to family GT33 of inverting mannosyltransferase. There is no GT33 gene in the *P. calidifontis* genome, which means that the enzyme adding the α-linked mannose to the chitobiose-like core belongs to a different GT family. The gene coding for *PcManGT* (Pcal_0472) could be a possible candidate for Alg1. Since *PcManGT* and Alg1 belong to different GT families, it is not meaningful to compare their sequences; however, both display a relatively simple transmembrane architecture: *PcManGT* has three TM helices; and Alg1, whose structure is unknown, is predicted to have 1–3 TMHs depending on the prediction software used (TOPCONS; http://topcons.cbr.su.se) [24]. Two GT4 genes (Pcal_0473 and Pcal_0483) in this cluster could be possible candidates for the euryarchaeal *alg2* gene (Table S4). The gene Pcal_2037 has been proposed to be an *alg11* ortholog [14] and is 23.8% identical in amino-acid sequence to human Alg11. While Pcal_2037 is not present in the proposed PNG cluster, Pcal_0483 shows 23.1% sequence identity to human Alg11 and could be an alternative. The GT responsible for transferring the N-glycan onto the protein asparagine substrate is referred to as AglB in archaea (OST in eukaryotes) and belongs to GT family 66. The only GT66 gene in *P. calidifontis* genome, Pcal_0997, is the likely candidate for this activity. Hyperthermophilic and crenarchaeal GT genes responsible for N-glycan biosynthesis do not typically display *aglB*-based *N*-glycosylation-gene clustering as in euryarchaeal genomes [25,26], and in agreement with this observation, Pcal_0997 is distantly located some 470,000 bp downstream from the proposed PNG gene cluster.

**Discussion**

Heretofore, the enzymatic glycosylation machinery of *P. calidifontis* has not been studied experimentally, but its N-glycan structure has been characterized and shows resemblance to the eukaryotic HMG-type N-glycan (Figure S1) [14]. Among the archaeal LLOs characterized to date, that of *P. calidifontis* is comparably large, comprising 11 sugar units with the sequence Asn-Glc(NAc)_{2}-GlcA(NAc)_{2}-(Man)$_{9}$. An α-1,4-linked mannosyl unit is attached directly to the diacetylated chitobiose core that branches into α-1,3- and α-1,6-mannosyl-linked antennae, an arrangement similar to that of the eukaryotic triantennary HMG structure. The α-1,6-mannosyl branch is extended by four α-1,2-linked mannose units and a terminal α-1,3-linked mannose, and the α-1,3-mannosyl branch is capped by one α-1,2-linked mannose unit. Beyond the chitobiose core, the *P. calidifontis* biantennary N-glycan structure consists exclusively of α-linked mannosyl residues.

Mannosyltransferases capable of catalyzing the formation of glycosidic bonds in α-configuration are either retaining GTs in the cytoplasm that use GDP-α-Man as donor, or non-cytoplasmic inverting GTs that use Dol-P-β-linked mannosyl as donor substrate. The amino-acid sequence of *PcManGT* has been classified in the CAZy database [9] as belonging to the GT2 family of inverting glycosyltransferases. This classification infers that the enzyme acts by inversion of configuration at the anomeric center, meaning that transfer of mannosyl from the donor GDP-α-Man will result in a mannosylated product with the mannosyl attached by a β-configured glycosidic bond. Our results show that GDP-α-Man is the natural donor substrate for *PcManGT*, and that the enzyme catalyzes the hydrolytic side-reaction in the absence of acceptor. The ability of *PcManGT* to readily accept water as acceptor in enzyme-catalyzed hydrolysis of the donor substrate is a feature normally associated with a retaining reaction mechanism. Whether *PcManGT* is indeed a bona fide GT2 member, or possibly a retaining GT2-like enzyme raises a relevant question for further studies.

Like eukaryotes, *P. calidifontis* assembles its nascent N-glycans on the membrane-bound lipid carrier Dol-PP. The Dol-PP produced by *P. calidifontis* has an average isoprenoid chain length of 50–55 carbon [11], while the isoprenoid chains of eukaryotic Dol-PP are longer and less saturated. In contrast to crenarchaea and eukaryotes, the LLOs of euryarchaea are assembled on single-phosphorylated dolichol carriers, Dol-P. The dependency of *PcManGT* on the isoprenoid chain length for structural stability and catalytic activity is evidenced by the ability of Dol55-P to stabilize the enzyme against unfolding to provide a notable increase in melting temperature of 30 °C (Figure 5(a)). Furthermore, Dol55-P added to a membrane-mimicking environment (bicelles) increases donor hydrolysis by almost 5-fold compared to an environment containing only detergent, or detergent with the supplementation of the phospholipid PS (Figure 6, bars 3, 4 and 7).

Several membrane proteins have been reported to depend on anionic lipids for function and stability [27,28]. Our results add important new knowledge regarding of the function of protein–lipid interactions by establishing that beyond the general stabilization
by anionic lipids, dolichylphosphate-dependent membrane proteins (e.g., \textit{Pc}ManGT and \textit{Pd}PMS) are also specifically stabilized by dolichylphosphates. This behavior was not observed for a dolichylphosphate-independent membrane protein (MFS transporter), or for soluble proteins. Considering the importance of the identity of the lipid-carrier, the lack of transfer products from the soluble acceptors tested here is not surprising. For example, in the case of \textit{Pd}PMS, we showed that precise positioning of the phosphate moiety of Dol-P is critical for the mannosyltransfer reaction \cite{[17]}

Taken together, these observations convincingly point to a Dol-\textit{PP}-linked sugar intermediate as the natural acceptor substrate for \textit{Pc}ManGT. Furthermore, we have shown by sequence and structure comparisons, ITC, crystal-structure analysis and donor-hydrolysis assays that \textit{Pc}ManGT is a Leloir glycosyltransferase that uses GDP-\textit{α-Man} as donor substrate, in contrast to non-Leloir GTs that use other types of activated glycosyl donors. This further implies that the transfer reaction catalyzed by \textit{Pc}ManGT would need to take place on cytoplasmic face of the \textit{P. calidifontis} plasma membrane.

That GDP-\textit{α-Man} serves as the natural donor substrate and that the enzyme requires a Dol-\textit{PP}-linked sugar acceptor establish that \textit{Pc}ManGT acts as a Dol-\textit{PP}-dependent mannosyltransferase, and thereby assigns the enzyme to the \textit{P. calidifontis} \textit{N}-glycan-biosynthesis pathway. Thus, the function would be to add a mannosyl residue to one of the mannosyl positions in the \textit{P. calidifontis} HMG intermediate. The cryo-electron microscopy structure of \textit{Saccharomyces cerevisiae} Alg6 was reported recently \cite{[29]}. Alg6 is an inverting GT non-Leloir enzyme that adds the first \textit{α}-1,3-linked glucose moiety to the long antenna of yeast \textit{N}-glycans (Figure S1). Rather than using a nucleotide sugar, Alg6 uses Dol-\textit{P}-Glc as the activated glycosyl donor, from which glucose is transferred to Dol-\textit{PP}-(\textit{GlcNAC})\textsubscript{2}-(\textit{Man})\textsubscript{6} on the lumenal side of ER. Alg6 has a GT-C fold that is unrelated to the GT-A fold of \textit{Pc}ManGT but shares the functional similarity of depending on a dolichyl-linked substrate.

Docking of a modeled Dol-\textit{PP} acceptor molecule to the TM domain in \textit{Pc}ManGT suggests that a dolichyl isoprenoid chain is likely to adopt a similar kinked conformation as observed for Dol-\textit{P}-Glc in Alg6. Furthermore, adding sugar units to the docked Dol-\textit{PP} molecule indicates that two sugar moieties corresponding to the chitobiose-type core can be accommodated (Figure S9). If this is the case, a mannosyl unit would be transferred from GDP-\textit{α-Man} to Dol-\textit{PP}-Glc(\textit{NAC})\textsubscript{2}-Glc\textit{A}(\textit{NAC})\textsubscript{2}, and thereby assign a role to \textit{Pc}ManGT equivalent to the eukaryotic chitobiosylidophosphodolichol mannosyltransferase Alg1. However, this hypothesis presupposes that \textit{Pc}ManGT is a retaining GT, or that the assignment of the chitobiose-mannosyl bond should be in \textit{β}-configuration rather than \textit{α}.

The assignment of \textit{Pc}ManGT to the \textit{N}-glycosylation pathway is further strengthened by the observation that the gene \textit{Pcal_0472} is present in a large 14-gene cluster in the \textit{P. calidifontis} genome that includes six GT genes (Table S4). Of the GT genes, four are GT4 enzymes, including possible candidates for the \textit{α}-1,3/1,6-mannosyltransferase \textit{Alg2} (\textit{Pcal_0473}, \textit{Pcal_0483} and Alg7) that adds the first GlcNAc unit to the Dol-\textit{PP} carrier (\textit{Pcal_0478}). Thus, it is likely that the gene cluster constitutes a crenarchaeal \textit{N}-glycosylation (PNG) gene cluster.

**Materials and Methods**

**Cloning and mutagenesis**

The gene coding for \textit{P. calidifontis} JCM 11548 glycosyltransferase 2 (\textit{Pc}ManGT; ABO07903.1; UniProt A3MTD6) was synthesized and codon-optimized by MR.GENE for host expression in \textit{Escherichia coli}. The gene insert was amplified using the following primers:

\textit{Pc}ManGT\_fwd (5′-\textit{AAGGAGATATACCA}\textit{TATGCTGCTGGAACATCGTACCGC}-3′)\textit{Pc}ManGT\_rev (5′-\textit{GAAGTACAGGTTCTGAGCCGGTTTTTTCCCTGTGTTTTCGCCAGACCCAG}-3′).

A PCR protocol using 70 ng plasmid DNA and Phusion High-Fidelity (Thermo Scientific) included 98 °C for 30 s, 30 cycles of 98 °C for 10 s, 65 °C for 30 s, 72 °C for 90 s with a final incubation at 72 °C for 10 min. Following the PCR products were purified using GeneJET PCR Purification Kit (Thermo Scientific). The amplified gene was cloned using restriction enzymes into the pWaldo-GFPd vector \cite{[30,31]}, which was digested with NdeI and BamHI before being ligated into the pWaldo-GFPd vector \cite{[30,31]} and chemically competent \textit{E. coli} BL21(DE3) cells. Five colonies were picked for screening by colony PCR with the insert primers. Site-directed mutagenesis to generate a catalytically deficient \textit{Pc}ManGT mutant (D135A/D137A) was performed using the following forward and reverse primers:

\textit{D135A/D137A}\_fwd (5′-\textit{CATCACTGCTGCCGTGCTG}\textit{CCCTGTGACCCCGG}-3′)\textit{D135A/D137A}\_rev (5′-\textit{CCCGCCACAGGGCGCCGCAGCAGCAGCAGCAGT}-3′).

Mutagenesis was performed using the QuikChange II site-directed mutagenesis kit (Agilent Technologies), as described by the manufacturer. The sequence of the mutated gene was confirmed by automated DNA sequencing, and plasmid with desired mutations was transformed into \textit{E. coli} BL21 (DE3) for protein production.
Gene expression and solubilization of membrane proteins

Transformed cells were grown on a Lysogeny Broth (LB; Sigma Aldrich) agar plate and incubated at 37 °C. Positive transformants were selected with 50 μg/ml kanamycin. An overnight LB seed culture (50 ml) was used to inoculate Terrific Broth (TB; Sigma Aldrich) medium (5 L) containing 50 μg/ml kanamycin. At an OD600 value of 1.0, overexpression of wild-type or D135A/D137A variant PcManGT-TEV-GFP-His8 was induced by adding IPTG to a 0.2 mM concentration and continued for 15–17 h at 18 °C. The cells were harvested in phosphate-buffered saline supplemented with 1 tablet of cOmplete™ Protease Inhibitor Cocktail (Merck). The bacterial cells were lysed using an Avestin EmulsiFlex C3 homogenizer, followed by sonication to fragment nucleic acids. The cell lysate was centrifuged at 10,000 r.p.m. (12,096 g, Avanti J-20XP centrifuge) for 10 min at 4 °C. The supernatant was centrifuged at 35,000 r.p.m. (125,749 g, 45Ti rotor, Optima LE-80K centrifuge, Beckman Coulter) at 4 °C for 1 h to separate the membrane fraction.

The membrane fraction was resuspended in phosphate-buffered saline, centrifuged a second time at 35,000 r.p.m., and resuspended in 25 mM K2HPO4 buffer (pH 7.2), 150 mM NaCl, 5% glycerol (v/v), and 1.5% detergent (w/v) (Anatrace, Molecular Dimensions) to solubilize membrane proteins. Detergents used were N-decyl-β-D-maltopyranoside (DM), 3-dodecylamido-N,N'-dimethylpropyl amine oxide (LAPAO), or N,N-dimethyldecylamine N-oxide (LDAO). Extracted membrane proteins were recovered in the supernatant by centrifugation at 10,000 r.p.m. (12,096 g, 45Ti rotor, Optima LE-80K centrifuge, Beckman Coulter) at 4 °C for 1 h to separate the membrane fraction.

Protein purification

The supernatant containing the membrane proteins was incubated with Ni-NTA agarose resin (Thermo Fisher scientific, packed in Bio-Rad Econo-Pac® columns) equilibrated in 25 mM K2HPO4 buffer (pH 7.2), 150 mM NaCl, 30 mM imidazole, 5% glycerol (v/v) for 90 min at 4 °C. The columns were first washed with 25 mM K2HPO4, (pH 7.2), 150 mM NaCl, 30 mM imidazole, 5% glycerol (v/v), 0.2% DM (or 0.3% LDAO, 0.08% LAPAO, w/v), then with 25 mM K2HPO4 buffer (pH 7.2), 300 mM NaCl, 50 mM imidazole, 5% glycerol (v/v), 0.2% DM (or 0.2% LDAO, 0.07% LAPAO, w/v). Bound protein was eluted with 25 mM K2HPO4 buffer (pH 7.2), 150 mM NaCl, 500 mM imidazole, 5% glycerol (v/v), 0.2% DM (or 0.2% LDAO, 0.07% LAPAO, w/v).

To remove GFP-His8, PcManGT-TEV-GFP-His8 (wild-type or D135A/D137A variant) was mixed with 0.5 mM EDTA, 0.5 mM DTT, and tobacco etch virus (TEV) protease [32], with a TEV to protein molar ratio of 1:10, and incubated overnight at 4 °C. PcManGT was recovered as flow through from reverse IMAC. When necessary the protein was treated with 10 mM ethylenediaminetetraacetic acid (EDTA) and 10 mM ethylene glycol-bis(β-aminoethyl ether)-N,N,N',N'-tetraacetic acid (EGTA). The sample was concentrated with Amicon Ultra-15, MWCO 100 kDa and loaded onto a HiLoad™ Superdex™ 200 16/60 column (GE Healthcare Life Sciences) that had been equilibrated in 50 mM 4-(2-hydroxyethyl)-1-piperazinethanesulfonic acid (Hepes; pH 7.5), 150 mM NaCl, 5% glycerol (v/v), 0.1% DM (or 0.05% LDAO, 0.07% LAPAO, w/v). Fractions containing PcManGT were pooled and concentrated to 5.5 mg/ml (to 10–25 mg/ml for the batches in LDAO and LAPAO). Typical yields from 1 g of wet cell pellet were approximately 0.3 mg for wild-type PcManGT and 0.25 mg for the D135A/D137A variant.

Crystallization and data collection

PcManGT stored in 50 mM Hepes (pH 7.5), 150 mM NaCl, 5% glycerol (v/v), and 0.1% DM (w/v) at a concentration of 5.5 mg/ml was used for crystallization. Briefly, 40% bicine mixture (DMPC: CHAPSO, 2.8:1, MemMagic™ Bicelle Screen kit, Molecular Dimensions) was mixed with the protein to a final concentration of 8% bicides, and incubated 30 min on ice. A volume of 1 μl of protein-bicelle solution was mixed with 0.3 μl of reservoir containing 0.1 mM Hepes (pH 7.5), 20% PEG 4000, 10% isopropanol, 10 mM MnCl2 and 2% 1,2,3-heptanetriol, and the plates (Greiner BIO-ONE CombiPlate™) were stored at 22 °C.

For the crystals of PcManGT in complex with GDP-ManβS and GDP-α-ManβS, the crystals were soaked for 5 min in the presence of either 10 mM GDP or GDP-α-Man. For experimental phasing, a PcManGT crystal was soaked for 5 min with 10 mM Na2WO4.

Intensity data for all crystals, derivatized and non-derivatized, were recorded using synchrotron radiation (Table S1), followed by merging and scaling using the XDS package [33]. The protein crystallized in space group P212121 with two molecules in the asymmetric unit and approximately 50% solvent content. Heavy-atom substructure solution and phasing of amplitudes (Table S1) were performed using a crystal derivatized with Na2WO4 using single isomorphous replacement with anomalous scattering (SIRAS) as implemented in autoSHARP [34] included in the SHARP package [35]. The electron density was improved and phase-extended to 2.7 Å by solvent flattening using the program SOLOMON [36] implemented in SHARP with an optimized solvent fraction of 0.502. Non-crystallographic symmetry (NCS) operators were deduced by manual inspection of the improved electron density in O [37], and used for 2-fold NCS averaging, solvent modification and extension of the modified
**Isothermal titration calorimetry**

ITC was used to identify a possible donor substrate for PcManGT. The strategy employed was to first analyze the interaction with nucleotide diphosphates (adenosine diphosphate, ADP; cytidine diphosphate, CDP; uridine diphosphate, UDP; guanosine diphosphate, GDP), and second, to evaluate possible nucleotide diphosphate sugars guided by the nucleotide preference obtained. Based on the resulting binding data, GDP-α-Man and GDP-α-Glc were selected for ITC analysis.

To avoid buffer mismatch, the enzyme was dialyzed against 50 mM Hepes (pH 7.5), 150 mM NaCl, 5% glycerol (v/v), 0.1% DM (w/v), and the same dialysis buffer was used for dissolving the substrates. Substrate binding was analyzed in the absence and presence of the divalent metal cation Mn²⁺, Mg²⁺ or Ca²⁺, with MnCl₂, MgCl₂, or CaCl₂ added at a final concentration of 1 mM to enzyme and substrate solutions. The measurements were performed using 50 μM of PcManGT and 500 μM of substrates (corresponding to 10 × molar excess). Titrations were performed by 25 injections using a MicroCal ITC200 system (GE Healthcare), and the raw data were fitted with the software Origin 7.0.

**Thermal stability assay**

To assess thermal stability of PcManGT, a thermal stability assay (TSA) was performed. Samples containing 40 μl of purified protein at 0.25 mg/ml (0.0063 mM) were incubated in the absence and presence of different lipids (0.1 mM) for 10 min at temperatures ranging from 45 °C to 100 °C in a PCR thermal cycler (Veriti 96-Well Thermal Cycler with VeriFlex™ blocks, Applied Biosystems™). The lipids tested were 1-palmitoyl-2-oleoyl-sn-glycero-3-phospho-(1'-rac-glycerol) 16:0–18:1 (PG), L-α-phosphatidylserine (PS), 1-palmitoyl-2-oleoyl-sn-glycero-3-phosphatidylcholine 16:0–18:1 (PC), or 1-palmitoyl-2-oleoyl-sn-glycero-3-phosphatidylethanolamine 16:0–18:1 (PE), C55-dolichyl monophosphate (Dol55-P), or C95-dolichyl monophosphate (Dol95-P). PG, PS, PC and PE were purchased from Avanti Polar Lipids, Inc., and the dolichols from Larodan Research Grade Lipids.

Following heat treatment, precipitated protein was removed from the samples by centrifugation at 14,500 r.p.m. (19,780 g, F-45-24-11 rotor, Eppendorf 5424 Microcentrifuges, Eppendorf) for 10 min at 4 °C. The supernatants, containing the protein remaining in solution, were analyzed by Coomassie-stained SDS-PAGE (Coomassie Brilliant Blue R-250). Densitometric analysis of the gel bands was performed with ImageJ (Rasband, W.S., ImageJ, http://imagej.nih.gov/ij/, 1997–2016). Prism 7 (GraphPad Software, San Diego California USA) was used to fit the data in the unfolding curves, and to determine the Tₘ ± SD and the ΔTₘ ± SEM values.

Three proteins were used as references: two membrane proteins and one soluble protein. PDPMS [17] was used as a reference for a dolichylphosphate-dependent membrane protein; a transporter belonging to the major facilitator superfamily (MFS) was used as a reference for a dolichylphosphate-independent membrane protein; and the soluble protein was used as a reference for a lipid-independent protein. The stability assays were performed as described for PcManGT.

**Activity assay**

The enzymatic activity of PcManGT was tested with GDP-α-Man, and GDP-α-Glc (Sigma-Aldrich) as glycosyl donors. The dolichylmonophosphates were first dried under a low flow of nitrogen gas for 2–3 h to remove remaining methanol and chloroform. After drying, the dolichylmonophosphates were dissolved in 50 mM Hepes (pH 7.5), 150 mM NaCl and 0.07% LAPAO (w/v) while stirring for 2–3 days and occasional incubation in a sonication bath for 30 min. The phospholipids were dissolved in water. Reactions with a volume of 50 μl contained 25 μM PcManGT (in 50 mM Hepes (pH 7.5), 150 mM NaCl, 5% (v/v) glycerol, 0.07% LAPAO (w/v)), 0.3 mM donor substrate, and 10 mM MnCl₂. When applicable PcManGT was pre-incubated with 0.2 mg/ml PS, 100 μM Dol55- P, 100 μM Dol95- P and/or 1.2% bicelles (v/v) for 30 min on ice.

The reaction samples were incubated at 60 °C for 3 h. Following incubation, precipitated protein was removed from the samples by centrifugation at 13,000 r.p.m. (15,900 g, F-45-32-5-PCR rotor, Eppendorf 5424 Microcentrifuges, Eppendorf). To quantify the hydrolytic activity, we used a protocol based on the GDP-Glo™ Glycosyltransferase Assay (Promega Biotech AB). From each sample, 25 μl was transferred to a 96-well plate (Sarstedt96, solid white), and an aliquot of 25 μl of GDP detection...
reagent (prepared according to supplier information, GDP-Glo™ Glycosyltransferase Assay, Promega Biotech AB) was added to each sample. The dye-containing samples were incubated at room temperature for 60 min, followed by recording the luminescence at 555 nm using a microplate reader (Omega Fluostar; BMG Labtech). All samples were run in triplicates and from at least two different protein batches. The data were analyzed with Prism 8 (GraphPad Software, San Diego California USA) and displayed as mean values ± SEM.

**HPLC analysis of products**

The donor substrate hydrolysis was analyzed by High-Performance Anion Exchange Chromatography and Pulsed Amperometric Detection (HPAEC-PAD), using a Dionex ICS 5000 system (Dionex, Sunnyvale, CA) with an autosampler, equipped with a PA-1 column and guard cartridge. The enzymatic reactions were performed as described above for the activity assays with GDP-α-Man as donor substrate, Dol95-P as stabilizing lipid (used instead of Dol55-P since the longer dolichol is more easily available from Larodan), and with the addition of different possible soluble acceptor substrates (Figure S4): α-D-Man-[1→2]-D-Man (2α-Mannobiose; Carbosynth Cat. No. OM05906); α-D-Man-[1→3]-D-Man (3α-Mannobiose; Carbosynth Cat. No. OM00767); α-D-Man-[1→6]-D-Man (α-Mannobiose; Carbosynth Cat. No. OM06489); α-D-Man-[1→3]-(α-D-Man-[1→6]-D-Man (3α,6α-Mannotriose; Carbosynth Cat. No. OM05762); and N,N-Diacetyldichitobiose (chitobiose; Sigma Aldrich Cat. No. D1523). The following solvents were used as eluents: 1 mM sodium hydroxide (E1) and 1 M sodium acetate in 1 mM sodium hydroxide (E2). An aliquot of the enzymatic reaction (10 μl) was injected into a CarboPac PA-1 column equilibrated (E2). Elution was performed by the following gradient: T0 = 20% (v/v) E2; T10 = 55% (v/v) E2; T15 = 55% (v/v) E2; T20 = 80% (v/v) E2; T25 = 100% (v/v) E2; T30 = 100% (v/v) E2.

**Data deposition**

The atomic coordinates and structure factor amplitudes have been deposited in the Protein Data Bank (www.rcsb.org) with the accession codes 6YV7 (PcManGT), 6YV8 (PcManGT-Mn2+-GDP), and 6YV9 (PcManGT-Mn2+-GDP-α-Man).

**Acknowledgments**

We thank the synchrotron beamline staff at Diamond Light Source beamline I24 and I02 (UK),SOLEIL beamline PROxima1 (France); ESRF Grenoble beamline ID23-1 (France), and Swiss Light Source beamline PXI X06SA (Switzerland) for technical assistance during data collection. We also thank Michael Raba for supplying the MFS transporter used as control. We acknowledge funding to R.G., T.C.T and C.D. from the Swedish Research Council FORMAS Grant 2008-495; the Swedish Research Council VR Grants 2011-5768, 2013-5717 and 2017-03877; and the Carl Trygers Foundation Grant CTS08:78. The research leading to these results has received funding from the European Community’s Seventh Framework Programme FP7/2007-2013 under BioStruct-X Grant agreement 283570. The work was facilitated by the Protein Science Facility at Karolinska Institutet/SciLifeLab (http://psf.ki.se).

**Appendix A. Supplementary data**

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jmb.2020.06.016.

Received 6 May 2020; Received in revised form 13 June 2020; Accepted 16 June 2020
Available online 19 June 2020

**Keywords:**
mannosyltransferase; crenarchaea; protein N-glycosylation; crystal structure; membrane-protein stabilization†Present address: R. Gandini, Department of Bioscience, Aarhus University, 8000 Aarhus C, Denmark.
†Present address: O. Spadiut, Institute of Chemical Engineering, Vienna University of Technology, 1070 Vienna, Austria.
‡Present address: T.-C. Tan, Institute of Structural and Molecular Biology, Birkbeck University of London, WC1E 6BT London, UK.

**Abbreviations used:**

ADP, adenosine diphosphate; CAZy, Carbohydrate-Active enZymes database; CDP, cytidine diphosphate; DM, n-decyl-β-D-maltopyranoside; Dol-P, dolichyl monophosphate; Dol-PP, dolichyl diphosphate; DPMS, dolichylphosphate mannose synthase; EDTA, ethylenediaminetetraacetic acid; ER, endoplasmic reticulum; GDP, guanosine diphosphate; GFP, green fluorescent protein; Glc, glucose; GlcNAc, N-acetyl glucosamine; GT, glycosyltransferase; Hepes, 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid; HMG, high-mannose N-glycan; HPAEC-PAD, High-Performance Anion Exchange Chromatography and Pulsed Amperometric Detection;
IFH, interface helix; ITC, isothermal titration; LAPAO, 3-dodecylamido-N,N-dimethylpropyl amine oxide; LDAO, N,N-dimethylglycerol N-oxide; LLO, lipid-linked oligosaccharide; Man, mannose; MWCO, molecular weight cutoff; OST, oligosaccharide transferase; PC, 1-palmitoyl-2-oleoyl-sn-glycero-3-phosphatidylcholine; PE, 1-palmitoyl-2-oleoyl-sn-glycero-3-phosphatidylethanolamine; PEG, polyethylene glycol; PG, 1-palmitoyl-2-oleoyl-sn-glycero-3-phospho-(1′-rac-glycerol); PS, L-α-phosphatidylserine; r.m.s.d., root mean square distance; r.p.m., revolutions per minute; S-layer, surface layer; SD, standard deviation; SDS-PAGE, sodium dodecyl sulfate–polyacrylamide gel electrophoresis; SEM, standard error of the mean; Sub-IFH, substratum-interface helix; TACK, Thaumarchaeota, Aigarchaeota, Crenarchaeota and Korarchaeota; TEV, tobacco etch virus; TMH, transmembrane helix; UDP, uridine diphosphate.

References

[1] Nothaft, H., Szymanski, C.M., (2010). Protein glycosylation in bacteria: sweeter than ever. Nat. Rev. Microbiol., 8, 765–778. https://doi.org/10.1038/nrmicro2383.

[2] Moreman, K.W., Tiemeyer, M., Naim, A.V., (2012). Vertebrate protein glycosylation: diversity, synthesis and function. Nat. Rev. Mol. Cell. Biol., 13, 448–462, https://doi.org/10.1038/nrm3383.

[3] Eichler, J., (2013). Extreme sweetness: protein glycosylation in Archaea. Nat. Rev. Microbiol, 11, 151–156, https://doi.org/10.1038/nrmicro2957.

[4] Eichler, J., Imperiali, B., (2019). Biogenesis of asparagine-linked glycoproteins across domains of life-similarities and differences. ACS Chem. Biol., 13, 833–837, https://doi.org/10.1021/acscbr.9b00163.

[5] Apweiler, R., Hermjakob, H., Sharon, N., (1999). On the frequency of protein glycosylation, as deduced from analysis of the SWISS-PROT database. Biochim. Biophys. Acta, 1473, 4–8, https://doi.org/10.1016/s0304-4165(99)00165-8.

[6] Lombard, J., (2016). The multiple evolutionary origins of the eukaryotic N-glycosylation pathway. Biol. Direct, 11, 36, https://doi.org/10.1186/s13062-016-0137-2.

[7] Meyer, B.H., Albers, S.V., (2013). Hot and sweet: protein glycosylation in Crenarchaeota. Biochem. Soc. Trans., 41, 384–392, https://doi.org/10.1042/BST20120298.

[8] Larkin, A., Imperiali, B., (2011). The expanding horizons of asparagine-linked glycosylation. Biochemistry, 50, 4411–4426, https://doi.org/10.1021/bi100346n.

[9] Lombard, V., Golaconda Ramulu, H., Dula, E., Coutinho, P.M., Henrisst, B., (2014). The carbohydrate-active enzymes database (CAZY) in 2013. Nucleic Acids Res., 42, D490–D495, https://doi.org/10.1093/nar/gkt1178.

[10] Guan, Z., Naparstek, S., Kaminski, L., Konrad, Z., Eichler, J., (2010). Distinct glycan-charged phosphodolichol carriers are required for the assembly of the pentasaccharide N-linked to the Haloflexax volcanii S-layer glycoprotein. Mol. Microbiol., 78, 1294–1303, https://doi.org/10.1111/j.1365-2958.2010.07405.x.

[11] Taguchi, Y., Fujinami, D., Kohda, D., (2016). Comparative analysis of archaeal lipid-linked oligosaccharides that serve as oligosaccharide donors for Asn glycosylation. J. Biol. Chem., 291, 11042–11054, https://doi.org/10.1074/jbc.M115.713156.

[12] Albers, S.V., Meyer, B.H., (2011). The archaeal cell envelope. Nat. Rev. Microbiol., 9, 414–426, https://doi.org/10.1038/nrmicro2576.

[13] Jarrell, K.F., Ding, Y., Meyer, B.H., Albers, S.V., Kaminski, L., Eichler, J., (2014). N-linked glycosylation in Archaea: a structural, functional, and genetic analysis. Microbiol. Mol. Biol. Rev., 78, 304–341, https://doi.org/10.1128/MMBR.00052-13.

[14] Fujinami, D., Taguchi, Y., Kohda, D., (2017). Asn-linked oligosaccharide chain of a crenarchaeon, Pyrobaculum calidifontis, is reminiscent of the eukaryotic high-mannose-type glycan. Glycobiology, 27, 701–712, https://doi.org/10.1093/glycob/cwx044.

[15] Guy, L., Ettema, T.J.G., (2011). The archaeal ‘TACK’ superfamily and the origin of eukaryotes. Trends Microbiol., 19, 580–587, https://doi.org/10.1016/j.tim.2011.09.002.

[16] Koonin, E.V., Yutin, N., (2014). The dispersed archaeal eukaryome and the complex archaeal ancestor of eukaryotes. Cold Spring Harb. Perspet. Biol., 6, a016188, https://doi.org/10.1101/cshperspect.a016188.

[17] Gandini, R., Reichenbach, T., Tan, T.C., Divine, C., (2017). Structural basis for dolichylphosphate mannose biosynthesis. Nat. Commun., 8, 120, https://doi.org/10.1038/s41467-017-00187-2.

[18] Matsumoto, S., Shimada, A., Nyirenda, J., Igure, M., Kawano, Y., Kohda, D., (2013). 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, 17688–17783, https://doi.org/10.1073/pnas.1309777110.

[19] Sindhuwinita, N., Munoz, E., Javier Munoz, F., Palcic, M.M., Peters, H., Peters, T., (2010). Binding of an acceptor substrate analog enhances the enzymatic activity of human N-linked sugar synthase. Glycobiology, 20, 128–138, https://doi.org/10.1093/glycob/cwp019.

[20] Schuman, B., Evans, S.V., Fyles, T.M., (2013). Geometric attributes of retaining glycosyltransferase enzymes favor an orthogonal mechanism. PLoS One, 8, e70177, https://doi.org/10.1371/journal.pone.0071077.

[21] Zhang, L., Lau, K., Cheng, J., Yu, H., Li, Y., Sugiarlto, G., et al., (2010). Helicobacter hepaticus Hn0072 gene encodes a novel α1-3-fucosyltransferase belonging to C AZy GT11 family. Glycobiology, 20, 1077–1088, https://doi.org/10.1093/glycob/cwp068.

[22] Morgan, J.L.W., Strumillo, J., Zimmer, J., (2013). Crystallographic snapshot of cellulose synthesis and membrane translocation. Nature, 493, 181–186, https://doi.org/10.1038/nature11744.

[23] Saxena, I.M., Brown Jr., R.M., Dandekar, T., (2001). Structure–function characterization of cellulose synthase: relationship to other glycosyltransferases. Phytochemistry, 57, 1135–1148, https://doi.org/10.1016/s0031-1827(01)00048-6.

[24] Tsirigos, K.D., Peters, C., Shu,N., Käll, L., Elofsson, A., (2015). N-glycosylation in Archaea: putative components of glycosyltransferases and oligosaccharyltransferases in Archaea: putative components of glycosyltransferases and oligosaccharyltransferases in Archaea: putative components of glycosyltransferases and oligosaccharyltransferases in Archaea: putative components of glycosyltransferases and oligosaccharyltransferases in Archaea: putative components of glycosyltransferases and oligosaccharyltransferases in Archaea: putative components of glycosyltransferases and oligosaccharyltransferases in Archaea: putative components of glycosyltransferases and oligosaccharyltransferases in Archaea: putative components of glycosyltransferases and oligosaccharyltransferases in Archaea: putative components of glycosyltransferases and oligosaccharyltransferases in Archaea: putative components.
of the N-glycosylation pathway in the third domain of life. 

FEMS Microbiol. Lett., 300, 122–130, https://doi.org/10.1111/j.1574-6968.2009.01775.x.

[26] Kaminski, L., Lurie-Weinberger, M.N., Allers, T., Gophna, U., Eichler, J., (2013). Phylogenetic- and genome-derived insight into the evolution of N-glycosylation in Archaea. Mol. Phylogenet. Evol., 68, 327–339, https://doi.org/10.1016/j.ympev.2013.03.024.

[27] Palsdottir, H., Hunte, C., (2004). Lipids in membrane protein structures. Biochim. Biophys. Acta, 1666, 2–18, https://doi.org/10.1016/j.bbamem.2004.06.012.

[28] Musatov, A., Sediak, E., (2017). Role of cardiolipin in stability of integral membrane proteins. Biochimie, 142, 102–111, https://doi.org/10.1016/j.bioch.2017.08.013.

[29] Bloch, J.S., Pesciullesi, G., Boilevin, J., Nosol, K., Irobalieva, R.N., Darbre, T., et al., (2020). Structure and mechanism of the ER-based glucosyltransferase ALG6. Nature, 579, 443–447, https://doi.org/10.1038/s41586-020-2444-z.

[30] Waldo, G.S., Standish, B.M., Berendzen, J., Terwilliger, T.C., (1999). Rapid protein-folding assay using green fluorescent protein. Nat. Biotechnol., 17, 691–695, https://doi.org/10.1038/10904.

[31] Drew, D.E., von Heijne, G., Nordlund, P., de Gier, J.W., (2001). Green fluorescent protein as an indicator to monitor membrane protein overexpression in Escherichia coli. FEBS Lett., 507, 220–224, https://doi.org/10.1016/S0014-5793(01)02980-5.

[32] van den Berg, S., Löfdahl, P.A., Härd, T., Berglund, H., (2006). Improved solubility of TEV protease by directed evolution. J. Biotechnol., 121, 291–298, https://doi.org/10.1016/j.jbiotec.2005.08.006.

[33] Kabsch, W., (2010). XDS. Acta Crystallogr. D Biol. Crystallogr., 66, 125–132, https://doi.org/10.1107/S0907444909047337.

[34] Vonrhein, C., Blanc, E., Roversi, P., Bricogne, G., (2007). Automated structure solution with autoSHARP. Methods Mol. Biol., 364, 215–230, https://doi.org/10.1385/1-59745-266-1:215.

[35] Bricogne, G., Vonrhein, C., Flensburg, C., Schiltz, M., Paciorek, W., (2003). Generation, representation and flow of phase information in structure determination: recent developments in and around SHARP 2.0. Acta Crystallogr. D Biol. Crystallogr., 59, 2023–2030, https://doi.org/10.1107/s0907444903017694.

[36] Abrahams, J.P., Leslie, A.G.W., (1996). Methods used in the structure determination of bovine mitochondrial F1 ATPase. Acta Crystallogr. D Biol. Crystallogr., 52, 30–42, https://doi.org/10.1107/S0907449995008754.

[37] Jones, T.A., Zou, J.-Y., Cowan, S.W., Kjeldgaard, M., (1991). Improved methods for building protein models in electron density maps and the location of errors in these models. Acta Crystallogr. A, 47, 110–119, https://doi.org/10.1107/s010876739100010224.

[38] Collaborative Computational Project Number 4, (1994). The CCP4 suite: programs for protein crystallography. Acta Crystallogr. D Biol. Crystallogr., 50, 760–763, https://doi.org/10.1107/S0907444994003112.

[39] Adams, P.D., Afonine, P.V., Bunkóczi, G., Chen, V.B., Davis, I.W., Echols, N., et al., (2010). PHENIX: a comprehensive Python-based system for macromolecular structure solution. Acta Crystallogr. D Biol. Crystallogr., 66, 213–221, https://doi.org/10.1107/S0907444909052925.

[40] Bricogne, G., Blanc, E., Brandl, M.T., Flensburg, C., Keller, P., Paciorek, W., et al., (2011). BUSTER Version 2.10.0. Global Phasing Ltd., Cambridge, UK, 2011.

[41] Emsley, P., Cowtan, K., (2004). Coot: model-building tools for molecular graphics. Acta Crystallogr. D Biol. Crystallogr., 60, 2126–2132, https://doi.org/10.1107/S0907444904019158.