Protein glycosylation is a critical protein modification. In biogenic membranes of eukaryotes and archaea, these reactions require activated mannose in the form of the lipid conjugate dolichylphosphate mannose (Dol-P-Man). The membrane protein dolichylphosphate man- nose synthase (DPMS) catalyzes the reaction whereby mannose is transferred from GDP-mannose to the dolichol carrier Dol-P, to yield Dol-P-Man. Failure to produce or utilize Dol-P-Man compromises organism viability, and in humans, several mutations in the human dpm1 gene lead to congenital disorders of glycosylation (CDG). Here, we report three high-resolution crystal structures of archaeal DPMS from Pyrococcus furiosus, in complex with nucleotide, donor, and glycolipid product. The structures offer snapshots along the catalytic cycle, and reveal how lipid binding couples to movements of interface helices, metal binding, and acceptor loop dynamics to control critical events leading to Dol-P-Man synthesis. The structures also rationalize the loss of dolichylphosphate mannose synthase function in dpm1-associated CDG.
Protein glycosylation is critical for protein folding, sorting, stability, and function in all domains of life. In biogenic membranes (i.e., self-synthesizing endoplasmic reticulum (ER) membrane in eukaryotes and cytoplasmic cell membrane in archaea), the activated mannose used for protein glycosylation and glycosylphosphatidylinositol anchor synthesis is provided in the form of the glycolipid intermediate dolichylphosphate β-D-mannose (Dol-P-Man). An integral membrane protein, Dol-P-Man synthase (DPMS; EC 2.4.1.83), is responsible for Dol-P-Man biosynthesis by catalyzing the transfer of mannose from guanosine diphosphate α-D-mannose (GDP-Man) to dolichylphosphate (Dol-P). In eukaryotes, Dol-P-Man synthesis takes place on the cytosolic face of the ER membrane, followed by translocation of the glycolipid across the membrane to expose the activated mannosyl group to the ER lumen where initial protein glycosylation and glycosylphosphatidylinositol biosynthesis take place. In archaea, Dol-P-Man is synthesized on the cytoplasmic face of the plasma membrane, and flipped to the cell exterior, i.e., the topological counterpart of the ER lumen, to serve as mannosyl donor for glycosylation of surface proteins of the paracrystalline glycoprotein cell envelope. Altered DPMS activity in humans is linked to cancer and congenital disorders of glycosylation.

Here we report the structure of a bona fide DPMS. High-resolution crystal structures of wild-type Pyrococcus furiosus
DPMS (P)DPMS in complex with donor and acceptor substrates, as well as with Dol-P-Man product, were determined. The enzyme complexes provide snapshots along the catalytic cycle, and offer a detailed picture of how Dol-P-Man is synthesized in _Pyrococcus_. We expose a sophisticated interplay between juxta-membrane interface (IF) helices, donor, metal, and acceptor that controls Dol-P-Man synthesis. While the IF helices are strictly required for Dol-P-Man synthesis, the transmembrane (TM) domain is expendable. Furthermore, our structural data rationalize the importance of several amino acids that are implicated in human disease. Considering the importance of DPMS function for organism biology, including several common pathogens, these results provide a valuable structural platform for design of therapeutics that target Dol-P-Man-dependent protein glycosylation.

We note that the TM domain shows an unusual arrangement of the two TM dimers. TMD1 is perpendicular to the membrane and spans ∼45 Å (Fig. 1a), while TMD2 is rotated by 90° with respect to TMD1, giving the impression of walking legs that associate at a junction. Several observations imply that this arrangement is authentic, rather than an artifact of detergent solubilization or crystal-packing interactions. First, the interface between TMD1 and TMD2 displays pseudo-two-fold symmetry at the junction (Fig. 1b), which is also manifested as sequence similarity between the TM dimers (Fig. 1c). The pseudo-twofold interface is optimally packed at the junction, and decreasing the relative angle between the TM domains would result in severe clashes, most seriously between Phe237 (TMH1) and Tyr342 (TMH4), Phe278 (TMH2) and Trp350 (TMH4), Phe278 (TMH2) and Phe301 (TMH3). Analysis of the interface between TMD1 and TMD2 using PISA provides a valuable structural platform for design of therapeutics that target Dol-P-Man-dependent protein glycosylation.

Results

**Overall structure.** We determined the crystal structures of P)DPMS in complex with metal-bound nucleotide (GDP•Mg<sup>2+</sup>) and donor (GDP-Man•Mn<sup>2+</sup>; Supplementary Table 1). The overall structure of P)DPMS features an N-terminal, cytoplasmic catalytic domain covalently attached to a TM domain with four TM helices (TMs) arranged as two TMH dimers, TMD1 and TMD2 (Fig. 1). The TM domains of DPMSs show varying topological complexity, and based on differences in the TM domain, two classes of DPMS have been proposed<sup>17</sup> (Supplementary Fig. 1a, b). The type-I enzymes are represented by DPMSs from *Saccharomyces cerevisiae*<sup>9</sup>, where the catalytic domain is covalently attached to a single C-terminal TMH (Supplementary Fig. 1a). In contrast, the human type-II DPMS forms a multi-subunit complex with a catalytic domain (DPM1) that lacks TMHs but associates with two separately encoded membrane proteins, DPM2 and DPM3, each of which comprises 80–90 amino acids that fold into two TMHs<sup>18</sup> (Supplementary Fig. 1b). The structure of the TM domain in P)DPMS places the pyrococcal and thermococcal enzymes in a new class of DPMSs, type-III (Supplementary Fig. 1c), which is topologically similar to the assembled mammalian DPM1/DPM2/DPM3 complex. Human DPM1 requires DPM3 for targeting to the ER membrane. In contrast, DPM2 enhances activity, although the exact role is unknown<sup>18</sup>. Topologically, DPM2 and DPM3 in human DPMS could correspond to the two TMH dimers in P)DPMS.

The catalytic domain of P)DPMS belongs to family 2 of inverting glycosyltransferases (GT2; [www.cazy.org])<sup>19</sup> with a GT-A fold (Supplementary Fig. 2) where a conserved DXD sequence motif (Supplementary Figs. 2 and 3) coordinates the catalytic metal ion and donor substrate<sup>20</sup>. The canonical GT-A fold was first described for the spore-coat protein SpSA from *Bacillus subtilis*<sup>21</sup>, which features a 7-stranded β-sheet flanked by α-helices on both sides (Supplementary Fig. 2). The GT-A fold can be further broken down into an N-terminal and C-terminal half. The N-terminal region is composed of a Rossmann-like four-stranded, parallel β-sheet (strands β1–β4) with two α-helices (α1 and α2) on one side of the sheet, and one α-helix (α3) on the other side. This Rossmann-like N-terminal motif extends into a C-terminal motif that folds as a three-stranded mixed β-sheet (strands β5–β7) with one α-helix on either side (α4 and α5). The β-strands of the N- and C-terminal halves form an extended β-sheet (Supplementary Fig. 2). In addition to the GT-A core structure, members of the GT2 family display various additional structural features relating to their unique functional requirements and subcellular localization. Besides the TM domain, two IF helices (IFH1 and IFH2) are appended to the cytoplasmic GT-A domain in P)DPMS (Supplementary Fig. 2a). The IF helices are amphipathic, and oriented to expose their hydrophobic faces towards the membrane (Supplementary Fig. 4).

**Binding of donor substrate.** The guanosine moiety of GDP is located in a pocket formed at the C-terminal ends of the parallel β-strands (β1–β4), where interactions are offered mainly with the β-α loop regions (Supplementary Fig. 2). Guanosine interacts with the protein through several possible hydrogen bonds: guanine O6 – Gly68 N (β3-α3 loop); guanine N1 and N2 – Asp39 OD2 (β2-α2 loop); ribose O2 – Tyr10 N (β1-α1 loop); and ribose O3 – Ala90 N (β4-α4 loop). In addition, the guanine ring stacks with the hydrophobic Leu69 side chain (β3-α3 loop).
diphosphate-binding site is formed at the canonical DXD motif located in the β4-α4 loop, which is defined by DAD in DPMSs (Supplementary Fig. 3), and forms the inner wall of the diphosphate-binding site. The active site in the catalytic domain is gated by a flexible loop (residues 202–210), which we refer to as the acceptor loop, or A-loop (Figs. 1a and 2). In the PfDPMS•GDP•Mg2+ and PfDPMS•GDP-Man•Mn2+ complexes, the A-loop is closed and folded over the diphosphate-binding site to create an outer wall (“front door”) of the active site (Fig. 2).

Asp91 and Gln93 in the 89DADLQH94 motif coordinate the α- and β-phosphate groups of GDP via the metal cation: α-phosphate O2A – Me2+ – Asp91 OD2, and α-phosphate O2B – Me2+ – Gln93 OE1 (Fig. 2a, b). The conformation of the A-loop couples to donor binding through a network of interactions: GDP α-phosphate O2A – Arg202 Nh1, and α-phosphate O2B – Lys208 N. The A-loop is further secured by ionic interactions: Arg202 Nh2 – Glu12 Oe1; Arg202 Nh2 – Oe2; Arg202 Nh2 – Glu206 O; and Glu12 Oe2 – Lys208 Nz. On the other side of the donor pocket, Phe177 assumes a conformation that together with complementary shielding from the A-loop prevents the donor from leaving through a “back door”. The conformation of Phe177 is controlled by the precise conformation and position of residues 212–219 in IFH2.

In the donor complex, PfDPMS•GDP-Man•Mn2+, the side chains of Asp89 and His94 are within hydrogen-bonding distance of the donor mannosyl C2 and C3 hydroxyl groups, respectively (Fig. 2b). The mannosyl O6 and O4 hydroxyl groups are also within hydrogen-bonding distance of the side-chain nitrogen atoms in Arg117 and Lys178. The mannosyl C1 atom is rotated relative to the expected point of attack by an incoming nucleophilic Dol-P phosphate group (Fig. 2b), but would be optimally oriented by means of a simple rotation in response to acceptor binding. The PfDPMS•GDP•Mg2+ and PfDPMS•GDP-Man•Mn2+ complexes are very similar with only minor shifts in side chains located in IFH1 and IFH2, which do not interact directly with GDP or GDP-Man (Fig. 2c). As expected for GT2 members, wild-type PfDPMS activity depends on divalent cations (Supplementary Fig. 6).

Binding of Dol-P acceptor and Dol-P-Man product. The archaeon P. furiosus prefers 65-carbon Dol-P carriers25. Due to
commercial unavailability of 65-carbon Dol-P, we tested two alternative carbon-chain lengths, C55 and C95. Both Dol55-P and Dol95-P serve as acceptors for PfDPMS in vitro (Supplementary Fig. 7a), where Dol95-P performs better (37% higher in turnover). Both acceptor lipids provide significant thermal stabilization to unfolding for detergent-solubilized PfDPMS with an increase in melting temperature (ΔTm) of 13.3 °C and 18.6 °C, respectively, compared with no addition of Dol-P (Supplementary Fig. 7b, c). The ΔTm value is 40% higher for Dol95-P compared with Dol55-P, which indicates that the higher activity (37%) with Dol95-P is mostly due to protein stabilization. Visualization of the Dol95-P-Man product with thin layer chromatography proved challenging due to co-migration of Dol95-P and Dol95-P-Man in the solvents tested. This problem could be solved by subjecting Dol95-P-Man to mild acid hydrolysis using hydrochloric acid, which allowed the mannose released from Dol95-P-Man to be unambiguously identified (Supplementary Fig. 8).

In an attempt to capture catalytically relevant structural states of wild-type PfDPMS, we crystallized a reaction mixture that had been incubated at 60 °C prior to crystallization. A crystal grown from this reaction showed a post-catalysis active-site state with both products, GDP and Dol55-P-Man, bound (Fig. 3). The mannosyl O3, O2 and O5 atoms are within hydrogen-bonding distance of His94 Ne2 and Lys178 Nz (Fig. 3, left inset). In contrast to the PfDPMS•GDP•Mg2+ and PfDPMS•GDP•Dol55-P-Man•Mn2+ complexes, the diphosphate group of the GDP product has moved away towards the “back door” to provide space for the mannosyl head group of Dol-P-Man. To accommodate the new conformation of the GDP diphosphate moiety, Phe177, which gates the back door in the nucleotide and donor complexes, has flipped down to position the aromatic ring immediately below the GDP β-phosphate group (Fig. 3b, c). The phosphate group in the glycolipid product is coordinated by Ser135, Arg117, and Arg131, and the phosphoester oxygen interacts with Lys178 (Fig. 3, left inset). The binding site for the Dol-P phosphate group in Dol-P-Man shows a distinctly positive electrostatic potential (Supplementary Fig. 9). The importance of Ser135 in IFH1 for binding and positioning the phosphate group in the acceptor lipid is further emphasized by a significant drop in product formation by the S135A variant (Supplementary Fig. 6c), which agrees with the observed effect for an alanine mutant of the equivalent Ser141 in S. cerevisiae DPMS26. The dolichyl chain is bound between IFH1 and IFH2 where the first two isoprene units interact with hydrophobic residues: Ile134, Ala138, Val141.

Fig. 3 Structure of the PfDPMS•GDP•Dol55-P-Man complex. a Enlarged views of the active site (left inset) and the TM domain (right inset). Left inset, the Dol-P-Man head group is shown bound in the active site of wild-type PfDPMS. The metal ion has departed and the GDP diphosphate group has flipped away from the transfer site towards the “back door”, to accommodate the glycolipid product. Right inset, the Dol-P-Man isoprenoid chain traces down to the TM domain where it rests on TMH4. The acceptor loop is open and Phe177 is in the “down” position. Superimposition of the crystal complexes b PfDPMS•GDP•Mg2+ (gold) and PfDPMS•GDP•Dol55-P-Man (blue), and c PfDPMS•GDP-Man•Mn2+ (gold) and PfDPMS•GDP•Dol55-P-Man (blue). Relevant side chains are shown.
e.m. (Nand aqueous phase, respectively; and they display catalytic activity that acts as a supporting platform (Fig. 3, right inset).

The involvement of IFH1 and IFH2 in acceptor binding assigns an important role to these helices, which prompted us to investigate the role of the TM domain (TMD1/TMD2) in Dol-P-Man synthesis. For this purpose, a truncation variant lacking the TM domain (TMD1/TMD2) in Dol-P-Man synthesis was used as acceptor substrate.Δ230–352 was purified from the membrane fraction and aqueous phase, respectively; and they display catalytic activity comparable with that of the wild type. Errors are given as mean values ± s.e.m. (N = 3)

Catalytic mechanism of PfDPMS. Based on the PfDPMS crystal complexes, we can outline a catalytic mechanism for Dol-P-Man synthesis (Fig. 4). When GDP-Man and metal ion are bound in the active site in the absence of acceptor lipid, the A-loop shields the donor substrate at the "front door", and Phe177 gates the "back door". Based on the post-catalysis GDP•Dol55–Man product complex, a Dol-P molecule would position itself with the phosphate group in the acceptor-phosphate binding site (Ser135, Arg117, and Arg131), with the first two isoprene units of the dolichyl chain between IFH1 and IFH2. The Dol-P phosphate group is pre-activated for nucleophilic attack, and not dependent on a catalytic base for activation.

The Dol-P phosphate group attacks with one of its activated oxygen atoms at the α-mannosyl C1 and a new bond is formed with inverted configuration at C127. This is in agreement with the mechanism proposed for S. cerevisiae DPMS based on theoretical modeling and fluorescence resonance energy transfer experiments28, however, our present data do not lend support for the suggestion that DPMSs are binuclear metalloenzymes. Only a single metal ion, Mg2+ or Mn2+, is consistently observed in the PfDPMS complexes, and the high sequence conservation of the active sites in DPMSs suggests that PfDPMS is representative with respect to metal-ion binding stoichiometry. Furthermore, the role of the conserved Arg212 in ScDPMS (Arg202 in the A-loop of PfDPMS) is not, as suggested based on modeling28, to coordinate the mannosyl O3 atom in GDP-Man and the Dol-P phosphate oxygen atom, but to couple metal binding to conformational control of the conserved A-loop. Coordination of the nucleophilic Dol-P phosphate group is assigned to Arg117 and Ser135 with assistance of Arg131 (Arg122, Ser141 and Arg137 in ScDPMS), all of which are conserved in bona fide DPMSs (Supplementary Fig. 3).

The involvement of IFH1 and IFH2 in acceptor binding assigns an important role to these helices, which prompted us to investigate the role of the TM domain (TMD1/TMD2) in Dol-P-Man synthesis. For this purpose, a truncation variant lacking the membrane fraction.
TM domain (residues 230–352) was produced. Interestingly, the Δ230–352 variant was still targeted to the membrane and displayed catalytic activity comparable with that of the wild type (Fig. 5). This strengthens further the role of the IF helices in acceptor recognition and binding, and emphasizes that Dol-P-Man synthesis does not per se require the TM domain, and that the topologically variable TM domains of different DPMS classes (Supplementary Fig. 1) perform functions that are distinct from catalyzing Dol-P mannosylation.

**Structural comparison with related enzymes.** There are no structures of acceptor-product complexes available for glycosyltransferases that use lipid substrates, however, crystal structures of donor product complexes (i.e., the donor without sugar) have been reported for two bacterial membrane glycosyltransferases, i.e., *Synechocystis* PfGtrB (complex with uridine diphosphate, UDP, and Mg2+; PDB code 5EKP)29 and the aminoarabinose transferase ArnT from *Cupriavidus metallidurans* (complex with undecaprenylphosphate, Und-P; PDB code 5F15).30

GtrB is a GT2 member that catalyzes the transfer of glucose (Glc) from UDP-Glc to Und-P in vitro. PjDPMS and GtrB share similar overall topology of the catalytic domain (Supplementary Fig. 2), but differ in the TM domain and IF helices (Supplementary Fig. 10). Compared with PjDPMS, the two TMHs in GtrB assume different positions relative to the catalytic domain, and participate in homotrimer assembly (Supplementary Fig. 10d), and there are no counterparts in GtrB to the TMH-dimer motifs in PjDPMS (Supplementary Fig. 10e). Another topological difference is an extra β-strand that hydrogen bonds with β3 of the GT-A core (Supplementary Figs. 2 and 10a).

Detergent-solubilized PjDPMS is likely to be functional as a monomer, which does not preclude the possibility of higher-order oligomers in a native membrane. To investigate the possibility of higher-order oligomers for PjDPMS, a dimer and tetramer were modeled in silico. Due to the tilted orientation of TMD2, the PjDPMS structure is incompatible with a homotrimer assembly; however, a dimer is possible, albeit predicted not to be stable according to PISA (complex formation significance score value 0.0). Thus, the TM domains in PjDPMS and GtrB probably have different functions. Due to the differences of the IFH and TMH elements (Supplementary Fig. 10e), a structural alignment of PjDPMS and the GtrB subunit is only meaningful for the GT-A domain. Of the 196 and 192 residues that comprise the GT-A domains in PjDPMS and GtrB (PDB code 5EKP), respectively, 146 CA positions can be aligned with an r.m.s.d. value of ~1.6 Å. Of the aligned positions, 46 amino acids are identical, which corresponds to a structural sequence identity of 31.5% (24% identity for the entire extent of the catalytic domain).

All known bona fide DPMSs use GDP-Man as donor substrate, whereas GtrB use UDP-Glc in vitro. In agreement with the observed activity, UDP is bound in the donor-binding pocket in the GtrB crystal structure. Overall, the inner region of the pocket is similar in the two enzymes (Supplementary Fig. 10f). The 89DADLQH49 motif in PjDPMS corresponds to 94DADLDQ99 in GtrB where Asp96 and Gln98 coordinate the metal ion. A number of differences are also noted (Supplementary Fig. 10f). To accommodate binding of the smaller uridine ring, Asp43 in GtrB (Asp39 in PjDPMS) adopts a different conformation to enable a hydrogen bond with the UDP N3 nitrogen atom. In GtrB, the loop connecting β6 and β7 (“back door” loop in PjDPMS) has a different conformation and is folded away from the donor pocket to pack against the C-terminal end of TMH2. The acceptor loop in GtrB (residues 198–204) was not modeled in the crystal structure, which prevents analysis of functionally relevant structural determinants, however, the loop is three residues shorter and shows a different amino acid sequence compared with DPMSs (Supplementary Figs. 3 and 10a). To assess the degree of donor-substrate promiscuity in PjDPMS, the catalytic activity was compared for GDP-Man, GDP-Glc and UDP-Glc (Supplementary Fig. 11). The preference for the mannosyl donor is evident, as shown by a four-fold higher product formation with GDP-Man compared with GDP-Glc. Some residual activity is observed also for UDP-Glc, but the value falls within the error estimate of GDP-Man conversion. Spatially, the smaller uridine ring can be accommodated in the guanine-binding site of PjDPMS, and similarly, glucose is compatible with binding at the transfer site. The different interactions with a uridine ring as opposed to guanine, and the presence of an equatorial glucosyl C2 hydroxyl group compared with its axial configuration in mannose, are the main discrepancies that are expected to account for the lower activity with UDP-Glc or GDP-Glc as donor. Equatorial versus axial O2 will mainly affect interactions associated with Asp89, Leu69, and Phe177, all of which appear to have distinct roles in Dol-P-Man synthesis.

As for PjDPMS, the IF helices in GtrB are amphipathic, and oriented with the hydrophobic sides facing the membrane, but as noted above, they assume different positions relative to the catalytic domain and active site (Supplementary Fig. 10b, c). IFH1 is also shorter in GtrB. The interface between IFH1 and IFH2 in GtrB is densely occupied by large side chains (Phe134, Phe139, Phe178, Met179, Phe183, and Trp213), and packs against a loop segment (residues 150–155) between IFH1 and β6 that blocks a possible opening between the IF helices. A similar block does not exist in PjDPMS where the corresponding loop adopts a different conformation. Unless a major conformation change takes place in the blocking loop and IF helices in GtrB, entry of the Und-P acceptor between IFH1 and IFH2 is highly improbable. A more likely entry point for the Und-P substrate in GtrB is between IFH2 and the β6–β7 loop, where a large opening can provide direct access to the bound donor substrate. This opening is close to the GtrB A/B (and C/D) subunit interface, and is one of the regions that show pronounced differences in sequence and structure compared with PjDPMS. Because of these discrepancies between the PjDPMS and GtrB structures, it is difficult to identify side chains in GtrB that can serve as counterparts to Arg131 and Ser135 that bind the acceptor Dol-P phosphate group in PjDPMS. Arg122 in GtrB probably corresponds to Arg117, which also participates in Dol-P phosphate binding in PjDPMS. Furthermore, we predict that Asp99 in the 94DADLDQ99 motif of GtrB should be able to interact with the glycosyl O3 hydroxyl group in an Und-P-Glc product.

Based on the overall similarity of the substrates and the architecture of the donor-binding pocket, a similar but not necessarily identical, catalytic reaction mechanism can be expected for the two enzymes. In GtrB, Asp157 has been suggested to function as a catalytic acid to protonate the distal β-phosphate group of the donor.29 This cannot apply to PjDPMS since the corresponding residue is Gly158. A soluble member of the GT2 family, *Staphylococcus aureus* TarS (PDB code 5TZE) has a proline residue in the equivalent position, which suggests that if Asp157 in GtrB plays a role as catalytic acid, it is not a general theme among inverting GT2 enzymes. Because of the absence of acceptor loop in GtrB, and the large differences in the acceptor-binding region, further evaluation of possible similarities and differences with respect catalytic mechanism becomes difficult.

Unlike PjDPMS and GtrB, which are membrane enzymes of the GT2 family, the inverting glycosyltransferase ArnT belongs to family GT83, which features a GT-C fold similar to oligosaccharyl transferases.30 The catalytic domain is associated with three IF helices (JM1–JM3) and a TM domain composed of 13 TMHs.
ArnT does not use a nucleotide sugar as donor substrate, but instead, aminoarabinose is transferred from an undecaprenylphospho- 
date donor to the acceptor substrate lipid A. The structure of 
ArnT was determined in the apo state (PDB code 5EZM) and in 
complex with the “unloaded” donor, i.e., Und-P30 (PDB code 
5F15), the latter that would correspond functionally to our 
GDP•Mg$^{2+}$ complex. The structures and substrate specificities 
are fundamentally different in PfDPMS and ArnT, but similarities 
are noted related to the IF helices. As for IFH2 in PfDPMS, the IF 
helix JM3 in ArnT is directly connected to a flexible loop (A-loop 
in PfDPMS). In ArnT, this loop is shorter than the A-loop in 
PfDPMS, and disordered in both ArnT structures. In analogy to 
IFH2 in PfDPMS, binding of Und-P to ArnT induces a 
conformational change of JM3 that triggers opening of the 
acceptor loop, which in turn leads to loss of metal-ion 
coordination and metal release from the active site. This situation 
appears similar to the coupling between IFH2 and the meta-
binding site in PfDPMS. Beyond this observation, the absence of 
additional relevant structural states for ArnT precludes further 
mechanistic comparison of the enzymes.

Discussion
As observed in the GDP•Dol55-P-Man product complex, several 
side chains in the IF helices participate in critical interactions 
with the glycolipid product. This shows that the IF helices, rather 
than the TM domain, constitute the basis for recognition and 
binding of the lipid substrate. Although dispensable for activity 
of detergent-solubilized PfDPMS, the arrangement of TMD1 and 
TMD2 still offers a binding platform for the Dol-P acceptor 
and Dol-P-Man product, which is likely to be more relevant for 
substrate binding and activity in native membranes than in 
protein-detergent complexes.

Based on the crystal complexes, we can propose a tentative 
model for substrate binding and product release orchestrated by 
an interplay between IF helices, donor, metal ion and acceptor. 
Prior to encounter with a Dol-P acceptor molecule, GDP-Man 
and metal ion are expected to be bound in the active site (Fig. 2b), 
with the A-loop secured in its closed conformation by a network 
of interactions formed between A-loop residues (Arg202 and 
Lys208), the metal-coordination center (Asp91), and Glu12 in the 
β1-α1 loop (Fig. 6, step 1). Displacement and conformational 
changes of the IF helices, most pronouncedly of IFH2, enable 
entry of the Dol-P acceptor between IFH1 and IFH2, and docking of 
the Dol-P phosphate group at Ser135 (assisted by Arg117 and 
Arg131), which is located immediately below the mannosyl to be 
transferred. The acceptor-induced changes in the IF helices 
(Fig. 3) leads to disruption of the A-loop interaction network at 
the DXD motif and dislodgement of the A-loop, as well as 
opening of the back door through a conformational change in 
Phe177 and its immediate environment (Fig. 6, step 2). We 
predict that collapse of the interaction network and opening of 
the A-loop coincides with concomitant loss of metal ion and 
attack by the pre-activated nucleophilic Dol-P oxygen on the 
mannosyl C1 atom. As observed in the Dol-P-Man complex, the 
acceptor product (GDP) moves towards the “front door” 
to accommodate the mannosyl group (Fig. 3). If GDP would 
leave prior to release of Dol-P-Man, GDP must exit through the 
back door, and the metal ion probably departs through the “front 
door” (Fig. 6, step 3). The cycle is completed by release of pro-
ducts, rearrangement of the IF helices, binding of a new donor 
molecule and metal ion, and lock-down of the A-loop over the 
donor-binding pocket (Fig. 6, step 4).

A 13-amino acid sequence in yeast glycosyltransferases, 
including ScDPMS, was suggested to participate in dolichol 
recognition32. This sequence (A246LFITFWSILFFY258), referred 
to the polyisoprenyl-recognition sequence (PIRS), maps to the 
only existing TMH in ScDPMS, which would correspond topolo-
gically to TMH1 in PfDPMS. However, the amino-acid 
sequences of the two TMHs show weak similarity at best (Sup-
lementary Fig. 12a). In contrast, PfDPMS TMH1 (first TMH in 
TMD1) and TMH3 (first TMH in TMD2) show more convincing 
sequence similarity to human DPM2 (Supplementary Fig. 12b, c). 
While we cannot disprove or confirm the possible importance 
of the proposed PIRS in ScDPMS, the key determinants of Dol-P 
recognition can be outlined for PfDPMS, which may also be valid 
for human DPMS (Supplementary Fig. 12d). If PfDPMS, the 
amino-acid side chains that are in direct contact with the iso-
prenoid chain and phosphate group are mainly located in IFH1, 
IFH2, and TMH4 (Supplementary Fig. 12d). This agrees with the 
observations that (i) the PfDPMS Δ230–352 variant targets to the 
membrane of the expression host, and (ii) that the catalytic 
activity of Δ230–352 variant is similar to that of the full-length 
enzyme, which emphasizes that dolichyl-chain recognition is 
intact in the absence of TM domain. In the crystal, the electron 
density beyond the first two isoprene units is weak, and although
the principal chain trajectory can be discerned, detailed interpretation of interactions between the TM domain and the iso-preneoid chain beyond the second IP unit becomes difficult. Moreover, to allow close contacts between PdPDMS molecules in the crystal lattice, the extensive C55 dolichyl chain adopts a U-shape-folded conformation, which is not necessarily the conformation preferred in the membrane environment.

Like other inverting members of the GT2 family, DPMSs (including PdPDMS) depend on metal ions for efficient catalysis. We still observe some residual catalytic activity for the D89A and D91A variants of the DxD motif (Supplementary Fig. 6). Asp89 is not directly involved in metal binding, but appears to be more important for interactions with the mannosyl group. The principal role of the metal ion in PdPDMS is to coordinate and position the diphasphate group of the donor for catalysis. Asp91 and Gln93 coordinate the α- and β-phosphate groups of GDP via the metal ion, respectively (Fig. 2). When the acceptor loop is in the closed state, Arg202 forms ionic interactions with the O2A oxygen atom of the α-phosphate group in GDP and GDP-Man, as well as with Asp91 in the DxD motif and with Glu12 (Fig. 2). This network of ionic interactions provides a means to couple donor binding to conformational control of the A-loop. We predict that Gln93 and Arg202 may be sufficient to rescue some of the metal-binding capacity in D91A to account for the observed residual activity.

The sequences of the PdPDMS catalytic domain and human DPM1 are 39% identical, and both contain the consensus signatures of DPMSs. The close relationship between the enzymes offers a structural basis of understanding human dpm1-linked mutations. Dysfunctional human DPM1 is associated with congenital disorder of glycosylation type Ie, which is caused by different mutations associated with the dpm1 gene (Supplementary Table 2). Of these mutations, three are single amino-acid replacements: R92G, S248P and G152V. The drastic effect of these mutations on DPMS stability and function can be readily inferred from the PdPDMS crystal structures. The conserved Arg63 (Arg92 in HsDPM1) is located at the end of β3 in PdPDMS (Supplementary Figs. 2 and 3) where it helps to secure the guanine-loop binding comprising residues 64–69 through a network of interactions. Insertion of a glycine residue in this position is expected to greatly compromise protein stability, as well as donor binding and catalysis. Glu216 in PdPDMS corresponds to Ser248 in HsDPM1, and is situated in IFH2. In PdPDMS, this region of the IFH2, 214FYYEL218 (246IVSFL250 in HsDPM1), undergoes significant conformational changes in response to lipid binding, and contributes important hydrophobic interactions (underlined) to the Dol-P and Dol-P-Man isoprenoid chain. These conformational changes regulate, together with Phe177, the coupled opening and closure of the acceptor loop and donor-binding pocket. Insertion of a proline residue in this position, as observed in the human DPM1 mutant S248P, would introduce a kink in IFH2 and compromise conformational responsiveness and acceptor binding, and probably also access of the donor to the active site and/or departure of GDP.

Gly122 (Gly152 in HsDPM1) is conserved in most DPMSs, and is located in the loop between β5 and IFH1. The Ca walks within van der Waals interaction distances of several backbone atoms and introduction of a valine in this position would introduce considerable steric clashes with the backbone regions 116–118 and 196–198. Since Arg117 is responsible (together with Ser135) for positioning the Dol-P phosphate group for nucleophilic attack, the G152V mutation is expected to cause severe structural perturbation in the transfer site. An additional four mutations in premature termination of the HsDPM1 polypeptide chain at various positions, and neither of these gene variants are likely to produce a foldable protein.

Aberant glycosylation is a hallmark of cancer progression. By analyzing salivary mRNA transcript levels for human dpm1 and three other genes, pancreatic cancer patients could be differentiated from healthy subjects. In addition, the human dpm1 gene is one of six glycosylation genes that have been established as significant prognostic markers for breast carcinomas. Angiogenesis is significantly enhanced in malignant breast tumors, and neoplastic transformation is correlated with changes in the glycan pattern on the cell surfaces. cAMP-dependent phosphorylation of Ser141 in S. cerevisiae DPMS (and Ser165 in human DPMS) has been observed to increase DPMS activity, and in the case of human DPMS, phosphorylation was associated with a concomitant increase in lipid-linked oligosaccharide, protein N-glycosylation and proliferation of capillary endothelial cells (angiogenesis). While angiogenesis is a complex process that is influenced by range of factors, our observation that the conserved Ser135 in PdPDMS plays a critical role by positioning the acceptor phosphate group for nucleophilic attack on the mannosyl substrate, rationalizes the importance of this serine for DPMS function. The correlation between DPMS and tumor progression is interesting and highlights DPMS as a possible drug target for development of cancer glycotherapeutics.

### Methods

#### Cloning, site-directed mutagenesis and protein production

The gene coding for DPMS (EC 2.4.1.83) from the euryarcheon Pyrococcus furiosus DSM 3638 (UniProt Q8U4M3) was synthesized and codon optimized for expression in Esherichia coli by DNA20 (DNA20, Inc. Menlo Park, CA, USA). The gene was cloned into a ligation-independent cloning vector into the vector pNiCN28-Bsa41. A PCR protocol using 70 ng plasmid DNA and Pfu DNA polymerase (Fermentas, Germany) included 30 cycles of 95 °C, 3 min; 94 °C, 30 s; 65 °C, 8 s, with a final incubation at 65 °C, 10 min. DpnI-digested PCR products were transformed into E. coli Mach1 cells (Invitrogen). Site-directed mutagenesis experiments to generate the PdPDMS variants D89A, D91A, S135A and Δ230–352 were carried out using the following primers/3′:

- **D89A_fwd** (CTGTGCTGTTGTAAATGGCGGGATCGACATGCACGATC)
- **D89A_rev** (GAGTGCTGAGATCGGCCATATCACAGACACCAGACAG)
- **D91A_fwd** (TTGGAATGCGGCGCGTCAACCGCAGACACCAGACAG)
- **D91A_rev** (GGGCAAGGTGCTAGGCGCGATCGATATACCAAGACAG)
- **S135_fwd** (CGTAAAATCGTATGCGGGCATTATGCTATGG)
- **S135_rev** (CACCAAATGGCCGCATTGCTATTAGC)
- **Δ230–352_fwd** (TACTCTCAACATCGATTAAGTGGCATATCCCCACAC)
- **Δ230–352_rev** (TATTCCACCTTTACGTCAACCTCTACCTTACTCAGAG)

The PCR reactions were carried out using 30 ng plasmid DNA and Fusion High-Fidelity DNA polymerase (Thermo Fisher), and the resulting PCR products were transformed into E. coli DH5α cells grown on LB agar supplemented with 50 μg mL⁻¹ kanamycin.

The recombinant plasmids were transformed into E. coli CA1 (DE3) competent cells. Cultures containing Terrific Broth medium supplemented with kanamycin (50 μg mL⁻¹) and with β-1-thiogalactopyranoside and methionine sulfoximine (1 mM) were harvested, and the bacterial pellet resuspended in 25 mM K₂HPO₄ (pH 7.2), 150 mM NaCl, 5% (v/v) glycerol, with complete protease inhibitor cocktail (Roche). The resuspended pellet was homogenized using an AVESTIN EmulsifyFlex-C system, followed by centrifugation at 10,000 r.p.m. (Beckman Coulter JA-25.50 fixed-angle rotor, 12,090g) with an Avanti J-22XP centrifuge (Beckman Coulter) for 10 min at 4 °C. The resulting supernatant was centrifuged using an Optima L-80K centrifuge (Beckman Coulter 45 Ti rotor) at 35,000 r.p.m. (125,749g) for 70 min (4 °C) to pellet the membrane fraction. The membrane fractions were resuspended in 25 mM K₂HPO₄ (pH 7.2), 150 mM NaCl, 5% (v/v) glycerol with a tissue grinder, flash frozen in liquid nitrogen and stored at −80 °C.

#### Protein purification

Membrane proteins were solubilized from membrane fractions in the presence of 1% n-dodecyl-β-d-maltoside (DDM, Anatrace) in 25 mM K₂HPO₄ (pH 7.2), 150 mM NaCl and 5% (v/v) glycerol at 4 °C. Extracted membrane proteins were recovered by centrifugation at 30,000 r.p.m. (92,387g) at 4 °C (45 Ti rotor, Optima E-80K centrifuge; Beckman Coulter). The supernatant was mixed with Ni-NTA agarose resin (Invitrogen) and packed in Bio-Rad Econo-Pac (45 Ti rotor, Optima E-80K centrifuge; Beckman Coulter). The supernatant was centrifuged with buffer containing 500 mM imidazole. Protein samples were treated with 10 mM ethylenediamine tetraacetic acid and 100 mM ethylenediamine tetraacetic acid, concentrated using 20K centrifugal concentrators (polyethersulfone filter, molecular weight cut-off, MWCO, 50 kDa), and loaded onto a HiLoad 16/60 Superdex 200 prep grade
Activity assays. For the glycosyltransferase (GT) reaction, PJDPMs (0.014 mM enzyme in 50 mM HEPES pH 7.5, 150 mM NaCl, 5% (v/v) glycerol, 0.02% DDM or 0.05% LDAO) was incubated at 75 °C in a volume of 45 μL containing 0.23 mM GDP-Man donor substrate (GDP-Mann, GDP-Glc, UDP-Glc, Sigma-Aldrich). 0.25 mM Dol55-P (C55-dolichyl monophosphate; Larodon, Solna, Sweden; Cat. No. 67-1055) or 0.16 mM Dol95-P (C55-dolichyl monophosphate; Larodon, Solna, Sweden; Cat. No. 67-1095), 10 mM MgCl₂ (or CaCl₂, or MnCl₂). To quantify conversion, we used the dual wavelength method. Baseline readings were made after 75 °C for 10 min, followed by two mild acid hydrolysis as follows: the apolar phase sample was re-dissolved in 50% (v/v) DMSO to 30% (v/v) and dried using a SpeedVac. The residues were dissolved in DMSO and subjected to mild acid hydrolysis as follows: the apolar phase sample was re-dissolved in 50% n-propanol and 50 mM HCl (50% n-propanol) was added, followed by incubation of the sample at 95 °C for 30 min. The treated sample was extracted to separate the mannose-containing polar phase using chloroform, and the apolar phase was washed twice with water. A reaction mixture without addition of PJDPMs and one without Dol55-P were treated identically and used as negative control. For all reaction samples, both apolar and polar phases were analyzed by thin-layer chromatography. Thin-layer chromatography was performed on glass HPTLC silica Gel 60 plate (Merck, Millipore) in chloroform:methanol:water (65:25:4, v:v:v), and stained with a solution of p-anisaldehyde in ethanol and sulfuric acid.

Thermal stability to unfolding. The thermostability of the WT-Dol55-P-Man. The data that support the findings of this study are available from the corresponding author upon request.

Received: 18 February 2017 Accepted: 6 June 2017 Published online: 25 July 2017

References

1. Moreman, K. W., Tiemeyer, M. & Nairn, A. V. Vertebrate protein glycosylation: diversity, synthesis and function. Nat. Rev. Mol. Cell Biol. 13, 448–462 (2012).
2. Eichler, J. Extreme sweetness: protein glycosylation in Archaea.Nat. Rev. Microbiol. 11, 151–156 (2013).
3. Nothauf, H. & Szymanski, C. M. Protein glycosylation in bacteria: sweeter than ever. Nat. Rev. Microbiol. 8, 765–778 (2010).
4. Larkin, A. & Imperials, B. The expanding horizons of asparaginyl-linked glycosylation. Biochemistry. 48, 6412–6423 (2009).
5. Aebi, M. N-linked protein glycosylation in the ER. Biochim. Biophys. Acta 1833, 2430–2437 (2013).
6. Jarrell, K. F. et al. N-linked glycosylation in Archaea: a structural, functional, and genetic analysis. Microbiol. Mol. Biol. Rev. 78, 304–341 (2014).
7. Kobayashi, T., Nishizaki, R. & Ikazawa, H. The presence of GPI-linked protein (s) in an archaeobacterium, Sulfolobus acidocaldarius, closely related to eukaryotes. Biochim. Biophys. Acta 1334, 1–4 (1997).
8. Hirschberg, C. B., Robbins, P. W. & Abeijon, C. Transporters of nucleotide sugars, ATP, and nucleotide sulfate in the endoplasmic reticulum and Golgi apparatus. *Ann. Rev. Biochem.* **67**, 49–69 (1998).

9. Orlean, P., Albright, C. & Robbins, P. W. Cloning and sequencing of the yeast gene for dolichyl phosphate mannosynase, an essential protein. *J. Biol. Chem.* **263**, 17499–17507 (1988).

10. Hirschberg, C. B. & Snider, M. D. Topography of glycosylation in the rough endoplasmic reticulum and Golgi apparatus. *Ann. Rev. Biochem.* **56**, 63–87 (1987).

11. Saney, S. & Menon, A. K. Stereoselective transblayer translocation of mannosyl phospholipid dolichol by an endoplasmic reticulum flipase. *Proc. Natl. Acad. Sci. USA* **107**, 11289–11294 (2010).

12. Albers, S. A. & Meyer, B. H. The archaeal cell envelope. *Nat. Rev. Microbiol.* **9**, 414–426 (2011).

13. Fujinami, D., Matsumoto, M., Noguchi, T., Sonomoto, K. & Kohda, D. Structural elucidation of an asparagine-linked oligosaccharide from the hyperthermophilic archaen *Pyrococcus furiosus*. *Carbohydr. Res.* **387**, 30–36 (2014).

14. Milde-Langosch, K. et al. Prognostic relevance of glycosylation-associated genes in breast cancer. *Breast Cancer Res. Treat.* **145**, 295–305 (2014).

15. Lefebre, D. J. et al. Deficiency of Dol-P-Man synthase subunit DPM3 breeds the congenital disorders of glycosylation with the dystroglycanopathies. *Am. J. Hum. Genet.* **85**, 76–86 (2009).

16. Shah, S. D. H. et al. Identification of embryonic lethal genes in humans by autozygozity mapping and exome sequencing in consanguineous families. *Genome Biol.* **16**, 116 (2015).

17. Colussi, P. A., Taron, C. H., Mack, J. C. & Orlean, P. Human and *Saccharomices cerevisiae* dolichyl phosphate mannosynase syntheses represent two classes of the enzyme, but both function in *Schizosaccharomyces pombe*. *Proc. Natl. Acad. Sci. USA* **94**, 7973–7978 (1997).

18. Maeda, Y. & Kinoshita, T. Dolichyl-phosphate mannosynase structure: Function, regulation and evolution. *Biochim. Biophys. Acta* **1870**, 861–888 (2008).

19. Lombard, V., Golaonca Ramulu, H., Drula, E., Coutinho, P. M. & Henrissat, B. The Carbohydrate-active enzymes database (CAZy) in 2013. *Nucleic Acids Res.* **42**, D490–D495 (2014).

20. Breton, C., Snajdrova, L., Jeanneau, C., Koca, J. & Imbert, A. Structures and mechanisms of glycosyltransferases. *Glycobiology* **16**, 29R–37R (2006).

21. Charnock, S. J. & Davies, G. J. Structure of the nucleotide-diphospho-sugar transferase, SpmA from *Bacillus subtilis*, in native and nucleotide-complexed forms. *Biochemistry* **38**, 6380–6385 (1999).

22. Kissin, E. & Henrici, K. Inference of macromolecular assemblies from crystalline state. *J. Mol. Biol.* **372**, 774–797 (2007).

23. Zhou, G. P. & Troy, F. A. II NMR study of the preferred membrane orientation of porins with the membrane morphologies of *Pyrococcus furiosus*. *J. Mol. Biol.* **347**, 615–628 (2015).

24. Baksi, K., Tavarez-Pagan, J. J., Martinez, J. A. & Banerjee, D. K. Mannosylphosphodolichol synthase, SpsA from *Bacillus subtilis* forms a dodecamer. *Biochim. Biophys. Acta* **1838**, 3499–3509 (2014).

25. Chang, M. M., Imperiali, B., Eichler, J. & Guan, Z. N-Linked glycans are essential for the congenital disorders of glycosylation with the dystroglycanopathies. *Proc. Natl Acad. Sci. USA* **112**, 15612–15617 (2015).

26. Banerjee, D. K. et al. Identification of archaeal cell envelope. *Nat. Rev. Microbiol.* **9**, 414–426 (2011).

27. Kabsch, W. XDS. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 125–132 (2010).

28. von Heijne, G., Blanc, E., Roversi, P. & Brígido, G. Automated structure solution with autoSHARP. *Methods Mol. Biol.* **364**, 215–230 (2007).

29. Brígido, G., von Heijne, G., Flensburg, C., Schiltz, M. & Paciorek, W. 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 (2003).

30. Abrahams, J. P. & Leslie, A. G. W. Methods used in the structure determination of bovine mitochondrial F1 ATPase. *Acta Crystallogr. D Biol. Crystallogr.* **52**, 30–42 (1996).

31. Emley, P. & Cowtan, K. Coot: model-building tools for molecular graphics. *Acta Crystallogr. D Biol. Crystallogr.* **60**, 2126–2132 (2004).

32. 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. C** **47**, 110–119 (1991).

33. Adams, P. D. et al. PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 213–221 (2010).

34. Wu, Z. L., Ethen, C. M., Prather, B., Machacek, M. & Jiang, W. Universal phosphatase-coupled glycosyltransferase assay. *Glyobiology* **21**, 727–733 (2011).

35. Baykov, A. A., Evtushenko, O. A. & Avava, S. M. A malachite green procedure leading to these results has received funding from the European Community’s Framework Programme (FP7/283570). Part of this work was facilitated by the Protein Science Facility at Karolinska Institute/ScilifeLab (http://psf.ks.se).

Acknowledgements

We thank the beamline staff scientists for support during data collection at beamlines PROXIMA 1 at SOLEIL (France) and IO3 at Diamond Light Source (UK). We acknowledge financial support to C.D. from the Swedish Research Councils Formas (grants no 2012–915) and VR (grants no’s 2011–5768 and 2013–5717). 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 no 283570). Part of this work was facilitated by the Protein Science Facility at Karolinska Institute/ScilifeLab (http://psf.ks.se).

Author contributions

T.-C.T. performed the original cloning and expression of the *P. furiosus* *dpm* gene; T.R., T.-C.T., and R.G. performed the mutagenesis experiments; R.G. performed gene expression, protein purification, protein characterization, protein crystallization, heavy-atom screening, and synchrotron data collection; C.D. performed heavy-atom phasing, structure determination, and supervised the work; R.G. and C.D. performed model refinement and analysis of structural and biochemical data; R.G. and C.D. designed the experiments and analyzed the data. R.G., T.R., T.-C.T., and C.D. wrote the manuscript.

Additional information

Supplementary Information accompanies this paper at doi:10.1038/s41467-017-00187-2.

Competing interests: The authors declare no competing financial interests.

Reprints and permission information is available online at http://npg.nature.com/reprintsandpermissions/
