

Dolichyl-Phosphate Alpha-N-Acetylglucosaminyl transferase (DPAGT1)

A Target Enabling Package

Gene ID / Uniprot ID / EC	9606 / Q9H3H5 / E.C. 2.7.8.15
Target Nominator	David Beeson (WIMM, Oxford)
Authors	Yin Yao Dong*, Ashley C.W. Pike*, Amy Chu, Simon Bushell, Leela Shrestha, Georgina Berridge, Shubhashish Mukhopadhyay, Nicola A. Burgess-Brown, Elisabeth Carpenter
Target PI	Liz Carpenter (Oxford)
Therapeutic Area	Neuropsychiatry and neuro genetic disorders
Disease Relevance	Mutations in DPAGT1 cause congenital disorder of glycosylation type 1j (CDG: OMIM 608093) and congenital myasthenic syndrome with tubular aggregates (CMS: OMIM: 614750)
Date	2 nd June 2017
DOI	https://doi.org/10.5281/zenodo.1219694 .
Document Version	Version 2

SUMMARY OF PROJECT

The ER integral membrane enzyme dolichyl-phosphate alpha-N-acetyl glucosaminyl phosphotransferase (DPAGT1) catalyses the first step in the synthesis of the oligosaccharide-P-P-dolichol unit which provides the glycans structure for N-glycosylation of proteins. Mutations in DPAGT1 cause two muscle weakness conditions, limb-girdle congenital myasthenic syndrome (CMS) and congenital disorder of glycosylation type 1j (CDG1j). DPAGT1 overexpression has also been implicated in oral cancer. We have produced and solved structures of this integral membrane enzyme, DPAGT1 with the V264G mutation found in a patient with CMS, and complexes with a 50 nM inhibitor, tunicamycin. We have developed enzymatic activity and thermostability assays which have allowed us to assess the activity and stability of DPAGT1 mutants and the effect of small molecules. There are > 20 DPAGT1 associated missense variants in patients with CMS and CDG1j. We have mapped these mutations to the structure, and we will use the assays described here to assess how the activity and stability of DPAGT1 is affected by these missense variants.

USEFUL LINKS



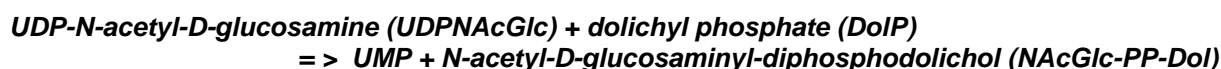
Open Targets



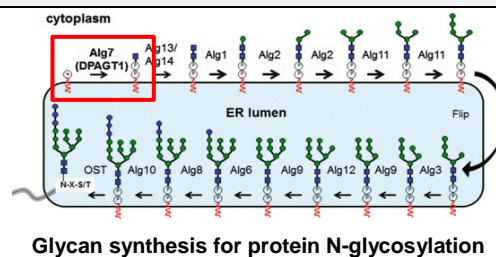
SCIENTIFIC BACKGROUND

Biology background

One of the most common post-translational modifications of proteins involves the addition of complex branched-chain glycans to asparagine residues. The first step in the preparation of the oligosaccharides is catalysed by the glycosyl transferase dolichyl-phosphate N-acetylglucosamine phosphate transferase (DPAGT1; E. C. 2.7.8.15). DPAGT1 transfers an N-acetyl glucosaminyl phosphate unit onto dolichyl phosphate (DoIP):



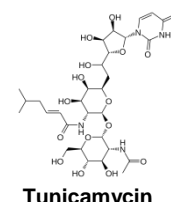
DoIP is the lipid carrier that anchors the oligosaccharides to the ER membrane as they are extended by a series of enzymes, flipped into the ER, further processed, then transferred onto asparagine residues. DPAGT1 is an integral membrane endoplasmic reticulum enzyme, with its catalytic site in the cytoplasm.



Genetic linkage - Exome sequencing, followed by Sanger sequencing identified mutations in DPAGT1 in patients with limb-girdle congenital myasthenia syndrome with tubular aggregates (Belaya *et al.*, 2012) and congenital disorders of glycosylation type Ij (Wu *et al.*, 2003, Wurde *et al.*, 2012, Iqbal *et al.*, 2013). These diseases are both autosomal recessive. Mutations are either homozygous or compound heterozygous. In general the lower the residual DPAGT1 activity, the more severe the disease. Complete loss of DPAGT1 activity is lethal *in utero* in mice (Thiel and Korner, 2011).

Link to oral cancer – Overexpression of DPAGT1 has been linked to oral cancer. Expression of DPAGT1 is controlled by a series of transcription factors including those involved in WNT signalling. Overexpression of DPAGT1 leads to aberrant glycosylation of both WNT and β -catenins, resulting in changes in the proliferation and adhesion of cancer cells (Varelas *et al.*, 2014).

Antibiotic development – Tunicamycin is a natural product made by streptomyces bacteria, which has powerful antibacterial effects, due to its ability to inhibit the bacterial cell wall synthesis protein MraY. However, tunicamycin is not useful as an antibiotic as it also inhibits DPAGT1, thus interfering with glycosylation of proteins, leading to cell cycle arrest and induction of the unfolded protein response. MraY transfers a pentapeptide from a UDP-pentapeptide to an undecaprenyl phosphate lipid carrier (Chung *et al.*, 2013).



RESULTS – The TEP

Purification

Full length human DPAGT1 with an N-terminal 6-His-TEV protease tag was expressed using the baculovirus/insect cell expression system. The protein was extracted and purified using the octyl glucose neopentyl glycol (OGNG) detergent, supplemented with cholesteryl hemisuccinate (CHS) (10:1 ratio, wt:wt) for extraction. Protein was purified in 0.18% OGNG, 0.018% CHS and 0.0036% cardiolipin containing buffer, using Talon resin, TEV cleavage, reverse purification and size exclusion chromatography. The final yield was 0.8 mg/Litre of cell culture and a final protein concentration of 27 mg/mL.

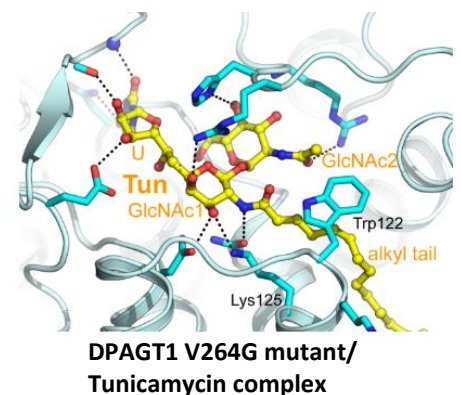
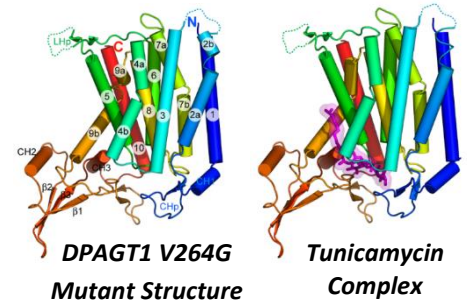
Crystal structures

DPAGT1 X-ray data and structures obtained:

- 3.2 Å V264G CMS mutation structure (PDB: 5LEV)
- 3.4 Å Tunicamycin complex (with V264G; PDB: 5O5E)

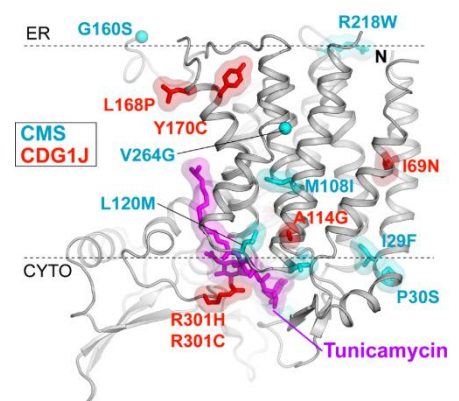
We have solved the structure of DPAGT1 with the point mutation V264G to 3.2 Å resolution (PDB: 5LEV). The structure shows 10 transmembrane helices, with the N- and C-termini in the ER lumen and the active site on the cytoplasmic face of the membrane. The overall fold resembles the bacterial cell wall synthesis protein MraY (Chung *et al.*, 2013), although the two proteins only have 23% sequence similarity. The DPAGT1 has an 82 residue domain (residues Leu293 to Ile374) between helices H9b and H10 that forms one side of the active site on the cytoplasmic face of the membrane and an luminal loop between helices H4b and H5 on the ER lumen face of the membrane, that may form part of the DolP binding surface.

The active site of DPAGT1 lies on the cytoplasmic face of the protein and is formed by the five loops between the TM helices. The complex with the substrate-analogue inhibitor tunicamycin (at 3.4 Å) shows how tunicamycin interacts with DPAGT1, occupying the substrate binding sites. Tunicamycin includes both a Uracil base, a GlcNAc moiety, linked by a ring structure that replaces the diphosphate bridge. It also has an aliphatic chain which lies in the DolP binding site. The uridine base lies between the extensive hairpin between transmembrane (TM) helices 1 and 2, and the cytoplasmic loops (CL) 5 and 7. The ribose moiety lies between CL3 and CL5, and the GlcNAc moiety lies between the extensive CL9 domain and CL5. The tunicamycin acyl chain extends across the protein surface between TM4, TM5 and TM9b, suggesting that the 90 carbon dolichol phosphate chain would pack against these helices, potentially also interacting with the luminal loop between H4 and H5. When tunicamycin binds, the sidechain of Trp122 rotates so that it covers the alkyl chain, thus locking the tunicamycin in place.



Disease associated mutations

More than 20 missense variants have been identified in DPAGT1 in patients with CMS or CDG1j. Mutations were first found in patients with CDG1j in 2003 (Wu *et al.*, 2003), with a total of 8 different mutations identified so far in 6 patients (Iqbal *et al.*, 2013; Wurde *et al.*, 2012). Clinical phenotypes can vary, but all patients had impaired cognitive and motor development. DPAGT1 missense variants were first identified in CMS patients in 2012 (Belaya *et al.*, 2012). Since then, 17 mutations have been published in 13 patients. Clinical phenotypes vary between patients, but they all have muscle weakness, most have limb girdle motor impairments, and some patients also have cognitive impairment. Complete loss of DPAGT1 function causes complete loss of N-glycosylation, which would be fatal. These mutations are likely to cause partial loss of function. In cases where different mutations occur in the two alleles, one mutation may cause complete loss of function whereas the other change may leave the enzyme partially active. Mutations such as R301H lie directly in the substrate binding site and they are likely to cause extensive loss of function. Mutations that are further from the active site, may cause smaller changes in activity. A



Mutations found in CMS (blue) and CDG1j (red) patients mapped onto the DPAGT1 structure

full analysis of the activity and stability of these mutants is in progress using the assays described below.

Chemical matter

- The antibiotic tunicamycin (natural product purified from *Streptomyces* sp. mixture of different alkyl chain lengths – n=8,9,10,11 (5%/32%/42%/20%)), which is an analogue of both substrates and has a K_i of 50 nM

Biochemical and biophysical assays

We have developed both thermostability and enzymatic activity assays for DPAGT1.

- **Thermostability assays using tryptophan fluorescence as an indication of protein unfolding**

- Equipment: the Prometheus from NanoTemper

In order to assess the thermostability of WT and mutant DPAGT1, as well as the effect of the addition of potential binders, we used the Prometheus system to follow changes in tryptophan and tyrosine fluorescence with temperature.

The $T_{m1/2}$ for DPAGT1 WT and V264G mutant are very similar: 51.6°C for the WT and 50.2°C for the V264G mutant, suggesting that the mutation does not lead to loss of protein stability.

The substrate UDP-GlcNAc causes an increase in stability of 4°C for both the WT and mutant proteins.

UDP and UMP do not change the thermostability of either protein.

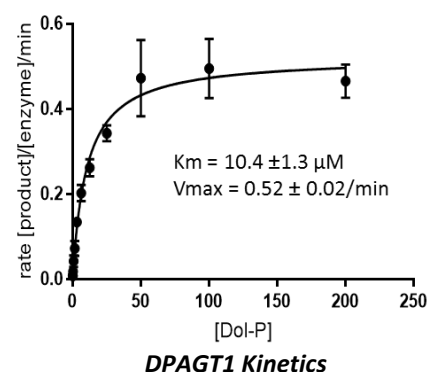
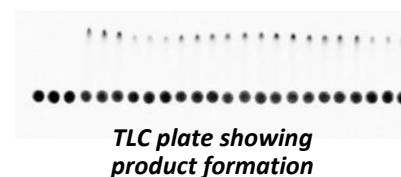
Tunicamycin causes an apparent increase in $T_{m1/2}$ of >30°C with both proteins.

- **Activity assays:**

We have now developed an activity assay using UDP-N-acetyl [1- 14 C] D-glucosamine as the substrate and detection of the labelled NAcGlc-PP-Dol product using of thin layer chromatography and a phosphorimager. This assay measures the transfer of the 1- 14 C labelled N-acetyl D-glucosamine monophosphate from the soluble UDP-N-acetyl [1- 14 C] D-glucosamine onto the 95-carbon DolP to form the more hydrophobic NAcGlc-PP-Dol product. Phosphor imaging and densitometry were used to quantify the relative amounts of substrate and product present.

Using this assay, we have determined that the K_m of UDP-GlcNAc is approximate 20 μ M and K_m of DolP is around 10 μ M. The V_{max} of the DPAGT1 is approximately 0.5 molecules of product formed per enzyme molecule per minute.

We have also demonstrated that tunicamycin inhibits 100% of DPAGT1 activity at a 1:1 molar ratio.



Future plans

- We are in the process of testing the enzymatic activity and thermostability of a series of mutated DPAGT1 proteins.
- We are working with Ben Davis's group in Chemistry, Oxford, to test novel tunicamycin analogues that are selective for MraY, while avoiding inhibition of DPAGT1.

CONCLUSION

Biological understanding

- We show how **tunicamycin** binds and **inhibits** DPAGT1
- The structures indicate the **binding sites of both substrates**
- There are **>20 mutations found in CMS and CDG patients**. Mapping the mutations on the structure has allowed us to identify the likely mechanisms for loss of activity

Medical relevance

- An understanding of how mutations affect DPAGT1 activity **will assist clinicians** in deciding on treatment.
- It is likely that in cases where missense mutations are found in patients, the mutations will cause only a modest drop in activity of the second allele. Using the TEP assays, the effect of mutations on activity and thermostability will be determined and it may be possible to identify activators.
- The antibiotic tunicamycin is highly toxic as it inhibits DPAGT1. The structure and assays of DPAGT1 will assist our collaborator Ben Davis (Chemistry, Oxford) with design and selection of derivatives of tunicamycin that do not inhibit DPAGT1, but retain activity on MraY and other bacterial homologues. As MraY is necessary for bacterial cell wall formation, this work has the potential to provide new leads for antibiotic therapy.

Collaborations

The following collaborations have been established.

- **David Beeson, (WIMM, Oxford)** – Studies the causes of CMS in patients at the Radcliffe Hospital, Oxford
- **Ben Davis, (Chemistry, Oxford)** – Chemist interested in design of novel tunicamycin analogues that will inhibit MraY, but not DPAGT1

FUNDING INFORMATION

The work done by SGC has been funded by a grant from the Wellcome [106169/ZZ14/Z].

ADDITIONAL INFORMATION

Materials & Methods	DPAGT1 (5LEV, 5O5E)	Procedures for expression, purification, crystallization (see below)
Structure files	PDB codes 5LEV ; 5O5E	Structures of DPAGT1

Supplemental Information

1. References

- Belaya, K., Finlayson, S., Slater, C.R., Cossins, J., Liu, W.W., Maxwell, S., McGowan, S.J., Maslau, S., Twigg, S.R., Walls, T.J., et al. (2012). *Mutations in DPAGT1 cause a limb-girdle congenital myasthenic syndrome with tubular aggregates*. Am J Hum Genet 91, 193-201.
- Chung, B.C., Zhao, J., Gillespie, R.A., Kwon, D.Y., Guan, Z., Hong, J., Zhou, P., and Lee, S.Y. (2013). *Crystal structure of MraY, an essential membrane enzyme for bacterial cell wall synthesis*. Science 341, 1012-1016.
- Iqbal, Z., Shahzad, M., Vissers, L.E., van Scherpenzeel, M., Gilissen, C., Razzaq, A., Zahoor, M.Y., Khan, S.N., Kleefstra, T., Veltman, J.A., et al. (2013). *A compound heterozygous mutation in DPAGT1 results in a congenital disorder of glycosylation with a relatively mild phenotype*. Eur J Hum Genet 21, 844-849.
- Thiel, C., and Korner, C. (2011). *Mouse models for congenital disorders of glycosylation*. J Inherit Metab Dis 34, 879-889.
- Varelas, X., Bouchie, M.P., and Kukuruzinska, M.A. (2014). *Protein N-glycosylation in oral cancer: dysregulated cellular networks among DPAGT1, E-cadherin adhesion and canonical Wnt signaling*. Glycobiology 24, 579-591.
- Wu, X., Rush, J.S., Karaoglu, D., Krasnewich, D., Lubinsky, M.S., Waechter, C.J., Gilmore, R., and Freeze, H.H. (2003). *Deficiency of UDP-GlcNAc:Dolichol Phosphate N-Acetylglucosamine-1 Phosphate Transferase (DPAGT1) causes a novel congenital disorder of Glycosylation Type Ij*. Hum Mutat 22, 144-150.
- Wurde, A.E., Reunert, J., Rust, S., Hertzberg, C., Haverkamper, S., Nurnberg, G., Nurnberg, P., Lehle, L., Rossi, R., and Marquardt, T. (2012). *Congenital disorder of glycosylation type Ij (CDG-Ij, DPAGT1-CDG): extending the clinical and molecular spectrum of a rare disease*. Mol Genet Metab 105, 634-641.

2. Protein expression and purification procedures

Vector: pFB-LIC-Bse (available from The Addgene Nonprofit Plasmid Repository)

Cell line: DH10Bac, Sf9 cells

Tags and additions: N-terminal, 6-His, TEV protease cleavable tag

Wild-type sequence:

MGHHHHHHSSGVDLGTENLYFQSMWAFSELPMPLLINLIVSLLGFVATVTLIPAFRGHFIAARLCG
QDLNKTSRQQIPESQGVISGAVFLILFCFIPFPFLNCFVKEQCKAFPHHEFVALIGALLAICCMIFLG
FADDVLNLRWRHKLPTAASLPLLMVYFTNFGNTTIVVPKPFPRPILGLHLDLGILYYVYMGLLAVF
CTNAINILAGINGLEAGQSLVISASIVFNLVELEGDCRDDHVFSLYFMIPFFFTTLGLLYHNWYPSR
VFVGDTFCYFAGMTFAVVGILGHFSKTMLLFFMPQVFNFYSLPQLLHIIPCPHRHRIPRLNIKTGKL
EMSYSKFKTKSLSFLGTFILKVAESLQLVTVHQSETEDGEFTECNMNTLINLLKVLGPIHERNLTL
LLLLLQILGSAITFSIRYQLVRLFYDV

The Val264Gly mutant:

MGHHHHHHSSGVDLGTENLYFQSMWAFSELPMPLLINLIVSLLGFVATVTLIPAFRGHFIAARLCG
QDLNKTSRQQIPESQGVISGAVFLILFCFIPFPFLNCFVKEQCKAFPHHEFVALIGALLAICCMIFLG
FADDVLNLRWRHKLPTAASLPLLMVYFTNFGNTTIVVPKPFPRPILGLHLDLGILYYVYMGLLAVF
CTNAINILAGINGLEAGQSLVISASIVFNLVELEGDCRDDHVFSLYFMIPFFFTTLGLLYHNWYPSR
VFVGDTFCYFAGMTFAVVGILGHFSKTMLLFFMPQVFNFYSLPQLLHIIPCPHRHRIPRLNIKTGKL
EMSYSKFKTKSLSFLGTFILKVAESLQLVTVHQSETEDGEFTECNMNTLINLLKVLGPIHERNLTL
LLLLLQILGSAITFSIRYQLVRLFYDV

Expression: The full length gene for human DPAGT1 was obtained from the Mammalian gene collection and the Val264Gly mutant gene was obtained from Prof. David Beeson, WIMM, Oxford. This gene was cloned into the pFB-LIC-Bse vector and baculoviruses were produced by transformation of DH10Bac cells. *Spodoptera frugiperda* (Sf9) insect cells in Sf-900 II SFM medium (Life Technologies) were infected with recombinant baculovirus and incubated for 65 h at 27°C in shaker flasks.

Cell Lysis and detergent extraction of membrane proteins:

Extraction Buffer, EXB: 50 mM HEPES (pH 7.5), 5mM MgCl₂, 200 mM NaCl, 5mM Imidazole, 2mM TCEP, 5% Glycerol, Roche protease inhibitors

The cell pellet from 1 litre of insect cell culture was resuspended in 40ml of lysis buffer in warm water, mixing constantly to keep the sample cold. Cells were lysed with a EmulsiFlex-C3 homogenizer (Aventin) (at room temperature, 2 passes.) 5 mL of 10%:1% (w:w) stock of OGNG/CHS was added per litre of cell culture and the volume was adjusted by the addition of EXB to a final volume of 50 mL/L of cell culture. The sample was rotated in the cold room for 1h, then unlysed cells and cell debris was removed by centrifugation at 35,000g for 45 min at 4°C.

Column 1: Co²⁺ talon resin (0.5 ml volume in a gravity-flow column):

Wash Buffer: 50 mM HEPES (pH 7.5), 5mM MgCl₂, 10 mM imidazole (pH 8.0), 200 mM NaCl , 2 mM TCEP, 5 % Glycerol, 0.18 % OGNG, 0.018 % CHS, 0.0036 % cardiolipin

Elution Buffer: 50 mM HEPES (pH 7.5), 5 mM MgCl₂, 250 mM imidazole (pH 8.0) , 200 mM NaCl, 2 mM TCEP, 5% Glycerol, 0.18 % OGNG, 0.018 % CHS, 0.0036 % cardiolipin

SEC Buffer: 20 mM HEPES (pH 7.5) , 5mM MgCl₂, 200 mM NaCl, 2 mM TCEP, 0.12 % OGNG, 0.012 % CHS, 0.0024 % cardiolipin

The detergent-extracted membrane protein from each litre of cells was combined with 1ml of pre-equilibrated slurry of 50% Co²⁺ talon resin (previously washed 2 times with H₂O and 4 times with EXB). The sample was rotated in the cold room for 1h then poured onto an econo column. The residual talon resin was washed with 5ml wash buffer and pipetted onto the econo column. The talon resin was washed with 30x talon resin volume of WB and the protein eluted with 2x Talon resin volume of elution buffer in 1ml fractions. The peak fractions were combined and the imidazole removed using two PD10 columns per purification (pre-equilibrated with SEC buffer). Protein was eluted from the PD10 columns using SEC buffer.

TEV protease cleavage and reverse purification

TEV protease was added at a ratio of 10:1 (DPAGT1:TEV protease (w:w)). Samples were incubated on ice in the cold room overnight. 2 mL of a 50 % slurry of Talon resin (pre-equilibrated as above) were added and the sample was rotated in the cold room for an hour, poured into an econo column and the flowthrough collected and centrifuged at 21,500 rpm in a Beckman JA25.5 rotor for 5 min at 4°C. The supernatant was removed and kept.

Column 2: Superdex 200 Size Exclusion Chromatography

The protein sample was concentrated in a 30 kDa PES concentrator (pre-equilibrated with SEC buffer) and concentrated, with mixing every 5 mins, to a final volume of 500 µL. After centrifugation at 20,000 g for 10 min the sample was subjected to size exclusion chromatography on a Superdex 200 column in SEC buffer. DPAGT1 protein was then concentrated with a Sartorius 2ml PES 50 kDa concentrator pre-equilibrated with SEC buffer (without detergent), at 3,220 g, with mixing every 5 mins. The protein was centrifuged at 20,000 g for 15 mins, then flash frozen in liquid nitrogen. The final concentration was ~30 mg/mL.

3. Assay protocols

Thermostability using the tryptophan fluorescence as readout

Samples with a volume of 40 μ L were prepared containing 0.5 mg/mL protein and 50 μ M compound or 5% DMSO in SEC buffer and a glass capillary was dipped into each sample, with the capillary held horizontally to ensure that it was full of the sample. The capillaries were placed on the capillary holder on the Prometheus. Triplicates of each sample were prepared. A discovery scan was performed to ensure that the values were roughly the same for all the samples. A melting curve from 20°C to 95°C at 5°C/min was performed. The minimum of the first derivative of the 330/350 nm ratio was used to determine the $T_{m1/2}$.

Activity assays using radiolabelled 14 C UDP-GlcNAc

2 μ L of 2 μ M protein in SEC buffer supplemented with 5mM extra $MgCl_2$ and 1% OGNG/CHS/cardiophilin and dolichyl monophosphate was combined with 2 μ L of UDP-N-acetyl [14 C] D-glucosamine in the same buffer and incubated at 37 °C for 21 min. The reaction was terminated by the addition of 6 μ L of 100% methanol and immediately transferred onto ice. 1 μ L of sample is spotted onto a silica coated TLC plate in triplicate and run with a mobile phase consisting of chloroform, methanol, and water at a 65:25:4 relative ratios. After running the TLC plate was thoroughly dried, wrapped in cling film, incubated with a phosphor imaging substrate for 4 days and phosphor was then imaged. The pixel density of the spot corresponding to the hydrophobic product is divided by combined pixel density of the product and the substrate and multiplied by the known concentration of substrate added to ascertain the concentration of product formed.

3. Structure determination procedures

All data were collected at Diamond Light Source (beamlines I24, I04-1 and I04) to resolutions between 3.2-3.6 Å. Data were processed, reduced and scaled using XDS and AIMLESS.

The crystals belong to space group $P6_522$ and contain a single DPAGT1 monomer in the asymmetric unit with 70 % solvent. Initial phase estimates were obtained using molecular replacement (MR). Briefly, an initial search model was built automatically using the PHYRE server based on the coordinates of the bacterial homolog MraY (23 % sequence identity; PDB: 4J72). However, this simple homology model failed to produce any meaningful MR solutions when used in isolation in PHASER. The PHYRE model was then subjected to model pre-refinement using the procedures implemented in MR-ROSETTA in PHENIX and the resultant five best-scoring output models were trimmed at their termini and the TM9/TM10 cytoplasmic loop region and superposed for use as an ensemble search model in PHASER. A marginal but consistent solution was obtained that exhibited sensible crystal packing in space group $P6_522$ but both the initial maps and model refinement were inconclusive. The model positioned using PHASER was converted to a poly-alanine trace and recycled into MR-ROSETTA, using model_already_placed=True option. The resultant MR-ROSETTA output model had an R/Rfree of 42/46 and the electron density maps showed new features not present in the input coordinates that indicated that the structure had been successfully phased.

Using the MR-ROSETTA solution as a starting point, the remaining regions of the DPAGT1 structure could be built manually using COOT and the wildtype 3.6 Å native data. However, the novel 54 amino acid cytoplasmic insertion domain between TM9 and TM10 was poorly ordered and proved difficult to trace. This region was primarily traced using the electron density maps for a UDP-GlcNAc complex as substrate binding results in partial stabilisation of the TM9/10 insertion domain. Sequence assignment was aided by using the sulphur anomalous signal from a dataset collected at a wavelength of 1.7 Å. Anomalous difference maps combined with anomalous substructure completion using PHASER-EP clearly revealed the location of 18 of the expected 22 sulphur positions and helped to confirm the sequence register. Additional experimental phasing information was provided by a Pr^{3+} derivative. The resultant model for the entire chain was then refined against both the apo V264G (3.2 Å) and V264G tunicamycin (3.4 Å). The representative final model comprises the entire polypeptide chain between residues Leu7 and Gln400 apart from a flexible loop connecting TM2 and TM3 and part of the poorly-ordered luminal hairpin (152-161).