Published in final edited form as:
Nature. 2017 November 23; 551(7681): 521–524. doi:10.1038/nature24464.

Structural basis of nucleotide sugar transport across the Golgi membrane

Joanne L Parker[1,*] and **Simon Newstead**[1,*]

[1]Department of Biochemistry, University of Oxford, South Parks Road, Oxford, OX1 3QU, UK

Abstract

Glycosylation is a fundamental cellular process that in eukaryotes occurs in the lumen of both the Golgi apparatus and endoplasmic reticulum 1. Nucleotide sugar transporters (NSTs) are an essential component of the glyscosylation pathway, providing the diverse range of substrates required for the glycosyltransferases 2,3. NSTs are linked to several developmental and immune disorders in humans and in pathogenic microbes play an important role in virulence 4–8. How NSTs recognise and transport activated monosaccharides however is currently unclear. Here we present the first crystal structure of an NST, the GDP-mannose transporter Vrg4, in both the substrate free and bound states. A hitherto unobserved requirement for short chain lipids in activating the transporter supports a model for regulation within the highly dynamic membranes of the Golgi apparatus. Our results provide a structural basis for understanding nucleotide sugar recognition and provide insights into the transport and regulatory mechanism for this family of intracellular transporters.

Within eukaryotes glycosylation is compartmentalised to maintain both high efficiency and tight regulation and occurs within the lumen of both the endoplasmic reticulum and Golgi apparatus 9. However, the substrates for the glycosylation enzymes, nucleotide sugars, are synthesised in either the cytoplasm or the nucleus and are unable to cross the organellular membranes. NSTs are therefore required to import into the Golgi and ER lumen all the nucleotide sugars required for glycosylation (Figure 1a). NSTs belong to the SLC35 family of solute carriers and their function is highly conserved from simple eukaryotes, fungi and parasites to plants and mammals 10. NSTs are amongst only a small number of transporters implicated in Mendelian diseases, such as Leukocyte adhesion deficiency II caused by point mutations within the transporter specific for GDP-fucose 4,8 and Congenital Disorder of Glycosylation IIf (CDG-IIf), which is caused through mutations in the CMP-sialic acid transporter 5.

[*]Correspondence and requests for materials should be addressed to: joanne.parker@bioch.ox.ac.uk or simon.newstead@bioch.ox.ac.uk.

Author contributions
JLP and SN designed and carried out the experiments, interpreted the data and wrote the paper.

Author Information
Reprints and permissions information is available at www.nature.com/reprints.

Competing financial interests
The authors declare no competing financial interests.

Europe PMC Funders Author Manuscripts

The transport of nucleotide sugars into the secretory pathway is also important for pathogenic fungi and trypanosomatid parasites, as these organisms contain a cell wall or surface glycocalyx predominantly formed of glycomannosylated conjugates that form a protective coat against the human immune system 6,7. In these organisms, GDP-mannose transport is fundamental for virulence making them attractive targets for inhibitor design 11,12. However, given the lack of any structural information, the molecular basis of nucleotide sugar recognition and transport remains poorly understood 13.

To obtain structural information about NSTs we determined the crystal structure of Vrg4 to 3.2 Å resolution (Extended data Table 1 & Extended Data Figure 1a and b). Vrg4 adopts a compact structure, approximately 30 Å in length, consisting of 10 transmembrane α-helices connected by short loops (Figure 1b & Extended Data Figure 2). The topology is similar to the drug and metabolite transporter (DMT) superfamily, to which the NSTs belong 14,15. The N- and C-termini are located on the same side of the transporter, which would correspond to the cytoplasmic side of the Golgi membrane 16. The Golgi retrieval sequence, K^{326}QKKQQ, that binds to COPI vesicles and enables localisation of Vrg4 to the Golgi membranes 17 forms a structured region at the very C-terminus. Interestingly TM10 is noticeably vertical with respect to the plane of the membrane, resulting in this sequence protruding above the transporter and likely to aid accessibility to the COP1 adaptor proteins 17.

TM6-7 are packed against TM8-9 to seal a large cavity from the cytoplasm, whereas TM1-2 and TM3-4 are opened out, providing a snapshot of the transporter in the state closed to the cytoplasm and open to the lumen of the Golgi (Figure 1c). The central cavity is positively charged suggestive of a potential site for sugar and phosphate binding. To analyse the substrate specificity, Vrg4 was reconstituted into proteoliposomes. The reconstituted protein acts as a strict antiporter 18 requiring either GDP-mannose or guanine monophosphate (GMP) to function, with poor transport observed with adenine monophosphate (AMP) and no significant transport with the pyrimidine base, cytosine monophosphate (CMP) (Figure 1d). Our assays show that the central binding site can discriminate between mono, di- and tri-phosphates (Extended Data Figure 1c and d). Intriguingly we noticed that the transport rate was faster when GDP mannose, as opposed to GMP, was used on the inside of the liposome (Figure 1e), implying a potential asymmetry within the transport mechanism.

The Golgi is known to have a different lipid composition in comparison to the plasma membrane, which gives rise to a thinner bilayer thickness 19. On reconstitution into different lipid vesicles we observed a marked effect on transport function (Figure 1f), highlighted by a severe reduction in transport with vesicles composed of pure POPE:POPG (3:1) lipids compared to yeast polar lipids (YPL). POPE:POPG have been extensively used to characterise a wide range of different membrane proteins 20,21, raising the question of why Vrg4 was not functional in these lipids. Phosphatidylglycerol comprises only a small percentage of the phospholipid in eukaryotes 22. We first addressed whether the lack of activity seen in POPE:POPG was due to the lack of phosphatidylcholine, the major structural lipid in eukaryotes. However, activity was still not observed in vesicles made either of pure POPC, POPC:POPE, or POPC spiked POPE:POPG vesicles. These results rule out the requirement of a specific lipid head group. Similarly, we observed no transport

upon addition of ergosterol. Activity was only observed when the proteoliposomes contained lipids from natural sources. We therefore reasoned that the lack of function of Vrg4 in POPE:POPG lipids was a result of hydrophobic mismatch 23. The 1-palmitoyl-2-oleoyl (PO) lipids consist of one 16 and one 18 carbon fatty acid (FA). We found that the addition of myristoyl FA (14 carbons) in the form of 1,2-dimyristoyl-sn-glycero-3-phosphocholine (DMPC) into the POPE:POPG vesicles could now support the function of Vrg4 (Figure 1f). The Golgi occupies a central position within the secretory pathway where it functions as a hub for vesicle trafficking. Many resident proteins are swept along the secretory pathway and require retrieval back to the Golgi complex 24,25. The extreme case of hydrophobic mismatch we observe for Vrg4 may reflect a mechanism within the NSTs to tightly regulate their function, ensuring maximal activity occurs only in the correct cellular context.

Vrg4 displays a strong preference for purine nucleotides (Figure 1d). However, the binding site can clearly discriminate between GMP and AMP (Figure 2a). The IC_{50} value for GDP-Mannose was the same as GMP (Extended Data Figure 3). This suggests that sugar binding negates the negative effect of two free phosphates. To understand the structural basis of ligand recognition, we determined the structure of Vrg4 in complex with GDP-Mannose to 3.6 Å (Extended Data Table 1). The nucleotide sugar binds in an extended conformation within the polar cavity observed previously (Figure 2b & Extended Data Figure 4). Both nucleotide and sugar are accommodated within separate binding pockets at either end of the transporter, with the sugar pocket being considerably more polar compared to the nucleotide pocket. The presence of a separate sugar binding pocket prompted us to test the recognition of different guanine-based nucleotide sugars (Figure 2c). We observed a higher apparent affinity for GDP-Mannose compared with GDP-fucose or GDP-glucose, supporting specific recognition of the sugar moiety.

The nucleotide pocket is formed by side chains from TM7 and 8, interacting via the O6 carbonyl group with Asn221 (TM7) and Ser266 (TM8) and via the N2 amine group with Asn220 (TM7) (Figure 3a). These specific interactions with the O6 carbonyl of the guanine ring help to explain the lower IC_{50} values we observed for GMP compared to AMP, which does not contain an oxygen atom at this position. Analysis of sequence alignments between NSTs reveals these asparagine residues are located within a conserved F^{218}YNN motif, which is only present in transporters that recognise guanine containing nucleotide sugars (Extended Data Figure 5). Transport assays conducted on alanine variants of either Asn220 or Asn221 show an equal reduction in overall transport, whereas removal of both resulted in complete loss of function, highlighting the importance of this motif for nucleotide recognition (Figure 3b). In contrast, the ribose group makes interactions with side chains from both the N- and C-terminal halves of the transporter, with the O2 oxygen sitting within hydrogen bond distance to Tyr28 (TM1), Ser269 (TM8) and Tyr281 (TM9). Whereas Tyr28 is strictly conserved and its replacement with alanine abolished transport, Tyr281 is only semi conserved, but is generally a bulky side chain (Extended Data Figure 5). Although transport was abolished in an alanine variant, Vrg4 was still functional with a phenylalanine substitution at this position, suggesting that a bulky side chain is required to optimally position either the ribose group and/or the nearby conserved Tyr28 (Figure 3b).

The first phosphate sits close to a conserved methionine, Met35 (TM1), whereas the second phosphate can be seen interacting with a conserved lysine, Lys289 (TM9). Lys289 forms part of the previously identified G^{285}ALNK motif, which was predicted to interact with the substrate 26. Our structure reveals that this residue interacts with both the beta-phosphate and the glycosidic bond oxygen. Both Lys289Ala and Gly285Ala reduced activity, although the replacement of the lysine was more severe (Figure 3b). Disruption of either site specifically reduces the affinity for GDP-Mannose whilst the affinity for GMP is unchanged, supporting the involvement of this motif in sugar recognition (Figure 3c). The other side of the pocket is constructed from TM4, which contributes a conserved tyrosine, Tyr114, which makes a hydrogen bond to the C2 hydroxyl on the mannose ring. Although a tyrosine at this position is present in the majority of GDP sugar transporters, it is replaced by a glutamine in the Leishmania proteins (Extended Data Figure 5). A Tyr114Phe variant displayed reduced transport highlighting the importance of a hydrogen bond donor at this position (Extended Data Figure 6).

Using the crystal structure of Vrg4 we made homology models of the human GDP-fucose (SLC35C1) and CMP-sialic acid (SLC35A1) transporters (Extended Data Figure 7). The homology model of the human GDP-fucose transporter shows the position of the conserved FYNN motif on TM7 that would coordinate the guanine ring. The main difference, as expected, is within the sugar pocket, where we see the GALNK motif being replaced by GTAKA, which is conserved amongst GDP-fucose transporters (Extended Data Figure 5). Mutations that result in Leukocyte adhesion deficiency II (LAD II) are Arg147Cys and Thr308Arg 4,8,27. Thr308 is located on the same helix as the GTAKA motif and packs against TM10, whereas Arg147 is equivalent to Lys118 in Vrg4, and sits close to the sugar binding site. It seems likely that these disease mutations result in disruption of the sugar binding site in the GDP-fucose transporter highlighting the importance for recognition of the sugar moiety.

In order to shuttle ligands across the membrane, solute carriers use a general alternating access mechanism 28. Many transporters contain symmetrically arranged structural repeats that are inverted relative to one another in the plane of the membrane 29 30. Specific interactions between these repeats and the ligand(s) are formed and broken allowing their rearrangement and facilitating transport across the membrane. Vrg4 contains two inverted repeats, TM1-5 and TM6-10 (Extended Data Figure 8). In the current structure, interactions are observed between these repeats in the vicinity of the bound ligand, specifically Ser32 (TM1) and Lys289 (TM9) and Asn191 (TM6) and Lys118 (TM4), all of these residues are conserved (Figure 4a & Extended Data Figure 5). We generated a cytoplasmic facing model to investigate the structural changes that facilitate transport. The axis on which the transporter alternates between Golgi and cytoplasmic facing states runs through the observed interaction network and results in TM1 and 3 packing together and TM6 and 8 moving apart opening the binding site to the cytoplasm (Figure 4a). Supporting the involvement of the interaction network in transport we found that alanine variants of these residues severely reduced activity (Figure 4b), whereas a conservative substitution of Lys118 to arginine retained ~ 50 % activity, indicating the importance of a positive charge on TM4. A key feature of any transporter is the coupling between substrate binding and conformational rearrangement. In Vrg4 the bound GDP-mannose spans the binding site interacting with

TMs from both repeats, facilitating substrate induced reorientation. If we model GMP into the nucleotide binding pocket on TM7, the phosphate sits close to Lys118 (TM4), potentially explaining the necessity of a positive charge at this position. The ribose would make similar interactions with the tyrosine on TM1 and serine on TM8, thereby coordinating movements between the two-inverted topology repeats.

Taken together these insights suggest a model for transport within NSTs. In the cytoplasmic facing state binding of nucleotide-sugar via the two specificity pockets triggers the reorganisation of the inverted topology repeats resulting in formation of the Golgi open state (Figure 4c). Disassociation of the nucleotide sugar would vacate the binding site for nucleotide monophosphate to bind, triggering reorientation back to the cytoplasmic state (Figure 4d). This mechanism contains an inherent asymmetry, with the nucleotide sugar making more interactions with the transporter than the nucleotide monophosphate. This may explain the rate difference we observed in Vrg4 for GDP-mannose vs GMP (Figure 1e). In the cellular context, where both GDP-mannose and GMP are present in the cytoplasm, this asymmetry may facilitate GDP-mannose import in preference over GMP, as GDP-mannose is more effective in reorientating the transporter, resulting in preferential nucleotide sugar import into the Golgi.

Methods

Cloning, Expression and Purification of Vrg4

The gene encoding *VRG4* was amplified from *Saccharomyces cerevisiae* genomic DNA and cloned into a modified form of a C-terminal GFP^{His} fusion yeast expression vector 31. Wild type and mutant forms of Vrg4 were expressed by growing an overnight culture in minus leucine media with 2 % glucose. This culture was diluted 10-fold in minus leucine with 2 % lactate in a 15-litre fermentation vessel, after 24 hours, expression was induced thorough the addition of 1.5 % galactose from a 25 % w/v stock. The yeast was harvested after a further 24 hours and membranes were prepared. Vrg4 and mutants were purified to homogeneity (Extended Data Figure 9), using standard IMAC protocols in n-dodecyl-β-D-maltopyranoside (DDM) detergent (Glycon) as described previously 32. The protein was concentrated to 40 $mg.ml^{-1}$ for crystallization. For functional characterisation, the detergent was exchanged through size exclusion chromatography using an Superdex 200 (10/300) SEC column to 0.3 % n-decyl-β-D-maltopyranoside (DM).

Crystallization and Structure Determination

The protein-laden mesophase was prepared by homogenizing monoolein (Sigma) and 40 $mg.ml^{-1}$ protein solution in a 60:40 ratio by weight using a dual syringe mixing device at 20°C 33. Crystallisation was carried out at 4 °C in 96-well glass sandwich plates with 50 nL mesophase and 0.8 μL precipitant solution. The crystallisation solution consisted of 26 - 30 % (v/v) PEG 400, 0.1 M sodium citrate pH 5.0 and 75 mM sodium chloride or sodium acetate. Crystals grew within four days and reached a maximum size of 30 μM x 5μM. Wells were opened using a tungsten-carbide glasscutter and the crystals were harvested using 30-50 μm micromounts (MiTeGen). Crystals were cryo-cooled directly in liquid nitrogen. Heavy atom screening was carried out by soaking the crystals *in situ* with 1-5 mM heavy

atom salts following manual removal of the glass cover plate followed by resealing using a glass cover slip. Successful Hg derivatives were obtained using 5 mM Methyl Mercury Chloride and soaking for 24 hours. > 200 complete datasets were collected before phases were successfully determined using long wavelength anomalous dispersion. Diffraction data for native crystals were collected at micro focus beamlines I24 (Diamond Light Source, UK), Proxima 2-A (Soleil, France) and BL1-A (Photon Factory, Japan). Data were processed and scaled using the Xia2 34 pipeline to XDS 35 and AIMLESS 36. The space group was initially determined to be $P2_12_12$ with 57 % non-protein content. Crystals routinely diffracted between 2.8 and 3.5 Å maximum resolution.

Following an extensive screen of heavy atoms at I24 phases were finally calculated using data collected at BL1-A, using long wavelength anomalous dispersion at 2.7 Å. The anomalous component of the atomic scattering factor for Hg is larger at longer wavelengths, resulting in a larger value of f" and therefore a stronger anomalous signal. The successful dataset used for phasing (Hg/S-SAD; Extended Data Table 1) was the result of merging and scaling data from seven different crystals, each containing 360 º of data collected on a single axis goniometer and measured using an Eiger detector (DECTRIS). Hg and sulphur sites were found using HySS 37 and phases calculated using the 'extreme density modification' function, as implemented in RESOLVE. The resulting maps were of excellent quality allowing all 10 helices from two molecules in the asymmetric unit to be clearly observed. A second higher resolution dataset processed to 3.2 Å resolution was used to verify the correct register of the structure (S-SAD). The Hg soaked crystals were not isomorphous with the native crystals, preventing the use of the higher resolution data during the initial phase calculations. In the crystal structure, we observed four of the eight molecules formed dimers (Extended Data Figure 2), as determined by the PDBePISA server (http://www.ebi.ac.uk/pdbe/pisa/).

GDP-mannose was soaked into the crystals using a similar procedure to the heavy atom soaking. In short, the top glass plate was removed using a glass scribe followed by application of 2 μL of crystallisation solution containing 20 mM GDP-mannose. The well was resealed using a thin glass cover slip and left overnight. Crystals were harvested using 30 μm micromounts prior to flash cooling in liquid nitrogen.

Model building and refinement

An initial Cα model was built into the density using O and Coot 38. Refinement of the model was initially carried out in Phenix in space group $P2_12_12$. However, failure to observe a drop in the crystallographic R factors from ~ 40 % prompted us to consider twinning within the crystals. Subsequent reprocessing of the data in P1, followed by molecular replacement to locate the eight molecules in the unit cell confirmed the crystals were pseudo merohedrally twinned, with an apparent higher symmetry of $P2_12_12$. Fortunately, the twinning in this case actually facilitated the location the heavy atoms in the higher symmetry space group. The crystal structure was only fractionally divergent from the twin operator; this was large enough to make refinement of the crystal structure stall, but small enough to be accommodated within the heavy atom refinement within HySS. The high- resolution S-SAD data set (Extended Data Table 1) was reprocessed in P1, resulting in a slightly lower

resolution of 3.22 Å (Apo, Extended Data Table 1). Subsequent refinement of the eight molecules was continued in P1 with non-crystallographic symmetry restraints applied in Phenix and manual building in Coot. A final round of refinement was carried out using Buster 39 resulting in the R factors reported in Extended Data Table 1.

Transport assays

Vrg4 was reconstituted into liposomes using the dilution method 20. Chloroform was removed from the yeast polar lipids (Avanti polar lipids) through the use of a rotary evaporator to obtain a thin film. The lipids were washed twice in pentane and then resuspended at 5 mg/ml in lipid buffer (20 mM Hepes pH 7.5, 100 mM KCl). These lipid vesicles were frozen and thawed twice in liquid nitrogen and stored at -80 °C until required. For reconstitution, the lipids were thawed and then extruded first through a 0.8 μm filter and then a 0.4 μm filter. Purified Vrg4 in DM (at 0.5 μg/μl concentration) was added to the lipids at a final lipid:protein ratio of 80:1 and incubated for 1 hour at room temperature, then a further hour on ice, or for the no protein liposome control the same volume of gel filtration buffer containing 0.3 % DM was added. After this time, the protein lipid mix was diluted rapidly into 65 ml of assay buffer (20 mM Hepes pH 7.5, 50 mM KCl, 2 mM $MgSO_4$) and proteoliposomes were harvested through centrifugation >200,000g for 2 hours. To remove trace detergent the proteoliposomes were dialysed overnight against a large volume of assay buffer. After dialysis, the proteoliposomes were harvested and resuspended in assay buffer to a final protein concentration of 0.25 μg/μl and subjected to three rounds of freeze thawing in liquid nitrogen before storage at -80 °C. The amount of protein (both wild type and mutant variants) reconstituted into the lipids was quantified by SDS-PAGE and densitometry (Extended Data Figure 9).

To analyse transport activity proteoliposomes were thawed and the desired concentration of internal substrate was added. To load the liposomes, they were subjected to six rounds of freeze thaw in liquid nitrogen and then extruded through a 0.4 μm membrane. The equivalent of 4 μg of protein was added to 50 μl of assay buffer containing 0.5 μM ^{3H}GMP which initiated the transport. The uptake of radiolabelled substrate was stopped at the desired time by rapidly filtering onto 0.22 μm filters which were then washed with 2 x 2ml cold water. For the time course experiments shown in figure 1e, the equivalent of 0.5 μg of protein was used per time point and a total of 20 μM GMP as substrate on the outside of the vesicles was used.

The amount of GMP transported inside the liposomes was calculated by scintillation counting in Ultima Gold™ (Perkin Elmer). Experiments were performed a minimal of four times to generate an overall mean and S.D.

Lipid specificity

Wild type Vrg4 was reconstituted into a range of different lipid vesicles. The lipids used were either yeast polar lipids, *E. coli* polar lipids, POPC (Avanti Polar Lipids), or a 3:1 ratio of POPE:POPG, POPC:POPE, or soybean PE:eggPG or a 3:1 ratio of POPE:POPG with 12 % egosterol, POPC, eggPC, YPL or DMPC added. All lipid vesicles were prepared as described in the transport assay section and wild type Vrg4 was reconstituted at an 80:1 total

lipid to protein ratio as described. Transport was assessed using 4 μg of protein per time point taken and 20 μM GMP as substrate on the outside and 1 mM GDP-mannose on the inside of the vesicles, or water as a negative control. Transport was monitored over 10 minutes, and figure 1f shows the relative uptake in comparison to yeast polar lipids at 5 minutes for each lipid type used. Experiments were performed a minimum of four times from two separate reconstitutions to generate an overall mean and S.D.

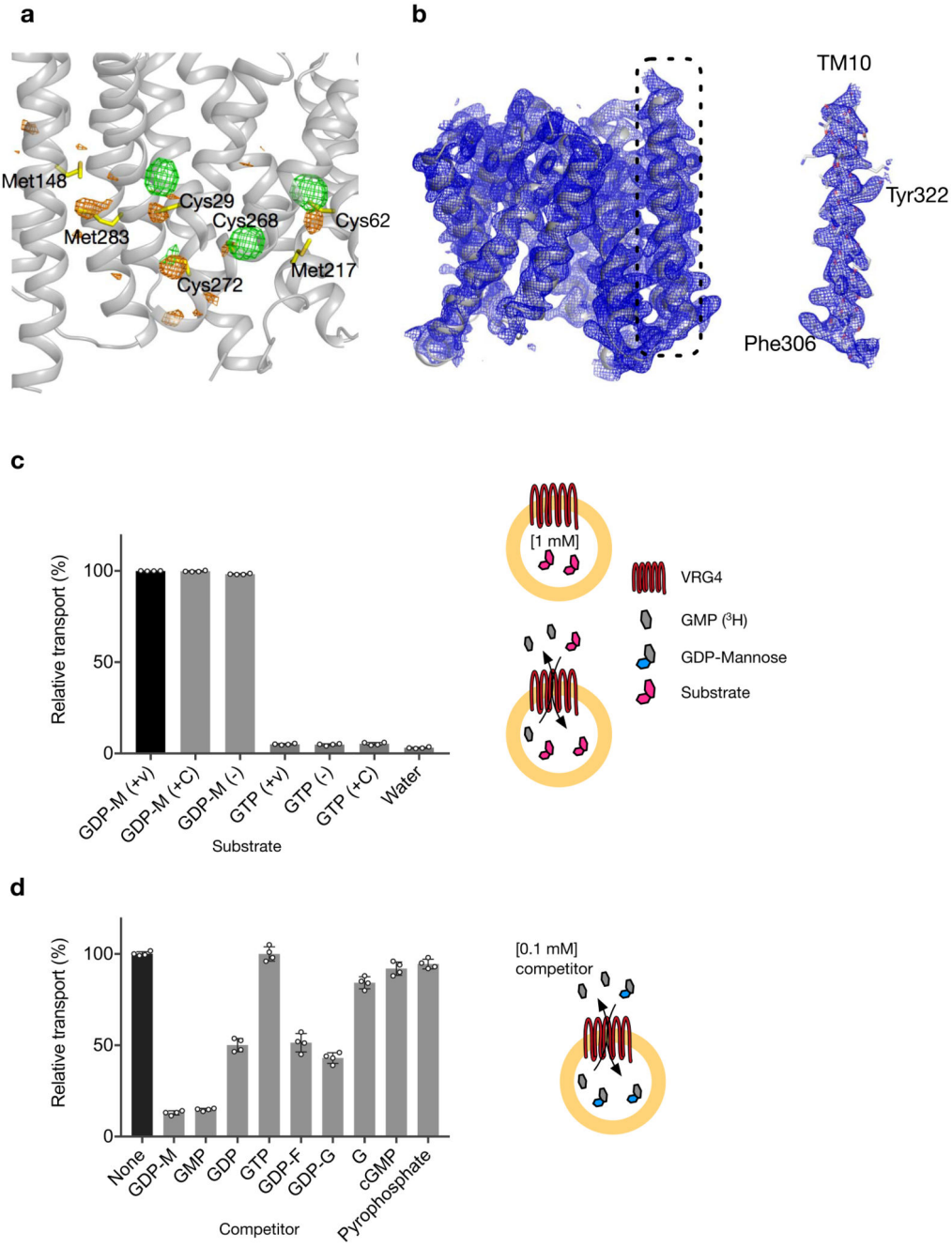
Data Availability

Atomic coordinates for the atomic models have been deposited in the Protein Data Bank under accession numbers 5OGE (Apo structure of Vrg4) and 5OGK (GDP-Mannose bound Vrg4). The expression plasmid has been deposited in Addgene, with ID 102334.

Europe PMC Funders Author Manuscripts

Europe PMC Funders Author Manuscripts

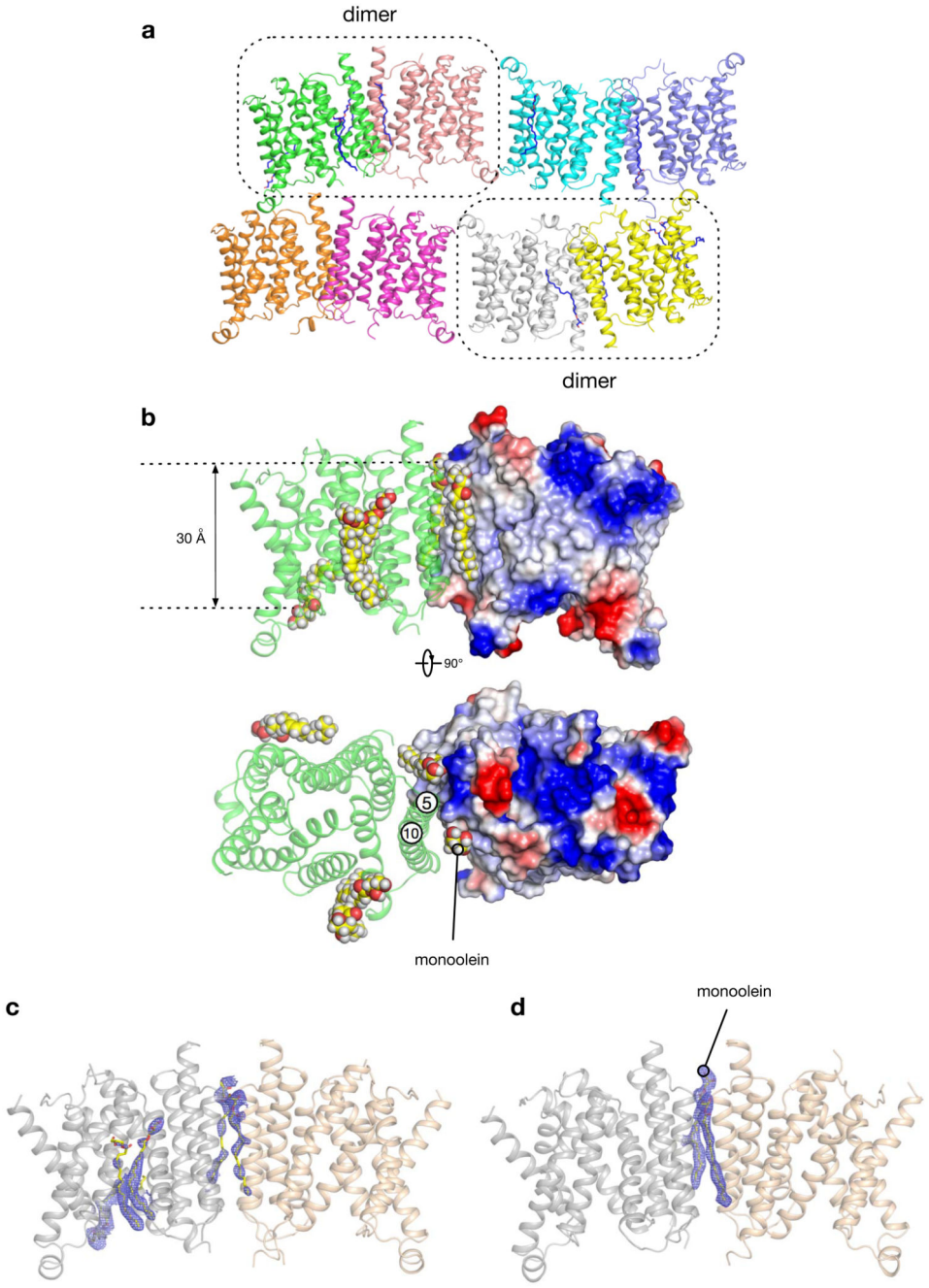
Extended Data



Extended Data Figure 1. Top panel. Analysis of anomalous electron density and final refined 2Fo-Fc map.

a, Anomalous difference peaks for the Hg (orange mesh) and sulphur (green mesh) atoms calculated using AnoDe and contoured at 3σ, and overlaid on the structure of Vrg4. **b**, Final refined 2mFo-DFc maps for Vrg4, contoured at 1σ. TM10 is highlighted along with side chain density. **Bottom panel. *In vitro* substrate specificity and recognition by Vrg4. c,** Transport cannot be driven by GTP in the presence or absence of CCCP (C) or valinomycin

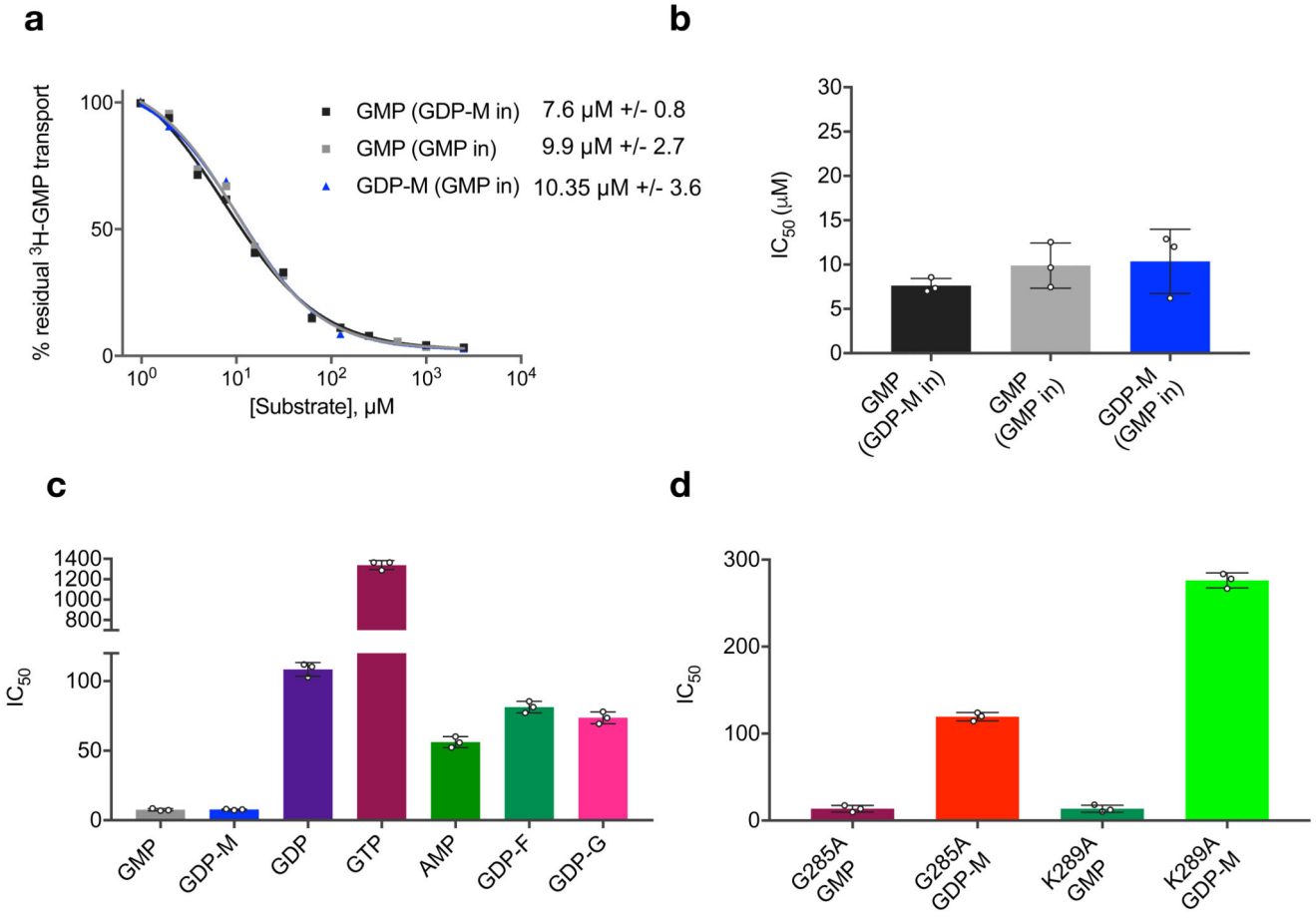
(v). n=4 independent experiments, measure of centre mean, error bars s.d. **d,** Competition assays demonstrate that GMP and GDP-mannose act as the most effective competitors of transport but other guanosine containing substrates, GDP, GDP-fucose, GDP-glucose can also compete and be recognised by the binding site. However, the free base guanosine is very poorly recognised as is the cyclic GMP. Pyrophosphate, which consists of a di-phosphate moiety and is an inhibitor of ATPases, cannot inhibit the transport activity of Vrg4. n=4 independent experiments, measure of centre mean, error bars s.d.



Europe PMC Funders Author Manuscripts

Extended Data figure 2. Lipid molecules are found at the dimer interface of Vrg4.

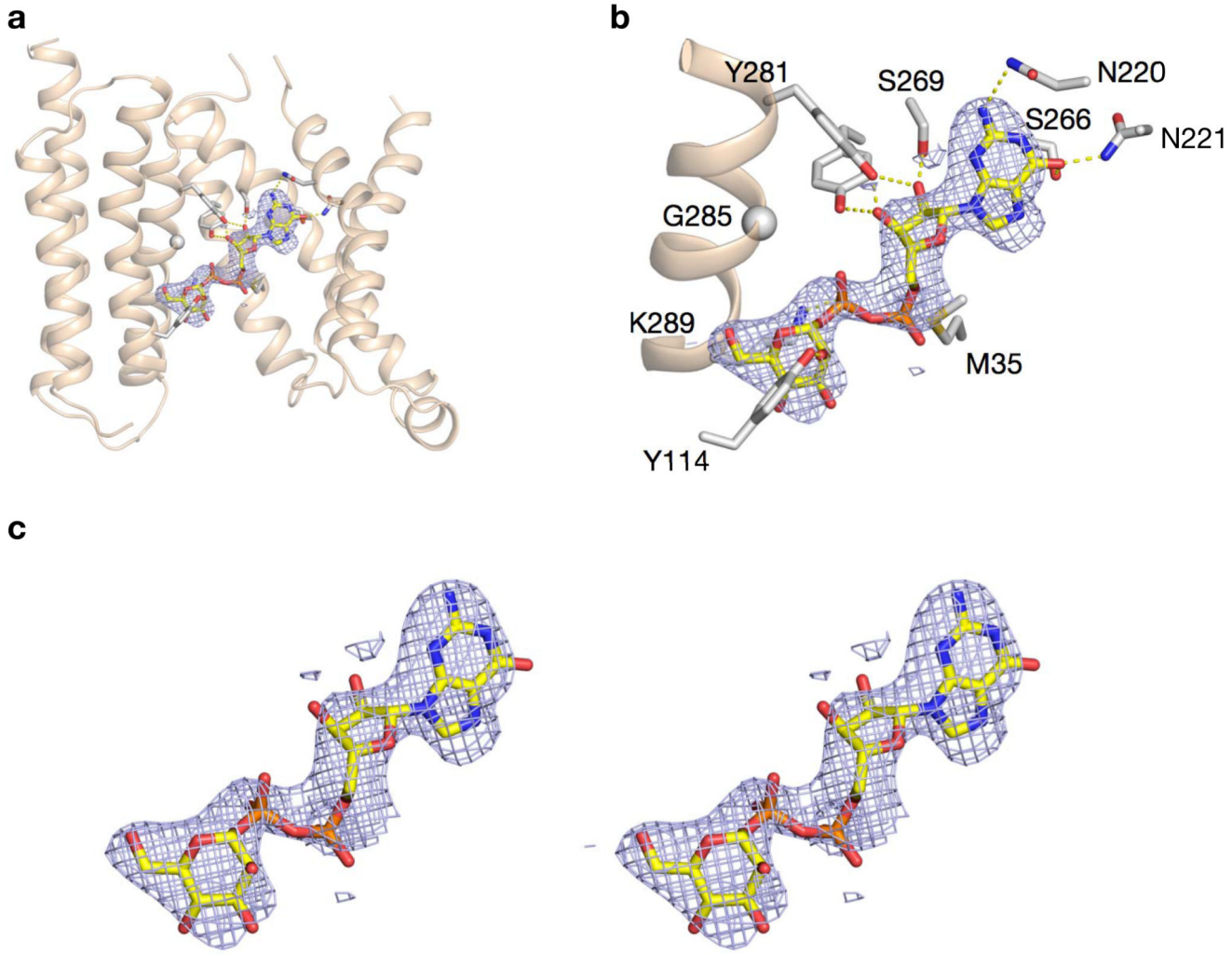
a, The eight molecules in the P1 unit cell are shown, four of which adopt a dimer arrangement with the monomers arranged in a physiological orientation. Several lipid molecules could be identified and are shown in blue. **b,** Zoomed in view of the Vrg4 dimer showing the electrostatic surface of the dimer interface between helices 5 and 10 and the two monoolein lipids bound. The hydrophobic belt is noticeably shorter (~30 Å) than observed for plasma or inner membrane transporters (~40 Å). Within the dimer interface the two well-ordered monoolein molecules contribute ~60 % to the total buried surface area of 1514 $Å^2$. In the remaining four monomeric molecules in the unit cell we do not observe any density for lipid molecules at these sites, suggesting an important role for lipid mediated dimerization. **c,** Refined 2mFo-DFc electron density for the lipid molecules shown in b, present in the apo crystal structure. **d,** The dimer interface lipids were more ordered in the GDP-mannose soaked crystals. All electron density maps are contoured at 1σ.



Extended Data Figure 3.

a, Vrg4 recognises both GMP and GDP-M with the same apparent affinity. (GDP-M in) refers to GDP-mannose loaded liposomes and (GMP in) are GMP loaded liposomes. Representative IC_{50} curves calculated for wild type Vrg4. Calculated IC_{50} values are the mean of three independent experiments (b), errors shown are s.d. **c,** Calculated IC50 values

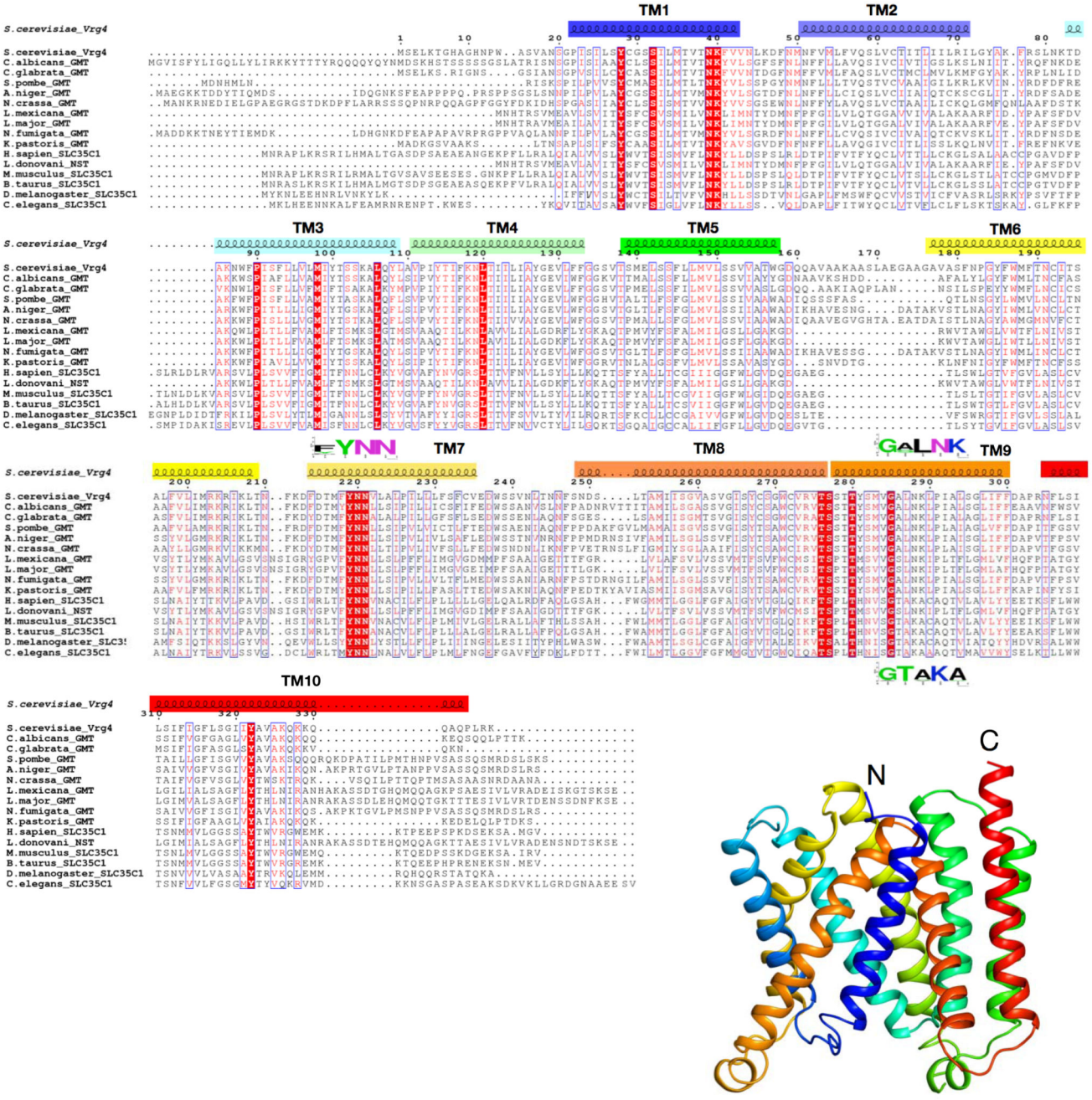
for Figure 2a,c. n=3 independent experiments, measure of centre mean, error bars s.d. **d,** Calculated IC50 values for Figure 3c. n=3 independent experiments, measure of centre mean, error bars s.d.



Extended Data Figure 4. Electron density maps for bound GDP-mannose.

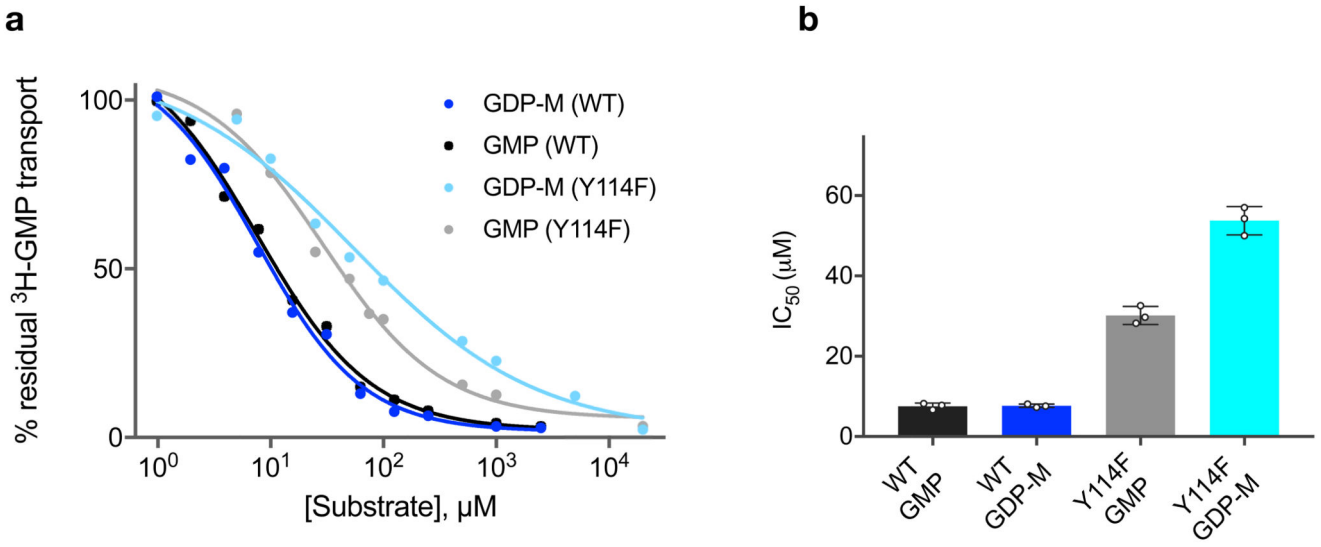
a, Crystal structure of Vrg4 bound to GDP-mannose in the plane of the membrane. GDP-mannose is shown as sticks with refined 2mFo-DFc electron density contoured at 1σ. **b,** Zoomed in view of the binding site showing key polar interactions. **c,** stereo view of bound GDP-mannose shown in b.

Europe PMC Funders Author Manuscripts



Extended Data Figure 5. Alignment of Vrg4 with GDP-mannose and GDP-fucose transporters.

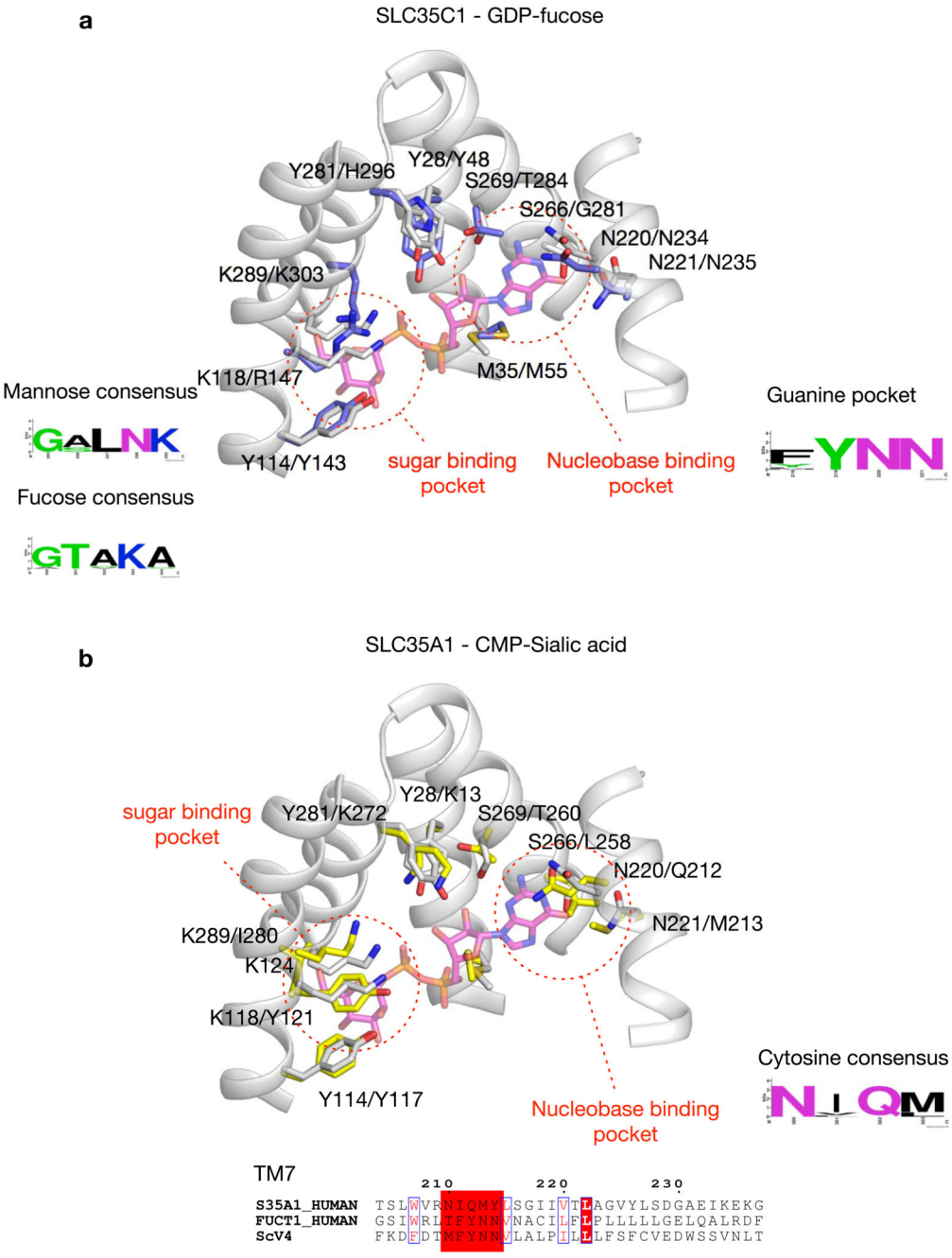
Amino acid alignment of *S. cerevisiae* Vrg4 with GDP-mannose transporters from various organisms (_GMT or_NST*) or GDP-fucose transporters (*SLC35C1). Identical residues are highlighted in red. The nucleotide binding, FYNN, and sugar recognition motifs are indicated as sequence logos. The positions of the 10 TMs are indicated and coloured from blue to red as in the crystal structure representation of Vrg4 shown on the right.



Extended Data Figure 6. Effect of Tyr114Phe mutation on substrate affinity.

a, Representative IC_{50} curves calculated for wild type Vrg4 (WT) and Tyr114Phe (Y114F) for GMP and GDP-mannose (GDP-M). **b,** Bar chart showing the IC_{50} values calculated from a. n=3 independent experiments, measure of centre mean, error bars s.d. Wild type Vrg4 recognises both GMP and GDP-M with the same affinity. Mutating Tyr114 to phenylalanine reduced the affinity for both substrates, demonstrating the importance of the hydroxyl group on this residue for high affinity binding. Unlike WT, Y114F now recognises GDP-M with a reduced affinity compared to GMP supporting the interaction we observed in the crystal structure (Figure 3a).

Europe PMC Funders Author Manuscripts

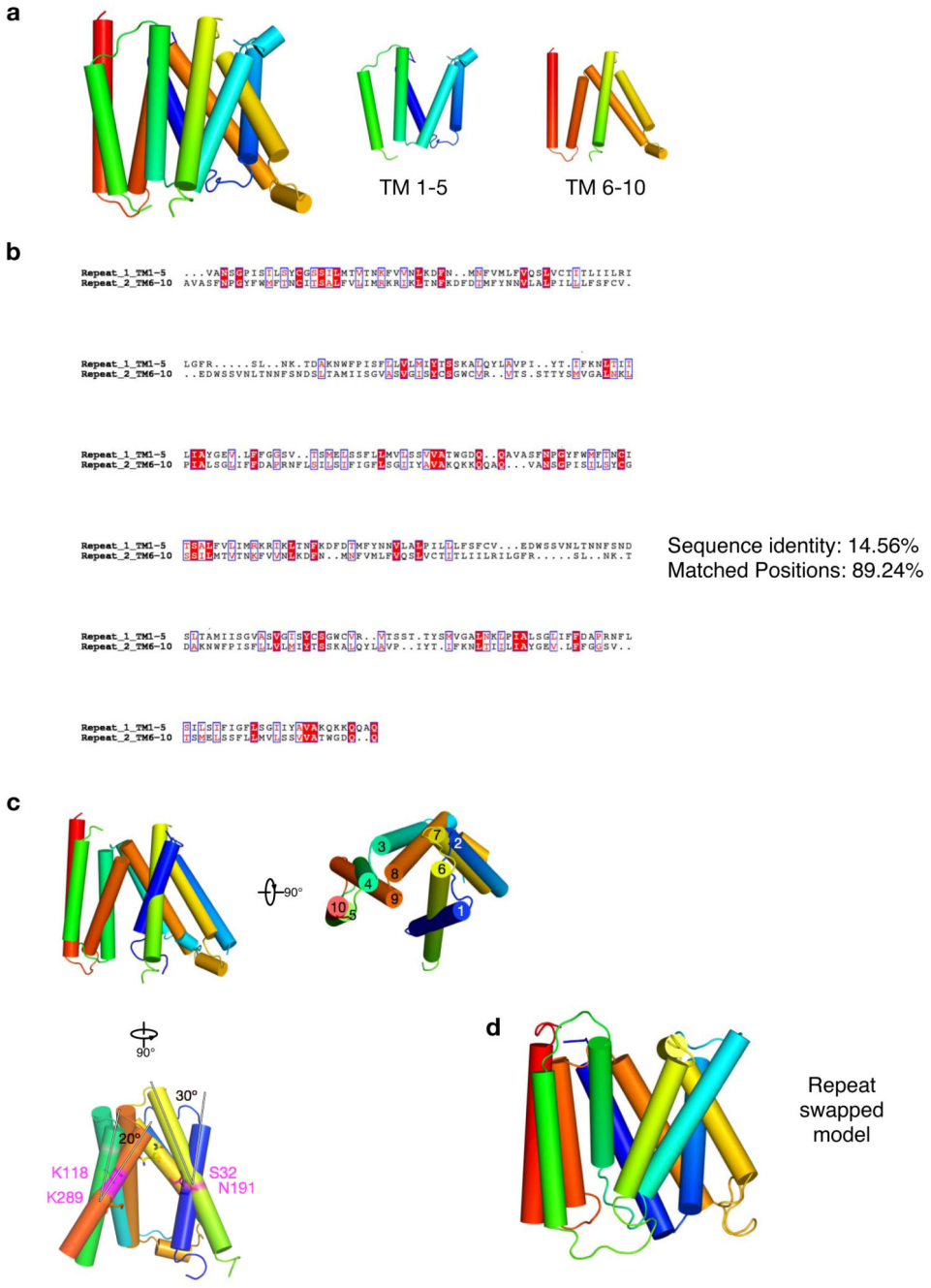


Extended Data Figure 7. Homology models of the human GDP-fucose transporter (SLC35C1) and CMP-sialic acid transporter (SLC35A1) conserved sequence features across the SLC35 family.

a, View of the binding of Vrg4 with bound GDP-mannose (grey) overlaid with the equivalent residues from the GDP-fucose homology model (blue). Residue numbers are shown first for Vrg4, then for the equivalent residue in the human transporter. The consensus motifs for mannose (GALNK) and fucose (GTAKA) binding are shown, as is the FYNN motif that dictates guanine binding. **b,** The equivalent analysis was carried out for the human CMP-sialic acid transporter (SLC35A1) which is shown in yellow. Our homology model

reveals important insights into the human CMP-sialic acid transporter (CST), which is mutated in the Congenital Disorder of Glycosylation IIf (CDG-IIf) and results in a lack of sialic acid on the surface of cells 5. The model of SLC35A1 reveals the nucleotide binding pocket has changed from the FYNN motif to the NIQM motif, and now contains a glutamine (Gln212) instead of the asparagine. The extension of the glutamine side chain over the asparagine may help to facilitate recognition of cytosine, which is smaller than guanine. Previous studies have identified Tyr214 as being important for ligand recognition in the CST 40. In our model this residue faces the membrane and likely anchors Glu212 to stabilise its interaction with the cytosine. Another significant difference between the CST and Vrg4 is the replacement of Tyr28 and Tyr281 with lysine side chains (Lys13 and Lys272). Indeed, the mutation of Lys272 results in complete inactivation of CST 41. The GALNK motif, in contrast, is replaced by a stretch of alanine side chains. These changes suggest the binding site is both larger and more positively charged than Vrg4, which would be required to accommodate the nine-carbon sialic acid sugar group, which unlike fucose or mannose contains an N-acetyl extension at the C-5 position in addition to an extended hydroxyl arm at C-6.

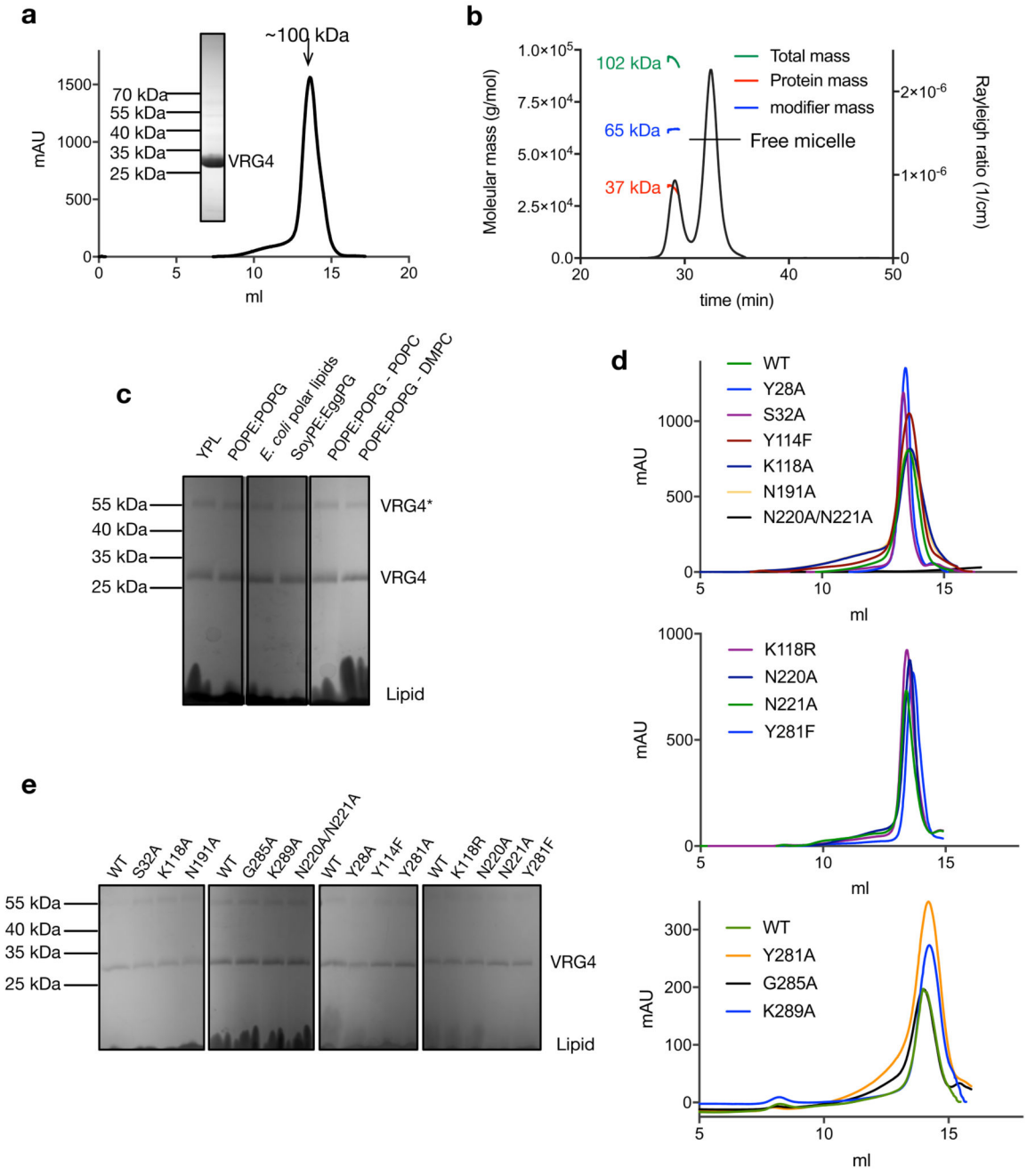
Europe PMC Funders Author Manuscripts



Extended Data Figure 8. Repeat swap modelling of the open to cytoplasm state of Vrg4.

a, Vrg4 contains two structural repeat units consisting of TM1-5 and TM6-10 which are related to each other via a two-fold rotation in the plane of the membrane. **b,** AlignMe42 sequence alignment between the two structural repeats reveals a high degree of sequence conservation making repeat swap modelling more accurate. Highlighted in magenta are the equivalent residues in each repeat that are seen making polar interactions in the crystal structure. **c,** The two repeats align with r.m.s.d. of 5 Å with the most notable deviation being between helices 4 and 9 and 1 and 6 which deviate from each other by 20 and 30 °

respectably. It is interesting to note that the axis on which these repeats deviate is where the residues highlighted in magenta overlap. **d,** the repeat swapped model of Vrg4 in the open to cytoplasm state which is achieved via a rotation of the structural repeats around an axis that runs through the residues highlighted in magenta.



Extended Data Figure 9. Analysis of purified Vrg4 in both detergent and lipid environments.

a, Representative gel filtration trace of the purified Vrg4 in the detergent decylmaltoside (DM) which was used for reconstitution, n=20. The protein elutes as a single peak and is

Europe PMC Funders Author Manuscripts

purified to homogeneity as judged by Coomassie stained SDS-PAGE where it runs as a single band at ~30 kDa. **b,** SecMALS analysis indicates that Vrg4 is monomeric in DM with an expected mass of 37 kDa (the dn/dc values for protein and DM detergent used were 0.185 and 0.147 respectively). n=3 independent experiments, producing similar results. **c,** Representative SDS-PAGE gel showing reconstitution of Vrg4 into different lipid environments. As shown the same amount of protein was used for each lipid type tested (YPL, yeast polar lipids, POPE:POPG in a 3:1 ratio, SoyPE:EggPG in a 3:1 ratio. POPE:POPG – POPC 3:1 ratio of POPE:POPG with 12 % POPC, POPE:POPG – DMPC, 3:1 ratio of POPE:POPG with 12 % DMPC added). n=5 independent experiments, producing similar results. * shows the presence of a protein band (which corresponds to a dimer in size) observed upon reconstitution into liposomes which is not seen when the protein is in detergent (a - inset) but is observed in the unit cell of the crystal structure. **d,** Representative gel filtration profiles of all the mutants used in this study, n=2 independent experiments, producing similar results. The traces show that all mutant variants behave similarly to wild type protein during purification, showing monodisperse peaks with little aggregation. With relation to the differing amounts of protein shown in these graphs this is due to the different volumes used to express the protein (either >10 l for upper two panels or 5 l for the lower panel). **e,** Representative SDS-PAGE gel showing reconstitution of Vrg4 and the mutant variants used in this study showing the same amount of protein was used in the assays. n=3 independent experiments, producing similar results.

Extended Data Table 1
Data collection and refinement statistics for Vrg4

	Apo PDB: 50GE	**GDP-Mannose PDB: 50GK**	**Apo Hg/S-SAD**	**Apo S-SAD**
Data collection				
Space group	PI	PI	$P2_i2_j2$	$P2_i2_j2$
Cell dimensions				
a, *b*, *c* (Å)	45.5, 102.9, 181.4	45.32, 101.69, 180.19	102.3, 178.8,45.8	102.9, 181.6, 45.6
α, *β*, *γ* (°)	89.92, 90.12, 90.06	89.96, 90.07, 90.11	90.0, 90.0, 90.0	90.0, 90.0, 90.0
Wavelength (Å)	2.70	0.9800	2.70	2.70
Resolution (Å)[a]	45.3-3.22	44.0-3.60	49.0-4.8	68.1-3.20
	(3.30-3.22)	(3.76-3.60)	(5.37-4.80)	(3.36-3.20)
R_{pim}	7.0 (57.5)	20.0(111)	6.4(17.5)	4.5 (37)
$I/\sigma I$	8.9(1.1)	3.8 (0.6)	18.3 (12.4)	14.4(1.5)
Completeness (%)	90.3(60.1)	96.8 (97.2)	99.9(100)	99.2 (88.6)
Anom. Completeness (%)			100(100)	100 (100)
Multiplicity	8.4 (2.9)	2.0(1.8)	86.2 (82.2)	22.2(12.5)
Anom. Multiplicity			48.5 (44.5)	
CC1/2 (%)	99.4 (56.4)	96.8 (30.3)	99.8 (97.8)	99.8 (75.3)
Anom. CC1/2 (%)			79.9 (40.2)	
Mid-Slope			2.468	1.041
Phasing Statistics				
Figure of Merit			51.6	

	Apo PDB: 50GE	GDP-Mannose PDB: 50GK	Apo Hg/S-SAD	Apo S-SAD
R_{dm}			37.99	
Refinement				
Resolution (Å)	45.35-3.22	45.02-3.60		
No. reflections	48149	36015		
R_{work}/ R_{free}	27.9/30.6	25.5/31.4		
No. atoms				
Protein	18836	18708		
Ramachandran favored	93.52	9.5		
Ramachandran outliers	0.68	0.09		
R.m.s deviations				
Bond lengths (Å)	0.009	0.004		
Bond angles (°)	1.02	0.714		

[a]Highest resolution shell is shown in parenthesis.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

We thank the staff of beamlines I24 Diamond Light Source, UK, Proxima-2a, Soleil, France and BL1-A, Photon Factory, Japan for assistance. We thank Professor O. Nureki and Associate Professor R. Ishitani, University of Tokyo, Japan, for assistance in accessing BL1-A and Dr G. Kuteyi, University of Oxford for technical support. This work was supported by Wellcome awards (102890/Z/13/Z & 102890/Z/13/A).

References

1. Dalziel M, Crispin M, Scanlan CN, Zitzmann N, Dwek RA. Emerging principles for the therapeutic exploitation of glycosylation. Science. 2014; 343:1235681.doi: 10.1126/science.1235681 [PubMed: 24385630]
2. Carey D, Sommers L, Hirschberg C. CMP-N-acetylneuraminic acid: isolation from and penetration into mouse liver microsomes. Cell. 1980; 19:597–605. [PubMed: 7363326]
3. Hirschberg CB, Robbins PW, Abeijon C. Transporters of nucleotide sugars, ATP, and nucleotide sulfate in the endoplasmic reticulum and Golgi apparatus. Annu Rev Biochem. 1998; 67:49–69. DOI: 10.1146/annurev.biochem.67.1.49 [PubMed: 9759482]
4. Luhn K, Wild MK, Eckhardt M, Gerardy-Schahn R, Vestweber D. The gene defective in leukocyte adhesion deficiency II encodes a putative GDP-fucose transporter. Nat Genet. 2001; 28:69–72. DOI: 10.1038/88289 [PubMed: 11326279]
5. Martinez-Duncker I, et al. Genetic complementation reveals a novel human congenital disorder of glycosylation of type II, due to inactivation of the Golgi CMP-sialic acid transporter. Blood. 2005; 105:2671–2676. DOI: 10.1182/blood-2004-09-3509 [PubMed: 15576474]
6. Descoteaux A, Luo Y, Turco SJ, Beverley SM. A specialized pathway affecting virulence glycoconjugates of Leishmania. Science. 1995; 269:1869–1872. [PubMed: 7569927]
7. Nishikawa A, Poster JB, Jigami Y, Dean N. Molecular and phenotypic analysis of CaVRG4, encoding an essential Golgi apparatus GDP-mannose transporter. J Bacteriol. 2002; 184:29–42. [PubMed: 11741841]

8. Lubke T, et al. Complementation cloning identifies CDG-IIc, a new type of congenital disorders of glycosylation, as a GDP-fucose transporter deficiency. Nat Genet. 2001; 28:73–76. DOI: 10.1038/88299 [PubMed: 11326280]
9. Dunphy WG, Fries E, Urbani LJ, Rothman JE. Early and late functions associated with the Golgi apparatus reside in distinct compartments. Proc Natl Acad Sci U S A. 1981; 78:7453–7457. [PubMed: 6801652]
10. Ishida N, Kawakita M. Molecular physiology and pathology of the nucleotide sugar transporter family (SLC35). Pflgers Archiv European Journal of Physiology. 2004; 447:768–775. DOI: 10.1007/s00424-003-1093-0 [PubMed: 12759756]
11. Dean N, Zhang Y, Poster J. The VRG4 gene is required for GDP-mannose transport into the lumen of the Golgi in the yeast, Saccharomyces cerevisiae. J Biol Chem. 1997; 272:31908–31914. [PubMed: 9395539]
12. Haanstra JR, Gonzalez-Marcano EB, Gualdron-Lopez M, Michels PA. Biogenesis, maintenance and dynamics of glycosomes in trypanosomatid parasites. Biochim Biophys Acta. 2016; 1863:1038–1048. DOI: 10.1016/j.bbamcr.2015.09.015 [PubMed: 26384872]
13. Liu L, Xu YX, Hirschberg CB. The role of nucleotide sugar transporters in development of eukaryotes. Semin Cell Dev Biol. 2010; 21:600–608. DOI: 10.1016/j.semcdb.2010.02.002 [PubMed: 20144721]
14. Jack DL, Yang NM, Saier MH Jr. The drug/metabolite transporter superfamily. Eur J Biochem. 2001; 268:3620–3639. [PubMed: 11432728]
15. Tsuchiya H, et al. Structural basis for amino acid export by DMT superfamily transporter YddG. Nature. 2016; 534:417–420. DOI: 10.1038/nature17991 [PubMed: 27281193]
16. Gao X, Dean N. Distinct protein domains of the yeast Golgi GDP-mannose transporter mediate oligomer assembly and export from the endoplasmic reticulum. J Biol Chem. 2000; 275:17718–17727. [PubMed: 10748175]
17. Abe M, Noda Y, Adachi H, Yoda K. Localization of GDP-mannose transporter in the Golgi requires retrieval to the endoplasmic reticulum depending on its cytoplasmic tail and coatomer. J Cell Sci. 2004; 117:5687–5696. DOI: 10.1242/jcs.01491 [PubMed: 15494368]
18. Capasso J, Hirschberg C. Mechanisms of glycosylation and sulfation in the Golgi apparatus: evidence for nucleotide sugar/nucleoside monophosphate and nucleotide sulfate/nucleoside monophosphate antiports in the Golgi apparatus membrane. Proc Natl Acad Sci U S A. 1984; 81:7051–7055. [PubMed: 6095266]
19. van Meer G. Lipids of the Golgi membrane. Trends Cell Biol. 1998; 8:29–33. [PubMed: 9695805]
20. Parker JL, Mindell JA, Newstead S. Thermodynamic evidence for a dual transport mechanism in a POT peptide transporter. Elife. 2014; 3doi: 10.7554/eLife.04273
21. Schmidt D, Jiang Q-X, Mackinnon R. Phospholipids and the origin of cationic gating charges in voltage sensors. Nature. 2006; 444:775–779. DOI: 10.1038/nature05416 [PubMed: 17136096]
22. van Meer G, Voelker DR, Feigenson GW. Membrane lipids: where they are and how they behave. Nature reviews Molecular cell biology. 2008; 9:112–124. DOI: 10.1038/nrm2330 [PubMed: 18216768]
23. Sharpe HJ, Stevens TJ, Munro S. A comprehensive comparison of transmembrane domains reveals organelle-specific properties. Cell. 2010; 142:158–169. DOI: 10.1016/j.cell.2010.05.037 [PubMed: 20603021]
24. Banfield DK. Mechanisms of protein retention in the Golgi. Cold Spring Harb Perspect Biol. 2011; 3:a005264.doi: 10.1101/cshperspect.a005264 [PubMed: 21525512]
25. Wong M, Munro S. Membrane trafficking. The specificity of vesicle traffic to the Golgi is encoded in the golgin coiled-coil proteins. Science. 2014; 346:1256898.doi: 10.1126/science.1256898 [PubMed: 25359980]
26. Gao X, Nishikawa A, Dean N. Identification of a conserved motif in the yeast golgi GDP-mannose transporter required for binding to nucleotide sugar. J Biol Chem. 2001; 276:4424–4432. [PubMed: 11067855]
27. Puglielli L, Hirschberg C. Reconstitution, identification, and purification of the rat liver golgi membrane GDP-fucose transporter. J Biol Chem. 1999; 274:35596–35600. [PubMed: 10585436]

28. Jardetzky O. Simple allosteric model for membrane pumps. Nature. 1966; 211:969–970. [PubMed: 5968307]
29. Drew D, Boudker O. Shared Molecular Mechanisms of Membrane Transporters. Annu Rev Biochem. 2016; 85:543–572. DOI: 10.1146/annurev-biochem-060815-014520 [PubMed: 27023848]
30. Forrest LR, Krämer R, Ziegler C. The structural basis of secondary active transport mechanisms. Biochimica et Biophysica Acta (BBA) - Bioenergetics. 2011; 1807:167–188. DOI: 10.1016/j.bbabio.2010.10.014 [PubMed: 21029721]

Additional References called in the Online Methods and Extended Data Figures

31. Parker JL, Newstead S. Method to increase the yield of eukaryotic membrane protein expression in Saccharomyces cerevisiae for structural and functional studies. Protein Sci. 2014; 23:1309–1314. DOI: 10.1002/pro.2507 [PubMed: 24947543]
32. Drewr D, et al. GFP-based optimization scheme for the overexpression and purification of eukaryotic membrane proteins in Saccharomyces cerevisiae. Nature Protocols. 2008; 3:784–798. DOI: 10.1038/nprot.2008.44 [PubMed: 18451787]
33. Caffrey M, Cherezov V. Crystallizing membrane proteins using lipidic mesophases. Nature Protocols. 2009; 4:706–731. DOI: 10.1038/nprot.2009.31 [PubMed: 19390528]
34. Winter G, Lobley CMC, Prince SM. Decision making in xia2. Acta Crystallographica Section D Biological Crystallography. 2013; 69:1260–1273. DOI: 10.1107/S0907444913015308 [PubMed: 23793152]
35. Kabsch W. XDS. Acta Crystallographica Section D Biological Crystallography. 2010; 66:125–132. DOI: 10.1107/S0907444909047337 [PubMed: 20124692]
36. Evans PR, Murshudov GN. How good are my data and what is the resolution? Acta Crystallographica Section D Biological Crystallography. 2013; 69:1204–1214. DOI: 10.1107/S0907444913000061 [PubMed: 23793146]
37. Adams PD, et al. PHENIX: a comprehensive Python-based system for macromolecular structure solution. Acta Crystallographica Section D Biological Crystallography. 2010; 66:213–221. DOI: 10.1107/S0907444909052925 [PubMed: 20124702]
38. Emsley P, Lohkamp B, Scott WG, Cowtan K. Features and development of Coot. Acta Crystallographica Section D Biological Crystallography. 2010; 66:486–501. DOI: 10.1107/S0907444910007493 [PubMed: 20383002]
39. Blanc E, et al. Refinement of severely incomplete structures with maximum likelihood in BUSTER-TNT. Acta Crystallographica Section D Biological Crystallography. 2004; 60:2210–2221. DOI: 10.1107/S0907444904016427 [PubMed: 15572774]
40. Eckhardt M, Gotza B, Gerardy-Schahn R. Mutants of the CMP-sialic acid transporter causing the Lec2 phenotype. The Journal of biological chemistry. 1998; 273:20189–20195. [PubMed: 9685366]
41. Chan KF, Zhang P, Song Z. Identification of essential amino acid residues in the hydrophilic loop regions of the CMP-sialic acid transporter and UDP-galactose transporter. Glycobiology. 2010; 20:689–701. DOI: 10.1093/glycob/cwq016 [PubMed: 20181793]
42. Stamm M, Staritzbichler R, Khafizov K, Forrest LR. AlignMe--a membrane protein sequence alignment web server. Nucleic Acids Res. 2014; 42:W246–251. DOI: 10.1093/nar/gku291 [PubMed: 24753425]

Europe PMC Funders Author Manuscripts

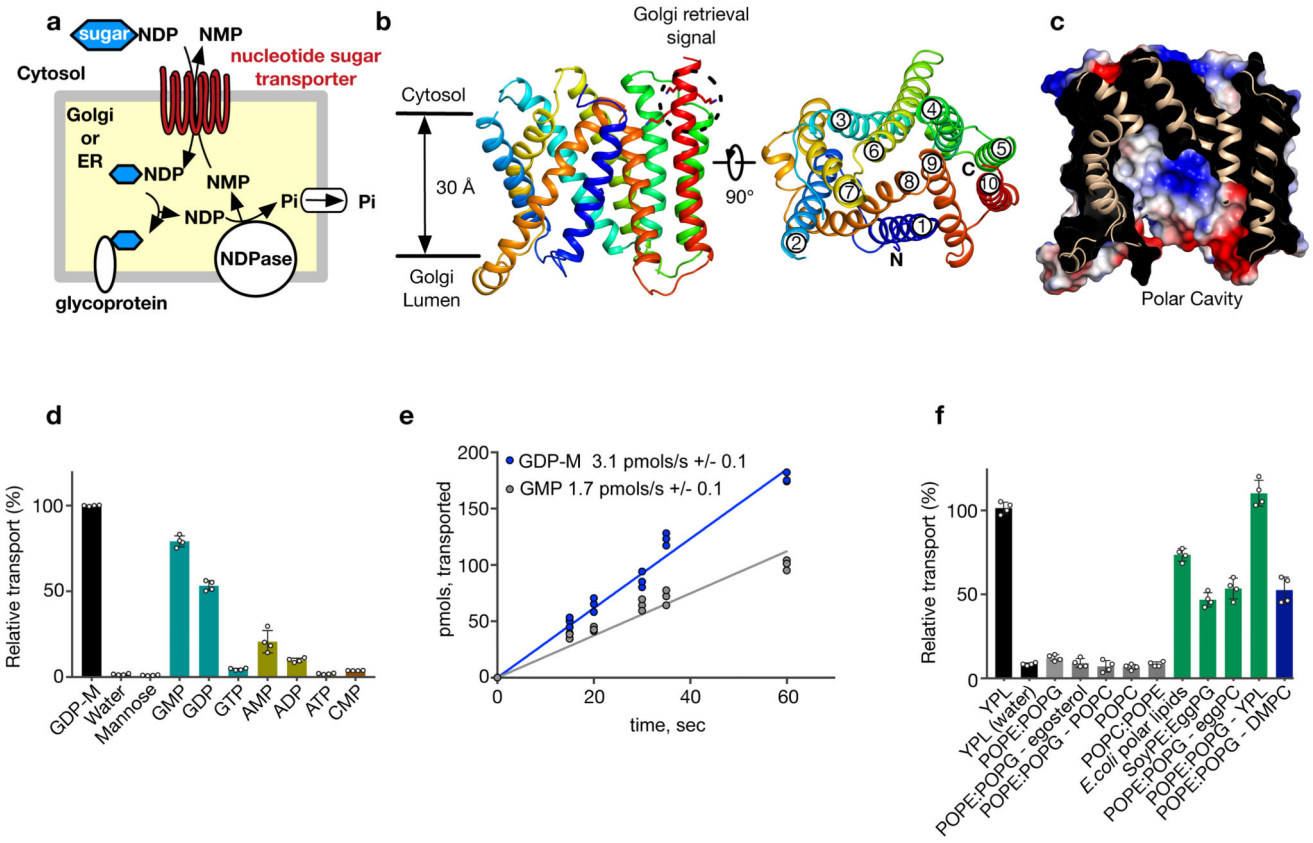


Figure 1. Crystal structure and functional analysis of Vrg4.

a, Nucleotide sugar transporters (NSTs) function to shuttle activated sugar donors (sugar-NDP) across the ER and Golgi membranes. **b**, Crystal structure of Vrg4 viewed from the Golgi membrane. **c**, Electrostatic surface representation of Vrg4 highlighting polar cavity. **d**, Vrg4 acts as a strict antiporter and has a strong preference for guanidine containing substrates. n=4 independent experiments, measure of centre mean, error bars s.d. **e**, Initial transport rates of GMP uptake into Vrg4 containing liposomes. n=3 independent experiments. **f,** Vrg4 requires short chain length lipids to function. n=4 independent experiments, measure of centre mean, error bars s.d.

Europe PMC Funders Author Manuscripts

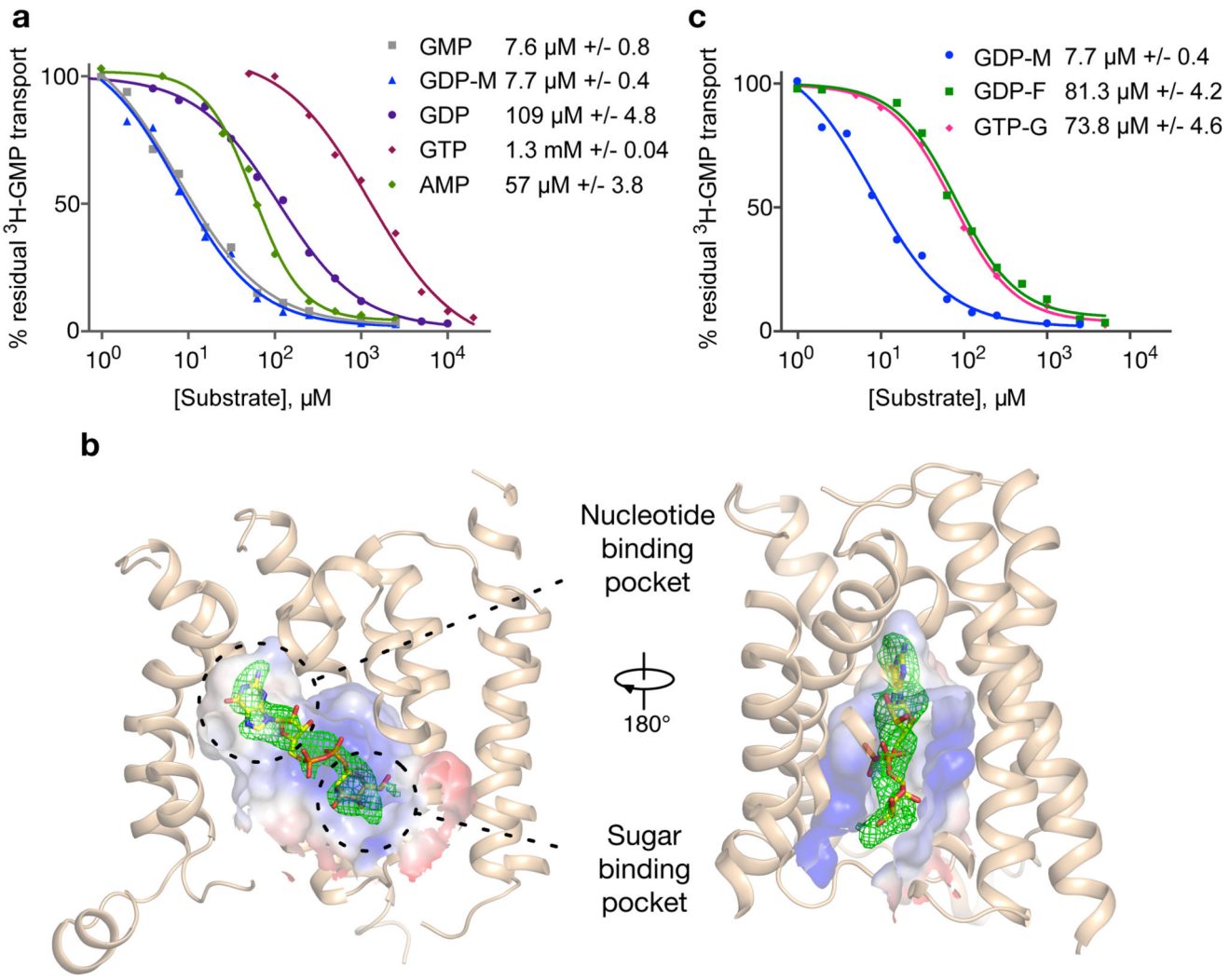


Figure 2. Recognition of GDP-mannose by Vrg4.

a, Representative IC_{50} curves calculated for wild type Vrg4. Calculated IC_{50} values are the mean of three independent experiments (Extended Data Fig 3), errors shown are s.d. **b,** The binding pocket in Vrg4 is shown in the plane of the membrane, coloured according to electrostatic surface, with bound GDP-mannose. The unbiased mFo-DFc electron density (green) is shown around the bound GDP-mannose ligand, contoured at 3σ. **c,** IC_{50} values for guanine conjugated nucleotide sugars. Calculated IC_{50} values are the mean of three independent experiments (Extended Data Fig 3), errors shown are s.d.

Europe PMC Funders Author Manuscripts

Europe PMC Funders Author Manuscripts

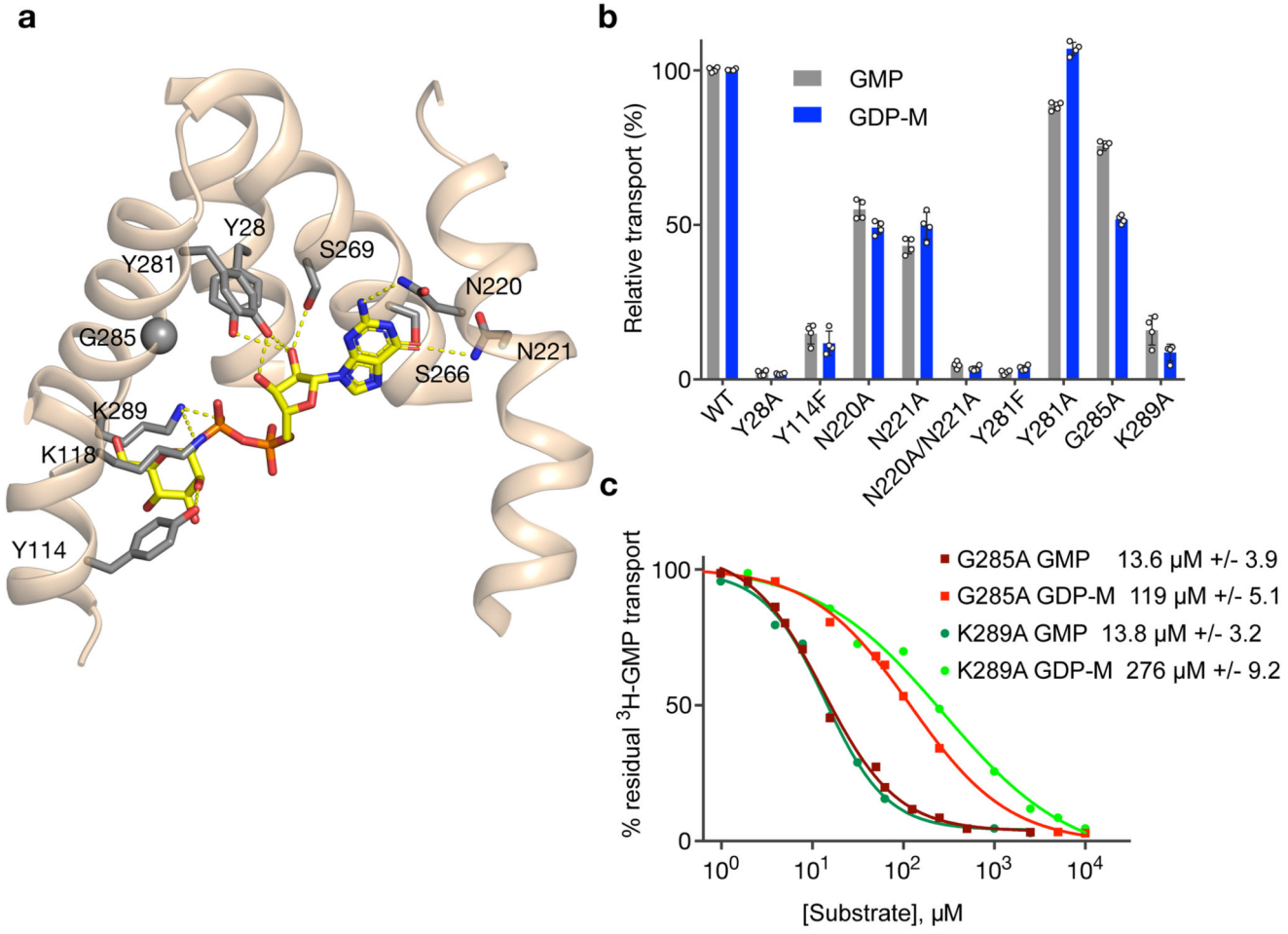


Figure 3. Molecular determinants of substrate specificity in Vrg4.

a, Binding site showing the main polar interactions made to GDP-mannose. The glycine of the GALNK motif is shown as a grey sphere, remaining binding site residues as sticks. **b,** Effect of binding site variants on transport of GDP-mannose or GMP substrates. n=4 independent experiments, measure of centre mean, error bars s.d. **c,** Effect of GALNK motif residues on GMP or GDP-mannose transport. Calculated IC_{50} values are the mean of three independent experiments (Extended Data Fig 3), errors shown are s.d.

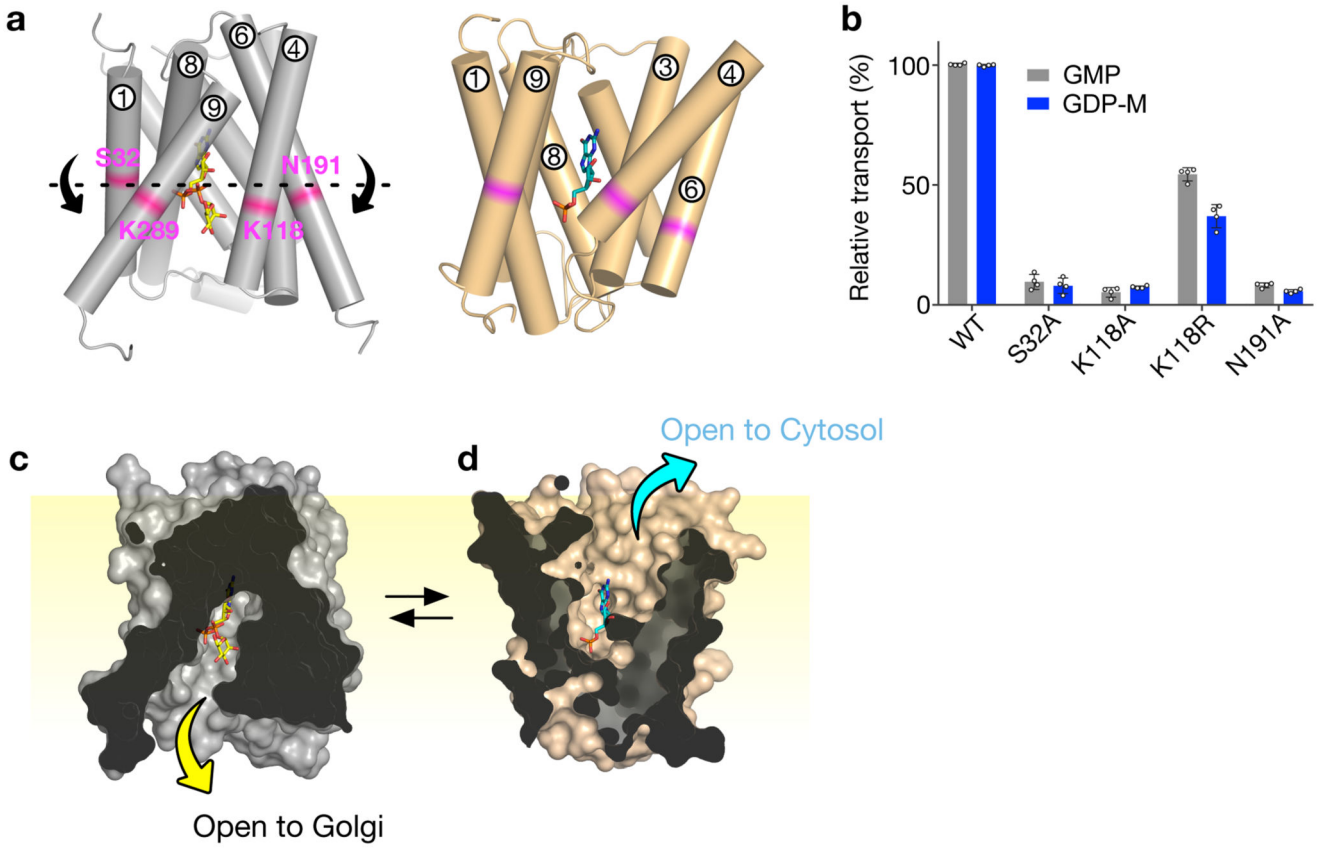


Figure 4. Alternating access transport mechanism of nucleotide sugar transporters.

a, Interacting residues (magenta) between the inverted topology repeats are highlighted. A model of the open to cytoplasm state is shown (wheat), revealing that reorientation of the transporter (black arrows) occurs at the same plane (dashed line) as the bound GDP-mannose (yellow) and GMP (cyan). **b,** Effect of replacing the interacting residues with alanine on transport. n=4 independent experiments, measure of centre mean, error bars s.d. **c,** Molecular surface representations showing entry and exit pathways to the central ligand binding site to the Golgi lumen in the crystal structure, and **d,** the repeat swapped model of the cytoplasmic facing state.

Europe PMC Funders Author Manuscripts