

**Classification: Biological Sciences: Medical Sciences**

**Structural elucidation of the *cis*-prenyltransferase NgBR/DHDDS complex reveals novel insights in regulation of protein glycosylation**

Ban H. Edani <sup>1,2</sup>, Kariona A. Grabińska <sup>1,2</sup>, Rong Zhang <sup>1,2</sup>, Eon Joo Park, <sup>1,2</sup>, Benjamin Siciliano <sup>1,2</sup>,  
Liliana Surmacz<sup>3</sup>, Ya Ha <sup>2\*</sup> and William C. Sessa <sup>1,2\*</sup>

<sup>1</sup> Vascular Biology and Therapeutics Program, <sup>2</sup>Department of Pharmacology, Yale University School of Medicine, 333 Cedar St, New Haven, CT 06520, USA

<sup>3</sup> Institute of Biochemistry and Biophysics, Polish Academy of Sciences, Pawinskiego 5a, Warsaw, Poland.

**\*Co-corresponding Authors:** Dr. Ya Ha, Department of Pharmacology, Yale University School of Medicine, 333 Cedar St, New Haven, CT 06520, E-mail: [ya.ha@yale.edu](mailto:ya.ha@yale.edu) or Dr. William C. Sessa, Department of Pharmacology, Vascular Biology and Therapeutics Program, Yale University School of Medicine, 10 Amistad Street, New Haven, CT 06520, E-mail: [william.sessa@yale.edu](mailto:william.sessa@yale.edu).

**Keywords:** dolichol, glycosylation, *cis*-prenyltransferase, X-ray crystallography

**Conflict of Interest Statement:** The authors have declared that no conflict of interest exists.

**Short Title:** Structure of NgBR/DHDDS complex

**Abstract:**

*Cis*-prenyltransferase (*cis*-PTase) catalyzes the rate-limiting step in the synthesis of glycosyl carrier lipids required for protein glycosylation in the lumen of endoplasmic reticulum. Here we report the crystal structure of the human NgBR/DHDDS complex, which represents the first atomic resolution structure for any heterodimeric *cis*-PTase. The crystal structure sheds light on how NgBR stabilizes DHDDS through dimerization, participates in the enzyme's active site through its C-terminal -RXG- motif, and how phospholipids markedly stimulate *cis*-PTase activity. Comparison of NgBR/DHDDS with homodimeric *cis*-PTase structures leads to a model where the elongating isoprene chain extends beyond the enzyme's active site tunnel, and an insert within the  $\alpha 3$  helix helps to stabilize this energetically unfavorable state to enable long chain synthesis to occur. These data provide unique insights into how heterodimeric *cis*-PTases have evolved from their ancestral, homodimeric forms to fulfill their function in long chain polyprenol synthesis.

**Significance Statement:**

The enzyme, *cis*-prenyltransferase (*cis*-PTase) composed of two subunits, NgBR and DHDDS, is essential for protein glycosylation reactions in all higher eukaryotes. Here we report the heterodimeric crystal structure of the complex and show novel features that impacts the stability, activity and lipid sensing of enzyme complex. Moreover, the structure rationalizes mutations that cause genetic disorders of glycosylation leading cognitive dysfunction, epilepsy and Parkinson's disease.

\body

## Introduction

Most members of the highly conserved *cis*-prenyltransferase (*cis*-PTase) family catalyze the sequential condensation of an allylic pyrophosphate with a variable number of 5-carbon isopentenyl pyrophosphates (IPPs), resulting in the formation of polyprenol pyrophosphates with *cis*-double bonds (1-4). In eukaryotes, the endoplasmic reticulum (ER) associated *cis*-PTase is the first enzyme committed to the synthesis of dolichol phosphate (DolP) (4-8), an indispensable lipid carrier for protein N-glycosylation, O-mannosylation, C-mannosylation and GPI-anchor formation (9, 10). In contrast to mammalian *cis*-PTase, undecaprenyl pyrophosphate synthase (UPPS), a bacterial *cis*-PTase, is essential for cell wall synthesis (11). Eukaryotic *cis*-PTases have the extraordinary ability to synthesize long-chain (14-24 C<sub>5</sub> isoprene units) or very-long-chain (>2,000 C<sub>5</sub>) products (12), whereas bacterial, some protistic, archaeal and plant enzymes mainly produce medium-chain (9-11 C<sub>5</sub>), or short-chain (2-5 C<sub>5</sub>), products (2, 3, 10, 13-16).

Although the structure and mechanism of homodimeric, bacterial *cis*-PTases, have been extensively studied (1, 17-22), the eukaryotic enzyme's unique mechanism has remained elusive until recently. Dehydrodolichyl diphosphate synthase (DHHDS) or its yeast orthologues Rer2 and Srt1, retain most of the active site residues in common with bacterial *cis*-PTase (4-6), however, as purified, have very little enzymatic activity (23). The full activity of *cis*-PTase requires the interaction of DHHDS with an additional subunit, originally identified as NgBR (Nogo-B receptor) or Nus1 in yeast (24-31). Within eukaryotic cells, DHHDS is stabilized by the association with NgBR (24). Part of NgBR shares sequence and structural homology with bacterial *cis*-PTase (30, 32, 33), but lacks most catalytic residues. Outside the *cis*-PTase homology domain, NgBR has a N-terminal membrane-binding region (24) (Fig. 1A). Various roles have been proposed for NgBR and NgBR-like proteins in the membrane association and function of eukaryotic *cis*-PTases and a subgroup of archaeal enzymes that share this heteromeric arrangement (24, 33-37). We have previously shown that NgBR and Nus1 are indispensable for *cis*-PTase activity in humans and in yeast, respectively (25, 33).

Exome sequencing of patients with congenital disorders of glycosylation (CDG) has identified a number of disease-causing mutations in both DHDDS and NgBR (25, 38-42). A homozygous mutation in NgBR at Arg-290 (NgBR R290H mutation), occurs in a conserved C-terminal -RXG- motif that is shared amongst NgBR orthologs as well as in bacterial homodimeric *cis*-PTases, but is absent in DHDDS and its orthologs, suggesting that the C tail of NgBR participates in *cis*-PTase function. Biochemical characterization of the R290H mutant revealed this mutation impairs IPP binding and markedly reduces catalytic activity but does not influence its interaction with DHDDS (25, 33) raising the possibility that NgBR could directly contribute to substrate binding and catalysis through its -RXG- motif (21, 43, 44). Another recent exome sequencing study identified multiple mutations in NgBR that appear to cause Parkinson's disease (45), although it is uncertain whether these mutations affect *cis*-PTase activity.

Here, we report the first crystal structure of a heteromeric, human *cis*-PTase NgBR/DHDDS complex solved at 2.3 Å resolution. Besides proving a catalytic role of NgBR's -RXG- motif, structural-functional analyses have unveiled several unique attributes of the complex that were not predicted based on the structures of UPPS. This includes a unique C-terminal clamp in DHDDS that contributes to heterodimerization; the mechanistic basis for a Parkinson's disease causing, loss of function mutation; a novel N-terminal membrane sensor critical for lipid activation of *cis*-PTase activity and a critical structural feature in DHDDS that impacts product polyprenol chain length. Thus, our crystal structure provides unique insights into how heterodimeric mammalian *cis*-PTases have evolved from their ancestral, homodimeric forms to synthesize long chain polyprenols and dolichol, lipids essential for protein glycosylation reactions.

## Results and Discussion

***The core catalytic domain of NgBR/DHDDS complex.*** The rate-limiting step in dolichol biosynthesis in the ER is catalyzed by *cis*-PTase, which consists of NgBR and DHDDS subunits (24, 25). Unlike homodimeric *cis*-PTase found in bacteria that synthesizes a medium-chain-length polyprenol pyrophosphate (e.g., *E. coli* UPPS; 11 isoprene units), human *cis*-PTase preferentially catalyzes the condensation of sixteen IPP molecules with a single farnesyl pyrophosphate (FPP), generating a long-chain reaction product (i.e. 19 isoprene units see Supporting Information (SI) in the SI appendix, Fig. S1) (10, 11). To facilitate structural and mechanistic characterization of the heteromeric *cis*-PTase, we co-expressed a polyHis- and SUMO-tagged human NgBR (46), which has the N-terminal 78 amino acids deleted (Fig. 1A), with full-length human DHDDS in *E. coli*, and purified the protein complex to homogeneity in milligram quantities (Fig. 1B). Consistent with the model where NgBR's N-terminal region mainly functions as a membrane anchor (24), the N terminal truncated NgBR/DHDDS complex is as active enzymatically as the full-length protein complex purified from mammalian cells (33). Also like the full-length protein, the activity of the truncated NgBR/DHDDS was potently stimulated by phosphatidylinositol (PI), a phospholipid abundantly present in the ER membrane (Fig. 1C) (33). Reverse phase thin layer chromatography (RPTLC) confirmed that NgBR N-terminal truncation did not alter the range, or relative abundance, of various long-chain polyprenols generated from the *in vitro* reaction (see below).

***The heterodimeric structure.*** The core catalytic domain of the NgBR/DHDDS complex was crystallized in the presence of IPP and Mg<sup>2+</sup>. The crystal belongs to space group R32, and the asymmetric unit contains a single NgBR<sup>79-293</sup>/DHDDS heterodimer (hereinafter referred to as NgBR/DHDDS). The structure was determined by molecular replacement using *E.coli* UPPS (PDB entry 1X06) and *S. cerevisiae* Nus1 (PDB entry 6JCN) as search probes for DHDDS and NgBR, respectively (Fig. 2A; SI appendix, Table S1) (32, 47). Clear electron densities enabled us to model several functionally important elements that were absent

in the search probes (SI appendix, Fig. S2-5). In NgBR, these structural elements include an N-terminal  $\alpha$ -helix ( $\alpha$ 1, residues 82-93) and an extended C-terminal segment (residues 286-293) that contains the catalytically essential -RXG- motif. Furthermore, the two outermost  $\beta$ -strands ( $\beta$ C,  $\beta$ C') observed in Nus1 are disordered (residues 230-244; SI appendix, Fig. S4). The peptide segment following the  $\beta$ -strands forms a new helix ( $\alpha$ 4, residues 179-186) and partially covers the hydrophobic cavity previously proposed to constitute the binding site for farnesylated Ras (48). In DHDDS, the additional structures include a conserved N-terminal segment (residues 1-24) that reaches into the hydrophobic tunnel of the active site, and a pair of long  $\alpha$ -helices (residues 251-333) toward the C-terminus that wrap around the protein complex.

The crystal structure of a heterodimer is consistent with earlier predictions that the two proteins stabilize each other (24, 25) and that the central region of the NgBR/DHDDS interface is similar to that observed in prokaryotic UPPS and yeast Nus1 homodimers (SI appendix, Fig. S6) (32, 47). In contrast to a recent modeling study (23), the crystal structure reveals that DHDDS's C-terminus wraps around the protein complex and makes additional contact with NgBR. To test the nature of the interaction between the two subunits, partial proteolysis experiments were performed. As seen in Fig 2B; (and in SI appendix, Fig. S7) the initial cleavage of NgBR occurs prior to cleavage of DHDDS consistent with the proteins stabilizing each other. Moreover, mutations of several amino acids at the interface clearly weaken the binding between NgBR and DHDDS when the constructs were transfected into HEK293T cells and proteins immunoprecipitated (Fig. 2C, SI appendix, Fig. S8). To study the contribution of DHDDS's extended C-terminus *in vivo*, two DHDDS truncation mutants ( $\Delta$ 256,  $\Delta$ 289), where part of  $\alpha$ 7 and the entire  $\alpha$ 8 helix was deleted, and a triple mutant (R306A/F313A/L317A), which disrupts the packing of the two helices, were generated (SI appendix, Fig. S3B). Both deletion mutants failed to support growth of *rer2 $\Delta$ /srt1 $\Delta$ /nus1 $\Delta$*  yeast cells (lacking endogenous genes critical for *cis*-PTase, (25), whereas the triple mutant delayed growth (Fig. 2D) implying a critical functional role for the C tail of DHDDS.

**The active site of the NgBR/DHDDS complex.** Both NgBR and DHDDS subunits are required for a functional *cis*-PTase (24, 25). Although NgBR, and its yeast homolog Nus1, have a similar fold as DHDDS and members of the homodimeric *cis*-PTase family, they lack multiple key catalytic residues and have a distorted “active site” unable to accommodate substrate. As we recently demonstrated in a functional study with purified human complex, the NgBR subunit contributes to *cis*-PTase activity through its C-terminal -RXG- motif, which is conserved amongst hetero- and homodimeric enzymes (33) and this critical -RXG- motif is disordered in the reported homodimeric Nus1 structure (32). Here, based on clear electron densities, we modeled the entire C-terminal region of NgBR, as well as two IPP molecules, one (IPP1) occupying the allylic substrate binding site S1, and the other (IPP2) occupying the homoallylic site S2, and a bridging Mg<sup>2+</sup> ion (Fig. 3A; SI appendix, Fig. S2B, S5 and S9). The C-terminus of NgBR has an extended conformation and traverses the dimeric interface to complete the active site (Fig. 3B): the mainchain amide groups of Leu-291 and Gly-292 of the -RXG- motif form critical hydrogen bonds with homoallylic substrate IPP2’s β-phosphate, while the side chain of Arg-290 interacts with a water molecule that coordinates the Mg<sup>2+</sup> ion. Gly-292 of the -RXG- motif enables the peptide to make a tight turn over IPP2 and further stabilizes the turn by forming a hydrogen bond with Arg-85, a conserved DHDDS residue that also participates in the binding of the allylic substrate. The function of the -RXG- motif is thus equivalent to that of the P-loop, which binds the β-phosphate group of the allylic substrate FPP.

The binding interactions within the S1 site are highly conserved within the *cis*-PTase family (SI appendix, Fig. S10). Despite weak electron density, we were confident in modeling IPP1 in such a way that its isopentenyl group is pointed toward the hydrophobic tunnel where the elongating product is bound (Fig. 3A). Unlike the S1 site, in previous studies, substrate binding to the S2 site was less consistent among various crystal structures (SI appendix, Fig. S11) and clear from our study. Some differences were caused by mutations introduced into the active site, but differences may also result from crystallization conditions that destabilized the -RXG- motif since the disordered motif is almost always associated with poorer occupancy of the S2 site, or with misplaced pyrophosphate groups. Amongst known *cis*-PTase structures,

the binding mode of IPP2 is most similar to that observed in the complex between MIDPPS and substrate analogs (SI appendix, Fig. S11). In both structures, the pyrophosphate is tightly bound by three conserved arginine residues (DHDDS Arg-85, Arg-205, Arg-211), a conserved serine (Ser-213), and the -RXG- motif from the dimerization partner. The magnesium ion is coordinated by phosphate groups from both allylic and homoallylic substrates, a conserved aspartate from the “P-loop” (Asp-34), and two water molecules. A conserved asparagine (Asn-82) is identically positioned near IPP’s C2 group for proton extraction during catalysis. These structural similarities reinforce the notion that hetero- and homo- dimeric *cis*-PTases share the same catalytic mechanism.

The majority of CDG-causing missense mutations in DHDDS (R37H, R38H, R211Q) and NgBR (R290H) affects active site residues directly involved in substrate binding and catalysis (SI appendix, Fig. S12A). Although uncharacterized biochemically, the DHDDS<sup>T206A</sup> mutation could also perturb the active site because the threonine hydroxyl is simultaneously hydrogen-bonded to the backbone amide and carbonyl of the metal-binding Asp-34 (SI appendix, Fig. S12B (49)). The only exception to this pattern is DHDDS<sup>K42E</sup>, which affects ~17% of Ashkenazi Jewish patients diagnosed with retinitis pigmentosa (42, 50). It was previously unclear how this mutation could affect enzyme activity because the corresponding residue in *E. coli* UPPS (Lys-34) is exposed and does not interact with any other residue that could link the mutation to the active site. Here, we show that instead of pointing toward solvent, Lys-42 forms a salt bridge with the highly conserved Glu-234 (Fig. 3C); and this interaction is equivalent to that between Trp-31 and Asp-223 in *E. coli* UPPS. Therefore, this charge reversal mutation could be disruptive to protein structure, affecting the short helix ( $\alpha$ 1) that contains Lys-42. Given the role of  $\alpha$ 1 and the preceding P-loop in FPP binding, this interpretation is consistent with the observation that K42E mutation increases the  $K_M$  for FPP, decreases  $k_{cat}$ , but has no effect on IPP binding (33). K42E also causes product chain shortening (25, 49), which will be discussed below.



**NgBR mutations and Parkinson's disease.** Recently, a number of mutations in the *NUS1* gene were discovered by exome sequencing in patients with Parkinson's disease (PD) (45). Amongst the 15 missense mutations, 11 occur in the region that we now have structural information for (Fig. 4A). In contrast to CDG-causing mutations, which cluster around the active site, most PD mutations are scattered throughout the three-dimensional structure and do not appear to have any direct effect on protein folding (7 of the 11 mutated residues are solvent-exposed; Fig. 4D), heterodimerization with DHDDS, or substrate binding. Therefore, for these mutations, it remains uncertain whether they altered *cis*-PTase activity *in vivo*, or impacted another biochemical mechanism involving NgBR that is unrelated to *cis*-PTase or dolichol function. It is not yet known whether any of the PD patients had symptoms overlapping with the CDG spectrum earlier in life (45), or if CDG patients with mild DHDDS mutations, e.g., K42E, could have a higher risk for developing Parkinson's disease.

One PD mutation (NgBR<sup>G91C</sup>), however, could indirectly affect the enzyme's active site through the -RXG- motif. As alluded to earlier, the NgBR structure has an N-terminal helix ( $\alpha$ 1) that is absent in either yeast Nus1 or bacterial UPPS structures (highlighted in green in Fig. 4A). A lysine residue (Lys-96) at the end of this helix plays an important structural role in stabilizing the NgBR's C-terminal segment by forming two critical hydrogen bonds with the backbone carbonyl of Tyr-284 and the side chain of Gln-289 (Fig. 4B). Gly-91 mediates the contact between  $\alpha$ 1 and the rest of the protein and introducing a cysteine at this position would be incompatible with the packing of the helix and could perturb, through Lys-96, the C-terminal segment that harbors the -RXG- motif. To investigate this possibility, we generated and biochemically characterized the NgBR<sup>G91C</sup> mutant. The purified, G91C complex had a ~40% reduction in the enzyme's specific activity *in vitro* (Fig. 4C). In comparison, the CDG mutation DHDDS<sup>K42E</sup> causes an 80% reduction in enzyme activity compared to wildtype activity. This finding is significant and raises the possibility that subtle changes of *cis*-PTase activity could contribute to the pathogenesis of PD. These data are consistent with the observation that a splice variant of NgBR reducing its mRNA levels by 50% also increases PD risk (45). Dolichol is the most abundant lipid associated with neuromelanin, a dark pigment

enriched in catecholaminergic neurons that are selectively vulnerable in Parkinson's disease (51, 52). The genetic finding of mutations affecting a critical component of the enzyme complex responsible for dolichol synthesis, and our biochemical characterization of one such mutation, suggest a possible link between dolichol metabolism and age-related neurodegeneration.

**Regulation of enzyme activity by membrane binding.** Heterodimeric *cis*-PTases are invariably found to be peripherally associated with large hydrophobic structures like membrane bilayers, lipid droplets and rubber particles (12, 24, 53). Two hydrophobic segments within the N-terminal region of NgBR is responsible for stable anchoring of the NgBR/DHDDS complex to the ER membrane. Our crystallographic analysis of the core catalytic domain reveals a unique N-terminal structure within DHDDS that is absent in bacterial UPPS. The N-terminal segment consists of a random coil (residues 1-10) and a short  $\alpha$ -helix ( $\alpha 0$ ; residues 11-21). The coil starts from the gap between  $\alpha 2$  and  $\alpha 3$ . A bulky side chain (Trp-3) is inserted between the two helices, constricting the hydrophobic tunnel that is predicted to house the polyisoprenyl chain (Fig. 5A, left pane; Fig. 5B). A conformational change displacing the coil from the tunnel is thus required to enable product elongation (Fig. 5A, right panel). The short helix ( $\alpha 0$ ) lies alongside  $\alpha 7$  and forms the outermost tip of the entire protein complex. Intriguingly, the side of  $\alpha 0$  that is exposed to the solvent contains three hydrophobic residues (Trp-12, Phe-15, Ile-19). Sequence analysis indicates that this continuous hydrophobic patch is conserved in all DHDDSs, raising the possibility that  $\alpha 0$  may have evolved specifically for membrane binding (Fig. 5C). We generated a triple DHDDS mutant (W12A/F15A/I19A) to examine the role of the hydrophobic patch in enzyme function. The mutant has similar activity as the wildtype enzyme in the absence of phospholipid. However, W12A/F15A/I19A is no longer potently activated by the addition of phosphatidylinositol (Fig. 5D) and the dominant polyprenol generated by the W12A/F15A/I19A mutant is also three isoprene units longer (Fig 5E). We hypothesize that membrane binding would trigger a conformational change that results in the unblocking of the hydrophobic tunnel during chain elongation and destabilizing the enzyme:product complex to facilitate product release (see

below). Therefore, DHDDS's  $\alpha 0$  may function as a membrane sensor, contributing to enzyme activation by phospholipid and to the release of polyisoprenyl product of the appropriate chain length.

**The mechanism determining product chain length.** Eukaryotic heterodimeric *cis*-PTases differ categorically from bacterial homodimeric enzymes in that they not only generate products with longer chain lengths but also tend to generate a range of products. The human NgBR/DHDDS complex preferentially synthesizes  $C_{95}$ , but as Fig. 5E illustrates,  $C_{95}$  is one of many polyprenols that differ by single isoprene unit and are present in varying amounts. Here we examine the mechanistic basis for this biochemical phenomenon in light of the new crystal structure.

The size of the hydrophobic tunnel cannot explain the longer product ( $C_{95}$ ) generated by the NgBR/DHDDS complex. After manually removing the N-terminal coil that blocks the hydrophobic tunnel (Fig. 5A), we used 3V web server to calculate the volume of DHDDS's tunnel to be around  $1,300 \text{ \AA}^3$  (54), which is comparable in size to that ( $1,200 \text{ \AA}^3$ ) of *E. coli* UPPS (Fig. 5B; UPPS generates  $C_{55}$  product). Therefore, as polyprenol intermediates are formed, either the protein tunnel has to expand significantly, or the tail of the polyprenol chain needs to be pushed out of the tunnel. A model of protein conformational change that involves a large movement of helix  $\alpha 3$  has been proposed (55). This type of movement, however, is unlikely to occur during product elongation because  $\alpha 3$ 's N-terminal half is involved in binding the allylic substrate's pyrophosphate group, and thus essential for subsequent rounds of reaction. We favor the second possibility where a large portion of the polyprenol chain exits the active site tunnel and becomes exposed on the protein surface. Mutagenesis studies performed on *E. coli* UPPS suggested that Leu-137, equivalent to DHDDS's Cys-148, located near the end of the tunnel could function as the barrier to such an exit since the UPPS<sup>L137A</sup> mutant synthesizes a range of long-chain products (up to  $C_{75}$ ) in the absence of detergent (Fig. 5B) (19).

Sequence alignment revealed that all eukaryotic heterodimeric *cis*-PTases carry a short insertion (human DHDDS Glu-107, Lys-108, Glu-109), in the middle of the  $\alpha 3$  helix (Fig. 6A). The extra sequence creates a greater bulge in  $\alpha 3$  but does not significantly increase the volume of the hydrophobic tunnel that lies beneath the helix (Fig. 5A; left panel and Fig. 6C). To examine the possibility that the insertion may impact product length, we generated an  $\Delta$ EKE deletion mutant of DHDDS and transformed it and wildtype DHDDS into the *S. cerevisiae* triple deletion strain (*rer2* $\Delta$ /*srt1* $\Delta$ /*nus1* $\Delta$ ) with wildtype NgBR. When expressed in yeast, the wildtype enzyme generated Dol-20 (C<sub>100</sub>) as the dominant product. Although the  $\Delta$ EKE mutant generated shorter products (Fig. 6B), it retained the two key characteristics of long chain *cis*-PTase (i.e. chain length distribution of products and each product > C<sub>55</sub>). Therefore, the wider opening between  $\alpha 2$  and  $\alpha 3$  created by the sequence insert, *per se*, is not a required structural feature of the long-chain enzyme.

If the tail of the elongating polyprenol product does exit the enzyme's hydrophobic tunnel, this is expected to generate several consequences (Fig. 6D). (i) Since there is no longer a physical barrier, the polyprenol chain can grow indefinitely in theory. This distinguishes the long-chain *cis*-PTases from the short-chain and medium-chain *cis*-PTases. In the latter two, the product chain length is strictly determined by the size of the hydrophobic tunnel (19, 21). (ii) The stability of the enzyme complex with product intermediates likely plays a major role in determining when elongation reaction stops. The exposed portion of the hydrophobic polyprenol could be increasingly disruptive to the enzyme:product complex through various mechanisms, e.g., interaction with detergent micelles, membrane bilayers, or lipid droplets. In the example cited above, *E. coli* UPPS<sup>L137A</sup> mutant failed to produce long-chain product in the presence of TX100 (19). (iii) The insert in  $\alpha 3$ , we thus argue, plays a role in stabilizing the enzyme:product complex. One possibility, as schematically illustrated in Fig. 6D, is that the insert may confer greater conformational flexibility to the  $\alpha 3$  helix and its surrounding regions. (iv) The subtle effect of the environment on the enzyme:product complex determines that the final product will invariably have a distribution of chain lengths, another defining feature of the long-chain *cis*-PTases. The most extreme example of a long chain

*cis*-PTase is rubber synthase, together with its protein cofactors, must have extraordinary stability in its product-bound form (12).

In summary, we provide the first atomic insights into how eukaryotes synthesize dolichol, the obligate carrier lipid for N-glycosylation, O- and C-mannosylation reactions and GPI anchor biosynthesis. The new features elucidated by this structure rationalize the stability of the heterodimeric complex, disease causing mutations and lipid regulation of *cis*-PTase activity. The mechanism suggested by the NgBR/DHDDS crystal structure provides a conceptual framework for understanding the unique enzymatic properties of the long-chain *cis*-PTases, including rubber synthase. Obligate heterodimerization with a membrane-binding partner probably ensures that the long-chain product is only robustly synthesized in close vicinity to the membrane as elsewhere the exposed polyprenol could be harmful to the cell.

## METHODS

**Materials :** Unless otherwise stated, all reagents were of analytical grade and purchased from Sigma-Aldrich, Thermo Fisher Scientific, and Zymo Research (Irvine, CA). Restriction enzymes were from New England Biolabs (Ipswich, MA). [1-<sup>14</sup>C] IPP (50 mCi/mmol) was purchased from American Radiolabeled Chemicals (St. Luis, MO). Reverse phase thin layer chromatography (RP18-HTLC) plates were from MilliporeSigma (cat# 1.51161.0001). Primary antibodies used in this study include Anti-HA High Affinity antibody (Roche, 11867423001) and Monoclonal anti-Flag M2 antibody (Sigma, F3165). HiFi DNA Assembly method (NEBuilder<sup>®</sup>, NEB) was used to construct expression vectors and perform site-directed mutagenesis. List of plasmids is in supplementary Table 2-4. Primers used in cloning are listed in supplementary Table 5 and primers used for mutagenesis are listed in supplementary Table 6.

**Cloning and Purification of NgBR/DHDDS Complex:** To express His-SUMO-NgBR<sup>79-293</sup> and untagged, full length DHDDS (1-333) in bacteria, His-SUMO and NgBR overlapping PCR fragments were first assembled into pRSF-DUET1 vector cut with NdeI/XhoI restriction enzymes. DHDDS PCR fragment was

then assembled into pRSF-DUET1-HIS-SUMO-NgBR cut with NcoI/NotI. The Recombinant NgBR/DHDDS complex was expressed in *Escherichia coli* Rosetta (DE3) cells (Novagen) and induced with 0.7 mM IPTG (OD<sub>600</sub> 0.6) overnight at 18 °C. Cells were harvested and then resuspended in lysis buffer containing 20 mM Tris-HCl pH 8.0, 500 mM NaCl, 20 mM imidazole, 10% glycerol, 0.5% triton X-100 and 2 mM 2-Mercaptoethanol, cOmplete protease inhibitors (Roche), lysozyme (100 µg/ml) and DNase I (10 µg/ml). Three cycles of freeze/thaw were conducted using ethanol/dry ice bath and cells were sonicated in 50 ml falcon tube for 2 minutes total. The samples were clarified by centrifugation at 20,000 rpm for 1 hour at 4 °C. Supernatant was then applied to a 1 ml HisTrap (GE Healthcare) nickel affinity column, and the protein was eluted with 6 ml lysis buffer containing 400 mM imidazole. The sample was then applied to size exclusion column (Superdex 200, GE Healthcare) equilibrated with buffer containing (50 mM Tris-HCl pH 8.0, 150 mM NaCl, 1 mM MgCl<sub>2</sub> and 2 mM TCEP). Fractions containing protein complex were collected and subjected to cleavage with SUMO protease overnight at 4°C to remove the His-SUMO tag. The cleaved protein was then re-applied to a HisTrap column and the flow-through was collected. Protein sample was then passed through another size exclusion column equilibrated with 50 mM Tris-HCl pH 8.0, 150 mM NaCl, 2.5 mM MgCl<sub>2</sub> and 2 mM TCEP. Fractions were collected and analyzed by 12% SDS-PAGE gel.

**Crystallization, Data Collection and Structural Determination:** The purified protein was concentrated to 3.2 mg/ml and incubated with 3.3 mM IPP (sigma) on ice for 2 hours. Crystallization screening was performed using the sitting-drop vapor diffusion method, and an initial hit was obtained from PEG screen (Hampton Research). Crystallization was optimized by grid screening and the best crystals were obtained by mixing 1 µl protein solution with 1 µl reservoir solution consisting of 0.1 M Bicine (pH 8.5), 10% v/v 2-propanol, 22% PEG 1500. Crystals appeared within 2 days and grew to maximum size within one week at room temperature. Crystals were cryoprotected with the reservoir solution supplemented with 20% glycerol and flash frozen in liquid nitrogen. Diffraction Data were collected on beamline 24-ID-E of the Advanced Photon Source at Argonne National Laboratory and processed using *HKL2000* (56). Although

individual reflections up to 2.2 Å resolution could be observed, after merging,  $CC_{1/2}$  quickly fell off beyond 2.3 Å. The structure of the complex was determined by molecular replacement and refined to 2.3 Å resolution using CCP4i (supplementary Table 1) (57). The *E. coli* UPPs (PDB entry 1X06) and *S. cerevisiae* NUS1 (PDB entry 6JCN) were used as search probes for the DHDDS and NgBR subunits, respectively. Model building were performed using *Coot* (58).

***cis*-PTase activity of NgBR/DHDDS:** The steady state activity of purified NgBR/DHDDS complex was assayed as before with minor modifications (33). Briefly, a standard incubation mixture contained, in a final volume of 25 µl, 100 µM [ $1-^{14}C$ ] IPP, 20µM FPP, 50 mM Tris-HCl pH 8.0, 1 mM  $MgCl_2$ , 10 mM KF, 20 mM 2-mercaptoethanol, 1 mg/ml BSA, 1 % Phosphatidylinositol and 100 ng of purified enzyme. The mixture was incubated for 1 hour at 37 °C and product was extracted with chloroform:methanol (3:2), followed by washing three times with 1/5 volume of 10 mM EDTA in 0.9% NaCl. In order to determine the chain length of the *cis*-PTase products, polyprenol diphosphates were chemically dephosphorylated by incubation of the lipids at 90° in 1 N HCl for 1 hr. Dephosphorylated prenols were extracted three times with two volumes of hexane. The organic fraction was washed with 1/3 volume of water, hexane was evaporated under stream of nitrogen and lipids were loaded onto HPTLC RP-18 precoated plates and run in acetone containing 50 mM  $H_3PO_4$ . The plates were exposed to film to visualize the products of IPP incorporation. As an internal and external standards Geranylgeraniol (Echelon Biosciences), Undecaprenol and Polyprenol 19 (*Institute of Biochemistry and Biophysics, PAS the Collection of Polyprenols*) and Prenols mixture (13-21) (Avanti Polar Lipids) were used. Prenol standards were visualized by exposing the TLC plate to iodine vapor.

### **Limited Proteolysis**

20 ul of 0.2 mg/ml of purified NgBR/DHDDS enzyme was incubated with 5 ul of a protease including thermolysin (*Sigma*), proteinase k (*Sigma*) and trypsin (*Sigma*) at different concentrations (0.005, 0.01, 0.02, 0.04, 0.08, 0.16 mg/ml). The reaction mixture was incubated at room temperature for 30 minutes and

stopped with SDS buffer containing 5.2 mM PMSF and 5.2 mM EDTA. Samples were boiled for 5 minutes and analyzed by 12% SDS-PAGE gel.

### **Co-Immunoprecipitation**

HEK293T cell was transfected using lipofectamine 2000 (Invitrogen) according to the manufacturer's protocol and harvested 48 hrs after transfection. Cells were collected and lysed in IP buffer (IP buffer: 50mM HEPES, 150mM NaCl, 1mM EDTA, 1% Triton X-100, Complete Protease Inhibitors (Roche)). Lysates were cleared by centrifugation at 12000 rpm for 10 min and 10 $\mu$ l of anti-Flag M2 magnetic beads (Sigma) was used to pulldown the Flag-tagged protein from 0.5-1 mg of cell lysate. After incubation for 2 hours at 4°C, magnetic beads were washed with IP buffer, resuspended in 2X Laemmli sample buffer and boiled for 5 min before western blot analysis.

### **Yeast Complementation Assay**

For yeast complementation analysis of *cis*-PTase, *S. cerevisiae* strains KG405 (*nus1 $\Delta$  rer2 $\Delta$  srt1 $\Delta$* ), carrying the *Glcis*-PTase encoding gene on a plasmid with a *URA3* marker was used (25). To phenotypically analyze human *cis*-PTase mutants, strain KG405 was transformed with vectors pKG-GW1 carrying DHDDS variants (leucine selection) and pKG-GW2 carrying NgBR variants (methionine selection) in combination or empty vectors as negative control. Transformed yeast cells were grown overnight at 30 °C in synthetic defined medium or lacking uracil, methionine, and leucine were streaked onto synthetic defined medium containing all amino acids, nucleotide supplements, and 1% (w/v) 5-FOA (Zymo Research) and onto YPD plates. The plates were incubated for up to 5 days at 30 °C. Colonies growing on the 5-FOA plates were streaked on synthetic defined medium lacking uracil and incubated at 30 °C for 3 days to verify the loss of the pNEV-*Glcis*-PTase plasmid. Yeast strain KG405 and its derivative carrying NgBR/DHDDS complex were cultured in 2% (w/v) Bacto peptone and 1% (w/v) yeast extract supplemented with 2% glucose (w/v) (YPD). Synthetic minimal media were made of 0.67% (w/v) yeast nitrogen base and 2% (w/v) supplemented with auxotrophic requirements. For solid medium, agar (BD



Biosciences, Sparks, MD) was added at a 2% (w/v) final concentration. Yeast cells were transformed using the Frozen-EZ yeast transformation II kit (Zymo Research).

### **HPLC Analysis**

To estimate the chain length of dolichols produced by EKE deletion mutant, total lipids from 3 g of yeast cells grown overnight till late logarithmic phase of growth (OD 3-4) were extracted by the modified Folch method. Lipids extracted from yeast cells were hydrolyzed in hydrolytic solution containing toluene/7.5% KOH/95% ethanol (20:17:3, v/v/v) for 1h at 90°C. Nonsaponifiable lipids were then extracted four times with hexane, purified on silica gel 60 columns using isocratic elution with 10% diethyl ether in hexane, evaporated to dryness in a stream of nitrogen and dissolved in isopropanol. Extracts were analyzed by HPLC using a Waters dual-pump HPLC device coupled with a Waters Photodiode Array Detector (spectrum range: 210 - 400 nm) and ZORBAX XDB-C18 (4.6 × 75 mm, 3.5 μm) reversed-phase column (Agilent, USA). Polyisoprenoids were eluted using the solvent mixtures A - methanol/water, 9:1 (v/v) and B - methanol/isopropanol/hexane, 2:1:1 (v/v/v) combined as follows from 0% B to 75% B in 20 min, from 75% B to 90% B in 5 min, from 90% B to 100% B in 2 min, 100% B maintained for 11 min, then from 100% B to 0% B in 1min at a flow rate of 1.5 mL/min. The chain length and identity of lipids were confirmed by comparison with external standards of a polyprenol (Pren-10 – Pren-24) and dolichol (Dol-17 – Dol-23) mixtures.

**Data Availability Statement:** All data are available in the paper and in SI appendix. Unique reagents will be readily available to the scientific community.

**Author contributions:** BE, KG, EJP, RZ, YA and WCS designed experiments; BE, EJP, KG, BS performed experiments; BE and RZ conducted protein purification and crystallization trials; BE and YH analyzed crystallographic data; LS performed HPLC analysis of dolichol species; BE, KG, YH and WCS wrote the manuscript. All authors contributed to the review of the manuscript. YH and WCS are co-

corresponding authors. The project emanated from the Sessa lab and where the proteins were purified, and enzymology studied. The Ha lab was essential for crystallization expertise and structural insights resulting in a logical arrangement of both being co-corresponding authors.

**Acknowledgements:** This work was supported by Grants R35 HL139945 (to WCS), an American Heart Association Scientist Development Grant (to EJP) and a National Science Centre of Poland Grant UMO-2016/21/B/NZ1/02793 (to LS). Additionally, the work was supported by the Northeastern Collaborative Access Team beamlines, which are funded by the National Institute of General Medical Sciences from the National Institutes of Health (P30 GM124165). The Pilatus 6M detector on 24-ID-C beam line is funded by a NIH-ORIP HEI grant (S10 RR029205). This research used resources of the Advanced Photon Source, a U.S. Department of Energy (DOE) Office of Science User Facility operated for the DOE Office of Science by Argonne National Laboratory under Contract No. DE-AC02-06CH11357.

## Figure Legends

### Figure 1. The catalytic core domain of human *cis*-PTase.

(A) Comparing the domain structure of NgBR, DHDDS and *E.coli* UPPS. The *cis*-PTase domain is colored yellow; N- and C-terminus of NgBR and DHDDS, gray; N-terminus of *E.coli* UPPS, blue.

(B) Purification of NgBR/DHDDS complex. Left Panel: Coomassie-stained SDS/PAGE showing purification steps. Lane 1: uncleaved 6HIS-SUMO-NgBR/DHDDS complex, Lane 2: cleaved NgBR/DHDDS complex and SUMO, Lane 3: NgBR/DHDDS complex after removing SUMO. Right panel, Size exclusion chromatography profile of the purified complex after cleavage with SUMO protease.

(C) Stimulation of *cis*-PTase activity of NgBR/DHDDS complex by phosphatidylinositol (PI). The values are means  $\pm$ S.D. of eight independent measurements from two independent purifications.

### Figure 2. The overall structure of NgBR/DHDDS heterodimer.

(A) Ribbon diagrams showing the front and back of the heteromeric complex. NgBR is colored in orange and DHDDS is colored in deep teal.  $Mg^{2+}$  ion is shown as a gray sphere, IPP molecules occupying S1 and S2 sites are shown in red and green, respectively.

(B) Limited proteolysis of the core domain was performed by incubating the protein with increasing amounts of thermolysin (0.005, 0.01, 0.02, 0.04, 0.08, 0.16 mg/ml). Untreated protein sample is denoted as (U). In this gel, thermolysin co-migrates with the full-length DHDDS.

(C) Co-immunoprecipitation of NgBR/DHDDS mutations introduced at the complex interface. HEK293T cells were co-transfected with NgBR-HA and Flag-DHDDS cDNAs; cells were lysed 48 hours post-transfection and immunoprecipitation performed using anti-flag magnetic beads. The lysate was analyzed by western blotting.

(D) Characterization of *cis*-PTase mutants in the C-terminal region of DHDDS using yeast complementation assay. The *nus1 $\Delta$  rer2 $\Delta$  srt1 $\Delta$*  deletion strain expressing *G. lamblia cis*-PTase from *URA3* plasmid was co-transformed with *MET15* bearing wildtype (WT) NgBR and the *LEU2* plasmid bearing either WT or mutant variants of DHDDS at the C-terminus. Three variants were analyzed including

a triple mutation (3A) corresponding to R306A, F313, L317 and two DHDDS truncation  $\Delta 256$  and  $\Delta 289$ . The cells were streaked onto complete plates (YPD) or synthetic complete medium containing 1% FOA. The Ura3 protein, which is expressed from the *URA3* marker converts FOA to toxic 5-fluorouracil, forcing the cells to lose the *G. lamblia cis*-PTase plasmid. Cell growth was monitored over time to assess phenotypic differences.

**Figure 3. The active site of the NgBR/DHDDS complex.**

(A) Omit difference map, countered at  $3.0 \sigma$  level, showing the two IPP molecules and  $Mg^{2+}$  ion bound at the active site. IPP1 is assigned to IPP molecule bound at S1 site, and IPP2 to that at S2 site. The oxygen atoms are colored red and phosphorus atoms are colored orange. The carbon atoms of IPP1 are colored salmon and those of IPP2 are in green.

(B) Detailed view of the -RXG- motif and the active site. The carbon atoms of NgBR -RXG- motif residues are colored orange and labeled, and those for DHDDS are colored cyan. Nitrogen atoms are colored blue and oxygen atoms are colored red. IPP1 and IPP2 are colored red and green, respectively.  $Mg^{2+}$  is shown as a gray sphere, and its co-ordination is indicated by the dashed lines. A coordinating water molecule is shown as red sphere.

(C) The K42E retinitis pigmentosa mutation in DHDDS. A cartoon representation indicating the locations of Lys-42 and Glu-234 relative to the P-loop and bound substrates. DHDDS is colored in deep teal, IPP1 in red, IPP2 in green and  $Mg^{2+}$  is shown as a gray sphere.

**Figure 4. Missense NgBR mutations associated with Parkinson's disease.**

(A) NgBR mutations related to Parkinson's disease are shown as purple spheres and labeled. DHDDS is colored in gray and NgBR is colored in orange except for the N-terminal helix ( $\alpha 1$ ) which is shown in green.

(B) Detailed view showing the location of G91C disease mutant within NgBR colored orange except for  $\alpha 1$  helix, shown in green. Gly-91 is involved in hydrophobic packing against Val-127. Lys-96 forms hydrogen bonds with Tyr-248 and Gln-289, which stabilize the C-terminal -RXG- motif.

(C) *cis*-PTase activity was measured using purified wildtype and NgBR disease mutant, G91C. The mutant exhibits ~ 40% reduction in *cis*-PTase activity compared to wildtype enzyme. The values are the mean  $\pm$ S.D. of three independent measurements.

(D) The majority of the mutated residues is solvent-exposed. The solvent accessible surface area ( $\text{\AA}^2$ ) were calculated using GETAREA server.

**Figure 5. DHDDS's helix  $\alpha 0$  functions as a membrane sensor.**

(A) A cartoon representation of DHDDS subunit colored in deep teal. The hydrophobic cavities of DHDDS before (left) and after (right) N-terminal loop deletion (residues 1-10) are shown in yellow. The N-terminal loop and  $\alpha 0$  helix are shown in red; the side chains of the three exposed hydrophobic residues are shown and labeled. The hydrophobic cavity was generated using the 3V web server (54).

(B) A cartoon representation of *E.coli* UPPs (PDB entry 1X06) monomer colored in purple. The hydrophobic cavity is shown in yellow. Leu-137 involved in chain length control in UPPS is located at the end of the cavity.

(C) Sequence alignment showing the conservation of three hydrophobic amino acids at the N-terminal helix  $\alpha 0$  among DHDDS orthologs (highlighted in red). Conserved residues are highlighted in yellow. Proteins represented in this alignment are orthologs of human DHDDS *cis*-PTase subunit as follows: hDHDDS (human, UniProtKB Q86SQ9-1), XlDhdds (*Xenopus laevis*; UniProtKB Q7ZYJ5), DrDhdds (*Danio rerio*, UniProtKB Q6NXA2), CeDHDDS (*Caenorhabditis elegans*, UniProtKB Q5FC21), ScRer2 (*Saccharomyces cerevisiae*, UniProtKB P35196), ScSrt1 (*Saccharomyces cerevisiae*, UniProtKB Q03175), SpRer2 (*Schizosaccharomyces pombe*, UniProtKB O14171), TrRER2 (*Trichoderma reesei*, UniProtKB - G0ZKV6), AfRer2 (*Aspergillus fumigatus* UniProtKB - Q4WQ28 ), SICPT3 (*Solanum*

*lycopersicum*, UniProtKB K7WCI9), AtCPT3 (*Arabidopsis thaliana*, UniProtKB Q8S2T1), AtCPT4 (*Arabidopsis thaliana*, UniProtKB Q8LAR7), AtCPT5 (*Arabidopsis thaliana*, UniProtKB Q8LED0)

(D) Phospholipid stimulation of wildtype and W12A/F15A/I19A triple mutant is shown. Stimulation was compared by measuring *cis*-PTase activity of purified wildtype and DHDDS triple mutant in the presence and absence of 1% phosphatidylinositol (PI). The values are the mean  $\pm$ S.D. of five to eight independent measurements.

(E) Reverse phase TLC separation of dephosphorylated products from *cis*-PTase activity of WT and DHDDS triple mutation (W12A/F15A/I19A) denoted as MUT. Numbers correspond to the dominant polyprenols in each sample are shown at the bottom of the plate. The position of the polyprenol standards is shown on the left.

### **Figure 6. Mechanism of chain elongation by heteromeric *cis*-PTases.**

(A) Multiple amino acid sequence alignment comparing  $\alpha$ 3 helix between DHDDS orthologs and homodimeric *cis*-PTases. Highly conserved residues are highlighted in red and less conserved ones are shown in yellow. Region corresponding to DHDDS EKE insert is highlighted in blue; the region is missing in homodimeric enzymes and a gap is present instead. Proteins represented in this alignment are: single subunit *cis*-PTs: EcUPPS (*Escherichia coli*, UniProtKB P60472), MIDPPS (*Micrococcus luteus*, UniProtKB O82827), SaUPPS (*Sulfolobus acidocaldarius*, UniProtKB Q9HH76). Orthologues of human DHDDS *cis*-PTase subunit: hDHDDS (human, UniProtKB Q86SQ9-1), XIDhdds (*Xenopus laevis*, UniProtKB Q7ZYJ5), DrDhdds (*Danio rerio*, UniProtKB Q6NXA2), CeDHDDS (*Caenorhabditis elegans*, UniProtKB Q5FC21), ScRer2 (*Saccharomyces cerevisiae*, UniProtKB P35196), SpRer2 (*Schizosaccharomyces pombe*, UniProtKB O14171), SlCPT3 (*Solanum lycopersicum*, UniProtKB K7WCI9), AtCPT3 (*Arabidopsis thaliana*, UniProtKB Q8S2T1), AtCPT4 (*Arabidopsis thaliana*, UniProtKB Q8LAR7), AtCPT5 (*Arabidopsis thaliana*, UniProtKB Q8LED0)

(B) HPLC analysis of chain length of dolichol generated by WT and  $\Delta$ EKE DHDDS mutant in yeast cells. The dolichol peaks were identified and labeled on top of the chromatogram. The WT cells yielded the main compound Dol-20 compared to Dol-17 in  $\Delta$ EKE mutant. The chain length and identity of lipids were confirmed by comparison with external standards of a polyprenol (Pren-10 – Pren-24) and dolichol (Dol-17 – Dol-23) mixtures.

(C) Structural comparison between *E. coli* UPPS (PDB ID 1X06) monomer and human DHDDS subunit. The EKE insert within  $\alpha$ 3 helix of DHDDS creates a bigger bulge that may contribute to the stabilization of the enzyme:product complex during chain elongation.

(D) Schematic diagram illustrating the proposed chain elongation mechanism for *cis*-PTases. Red arrow indicates the direction of product elongation. Exposed hydrophobic isoprene units may increasingly destabilize the enzyme:product complex by interacting with detergent micells (blue) or lipid bilayers (grey).

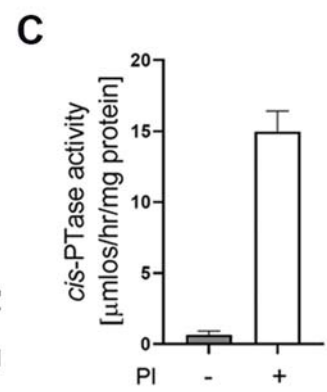
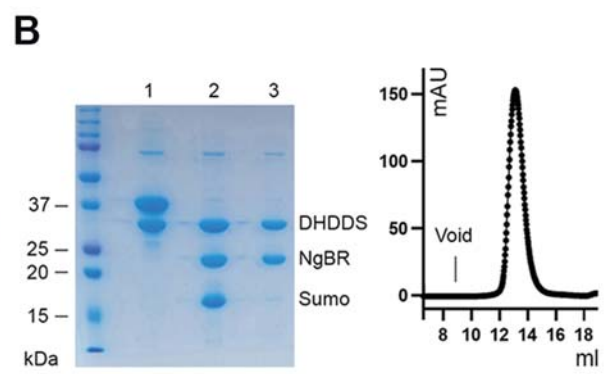
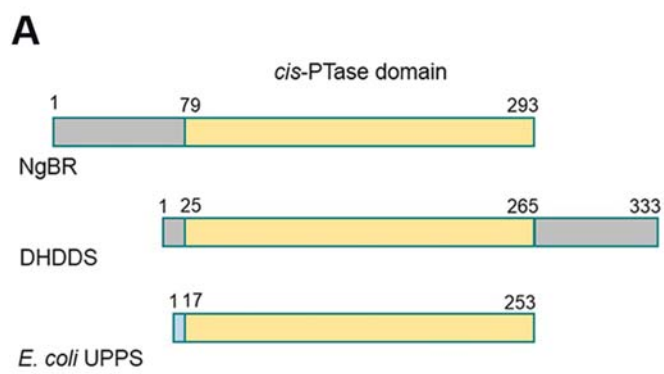
## References

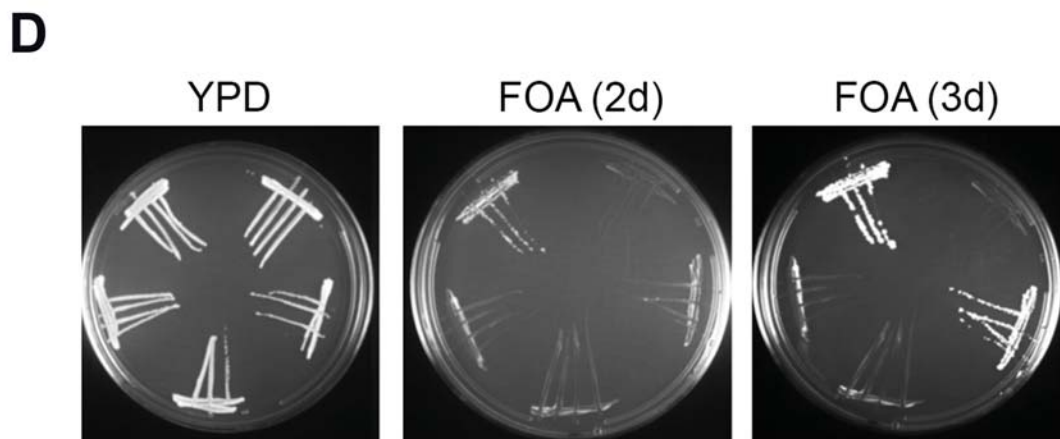
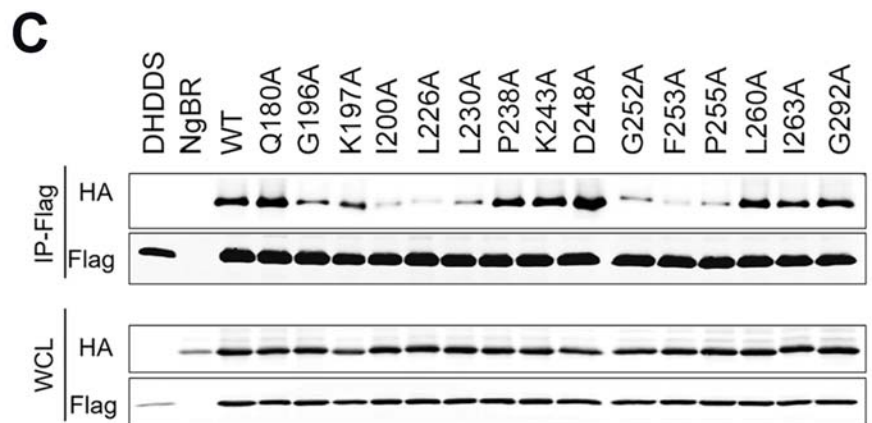
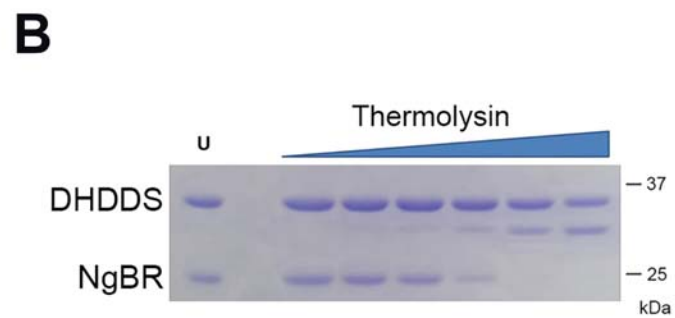
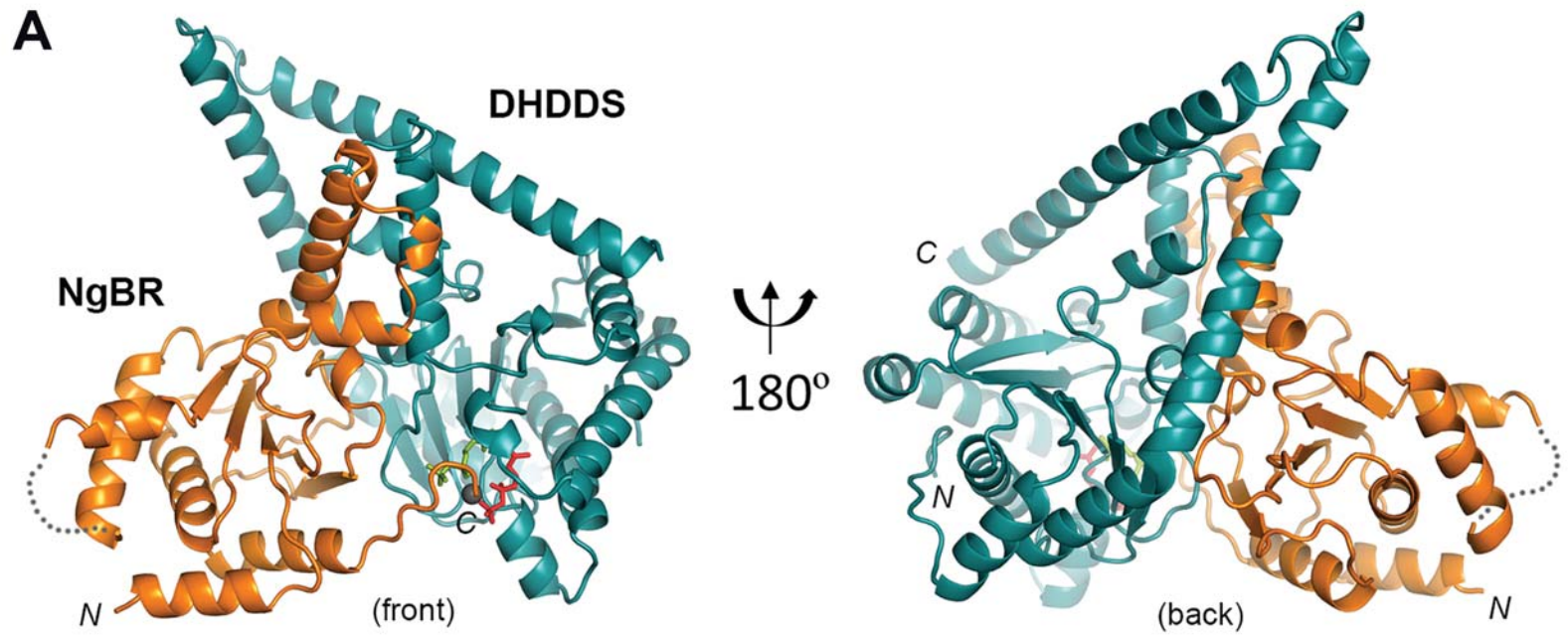
1. Takahashi S & Koyama T (2006) Structure and function of cis-prenyl chain elongating enzymes. *Chem Rec* 6(4):194-205.
2. Shimizu N, Koyama T, & Ogura K (1998) Molecular cloning, expression, and purification of undecaprenyl diphosphate synthase. No sequence similarity between E- and Z-prenyl diphosphate synthases. *J Biol Chem* 273(31):19476-19481.
3. Apfel CM, Takacs B, Fountoulakis M, Stieger M, & Keck W (1999) Use of genomics to identify bacterial undecaprenyl pyrophosphate synthetase: cloning, expression, and characterization of the essential uppS gene. *J Bacteriol* 181(2):483-492.
4. Sato M, *et al.* (1999) The yeast RER2 gene, identified by endoplasmic reticulum protein localization mutations, encodes cis-prenyltransferase, a key enzyme in dolichol synthesis. *Mol Cell Biol* 19(1):471-483.
5. Endo S, Zhang YW, Takahashi S, & Koyama T (2003) Identification of human dehydrodolichyl diphosphate synthase gene. *Biochim Biophys Acta* 1625(3):291-295.
6. Shridas P, Rush JS, & Waechter CJ (2003) Identification and characterization of a cDNA encoding a long-chain cis-isoprenyltransferase involved in dolichyl monophosphate biosynthesis in the ER of brain cells. *Biochem Biophys Res Commun* 312(4):1349-1356.
7. Cunillera N, Arro M, Fores O, Manzano D, & Ferrer A (2000) Characterization of dehydrodolichyl diphosphate synthase of *Arabidopsis thaliana*, a key enzyme in dolichol biosynthesis. *FEBS Lett* 477(3):170-174.
8. Oh SK, Han KH, Ryu SB, & Kang H (2000) Molecular cloning, expression, and functional analysis of a cis-prenyltransferase from *Arabidopsis thaliana*. Implications in rubber biosynthesis. *J Biol Chem* 275(24):18482-18488.
9. Buczkowska A, Swiezewska E, & Lefeber DJ (2015) Genetic defects in dolichol metabolism. *J Inherit Metab Dis* 38(1):157-169.
10. Grabinska KA, Park EJ, & Sessa WC (2016) cis-Prenyltransferase: New Insights into Protein Glycosylation, Rubber Synthesis, and Human Diseases. *J Biol Chem* 291(35):18582-18590.
11. Hartley MD & Imperiali B (2012) At the membrane frontier: a prospectus on the remarkable evolutionary conservation of polyprenols and polyprenyl-phosphates. *Arch Biochem Biophys* 517(2):83-97.
12. Cherian S, Ryu SB, & Cornish K (2019) Natural rubber biosynthesis in plants, the rubber transferase complex, and metabolic engineering progress and prospects. *Plant Biotechnol J* 17(11):2041-2061.
13. Schulbach MC, Brennan PJ, & Crick DC (2000) Identification of a short (C15) chain Z-isoprenyl diphosphate synthase and a homologous long (C50) chain isoprenyl diphosphate synthase in *Mycobacterium tuberculosis*. *J Biol Chem* 275(30):22876-22881.
14. Kaur D, Brennan PJ, & Crick DC (2004) Decaprenyl diphosphate synthesis in *Mycobacterium tuberculosis*. *J Bacteriol* 186(22):7564-7570.
15. Akhtar TA, *et al.* (2013) The tomato cis-prenyltransferase gene family. *Plant J* 73(4):640-652.
16. Akhtar TA, *et al.* (2017) Polyprenols Are Synthesized by a Plastidial cis-Prenyltransferase and Influence Photosynthetic Performance. *Plant Cell* 29(7):1709-1725.
17. Teng KH & Liang PH (2012) Structures, mechanisms and inhibitors of undecaprenyl diphosphate synthase: a cis-prenyltransferase for bacterial peptidoglycan biosynthesis. *Bioorg Chem* 43:51-57.
18. Fujihashi M, *et al.* (2001) Crystal structure of cis-prenyl chain elongating enzyme, undecaprenyl diphosphate synthase. *Proc Natl Acad Sci U S A* 98(8):4337-4342.
19. Ko TP, *et al.* (2001) Mechanism of product chain length determination and the role of a flexible loop in *Escherichia coli* undecaprenyl-pyrophosphate synthase catalysis. *J Biol Chem* 276(50):47474-47482.
20. Guo RT, *et al.* (2007) Bisphosphonates target multiple sites in both cis- and trans-prenyltransferases. *Proc Natl Acad Sci U S A* 104(24):10022-10027.



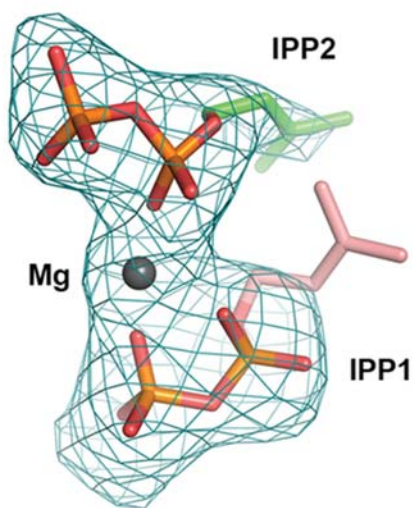
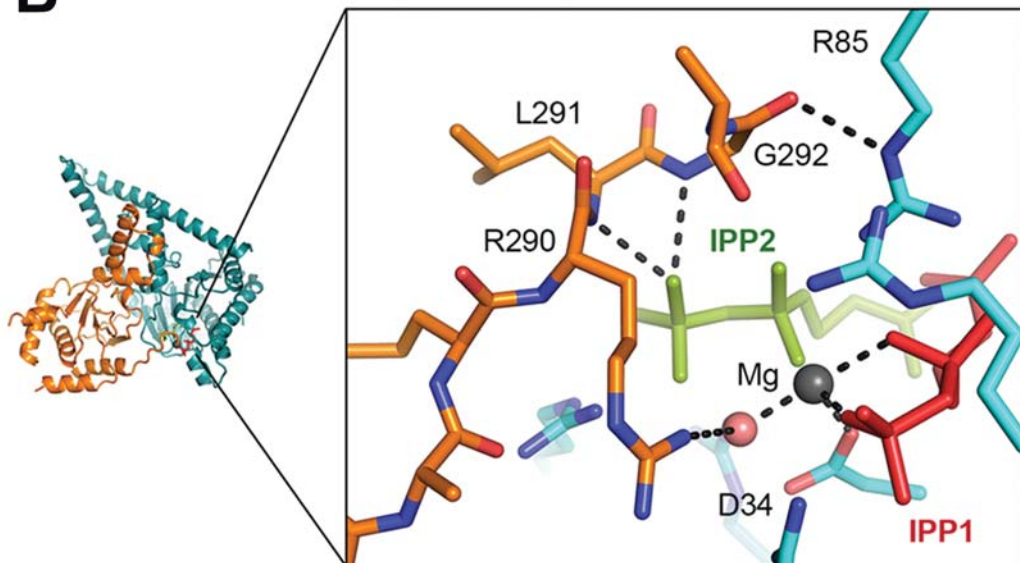
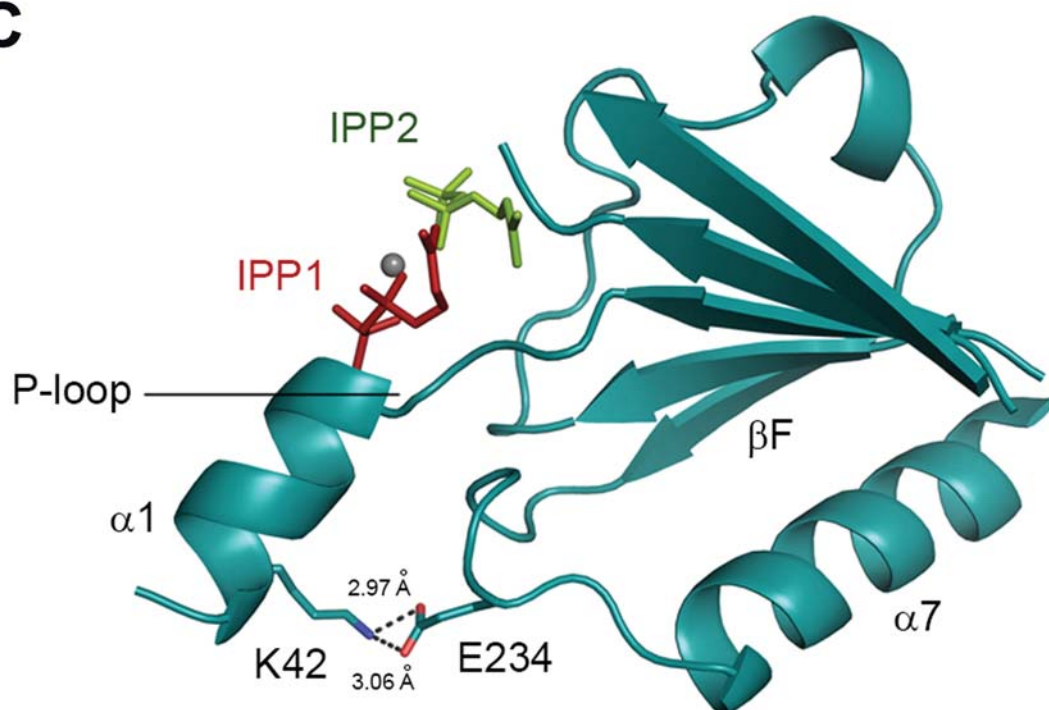
21. Wang W, *et al.* (2008) The structural basis of chain length control in Rv1086. *J Mol Biol* 381(1):129-140.
22. Chan YT, *et al.* (2017) Crystal Structure and Potential Head-to-Middle Condensation Function of a Z,Z-Farnesyl Diphosphate Synthase. *ACS Omega* 2(3):930-936.
23. Lisnyansky Bar-El M, *et al.* (2019) Structural Characterization of Full-Length Human Dehydrodolichyl Diphosphate Synthase Using an Integrative Computational and Experimental Approach. *Biomolecules* 9(11).
24. Harrison KD, *et al.* (2011) Nogo-B receptor is necessary for cellular dolichol biosynthesis and protein N-glycosylation. *EMBO J* 30(12):2490-2500.
25. Park EJ, *et al.* (2014) Mutation of Nogo-B receptor, a subunit of cis-prenyltransferase, causes a congenital disorder of glycosylation. *Cell Metab* 20(3):448-457.
26. Brasher MI, *et al.* (2015) A two-component enzyme complex is required for dolichol biosynthesis in tomato. *Plant J* 82(6):903-914.
27. Epping J, *et al.* (2015) A rubber transferase activator is necessary for natural rubber biosynthesis in dandelion. *Nature Plants* 1(5):15048.
28. Kwon M, Kwon EJ, & Ro DK (2016) cis-Prenyltransferase and Polymer Analysis from a Natural Rubber Perspective. *Methods Enzymol* 576:121-145.
29. Qu Y, *et al.* (2015) A lettuce (*Lactuca sativa*) homolog of human Nogo-B receptor interacts with cis-prenyltransferase and is necessary for natural rubber biosynthesis. *J Biol Chem* 290(4):1898-1914.
30. Miao RQ, *et al.* (2006) Identification of a receptor necessary for Nogo-B stimulated chemotaxis and morphogenesis of endothelial cells. *Proc Natl Acad Sci U S A* 103(29):10997-11002.
31. Yu L, Pena Castillo L, Mnaimneh S, Hughes TR, & Brown GW (2006) A survey of essential gene function in the yeast cell division cycle. *Mol Biol Cell* 17(11):4736-4747.
32. Ma J, *et al.* (2019) Structural insights to heterodimeric cis-prenyltransferases through yeast dehydrodolichyl diphosphate synthase subunit Nus1. *Biochem Biophys Res Commun* 515(4):621-626.
33. Grabinska KA, Edani BH, Park EJ, Kraehling JR, & Sessa WC (2017) A conserved C-terminal RXG motif in the NgBR subunit of cis-prenyltransferase is critical for prenyltransferase activity. *J Biol Chem* 292(42):17351-17361.
34. Yamashita S, *et al.* (2016) Identification and reconstitution of the rubber biosynthetic machinery on rubber particles from *Hevea brasiliensis*. *Elife* 5.
35. Lakusta AM, *et al.* (2019) Molecular Studies of the Protein Complexes Involving Cis-Prenyltransferase in Guayule (*Parthenium argentatum*), an Alternative Rubber-Producing Plant. *Front Plant Sci* 10:165.
36. Emi KI, Sompiyachoke K, Okada M, & Hemmi H (2019) A heteromeric cis-prenyltransferase is responsible for the biosynthesis of glycosyl carrier lipids in *Methanosarcina mazei*. *Biochem Biophys Res Commun* 520(2):291-296.
37. Zhang H, *et al.* (2008) Dolichol biosynthesis and its effects on the unfolded protein response and abiotic stress resistance in *Arabidopsis*. *Plant Cell* 20(7):1879-1898.
38. Den K, *et al.* (2019) Recurrent NUS1 canonical splice donor site mutation in two unrelated individuals with epilepsy, myoclonus, ataxia and scoliosis - a case report. *BMC Neurol* 19(1):253.
39. Hamdan FF, *et al.* (2017) High Rate of Recurrent De Novo Mutations in Developmental and Epileptic Encephalopathies. *Am J Hum Genet* 101(5):664-685.
40. Sabry S, *et al.* (2016) A case of fatal Type I congenital disorders of glycosylation (CDG I) associated with low dehydrodolichol diphosphate synthase (DHDDS) activity. *Orphanet J Rare Dis* 11(1):84.
41. Zelinger L, *et al.* (2011) A missense mutation in DHDDS, encoding dehydrodolichyl diphosphate synthase, is associated with autosomal-recessive retinitis pigmentosa in Ashkenazi Jews. *Am J Hum Genet* 88(2):207-215.

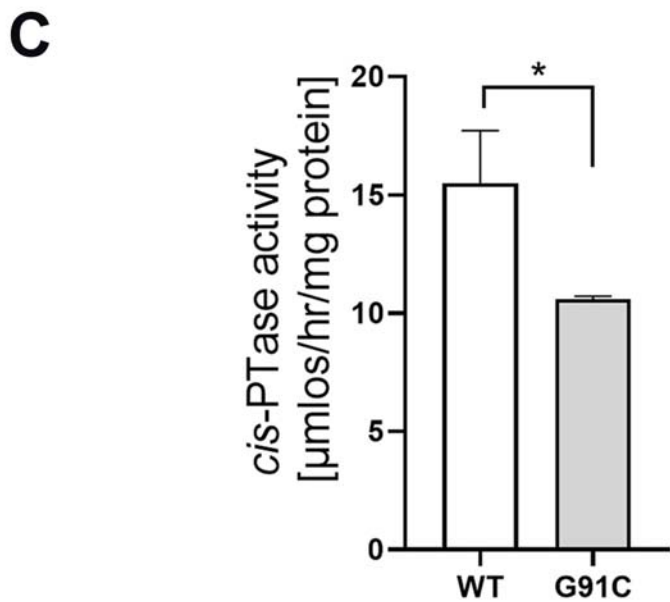
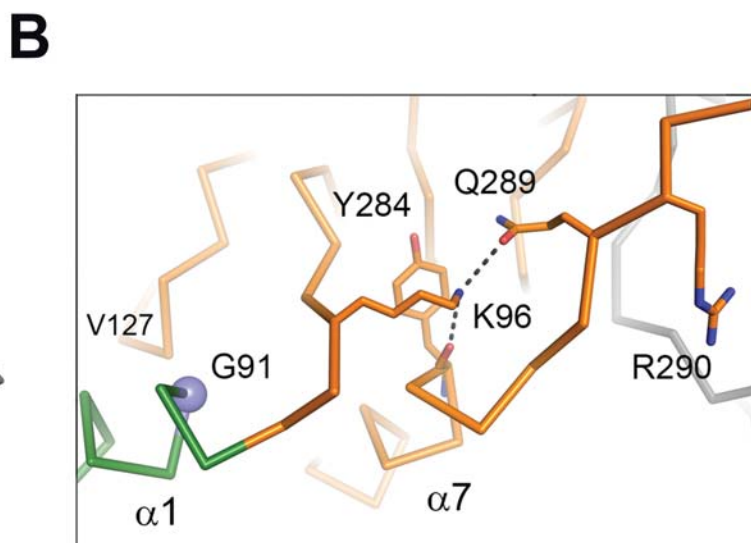
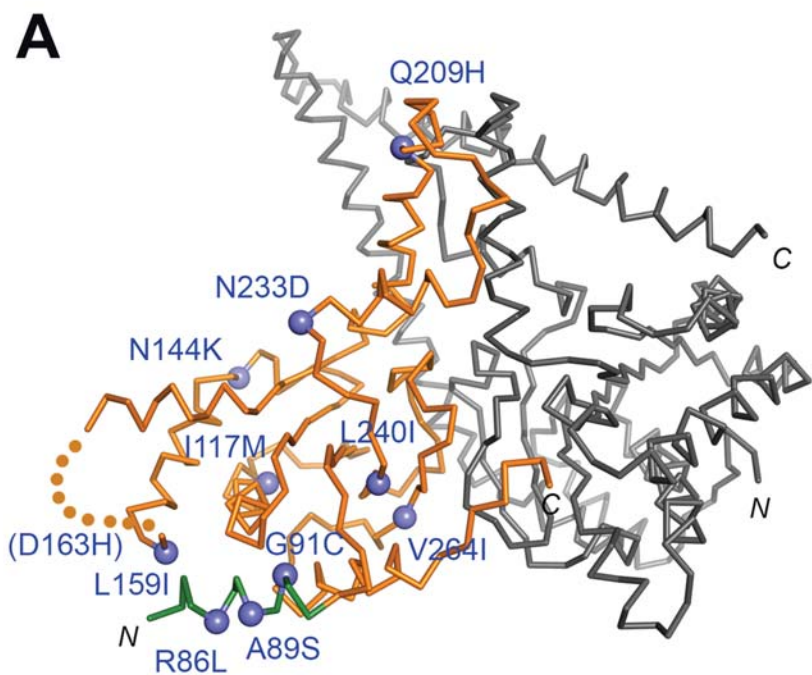
42. Zuchner S, *et al.* (2011) Whole-exome sequencing links a variant in DHDDS to retinitis pigmentosa. *Am J Hum Genet* 88(2):201-206.
43. Gao J, *et al.* (2018) "Head-to-Middle" and "Head-to-Tail" cis-Prenyl Transferases: Structure of Isosesquilandulyl Diphosphate Synthase. *Angew Chem Int Ed Engl* 57(3):683-687.
44. Ko TP, *et al.* (2019) Substrate-analogue complex structure of Mycobacterium tuberculosis decaprenyl diphosphate synthase. *Acta Crystallogr F Struct Biol Commun* 75(Pt 4):212-216.
45. Guo JF, *et al.* (2018) Coding mutations in NUS1 contribute to Parkinson's disease. *Proc Natl Acad Sci U S A* 115(45):11567-11572.
46. Holcomb J, *et al.* (2018) SAXS analysis of a soluble cytosolic NgBR construct including extracellular and transmembrane domains. *PLoS One* 13(1):e0191371.
47. Guo RT, *et al.* (2005) Crystal structures of undecaprenyl pyrophosphate synthase in complex with magnesium, isopentenyl pyrophosphate, and farnesyl thiopyrophosphate: roles of the metal ion and conserved residues in catalysis. *J Biol Chem* 280(21):20762-20774.
48. Zhao B, *et al.* (2017) The Nogo-B receptor promotes Ras plasma membrane localization and activation. *Oncogene* 36(24):3406-3416.
49. Wen R, Lam BL, & Guan Z (2013) Aberrant dolichol chain lengths as biomarkers for retinitis pigmentosa caused by impaired dolichol biosynthesis. *J Lipid Res* 54(12):3516-3522.
50. Kimchi A, *et al.* (2018) Nonsyndromic Retinitis Pigmentosa in the Ashkenazi Jewish Population: Genetic and Clinical Aspects. *Ophthalmology* 125(5):725-734.
51. Fedorow H, *et al.* (2005) Dolichol is the major lipid component of human substantia nigra neuromelanin. *J Neurochem* 92(4):990-995.
52. Ward WC, *et al.* (2007) Identification and quantification of dolichol and dolichoic acid in neuromelanin from substantia nigra of the human brain. *J Lipid Res* 48(7):1457-1462.
53. Hoffmann R, Grabinska K, Guan Z, Sessa WC, & Neiman AM (2017) Long-Chain Polyprenols Promote Spore Wall Formation in *Saccharomyces cerevisiae*. *Genetics* 207(4):1371-1386.
54. Voss NR & Gerstein M (2010) 3V: cavity, channel and cleft volume calculator and extractor. *Nucleic acids research* 38(Web Server issue):W555-562.
55. Chang SY, Ko TP, Liang PH, & Wang AH (2003) Catalytic mechanism revealed by the crystal structure of undecaprenyl pyrophosphate synthase in complex with sulfate, magnesium, and triton. *J Biol Chem* 278(31):29298-29307.
56. Otwinowski Z & Minor W (1997) [20] Processing of X-ray diffraction data collected in oscillation mode. *Methods in enzymology* 276:307-326.
57. Winn MD, *et al.* (2011) Overview of the CCP4 suite and current developments. *Acta crystallographica. Section D, Biological crystallography* 67(Pt 4):235-242.
58. Emsley P, Lohkamp B, Scott WG, & Cowtan K (2010) Features and development of Coot. *Acta crystallographica. Section D, Biological crystallography* 66(Pt 4):486-501.





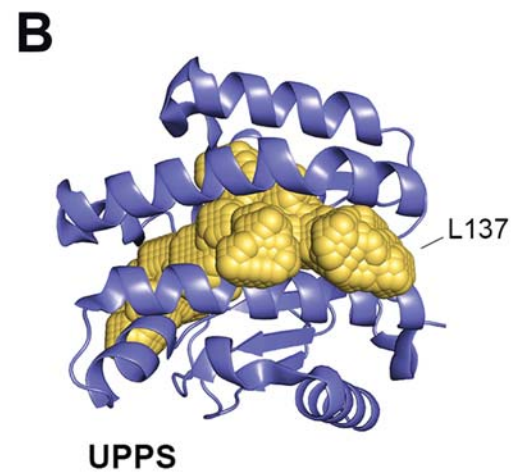
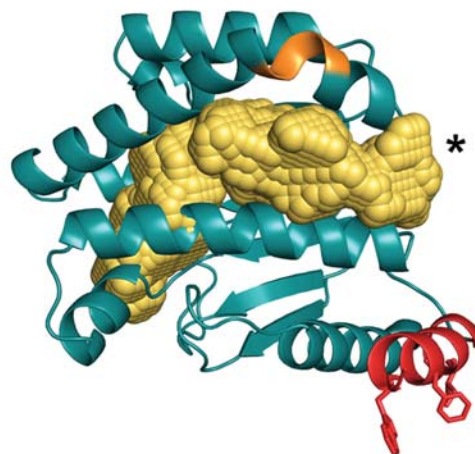
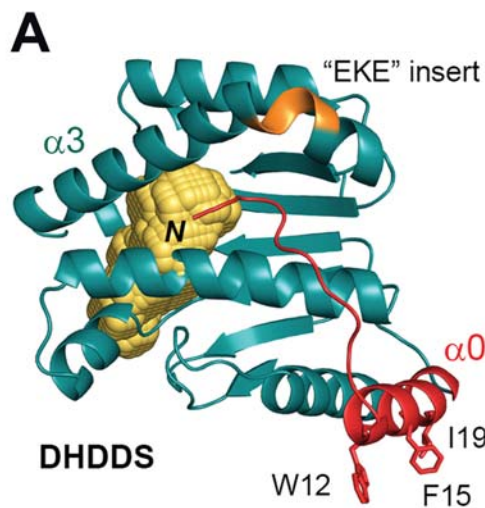


**A****B****C**



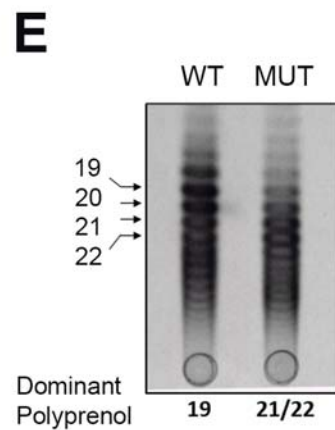
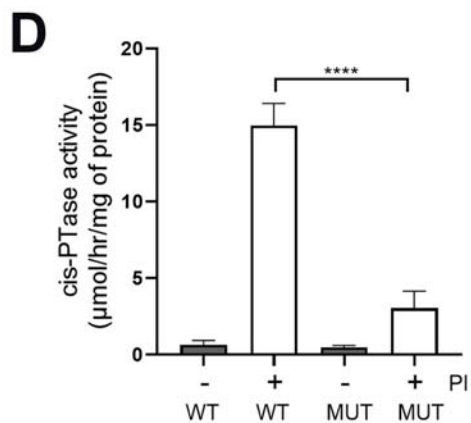
**D**

Mutation	Solvent accessible surface area ( $\text{\AA}^2$ )
R86L	49.20
A89S	60.29
G91C*	0.90
I117M	0.00
N144K	49.42
L159I	80.98
D163H	-
Q209H	92.03
N233D	108.94
L240I	0.00
V264I	11.00



**C**

		$\alpha 0$									
hDHDDS	9	LSL	WER	FCANI	I	KAGP					
XlDhdds	9	LSL	I	ERLCAN	VLKAGP						
DrDhdds	9	ESL	L	ERLTANI	LKAGP						
SpRer2	18	LDQL	ET	MMI	NTI	KRGK					
AfRer2	24	I	AKL	RELLI	GALKQGP						
TRRER2	22	VNQL	RE	LLI	GALKQGP						
ScSrt1	50	YVNL	QNI	LI	KALRVGP						
AtCPT4	19	MSL	MRKFI	FKVLRVGP							
AtCPT5	18	MSL	MRKFL	FRVLCVGP							
SlCPT3	17	SSF	V	RQCI	FSI	LSVGP					
CeDHDDS	20	QPWW	QWLL	RRFI	ASGP						
ScRer2	14	L	KWT	KNIF	SRTL	RASN					
AtCPT3	23	YGF	S	RRSL	FRVI	SMGP					

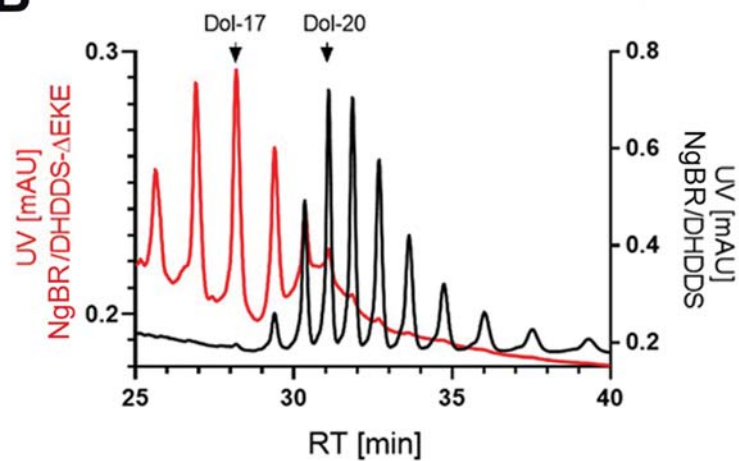
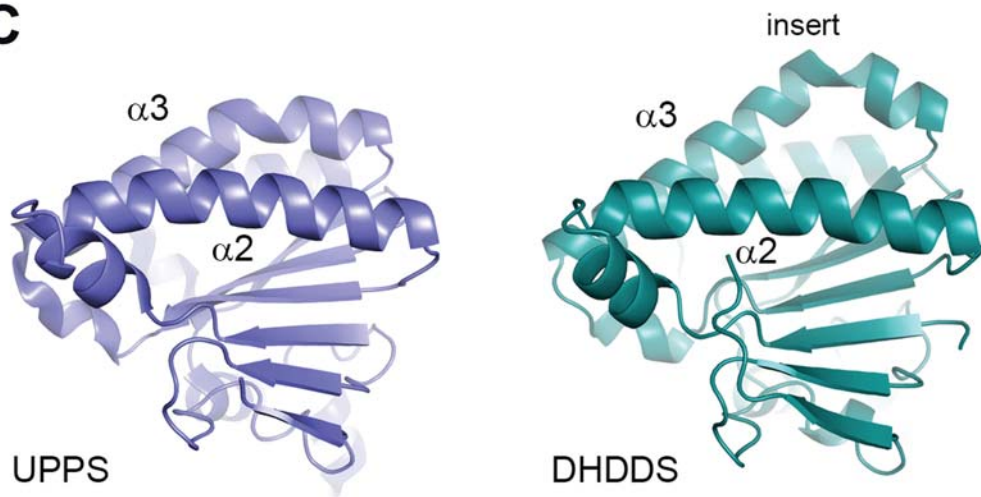




**A**

		$\alpha 3$																																			
hDHDDS	84	K	R	S	K	S	E	V	D	G	L	M	D	L	A	R	Q	K	F	S	R	L	M	E	E	K	E	K	L	Q	K	H	G	V	C	I	R
XlDhdds	84	K	R	S	A	E	E	V	E	G	L	L	E	L	A	R	Q	K	F	T	R	L	L	E	E	K	D	K	L	E	K	H	G	V	R	I	R
DrDhdds	84	K	R	S	E	E	V	K	D	L	L	D	L	A	R	E	K	F	L	N	I	I	N	D	E	A	K	L	N	E	H	G	V	R	I	L	
CeDHDDS	96	K	R	S	E	E	V	S	G	L	M	R	L	A	E	E	K	F	Q	K	L	L	N	D	S	E	K	L	D	E	K	R	I	C	F	R	
ScRer2	90	K	R	S	S	R	E	V	E	S	L	M	T	L	A	R	E	R	I	R	Q	I	T	E	R	G	E	L	A	C	K	Y	G	V	R	I	K
AtCPT3	98	R	R	K	P	E	E	V	E	S	V	M	D	L	M	L	E	K	I	K	S	L	L	E	K	E	S	I	V	H	Q	Y	G	I	R	V	Y
AtCPT4	95	K	R	G	P	E	E	V	K	C	V	M	D	L	M	L	E	K	V	E	L	T	I	D	Q	A	V	S	G	N	M	N	G	V	R	I	I
AtCPT5	94	K	R	G	P	E	E	V	K	C	V	M	D	L	M	L	E	K	V	E	L	A	I	D	Q	A	V	S	G	N	M	N	G	V	R	I	I
SlCPT3	92	K	R	R	P	E	E	V	S	L	M	K	L	M	Q	E	K	I	D	E	L	T	K	E	E	S	I	V	N	R	L	G	I	R	I	Y	
SpRer2	93	K	R	S	K	Y	E	V	D	M	L	M	E	I	A	K	N	S	L	T	Q	I	T	A	H	G	D	L	V	D	Q	Y	G	I	R	I	R
MlDPPS	79	S	R	P	K	D	E	V	N	Y	L	M	K	L	P	G	D	F	L	N	T	F	L	P	-	-	-	E	L	I	E	K	N	V	K	V	E
SaUPPS	83	S	R	P	E	S	E	V	N	Y	I	M	N	L	P	V	N	F	L	K	T	F	L	P	-	-	-	E	L	I	E	K	N	V	K	V	E
EcUPPS	76	N	R	P	A	Q	E	V	S	A	L	M	E	L	F	V	W	A	L	D	S	E	V	K	-	-	-	S	L	H	R	H	N	V	R	L	R

insert

**B****C****D**