

# Integrative modelling of the full-length human dehydrololichyl diphosphate synthase using a hybrid computational and experimental approach

Michal Lisnyansky Barel<sup>1</sup>, Su Youn Lee<sup>2</sup>, Ah Young Ki<sup>2</sup>, Noa Kapelushnik<sup>3</sup>, Anat Loewenstein<sup>4</sup>, Ka Young Chung<sup>2</sup>, Dina Schneidman-Duhovny<sup>5,6</sup>, Moshe Giladi<sup>1,7,¶\*</sup>, Hadas Newman<sup>4,¶</sup>, Yoni Haitin<sup>1,¶</sup>,

<sup>1</sup> Department of Physiology and Pharmacology, Sackler Faculty of Medicine, Tel-Aviv University, Tel-Aviv, Israel

<sup>2</sup> School of Pharmacy, Sungkyunkwan University, Jangan-gu, Suwon, South Korea

<sup>3</sup> Department of Ophthalmology, Sheba Medical Center, Ramat Gan, Israel

<sup>4</sup> Department of Ophthalmology, Tel-Aviv Sourasky Medical Center, Tel Aviv, Israel

<sup>5</sup> Department of Biological Chemistry, Institute of Life Sciences, The Hebrew University of Jerusalem, Jerusalem, Israel

<sup>6</sup> School of Computer Science and Engineering, The Hebrew University of Jerusalem, Jerusalem, Israel

<sup>7</sup> Tel Aviv Sourasky Medical Center, Tel Aviv, Israel

¶ These authors contributed equally to this work.

\* Correspondence: [moshegil@post.tau.ac.il](mailto:moshegil@post.tau.ac.il) (MG), [yhaitin@tauex.tau.ac.il](mailto:yhaitin@tauex.tau.ac.il) (YH)

**Abstract:** Dehydrololichyl diphosphate synthase (DHDDS) and Nogo-B receptor (NgBR) form the heteromeric human *cis*-prenyltransferase complex, synthesizing the precursor for the glycosyl carrier involved in N-linked protein glycosylation. In line with the important role of N-glycosylation in protein biogenesis, mutations in DHDDS, the catalytic subunit of the complex, were shown to result in human diseases. Importantly, well-characterized DHDDS homologs function as homodimers and not as heteromeric complexes. Moreover, DHDDS encompasses a C-terminal region, which does not converge with any known conserved domains. Therefore, despite the clinical importance of DHDDS, our understanding of its structure-function relations remains poor. Here, we provide a structural model for the full-length human DHDDS using a multidisciplinary experimental and computational approach. Our model suggests that the C-terminal domain of DHDDS forms a helix-loop-helix motif, tightly packed against the core catalytic *cis*-prenyltransferase domain. This model is consistent with small-angle X-ray scattering data, indicating that the full-length DHDDS maintains a similar conformation in solution. Moreover, hydrogen-deuterium exchange mass-spectrometry experiments show time-dependent deuterium uptake in the C-terminal domain, consistent with its overall folded state. Finally, we provide a model for the DHDDS-NgBR heterodimer, offering a structural framework for future structural and functional studies of the human *cis*-prenyltransferase complex.

**Keywords:** *cis*-prenyltransferase, computational modelling, hydrogen-deuterium exchange mass-spectrometry, small-angle X-ray scattering, enzyme kinetics, DHDDS

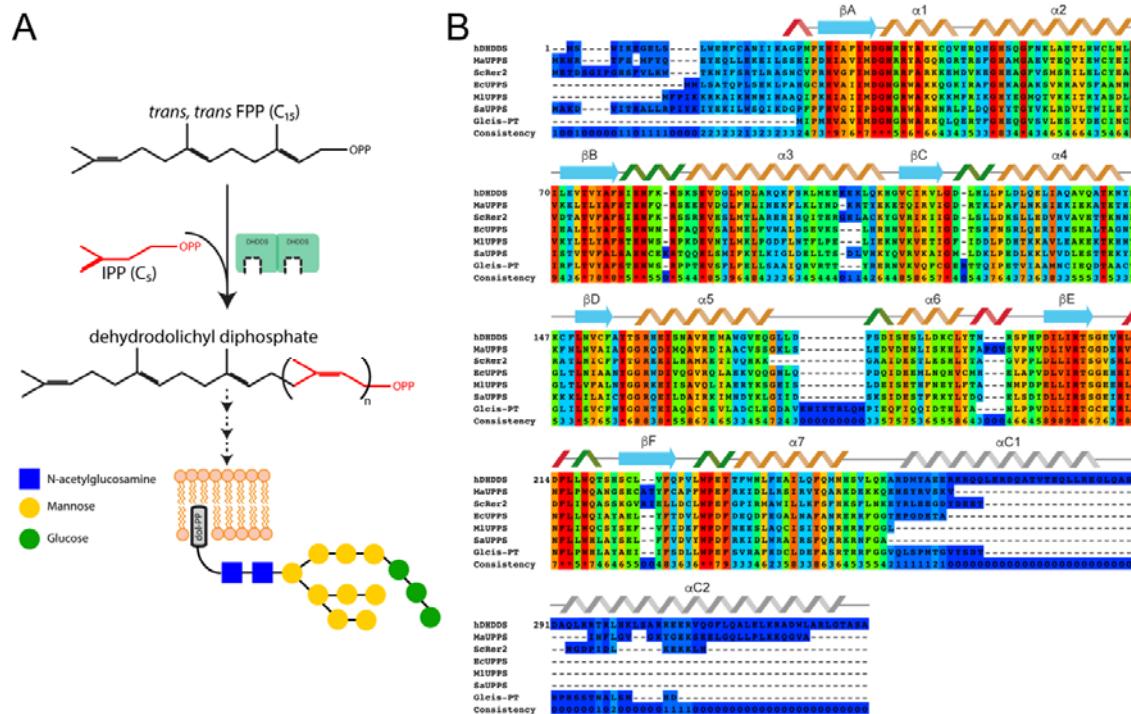
## 1. Introduction

Protein N-glycosylation is a critical post-translational modification, in which an oligosaccharide moiety is transferred from the glycosyl carrier dolichol-phosphate (Dol-P) and covalently attaches to a polypeptide chain via an asparagine residue [1]. Following protein conjugation, the oligosaccharide is further modified, increasing the functional and structural diversity of the proteome [2]. Indeed, this modification is crucial for proper protein folding, oligomerization, quality control, sorting, and transport [2]. Accordingly, even slight variations in protein glycosylation can markedly affect proteins structure and function [3] and thus, glycosylation disorders result in a wide range of clinical syndromes affecting practically every organ system [4,5].

Dehydrololichyl diphosphate synthase (DHDDS) together with Nogo-B receptor (NgBR) form the human *cis*-prenyltransferase complex [6,7]. This complex catalyzes the formation of dehydrololichyl diphosphate (DHDD), a long-chain polypropenyl serving as a precursor for Dol-P. DHDD is synthesized by

chain elongation of farnesyl diphosphate (FPP) via multiple condensations with isopentenyl diphosphate (IPP) (Figure 1A) [6]. Importantly, while previous studies revealed that DHDDS is required and sufficient for this catalytic activity, NgBR induces an increase in the expression and activity of the complex but exhibits no catalytic activity [6–8]. In line with the important role of DHDDS in protein glycosylation, mutations in DHDDS were recently shown to cause autosomal recessive retinitis pigmentosa [9,10], a fatal glycosylation disorder [11], and developmental epileptic encephalopathies [12].

Despite the clinical importance of DHDDS, it was not structurally and functionally characterized to date. Conserved domains search analysis [13] reveals that the *cis*-prenyltransferase catalytic domain of DHDDS (residues 24-248) is homologous to undecaprenyl pyrophosphate synthase (UPPS), a well characterized bacterial *cis*-prenyltransferase [14]. Consisting of 7  $\alpha$ -helices and 6  $\beta$ -strands, this domain forms dimers, with a hydrophobic catalytic tunnel surrounded by 2  $\alpha$ -helices and 4  $\beta$ -strands within each monomer [15]. In contrast with the N-terminal *cis*-prenyltransferase homology domain, the C-terminus of human DHDDS does not converge with any known conserved domains. Intriguingly, despite the low level of conservation, a C-terminal domain is common to long-chain *cis*-prenyltransferases and is absent in short and medium chain *cis*-prenyltransferases [7] (Figure 1B). Nevertheless, its structure and functional role remains elusive.



**Figure 1. Catalytic cycle of human DHDDS. (A)** Multiple condensations of an FPP primer with IPP result in the formation DHDD by DHDDS. DHDD is further processed to form Dol-P, the carrier of the oligosaccharide used for N-glycosylation. **(B)** Sequence alignment of *cis*-prenyltransferases, with conservation scores calculated by PRALINE. Human DHDDS (hDHDDS), *M. acetivorans* UPPS (MaUPPS), *S. cerevisiae* Rer2 (ScRer2) are components of heteromeric *cis*-prenyltransferase complexes; *E. coli* UPPS (EcUPPS), *M. luteus* UPPS (MiUPPS), *S. acidocaldarius* UPPS (SaUPPS) and *G. lamblia* cis-PT (Glcis-PT) are homodimeric *cis*-prenyltransferases. The secondary structure of the *cis*-prenyltransferase domain is depicted according to the structure of EcUPPS. The secondary structure of the C-terminus is depicted according to our structural prediction.

Recently, we have devised an efficient procedure for overexpression and purification of human DHDDS in *E. coli* [16,17]. Similar to the bacterial homolog UPPS, we showed that human DHDDS forms active homodimers [17]. However, the molecular details of the interaction interface and the overall architecture of the enzyme remained obscure. Here, we determine a structural model for the full-length human DHDDS homodimer using an integrative approach that combines multiple sources of

information [18], including a radioligand-based enzyme kinetics assay, size-exclusion chromatography multi-angle light scattering (SEC-MALS), small-angle X-ray scattering (SAXS), and hydrogen-deuterium exchange mass-spectrometry (HDX-MS).

## 2. Materials and Methods

**2.1. Cloning** – Full-length human DHDDS (UniProt Q86SQ9) was synthesized and cloned into pET-32b plasmid between the NdeI and BamHI restriction sites (GenScript, USA) as a thioredoxin (TRX) fusion protein, as previously described [17]. The construct includes a 6xHis-tag to facilitate protein purification and a TEV-protease (Tobacco Etch Virus) cleavage site to remove the 6xHis-tag and TRX fusion.

**2.2. Protein expression and purification** – Protein expression and purification was carried out as previously described [16,17]. Briefly, *E. coli* T7 express competent cells were transformed, grown in 2xYT medium at 37°C until reaching OD<sub>595nm</sub> = 0.5 and induced at 16°C by adding 0.5 mM isopropyl β-D-1-thiogalactopyranoside (IPTG). Proteins were expressed at 16°C for 16-20 h, harvested by centrifugation (10,000xg for 10 min), and then resuspended in buffer with 1 µg/ml DNase I (type DN-25, Sigma) and a protease inhibitor mixture (Roche Applied Science, Germany). Resuspended cells were homogenized and disrupted in a microfluidizer (Microfluidics, USA). Soluble proteins were recovered by centrifugation at ~ 40,000xg for 45 min at 4°C. Overexpressed proteins were purified on a TALON-superflow column (GE healthcare, UK), followed by TEV protease cleavage of the 6xHis-tag. The reaction mixture was buffer exchanged to remove imidazole and loaded again on a TALON-superflow column. The unbound protein was collected and loaded onto a Superdex-200 preparative size-exclusion column pre-equilibrated with Tris-HCl, pH 7.5, 150 mM NaCl, 20 mM β-mercaptoethanol and 0.02% Triton-X100. Purified proteins were flash-frozen in liquid nitrogen and stored at -80° C until use. Protein purity was >90%, as judged by SDS-PAGE.

**2.3. Size-exclusion chromatography multi-angle light scattering (SEC-MALS)** – Experiments were carried out using a size-exclusion chromatography column (superdex-200 increase 10/300 GL column) pre-equilibrated with Tris-HCl, pH 7.5, 150 mM NaCl, 20 mM β-mercaptoethanol and 0.02% Triton-X100. Samples (2 mg/ml; 50 µl) were injected onto an HPLC, connected to an 8-angle light-scattering detector, followed by a differential refractive-index (RI) detector (Wyatt Technology, Santa Barbara, CA, USA). RI and MALS readings were analyzed with the ASTRA software package (Wyatt Technology, Santa Barbara, CA, USA) to determine molecular mass.

**2.4. Enzyme kinetics** – The activity of purified DHDDS was assayed as previously described [16,17]. Briefly, 0.1-1 µM of purified protein were mixed with FPP and [<sup>14</sup>C]-IPP to initiate the reaction in buffer composed of Tris-HCl, pH 7.5, 150 mM NaCl, 20 mM β-mercaptoethanol, 0.1% Triton-X100 at 30°C. 10 mM EDTA (final concentration) were added to quench the reaction and 1 mL of water-saturated 1-butanol was added to extract the reaction products by thorough vortexing. Initial rates were measured by quenching the reaction at 10% or lower substrate consumption. The K<sub>m</sub> value of FPP was determined in the presence of 0.1 µM protein by varying [FPP] while holding [IPP] constant at 20 µM, and the K<sub>m</sub> value of IPP was determined in the presence of 0.5 µM protein by varying [IPP] while holding [FPP] constant at 10 µM. The product, encompassing <sup>14</sup>C, was quantitated using a scintillation counter. Kinetic constants were obtained by fitting the data to the Michaelis-Menten equation using Origin 7.0 (OriginLab, USA).

**2.5. Small-angle X-ray scattering (SAXS)** – SAXS data were measured at beamline BM29 of the European Synchrotron Radiation Facility (ESRF), Grenoble, France. 50 µL of purified protein concentrated to ~8 mg/mL were chromatographed using a size-exclusion chromatography column (superdex-200 increase 10/300 GL column) pre-equilibrated with Tris-HCl, pH 7.5, 150 mM NaCl, 20 mM β-mercaptoethanol and 0.02% Triton-X100 on an ultra-performance liquid chromatograph (Shimadzu Corporation). Data were collected in-line at 20°C with X-ray beam at wavelength  $\lambda=1.0$  Å, and the distance from the sample to detector (PILATUS 1M, Dectris Ltd) was 2.867 meters, covering a scattering vector range ( $q=4\pi\sin\theta/\lambda$ ) from 0.0025 to 0.5 Å<sup>-1</sup>, with an exposure time of 1 second per frame. The 2D images were reduced to one-dimensional scattering profiles using the software on site. Frames corresponding to dimeric DHDDS were averaged and subtracted from buffer frames using

ScÅtter (<http://www.bioisis.net/tutorial/9>). The experimental radius of gyration ( $R_g$ ) was calculated from data at low  $q$  values in the range of  $qR_g < 1.3$ , according to the Guinier approximation:  $\ln(I(q)) \approx \ln(I(0)) - R_g^2 q^2/3$  using PRIMUS. The  $D_{max}$  value and the Porod volume were derived from the paired-distance distribution function (PDDF or  $P(r)$ ) calculated using GNOM. The flexibility analysis was performed using ScÅtter [19].

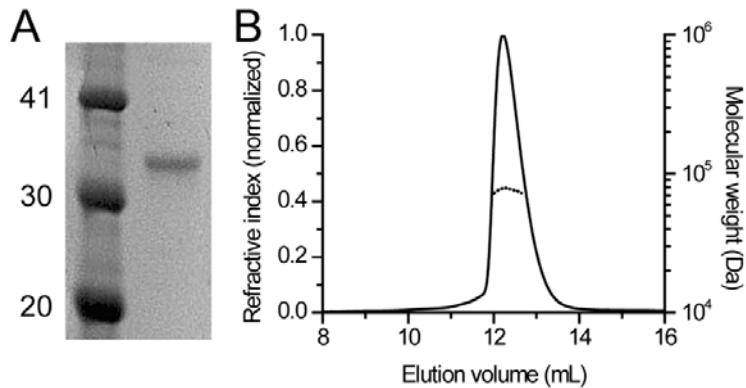
**2.6. Hydrogen-deuterium exchange mass-spectrometry** – The purified proteins were prepared at a concentration of 50 pmole/ $\mu$ l (50  $\mu$ M) in a buffer composed of 25 mM Tris/HCl, pH 7.4, 150 mM NaCl, 10 mM  $\beta$ -mercaptoethanol and 0.02% Triton X-100. The hydrogen deuterium exchange was initiated by mixing 3  $\mu$ l of purified protein with 27  $\mu$ l of the  $D_2O$  buffer (25 mM Tris/HCl, pH 7.4, 150 mM NaCl, 10 mM  $\beta$ -mercaptoethanol, 0.02% Triton X-100 and 10% glycerol in  $D_2O$ ) containing 2 mM EDTA; the mixtures were incubated at various time intervals including 10 sec, 100 sec, 1000 sec and 10000 sec at 4°C. At the indicated time-points, the mixtures were quenched by adding 30  $\mu$ l of ice-cold quench buffer (100 mM  $NaH_2PO_4$ , pH 2.01, 20 mM tris (2-carboxyethyl) phosphine, 1 M guanidine). For non-deuterated samples, 3  $\mu$ l of the purified protein were mixed with 27  $\mu$ l of  $H_2O$  buffer (25 mM Tris/HCl, pH 7.4, 150 mM NaCl, 10 mM  $\beta$ -mercaptoethanol, 0.02% Triton X-100 and 10% glycerol in  $H_2O$ ) containing 2 mM EDTA, and 30  $\mu$ l of ice-cold quench buffer was added. The quenched samples were digested online by passing them through an immobilized pepsin column (2.1 mm  $\times$  30 mm) (Life Technologies) at a flow rate of 100  $\mu$ l/min with 0.05% formic acid in water at 12 °C. Peptide fragments were subsequently collected on a C<sub>18</sub> VanGuard trap column (1.7  $\mu$ m  $\times$  30 mm) (Waters) for desalting with 0.05% formic acid in water and then separated by ultra-performance liquid chromatography (UPLC) using an ACQUITY UPLC C<sub>18</sub> column (1.7  $\mu$ m, 1.0 mm  $\times$  100 mm) (Waters) at a flowrate of 40  $\mu$ l/min with an acetonitrile gradient by using two pumps, which started with 8 % B and increased to 85% B over the next 8.5 min. The mobile phase A was 0.15% formic acid in water and solvent B was acetonitrile containing 0.15% formic acid. To minimize the back-exchange of deuterium to hydrogen, the sample, solvents, trap, and UPLC column were all maintained at pH 2.5 and 0.5°C during analysis. Mass spectral analyses were performed with a Xevo G2 Quadrupole-time of fly (Q-TOF) equipped with a standard ESI source in MS<sup>E</sup> mode (Waters, Milford, MA, USA). The mass spectra were acquired in the range of m/z 100-2000 for 12 minutes in positive ion mode. Peptides were identified in non-deuterated samples with ProteinLynx Global Server (PLGS) 2.0 (Waters, Milford, MA, USA). The following parameters were applied: monoisotopic mass, nonspecific for the enzyme while allowing up to 1 missed cleavage, MS/MS ion searches, automatic fragment mass tolerance, and automatic peptide mass tolerance. Searches were performed with the variable methionine oxidation modification, and the peptides were filtered with a peptide score of 6. To process HDX-MS data, the amount of deuterium in each peptide was determined by measuring the centroid of the isotopic distribution using DynamX 2.0 (Waters, Milford, MA, USA). All data were derived from 3 independent experiments. The back-exchange level was not corrected because the fully-deuterated samples aggregated.

**2.7. Structural modelling** – The cis-prenyltransferase domain was modeled by comparative modeling [20] using Z,Z-Farnesyl Diphosphate Synthase (PDB 5hxn, sequence identity 41%) and undecaprenyl pyrophosphate synthase (PDB 6acs, sequence identity 35%) as templates. The C-terminal domain (residues 261-340) was modeled by RaptorX-Contact [21]. Residues 18-29 were modeled as  $\alpha$ -helix based on the secondary structure prediction. The dimer was assembled in MODELLER v9.18 [20] using known dimerization interface of the cis-prenyltransferase domains.

### 3. Results

**3.1. Human DHDDS forms a monodisperse homodimer** – Short- and medium-chain cis-prenyltransferases form homodimers via a conserved dimerization interface throughout the phylogenetic tree [7,14]. The structure of these homodimers was extensively studied at high-resolution [22]. However, while sharing the conserved cis-prenyltransferase homology domain, including the dimerization interface, the stoichiometry of long-chain cis-prenyltransferases, such as human DHDDS, was not comprehensively studied [7,23]. Previously, in order to assess the oligomeric state of human DHDDS, we used size exclusion chromatography (SEC). By comparing with a BSA standard, we concluded that the elution volume corresponds to that expected for a homodimer [17]. However,

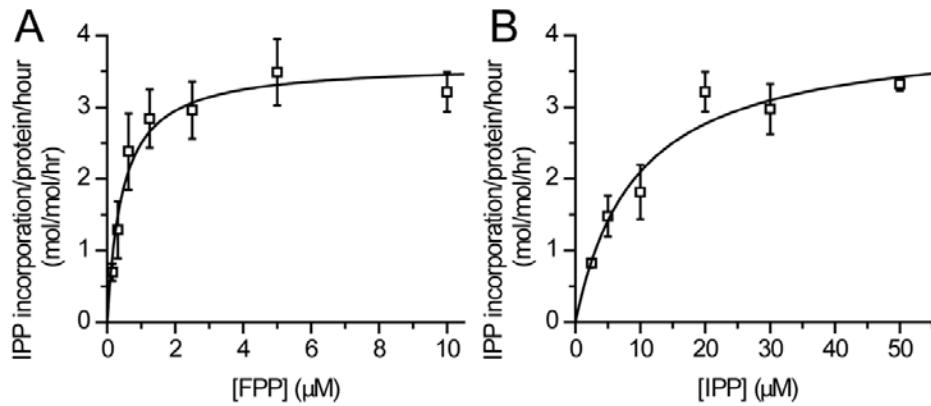
relying on the elution volume, which depends on the hydrodynamic radius, provides a mere surrogate of the actual mass. For example, lower than expected elution volume, corresponding to a higher mass estimate, can result from an elongated conformation or high flexibility of the previously uncharacterized C-terminus [24].



**Figure 2. Human DHDDS forms monodispersed homodimers.** (A) SDS-PAGE analysis of purified DHDDS. Left lane, molecular weight marker; right lane, purified DHDDS. Numbers indicated weight in kDa. (B) SEC-MALS analysis of purified DHDDS. The solid line represents the normalized refractive index and the dashed line represents the calculated molecular weight in Da.

Thus, following overexpression and purification of full-length human DHDDS in *E. coli* (Figure 2A), we used SEC-MALS (Figure 2B) [25,26]. MALS is based on the formation of an oscillating dipole within molecules resulting from interaction of the molecule with light. The molecular oscillating dipole is then re-radiated and measured in multiple angles, where the intensity of the scattered light depends on the size and shape of the molecule. Thus, this method allows to accurately determine both absolute molar mass and the average size of molecules in solution by detection of light scattering in solution [25,26]. In agreement with our previous studies, the calculated mass is  $77.0 \pm 4.0$  kDa, closely matching the calculated mass of a homodimer (78.3 kDa). Notably, the protein is monodispersed ( $M_w/M_n = 1.000 \pm 0.07$ ) in solution and no other oligomeric species are observed (Figure 2B), making this preparation suitable for downstream structural studies.

**3.2. Catalytic activity of purified DHDDS** – Previous studies showed that NgBR-containing heteromeric *cis*-prenyltransferase complexes are formed in cells, and NgBR was suggested to increase the activity of DHDDS within the complex [6]. Interestingly, NgBR encompasses a *cis*-prenyltransferase homology domain, which includes the conserved dimerization interface but lacks residues which are essential for catalysis [8]. Thus, it was suggested that DHDDS and NgBR form heterodimers [8]. However, several combinatorial dimeric complexes may co-exist physiologically – DHDDS homodimers, NgBR homodimers, and DHDDS-NgBR heterodimers. Here, we sought to determine the catalytic activity of DHDDS in the absence of NgBR (Figure 3), to gain insight into the possible mechanism that leads to the increased activity exerted by NgBR. The turnover number ( $k_{cat}$ ) of homodimeric DHDDS is  $1.1 \cdot 10^{-3} \pm 0.1 \cdot 10^{-3} \text{ s}^{-1}$ , with  $K_m$  values of  $9.3 \pm 2.8 \mu\text{M}$  and  $0.45 \pm 0.1 \mu\text{M}$  for IPP and FPP, respectively (Figure 3A,B). Interestingly, the  $K_m$  values obtained for the homodimeric DHDDS are similar to those obtained for the heteromeric *cis*-prenyltransferase complex, while the  $k_{cat}$  value is  $\sim 500$  fold lower in the absence of NgBR [23]. Thus, our results suggest that co-assembly with NgBR does not alter the substrates affinity, but allosterically increases the catalytic activity of DHDDS.



**Figure 3. *In vitro* catalytic activity of purified DHDDS.** *In vitro* activity of purified human DHDDS with FPP and <sup>14</sup>C-IPP as substrates, assessed as IPP incorporation. **(A)** FPP-dependent activity in the presence of 20 μM <sup>14</sup>C-IPP. **(B)** IPP-dependent activity in the presence of 10 μM FPP. Initial rates were obtained by quenching the reaction at <10% substrate consumption. Data are presented as mean ± SEM (n = 3).

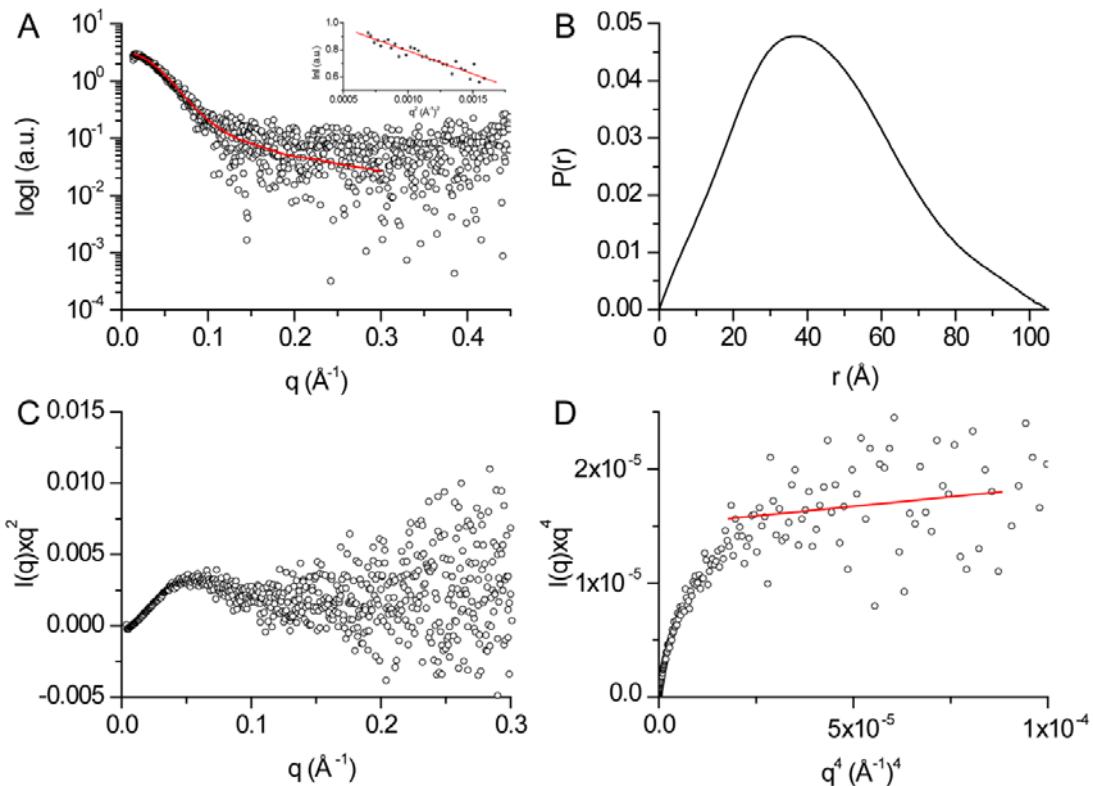
**3.3. Structural features of DHDDS in solution** – Conserved domains search of human DHDDS readily identified the canonical *cis*-prenyltransferase domain but yielded no hits for the C-terminus [13]. However, as a C-terminal extension is common to long-chain *cis*-prenyltransferases (Figure 1B), we hypothesized that it may form a structured domain which plays a role in enzyme assembly and function. Thus, in order to gain insight into the spatial organization of the C-terminus, we used SAXS (Table 1). SAXS provides comprehensive structural information about macromolecules in solution including molecular shape, size, and structural variability [19,27]. Moreover, SAXS profile can be converted into an approximate distribution of pairwise electron distances in the macromolecule (i.e. the paired-distance distribution function, PDDF) via a Fourier transform.

**Table 1. SAXS Data collection and analysis.**

Data collection parameters	
Beamline	ESRF BM29
Beam geometry (mm <sup>2</sup> )	0.7 x 0.7
Wavelength (Å)	1.0
Q range (Å <sup>-1</sup> )	0.0025 – 0.5
Exposure per frame (seconds)	1
Temperature (°C)	20
Structural parameters	
R <sub>g</sub> (Å) <sup>1</sup>	39.3 ± 0.1
D <sub>max</sub> (Å) <sup>2</sup>	105 ± 11
Porod volume [from P(r)] (10 <sup>3</sup> Å <sup>3</sup> )	132.8
Estimated mass (kDa) <sup>2</sup>	78.1 ± 7.8
Porod exponent	3.6
χ <sup>2</sup> model	1.03
Software employed	
Primary data reduction	AUTOMAR
Data processing	PRIMUS, GNOM, ScÅtter

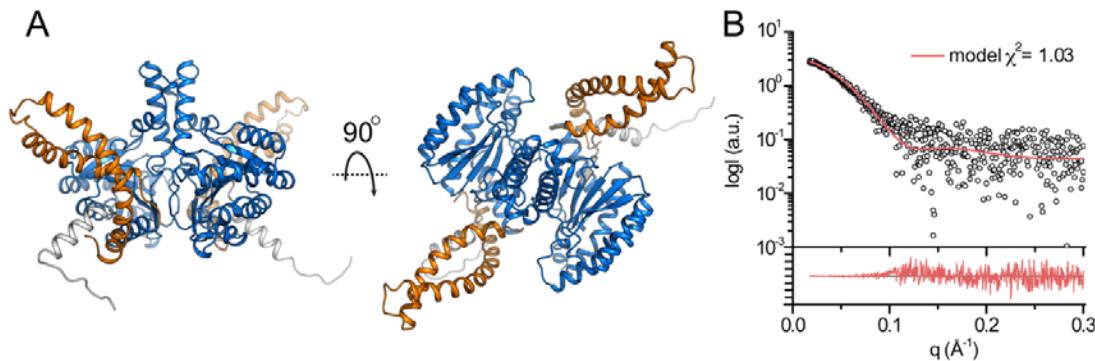
<sup>1</sup> ± S.E. <sup>2</sup> ± 10% (estimated range).

SAXS data were recorded in-line with SEC, and frames corresponding to homodimeric DHDDS were averaged and background subtracted (Figure 4A). The protein exhibits  $R_g = 31.8 \pm 1.4 \text{ \AA}$  (radius of gyration, the root-mean-square-distance from the center of mass) and  $D_{\max} = 105 \pm 11 \text{ \AA}$  (maximal intramolecular distance). The PDDF indicates that the protein assumes an overall elongated conformation in solution (Figure 4B), and the derived Porod volume matches that expected for a homodimer (Table 1) [28]. Importantly, Kratky plot analysis demonstrates that the protein is well folded, as reflected by the peak observed at low angles and convergence with the X-axis at high angles (Figure 4C) [19,27]. Next, the protein flexibility was assessed using the Porod-Debye law (Figure 4D). The scattering curve shows a clear plateau within the Porod-Debye region with a Porod exponent of 3.6, consistent with an overall compact and rigid particle [19,27]. In summary, our SAXS analysis is consistent with a well-folded and compactly packed C-terminus.



**Figure 4. SAXS analysis of purified human DHDDS.** (A) Experimental SAXS curves of human DHDDS (circles) along with the fit obtained by GNOM (red line). Inset: Guinier analysis. (B) Paired-distance distribution function of human DHDDS determined using GNOM. (C) Kratky plot analysis. (D) Porod-Debye plot. The red line represents the fit used to determine the Porod exponent.

**3.4. Structural modelling of full-length DHDDS** – Following our SAXS analysis, which suggested that the C-terminus forms a structured domain, we sought to determine the structure of full-length DHDDS using a computational modelling approach. While the *cis*-prenyltransferase domain shows high structural conservation, allowing comparative modelling [20], the N-terminus (residues 18-29) was modeled as  $\alpha$ -helix based on the secondary structure prediction and the C-terminal domain (residues 254-333) was modeled by co-evolution analysis [21]. Finally, the dimeric assembly was generated using the canonical dimerization interface of the *cis*-prenyltransferase domains (Figure 5A). Importantly, fitting the model against the SAXS data reveals high agreement with  $\chi^2=1.03$  (Figure 5B) [29].



**Figure 5. Computational modelling of full-length human DHDDS.** (A) The model of full-length human DHDDS is shown in cartoon representation. The N-terminus, the *cis*-prenyltransferase homology domain and the C-terminal helix-turn-helix motif are colored gray, blue and orange, respectively. (B) Fit of the theoretical SAXS curve of the model (red line) to the measured SAXS profile (circles). The lower panel shows the residuals plot.

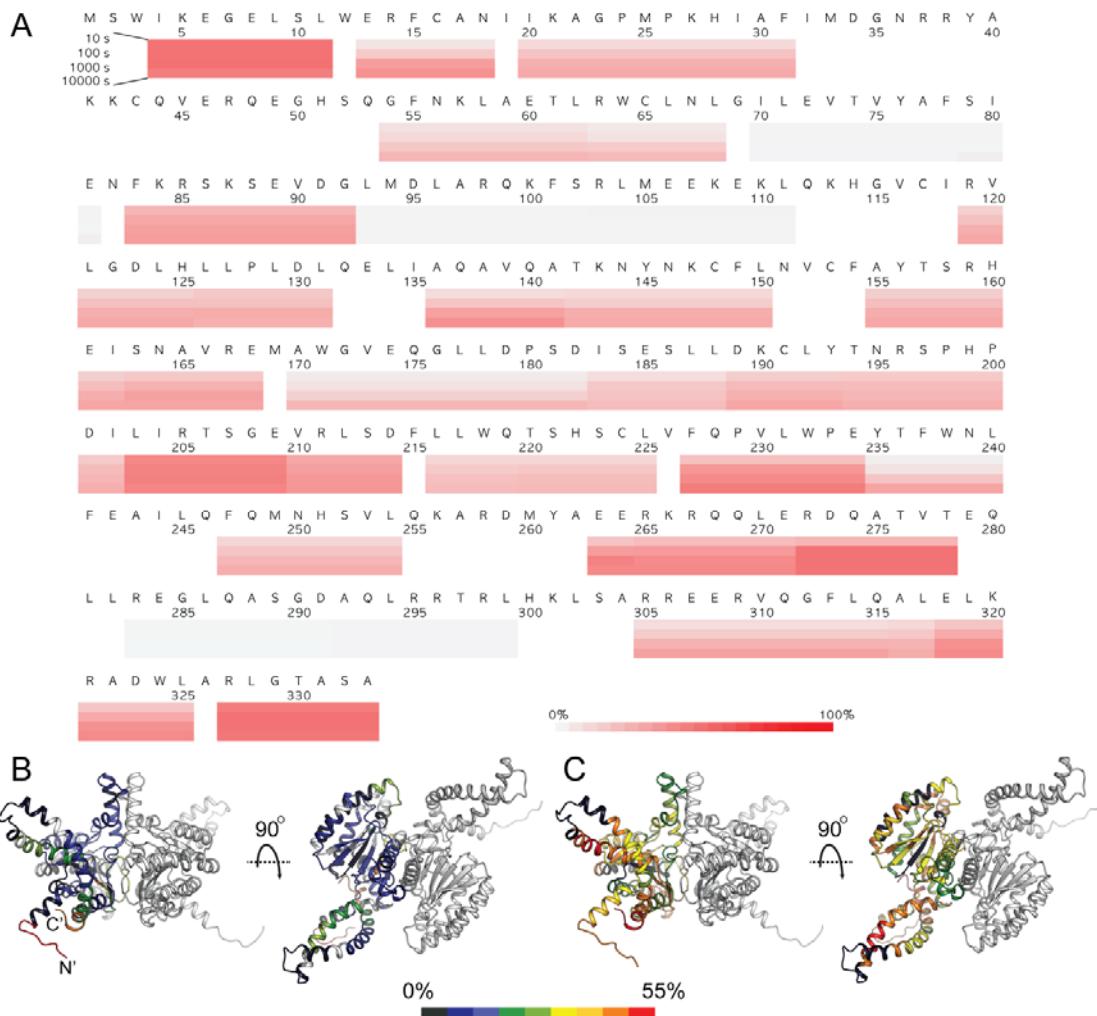
**3.5. Structural dynamics of DHDDS** – Our SAXS analysis is consistent with our computational model, suggesting that the *cis*-prenyltransferase homology domains dimerize via the canonical dimerization interface, tightly packed against the C-terminus which forms a helix-turn-helix fold (Figure 5). To further experimentally validate our model, and to gain insight into the structural dynamics of DHDDS, we used HDX-MS. In HDX-MS experiments, the protein sample is diluted into a D<sub>2</sub>O containing buffer. Following dilution, deuterium in the buffer is exchanged with amide hydrogens within the protein peptide backbone [30]. The reaction is quenched at different time points, followed by proteolytic digestion and peptide separation using reverse-phase chromatography. The eluted peptides are detected using a mass spectrometer, and the degree of deuterium incorporation within each peptide is determined. The exchange rate depends on the conformational dynamics and solvent accessibility of each protein region, where folded and solvent inaccessible regions exhibit low HDX.

The peptic peptides analyzed by a mass spectrometer covers 83.5% of DHDDS (Figure S1). The levels of deuterium uptake at 10, 100, 1000 and 10000 seconds are presented on the protein sequence (Figure 6A) and the levels of deuterium uptake at 10 and 10000 seconds are color-coded on our model structure (Figure 6 B,C). Because we could not correct the back-exchange level due to the aggregation issue of the fully-deuterated samples (as discussed in the method section), the highest deuterium uptake was approximately 52% (Figure S1). The average back-exchange level of our HDX-MS system is approximately 30-40%. Therefore, 40-50% deuterium uptake was considered as very high-level deuterium uptake. We carefully examined the deuterium uptake profiles; the residues with uptake plots with high uptake levels (40-50%) even for the short D<sub>2</sub>O incubation (10 sec) were considered to be as highly flexible and/or exposed to the buffer while the residues with uptake plots with low uptake levels (less than 3%) even for the long D<sub>2</sub>O incubation (10,000 sec) were considered to be as highly rigid and/or excluded from the buffer; the residues showing gradual increase in deuterium uptake were considered to have relatively dynamic secondary structures and/or relatively less exposed to the buffer.

As expected, the N-terminus (peptide 4-12), the C-terminus (peptide 327-334), and relatively long loop (peptide 203-210) regions showed highest deuterium uptake (more than 45% at 10 sec) as these regions are exposed to the buffer and flexible. The core of the *cis*-prenyltransferase homology domain (peptides 70-79, 71-82, 93-103, 101-112) showed lowed deuterium uptake (less than 3% at 10,000 sec), which suggests that the core of the *cis*-prenyltransferase homology domain is highly rigid. On the other hand, other regions of the *cis*-prenyltransferase homology domain showed gradual increase of the deuterium uptake suggesting that these regions form secondary structures with relatively dynamic nature.

Consistent with our structural model, the dimerization interface (peptides 170-182, 173-189, 216-226, 220-226) exhibits medium deuterium incorporation with gradual increase (3 to 25%) although the peptides contain flexible loops, indicating their relative inaccessibility to the bulk solvent. Importantly, highlighting its folded state, the C-terminus (residues 254-333) exhibits an overall time-dependent deuterium incorporation. Thus, as predicted by our computational model, the HDX-MS

results support the notion that the monomers interact via the conserved dimerization interface and that the C-terminus forms a folded domain.



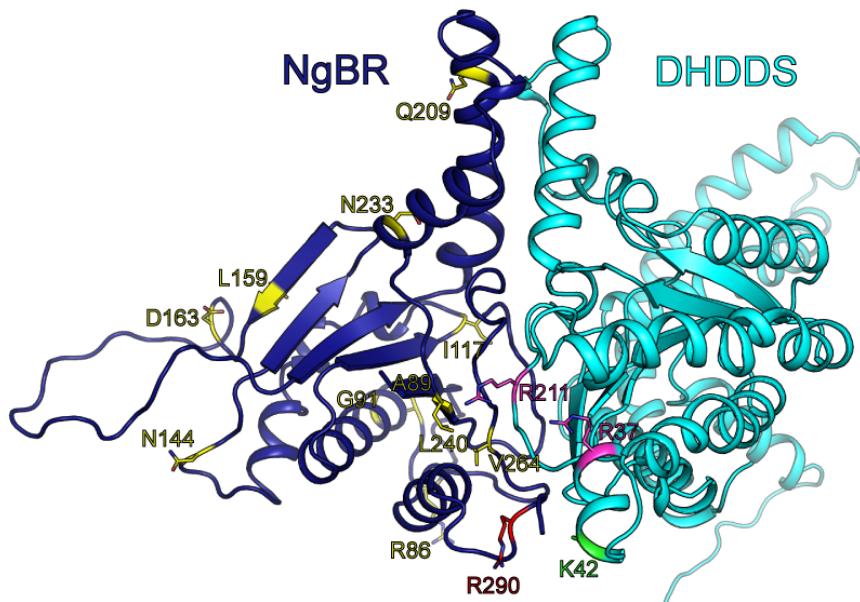
**Figure 6. HDX-MS analysis of human DHDDS.** (A) Deuterium uptake profiles. Incorporation of deuterium at 10, 100, 1000, and 10,000 sec are indicated by color-coded blocks underlining the amino acid sequence. The color legend shows the deuterium uptake levels. (B,C) The heat maps at 10 sec (B) and 10,000 sec (C) are overlaid on one monomer of the DHDDS model. The color legend shows the deuterium uptake level. Regions not covered are colored gray.

#### 4. Discussion

Here we provide a structural model for the full-length human DHDDS using a hybrid experimental and computational approach. Following expression and purification (Figure 2A), we showed using SEC-MALS that DHDDS exists as a monodispersed dimer in solution (Figure 2B), making the preparation suitable for downstream experimentation. The purified homodimer is active, exhibiting similar  $K_m$  values to the values reported for the heteromeric *cis*-PT complex but markedly reduced  $k_{cat}$  (Figure 3) [7]. Despite the lack of matching homologous domains, SAXS analysis revealed that the protein is compact and rigid, supporting a notion that the C-terminus forms a folded domain (Figure 4). Co-evolution analysis predicted a helix-loop-helix structure for the C-terminal domain (Figure 5A), and the model was verified by two independent approaches. The model fits the SAXS data very well (Figure 5B), suggesting that the global structure is similar to that existing in solution. Moreover, HDX-MS analysis supports the structural conservation of the dimerization interface and the folded state of the C-terminus (Figure 6).

In contrast to the well characterized short- and medium-chain prenyltransferases, which function as homodimers [15], long-chain *cis*-prenyltransferases and rubber synthases are heteromeric enzymes [7]. The heteromeric human long-chain *cis*-prenyltransferase, composed of DHDDS and NgBR, is gaining increasing biomedical interest due to the identification of mutations in both its components leading to human diseases [9,10,12,31,32]. Rubber synthases are of significant biotechnological and industrial interest, and a recent study showed that, in *Hevea brasiliensis*, the industrial source for latex, the complex is composed of at least three different subunits [33]. Despite the importance of the heteromeric *cis*-prenyltransferases, the molecular mechanisms involved in subunits engagement, as well as the mechanisms determining their product chain length, remain elusive due to lack of structural information [7]. Thus, the goal of the present study was to provide structural insights into the organization of the entire catalytic subunit of the human *cis*-prenyltransferase complex, using an experimentally-validated computational approach.

Importantly, our model predicts a novel C-terminal motif, absent from short- and medium-chain homodimeric *cis*-prenyltransferases. However, its structural role in the context of the heteromeric human long-chain *cis*-prenyltransferase complex remains to be determined. According to the structural conservation of the canonical dimerization interface, DHDDS and NgBR may form heterodimers [8]. Therefore, we constructed a homology model of the *cis*-prenyltransferase homology domain of NgBR based on the recently resolved structure of the yeast ortholog, Nus1 [8], and used it to replace one of the DHDDS subunits in our model (Figure 7). This model of the heterodimeric complex suggests that the C-terminal domain is not directly engaged in subunits interactions. Intriguingly, this domain is somewhat reminiscent of the dimerization interface, formed by two anti-parallel packed helices. This arrangement may suggest that the C-terminal domain can interact with additional proteins, as in the case of rubber synthase. Future studies, focused on the functional role of the C-terminus, are required to understand its role in the context of the heteromeric complex.



**Figure 7. A model for the heterodimeric *cis*-prenyltransferase complex.** The model of the heterodimeric complex is shown in cartoon representation. DHDDS and NgBR are colored cyan and blue, respectively. Positions with disease-associated mutations are shown as sticks and colored as follows: K42, related to retinitis pigmentosa, green; R37 and R211, related to developmental epileptic encephalopathies, magenta; R290, related to a congenital glycosylation disorder, red; R86, A89, G91, I117, N144, L159, D163, Q209, N233, L240 and V264, related to Parkinson's disease, yellow.

Finally, the DHDDS mutations known to result in autosomal recessive retinitis pigmentosa [9,10] and developmental epileptic encephalopathies [12], and NgBR mutations associated with Parkinson's disease [32] and a congenital glycosylation disorder [31] were mapped on the heterodimer model (Figure 7). Interestingly, all the DHDDS mutations, which are positively charged (K42E, R37H, R211Q),

reside in close proximity to the dimeric interface. Moreover, the NgBR mutation resulting in congenital glycosylation disorder, R290H, is also charged and lies in close proximity to the DHDDS disease-causing mutations and catalytic pocket of the enzymatic complex. R290 is part of the conserved C-terminal RxG motif of NgBR [23], which was previously shown to play a critical role in the catalytic activity of the heteromeric human complex. Indeed, the location of R290 on an unstructured loop may provide it with the structural flexibility needed for interacting with the catalytic pocket and, in turn, to allosterically activate DHDDS. As we show here, the DHDDS homodimers exhibit markedly reduced  $k_{cat}$  compared to the heteromeric complex [23] (Figure 3). Taken together, the structural proximity of the NgBR C-terminus to the catalytic pocket, its predicted flexibility, and the reduced activity of homodimeric DHDDS compared with the heteromeric complex harboring NgBR, are consistent with the role of NgBR in allosteric activation of DHDDS.

In summary, this study presents the first model of full-length human DHDDS, experimentally validated using two independent approaches. This model sheds light on the previously uncharacterized C-terminus of DHDDS and, together with the homology model of NgBR, provides a structural framework for future structural and functional studies of the human *cis*-prenyltransferase complex, as well as other long-chain prenyltransferases and rubber synthases.

**Funding:** This research was funded by Israel Science Foundation, grant numbers 1721/16, 1775/12 (Y.H.), Israel Cancer Research Foundation, grant number 01214 (Y.H.), German-Israeli Foundation, grant number I-2425-418.13/2016 (Y.H.), Stoltz foundation (H.N.), Tel-Aviv Sourasky Medical Center (A.L. and M.G.), the Clair and Amedee Maratier Institute for the study of Blindness and Visual Disorders, Sackler Faculty of Medicine, Tel-Aviv University (Y.H. and M.G.), and the Korean government (NRF-2018R1A2B6001554) (K.Y.C.).

**Acknowledgments:** We thank the staff of BM29 at the ESRF for assistance with SAXS experimentation. This work was performed in partial fulfillment of the requirements for a Ph.D. degree of M.L.B., Sackler Faculty of medicine, Tel Aviv University, Israel.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Schwarz, F.; Aebi, M. Mechanisms and principles of N-linked protein glycosylation. *Curr. Opin. Struct. Biol.* 2011, **21**, 576–582.
2. Helenius, A.; Aebi, M. Intracellular functions of N-linked glycans. *Science (80-.)* 2001, **291**, 2364–2369.
3. Stowell, S.R.; Ju, T.; Cummings, R.D. Protein glycosylation in cancer. *Annu. Rev. Pathol.* **2015**, **10**, 473–510.
4. Ng, B.G.; Freeze, H.H. Perspectives on Glycosylation and Its Congenital Disorders. *Trends Genet.* 2018, **34**, 466–476.
5. Freeze, H.H. Understanding human glycosylation disorders: Biochemistry leads the charge. *J. Biol. Chem.* 2013, **288**, 6936–6945.
6. Harrison, K.D.; Park, E.J.; Gao, N.; Kuo, A.; Rush, J.S.; Waechter, C.J.; Lehrman, M.A.; Sessa, W.C. Nogo-B receptor is necessary for cellular dolichol biosynthesis and protein N-glycosylation. *EMBO J.* **2011**, **30**, 2490–2500.
7. Grabińska, K.A.; Park, E.J.; Sessa, W.C. CIS-Prenyltransferase: New insights into protein glycosylation, rubber synthesis, and human diseases. *J. Biol. Chem.* 2016, **291**, 18582–18590.
8. Ma, J.; Ko, T.P.; Yu, X.; Zhang, L.; Ma, L.; Zhai, C.; Guo, R.T.; Liu, W.; Li, H.Z.; Chen, C.C. Structural insights to heterodimeric *cis*-prenyltransferases through yeast dehydrololichyl diphosphate synthase subunit Nus1. *Biochem. Biophys. Res. Commun.* **2019**, **515**, 621–626.
9. Zelinger, L.; Banin, E.; Obolensky, A.; Mizrahi-Meissonnier, L.; Beryozkin, A.;

Bandah-Rozenfeld, D.; Frenkel, S.; Ben-Yosef, T.; Merin, S.; Schwartz, S.B.; et al. A missense mutation in DHDDS, encoding dehydrololichyl diphosphate synthase, is associated with autosomal-recessive retinitis pigmentosa in ashkenazi jews. *Am. J. Hum. Genet.* **2011**, *88*, 207–215.

10. Züchner, S.; Dallman, J.; Wen, R.; Beecham, G.; Naj, A.; Farooq, A.; Kohli, M.A.; Whitehead, P.L.; Hulme, W.; Konidari, I.; et al. Whole-exome sequencing links a variant in DHDDS to retinitis pigmentosa. *Am. J. Hum. Genet.* **2011**, *88*, 201–206.

11. Sabry, S.; Vuillaumier-Barrot, S.; Mintet, E.; Fasseu, M.; Valayannopoulos, V.; Héron, D.; Dorison, N.; Mignot, C.; Seta, N.; Chantret, I.; et al. A case of fatal Type i congenital disorders of glycosylation (CDG I) associated with low dehydrololichol diphosphate synthase (DHDDS) activity. *Orphanet J. Rare Dis.* **2016**, *11*.

12. Hamdan, F.F.; Myers, C.T.; Cossette, P.; Lemay, P.; Spiegelman, D.; Laporte, A.D.; Nassif, C.; Diallo, O.; Monlong, J.; Cadieux-Dion, M.; et al. High Rate of Recurrent De Novo Mutations in Developmental and Epileptic Encephalopathies. *Am. J. Hum. Genet.* **2017**, *101*, 664–685.

13. Marchler-Bauer, A.; Derbyshire, M.K.; Gonzales, N.R.; Lu, S.; Chitsaz, F.; Geer, L.Y.; Geer, R.C.; He, J.; Gwadz, M.; Hurwitz, D.I.; et al. CDD: NCBI's conserved domain database. *Nucleic Acids Res.* **2015**, *43*, D222-6.

14. Takahashi, S.; Koyama, T. Structure and function of *cis*-prenyl chain elongating enzymes. *Chem. Rec.* **2006**, *6*, 194–205.

15. Liang, P.H.; Ko, T.P.; Wang, A.H.J. Structure, mechanism and function of prenyltransferases. *Eur. J. Biochem.* **2002**, *269*, 3339–3354.

16. Edri, I.; Goldenberg, M.; Lisnyansky, M.; Strulovich, R.; Newman, H.; Loewenstein, A.; Khananshvili, D.; Giladi, M.; Haitin, Y. Overexpression and Purification of Human *cis*-prenyltransferase in *Escherichia coli*. *J. Vis. Exp.* **2017**, *2017*.

17. Giladi, M.; Edri, I.; Goldenberg, M.; Newman, H.; Strulovich, R.; Khananshvili, D.; Haitin, Y.; Loewenstein, A. Purification and characterization of human dehydrololichil diphosphate synthase (DHDDS) overexpressed in *E. coli*. *Protein Expr. Purif.* **2017**, *132*, 138–142.

18. Rout, M.P.; Sali, A. Principles for Integrative Structural Biology Studies. *Cell* **2019**, *177*, 1384–1403.

19. Rambo, R.P.; Tainer, J.A. Characterizing flexible and intrinsically unstructured biological macromolecules by SAS using the Porod-Debye law. *Biopolymers* **2011**, *95*, 559–571.

20. Webb, B.; Sali, A. Comparative protein structure modeling using MODELLER. *Curr. Protoc. Bioinforma.* **2016**, *2016*, 5.6.1-5.6.37.

21. Wang, S.; Sun, S.; Li, Z.; Zhang, R.; Xu, J. Accurate De Novo Prediction of Protein Contact Map by Ultra-Deep Learning Model. *PLoS Comput. Biol.* **2017**, *13*.

22. Guo, R.T.; Ko, T.P.; Chen, A.P.C.; Kuo, C.J.; Wang, A.H.J.; Liang, P.H. 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.* **2005**, *280*, 20762–20774.

23. Grabińska, K.A.; Edani, B.H.; Park, E.J.; Kraehling, J.R.; Sessa, W.C. A conserved C-terminal RXG motif in the NgBR subunit of *cis*-prenyltransferase is critical for prenyltransferase activity. *J. Biol. Chem.* **2017**, *292*, 17351–17361.
24. Hall, M. Size Exclusion Chromatography (SEC). In *Biopharmaceutical Processing: Development, Design, and Implementation of Manufacturing Processes*; 2018; pp. 421–432 ISBN 9780128125526.
25. Andersson, M.; Wittgren, B.; Wahlund, K.G. Accuracy in multiangle light scattering measurements for molar mass and radius estimations. Model calculations and experiments. *Anal. Chem.* **2003**, *75*, 4279–4291.
26. Folta-Stogniew, E. Oligomeric states of proteins determined by size-exclusion chromatography coupled with light scattering, absorbance, and refractive index detectors. *Methods Mol. Biol.* **2006**, *328*, 97–112.
27. Dyer, K.N.; Hammel, M.; Rambo, R.P.; Tsutakawa, S.E.; Rodic, I.; Classen, S.; Tainer, J.A.; Hura, G.L. High-throughput SAXS for the characterization of biomolecules in solution: A practical approach. *Methods Mol. Biol.* **2014**, *1091*, 245–258.
28. Koch, M.H.J.; Vachette, P.; Svergun, D.I. Small-angle scattering: A view on the properties, structures and structural changes of biological macromolecules in solution. *Q. Rev. Biophys.* **2003**, *36*, 147–227.
29. Schneidman-Duhovny, D.; Hammel, M.; Tainer, J.A.; Sali, A. Accurate SAXS profile computation and its assessment by contrast variation experiments. *Biophys. J.* **2013**, *105*, 962–974.
30. Masson, G.R.; Jenkins, M.L.; Burke, J.E. An overview of hydrogen deuterium exchange mass spectrometry (HDX-MS) in drug discovery. *Expert Opin. Drug Discov.* **2017**, *12*, 981–994.
31. Park, E.J.; Grabińska, K.A.; Guan, Z.; Stránecký, V.; Hartmannová, H.; Hodaňová, K.; Barešová, V.; Sovová, J.; Jozsef, L.; Ondrušková, N.; et al. Mutation of Nogo-B receptor, a subunit of *cis*-prenyltransferase, causes a congenital disorder of glycosylation. *Cell Metab.* **2014**, *20*, 448–457.
32. Guo, J.F.; Zhang, L.; Li, K.; Mei, J.P.; Xue, J.; Chen, J.; Tang, X.; Shen, L.; Jiang, H.; Chen, C.; et al. Coding mutations in NUS1 contribute to Parkinson's disease. *Proc. Natl. Acad. Sci. U. S. A.* **2018**, *115*, 11567–11572.
33. Yamashita, S.; Yamaguchi, H.; Waki, T.; Aoki, Y.; Mizuno, M.; Yanbe, F.; Ishii, T.; Funaki, A.; Tozawa, Y.; Miyagi-Inoue, Y.; et al. Identification and reconstitution of the rubber biosynthetic machinery on rubber particles from *Hevea brasiliensis*. *Elife* **2016**, *5*.