Structural Characterization of Full-Length Human Dehydrodolichyl Diphosphate Synthase Using an Integrative Computational and Experimental Approach

Michal Lisnyansky Bar-El [1], Su Youn Lee [2], Ah Young Ki [2], Noa Kapelushnik [3], Anat Loewenstein [4], Ka Young Chung [2], Dina Schneidman-Duhovny [5,6], Moshe Giladi [1,7,*,†], Hadas Newman [4,†] and Yoni Haitin [1,*,†]

[1] Department of Physiology and Pharmacology, Sackler Faculty of Medicine, Tel-Aviv University, Tel Aviv 6997801, Israel; michallis505@gmail.com

[2] School of Pharmacy, Sungkyunkwan University, Jangan-gu, Suwon 16419, Korea; youn3887@hanmail.net (S.Y.L.); kay9039@naver.com (A.Y.K.); kychung2@gmail.com (K.Y.C.)

[3] Department of Ophthalmology, Sheba Medical Center, Ramat Gan 5265601, Israel; kapelushniknoa@gmail.com

[4] Department of Ophthalmology, Tel-Aviv Sourasky Medical Center, Tel Aviv 6423906, Israel; anatl@tlvmc.gov.il (A.L.); hadasng@gmail.com (H.N.)

[5] Department of Biological Chemistry, Institute of Life Sciences, The Hebrew University of Jerusalem, Jerusalem 9190401, Israel; dina.schneidman@mail.huji.ac.il

[6] School of Computer Science and Engineering, The Hebrew University of Jerusalem, Jerusalem 9190401, Israel

[7] Tel Aviv Sourasky Medical Center, Tel Aviv 6423906, Israel

* Correspondence: moshegil@post.tau.ac.il (M.G.); yhaitin@tauex.tau.ac.il (Y.H.)

† These authors contributed equally to this work.

check for updates

Abstract: Dehydrodolichyl diphosphate synthase (DHDDS) is the catalytic subunit of the heteromeric human *cis*-prenyltransferase complex, synthesizing the glycosyl carrier precursor for N-linked protein glycosylation. Consistent with the important role of N-glycosylation in protein biogenesis, DHDDS mutations result in human diseases. Importantly, DHDDS encompasses a C-terminal region, which does not converge with any known conserved domains. Therefore, despite the clinical importance of DHDDS, our understating 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. Size-exclusion chromatography multi-angle light scattering revealed that DHDDS forms a monodisperse homodimer in solution. Enzyme kinetics assays revealed that it exhibits catalytic activity, although reduced compared to that reported for the intact heteromeric complex. Our model suggests that the DHDDS C-terminus forms a helix–turn–helix motif, tightly packed against the core catalytic 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 complex.

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].

Dehydrodolichyl diphosphate synthase (DHDDS) together with Nogo-B receptor (NgBR) form the human *cis*-prenyltransferase complex [6,7]. This complex catalyzes the formation of dehydrodolichyl diphosphate (DHDD), a long-chain polyprenyl 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].
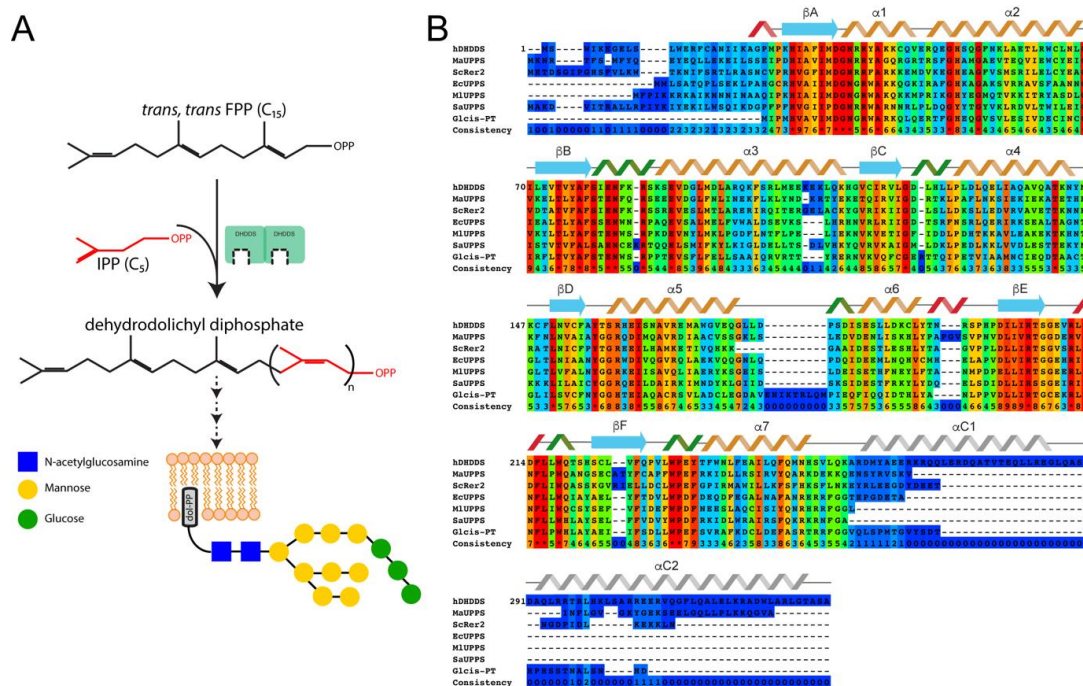

Figure 1. Catalytic cycle of human dehydrodolichyl diphosphate synthase (DHDDS). (**A**) Multiple condensations of a farnesyl diphosphate (FPP) primer with isopentenyl diphosphate (IPP) result in the formation DHDD by DHDDS. DHDD is further processed to form Dol-P, the carrier of the oligosaccharide used for N-glycosylation. (**B**) Sequence alignment of *cis*-prenyltransferases, with conservation scores calculated by PRALINE. Human DHDDS (hDHDDS), *Methanosarcina acetivorans* UPPS (MaUPPS), *Saccharomyces cerevisiae* Rer2 (ScRer2) are components of heteromeric *cis*-prenyltransferase complexes; *Escherichia coli* UPPS (EcUPPS), *Micrococcus luteus* UPPS (MlUPPS), *Sulfolobus acidocaldarius* UPPS (SaUPPS), and *Giardiasis lambia 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.

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 seven α-helices and six β-strands, this domain forms dimers, with a hydrophobic catalytic tunnel surrounded by two α-helices and four β−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.

Recently, we have devised an efficient procedure for overexpression and purification of human DHDDS in *Escherichia 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 $OD_{595nm} = 0.5$ 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 (10,000× *g* 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 ~40,000× *g* 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 X-100. 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. 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 X-100. Samples (2 mg/mL; 50 μL) were injected onto an HPLC, connected to an eight-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, 0.1–0.5 μM of purified protein was mixed with FPP and [^{14}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 X-100 at 30 °C. Then, 10 mM EDTA (final concentration) was 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_m 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_m 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 ^{14}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. SAXS

SAXS data were measured at beamline BM29 of the European Synchrotron Radiation Facility (ESRF), Grenoble, France. Fifty microliters of purified protein concentrated to ~8 mg/mL was 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 X-100 on an ultra-performance liquid chromatograph (Shimadzu Corporation). Data were collected in-line at 20 °C with X-ray beam at wavelength λ = 1.0 Å, and the distance from the sample to detector (PILATUS 1M, Dectris Ltd.) was 2.867 m, covering a scattering vector range (q = $4\pi\sin\theta/\lambda$) from 0.0025 to 0.5 Å^{-1}, with an exposure time of 1 s 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. HDX-MS

The purified proteins were prepared at a concentration of 50 pmole/μL (50 μM) in a buffer composed of 25 mM Tris/HCl, pH 7.4, 150 mM NaCl, 10 mM β-mercaptoethanol, and 0.02% Triton X-100. The hydrogen–deuterium exchange was initiated by mixing 3 μL of purified protein with 27 μL of the D$_2$O buffer (25 mM Tris/HCl, pH 7.4, 150 mM NaCl, 10 mM β-mercaptoethanol, 0.02% Triton X-100, and 10% glycerol in D$_2$O) containing 2 mM EDTA; the mixtures were incubated at various time intervals including 10, 100, 1000, and 10,000 s at 4 °C. At the indicated time-points, the mixtures were quenched by adding 30 μL of ice-cold quench buffer (100 mM NaH$_2$PO$_4$, pH 2.01, 20 mM tris (2-carboxyethyl) phosphine, 1 M guanidine). For non-deuterated samples, 3 μL of the purified protein were mixed with 27 μL of H$_2$O buffer (25 mM Tris/HCl, pH 7.4, 150 mM NaCl, 10 mM β-mercaptoethanol, 0.02% Triton X-100, and 10% glycerol in H$_2$O) containing 2 mM EDTA, and 30 μL of ice-cold quench buffer was added. The quenched samples were digested online by passing them through an immobilized pepsin column (2.1 mm × 30 mm) (Life Technologies) at a flow rate of 100 μL/min with 0.05% formic acid in water at 12 °C. Peptide fragments were subsequently collected on a C$_{18}$ VanGuard trap column (1.7 μm × 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$_{18}$ column (1.7 μm, 1.0 mm × 100 mm) (Waters) at a flowrate of 40 μL/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 Quadruple-Time of Flight (Q-TOF) equipped with a standard electrospray ionization (ESI) source in MS^E mode (Waters, Milford, MA, USA). The mass spectra were acquired in the range of m/z 100–2000 for 12 min 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, 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 three independent experiments. The back-exchange level was not corrected because the fully deuterated samples aggregated during the long incubation required.

2.7. Structural Modeling

The *cis*-prenyltransferase domain was modeled by comparative modeling [20] using Z,Z-farnesyl diphosphate synthase (PDB 5hxn [21], sequence identity 41%) and undecaprenyl pyrophosphate synthase (PDB 6acs [22], sequence identity 35%) as templates. The C-terminal domain (residues 261–340) was modeled by RaptorX-Contact, which is using co-evolutionary information to predict residue–residue contacts [23]. Residues 18–29 were modeled as α-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 [24]. 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 [7,25]. Previously, in order to assess the oligomeric state of human DHDDS, we used size-exclusion chromatography (SEC). By comparing with a bovine serum albumin (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 [26].

Thus, following overexpression and purification of full-length human DHDDS in *E. coli* (Figure 2A), we used SEC-MALS (Figure 2B) [27,28]. Importantly, MALS allows determination of absolute molar mass and average size of molecules based on the intensity of light scattered at different angles, while SEC only serves as a fractionation step in this setup [27,28]. In agreement with our previous studies, the calculated mass is 77.0 ± 4.0 kDa, closely matching the calculated mass of a homodimer (78.3 kDa). Notably, the protein is monodisperse (Mw/Mn = 1.000 ± 0.07) in solution and no other oligomeric species are observed (Figure 2B), making this preparation suitable for downstream structural studies.
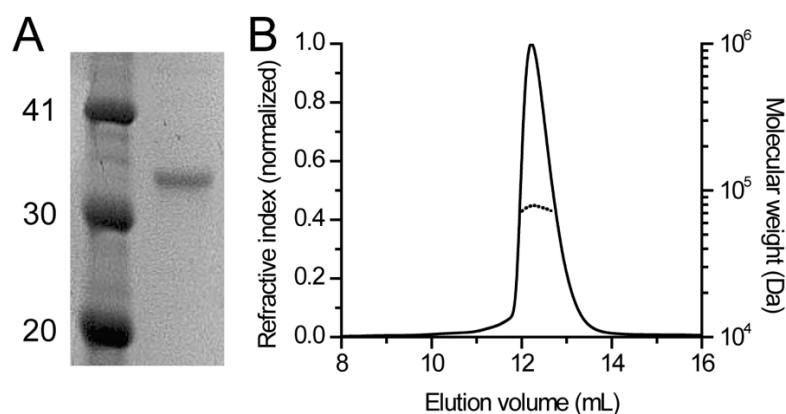
Figure 2. Human DHDDS forms monodisperse homodimers. (**A**) SDS-PAGE analysis of purified DHDDS. Left lane, molecular weight marker; right lane, purified DHDDS. Numbers indicate weight in kDa. (**B**) SEC-MALS analysis of purified DHDDS. The solid line represents the normalized refractive index and the dashed line represents the calculated molecular weight in Da.

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 $1.1 \times 10^{-3} \pm 0.1 \times 10^{-3}$ s^{-1}, with K_m values of 9.3 ± 2.8 and 0.45 ± 0.1 µM 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 ~ 500-fold lower in the absence of NgBR [25]. Thus, our results suggest that co-assembly with NgBR does not alter the substrates' affinity, but allosterically increases the catalytic activity of DHDDS.
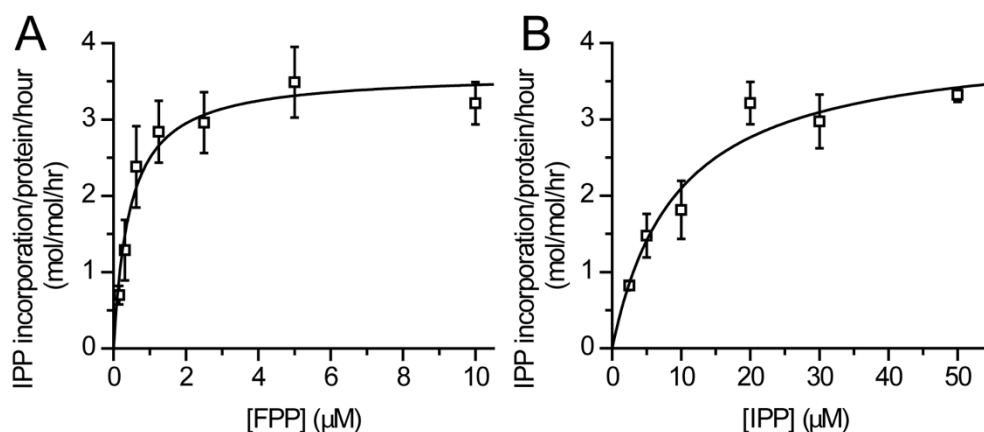

Figure 3. In vitro catalytic activity of purified DHDDS. (**A**,**B**) In vitro activity of purified human DHDDS with FPP and ^{14}C-IPP as substrates, assessed as IPP incorporation. (**A**) FPP-dependent activity in the presence of 20 µM ^{14}C-IPP. (**B**) IPP-dependent activity in the presence of 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 [19,29]. Moreover, the SAXS profile can be converted into an approximate distribution of pairwise electron distances in the macromolecule (i.e., PDDF) via a Fourier transform.

Table 1. SAXS data collection and analysis.

Data Collection Parameters	
Beamline	ESRF BM29
Beam geometry (mm^2)	0.7×0.7
Wavelength (Å)	1.0
Q range (Å^{-1})	0.0025–0.5
Exposure per frame (seconds)	1
Temperature (°C)	20
Structural parameters	
R_g (Å)[1]	31.8 ± 1.4
D_{max} (Å)[2]	105 ± 11
Porod volume [from $P(r)$] (10^3 Å^3)	132.8
Estimated mass (kDa)[2]	78.1 ± 7.8
Porod exponent	3.6
χ^2 model	1.03
Software employed	
Primary data reduction	AUTOMAR
Data processing	PRIMUS, GNOM, ScÅtter

[1] $\pm$ S.E. [2] $\pm$ 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 ± 1.4 Å (radius of gyration, the root-mean-square-distance from the center of mass) and D_{max} = 105 ± 11 Å (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) [30]. 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) [19,29]. 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 [19,29]. In summary, our SAXS analysis is consistent with a well-folded and compactly packed C-terminus.

3.4. Structural Modeling 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 modeling approach. While the *cis*-prenyltransferase domain shows high structural conservation, allowing comparative modeling [20], the N-terminus (residues 18–29) was modeled as α-helix based on the secondary structure prediction and the C-terminal domain (residues 254–333) was modeled by co-evolution analysis [23]. 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) [31].
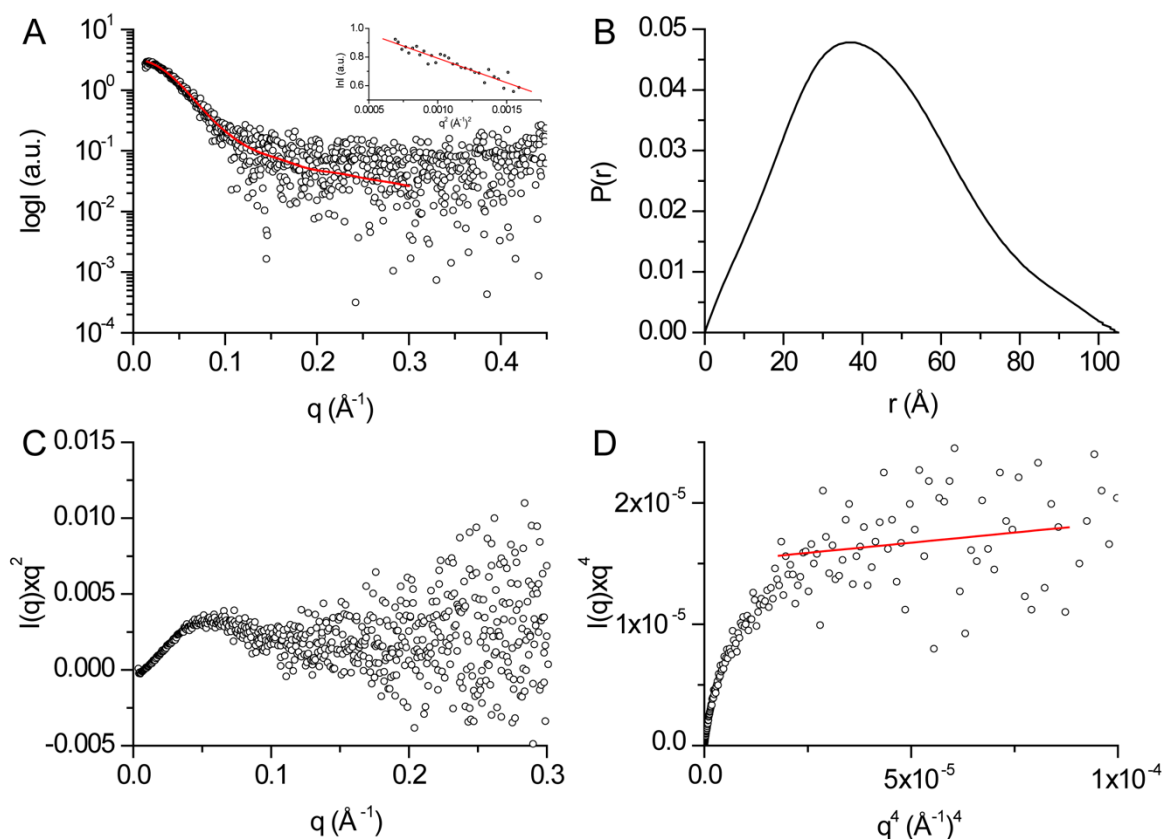

Figure 4. SAXS analysis of purified human DHDDS. (**A**) Experimental SAXS curves of human DHDDS (circles) along with the fit obtained by GNOM (red line). Inset: Guinier analysis. (**B**) Paired-distance distribution function of human DHDDS determined using GNOM. (**C**) Kratky plot analysis. (**D**) Porod–Debye plot. The red line represents the fit used to determine the Porod exponent.
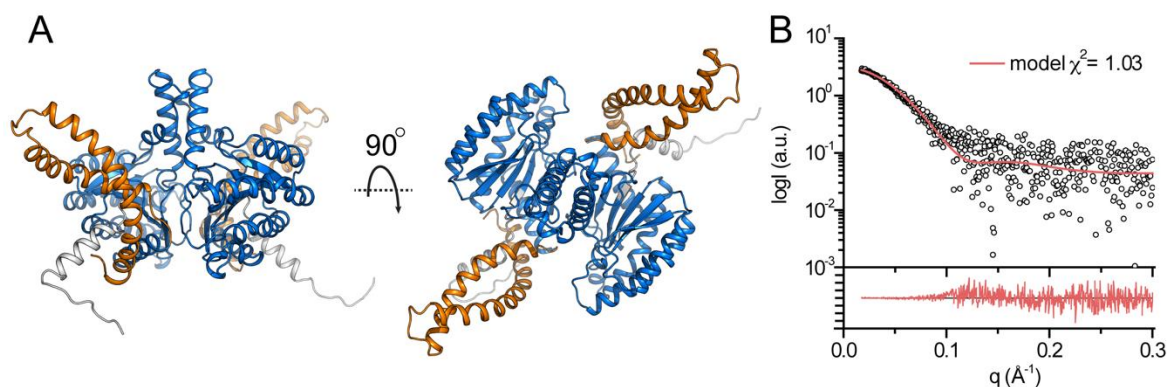
Figure 5. Computational modeling of full-length human DHDDS. (**A**) The model of full-length human DHDDS is shown in cartoon representation. The N-terminus, the *cis*-prenyltransferase homology domain, and the C-terminal helix–turn–helix motif are colored gray, blue, and orange, respectively. (**B**) 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. HDX-MS Analysis of Human DHDDS

Our SAXS analysis is consistent with our computational model, suggesting that the *cis*-prenyltransferase homology domains dimerize via the canonical dimerization interface. Moreover, it is consistent with tight packing of the C-terminus, which forms a helix–turn–helix fold, against the homology domain (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_2O-containing buffer. Following dilution, deuterium in the buffer is exchanged with amide hydrogens within the protein peptide backbone [32]. The reaction is quenched at different time points, followed by proteolytic digestion with pepsin and separation of the peptic peptides 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 cover 83.5% of DHDDS (Figure S1). The levels of deuterium uptake at 10, 100, 1000, and 10,000 s are presented on the protein sequence (Figure 6A) and the levels of deuterium uptake at 10 and 10,000 s are color-coded on our model structure (Figure 6B,C). Because we could not correct the back-exchange level due to the aggregation issue of the fully deuterated samples (as discussed in Section 2.6), 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_2O 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_2O incubation (10,000 s) 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 the highest deuterium uptake (more than 45% at 10 s) 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 the lowest deuterium uptake (less than 3% at 10,000 s), which suggests that the core of this 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 (3–25%) 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.
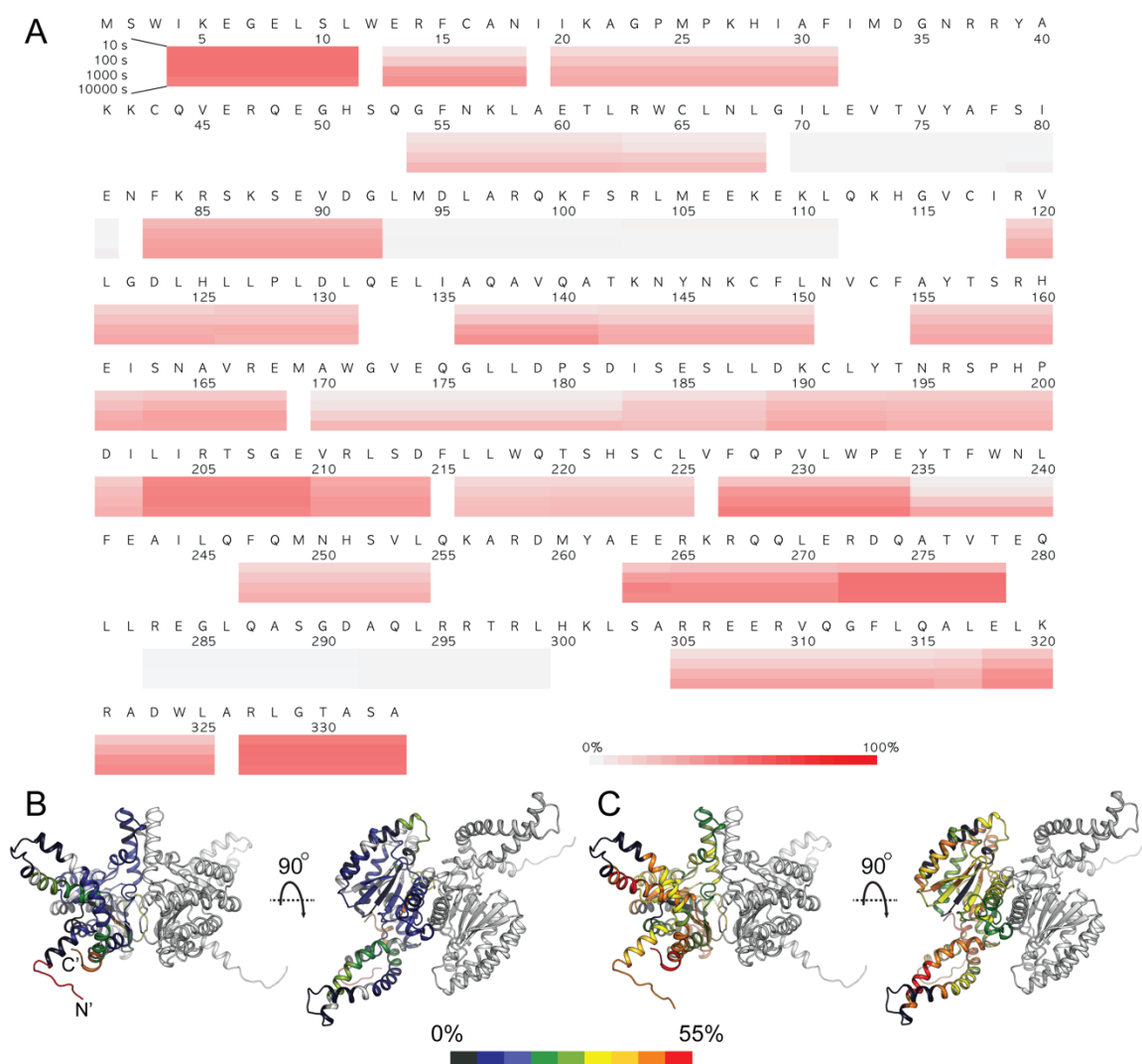
Figure 6. HDX-MS analysis of human DHDDS. (**A**) Deuterium uptake profiles. Incorporation of deuterium at 10, 100, 1000, and 10,000 sec are indicated by color-coded blocks underlining the amino acid sequence. The color legend shows the deuterium uptake levels. (**B**,**C**) The heat maps at 10 sec (**B**) and 10,000 s (**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 a 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–turn–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 *cis*-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 [9,10,12,33,34]. 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 [35]. Despite the importance of the heteromeric *cis*-prenyltansferases, the molecular mechanisms involved in subunit 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 subunit interactions. 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.
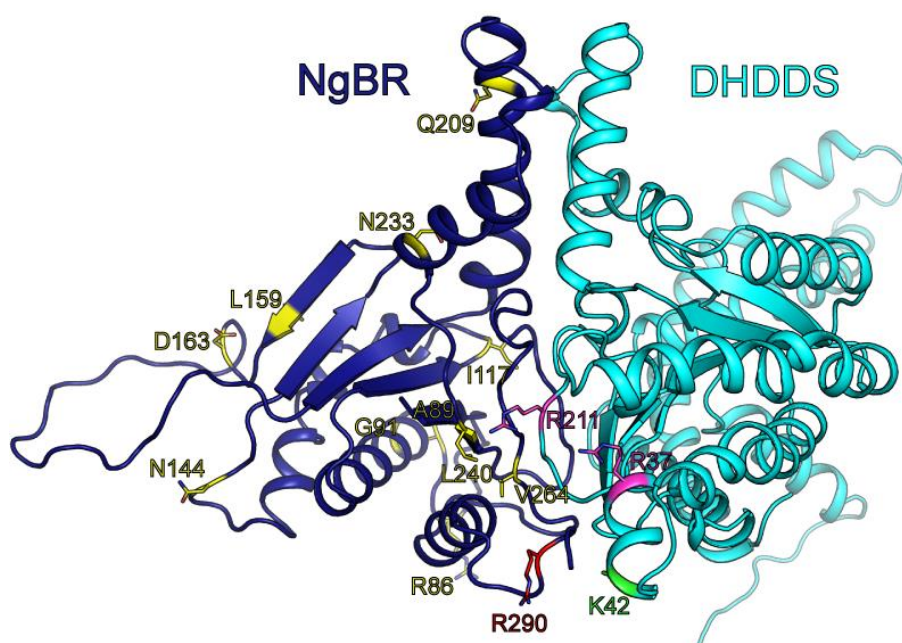

Figure 7. A model for the heterodimeric *cis*-prenyltransferase complex. The model of the heterodimeric complex is shown in cartoon representation. DHDDS and NgBR are colored cyan and blue, respectively. Positions with disease-associated mutations are shown as sticks and colored as follows: K42, related 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 [34] and a congenital glycosylation disorder [33] 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 [25], which was previously shown to play a critical role in the catalytic activity of the heteromeric human complex. Indeed, the location of R290 on a surface 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 [25] (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.

References

1. Schwarz, F.; Aebi, M. Mechanisms and principles of N-linked protein glycosylation. *Curr. Opin. Struct. Biol.* **2011**, *21*, 576–582. [CrossRef] [PubMed]

2. Helenius, A.; Aebi, M. Intracellular functions of N-linked glycans. *Science* **2001**, *291*, 2364–2369. [CrossRef]

3. Stowell, S.R.; Ju, T.; Cummings, R.D. Protein glycosylation in cancer. *Annu. Rev. Pathol.* **2015**, *10*, 473–510. [CrossRef] [PubMed]

4. Ng, B.G.; Freeze, H.H. Perspectives on Glycosylation and Its Congenital Disorders. *Trends Genet.* **2018**, *34*, 466–476. [CrossRef] [PubMed]

5. Freeze, H.H. Understanding human glycosylation disorders: Biochemistry leads the charge. *J. Biol. Chem.* **2013**, *288*, 6936–6945. [CrossRef] [PubMed]

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. [CrossRef]

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. [CrossRef]

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 dehydrodolichyl diphosphate synthase subunit Nus1. *Biochem. Biophys. Res. Commun.* **2019**, *515*, 621–626. [CrossRef]

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 dehydrodolichyl diphosphate synthase, is associated with autosomal-recessive retinitis pigmentosa in ashkenazi jews. *Am. J. Hum. Genet.* **2011**, *88*, 207–215. [CrossRef]

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. [CrossRef]

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 dehydrodolichol diphosphate synthase (DHDDS) activity. *Orphanet J. Rare Dis.* **2016**, *11*. [CrossRef]

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. [CrossRef] [PubMed]

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*, 222–226. [CrossRef] [PubMed]

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

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

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**. [CrossRef]

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

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

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. [CrossRef]

20. Webb, B.; Sali, A. Comparative protein structure modeling using MODELLER. *Curr. Protoc. Bioinforma.* **2016**. [CrossRef]

21. Chan, Y.T.; Ko, T.P.; Yao, S.H.; Chen, Y.W.; Lee, C.C.; Wang, A.H.J. Crystal Structure and Potential Head-to-Middle Condensation Function of a Z,Z-Farnesyl Diphosphate Synthase. *ACS Omega* **2017**, *2*, 930–936. [CrossRef] [PubMed]

22. Ko, T.P.; Huang, C.H.; Lai, S.J.; Chen, Y. Structure of undecaprenyl pyrophosphate synthase from Acinetobacter baumannii. *Acta Crystallogr. Sect. F Struct. Biol. Commun.* **2018**, *74*, 765–769. [CrossRef] [PubMed]

23. 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*. [CrossRef] [PubMed]

24. 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. [CrossRef] [PubMed]

25. 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. [CrossRef] [PubMed]

26. Hall, M. Size Exclusion Chromatography (SEC). In *Biopharmaceutical Processing: Development, Design, and Implementation of Manufacturing Processes*; Elsevier: Amsterdam, The Netherlands, 2018; pp. 421–432. ISBN 9780128125526.

27. 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. [CrossRef] [PubMed]

28. 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. [PubMed]

29. 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.

30. 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. [CrossRef]

31. 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. [CrossRef]

32. 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. [CrossRef] [PubMed]

33. 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. [CrossRef] [PubMed]

34. 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. USA.* **2018**, *115*, 11567–11572. [CrossRef] [PubMed]

35. 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*, e19022. [CrossRef] [PubMed]