Conserved intramolecular networks in GDAP1 are closely connected to CMT-linked mutations and protein stability

Aleksi Sutinen[1], Dirk Paffenholz[2], Giang Thi Tuyet Nguyen[1], Salla Ruskamo[1], Andrew E. Torda[2], and Petri Kursula[1,3,#]

[1]Faculty of Biochemistry and Molecular Medicine & Biocenter Oulu, University of Oulu, Oulu, Finland

[2]Centre for Bioinformatics, University of Hamburg, Hamburg, Germany

[3]Department of Biomedicine, University of Bergen, Bergen, Norway

[#]*Corresponding author* email: petri.kursula@uib.no

Keywords: Charcot-Marie-Tooth disease, GDAP1, GST superfamily, protein structure, neuropathy, conservation

ABSTRACT

Charcot-Marie-Tooth disease (CMT) is the most common inherited peripheral polyneuropathy in humans, and its subtypes are linked to mutations in dozens of different genes, including the gene coding for ganglioside-induced differentiation-associated protein 1 (GDAP1). The main GDAP1-linked CMT subtypes are the demyelinating CMT4A and the axonal CMT2K. Over a hundred different missense CMT mutations in the *GDAP1* gene have been reported. However, despite implications for mitochondrial fission and fusion, cytoskeletal interactions, and response to reactive oxygen species, the etiology of GDAP1-linked CMT is poorly understood at the protein level. Based on earlier structural data, CMT-linked mutations could affect intramolecular interaction networks within the GDAP1 protein. We carried out structural and biophysical analyses on several CMT-linked GDAP1 protein variants and describe new crystal structures of the autosomal recessive R120Q and the autosomal dominant A247V and R282H GDAP1 variants. These mutations reside in the structurally central helices α3, α7, and α8. In addition, solution properties of the CMT mutants R161H, H256R, R310Q, and R310W were analysed. All disease variant proteins retain close to normal structure and solution behaviour. All mutations, apart from those affecting Arg310 outside the folded GDAP1 core domain, decreased thermal stability. In addition, a bioinformatics analysis was carried out to shed light on the conservation and evolution of GDAP1, which is an outlier member of the GST superfamily. Many CMT mutation sites are highlighted in the bioinformatics analyses focusing on sequence conservation and entropy, and the analyses support the outlier nature of GDAP1 in the GST superfamily. A central role for the α6-α7 loop, within a conserved interaction network, is identified for GDAP1 protein stability. To conclude, we have expanded the structural analysis on GDAP1, strengthening the hypothesis that alterations in conserved intramolecular interactions may alter GDAP1 stability and function, eventually leading to mitochondrial dysfunction, impaired protein-protein interactions, and neuronal degeneration.

INTRODUCTION

The demand for sufficient energy supply *via* the aerobic process is elevated in neurons compared to other organelles and tissues, such as muscles [1]. Mitochondria are responsible for cellular respiration and linked to Ca^{2+} signalling and reactive oxygen species metabolism [2-5]. Since neurons depend on aerobic energy, their demand for oxidative phosphorylation is high, and 20% of the net oxygen consumed by the body is used for oxidative phosphorylation in neurons. Therefore, neurons are sensitive to alterations in mitochondrial function, and disruptions in mitochondrial dynamics can have severe consequences on neuronal functions.

Mitochondria are not isolated organelles, but interact with other cellular compartments, such as the endoplasmic reticulum, lysosomes, and peroxisomes, exchanging metabolites [2, 6, 7]. Mitochondria are renewed *via* fission and fusion, which are driven by proteins on the mitochondrial outer membrane (MOM), such as mitofusin 1 and 2 (MFN1/2), and the mitochondrial inner membrane (MIM), such as OPA1 and FIS1. Auxiliary proteins may bind to either MOM or MIM to enhance the process. The ganglioside-induced differentiation-associated protein 1 (GDAP1) is an integral MOM protein, proposed to have an auxiliary role in mitochondrial fission and fusion [8], possibly *via* redox-dependent interactions with cytoskeletal components [9]. However, the molecular basis of GDAP1 function and its exact relation to disease are currently not known.

Structurally, GDAP1 resembles glutathione *S*-transferases (GST), and it contains unique flexible loops [10, 11]. GDAP1 is constructed of two GST-like domains in the N and C terminus, followed by a transmembrane helix, which anchors the protein into the MOM. Structural data have shown a covalently bound dimer interface in GDAP1 [11, 12], and while dimerization is a common feature in catalytic GSTs [13], the GDAP1 dimer is formed differently [11]. Enzymatic activity of GDAP1 has not been convincingly demonstrated, nor has any substrate been identified *in vivo*. The members of the GST superfamily have similar fold properties, whereas distinct differences and low sequence conservation result in a diverse group of substrates; hence, it is possible that GDAP1 is an enzyme, but the substrate and reaction mechanism remain unidentified.

Increasing numbers of genes related to mitochondrial functions have been linked to neuropathophysiological conditions. Inherited polyneuropathies are a genetically and clinically diverse group of neurodegenerative diseases, which affect the outer motor and sensory neurons in the peripheral nervous system (PNS) [14, 15], the most common being Charcot-Marie-Tooth disease

(CMT). Clinical profiling divides CMT into three classes: demyelinating, axonal, and intermediate [16, 17]. The phenotype often implies insufficient mitochondrial fission and fusion, and mitochondria appear fragmented and elongated [18]. The etiology of CMT is linked to the hereditary pattern, whereby the autosomal recessive form has an earlier onset and more severe symptoms than the autosomal dominant form [19-21]. In the case of GDAP1, both autosomal dominant (axonal type CMT2) and recessive (demyelinating type CMT4) modes of inheritance are found. The severity is often correlated with the location of the causative mutation in the protein. The *GDAP1* gene is one of the most common missense mutation targets linked to CMT [8, 22, 23]. GDAP1 is ubiquitously expressed in tissues, but most of the expression is confined to neuronal tissues [8, 24]. The most accurate structural data thus far cover the dimeric core GST-like domain of human GDAP1, including the GDAP1-specific insertion [11]. In addition, a structure of a construct missing the large GDAP1-specific insertion is available in monomeric form [10]. In full-length GDAP1, an amphipathic extension – originally termed the hydrophobic domain – links the transmembrane helix to the GST-like domain.

GSTs often contribute to mechanisms of neurodegenerative disease [25, 26]. GST superfamily members function in prokaryotic and eukaryotic metabolism by utilizing reduced glutathione to catalyse a range of chemically diverse reactions. In comparison to other enzyme superfamilies, GSTs are unique in that sequence conservation appears to be driven by fold stability instead of catalytic features, as reflected in the broad spectrum of GST substrates [27, 28]. Using X-ray crystallography and complementary biophysical and computational techniques, we carried out structural analysis on selected GDAP1 mutants linked to CMT. We also analysed GDAP1 sequence conservation to investigate its GST-linked ancestry and to get clues into its molecular function and the relationship between conserved residue interaction networks and disease mutations.

MATERIALS AND METHODS

Recombinant protein production and purification

The GDAP1Δ303-358 and GDAP1Δ319-358 constructs, with an N-terminal His_6 tag and a Tobacco Etch Virus (TEV) protease digestion site, for producing soluble recombinant human GDAP1 in *E. coli*, have been described [11]. The point mutations R120Q, R161H, A247V, H256R, and R282H were generated in GDAP1Δ303-358, and the mutations R310Q and R310W in GDAP1Δ319-358, by a site-directed mutagenesis protocol with Pfu polymerase. All constructs were verified by DNA sequencing.

Recombinant GDAP1 variants were expressed in *E. coli* BL21(DE3) using autoinduction [29], and purified as described [11]. Briefly, GDAP1 was separated from the lysate by Ni^{2+}-NTA chromatography, and the affinity tag was cleaved using TEV protease. Another Ni^{2+}-NTA affinity step removed the tag and TEV protease. Size exclusion chromatography (SEC) was performed on a Superdex 75 10/300 GL increase column (Cytiva) using 25 mM HEPES (pH 7.5) and 300 mM NaCl (SEC buffer) as mobile phase. SEC peak fractions were analysed with SDS-PAGE and concentrated by centrifugal ultrafiltration.

X-ray crystallography

Mutant GDAP1Δ303-358 crystals were obtained using sitting-drop vapour diffusion at +4 °C. Proteins were mixed with mother liquor on crystallisation plates using a Mosquito LCP nano-dispenser (TTP Labtech). The protein concentration was 10-30 mg/ml in 75 nl, and 150 nl of reservoir solution were added. R120Q crystals were obtained in 0.1 M succinic acid, 15% (w/v) PEG 3350. A247V crystals were obtained in 0.15 M *DL*-malic acid (pH 7.3), 20% (w/v) PEG3500. R282H crystals were obtained in 0.1 M succinic acid; 20% (w/v) PEG 3350. For cryoprotection, crystals were briefly soaked in a mixture containing 10% PEG200, 10% PEG400, and 30% glycerol, before flash cooling in liquid N_2.

X-ray diffraction data were collected at the PETRA III synchrotron (DESY, Hamburg, Germany) on the P11 beamline [30, 31] and the EMBL/DESY P13 beamline [32] at 100 K and processed using XDS [33]. The structure of wild-type GDAP1Δ303-358, PDB entry 7ALM [11], was used as the search model for molecular replacement in Phaser [34]. The models were refined using Phenix.Refine [35] and rebuilt using COOT [36]. The structures were validated using MolProbity [37]. The data processing and structure refinement statistics are in **Table 1**, and the refined coordinates and structure factors were

deposited at the Protein Data Bank with entry codes 7B2G (R120Q), 8A4J (A247V), and 8A4K (R282H).

Table 1. Data processing and refinement statistics. *Data in parentheses refer to the highest-resolution shell.*

Variant	**R120Q**	**A247V**	**R282H**
Space group	$P6_322$	$P2_12_12_1$	$P2_12_12_1$
Unit cell	148.1, 148.1, 114.538 Å 90, 90, 120 °	73.3, 115.7, 115.9 Å 90, 90, 90 °	73.3, 113.4, 115.2 Å 90, 90, 90 °
Resolution range (Å)	100-3.0 (3.1-3.0)	50-2.68 (2.84-2.68)	50-1.95 (2.07-1.95)
Completeness (%)	100 (100)	99.7 (99.3)	99.8 (99.9)
Redundancy	38.7 (39.0)	13.1 (13.5)	6.7 (6.7)
⟨I/σI⟩	17.0 (1.0)	16.9 (1.9)	15.4 (0.7)
R_{sym} (%)	27.3 (505.3)	8.8 (162.7)	4.9 (273.8)
R_{meas} (%)	27.7 (511.9)	9.1 (169.0)	5.3 (296.9)
$cc_{1/2}$ (%)	99.9 (48.0)	99.8 (81.4)	99.9 (41.6)
R_{cryst}/R_{free} (%)	23.0/26.5	25.2/28.3	22.5/25.1
RMSD bond lengths (Å)	0.002	0.015	0.014
RMSD bond angles (°)	0.5	1.5	1.4
MolProbity score / percentile	1.36 / 100^{th}	2.04 / 97^{th}	1.63 / 92^{nd}
Ramachandran favoured/outliers (%)	95.1/0.00	96.4 / 0.2	97.9 / 0.4
PDB entry	7B2G	8A4J	8A4K

Modelling

A model for full-length human GDAP1 was obtained from AlphaFold2 [38]. In addition, missing loops of the human wild-type GDAP1 crystal structure were built with CHARMM-GUI [39, 40]. The two models were used to shed light on conformational changes and sequence conservation in GDAP1. Earlier structure-based bioinformatics results [12] were analysed further with respect to the mutational spectrum of GDAP1.

Synchrotron small-angle X-ray scattering

The structure and oligomeric state of the GDAP1 mutants were analysed with SEC-coupled small-angle X-ray scattering (SAXS). SEC-SAXS experiments were performed on the SWING beamline [41] (SOLEIL synchrotron, Saint Aubin, France). Samples were dialyzed against SEC buffer and centrifuged at >20000 g for 10 min at +4 °C to remove aggregates. 100 µl of each protein sample at 1.7-36 mg/ml were injected onto a BioSEC3-300 column (Agilent), run at a 0.3 ml/min flow rate. SAXS data were collected at +15 °C, over a q-range of 0.003–0.5 $Å^{-1}$. SAXS data analysis, processing, and modelling were done in ATSAS 3.0 [42]. Scattering curves were analysed and particle dimensions determined

using PRIMUS [43] and GNOM [44]. Chain-like *ab initio* models were generated using GASBOR [45], dummy atom models were built with DAMMIN [46], and model fitting to data was analysed with CRYSOL [47].

Synchrotron radiation circular dichroism spectroscopy

Synchrotron radiation circular dichroism (SRCD) spectra were collected from 0.5 mg/ml samples on the AU-SRCD beamline at the ASTRID2 synchrotron (ISA, Aarhus, Denmark). The samples were prepared in a buffer containing 10 mM HEPES pH 7.5 and 100 mM NaF, equilibrated to room temperature, and applied into 0.1-mm closed circular quartz cuvettes (Suprasil, Hellma Analytics). SRCD spectra were recorded from 170 nm to 280 nm at +25 °C. Three scans per measurement were repeated and averaged. The spectra were processed using CDToolX [48].

Thermal stability

Thermal stability of GDAP1 variants was studied by nanoDSF using a Prometheus NT.48 instrument (NanoTemper), in SEC buffer. Tryptophan fluorescence was excited at 280 nm, and emission was recorded at 330 and 350 nm, while the samples were heated from +20 to +90 °C at a rate of 1 °C/min. Changes in the fluorescence ratio (F_{350}/F_{330}) were used to determine apparent melting points. The data were analyzed using Origin (OriginLab Corporation, Northampton, MA, USA).

Sequence entropy

Starting from the human GDAP1 reference sequence (NP_061845.2), iterative PSI-BLAST [49] searches were performed, initially accepting sequences with an e-value of 10^{-99} or smaller. From the first search, but the second iteration of PSI-BLAST, sequences were arbitrarily collected with $e < 10^{-07}$. This resulted in 5986 sequences. This original data set was further subdivided into a first set that contains only sequences labeled GDAP1, a second set containing only sequences labeled GDAP1L1, and a third set containing only GST-labeled sequences. In this process, unnamed sequences, hypothetical sequences, and those from other proteins were discarded. Although GDAP1 and GDAP1L1 were initially separated, they will here be treated as one group (GDAP1/GDAP1L1). This is justifiable, since GDAP1L1 is by far the closest paralogue to GDAP1, and together they form an outlier group in the GST family. This was also a necessity, because the total number of sequences was too small compared to the GST data set. Even after merging the GDAP1/GDAP1L1 sets, there are still about twice as many sequences in the GST data set.

Lastly, all sets were combined to calculate the entropy across all data sets (GDAP1/GDAP1L1 + GST), which included 5065 sequences. For each set, a multiple sequence alignment was generated using MAFFT [50], and each sequence was mapped to a reference sequence. This was done by removing any columns, which correspond to a gap in the reference sequence. At this point, some information might get lost in the process. The first hit of human GDAP1L1 in the original data set was selected as reference. For the GST reference, a sequence was selected, for which there is a structure in the PDB database. This was not in the original data set and was added later (before alignment). **Supplementary Table 1** gives a detailed overview of the generated data sets.

For calculating the per-site entropy for a multiple sequence alignment, we used

$$f_i = \frac{count(aa)}{N}$$

$$S = \sum_{i=1}^{N} |f_i \log (f_i)|$$

Where f_i is the normalized frequency for each amino acid in each column of a MSA, and N is the number of sequences in the MSA, hence the number of rows. The basis of the logarithm depends on the alphabet (number of different symbols) and if gaps are treated as valid characters or should be ignored. For example, for protein sequences, the size of the alphabet is 20 without gaps and 21 with gaps. The entropy score S ranges from 0 (only one residue is present at that position) to 1 (all 20 residues are equally represented at the position). In this study, positions with S > 0.2 are considered variable, whereas those with S < 0.2 are considered conserved. Highly conserved positions are those with S < 0.1. Gaps are treated as non-valid characters.

Kullback-Leibler divergence

The above chapter focused on the similarities between GDAP1/GDAP1L1 and GST. In addition to the similarities, differences are of particular importance. In order to capture these differences, the Kullback-Leibler (KL) divergence was used. The KL divergence is a statistical distance, a measure to quantify the differences between two probability distributions [51]. The KL divergence can be calculated as the sum of probability of each event in probability distribution P multiplied by the log of the probability of the event in probability distribution Q over the probability of the event in P [52]. Similarly to entropy calculation, the KL divergence score is derived from the normalized amino acid frequencies as probabilities:

$$D_{KL}(P\|Q) = \sum_{i}^{N} P_i \, log\left(\frac{P_i}{Q_i}\right)$$

The intuition for the KL divergence score is that when the probability for an event from P is large, but the probability for the same event in Q is small, there is a large divergence. When the probability from P is small and the probability from Q is large, there is also a large divergence, but not as large as the first case. Importantly, the KL divergence score is asymmetrical:

$$D_{KL}(P\|Q) \neq D_{KL}(Q\|P)$$

Hence, the KL divergence is calculated with one file as the P distribution and also as the Q distribution. In contrast to the above section, gaps are treated as valid characters. The goal is to find residues or regions that are highly conserved in both multiple sequence alignment files, but differ from each other. This should further shed light on the fundamental differences between GDAP1 and conventional GSTs.

Phylogenetic analyses

To model the evolutionary history of GDAP1, the program MrBayes [53] was used. Visualizations were made with Interactive Tree of Life (iTOL) [54]. MrBayes is a Bayesian inference model that uses Markov chain Monte Carlo (MCMC) methods to estimate the posterior probability distribution of model parameters. Inference of phylogeny is based upon the posterior probability distribution of the constructed trees T_1, T_2, ..., T_n. The posterior probability of the i-th tree is calculated using Bayes's theorem:

$$\mathrm{P}[\mathrm{T}_i \mid X] = \frac{P[X \mid T_i\,]\,P[T_i\,]}{\sum_{i=1}^{N} P[X \mid T_i\,]\,P[T_i\,]}$$

$P[T_i \mid X]$ is the posterior probability of the i-th tree, $P[X \mid T_i]$ is the likelihood of the i-th tree, and $P[T_i]$ is the prior probability of the i-th tree. The denominator is a summation over all possible trees N. The likelihood $P[X \mid T_i]$ is calculated as a multidimensional integral over all possible combinations of branch lengths and substitution model.

The trees (with branch lengths) that were sampled by the MCMC procedure of each analysis were saved. The final tree is a consensus tree, an average over all calculated trees [55]. The settings used for the calculations are shown in **Table 2**. `Lset` is used to define the structure of the evolutionary

model and `prset` is required to define the prior probabilities of the model. In this case (`prset aamodel=mixed`) the program integrates over a predetermined set of fixed rate matrices. `Ngen` sets the number of generations for which the analysis will be run. `Nruns` is the number of simultaneous, completely independent analyses starting from different random trees. `Nchains` is the number of chains used in each analysis. By setting `nchains=4`, MrBayes will use 3 heated chains and one "cold" chain. `Relburnin=yes` activates “burning”, which means that MrBayes will discard the first 25 % samples (`burninfrac=0.25`) from the cold chain. `Beagle` is a high-performance library for parallel computation.

Table 2. Parameter settings for MCMC analyses.

prset aamodel= mixed
lset nucmodel=protein
Lset nst=1
Set autoclose=yes
Set usebeagle=yes
Set beagledevice=cpu
Mcmcp nruns=4 nchains =4
Mcmcp ngen= 2000000
Mcmp relburning= yes burninfrac=0.25
Mcmcp printfreq=1000000
Mcmcp savebrlens= yes

For phylogenetic analyses, PSI-BLAST was run as above, and from the 5986 sequences, GDAP1 sequences were extracted, resulting in 799 sequences, and the first hit of human GDAP1L1 (NP_001243666.1) was determined as the reference sequence for the GDAP1L1 set, and the first bacterial GST (TNE50161.1) was set as the bacterial GST reference sequence. Additionally, from the bacterial GST BLAST search, a sequence with a low e-value was set as the root sequence (WP_173192278.1). Since the BLAST search quickly arrives at bacterial GSTs, but it is also of interest where eukaryotic GSTs are found in the phylogenetic history, a human GST (PDB: 1PKZ), for which there is a crystal structure available, was set as a reference for eukaryotic GST. With each of these reference sequences (GDAP1L1, bacterial GST, eukaryotic GST), another BLAST search was performed, and the first 1000 hits of each BLAST search were extracted.

To study the relation of GDAP1 to archaea, a BLAST search with the GDAP1 reference was performed, but the output was restricted to archaeal sequences. With the first hit, an archaeal GST (MAE98075.1), another BLAST search was performed and the first 1000 hits were extracted, defining the last dataset (second_GST_set). Table 3 shows an overview over the separated datasets. Bacterial

GST and eukaryotic GSTs were combined into the "first_GST_set", to avoid oversampling. For each dataset, a multiple sequence alignment (MSA) was generated using MAFFT [Katoh et al., 2002]. Each set was then reduced to 100 sequences with the *reduce* program (*https://github.com/andrew-torda/sequtils*). *Reduce* generates an evenly distributed set of sequences by scanning the distance matrix and discarding very closely related sequences. The remaining 100 sequences of each dataset were combined, possible duplicated sequences were removed, and finally, together with the root sequence, re-aligned.

Table 3. Dataset overview.

Dataset	**# Sequences**
GDAP1	799
GDAP1L1	1000
First_GST_set	1000 + 1000 = 2000
Second_GST_set	1000
Final_set	400 (397 after removing duplicates)

RESULTS

Building upon earlier work on GDAP1 structure [10-12], we focused here on several CMT-linked variants that reside on different secondary structure elements on the GDAP1 crystal structure. While we earlier specifically looked at R120W and H123R on helix α3 [12], here we produced and characterised the variants R120Q, R161H, A247V, H256R, R282H, R310Q, and R310W. The stability and solution structure were studied for all variants, while crystal structures were determined for three of them: R120Q, A247V, and R282H. The location of the studied mutation sites in the wild-type GDAP1 structure is shown in **Fig. 1A-B**.

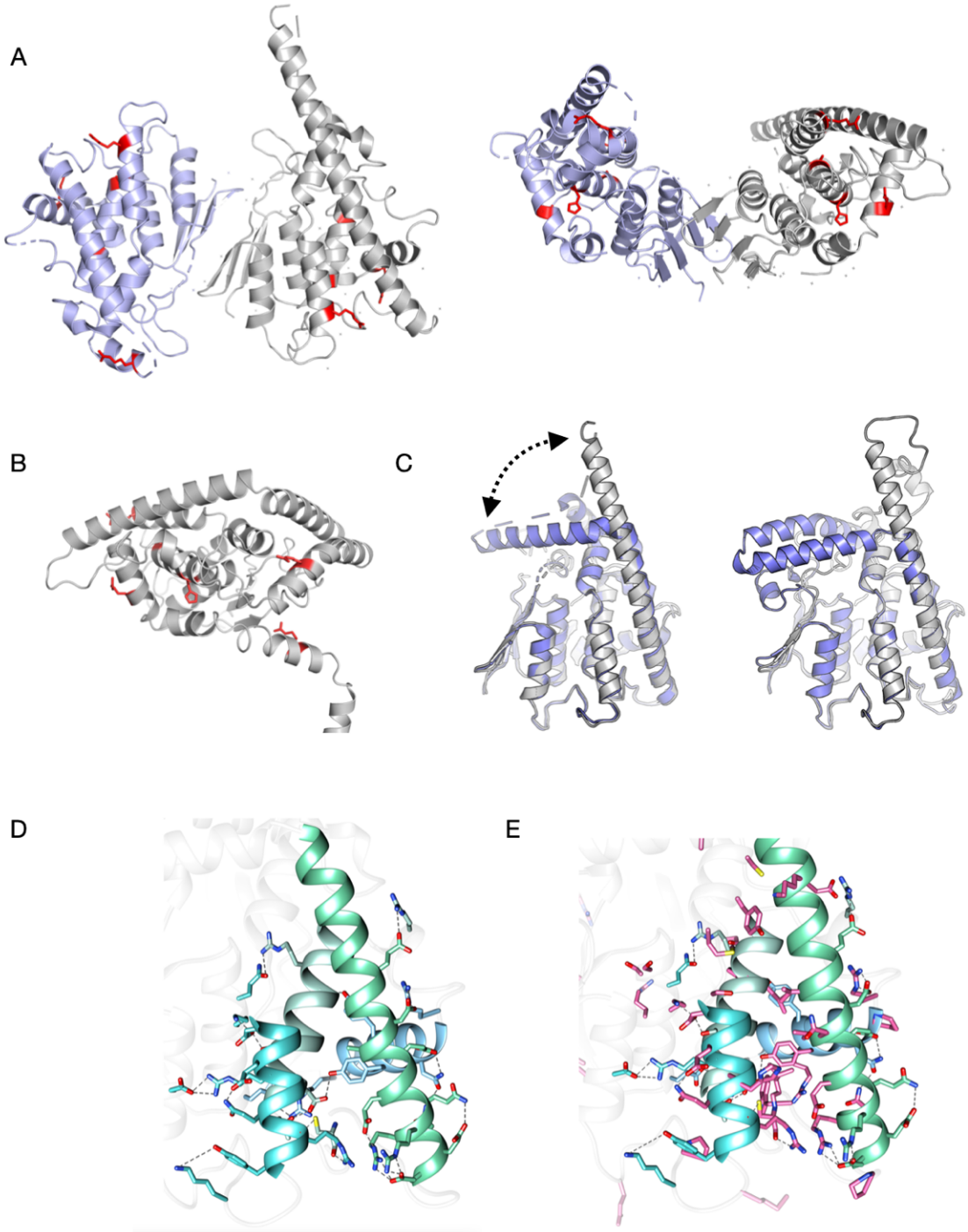


Figure 1. Overall structure of GDAP1. *A. Location of the mutations studied here, mapped onto the crystal structure of wild-type GDAP1 [11] in two different orientations. B. The mapping of the mutations onto the AlphaFold2 model, to include those not visible in the crystallised construct (R161H, R310Q, R310W). C. Open/close conformations involving the long helix α6 have now been observed both experimentally (left) and using structure prediction (right). D. Hydrogen bonding network of residues on the core helices of GDAP1. E. Same view as E, but known sites of CMT mutations have been added in magenta. CMT mutations are clustered on the core helices.*

Helices α3, α6, α7, and α8 form the basis of GDAP1 intramolecular networks

The majority of CMT-linked missense mutations in GDAP1 are located within the vicinity of the hydrophobic clusters of the GST-like domains and the dimer interface [11], and the variants may induce changes in intramolecular hydrogen bonding networks [12]. In addition to the R120W and H123R studied earlier [12], we now determined three new mutant crystal structures: R120Q, A247V, and R282H. These mutations reside in helices α3 (R120Q), α7 (A247V), and α8 (R282H), which are core elements of the GDAP1 fold. We shall first look at the central helices regarding GDAP1 folding.

The GST-like core fold of GDAP1 is supported by the α7 helix, surrounded by helices α3, α6, and α8. The helix α3 is connected to α6 *via* the α6-α7 loop, and Cys240 in this loop – itself being a CMT mutation site – is central to many interactions. The α6 helix can either extend or turn back towards the dimer interface, as seen in earlier crystal structures and models [12]. A new model built here indicates that the extended conformation is predictable (**Fig. 1C**). Open/close movements of α6 can be functionally relevant for GDAP1 interactions with other proteins, such as cofilin and tubulin [9]. The α8 helix is positioned perpendicular to the others, and its orientation could be related to the transmembrane helix position, as it is expected to face the membrane surface. Together, these helices form an intramolecular network of polar (**Fig. 1D**) and non-polar contacts, and many CMT-linked missense mutations are found on these helices (**Fig. 1E).**

There are 15 designated hydrogen bond contacts or ionic interactions between helices α3, α6, α7, and α8. Helix α7, which is central to the GDAP1 fold, makes only a few hydrogen bonds to the surrounding helices. This indicates that hydrophobic residue clustering houses the α7 helix in the core of the fold; this is in line with centrality analyses of the GDAP1 fold, showing helix α7 to be the most central part of the 3D structure [12]. Some of the disease mutations on these helices correspond to solvent-exposed residues. Accordingly, helices α3, α6, and α8 show more polar interactions and higher flexibility compared to α7.

Structural effects of the individual CMT mutations

Earlier, we analysed the effects of R120W and H123R, both on helix α3 [12]. Here, we extend the crystal structure analyses to three more CMT mutations: R120Q, A247V, and R282H (**Fig. 2, Table 1**).

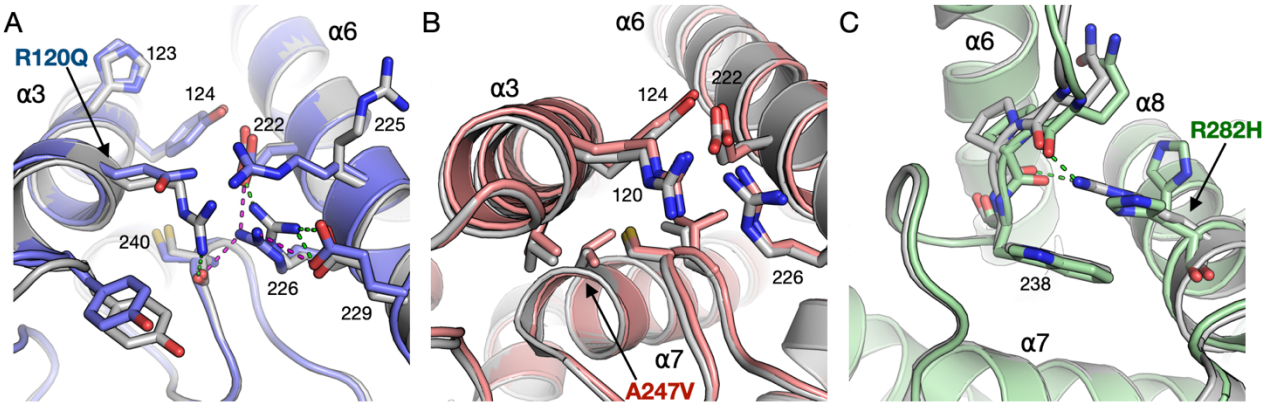


Figure 2. Structural details for each point mutation. *A. Immediate environment of the R120Q mutation. B. Effects of the A247V mutation. C. The R282H mutation.*

The R120Q mutation is located close to the N-terminal end of the α3 helix. Due to the mutation, residue 120 loses contact with the α6-α7 loop tip at Cys240, whereby the hydrogen bond between the Arg120 side chain and the backbone carbonyl of Cys240 is lost (**Fig. 2A**). This is further linked to alterations in surrounding side chain conformations**.**

The A247V mutation site resides at the N terminus of helix α7, tightly surrounded by helix α3 and the α6-α7 loop and having van der Waals contacts to Val121, Tyr124, Cys240, and Thr245. Comparing to the wild-type protein, changes in the crystal structure are minimal, but an overall movement of surrounding protein segments is caused by the presence of Val in this position. Both helices α3 and α6 move slightly away, without altering hydrogen-bonding patterns (**Fig. 2B**).

R282H is located on helix α8, and the side chain of Arg282 points inwards in the wild-type GDAP1 structure; it is stacked against Trp238 and makes hydrogen bonds to the backbone carbonyl groups of residues 236 and 237 in the α6-α7 loop (**Fig. 2C**). All these interactions are lost upon the mutation, and the His residue in the mutant is observed in a double conformation, making no side-chain hydrogen bond contacts.

Conformation and stability in solution

Computational predictions suggested an overall destabilising effect of CMT-linked mutations in the GDAP1 protein [12]. Similarly, protein destabilisation was observed experimentally for the myelin protein P2 in the context of all identified CMT mutations [56, 57]. A comparative analysis of seven GDAP1 variants in solution was therefore carried out here, to support the crystal structures and other complementary data from the current and earlier studies, and to identify general trends linking CMT mutations and GDAP1 stability *in vitro*. While the 3D shape and dimensions were studied using SAXS, SRCD was employed to follow secondary structure content and nanoDSF to compare thermal stability.

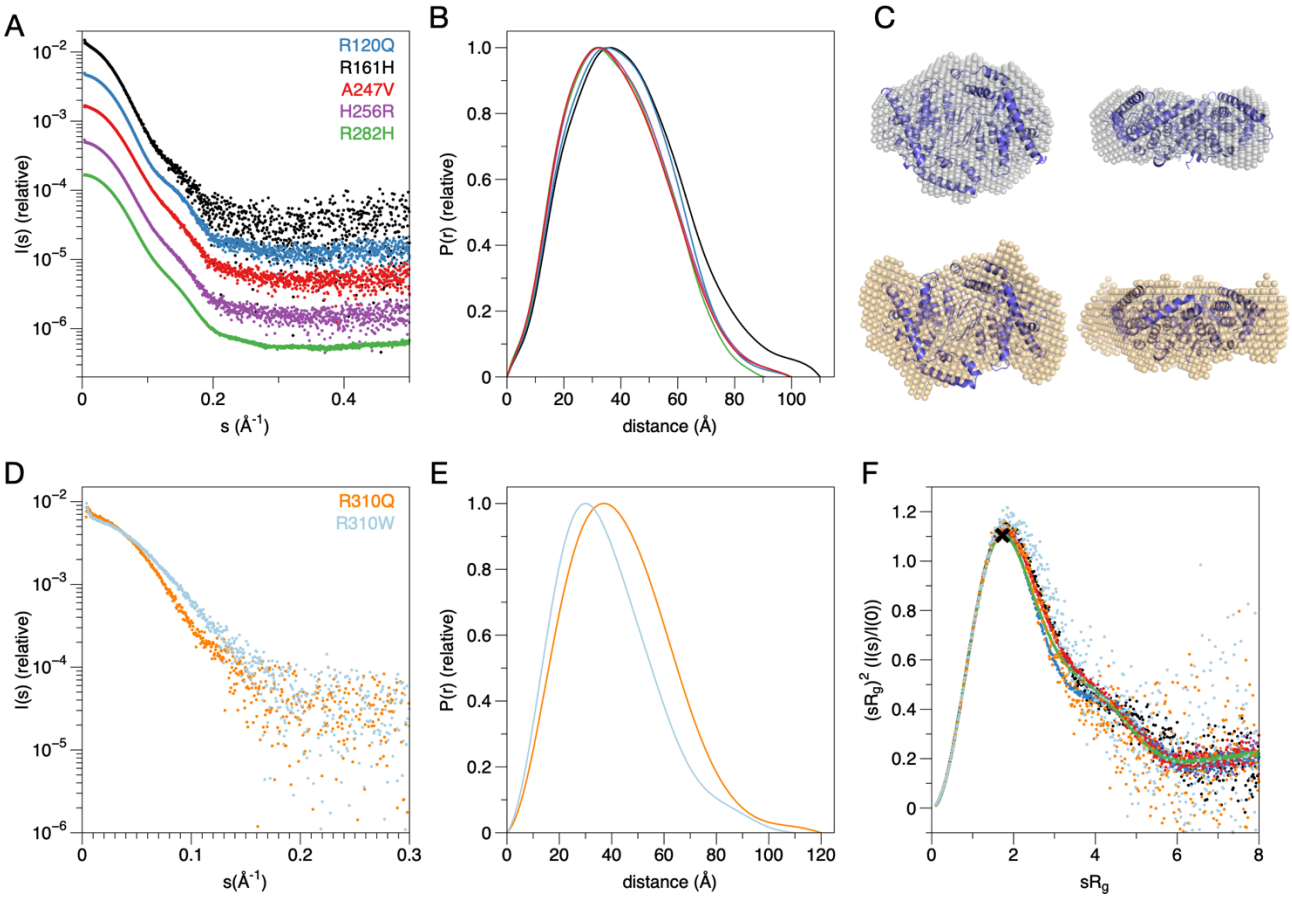


Figure 3. SAXS analysis. *A. SAXS curves for the mutants in the GDAP1Δ303-358 construct. B. Distance distributions for the curves in A. C. Dummy atom model of R282H superimposed with a structure based on the collapsed conformations of wild-type GDAP1Δ295-358 [12] (top). Dummy atom model of the R310Q mutant in GDAP1Δ319-358, superimposed on the same structure. D. SAXS curves for the mutants in the GDAP1Δ319-358 construct. E. Distance distributions for the curves in D. F. Dimensionless Kratky plots for all constructs show similar levels of rigidity and globularity.*

Table 4. SAXS parameters. *Data for wild-type GDAP1 as well as the monomeric mutation Y29EC88E are taken from [11]. The MW estimate corresponds to the Bayesian estimate from PRIMUS.*

Variant	Δ303-358 dimer/ Y29EC88E monomer	R120Q	R161H	A247V	H256R	R282H	R310Q main peak	R310W main peak	Δ319-358 dimer/monomer
R_g (Guinier) (Å)	30.7/24.5	30.9	33.4	29.8	30.3	29.3	33.3	29.3	34.7/27.1
R_g (GNOM) (Å)	30.6/24.59	30.7	33.6	30.0	30.3	29.4	33.6	29.9	33.5/27.3
D_{max} (Å)	99/86.7	99.7	110	100	99.2	91.0	120	110	107.9/89.6
V_p (nm^3)	105.8/58.7	102.0	127.4	93.1	95.8	89.4	101.2	58.2	129.5/69.4
Estimated MW (kDa)	72.4/35.4	67.1	91.2	62.4	62.4	59.5	94.2	58.2	94.2/46.7

The SAXS analysis of GDAP1Δ303-358 showed that the protein particle dimensions in solution correspond to a dimer, and the scattering profiles showed only minor shape differences, ruling out large-scale conformational differences or aggregation (**Fig. 3A**). The largest deviation was observed for R161H, which had a larger R_g than the other variants, suggesting a more open structure. Also R120Q had a slightly different conformation, most clearly seen in the distance distribution. The pair distance distribution function showed that the maximum particle dimension in all samples was ~90-110 Å, which corresponds to a dimer (**Fig. 3B**). *Ab initio* modelling, here carried out on the R282H mutant, which was measured at the highest concentration, agreed with the presence of a dimeric GDAP1, corresponding to the conformation of the dimer we previously showed to fit the solution SAXS data for wild-type GDAP1 [12] (**Fig. 3C**).

For the two mutations affecting Arg310, we used the longer construct GDAP1Δ319-358, and both R310Q and R310W eluted as two peaks in SEC-SAXS, corresponding to a dimer and monomer (**Fig. 3D-E**). This behaviour has previously been linked to protein concentration, in that monomers can be observed at low concentration [11], and the Arg310 mutants were here studied at lower concentration than all the other mutants. R310Q mainly eluted as a dimer, while the highest peak for R310W was monomer; since the protein concentrations were very similar, this result indicates that R310W may affect GDAP1 dimerisation. The R310Q dimer was more elongated than R282H of the shorter construct (**Fig. 3C**), reflecting the presence of the C-terminal amphipathic domain in the construct. Comparing the Kratky plot for all variants studied here, it is evident that they all are rigidly folded, with little flexibility (**Fig. 3F**).

For thermal stability and secondary structure analysis, nanoDSF and SRCD were performed. The SRCD spectra showed nearly identical secondary structure composition for all variants and wild-type

GDAP1. Thus, none of the studied mutations affect the overall folding of GDAP1 (**Fig. 4A-B**). The DSF experiment revealed a ~1-12 °C decrease in apparent melting temperature for all core domain mutants, compared to the wild-type GDAP1Δ303-358 (**Fig. 4C**). This is similar to the previously studied R120W and H123R mutants [12]. All in all, the biophysical data indicate that the structural effects of all studied CMT variants in the core domain are local, and the mutations do not large-scale conformational changes at a level detectable with SRCD or SAXS. However, all studied mutations in the core domain caused a decrease in thermal stability, suggesting a breakdown of stabilising intramolecular interactions within the GDAP1 molecule. As the most dramatic example, the apparently harmless mutation A247V decreased the stability by 12.4 °C.

For the mutations at position 310, it was observed that both R310Q and R310W in fact increased the stability of GDAP1Δ319-358 by ~1 °C (**Fig. 4D**). This can be explained by the fact that the segment carrying these mutations is no longer part of the core domain, but rather likely to represent an α helix attached to the membrane surface, possibly linking the membrane thus to the core domain. Overall, one can say GDAP1 stability correlates with the dimer-monomer ratio observed in solution; the most destabilised mutants also show a higher fraction of monomeric protein on non-reducing SDS-PAGE (**Fig. 4E**), even though in the 3D structure, they are far from the dimer interface.

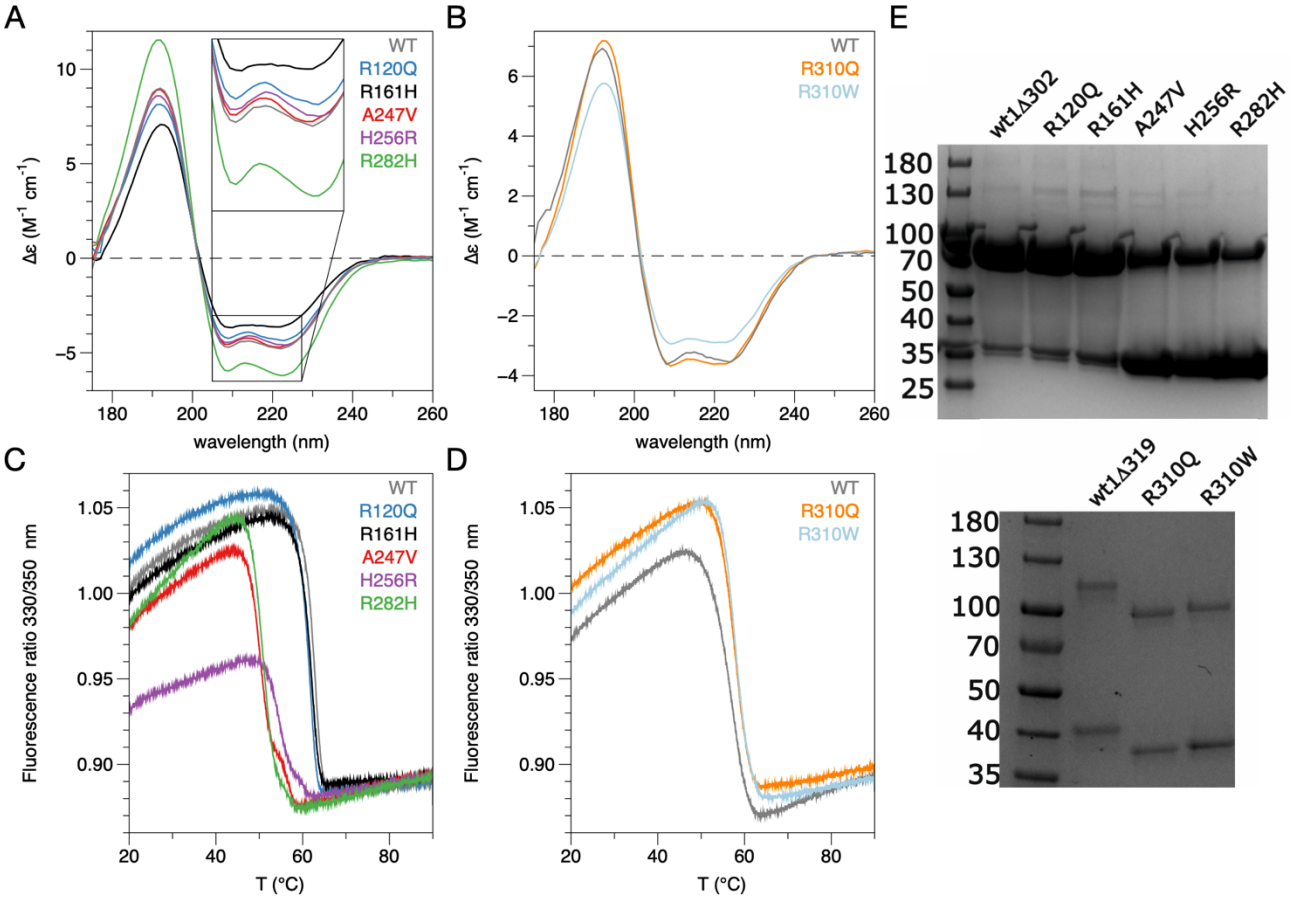


Figure 4. Folding and stability. *A,B. nanoDSF analysis for the mutants in GDAP1Δ303-358 (A) and GDAP1Δ319-358 (B). C. SRCD spectra for the mutants in the GDAP1Δ303-358 construct. D. SRCD spectra for the mutants in GDAP1Δ319-358. E. SDS-PAGE analysis of all studied variants.*

Table 5. nanoDSF apparent melting points. All values are average from 3 replicates.

Sample	**T_m (°C)**
GDAP1Δ303-358	62.6
R120Q	61.3
R161H	61.8
A247V	50.2
H256R	54.6
R282H	50.8
GDAP1Δ319-358	56.9
R310Q	57.9
R310W	57.8

Implications of the point mutations for the whole protein

The studied mutations cause subtle variations in hydrogen bond and van der Waals distances in nearby residues in the crystal structures, while there are no drastic structural differences, when the mutant structures are superposed to the wild-type crystal structure. However, when comparing the variation of the residues in helices α3, α6, α7, and α8, in comparison to the wild-type Cα-atoms, a specific pattern arises. We analysed the residues participating a hydrogen-binding network in the mutant structures (**Fig. 5A**), and the mapping shows that the residues pointing towards the dimer interface have only minor variations compared to the wild-type protein. In contrast, at the C-terminal end of helix α6, in all mutant crystal structures, the variation of the Cα atoms is high (**Fig. 5B**). This suggests that flexibility of helix α6 could arise from altered intramolecular contacts in the vicinity. The C-terminal end of the α6 helix, the conformation of which apparently is affected by the mutations studied here, is itself a target for multiple CMT mutations affecting Gln218, Va219, Glu222, Arg226, and Glu227 [19, 58-61].

We additionally investigated the more global mutational effects using bioinformatics tools. In our previous study [12], we analysed the wild-type GDAP1 core domain structure and the full-length AlphaFold2 coordinates with CUPSAT and MAESTRO [62, 63]. This analysis provided predicted ΔΔG values and geometry properties; the results from CUPSAT analyses are further depicted in **Fig. 5C**. Clearly, on the average most CMT mutation sites are predicted to cause destabilisation and the few residues predicted to be stabilised by mutations lie on the protein surface.

While the above analyses indicate an average effect that may destabilise GDAP1 structure, we looked at predictions for each mutation studied here in more detail (**Fig. 5D**). In essence, all substitutions to Arg120 are unfavourable; this shows that the interactions made by Arg120 are important for folding and stability. This is in line with our experimental data for both R120Q and R120W, which indicate minor changes in structure, but destabilisation of the fold. For Ala247, a mutation into Val is actually predicted to be slightly stabilising, indicating that such a replacement, in a tightly confined pocket and with small but long-range effects, is difficult for the prediction algorithm. In the case of Arg282, mutation to His is indeed one of the most destabilising variants in the prediction.

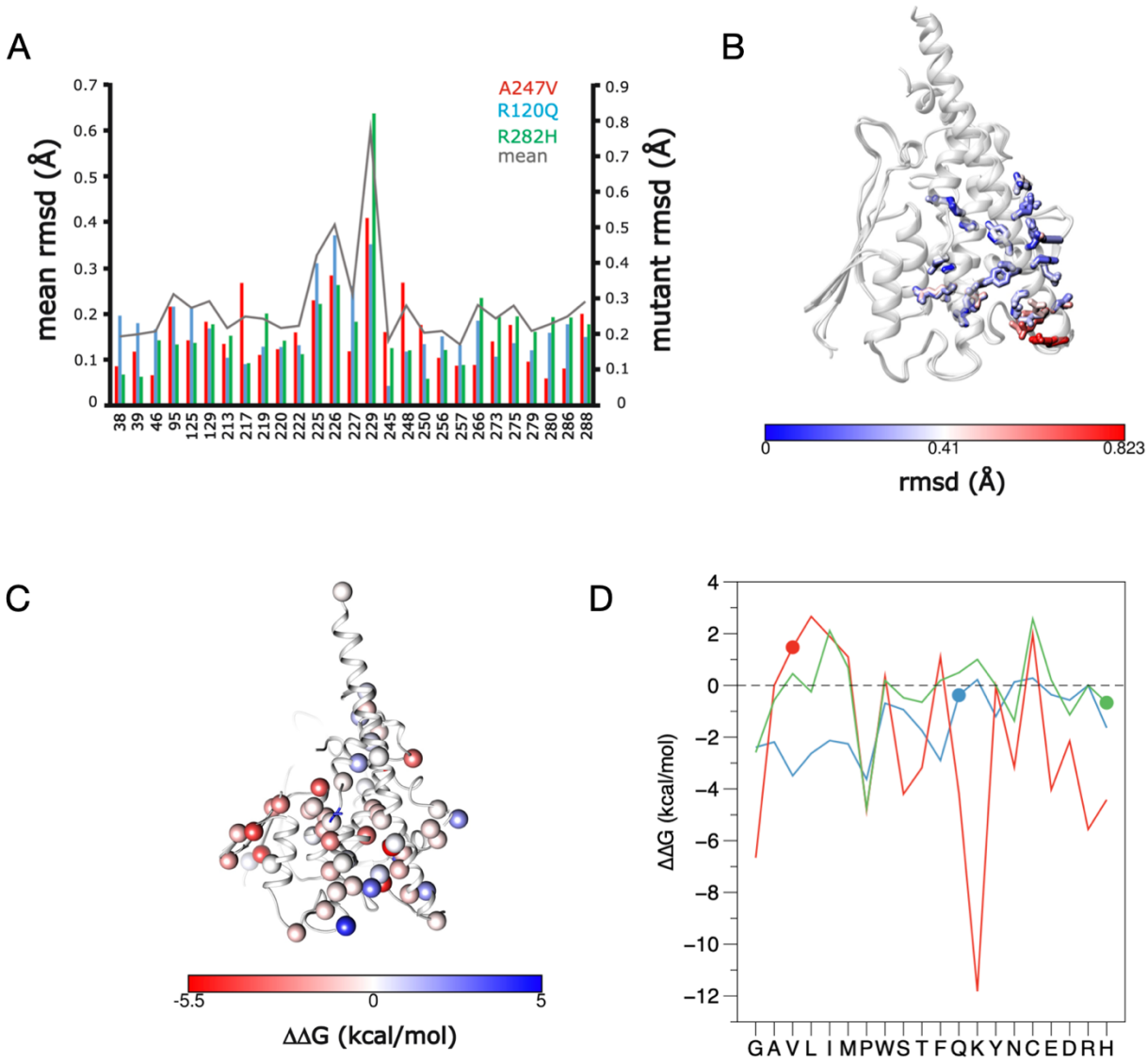


Figure 5. Structural bioinformatics. *A. The Cα deviation of each mutant vs. wild-type GDAP1 structure. Shown are only the residues participating in the hydrogen bonding network. B. Mapping the results onto the structure, it becomes evident that the C-terminal end of helix α6 deviates the most from wild-type GDAP1 on average. C. Average ΔΔG effect of mutations at CMT sites, as defined by CUPSAT. D. CUPSAT predictions for R120Q, A247V, and R282H being mutated into all possible amino acids. Note that a negative ΔΔG in CUPSAT means destabilisation. The mutations studied here are marked with spheres.*

Phylogenetic analysis of GDAP1 and the GST superfamily

Since GSTs are universally widespread across almost all organisms, the GST superfamily is inherently very large and diverse. There are massive differences between families and even within the subfamilies. Starting a BLAST search with the reference sequence will quickly end up in GSTs. This

leads to the fact that GSTs are oversampled compared to GDAP1 and GDAP1L1; respectively, GDAP1/GDAP1L1 are undersampled. Therefore, the phylogenetic tree calculates the evolutionary history of GSTs, rather than the evolutionary history of GDAP1. To avoid this, GDAP1, GDAP1L1 and GSTs were considered separately.

As a first step in a sequence-based bioinformatics approach, we carried out a phylogenetic analysis of the GST family with a specific focus on GDAP1 (**Fig. 6**). Intriguingly, based on large-scale sequence alignments, GDAP1 turns out to be more closely related to prokaryotic GSTs than eukaryotic proteins. In this situation, we must remember that sequence conservation levels in the GST superfamily are overall low. Based on the phylogenetic analysis, it is still not possible to depict the exact pathways of GDAP1 evolution. In general terms, one can say GDAP1 is at least as far from generic eukaryotic GSTs as it is from prokaryotic proteins. The analysis provides little clues as to any enzymatic function of GDAP1, but it does give more support to the findings that it has no conventional GST activity.

Fig. 6B and 6C show that eukaryotic GSTs split off from other GSTs very early in evolutionary history. This was not to be expected. In addition, **Fig. 6B** shows that the evolutionary history at this level could not be resolved (represented by a red line) despite exhausting CPU time (**Table 2**). One can see that it is a bifurcating tree. Bacterial GST sequences are shown in the right branch and eukaryotic GST sequences in the left branch. The remaining sequences in the tree are represented by the grey circle.

Fig. 6C shows the last level of the phylogenetic tree. Considering only this part, this tree is a bifurcating tree. Towards the right side there is a branching to bacterial GST sequences (purple) and in the left branching the GDAP1/GDAP1L1 sequences are found again. The figure also shows that orthologs of GDAP1 and GDAP1L1 are found in mammals, birds, amphibians, and fishes and likely orthologs of those genes in invertebrates as well.

A first assumption of the evolutionary history is based on the endosymbiotic theory and the current perception of the phylogenetic tree of life that archaea are evolutionarily closer to eukaryotes than bacteria [64]. Since GDAP1 is a protein of the outer mitochondrial membrane, it is possible that GDAP1 is closely related to archaeal proteins. Since the BLAST search with an archaeal sequence almost exclusively finds bacterial GST sequences, and the fact that the times could not be fully resolved, this tree does not make any assumptions whether GDAP1 are derived from archaea or

not. However, it does suggest early multiple gene duplication events that split up eukaryotic GST sequences.

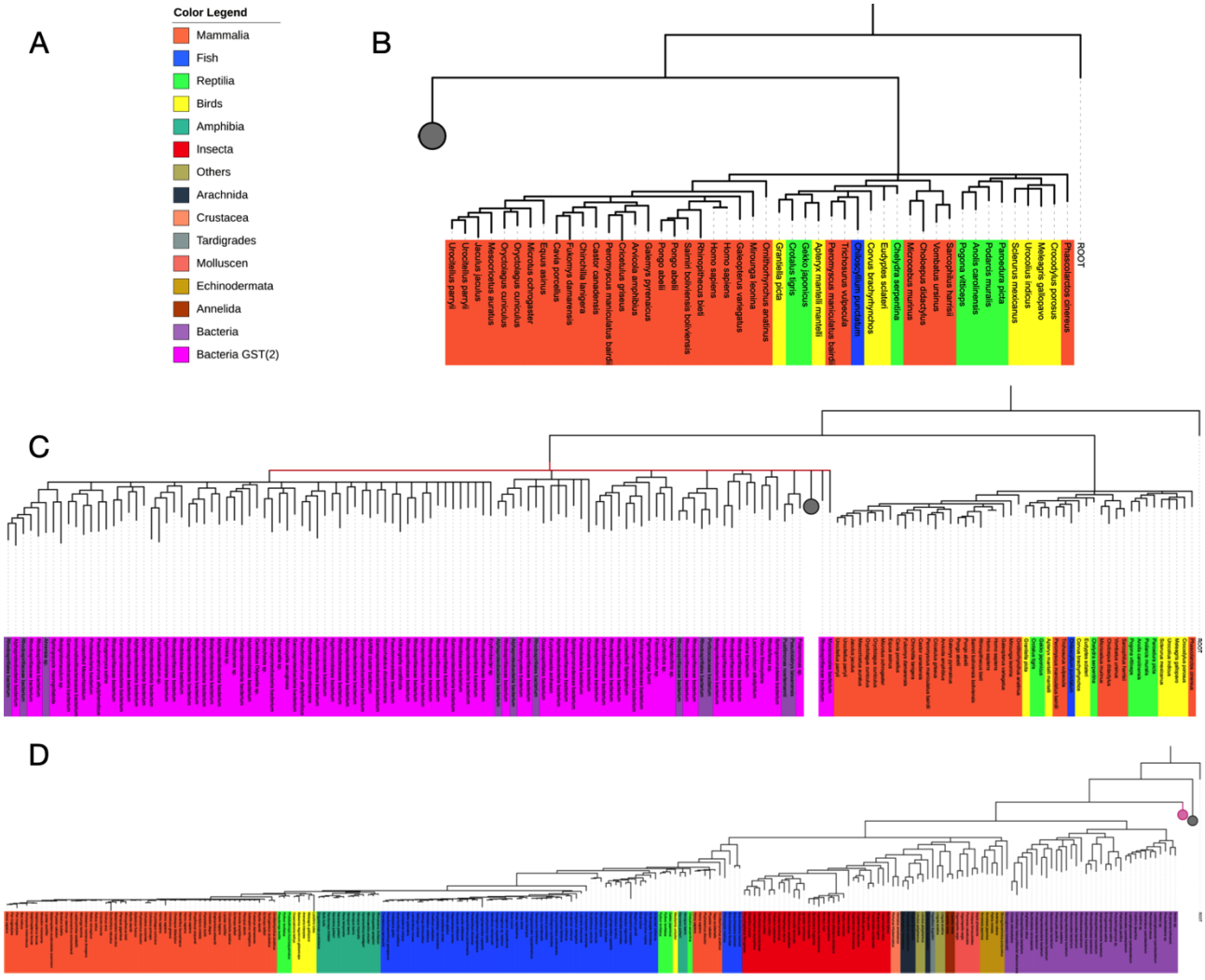


Figure 6. Phylogenetic analyses. *A. Colour scheme for phylogenetic trees. Note that not all organisms seen in the colour legend are present in every picture. This is because the trees are very large and to increase visibility parts of the tree will be collapsed (denoted by a grey circle), to hide underlying branches. B. The first branch of the phylogenetic tree. The tree was rooted at the root sequence. On the right side are eukaryotic GST sequences (mammals, fish, birds, reptiles). The branch includes 48 sequences. The rest of the 398 sequences that are represented are at hierarchically lower levels of the phylogenetic tree. The corresponding branches were collapsed (grey circle). C. In contrast to panel B, the left branch is partially extended. The right-sided branch shows eukaryotic GST sequences, while the left side shows a total of 110 bacterial GST sequences. The remaining sequences are at hierarchically lower levels of the phylogenetic tree. The corresponding branches were collapsed (grey circle). The red line indicates that the time could not be resolved at this level. D. The last hierarchical level of the phylogenetic tree. The tree was rooted at the root sequence. Purple sequences on the right side correspond to bacterial GST sequences. Sequence for eukaryotic GDAP1 and GDAP1L1 are in the left side branches. The grey circle denotes the collapsed branch for the eukaryotic GSTs (see panel B) and the magenta circle denotes the second group of bacterial GST sequences (see panel C).*

Sequence entropy in GDAP1

The entropy analysis was carried out on both the whole GST family, only the GDAP1/GDAP1L1 sequences, or only on non-GDAP1 sequences. We will first look at the results from the GDAP1 subset of sequences (**Fig. 7A**). A total of 44 residues were found with an entropy score ≤ 0.10 and a percentage of non-gap characters present in the alignment > 70 (**Table 6**); the entropy is mapped onto the GDAP1 structure in **Fig. 7B**. The data highlight several interesting residues, which are known to be sites of CMT mutations, relevant for folding and structure, or with possible functions in ligand binding. These aspects are discussed more below.

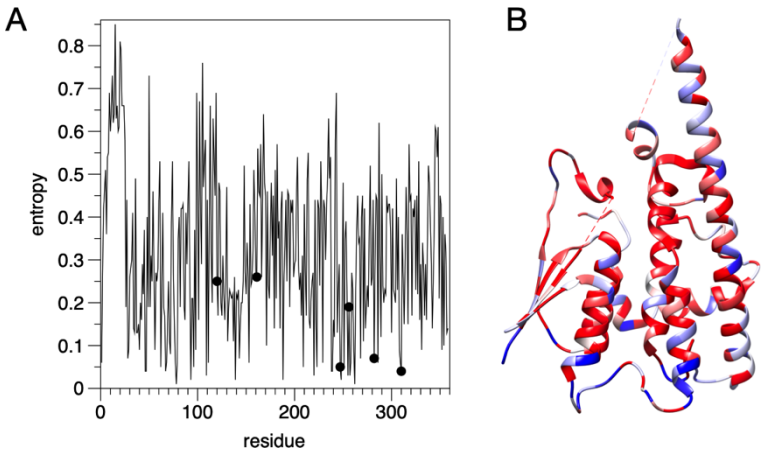


Figure 7. Sequence entropy analysis. *A. Entropy plot for GDAP1. The positions of the mutations studied here are shown with dots. B. Mapping of entropy onto the GDAP1 crystal structure monomer. Blue indicates high entropy and red low.*

Table 6. Entropy analysis of the GDAP1 subfamily. *Shown are the positions (human GDAP1 reference sequence numbering) with the lowest entropy (S < 0.1).*

residue	entropy	frac	Amino acid	residue	entropy	frac	Amino acid	residue	entropy	frac	Amino acid
28	0.07	0.78	L	98	0.09	0.94	E	248	0.02	0.97	D
29	0.09	0.78	Y	109	0.03	0.87	L	251	0.05	0.97	L
40	0.09	0.83	V	111	0.06	0.91	P	255	0.03	0.97	L
46	0.04	0.83	E	139	0.02	0.91	G	257	0.03	0.97	R
47	0.04	0.83	K	143	0.07	0.91	H	258	0.06	0.97	L
56	0.09	0.83	V	153	0.06	0.91	P	261	0.09	0.97	L
62	0.05	0.84	E	171	0.1	0.94	L	262	0.01	0.96	G
67	0.04	0.84	W	188	0.02	0.94	K	268	0.07	0.96	W
68	0.06	0.84	F	209	0.05	0.94	L	282	0.07	0.96	R
72	0.08	0.93	N	220	0.03	0.94	E	286	0.06	0.96	R
77	0.06	0.93	V	223	0.05	0.94	L	308	0.09	0.84	A
78	0.01	0.93	P	229	0.06	0.82	E	309	0.06	0.82	F
79	0.06	0.93	V	238	0.04	0.97	W	310	0.04	0.82	R
93	0.02	0.93	I	239	0.04	0.97	L	331	0.09	0.89	G

96	0.05	0.94	Y	247	0.05	0.97	A				

508

509 One intuitive approach to search for functionally important sites is to focus on residues involved in
510 ligand binding, *i.e.* residues that are located in the active centre. However, the question of whether
511 GDAP1 is an active GST, or even lacks glutathione binding completely, has not been clearly resolved.
512 Hexadecanedioic acid was identified as possible ligand and subsequently crystallized in complex
513 with GDAP1 [11]. Gln235, Trp238, Arg282, Arg286, and Lys287 contact the ligand; Trp238, Arg282
514 and Arg286 are highly conserved (**Table 6**). This suggests that they could be important for ligand
515 binding and that mutations in these positions could affect GDAP1 function; indeed, Trp238 and
516 Arg282 are CMT mutation sites. These 3 residues cluster on the side of GDAP1, which must face the
517 membrane surface. One option is that the bound fatty acid, in fact, mimics the membrane surface,
518 and that these residues are directly involved in membrane surface binding at the outer surface of
519 the MOM.

520 **Table 7** shows the results of conservation entropy based on the combined data set (GDAP1,
521 GDAP1L1, and GST). Since the GST superfamily is diverse by nature, with a conserved fold but low
522 sequence conservation, the limits were relaxed. All residues with an entropy ≤ 0.20 and a
523 percentage of non-gap character > 70 are listed in the table. Overall, 14 residues were derived by
524 entropy calculations. Note that these include Leu239 and Gly241 in the α6-α7 loop and Ala247 and
525 Asp 248 on helix α7. The effect of the A247V mutation on protein stability may therefore be linked
526 to its high conservation in the GST family.

527 ***Table 7. Entropy analysis for the large data set including the GST dataset.*** *Shown are residue*
528 *positions with the lowest entropy (S < 0.2).*

residue	entropy	frac	Amino acid
28	0.12	0.92	L
72	0.15	0.98	N
77	0.17	0.98	V
78	0.01	0.98	P
93	0.05	0.98	I
96	0.1	0.98	Y
109	0.17	0.9	L
111	0.14	0.81	P
223	0.1	0.89	L
239	0.2	0.9	L
241	0.14	0.9	G
247	0.12	0.9	A

248	0.02	0.9	D
286	0.17	0.86	R

529

530 **Kullback-Leibler divergence**

531 The results of the KL divergence are shown in **Table 8**. Many listed residues are insertions because
532 they correspond to a gap symbol in the alignment. Of the GDAP1 residues Cys51, Gly83, Tyr124, and
533 Glu228 that are not assigned to a gap symbol, Cys51 and Gly83 are located close to the dimerization
534 site of GDAP1. Gly83 is a target of a non-pathological polymorphism, G83A [65]. The β sheet is the
535 dimerization site of GDAP1, and the larger amino acids Phe and Asp may lead to steric hindrance
536 during dimer formation. On the other hand, Phe and Asp are predestined for stacking and polar
537 interactions. Since the binding site for glutathione in GST takes place near this segment, these
538 residues could be important for ligand binding in the GSTs, which may not to be relevant for GDAP1
539 due to the apparent lack of glutathione binding. The observations support the assumption that the
540 binding site for a small-molecule ligand, if any, in GDAP1 may be located elsewhere.

541 On the other hand, Tyr124 and Glu228 are close to each other in 3D space, being located on the
542 core helices α3 and α6, respectively. The conversion of tyrosine to tryptophan (Tyr124 → Trp) and
543 glutamic acid to glutamine (Glu228 → Gln) between GDAP1 and canonical GSTs causes only small
544 changes in physicochemical properties. However, the central location of these residues in the
545 GDAP1 fold suggests this finding may reflect a structural or functional aspect of GDAP1.

546 Of additional interest are Lys193 and Asn166, which are both close to the GDAP1-specific insertion.
547 In contrast to the crystal structure, where Asn166 is not visible, the AlphaFold2 model predicts that
548 the helix α6 is bent, which brings Lys193 and Asn166 into close proximity. This could enable
549 intramolecular interactions or provide a specific site for protein-protein interactions. Overall, the
550 results indicate that certain subdomains in GDAP1 may have evolved to fulfill a different function
551 than found in canonical GSTs.

552 ***Table 8. KL divergence.***

res num	klP	klQ	Sp	Sq	GDAP1	GST
51	1.0292121	0.776726	0.1934639	0.46464223	C	F
83	0.711117	0.8625605	0.4245107	0.37385303	G	D
106	1.608235	1.8079219	0.47457752	0.26693898	T	-
124	1.1740065	1.2500156	0.19621734	0.38821948	Y	W
131	0.9714733	2.912868	0.11059844	0.16480081	L	-
132	0.94019866	1.8085836	0.19164778	0.39762115	P	-

149	0.7277371	2.1168306	0.27086747	0.26017815	D	-
150	0.703265	1.8003652	0.21712388	0.2821153	S	-
151	1.8435297	2.2269633	0.3448432	0.3719185	M	-
166	1.2843894	2.3159494	0.2552226	0.25735173	N	-
186	1.1256604	2.4295797	0.41708586	0.2931384	I	-
192	1.0507396	1.881601	0.24847703	0.4722129	L	-
193	0.86935514	2.3180199	0.4641341	0.4705446	K	-
228	1.1552744	1.9867854	0.40772098	0.3858474	E	Q
229	1.6037626	1.9517808	0.06304588	0.48370922	E	-
233	2.1693683	2.274921	0.42683157	0.312741	E	-
234	2.1514575	2.6264071	0.48738602	-0.0	G	-
236	2.176767	2.9161973	0.48846936	-0.0	Q	-
295	1.5544128	2.371196	0.45351568	0.49149024	H	-
300	2.4423957	2.5055954	0.12804487	0.4074487	L	-
319	1.0532062	2.999142	0.4438014	0.21979499	V	-
339	1.537214	1.3012601	0.48322365	0.352141	L	-
352	2.5170794	1.6161157	0.094565034	0.42042798	R	-
353	1.3202384	1.7585322	0.39914683	0.4678713	P	-
355	2.271911	1.898613	0.3553678	-0.0	P	-
356	2.2988195	1.9953156	0.32429653	0.41148844	N	-
358	2.5016084	2.5750391	0.1393856	-0.0	F	-

553

554

DISCUSSION

Mutations in dozens of genes are causative for various subtypes of CMT disease; for most of them, effects of missense mutations at the protein level remain poorly characterised. This is partly because the biological functions of the corresponding proteins are poorly understood at the molecular level. One of the characterised examples is myelin protein P2, which loses thermal stability upon all the known 6 CMT mutations in the protein, while the crystal structure remains nearly unaltered [56, 57]. Furthermore, the disordered tail of myelin protein P0 is a target for P0 mutations, and its membrane interactions may be compromised upon CMT [66]. Periaxin carries several CMT-linked truncation mutations that abolish protein-protein interactions [67, 68], most notably with β4 integrin [69]. While these proteins, highly enriched in myelinating Schwann cells, are involved in the classical Schwann cell phenotypes related to myelination, it is evident that compromised mitochondrial function is one underlying cause of especially axonal CMT subtypes, and mutations in GDAP1 are linked to mitochondrial dysfunction. Linked to this mechanism, recent data show that GDAP1 may be involved in interactions with the actin and tubulin cytoskeletons [9].

Mutations in the GST-like domain of GDAP1 have a broad pathological spectrum. The molecular basis is still unknown despite the cellular observations confirming the causality of impaired mitochondrial dynamics, and accurate structural information is required to support these findings. Below, we shall discuss some implications of our findings to understand GDAP1 function at the molecular level and the effects of missense mutations therein.

CMT mutations at the GDAP1 protein level

CMT-linked missense mutations are relatively common in the *GDAP1* gene compared to other CMT target genes, especially when the size of the protein is taken into consideration [23, 70]. With careful examination of structural models and using them as inputs for further bioinformatics analyses, the mutations are observed to cause subtle changes in residue interaction networks, in line with experimental data. A comprehensive understanding of the effects of single mutations requires local observation of the structure coupled to experimental data at the protein and cellular levels.

Of the residues highlighted by the entropy analysis either within the GDAP1 sequence set or the entire GST set, several are targets for missense mutations. Here, we shall briefly compare selected

584 CMT mutations in GDAP1. The immediate environment of each residue is considered to shed light
585 on local effects of each mutation.

586 Tyr29 is conserved at the symmetry axis of the GDAP1 dimer, forming a H bond between OH groups
587 of Tyr29 from the two protomers. Y29S [71] would both remove this polar interaction and make the
588 dimer interface much less hydrophobic.

589 Leu239 is at the tip of the α6-α7 loop, inserted into the structural core, and its mutation to Phe has
590 been reported in CMT [72]. The side chain is close to those of Cys240 and Ala247, and hence, a larger
591 residue at this position could similarly affect protein stability as A247V. Similar effects could be
592 foreseen for the C240Y mutation [73]. Pro111 is in the α2-α3 loop, close in 3D space to the N
593 terminus of helix α7 and the α6-α7 loop. Its mutation to His [74] would cause steric hindrance and
594 alter the conformation of the α2-α3 loop, possibly destabilising the fold.

595 Ala247, which was studied in the current work, lies centrally in the GDAP1 fold on helix α7. A247V
596 is linked to CMT, showing that no larger residue fits into this tight space. A247V caused
597 destabilisation of the GDAP1 fold. Ala247 is one of the most conserved residues in the GST
598 superfamily, indicating a role in the GST fold. Other residues conserved in the GDAP1 set include
599 Leu255 and Gly262, which lie in the middle and at the end of helix α7, respectively. L255F [75] would
600 be expected to cause similar steric hindrance as A247V, and G262E [71] will likely disturb the tight
601 turn right after helix α7 and cause steric clashes. Both Leu255 and Gly262 are highly conserved in
602 GDAP1 sequences.

603 Arg282, studied here in the form of the CMT mutation R282H, is one of the most conserved residues
604 in the GDAP1/GDAP1L1 subfamily. Its strong hydrogen bonding to the carbonyl groups in the α6-α7
605 loop indicates a central role in GDAP1 structure. In addition to R282H, the mutation R282C has been
606 reported [76], and we can expect it to have the same kind of effect as R282H, losing the side-chain
607 interactions of Arg282 to the backbone of the α6-α7 loop. Arg282 is stacked against the aromatic
608 side chain of Trp238, which is both as a conserved residue and a CMT mutation target [77]. This Arg-
609 Trp-(α6-α7) unit is likely important for stable folding of GDAP1.

610 Arg310 is not present in the constructs we used for crystallisation, and hence, we have no high-
611 resolution data on its conformation. However, the AF2 model of GDAP1 extends our crystal
612 structures and predicts that Arg310 points away from the membrane and forms two salt bridges to
613 acidic residues in the GDAP1 fold (**Fig. 8**). Interestingly, Arg310 resides in the segment originally

coined the hydrophobic domain; in our view, this fragment more likely corresponds to an amphipathic helix, which could bind onto a phospholipid membrane surface. The AlphaFold2 prediction supports this view. Arg310 is likely to link this membrane-bound helix to the folded core of GDAP1 and therefore directly affect the conformation of full-length, membrane-bound GDAP1. This is supported by the lack of a destabilising effect on recombinant GDAP1 structure in our experiments.

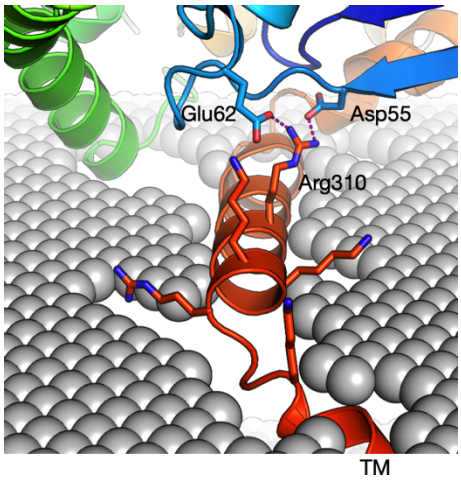


Figure 8. Location of Arg310 outside the GST-like core. *Arg310 within the amphipathic helix preceding the transmembrane domain is predicted to make salt bridges back towards the core domain.*

The clinical profiling of CMT patients and experiments in cell-based models have shown that residues located within or near the vicinity of the transmembrane helix are severe [78, 79]. These mutations likely affect the proper localization into the MOM, leading to impaired GDAP1 folding and function. So far, studies *in vitro* on the GDAP1 protein have been done only on soluble constructs, lacking the transmembrane domain. The stability analysis of the mutant form of the full-length GDAP1 could have a different outcome due to the lipid bilayer component, and this is a future research direction, given that full-length recombinant GDAP1 can be expressed and purified.

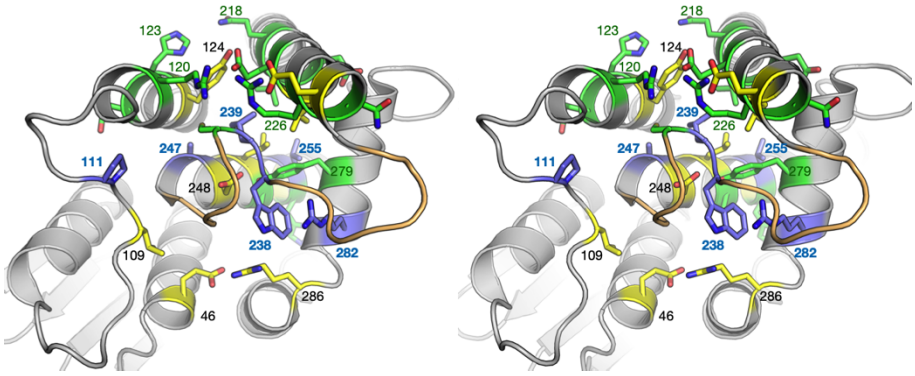
The central role of the α6-α7 loop

Despite the broad spectrum of disease mutations, certain general conclusions hold in terms of the mutations. The interaction network described here shows that the critical helices in the C-terminal

637 GST-like core are linked by residues that correspond to CMT target locations and/or are conserved.
638 Therefore, based on evolutionary reasoning, these residues are likely essential for the integrity of
639 GDAP1 and its function.

640 The α6-α7 loop is a central feature in the structure of GDAP1, but also in canonical GSTs. This loop
641 inserts itself back into the protein, being a central interaction hub between helices α3, α6, α7, and
642 α8 (**Fig. 9**). In GDAP1, the 3 residues at the tip of the loop (238-240) are targets for CMT disease
643 mutations, showing the crucial importance of this segment. Our entropy analysis confirms the
644 strong conservation of this segment. We can then take a broader scope and further look at residues
645 interacting with the α6-α7 loop in GDAP1. Of note, several of the residues in direct contact with the
646 α6-α7 loop are known CMT mutation target sites and/or highly conserved (**Table 6**). Hence, the
647 inward-bending α6-α7 loop appears to be a central structural feature in GDAP1, and its alteration
648 either directly or *via* intramolecular interactions may be a common mechanism for CMT.

649



650

651 ***Figure 9. An overview of the α6-α7 loop and its surroundings in light of CMT mutations and***
652 ***conserved positions.*** *In this stereo view, the α6-α7 loop is coloured orange, green shows positions*
653 *for CMT mutations, yellow positions that are highly conserved, and blue the positions that are both*
654 *targeted by CMT mutations and highly conserved. Selected residues are labelled for clarity.*

655

656 **GDAP1 as a member of the GST superfamily**

657 The predictions that GDAP1 would be structurally related to GSTs have been confirmed by recent
658 structural studies by us and others [10-12]. Important differences to canonical GSTs exist, however.
659 The sequence identity to any enzyme with known GST activity is very low, and contradictory results
660 have been obtained as to the GST activity of GDAP1. In our hands, GDAP1 does not bind glutathione
661 or act as a GST [11]. The latter is logical, since the GDAP1 dimers are formed completely differently

662 from canonical GST, in which the active site in fact lies at the dimer interface [80]. We hypothesise
663 that the residues with low entropy and high conservation in the whole GST set are crucial for correct
664 folding. On the other hand, residues additionally highlighted in the GDAP1/GDAP1L1 dataset may
665 relate to GDAP1-specific functions and/or unique structural aspects of this subfamily.

666 Our conservation analyses highlight the low conservation of GDAP1 (and GDAP1L1) in the GST
667 superfamily. While the entire evolutionary pathway cannot be tracked based on the analyses, it is
668 intriguing that GDAP1 is closer to bacterial than eukaryotic GSTs. It is possible that during evolution,
669 GDAP1 has lost the characteristic GST activity, while becoming an integral membrane protein of the
670 MOM. Its functions could, therefore, be mediated through protein-protein interactions instead of
671 enzymatic activity. The functions can be redox-regulated, which could explain the observation that
672 the human GDAP1 dimer is mediated by a disulphide bridge *via* Cys88 [11].

673 To complement the above analyses, we superposed crystal structures of human GDAP1 and a
674 canonical GST, that from *S. japonicum* [80], and analysed the current crystal structures with respect
675 to GST. This is the GST widely used in molecular biology applications as a fusion tag for affinity
676 purification. We were interested in the residues affected by CMT mutations, especially those
677 crystallised here. Hence, of specific interest were the apparent non-conservation of Arg120 and the
678 conservation of Ala247 and Arg282 (**Fig. 10A**).

679

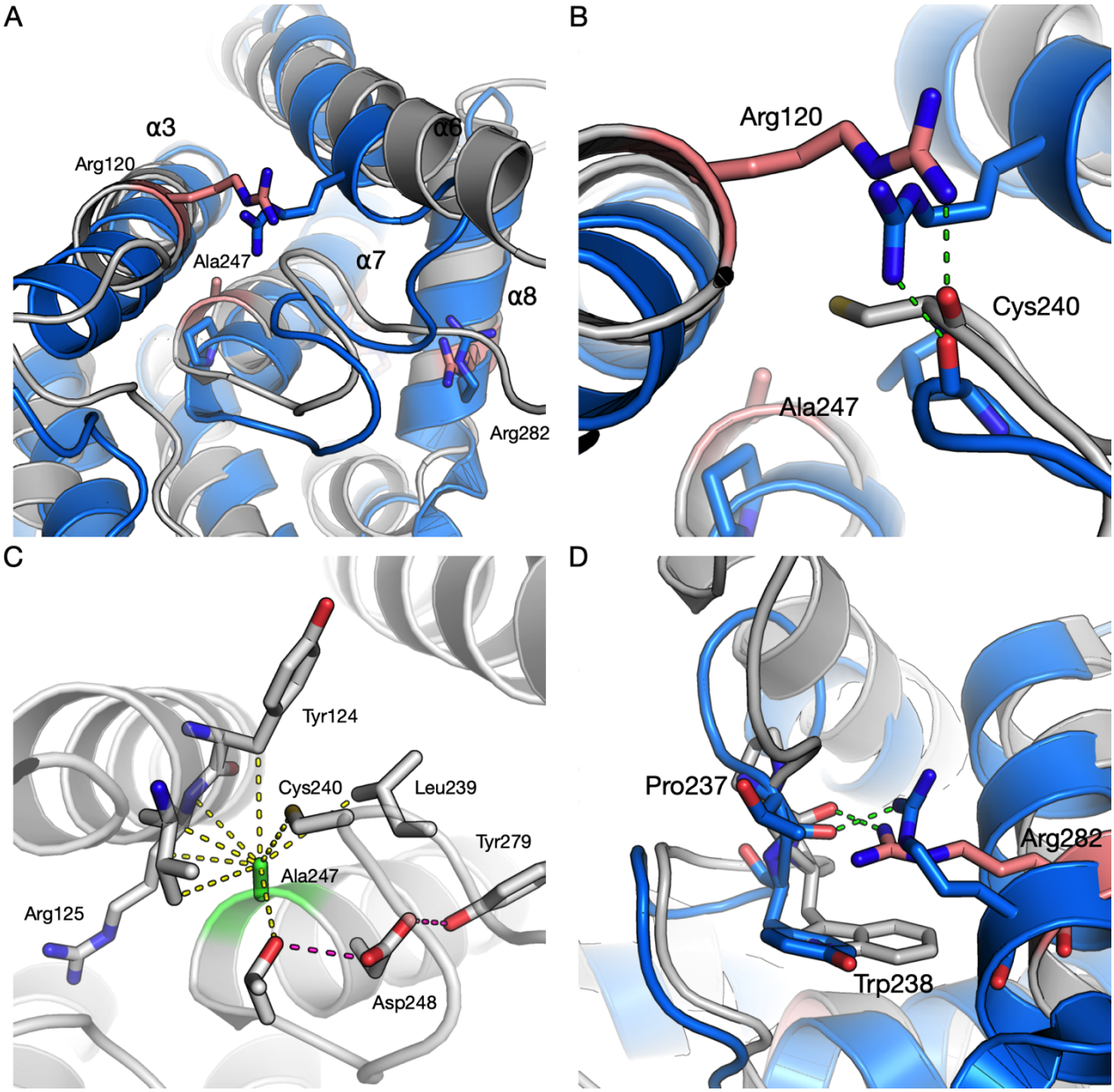


Figure 10. Comparisons of the studied mutations to a canonical GST from S. japonicum. *GDAP1 is in gray and GST in blue. A. Overall view of the GDAP1-GST superposition, with the CMT mutation sites crystallised here highlighted. B. Arg120 (pink) in GDAP1 and the corresponding interaction in GST. C. Tight packing of Ala247 (green) in the GDAP1 structure; Ala at this position is highly conserved across the whole GST family. Note Tyr124, which was highlighted in the KL divergence analysis, making direct contact with Ala247. D. The interactions of Arg282 (pink) towards the α6-α7 loop are conserved in GST.*

Arg120 in GDAP1 is a target of several CMT mutations [81-83], and its interactions with the backbone carbonyl of Cys240 in the α6-α7 loop appear central. Structures of the mutants R120Q and R120W indicate loosening of the structure locally, as well as loss of the hydrogen bond, which is accompanied by a decrease in observed heat stability in R120Q and R120W [12]. Arg120 is not

conserved in GSTs at the sequence level; however, an Arg residue from the neighbouring helix in GST reaches to the same position and makes similar interactions to the same carbonyl group in the α6-α7 loop (**Fig. 10B**). These observations highlight the importance of the central α6-α7 loop in intramolecular interaction networks and fold stability of GDAP1.

Ala247 of GDAP1, although perhaps a mundane residue *per se*, appears surprisingly conserved in the GST superfamily in our dataset (**Table 6**). In the superposed individual GST structure (**Fig. 10A**), this residue is Pro, which fits well into the structure due to slightly different conformations of the main helices in GST. In GDAP1, Ala247 is so snugly packed (**Fig. 10C**) that even the addition of two methyl groups in the A247V variant causes protein instability and disease, despite minor effects in the crystal structure. A247V is an example of a mutation introduced into the hydrophobic core that may seem relatively harmless, but in fact has long-range effects on the entire protein structure stability. One of the residues in GDAP1 highlighted by the KL divergence analysis, Tyr124, closely interacts with Ala247, indicating a GDAP1-specific arrangement at this site. Additionally, one of the most conserved GDAP1 residues, Asp248, is central in a hydrogen bonding network linking the conserved CMT target Tyr279 into the picture (**Fig. 10C**).

GST has an Arg corresponding to Arg282, and this residue is highly conserved in the GST family, as shown by our bioinformatics analyses (**Table 7**). Arg282 in GDAP1 interacts directly with the backbone of the α6-α7 loop, and a similar interaction is observed in GST (**Fig. 10D**). This is another example of an Arg – loop backbone interaction conserved and relevant for human disease mutations.

Biological implications

The function of GDAP1 at the molecular level remains enigmatic to date. While indications exist from functional studies [2, 3, 79], evidence is still incomplete for any enzymatic activity as well as direct protein-protein interactions. Redox regulation seems to play a role [9, 73, 84], but is this related to an enzymatic GST-like activity, or regulation of oligomeric state and/or protein-protein interactions?

At the molecular level, we believe to have identified important residue interaction networks between the core helices in the GST-like domain of GDAP1, strongly interacting with the inward α6-α7 loop. These networks could be important for both GDAP1 stability and its interactions with other

molecules, such as the membrane or the cytoskeleton. The correct conformation of GDAP1 on the MOM, as well as its interactions with other proteins, will then directly or indirectly affect mitochondrial dynamics to promote correct development and function of the nervous system. The disease mutations may – due to their involvement in the same intramolecular networks – cause similar overall effects on GDAP1 stability and properties, which then leads to the CMT disease phenotype in patients. In line with this, all the missense mutations we have studied at the protein level allow for GDAP1 folding, while at the same time decreasing the stability of the fold.

It should be noted that due to the large pool of GDAP1 mutants causing CMT, our experimental sample size is still relatively small, and the hypothesis may not be correct in all cases. However, we have picked mutations from different core secondary structure elements for experimental analyses to account for an incomplete dataset, and predictions have been done for all mutations [12], indicating a general trend of structural destabilisation upon CMT missense mutations in GDAP1.

Concluding remarks

Considering mitochondrial dynamics and interactions with other organelles, the homology to GSTs brings attractive prospects for GDAP1 function. GDAP1 takes part not only in mitochondrial fission and fusion, but in interactions with the endoplasmic reticulum, peroxisomes, Golgi, and the cytoskeleton. These findings coupled with structural and biophysical data shall aid the understanding of the pathophysiological mechanism of GDAP1-linked CMT and may affect future treatment approaches. Future studies are needed to identify proteins and small molecules directly interacting with full-length GDAP1 in the physiological setting, allowing further structural investigations on the related molecular processes in nervous system function and disease. In a wider setting, we hypothesise that decrease in overall protein stability upon missense mutations is one common mechanism for CMT at the molecular level, and many of the related interaction networks are linked to the core helices and the α6-α7 loop of GDAP1.

751 **Acknowledgements**

752 This work was funded by the Academy of Finland, project number 24302881. We wish to
753 acknowledge the availability of synchrotron beamtime and excellent beamline support at both
754 SOLEIL, ISA, and DESY. The SAXS and SRCD experiments were supported by CalipsoPlus, which is
755 funding from the European Union Horizon 2020 research and innovation programme under grant
756 agreement No 730872. SRCD experiments further received funding from the European Union
757 Horizon 2020 research and innovation programme under grant agreement No 101004806 (MOSBRI-
758 2021-24).

759

760

761

References

1. Watts ME, Pocock R, Claudianos C. Brain Energy and Oxygen Metabolism: Emerging Role in Normal Function and Disease. Front Mol Neurosci. 2018;11:216. Epub 20180622. doi: 10.3389/fnmol.2018.00216. PubMed PMID: 29988368; PubMed Central PMCID: PMCPMC6023993.
2. Barneo-Munoz M, Juarez P, Civera-Tregon A, Yndriago L, Pla-Martin D, Zenker J, et al. Lack of GDAP1 induces neuronal calcium and mitochondrial defects in a knockout mouse model of charcot-marie-tooth neuropathy. PLoS Genet. 2015;11(4):e1005115. Epub 20150410. doi: 10.1371/journal.pgen.1005115. PubMed PMID: 25860513; PubMed Central PMCID: PMCPMC4393229.
3. Miressi F, Benslimane N, Favreau F, Rassat M, Richard L, Bourthoumieu S, et al. GDAP1 Involvement in Mitochondrial Function and Oxidative Stress, Investigated in a Charcot-Marie-Tooth Model of hiPSCs-Derived Motor Neurons. Biomedicines. 2021;9(8). Epub 20210802. doi: 10.3390/biomedicines9080945. PubMed PMID: 34440148; PubMed Central PMCID: PMCPMC8393985.
4. Pla-Martin D, Rueda CB, Estela A, Sanchez-Piris M, Gonzalez-Sanchez P, Traba J, et al. Silencing of the Charcot-Marie-Tooth disease-associated gene GDAP1 induces abnormal mitochondrial distribution and affects Ca2+ homeostasis by reducing store-operated Ca2+ entry. Neurobiol Dis. 2013;55:140-51. Epub 20130328. doi: 10.1016/j.nbd.2013.03.010. PubMed PMID: 23542510.
5. Rossi A, Pizzo P, Filadi R. Calcium, mitochondria and cell metabolism: A functional triangle in bioenergetics. Biochim Biophys Acta Mol Cell Res. 2019;1866(7):1068-78. Epub 20181026. doi: 10.1016/j.bbamcr.2018.10.016. PubMed PMID: 30982525.
6. Niemann A, Ruegg M, La Padula V, Schenone A, Suter U. Ganglioside-induced differentiation associated protein 1 is a regulator of the mitochondrial network: new implications for Charcot-Marie-Tooth disease. J Cell Biol. 2005;170(7):1067-78. Epub 20050919. doi: 10.1083/jcb.200507087. PubMed PMID: 16172208; PubMed Central PMCID: PMCPMC2171517.
7. Niemann A, Wagner KM, Ruegg M, Suter U. GDAP1 mutations differ in their effects on mitochondrial dynamics and apoptosis depending on the mode of inheritance. Neurobiol Dis. 2009;36(3):509-20. Epub 20090925. doi: 10.1016/j.nbd.2009.09.011. PubMed PMID: 19782751.
8. Cuesta A, Pedrola L, Sevilla T, Garcia-Planells J, Chumillas MJ, Mayordomo F, et al. The gene encoding ganglioside-induced differentiation-associated protein 1 is mutated in axonal Charcot-Marie-Tooth type 4A disease. Nat Genet. 2002;30(1):22-5. Epub 20011217. doi: 10.1038/ng798. PubMed PMID: 11743580.
9. Wolf C, Pouya A, Bitar S, Pfeiffer A, Bueno D, Rojas-Charry L, et al. GDAP1 loss of function inhibits the mitochondrial pyruvate dehydrogenase complex by altering the actin cytoskeleton. Commun Biol. 2022;5(1):541. Epub 20220603. doi: 10.1038/s42003-022-03487-6. PubMed PMID: 35662277.
10. Googins MR, Woghiren-Afegbua AO, Calderon M, St. Croix CM, Kiselyov KI, VanDemark AP. Structural and functional divergence of GDAP1 from the glutathione S-transferase superfamily. The FASEB Journal. 2020;34:7192-207. doi: 10.1096/fj.202000110R.
11. Nguyen GTT, Sutinen A, Raasakka A, Muruganandam G, Loris R, Kursula P. Structure of the Complete Dimeric Human GDAP1 Core Domain Provides Insights into Ligand Binding and Clustering of Disease Mutations. Frontiers in Molecular Biosciences. 2020;7:631232. doi: 10.3389/fmolb.2020.631232.
12. Sutinen A, Nguyen GTT, Raasakka A, Muruganandam G, Loris R, Ylikallio E, et al. Structural insights into Charcot-Marie-Tooth disease-linked mutations in human GDAP1. FEBS Open Bio. 2022. Epub 20220504. doi: 10.1002/2211-5463.13422. PubMed PMID: 35509130.
13. Mannervik B, Alin P, Guthenberg C, Jensson H, Tahir MK, Warholm M, et al. Identification of three classes of cytosolic glutathione transferase common to several mammalian species: correlation between structural data and enzymatic properties. Proceedings of the National Academy of Sciences of the United States of America. 1985;82(21):7202-6. doi: 10.1073/pnas.82.21.7202.
14. Fontés M. Charcot Marie Tooth Disease. A Single Disorder? International Journal of Molecular Sciences. 2018;19(12). doi: 10.3390/ijms19123807.
15. Rossor AM, Tomaselli PJ, Reilly MM. Recent advances in the genetic neuropathies. Current Opinion in Neurology. 2016;29(5):537-48. doi: 10.1097/WCO.0000000000000373.

16. DiVincenzo C, Elzinga CD, Medeiros AC, Karbassi I, Jones JR, Evans MC, et al. The allelic spectrum of Charcot-Marie-Tooth disease in over 17,000 individuals with neuropathy. Mol Genet Genomic Med. 2014;2(6):522-9. Epub 20140821. doi: 10.1002/mgg3.106. PubMed PMID: 25614874; PubMed Central PMCID: PMCPMC4303222.
17. Szigeti K, Lupski JR. Charcot–Marie–Tooth disease. European Journal of Human Genetics. 2009;17(6):703-10. doi: 10.1038/ejhg.2009.31.
18. Niemann A, Ruegg M, La Padula V, Schenone A, Suter U. Ganglioside-induced differentiation associated protein 1 is a regulator of the mitochondrial network: new implications for Charcot-Marie-Tooth disease. The Journal of cell biology. 2005;170(7):1067-78. doi: 10.1083/jcb.200507087.
19. Cassereau J, Chevrollier A, Gueguen N, Desquiret V, Verny C, Nicolas G, et al. Mitochondrial dysfunction and pathophysiology of Charcot-Marie-Tooth disease involving GDAP1 mutations. Experimental neurology. 2011;227(1):31-41. doi: 10.1016/j.expneurol.2010.09.006.
20. Mai P-T, Le D-T, Nguyen T-T, Le Gia H-L, Nguyen Le T-H, Le M, et al. Novel GDAP1 Mutation in a Vietnamese Family with Charcot-Marie-Tooth Disease. BioMed research international. 2019;2019:7132494. doi: 10.1155/2019/7132494.
21. Zimon M, Baets J, Fabrizi GM, Jaakkola E, Kabzinska D, Pilch J, et al. Dominant GDAP1 mutations cause predominantly mild CMT phenotypes. Neurology. 2011;77(6):540-8. doi: 10.1212/WNL.0b013e318228fc70.
22. Baxter RV, Ben Othmane K, Rochelle JM, Stajich JE, Hulette C, Dew-Knight S, et al. Ganglioside-induced differentiation-associated protein-1 is mutant in Charcot-Marie-Tooth disease type 4A/8q21. Nature genetics. 2002;30(1):21-2. doi: 10.1038/ng796.
23. Rzepnikowska W, Kochanski A. A role for the GDAP1 gene in the molecular pathogenesis of CharcotMarieTooth disease. Acta neurobiologiae experimentalis. 2018;78(1):1-13.
24. Fagerberg L, Hallstrom BM, Oksvold P, Kampf C, Djureinovic D, Odeberg J, et al. Analysis of the human tissue-specific expression by genome-wide integration of transcriptomics and antibody-based proteomics. Molecular & cellular proteomics : MCP. 2014;13(2):397-406. doi: 10.1074/mcp.M113.035600.
25. Dasari S, Gonuguntla S, Ganjayi MS, Bukke S, Sreenivasulu B, Meriga B. Genetic polymorphism of glutathione S-transferases: Relevance to neurological disorders. Pathophysiology. 2018;25(4):285-92. Epub 20180611. doi: 10.1016/j.pathophys.2018.06.001. PubMed PMID: 29908890.
26. Kumar A, Dhull DK, Gupta V, Channana P, Singh A, Bhardwaj M, et al. Role of Glutathione-S-transferases in neurological problems. Expert Opin Ther Pat. 2017;27(3):299-309. Epub 20161110. doi: 10.1080/13543776.2017.1254192. PubMed PMID: 27785931.
27. Pandey T, Chhetri G, Chinta R, Kumar B, Singh DB, Tripathi T, et al. Functional classification and biochemical characterization of a novel rho class glutathione S-transferase in Synechocystis PCC 6803. FEBS Open Bio. 2015;5:1-7. doi: 10.1016/j.fob.2014.11.006.
28. Pflugmacher S, Wiegand C, Werner S, Schröder H, Kankaanpää H. Activity and substrate specificity of cytosolic and microsomal glutathione S-transferase in Australian black tiger prawns (Penaeus monodon) after exposure to cyanobacterial toxins. Environmental Toxicology. 2005;20(3):301-7. doi: https://doi.org/10.1002/tox.20121.
29. Studier FW. Protein production by auto-induction in high density shaking cultures. Protein expression and purification. 2005;41(1):207-34. doi: 10.1016/j.pep.2005.01.016.
30. Burkhardt A, Pakendorf T, Reime B, Meyer J, Fischer P, Stübe N, et al. Status of the crystallography beamlines at PETRA III. The European Physical Journal Plus. 2016;131(3):56. doi: 10.1140/epjp/i2016-16056-0.
31. Meents A, Reime B, Stuebe N, Fischer P, Warmer M, Goeries D, et al., editors. Development of an in-vacuum x-ray microscope with cryogenic sample cooling for beamline P11 at PETRA III. SPIE Optical Engineering + Applications; 2013 2013-9-26. San Diego, California, United States.
32. Cianci M, Bourenkov G, Pompidor G, Karpics I, Kallio J, Bento I, et al. P13, the EMBL macromolecular crystallography beamline at the low-emittance PETRA III ring for high- and low-energy phasing with variable beam focusing. J Synchrotron Radiat. 2017;24(Pt 1):323-32. Epub 20170101. doi: 10.1107/S1600577516016465. PubMed PMID: 28009574; PubMed Central PMCID: PMCPMC5182027.

862 33. Kabsch W. XDS. Acta crystallographica Section D, Biological crystallography. 2010;66(Pt 2):125-
863 32. doi: 10.1107/S0907444909047337.
864 34. McCoy AJ, Grosse-Kunstleve RW, Adams PD, Winn MD, Storoni LC, Read RJ. ıt Phaser
865 crystallographic software. Journal of Applied Crystallography. 2007;40(4):658–74. doi:
866 10.1107/S0021889807021206.
867 35. Afonine PV, Grosse-Kunstleve RW, Echols N, Headd JJ, Moriarty NW, Mustyakimov M, et al.
868 Towards automated crystallographic structure refinement with ıt phenix.refine. Acta Crystallographica
869 Section D. 2012;68(4):352–67. doi: 10.1107/S0907444912001308.
870 36. Emsley P, Lohkamp B, Scott WG, Cowtan K. Features and development of Coot. Acta
871 crystallographica Section D, Biological crystallography. 2010;66(Pt 4):486-501. doi:
872 10.1107/S0907444910007493.
873 37. Chen VB, Arendall WBr, Headd JJ, Keedy DA, Immormino RM, Kapral GJ, et al. MolProbity: all-
874 atom structure validation for macromolecular crystallography. Acta crystallographica Section D, Biological
875 crystallography. 2010;66(Pt 1):12-21. doi: 10.1107/S0907444909042073.
876 38. Jumper J, Evans R, Pritzel A, Green T, Figurnov M, Ronneberger O, et al. Highly accurate protein
877 structure prediction with AlphaFold. Nature. 2021;596(7873):583-9. Epub 20210715. doi: 10.1038/s41586-
878 021-03819-2. PubMed PMID: 34265844; PubMed Central PMCID: PMCPMC8371605.
879 39. Huang J, Rauscher S, Nawrocki G, Ran T, Feig M, de Groot BL, et al. CHARMM36m: an improved
880 force field for folded and intrinsically disordered proteins. Nat Methods. 2017;14(1):71-3. Epub 20161107.
881 doi: 10.1038/nmeth.4067. PubMed PMID: 27819658; PubMed Central PMCID: PMCPMC5199616.
882 40. Jo S, Cheng X, Lee J, Kim S, Park SJ, Patel DS, et al. CHARMM-GUI 10 years for biomolecular
883 modeling and simulation. J Comput Chem. 2017;38(15):1114-24. Epub 20161114. doi: 10.1002/jcc.24660.
884 PubMed PMID: 27862047; PubMed Central PMCID: PMCPMC5403596.
885 41. Thureau A, Roblin P, Pérez J. BioSAXS on the SWING beamline at Synchrotron SOLEIL. J Appl
886 Crystallogr. 2021;54:1698-710.
887 42. Manalastas-Cantos K, Konarev PV, Hajizadeh NR, Kikhney AG, Petoukhov MV, Molodenskiy DS,
888 et al. ATSAS 3.0: expanded functionality and new tools for small-angle scattering data analysis. J Appl
889 Crystallogr. 2021;54(Pt 1):343-55. Epub 20210201. doi: 10.1107/S1600576720013412. PubMed PMID:
890 33833657; PubMed Central PMCID: PMCPMC7941305.
891 43. Konarev PV, Volkov VV, Sokolova AV, Koch MHJ, Svergun DI. PRIMUS: a Windows PC-based
892 system for small-angle scattering data analysis. Journal of Applied Crystallography. 2003;36(5):1277–82. doi:
893 10.1107/S0021889803012779.
894 44. Svergun DI. Determination of the regularization parameter in indirect-transform methods
895 using perceptual criteria. Journal of Applied Crystallography. 1992;25(4):495-503. doi:
896 10.1107/S0021889892001663.
897 45. Svergun DI, Petoukhov MV, Koch MHJ. Determination of domain structure of proteins from x-
898 ray solution scattering. Biophys J. 2001;80:2946-53. doi: 10.1016/S0006-3495(01)76260-1. PubMed PMID:
899 11371467.
900 46. Svergun DI. Restoring low resolution structure of biological macromolecules from solution
901 scattering using simulated annealing. Biophysical journal. 1999;76(6):2879-86. doi: 10.1016/S0006-
902 3495(99)77443-6.
903 47. Svergun DI, Barberato C, Koch MH. CRYSOL - A program to evaluate X-ray solution scattering
904 of biological macromolecules from atomic coordinates. J Appl Crystallogr. 1995;28:768-73. doi:
905 10.1107/S0021889895007047.
906 48. Miles AJ, Wallace BA. CDtoolX, a downloadable software package for processing and analyses
907 of circular dichroism spectroscopic data. Protein science : a publication of the Protein Society.
908 2018;27(9):1717-22. doi: 10.1002/pro.3474.
909 49. Altschul SF, Madden TL, Schaffer AA, Zhang J, Zhang Z, Miller W, et al. Gapped BLAST and PSI-
910 BLAST: a new generation of protein database search programs. Nucleic Acids Res. 1997;25(17):3389-402. doi:
911 10.1093/nar/25.17.3389. PubMed PMID: 9254694; PubMed Central PMCID: PMCPMC146917.

912 50. Katoh K, Misawa K, Kuma K, Miyata T. MAFFT: a novel method for rapid multiple sequence
913 alignment based on fast Fourier transform. Nucleic Acids Res. 2002;30(14):3059-66. doi: 10.1093/nar/gkf436.
914 PubMed PMID: 12136088; PubMed Central PMCID: PMCPMC135756.
915 51. Kullback S, Leibler RA. On information and sufficiency. Ann Math Stat. 1951;22:79-86.
916 52. Mackay DJC. Information theory, inference, and learning algorithms: Cambridge University
917 Press; 2003.
918 53. Huelsenbeck JP, Ronquist F. MRBAYES: Bayesian inference of phylogenetic trees.
919 Bioinformatics. 2001;17:754-5
920 .
921 54. Letunic I, Bork P. Interactive Tree Of Life (iTOL) v5: an online tool for phylogenetic tree display
922 and annotation. Nucleic Acids Res. 2021;49(W1):W293-W6. doi: 10.1093/nar/gkab301. PubMed PMID:
923 33885785; PubMed Central PMCID: PMCPMC8265157.
924 55. Ronquist F, Teslenko M, van der Mark P, Ayres DL, Darling A, Hohna S, et al. MrBayes 3.2:
925 efficient Bayesian phylogenetic inference and model choice across a large model space. Syst Biol.
926 2012;61(3):539-42. Epub 20120222. doi: 10.1093/sysbio/sys029. PubMed PMID: 22357727; PubMed Central
927 PMCID: PMCPMC3329765.
928 56. Ruskamo S, Nieminen T, Kristiansen CK, Vatne GH, Baumann A, Hallin EI, et al. Molecular
929 mechanisms of Charcot-Marie-Tooth neuropathy linked to mutations in human myelin protein P2. Sci Rep.
930 2017;7(1):6510. Epub 20170726. doi: 10.1038/s41598-017-06781-0. PubMed PMID: 28747762; PubMed
931 Central PMCID: PMCPMC5529448.
932 57. Uusitalo M, Klenow MB, Laulumaa S, Blakeley MP, Simonsen AC, Ruskamo S, et al. Human
933 myelin protein P2: from crystallography to time-lapse membrane imaging and neuropathy-associated
934 variants. FEBS J. 2021;288(23):6716-35. Epub 20210714. doi: 10.1111/febs.16079. PubMed PMID: 34138518.
935 58. Chung KW, Kim SM, Sunwoo IN, Cho SY, Hwang SJ, Kim J, et al. A novel GDAP1 Q218E mutation
936 in autosomal dominant Charcot-Marie-Tooth disease. J Hum Genet. 2008;53(4):360-4. Epub 20080131. doi:
937 10.1007/s10038-008-0249-3. PubMed PMID: 18231710.
938 59. Crimella C, Tonelli A, Airoldi G, Baschirotto C, D'Angelo MG, Bonato S, et al. The GST domain
939 of GDAP1 is a frequent target of mutations in the dominant form of axonal Charcot Marie Tooth type 2K. J
940 Med Genet. 2010;47(10):712-6. Epub 20100803. doi: 10.1136/jmg.2010.077909. PubMed PMID: 20685671.
941 60. Kabzinska D, Kotruchow K, Cegielska J, Hausmanowa-Petrusewicz I, Kochanski A. A severe
942 recessive and a mild dominant form of Charcot-Marie-Tooth disease associated with a newly identified
943 Glu222Lys GDAP1 gene mutation. Acta Biochim Pol. 2014;61(4):739-44. Epub 20141022. PubMed PMID:
944 25337607.
945 61. Moroni I, Morbin M, Milani M, Ciano C, Bugiani M, Pagliano E, et al. Novel mutations in the
946 GDAP1 gene in patients affected with early-onset axonal Charcot-Marie-Tooth type 4A. Neuromuscul Disord.
947 2009;19(7):476-80. Epub 20090604. doi: 10.1016/j.nmd.2009.04.014. PubMed PMID: 19500985.
948 62. Laimer J, Hiebl-Flach J, Lengauer D, Lackner P. MAESTROweb: a web server for structure-based
949 protein stability prediction. Bioinformatics. 2016;32(9):1414-6. doi: 10.1093/bioinformatics/btv769. PubMed
950 PMID: WOS:000376106100021.
951 63. Parthiban V, Gromiha MM, Schomburg D. CUPSAT: prediction of protein stability upon point
952 mutations. Nucleic Acids Res. 2006;34(Web Server issue):W239-42. doi: 10.1093/nar/gkl190. PubMed PMID:
953 16845001; PubMed Central PMCID: PMCPMC1538884.
954 64. Spang A, Saw JH, Jorgensen SL, Zaremba-Niedzwiedzka K, Martijn J, Lind AE, et al. Complex
955 archaea that bridge the gap between prokaryotes and eukaryotes. Nature. 2015;521(7551):173-9. Epub
956 20150506. doi: 10.1038/nature14447. PubMed PMID: 25945739; PubMed Central PMCID:
957 PMCPMC4444528.
958 65. Pedrola L, Espert A, Valdes-Sanchez T, Sanchez-Piris M, Sirkowski EE, Scherer SS, et al. Cell
959 expression of GDAP1 in the nervous system and pathogenesis of Charcot-Marie-Tooth type 4A disease. J Cell
960 Mol Med. 2008;12(2):679-89. Epub 20071116. doi: 10.1111/j.1582-4934.2007.00158.x. PubMed PMID:
961 18021315; PubMed Central PMCID: PMCPMC2570022.

962 66. Raasakka A, Ruskamo S, Barker R, Krokengen OC, Vatne GH, Kristiansen CK, et al. Neuropathy-
963 related mutations alter the membrane binding properties of the human myelin protein P0 cytoplasmic tail.
964 PLoS One. 2019;14(6):e0216833. Epub 20190607. doi: 10.1371/journal.pone.0216833. PubMed PMID:
965 31173589; PubMed Central PMCID: PMCPMC6555526.
966 67. Sherman DL, Brophy PJ. A murine model of Charcot-Marie-Tooth disease 4F reveals a role for
967 the C-terminus of periaxin in the formation and stabilization of Cajal bands. Wellcome Open Res. 2018;3:20.
968 Epub 20180301. doi: 10.12688/wellcomeopenres.13673.1. PubMed PMID: 29623298; PubMed Central
969 PMCID: PMCPMC5861512.
970 68. Marchesi C, Milani M, Morbin M, Cesani M, Lauria G, Scaioli V, et al. Four novel cases of
971 periaxin-related neuropathy and review of the literature. Neurology. 2010;75(20):1830-8. doi:
972 10.1212/WNL.0b013e3181fd6314. PubMed PMID: 21079185.
973 69. Raasakka A, Linxweiler H, Brophy PJ, Sherman DL, Kursula P. Direct Binding of the Flexible C-
974 Terminal Segment of Periaxin to beta4 Integrin Suggests a Molecular Basis for CMT4F. Front Mol Neurosci.
975 2019;12:84. Epub 20190409. doi: 10.3389/fnmol.2019.00084. PubMed PMID: 31024253; PubMed Central
976 PMCID: PMCPMC6465933.
977 70. Azzedine H, Ruberg M, Ente D, Gilardeau C, Perie S, Wechsler B, et al. Variability of disease
978 progression in a family with autosomal recessive CMT associated with a S194X and new R310Q mutation in
979 the GDAP1 gene. Neuromuscul Disord. 2003;13(4):341-6. PubMed PMID: 12868504.
980 71. Nykamp K, Anderson M, Powers M, Garcia J, Herrera B, Ho YY, et al. Sherloc: a comprehensive
981 refinement of the ACMG-AMP variant classification criteria. Genet Med. 2017;19(10):1105-17. Epub
982 20170511. doi: 10.1038/gim.2017.37. PubMed PMID: 28492532; PubMed Central PMCID: PMCPMC5632818.
983 72. Ammar N, Nelis E, Merlini L, Barisic N, Amouri R, Ceuterick C, et al. Identification of novel
984 GDAP1 mutations causing autosomal recessive Charcot-Marie-Tooth disease. Neuromuscul Disord.
985 2003;13(9):720-8. doi: 10.1016/s0960-8966(03)00093-2. PubMed PMID: 14561495.
986 73. Cassereau J, Chevrollier A, Codron P, Goizet C, Gueguen N, Verny C, et al. Oxidative stress
987 contributes differentially to the pathophysiology of Charcot-Marie-Tooth disease type 2K. Experimental
988 neurology. 2020;323:113069. doi: 10.1016/j.expneurol.2019.113069.
989 74. Chung KW, Hyun YS, Lee HJ, Jung HK, Koo H, Yoo JH, et al. Two recessive intermediate Charcot-
990 Marie-Tooth patients with GDAP1 mutations. J Peripher Nerv Syst. 2011;16(2):143-6. doi: 10.1111/j.1529-
991 8027.2011.00329.x. PubMed PMID: 21692914.
992 75. Volodarsky M, Kerkhof J, Stuart A, Levy M, Brady LI, Tarnopolsky M, et al. Comprehensive
993 genetic sequence and copy number analysis for Charcot-Marie-Tooth disease in a Canadian cohort of 2517
994 patients. J Med Genet. 2021;58(4):284-8. Epub 20200506. doi: 10.1136/jmedgenet-2019-106641. PubMed
995 PMID: 32376792.
996 76. Nelis E, Erdem S, Van Den Bergh PY, Belpaire-Dethiou MC, Ceuterick C, Van Gerwen V, et al.
997 Mutations in GDAP1: autosomal recessive CMT with demyelination and axonopathy. Neurology.
998 2002;59(12):1865-72. doi: 10.1212/01.wnl.0000036272.36047.54. PubMed PMID: 12499475.
999 77. Antoniadi T, Buxton C, Dennis G, Forrester N, Smith D, Lunt P, et al. Application of targeted
1000 multi-gene panel testing for the diagnosis of inherited peripheral neuropathy provides a high diagnostic yield
1001 with unexpected phenotype-genotype variability. BMC Med Genet. 2015;16:84. Epub 20150921. doi:
1002 10.1186/s12881-015-0224-8. PubMed PMID: 26392352; PubMed Central PMCID: PMCPMC4578331.
1003 78. Kabzinska D, Niemann A, Drac H, Huber N, Potulska-Chromik A, Hausmanowa-Petrusewicz I,
1004 et al. A new missense GDAP1 mutation disturbing targeting to the mitochondrial membrane causes a severe
1005 form of AR-CMT2C disease. Neurogenetics. 2011;12(2):145-53. Epub 20110302. doi: 10.1007/s10048-011-
1006 0276-7. PubMed PMID: 21365284.
1007 79. Rzepnikowska W, Kaminska J, Kabzinska D, Kochanski A. Pathogenic Effect of GDAP1 Gene
1008 Mutations in a Yeast Model. Genes (Basel). 2020;11(3). Epub 20200314. doi: 10.3390/genes11030310.
1009 PubMed PMID: 32183277; PubMed Central PMCID: PMCPMC7140815.
1010 80. Kursula I, Heape AM, Kursula P. Crystal structure of non-fused glutathione S-transferase from
1011 Schistosoma japonicum in complex with glutathione. Protein Pept Lett. 2005;12(7):709-12. doi:
1012 10.2174/0929866054696154. PubMed PMID: 16522189.

1013 81. Boerkoel CF, Takashima H, Nakagawa M, Izumo S, Armstrong D, Butler I, et al. CMT4A:
1014 identification of a Hispanic GDAP1 founder mutation. Ann Neurol. 2003;53(3):400-5. doi: 10.1002/ana.10505.
1015 PubMed PMID: 12601710.
1016 82. Claramunt R, Pedrola L, Sevilla T, Lopez de Munain A, Berciano J, Cuesta A, et al. Genetics of
1017 Charcot-Marie-Tooth disease type 4A: mutations, inheritance, phenotypic variability, and founder effect.
1018 Journal of medical genetics. 2005;42(4):358-65. doi: 10.1136/jmg.2004.022178.
1019 83. Manganelli F, Pisciotta C, Nolano M, Capponi S, Geroldi A, Topa A, et al. A novel autosomal
1020 dominant GDAP1 mutation in an Italian CMT2 family. J Peripher Nerv Syst. 2012;17(3):351-5. doi:
1021 10.1111/j.1529-8027.2012.00414.x. PubMed PMID: 22971097.
1022 84. Cantarero L, Juárez-Escoto E, Civera-Tregón A, Rodríguez-Sanz M, Roldán M, Benítez R, et al.
1023 Mitochondria-lysosome membrane contacts are defective in GDAP1-related Charcot-Marie-Tooth disease.
1024 Human Molecular Genetics. 2021;29(22):3589-605. doi: 10.1093/hmg/ddaa243.

1025