

RESEARCH ARTICLE

The effect of the pathological V72I, D109N and T190M missense mutations on the molecular structure of α -dystroglycan

Sonia Covaceuszach^{1☯*}, Manuela Bozzi^{2,3☯}, Maria Giulia Bigotti⁴, Francesca Sciandra³, Petr V. Konarev^{5,6}, Andrea Brancaccio^{3,4*}, Alberto Cassetta¹

1 Istituto di Cristallografia—CNR, Trieste Outstation, Trieste, Italy, **2** Istituto di Biochimica e Biochimica Clinica, Università Cattolica del Sacro Cuore, Roma, Italy, **3** Istituto di Chimica del Riconoscimento Molecolare—CNR c/o Università Cattolica del Sacro Cuore, Roma, Italy, **4** School of Biochemistry, University of Bristol, Bristol, United Kingdom, **5** A.V. Shubnikov Institute of Crystallography of Federal Scientific Research Centre “Crystallography and Photonics” of Russian Academy of Sciences, Moscow, Russia, **6** National Research Centre “Kurchatov Institute”, Moscow, Russia

☯ These authors contributed equally to this work.

* sonia.covaceuszach@ts.ic.cnr.it (SC); andrea.brancaccio@icrm.cnr.it (AB)



OPEN ACCESS

Citation: Covaceuszach S, Bozzi M, Bigotti MG, Sciandra F, Konarev PV, Brancaccio A, et al. (2017) The effect of the pathological V72I, D109N and T190M missense mutations on the molecular structure of α -dystroglycan. PLoS ONE 12(10): e0186110. <https://doi.org/10.1371/journal.pone.0186110>

Editor: Danilo Roccatano, University of Lincoln, UNITED KINGDOM

Received: June 12, 2017

Accepted: September 25, 2017

Published: October 16, 2017

Copyright: © 2017 Covaceuszach et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: The crystallographic structures reported in this article have been deposited within the Protein Data Bank (<http://www.rcsb.org/pdb/home/home.do>) with codes 5N30 (V72I) and 5N4H (D109N).

Funding: Funded by AFMTELETHON grant n. 20009 to AB; Wellcome Trust Career Re-entry fellowship to MGB. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Abstract

Dystroglycan (DG) is a highly glycosylated protein complex that links the cytoskeleton with the extracellular matrix, mediating fundamental physiological functions such as mechanical stability of tissues, matrix organization and cell polarity. A crucial role in the glycosylation of the DG α subunit is played by its own N-terminal region that is required by the glycosyltransferase LARGE. Alteration in this O-glycosylation deeply impairs the high affinity binding to other extracellular matrix proteins such as laminins. Recently, three missense mutations in the gene encoding DG, mapped in the α -DG N-terminal region, were found to be responsible for hypoglycosylated states, causing congenital diseases of different severity referred as primary dystroglycanopathies. To gain insight on the molecular basis of these disorders, we investigated the crystallographic and solution structures of these pathological point mutants, namely V72I, D109N and T190M. Small Angle X-ray Scattering analysis reveals that these mutations affect the structures in solution, altering the distribution between compact and more elongated conformations. These results, supported by biochemical and biophysical assays, point to an altered structural flexibility of the mutant α -DG N-terminal region that may have repercussions on its interaction with LARGE and/or other DG-modifying enzymes, eventually reducing their catalytic efficiency.

Introduction

Dystroglycan (DG), a ubiquitous membrane receptor belonging to the glycoprotein complex associated to dystrophin, plays a crucial role in the stability of the plasma membrane, especially in skeletal muscle tissues where it is highly expressed [1]. A single gene, *DAG1*, encodes for a protein precursor, which is proteolytically cleaved in two subunits, alpha-dystroglycan (α -DG), located at the extracellular side of the plasma membrane where it binds several extracellular

Competing interests: The authors have declared that no competing interests exist.

matrix proteins, and beta-dystroglycan (β -DG), which is a transmembrane protein that interacts with dystrophin in the cytoplasm [2]. Many studies have been devoted to the structural characterization of the α -DG subunit. A pioneering electron microscopy study of α -DG revealed its dumbbell-like shape [3]. The molecular structure of the N-terminal region of α -DG (α -DG-Nt) was later determined at high resolution by X-ray crystallography, revealing the presence of two domains, an immunoglobulin-like (Ig-like) domain and a domain similar to the small subunit ribosomal protein S6 of *T. thermophilus* (S6 domain), connected by a flexible loop [4]. The high resolution structure of the α -DG C-terminal domain is still elusive, although a computational study envisaged that this domain is also likely to possess an Ig-like fold, followed by a disordered part at the C-terminus [5]. A biochemical characterization confirmed a highly disordered conformation for the mucin-like central domain of α -DG [6], that is decorated and stabilized *in vivo* by specific O-mannosyl glycans built by a complex array of enzymes [7]. Among these, particularly worth of mention is the bifunctional Like-acetylglucosaminyltransferase (LARGE), which adds the repeating heterodisaccharide [-glucuronic acid- β 1,3-xylose- α 1,3-]_n to a glycan anchored at Thr317 and Thr319 in the mucin-like domain [8]. This region of the protein plays a crucial functional role, as its protruding carbohydrate moieties are known to mediate the interactions with laminin and other extracellular proteins containing laminin globular (LG) domains [9]. Indeed, many congenital muscular dystrophies, due to mutations affecting the genes encoding for glycosyltransferases belonging to the glycosylation pathway of α -DG, are characterized by hypoglycosylated forms of α -DG that are unable to interact with laminin, compromising the stability of sarcolemma [10]. Noteworthy, it has been shown that α -DG-Nt is necessary for recruiting LARGE [11] and some pathological conditions that are caused by missense mutations in the DAG1 map within the α -DG-Nt.

The first identified mutation (T192M), associated to a form of limb-girdle muscular dystrophy and cognitive impairment, prevents the correct glycosylation of α -DG resulting in impaired laminin binding. It was shown that the T192M mutation weakens the interaction between α -DG-Nt and LARGE, strongly reducing the ability of the latter to decorate α -DG with the proper carbohydrate moiety [12]. More recently, two additional mutations, namely V74I and D111N, located on the Ig-like domain surface facing the S6 domain within α -DG-Nt, have been found in a seven years old compound heterozygous patient, who displays mild muscular dystrophy and asymptomatic hyperCKemia. Also in this case, a biochemical analysis carried out on the muscular tissues of the patient revealed the presence of a hypoglycosylated form of α -DG that cannot bind to laminin [13].

In order to investigate the molecular mechanisms underlining these diseases, we analyzed the impact of the three afore mentioned pathological missense mutations on the conformational stability and on the overall structure of α -DG-Nt. In particular, we focused on the murine α -DG-Nt region, fully validated as a model for the human counterpart [14], carrying the three pathological missense mutations V72I, D109N and T190M that correspond to their topological counterparts V74I, D111N and T192M in human DG.

Even though the missense mutations V72I and D109N do not alter the overall fold of α -DG-Nt, as previously assessed by X-ray crystallography for the mutant T190M [15], Small Angle X-ray Scattering (SAXS) analysis (supported by limited proteolysis and Differential Scanning Fluorimetry (DSF) experiments) highlights a more complex dynamic behavior of α -DG-Nt in solution that is remarkably affected by the three pathological mutations. The resulting altered conformations may impact on the interactions of α -DG with LARGE, negatively influencing LARGE recruiting and ultimately the proper maturation of α -DG.

Materials and methods

DNA constructs, site-directed mutagenesis and protein expression and purification

The point mutations V72I and D109N were introduced within the murine α -DG(50–313) construct, carrying the additional mutation R166H, to improve its proteolytical stability, (hereinafter WT), cloned in pHis-Trx [4], by the QuikChange site-directed mutagenesis kit (Stratagene) as previously reported for T190M [15]. The following primers were employed to insert the two mutations:

V72I Forward: 5' -CCTGATGGCACGGCTGTCATCGGGCGCTCATTTTCGAGTG-3'

V72I Reverse: 5' -CACTCGAAATGAGCGCCCGATGACAGCCGTGCCATCAGG-3'

D109N Forward: 5' -CCATCTTGGCTGCACTGGAAGTCCAGAGCCACACCCTG-3'

D109N Reverse: 5' -CAGGGTGTGGCTCTGTGAGTTCAGTGCAGCCAAGATGG-3'.

All constructs were verified by automated sequencing.

WT and the three pathological mutant proteins were expressed as N-terminal His₆-tagged thioredoxin fusion products, containing a thrombin cleavage site, and purified before and after thrombin cleavage according to the previously described protocol [15].

Differential scanning fluorimetry (DSF)

DSF experiments were performed using an excitation wavelength of 470–505 nm and an emission wavelength of 540–700 nm in a CFX96 Touch Biorad real-time PCR instrument (Bio-Rad) with a temperature gradient from 20 to 90°C in 0.2°C/min increments. The final protein concentration was 0.5 mg/mL in 20 mM Tris, 150 mM NaCl pH 7.5 (in the case of T190M the buffer was supplemented with 2.5% glycerol) and 90 \times SYPRO Orange (Sigma). Experiments were carried out in triplicate and the averaged curves have been analyzed: melting temperatures (T_m) were calculated by fitting the sigmoidal melt curves to the Boltzmann equation (S1 Fig) [16].

Limited proteolysis

WT and its pathological mutants V72I and D109N, at a final concentration of 0.85 mg/mL, were subjected to limited proteolysis with either α -chymotrypsin and trypsin at a final concentration of 2 μ g/mL. Proteolytic digestions were performed at 37°C and were stopped after 1, 5, 10, 20, 40 and 60 min by adding SDS sample buffer to aliquots of the reaction mixtures. The samples were analyzed by 15% SDS-PAGE [17] and Coomassie staining.

Crystallization, data collection, structure solution and refinement

Crystals of D109N were grown by using the hanging-drop vapour diffusion method, exploring conditions similar to those used for both WT and T190M [4,15]. Drops were prepared by mixing 1 μ L of protein solution (5.5 mg/mL in 25 mM Tris, 150 mM NaCl, pH 7.5) with 1 μ L of precipitant solution (0.6–1.4 M citrate buffer; pH 6.8–7.2) and equilibrated against the reservoirs containing 0.7 mL of the precipitant solution. Fully-grown crystals were obtained after two weeks at 0.7 M citrate buffer and pH 7.0, at the temperature of 277 K. Crystals of V72I did not grow under the same conditions of D109N and were obtained by using the cross-streak-seeding method. D109N crystals were used as seeds source and crystallization conditions between 0.6–1.4 M citrate buffer and pH between 6.8–7.2 were explored. Drops were prepared

by mixing 1 μ L of protein solution (5.0 mg/ mL in 25 mM Tris, 150 mM NaCl and pH 7.5) with 1 μ L of precipitant solution; the drops were equilibrated against 0.7 mL of the precipitant solution at 277 K for 3–6 days, before seeding. Fully-grown crystals were obtained after 10–15 days after seeding (best crystals obtained at 0.8 M citrate buffer, pH 7.2). Repeated streak-seeding (2–3 times) at optimal precipitant conditions improved the crystal quality.

Data collections were carried out at the XRD1 beamline at ELETTRA (Trieste, Italy) [18,19] using Pilatus 2M (Dectris) detector and 1.00 Å (D109N) and 0.976 Å (V72I) wavelength. Data collection were carried out at 100 K. Crystals were quickly dipped into a cryoprotectant solution (25% v/v ethylene glycol added to the precipitant solution) and then frozen directly under a 100 K nitrogen gas stream.

Indexing, integration and data reduction of the diffraction data were carried out by using the XDS program [20]. Two data-set of highly isomorphous V72I crystals were merged with XSCALE [20] for higher data completeness. Data reduction statistics of D109N and V72I data-sets are reported in Table 1.

The structure solution of both D109N and V72I were obtained by Molecular Replacement, using the WT crystal structure (PDB ID: 1U2C [4]) as search-model and PHASER [21] computer program as implemented in PHENIX [22] crystallographic package. Rigid-body refinement was initially carried out, followed by a simulated-annealing step. Several cycles of crystallographic refinement, including positional refinement, Translation-Libration-Screw (TLS) and individual B-factors refinement were alternated with the manual rebuilding of the structure by using the COOT software [23]. All the refinement cycles were carried out by using *phenix.refine* [24]. Solvent molecules were added to the model by using the automatic search protocol available in *phenix.refine* and manually checked before being included in the final model. Protein stereochemistry was monitored throughout the refinement process and during manual rebuilding with MolProbity [25]. Statistics of the crystallographic refinement are reported in Table 1. Various CCP4 [26] utility programs were used throughout the crystallographic study. Molecular diagrams were prepared using the PyMol Molecular Graphics System [27]. The PDB codes of the deposited structures are 5N30 (V72I) and 5N4H (D109N).

Small-angle X-ray scattering measures and data processing

As previously reported for the WT protein [14], SAXS experiments for T190M were carried out at the BM29 beamline [28] of the European Synchrotron Radiation Facility (Grenoble, France) as 10 x 1 s exposures using a Pilatus 1M (Dectris) detector, with a sample-detector distance of 2.87 m and a wavelength of 0.99 Å. SAXS data for V72I and D109N were collected on the P12 beamline EMBL SAXS-WAXS at PETRAIII/DESY [29] (Hamburg, Germany) as 20 x 0.05 s exposures using a Pilatus 2M (Dectris) detector, sample-detector distance 3.00 m, wavelength 1.24 Å. Measurements were performed at six different concentrations (the ranges are reported in S2 Table) in 20 mM Tris, 150 mM NaCl pH 7.5 (supplemented with 2.5% glycerol in the case of T190M); the protein concentration was calculated using molar extinction coefficient at 280 nm ($26595 \text{ M}^{-1} \text{ cm}^{-1}$) and measuring each sample dilution by a nanodrop spectrophotometer. No radiation damage effects were detected comparing the scattering curves of the collected frames.

Data were merged for each sample after normalization to the intensity of the transmitted beam. Subtraction of the scattering of the buffer and the following processing steps were carried out with PRIMUS [30] from the ATSAS 2.6.0 program package [31].

The radius of gyration R_g of the solute proteins and the forward scattering $I(0)$ were evaluated by the Guinier approximation (1939) at very small angles ($s < 1.3/R_g$), assuming that the intensity is represented as $I(s) = I(0) \cdot \exp(-1/3(R_g \cdot s)^2)$ (S2 Fig), and from the entire scattering

Table 1. X-ray diffraction: Data collection and model refinement statistics.

	D109N	V72I
Space group	H3	H3
Unit-cell parameters (Å)		
<i>a</i>	71.75	71.81
<i>c</i>	144.00	143.99
Molecules per asymmetric unit	1	1
Wavelength (Å)	1.0000	0.9763
Resolution (Å)	48.0–1.70	35.9–1.80
Highest resolution shell (Å) ^a	1.76–1.70	1.86–1.80
Total observations	237820	294745
Unique reflections	30354	25641
R _{merge} (%) ^b	0.096 (1.098) ^a	0.105 (1.301) ^a
CC _{1/2} ^c	0.997 (0.883) ^a	0.999 (0.862) ^a
<I/σ(I)>	12.0 (1.9) ^a	15.8 (2.1) ^a
Completeness (%)	99.7 (97.9) ^a	99.9 (99.4) ^a
Redundancy	7.8 (7.5) ^a	11.5 (11.3) ^a
Refinement		
Number of reflections (work-set)	30326 (3042) ^a	25623 (2540) ^a
Number of reflections (test-set)	1516 (152) ^a	1282 (127) ^a
R _{work} ^d	0.159 (0.349) ^a	0.154 (0.292) ^a
R _{free} ^e	0.183 (0.360) ^a	0.182 (0.315) ^a
Number of non-H atoms		
Protein	1765	1738
Waters	196	172
Organic	15	19
Average isotropic B factors (Å ²)		
Protein	33.5	36.2
Solvent	39.9	42.8
R.m.s. deviation		
Bond length (Å)	0.013	0.016
Angle (deg)	1.19	1.34
Ramachandran plot		
favoured regions (%)	98	98
allowed regions (%)	2	2
disallowed regions (%)	0	0

^a Values in parenthesis are given for the highest resolution shell

$$^b R_{\text{merge}} = \frac{\sum_{hkl} \sum_j |I_{hkl,j} - \langle I_{hkl} \rangle|}{\sum_{hkl} \sum_j I_{hkl,j}}$$

^c CC_{1/2} values for *I*_{mean} are calculated by splitting the data randomly in half

$$^d R_{\text{work}} = \frac{\sum_{\text{work-set}} F_{\text{obs}} - F_{\text{cal}}}{\sum_{\text{work-set}} F_{\text{obs}}}$$

^e R_{free} = $\frac{\sum_{\text{test-set}} F_{\text{obs}} - F_{\text{cal}}}{\sum_{\text{test-set}} F_{\text{obs}}}$; Test-set: Randomly selected 5% of reflections excluded from the structure refinement.

<https://doi.org/10.1371/journal.pone.0186110.t001>

pattern by the program GNOM [32], that was used to compute also pair distance distribution functions of the particles *p*(*r*) and the maximum sizes *D*_{max}. Molecular weights (MM) were evaluated by comparison of the calculated *I*(0) values with that of the standard solution of bovine serum albumin (MM 66 kDa). The excluded volume of the hydrated protein molecule (*V*_p) was obtained using the Porod approximation [33]: $V_p = \frac{2\pi^2 I(0)}{\int I_{\text{exp}}(s) s^2 ds}$.

Low resolution shapes of the three mutants were produced by the *ab initio* program DAMMIN [34] that employs a simulated annealing procedure to build a compact dummy atoms (beads) model that fits the experimental data $I_{\text{exp}}(s)$ to minimize the discrepancy:

$$\chi^2 = \frac{1}{N-1} \sum_j \left[\frac{I_{\text{exp}}(s_j) - I_{\text{calc}}(s_j)}{\sigma(s_j)} \right]^2.$$

Ten independent DAMMIN runs were performed for each data set in the “slow” mode with no symmetry assumptions (P1 symmetry); the resulting models were superimposed using the program SUPCOMB [35] and averaged using DAMAVER [36] to identify the most typical models representing the global shape of the three mutants in solution. The normalized spatial discrepancy parameter (NSD) [35] obtained from DAMAVER indicated the similarity between models used for average calculations. NSD values ≤ 1.0 are expected for similar models.

Rigid-body modeling was performed using CORAL [31], where the high-resolution X-ray crystal structures of the Ig-like and S6 domains of the three mutants were used as inputs in the refinement calculations. This program was employed to refine the spatial arrangement of the two domains for the three mutants and to model clash-free configurations of the missing portions of polypeptide chains (around 10 aminoacids at both N-terminal and C-terminal and the missing linker between the Ig-like and S6 domains). CRY SOL [37] was used to evaluate the fits to the experimental data of the respective X-ray crystal structures; CORAL was also used to generate the approximate conformations of the missing regions keeping fixed the two domains as in the respective X-ray crystal structures.

Inter-domain flexibility and size distribution of possible conformers for the three mutants were quantitatively assessed by the ensemble optimization method (EOM) [38]. This method assumes the existence of a mixture of conformations in solution; the average scattering of the mixture fits the experimental data. In EOM, an initial random pool of 10000 conformers was generated. In these conformers the modeled linker residues and the modeled N-terminal and C-terminal stretches were allowed to have random-coil conformations and the S6 domain and the Ig-like domains, obtained by the respective crystal structures of the three mutants, were used as rigid bodies. The theoretical scattering was calculated for each generated model by CRY SOL. A genetic algorithm (GAJOE) was used to select an ensemble of conformations whose mixture best fitted the experimental data. Multiple runs of EOM were performed and the obtained subsets were analyzed to yield the R_g distributions in the selected ensembles. Once each ensemble is determined, the corresponding Shannon Entropy, reported as R_{flex} , provides a quantitative measure of flexibility [39]. Using R_{flex} each ensemble distribution can be numerically compared to that of the respective random pool, the latter representing a reference for flexibility. The complementary metric R_σ (i.e. the ratio of the standard deviation between the ensemble and the pool distributions) allows to identify potential spurious solutions: values close to 1.0 are obtained when the ensemble distribution largely reproduces the conformational space of the random pool.

Sequence alignment

Multiple sequence alignments of protein sequences were constructed in Clustal omega via the resources of EMBL/EBI (<http://www.ebi.ac.uk/Tools/msa/clustalo/>).

Results and discussion

V72I and D109N mutations determine small local variations into the crystal structure of N-terminal region of α -DG

The crystal structures of D109N and V72I have been determined to a resolution of 1.70 Å and 1.80 Å, respectively. Similarly to what observed for T190M [15], residues 50–58 and 305–315,

as well as residues 163–179 (V72I) or 164–179 (D109N), were missing in the D109N and V72I final models. Residues 89–91 and 181–185 belong to highly mobile loops thus showing an inherent lower quality electron density. In addition, the region encompassing residues 159–162 displays evidences of less populated conformations in both structures, which could not be confidently modeled during the structure refinement. That said, both D109N and V72I structures show the typical overall folding of α -DG-Nt [4]. The Ig-like and S6 domain are linked by a flexible loop encompassing residues 159–179, which is only partially visible in the present models. The two pathological point mutations introduced in α -DG-Nt were identified according to the $2F_o-F_c$ and F_o-F_c maps (Fig 1A and 1B) and by their effects on the local molecular geometry.

According to the crystal structures, the two point mutations do not induce significant changes in the domains fold nor in their mutual orientation when compared with WT. Indeed, the superposition of the structure of WT with those of D109N and V72I gives root mean square deviations of 0.353 Å (calculated on 227 C α) and 0.417 Å (calculated on 226 C α), respectively, in line with the superposition of the WT and T190M structures (0.758 Å) [15].

The Asp109→Asn mutation introduces local changes on the molecular geometry that are quite evident in the final refined model. The Asn109 side chain rotates nearly 120° around C α -C β with respect to the orientation held by Asp109 in the WT. The conformation assumed by Asn109 allows for the establishment of H-bonds with Lys203, Glu116 and His107. Accordingly, Asn109 is not engaged anymore in H-bonds with Ser112 and His111 (Fig 1A) as Asp109 is in WT. Quite surprisingly, this newly established H-bonds network does not affect the geometry of the now untied turn including His111 and Ser112 that connect the two β -strands D and E.

Furthermore, Asn109 also interacts with Lys203, which belongs to the S6 domain. According to its electron density and B-factors, Lys203 is inherently rather flexible and it is plausible that the strength of the interaction between Lys203 and Asn109 would be rather weak. In addition, the H-bond network involving Asn109, Glu116 and Lys203 in D109N abolishes the interaction between Lys203 and Glu116 observed in the WT structure. Hence, it is likely that the mutated H-bond network in D109N alters the tightness of the interaction between the Ig-like and the S6 domains. In this respect, it must be emphasized that the Asp109→Asn mutation not only changes the H-bond network involving residue 109, but also modifies the local electrostatic potential (Fig 1C and 1D), so that the interaction of Asn109 with the surrounding positively charged residues like Lys203 and Arg77 is affected. This is confirmed by the presence of two water molecules (HOH 548 and 621) absent in WT structure and coordinating Arg77, thus filling the void left by the Asn109 side-chain relocation. It is therefore likely that the change in electrostatic potential observed in D109N also affects the tightness of the interaction between the Ig-like and the S6 domains.

According to the refined crystallographic model, the pathological mutation Val72→Ile has no effect on the overall conformation of α -DG-Nt, but in this case also the local effects on the structure around the mutated residue are quite limited. The bulkier isoleucine residue might perturb the N-terminal part of the flexible linker connecting the two domains, with some effects on Lys124 and Asp121 (Fig 1B and S3 Fig). These residues are solvent exposed and display, in WT, a remarkable mobility. It is worth noting that the linker region is one of the structural elements of α -DG-Nt with the highest flexibility: it is therefore difficult to confidently assess the effect of the Val72→Ile mutation on the linker conformation, considering that it is only partially defined in the refined V72I model. It is also worth of note that the conformation of the modeled part of the linker in V72I is essentially the same as that observed in T190M, but rather different from what observed in WT, pointing to an effect not related to the specific presence of Ile72.

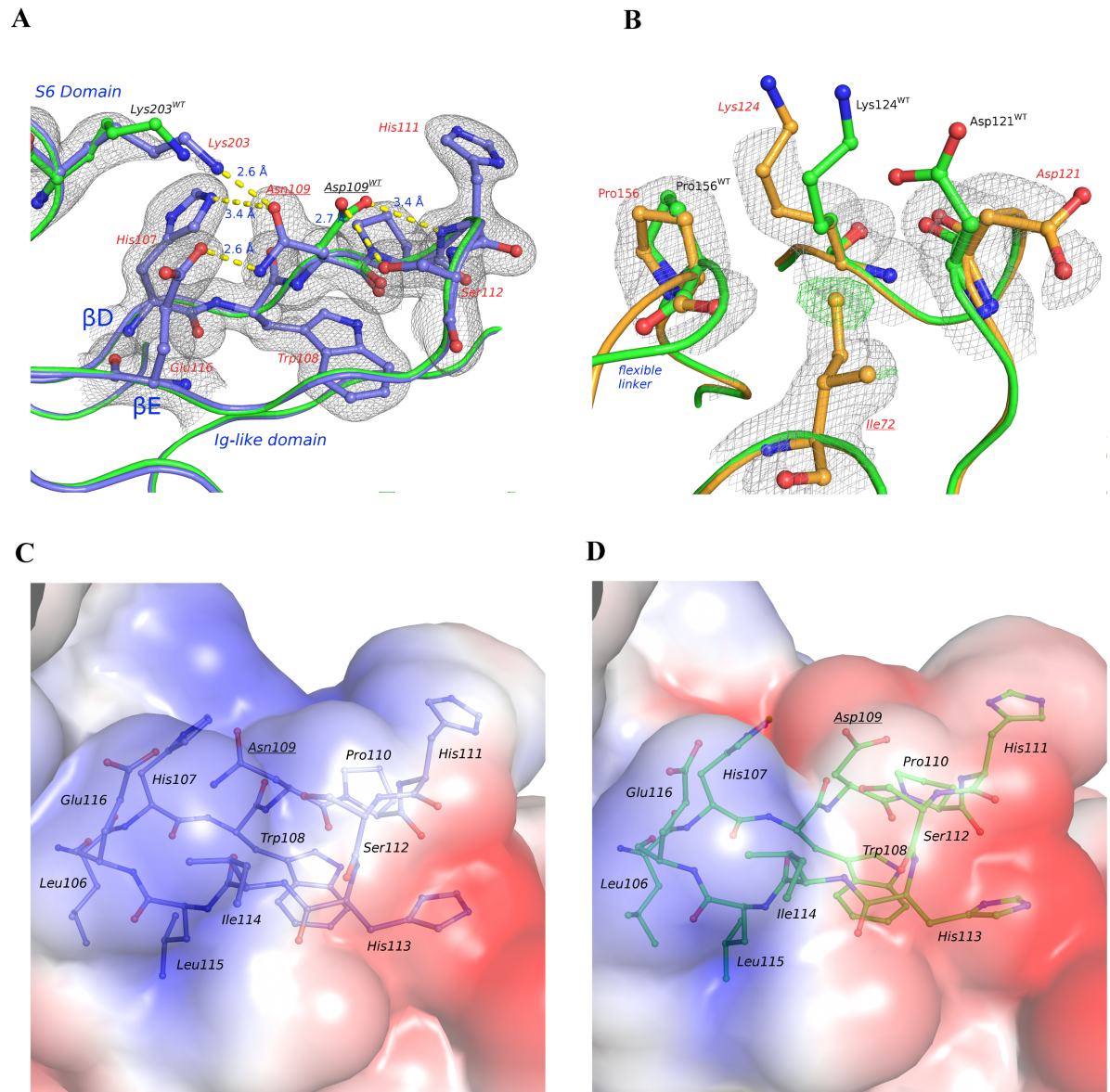


Fig 1. V72I and D109N crystal structures. Details of the superimposition of the WT with D109N (A) and V72I crystal structures (B). WT, D109N and V72I are represented as stick-and-ball. Selected residues are represented as stick-and-ball, with WT (PDB ID: 1U2C) colored in green, D109N colored in blue and V72I colored in orange. Residues belonging to WT are labelled with WT superscript in panels (A) and (B). The models are overlaid by D109N σ_A -weighted $2F_o-F_c$ map (contoured at 1.0σ and colored in grey) (A) and by V72I σ_A -weighted $2F_o-F_c$ map (contoured at 1.0σ and colored in grey) and F_o-F_c map (contoured at 3.0σ and colored in green) (B). Electrostatic potential maps of WT and D109N are represented in panel (C) and (D) respectively. The electrostatic potential (in $k_B T/e_c$ units) is mapped on solvent-accessible surface of the WT and D109N pathological mutant accessible surfaces. Negative potential is colored in blue, positive potential in red. Color scale varies between -2 and +2. Molecular models (stretch 105–116) are represented as ribbons with selected residues depicted as stick-and-ball.

<https://doi.org/10.1371/journal.pone.0186110.g001>

Conformational stability of the pathological mutants

Conformational stabilities of the three pathological mutants as compared to WT were assessed by limited proteolysis and DSF assays.

Considering that conformational parameters such as solvent accessibility and segmental mobility are known to correlate with exposed proteolytic sites [40], limited proteolysis analysis

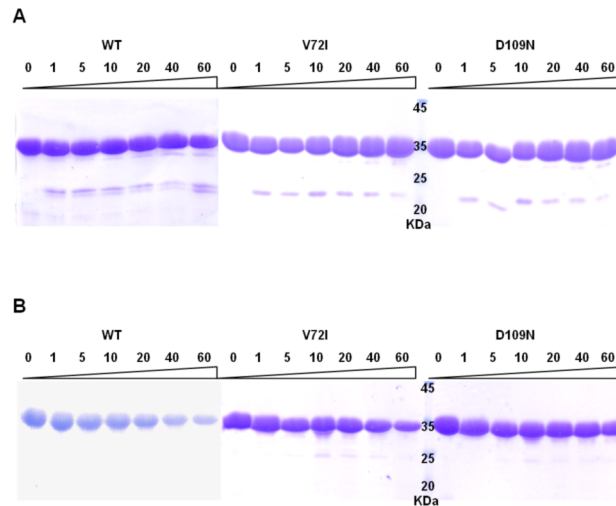


Fig 2. Limited proteolysis analysis. Protease digestion of α -DG-Nt WT and its pathological mutants V72I and D109N by α -chymotrypsin (A) and trypsin (B) for 1, 5, 10, 20, 40 and 60 min. Undigested samples served as the zero time point (0).

<https://doi.org/10.1371/journal.pone.0186110.g002>

was performed with two different proteases (respectively α -chymotrypsin and trypsin) in order to identify potential flexible and exposed regions. The results shown in Fig 2 suggest only slight differences in susceptibility to proteolysis of V72I and D109N with respect to WT protein, not as striking as those recorded in a comparative tryptic analysis of the T190M mutant [15].

In order to investigate further whether the pathological mutations are associated with variation in conformational stability, DSF was performed. This assay monitors thermally-induced protein unfolding by the binding of a fluorescent dye to the hydrophobic core of the protein as it becomes exposed, with an increase in fluorescence emission as the dye binds. From the unfolding curve thus the T_m of a protein can be calculated, which is indicative of its thermal and conformational stability: a shift towards lower temperatures in the T_m of a protein variant relative to that of the WT is evidence of its destabilization [16]. Therefore, the analysis of the DSF curves and T_m s are especially useful for a qualitative comparison of the mutants thermal stabilities. Fig 3 compares the changes in fluorescent signal of the WT and mutant proteins during thermal unfolding in the presence of the dye. WT, V72I and D109N show very low and flat background fluorescence in the pre-transition region, while T190M is characterized by a quite high fluorescence in the pre-transition region, often symptom of exposed hydrophobic residues [41]. This finding is in accordance with the substitution of the polar side chain of Thr with a bulkier and apolar group such as the methylthio group of the Met that can create a binding site for the hydrophobic dye in the native state. Even if all the resulting denaturation curves point to a two-transitions unfolding process, as it has been observed in several proteins composed by two domains that independently fold, the profiles are dramatically different, suggesting that the single point mutations deeply influence the protein conformation/stability in solution. In particular, the first transition is much sharper (i.e. more cooperative) for V72I and D109N than for T190M and WT. The T_m values (summarized in the insert of Fig 3) could be calculated, by fitting the data to a Boltzmann Sigmoid only for the first transition since the second is not sufficiently defined. The fitting was especially problematic for the T190M case, due to its large pre-transition fluorescence. Two mutants (D109N and T190M) were characterized by an increase in T_m relative to WT, while for V72I a small decrease was observed. The second

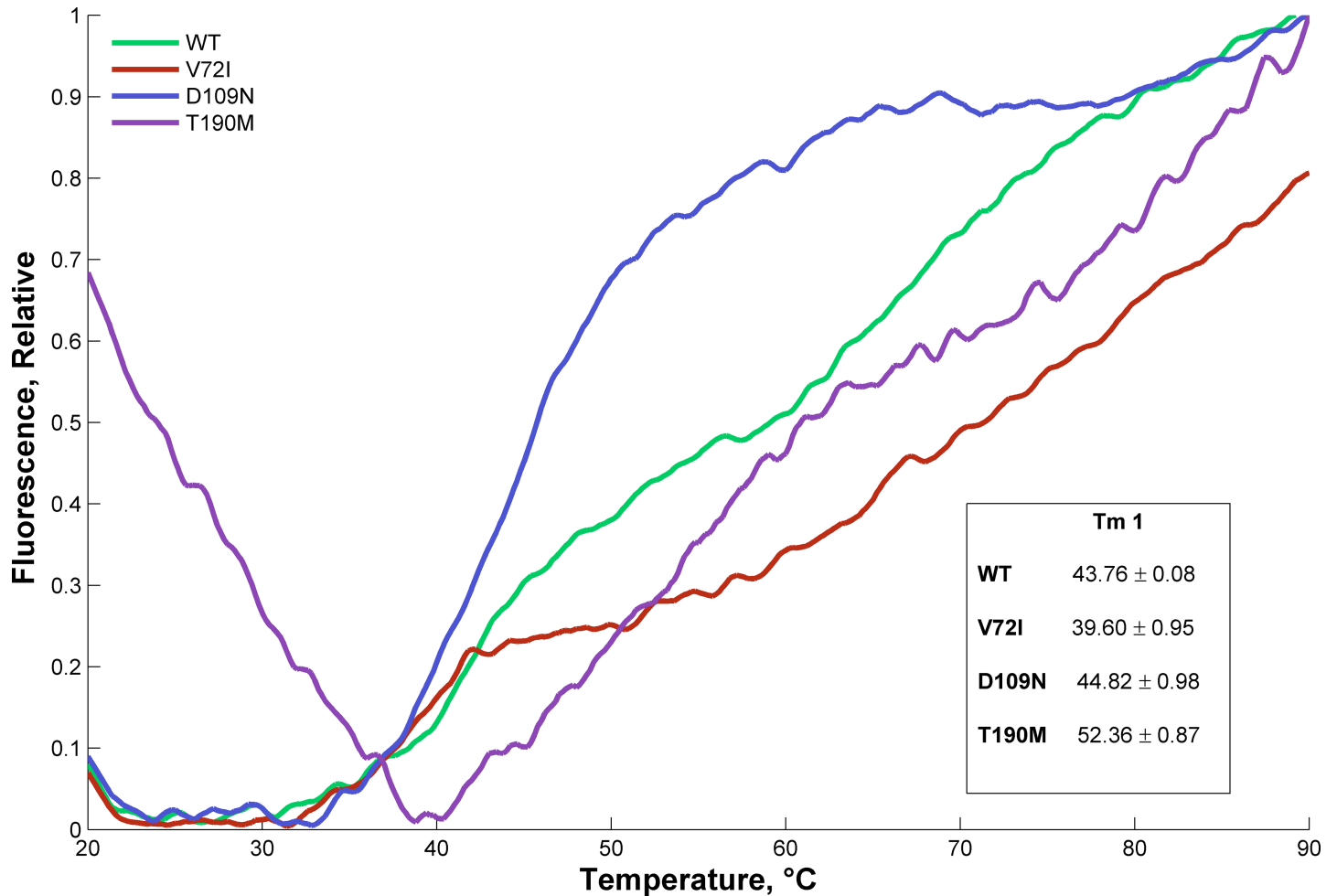


Fig 3. Thermal denaturation assay using DSF with SYPRO dye. Comparison of unfolding fluorescent curves for the WT α -DG-Nt and its pathological mutants; the melting temperature (T_m) values from the Boltzmann Sigmoid fitting of the first transition of the curves are reported in the insert. Experiments were performed in triplicate and the averaged curves are shown.

<https://doi.org/10.1371/journal.pone.0186110.g003>

transition profiles suggest an increased T_m of the mutants with respect to the WT, although in this cases the T_m could not be calculated because it was not possible to reach the aggregation region (*i.e.* the endpoint of the transition), even increasing the final temperature to the upper range value of the instrument.

Association state and overall size parameters of the pathological mutants in solution

In order to assess whether the single point mutations associated to pathological states impact on the conformation of α -DG-Nt in solution, SAXS experiments were performed on the three mutants at different concentrations (S1 Table) without observing systematic changes due to solute concentrations or any interparticle interaction (S4 Fig): therefore the analysis was performed on the curves collected for most concentrated samples. Fig 4 displays the processed scattering data collected for the highest concentration of each mutant, and S1 Table compares the resulting overall size parameters to previously collected WT data [14] (additional details on SAXS structural parameters are reported in S2 Table).

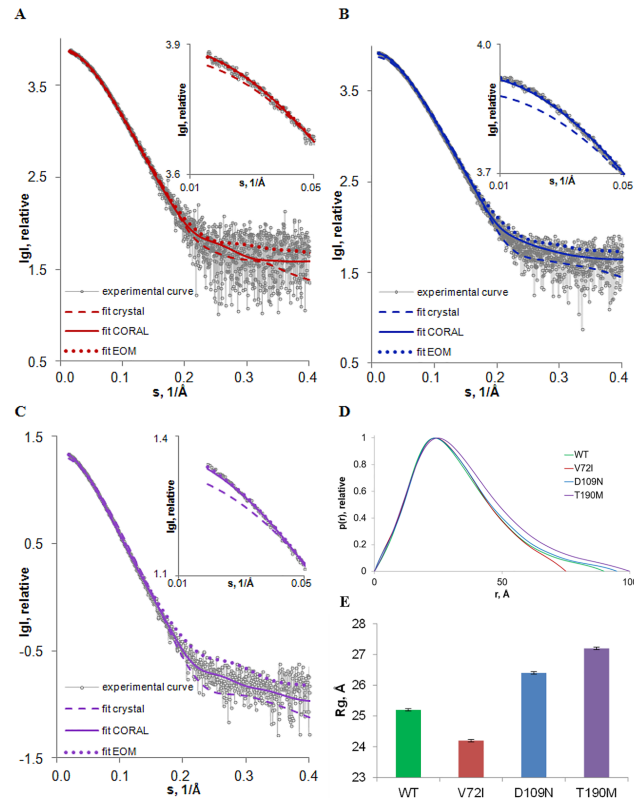


Fig 4. Experimental X-ray scattering data and the obtained fits of the three pathological mutants of α -DG-Nt: A) V72I; B) D109N and C) T190M. Experimental SAXS patterns, scattering calculated from the respective crystallographic models with added missing regions obtained by CORAL (fit crystal) keeping fixed the two domains and scattering calculated from models obtained by CORAL (fit CORAL) are plotted as indicated. The plots display the logarithm of the scattering intensity as a function of momentum transfer $s = [4\pi\sin(\theta/2)]/\lambda$ (\AA^{-1}), where θ is the scattering angle and λ is the X-ray wavelength. The zoomed regions of these graphs at low angles are presented in the inserts. The comparison of the distance distribution functions obtained for the WT and the three pathological mutants of α -DG-Nt are presented in panel (D), the corresponding R_g values are plotted in panel (E).

<https://doi.org/10.1371/journal.pone.0186110.g004>

The V_p and the MM calculated for all mutants (S1 Table) were consistent with the values expected for a monomeric species even at relatively high (up to 7 mg/ml) concentrations and are in agreement with the MM estimated from the primary sequences (around 28.5 kDa).

The computed distance distribution functions $p(r)$ are compared to the profile obtained for the WT protein in Fig 4D. All the profiles display a single peak with a tail, a pattern indicative of proteins having elongated structures, but the significant change in the maximum dimensions of the mutants, combined with a significant variation in the radius of gyration (R_g) (Fig 4E and S1 Table), are further evidence that the pathological point mutations affect the conformation of α -DG-Nt.

In details, the mutations Asp109 \rightarrow Asn and Thr190 \rightarrow Met led to a significant increase in the maximum dimensions (with D_{max} increased from $90 \pm 3\text{\AA}$ for WT to $95 \pm 3\text{\AA}$ and to $100 \pm 3\text{\AA}$ for D109N and T190M respectively) and in the R_g values (from $25.20 \pm 0.04\text{\AA}$ for the WT to $26.40 \pm 0.04\text{\AA}$ and to $27.20 \pm 0.04\text{\AA}$ for D109N and T190M, respectively), suggesting a more extended shape of the two pathological mutants compared to the WT protein. On the contrary, the single mutant V72I shows a decrease both in D_{max} (from $90 \pm 3\text{\AA}$ for WT to $80 \pm 2\text{\AA}$ for V72I) and in the R_g value (from $25.20 \pm 0.04\text{\AA}$ for the WT to $24.20 \pm 0.04\text{\AA}$ for V72I), pointing to a more compact shape as compared to the WT protein.

Similarly to what determined by the SAXS experiments performed on the WT protein [14], all the mutants in solution have a significantly more elongated conformation than the respective crystallographic models (S5 Fig). Indeed, the scattering curves computed by the CRY SOL program [37] from the crystallographic models (PDB_IDs: 5N30 for V72I, 5N4H for D109N and 4WIK for T190M [15]) give a poor fit to the experimental data (not shown), even after the reconstruction of the missing regions (around 10 aminoacids at both N-terminal and C-terminal and the missing linker between the two domains that are kept fixed) using the program CORAL [31] (S1 Table X_{crystal} , Fig 4 $\text{fit}_{\text{crystal}}$ with the zoomed portions at low angles in the inserts). Despite SAXS cannot characterize a molecular structure at a resolution level comparable to that provided by X-ray crystallography, it allows to define structural models in solution devoid of the constraints imposed by the packing forces featuring crystal structures. In our models packing forces may influence the relative orientation of the Ig-like and S6 domains that in solution gain a certain degree of freedom assured by the highly flexible loop connecting them. This may have relevant effects on the overall conformation of the α -DG N-terminal domain that is expected to influence the shape of the molecule and, accordingly, the very low-resolution part of the scattering curve.

Effects of the pathological mutations on the shape of the N-terminal region of α -DG in solution

The macromolecular shapes of the three mutant proteins in solution have been reconstructed in parallel by *ab initio* modeling and by rigid-body modeling.

All the 10 independent solutions for each mutant (Fig 5), reconstructed from the X-ray scattering data using DAMMIN [34], showed a good fit to the experimental curves ($X_{\text{ab-initio}}$ in S1 Table). Thus they were averaged to obtain the final low-resolution models of the three mutant proteins with quite low NSD values (0.698 ± 0.018 , 0.570 ± 0.039 and 0.533 ± 0.031 , for V72I, D109N and T190M, respectively), indicating that the multiple solutions built by the program are very similar to each other. The comparison of the resulting *ab initio* models with that of the WT suggests that small but significant rearrangements (at low resolution) in the orientation of the two domains occur, leading to a more extended structure for D109N and T190M and to a more compact conformation for V72I (Fig 5).

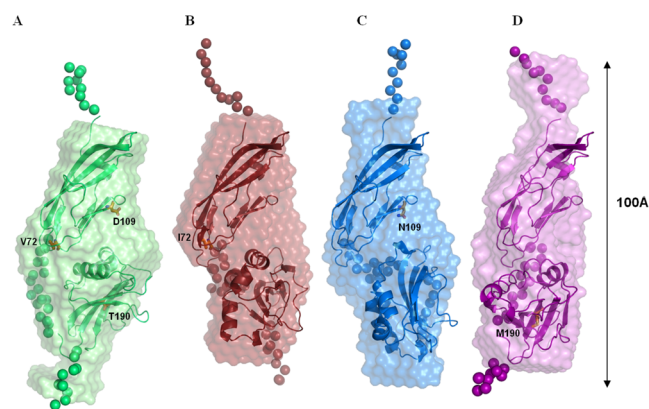


Fig 5. Comparison of the structural models of the WT and the three pathological mutants of α -DG-Nt. Averaged and filtered *ab initio* bead models obtained with DAMMIN (semitransparent surfaces) overlaid to the typical CORAL models (green, red, blue and violet cartoon representations for the folded Ig-like and S6 domains, spheres for the restored missing fragments of WT (A), V72I (B), D109N (C) and T190M (D), respectively). Residues involved in the three pathological mutations are represented by sticks and colored in orange.

<https://doi.org/10.1371/journal.pone.0186110.g005>

In order to get more detailed information exploiting the respective high-resolution crystal structures, with the Ig-like and S6 domains treated as rigid bodies, the program CORAL [31] was used to optimize the relative orientations of the two domains and to reconstruct the missing regions. Multiple runs were performed and yielded variable conformations, all providing good fits to the experimental data (i.e. X in the range of 1.0–1.1 for V72I, 1.1–1.6 for D109N and 1.1–1.7 for T190M). Even if this variety of configurations suggests a significant flexibility of this region (see below for the discussion of the inter-domain flexibility), the good fit of the best rigid-body models of the mutants (X_{CORAL} in S1 Table and $\text{fit}_{\text{CORAL}}$ in Fig 4) indicates that these models provide a good representation of the average conformations. These results confirm that the pathological mutant proteins have a more extended shape in solution than that observed in the crystal lattice (as previously assessed for the WT protein). One of the possible reasons for that is the absence of the packing forces in solution together with the two-domain structure of the proteins which are connected by the flexible linker. In this respect it is interesting to note that the positions of the centers of masses of the S6 domains in all the crystallographic models nicely overlap (S6 Fig) and are much closer to the center of mass of the Ig-like domain than the respective CORAL models. As a quantitative measure of structure compactness, the distances between centers of masses of the two domains in the crystallographic models (29.9 Å, 29.7 Å, 29.6 Å and 29.6 Å in the WT, D109N, V72I and T190M models, respectively) have been compared to the respective CORAL models (34.4 Å, 36.6 Å, 33.0 Å and 37.4 Å in the WT, D109N, V72I and T190M models, respectively), confirming the existence in solution of rather elongated conformations that are more flexible than what their crystal structures suggested. Moreover, this comparison highlights that the three mutations significantly impact on the protein conformation, in agreement with what observed for the respective SAXS model envelopes.

It is interesting to note that, comparing the CORAL models of the three mutants to that of the WT one, the S6 domain appears to be rotated to different extents around the Ig-like domain. In the case of both V72I and D109N, the mutated residues (whose lateral chains are highlighted by orange sticks in Fig 5) are mapped on the surface of the Ig-like domain facing the S6 domain, and might induce a rearrangement in the mutual orientations of the two domains either directly or by an indirect effect mediated by the linker. Such an explanation cannot be invoked for the T190M mutation, which is located on the external surface of the S6 domain.

Pathological mutations alter the flexibility of the N-terminal region of α -DG in solution

The presence of disordered regions in the crystal structure of the WT protein, has been reported to indicate a certain degree of inter-domain flexibility that could account for the observed conformational variability in solution [14].

The occurrence of possibly increased inter-domain flexibility in the three pathological mutants with respect to the WT was investigated using two different approaches. An essentially qualitative approach, the so called normalized Kratky plot [42], allows to directly compare objects of different shape and size. In such a plot, folded compact globular proteins provide a bell-shaped curve at low angles with a maximum at s^*R_g 1.75 [42]. Deviations from this behavior point to particle flexibility as in the case of the WT protein, whose maximum falls at s^*R_g 2.1 (Fig 6). The plots for D109N and T190M both show a broadening of the bell-shaped curve and a shift of the maxima to larger s^*R_g values, expected for more extended and flexible particles, while for V72I the curve is slightly sharper. In any case, the plots of the three mutant proteins are characterized by upward trends at higher s^*R_g values (i.e. higher scattering

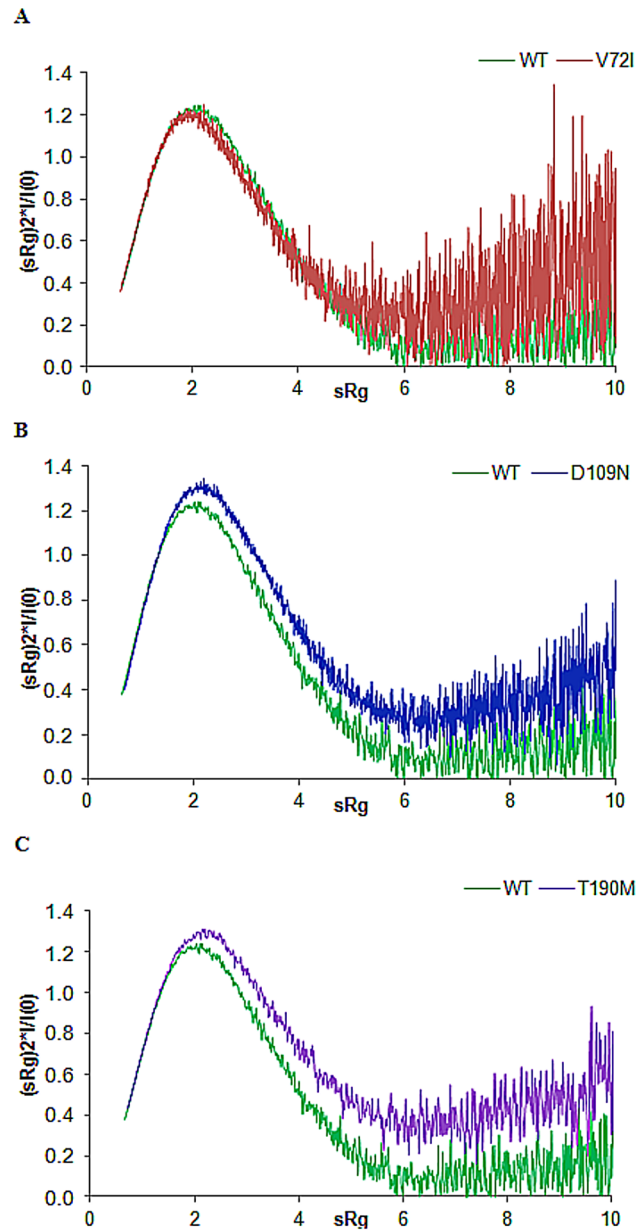


Fig 6. Normalized Kratky plots. Comparison of Kratky plots for the pathological mutants A) V72I, B) D109N and C) T190M with that of WT α -DG-Nt.

<https://doi.org/10.1371/journal.pone.0186110.g006>

angles) as compared to the more downward trend observed for the WT, an indication of increased flexibility in these former.

In the second, more quantitative approach, the ensemble optimization method (EOM) [38] was employed in order to analyze inter-domain flexibility and size distribution of possible multiple configurations in solution and to obtain typical optimized ensembles with a good fitting to the experimental scattering data (fit_{EOM} in Fig 4 and X_{EOM} S1 Table).

The EOM analysis for the three mutants are compared to WT in Fig 7 as a size distribution, plotting the R_g of the structures forming the initial random pool and the selected ensembles. The R_g distributions of these ensembles (Fig 7 solid lines) are nearly as broad as the distribution

of randomly generated models (Fig 7 dashed lines) supporting the hypothesis of considerable inter-domain flexibility. Moreover the quantification of the flexibility of the WT protein (ensemble $R_{flex} = 83.9\%$ versus pool $R_{flex} = 88.5\%$) and of the three mutants V72I (ensemble $R_{flex} = 81.9\%$ versus pool $R_{flex} = 87.7\%$), D109N (ensemble $R_{flex} = 85.8\%$ versus pool $R_{flex} = 86.6\%$) and T190M (ensemble $R_{flex} = 82.7\%$ versus pool $R_{flex} = 88.8\%$) confirmed random motion of the Ig-like domain with respect to the S6 domain. The quality of the ensemble solutions were further validated by the values of R_o that were close to 1.

However, it is interesting to note that the three mutants display alterations in the bimodal profile that characterizes the distribution of size observed for the WT protein (Fig 7 green solid lines). In detail for V72I the profile is mainly unimodal characterized by a broader single peak: even if its maximum is slightly shifted to more compact models with R_g about 21–25 Å (60.4% of the total population), there is an increased fraction of extended models with intermediate R_g around 25–30 Å. The distribution for D109N and T190M is still bimodal like the WT but beside the predominant fraction of relatively compact models, whose peak is sharper for D109N with R_g about 21–24 Å (43.1%), while slightly shifted to less compact models for T190M, with R_g about 21–27 Å (61.4%), both mutants show a significant increase in the population of elongated models with intermediate R_g around 25–30 Å. Indeed, although the small fraction of more extended models with R_g about 30–37 Å present in the WT (16.5%) can be still observed in D109N (15.3%) and is even increased in T190M (27.5%), but is scarcely represented in V72I (8.1%), all three mutants share a significant fraction with intermediate R_g around 25–30 Å (25.8%, 29.2% and 22.5% for V72I, D109N and T190M, respectively) that is remarkably less populated in the distribution of the WT protein (18.6%).

Functional implications of the pathological mutations of α -DG

The present analysis of the conformational variability of the mutants of α -DG-Nt in solution at low resolution pinpoints to inter-domain flexibility as an important structural determinant mediating the pathological alterations induced by the missense point mutations, that in human result into primary dystroglycanopathies [12,13,43].

To this respect it is worth noting that in the presence of a single wild-type copy of the DG allele no haploinsufficiency, or in any case no phenotype involving DG has been reported in the parents and/ or family of the compound heterozygous patient studied by Dong and colleagues. This finding suggests there are no relevant dominant effects in this case [13] nor in the heterozygous relatives carrying the T192M mutation [12]. In accordance with the heterozygous compound human phenotype, apparently they have never been co-selected [44]. No single nucleotide polymorphisms have been reported in the human DG gene within the region corresponding to its N-terminal Ig-like domain [45,46]. However, it is interesting to note that the sequence alignments for representative members of vertebrate α -DG regions containing these three mutations (S3A Table for Val72 and Asp109 according to the murine Ig-like domain of α -DG and S3B Table for Thr 190) reveal that Val72 and Thr190 are highly conserved across evolution, at least in higher animal groups. On the contrary, Asp109 has been apparently subjected to a less stringent negative selection pressure. Indeed Reptiles and Amphibians show a conservative substitution (Asp→Glu in green in S3A Table) in position 109 corresponding to the murine α -DG; in Fish beside the same conservative substitutions, also not conservative mutations can be observed (in red in S3A Table); in Birds the same pathological mutation Asp→Asn seems to be quite common (in yellow in S3A Table). However, in chicken [47] no missense mutation hitting the α -DG-Nt has been described to cause dystroglycanopathies. Among Fish, in zebrafish, where DG knock-down has been reported to cause a muscular dystrophy [48], all these three positions are conserved, even if a missense mutation (V567D) has been reported only within the second C-

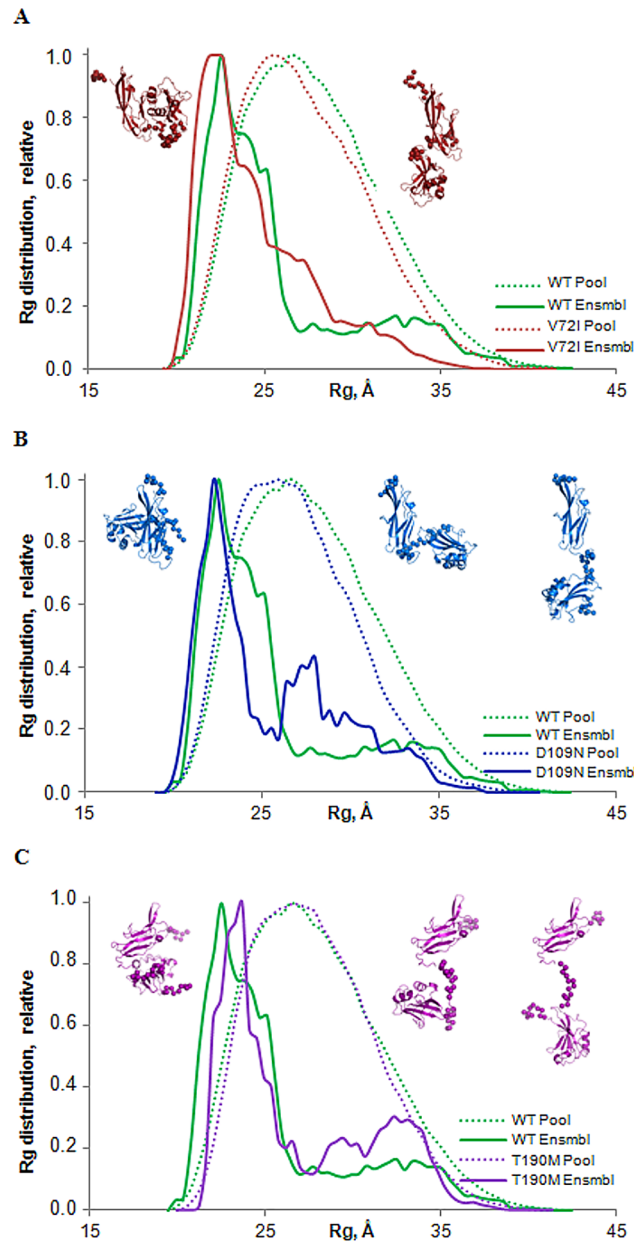


Fig 7. R_g distributions for the EOM models. Comparison of R_g distributions of the pathological mutants A) V72I, B) D109N and C) T190M with that of WT α -DG-Nt. The distributions for the initial random pools of models are shown by dot lines and the solid lines correspond to the selected ensembles. The representative conformations are shown near the distributions: compact on the left, intermediate in the middle and extended conformation on the right. Folded domains are depicted as cartoons (V72I in red, to D109N in blue and to T190M in violet); linkers and the N-terminal and C-terminal reconstructed regions are represented by spheres.

<https://doi.org/10.1371/journal.pone.0186110.g007>

terminal Ig-like domain of α -DG, completely abolishing its presence [49]. Most importantly these three positions are highly conserved in Mammals and in particular in dog and cat, where muscular dystrophies have been reported in which α -DG glycosylation [50] or expression levels [51] are respectively altered. Finally in humans these three α -DG pathological mutations are known to lead to hypoglycosylated α -DG and thus to primary dystroglycanopathies of different severity [12,13,43]. To bind its physiological ligands in ECM (laminin, agrin, perlecan) mature α -DG

requires a correct glycosylation that depends on a complex pattern of post-translational modifications. Indeed, during maturation a cascade of enzymatic reactions leads to matriglycan [7], recently proposed name for the glycan section of α -DG that actually forms a bridge with the laminin LG domains. The impact of these pathological mutations on α -DG-Nt, which is known to assist the bifunctional glycosyltransferase LARGE during the matriglycan extension stage [7,11] might depend on their ability to negatively affect the interaction of α -DG with LARGE and/ or make α -DG-Nt unable to properly assist LARGE in its enzymatic activity. From a structural point of view the crystallographic structures here discussed point to local differences with respect to the WT structure. It cannot be excluded that the local alterations observed in D109N and V72I may negatively affect the molecular determinants of LARGE recognition and/ or binding by α -DG. Such a hypothesis clearly requires a more detailed structural study in order to be correctly addressed.

Despite showing similar crystal structures, the three pathological mutants here discussed display a strikingly different behavior in solution, not only from each other but also from the WT protein. It has already been observed that the crystallographic models do not fully account for α -DG-Nt structure in solution nor for its conformational variability [14]. The framework resulting from a previous SAXS study suggests that multiple conformations and their associated populations observed in solution are shared features of both human and mouse α -DG [14], pointing to functional implications for the α -DG structural plasticity.

Indeed, LARGE action requires different catalytic steps involving two distinct glycosyltransferases domains that in turn extend matriglycan [7,11,52]. LARGE activity leads to the synthesis of a glycan of remarkably high molecular weight, due to the decoration of the mucin-like region of α -DG with several [-glucuronic acid- β 1,3-xylose- α 1,3-] units, that in mature α -DG spans more than 100 nm in length [8]. It is safe to assume that α -DG-Nt would need a certain degree of conformational variability in order to assist LARGE into such a complex matriglycan elongation. The SAXS study here presented clearly indicate that the pathological point mutations studied, while not largely affecting the α -DG-Nt structure in crystal, do affect its conformation in solution. The effects on the average conformations of the three pathological α -DG mutants T192M, V72I and D109N clearly emerged from shape reconstruction and from rigid-body fitting. The impact of the pathological mutations on the α -DG-Nt conformation are further confirmed by the analysis of the conformational ensembles variability. EOM analysis reveals that the mutations do affect the bimodal partition observed in human and mouse α -DG-Nt, with a general increase in the mutants inter-domain flexibility.

It is worth noting that none of the mutations under analysis seem to affect significantly the overall stability of the DG core protein as far as its expression, trafficking and membrane targeting is concerned [12,13]. Interestingly, the increased conformational flexibility of the mutant proteins seems to be related to a gain in their thermal stability (with the exception of V72I), quantified by DSF. This result is quite unexpected as thermal stability of a protein is commonly associated to its structural rigidity. Nevertheless, an increasing number of studies arises the need to reconsider such a notion [53]. For example, comparing thermodynamic parameters of the unfolding process of thermostable and mesostable proteins, it has been found that thermostability may be attributed to reduced entropy changes between unfolded and folded states, due to an enhanced number of conformations that can be sampled by the folded protein, especially at higher temperatures. In other words, an increased conformational mobility of the folded state reduces the entropic contribution to the ΔG of the folded \rightarrow unfolded transition, shifting melting temperature T_m towards higher values [54,55]. This may explain the increased T_m values measured for D109N and T190M, and the reduced T_m value of V72I, with respect to WT. Indeed, V72I populates conformational states that are more densely packed than those sampled by WT and its average conformation is more compact, as

also supported by its lower maximum dimension (D_{\max}) and radius of gyration (R_g) compared to WT. Moreover, the relative increased resistance to proteolysis, displayed by the mutant proteins, despite their enhanced protein flexibility, may be explained assuming that the exchange among compact and extended protein conformations occurs at a frequency that is higher with respect to the time scale of the proteolysis. Conversely, the increased conformational flexibility of the mutants probably perturbs the delicate equilibrium between structural rigidity and flexibility that must be assured to properly bind other proteins; this may account for the presence of some relevant post-translational effect of the mutations, i.e. a reduced affinity toward LARGE, resulting in a significant hypoglycosylation of the α -DG with loss of its functionality. Interestingly, we have recently reported that the mutation T190M reduces the mobility of α -DG within the membrane, as well as its clustering within the actin-rich domain, perturbing cell migration [56].

Although the mutants show remarkable differences in their conformations in solution, it is not easy to identify a correlation with different phenotypes featuring the related diseases. Indeed, the T192M mutation is associated to a severe form of limb-girdle muscular dystrophy with cognitive impairment, whereas the mutations V74I and D111N are related to a mild form of muscular dystrophy with asymptomatic hyperCKemia. Remarkably, the two diseases displayed apparently similar hypoglycosylated forms of α -DG. It can be hypothesized that such different phenotypes could be due to a different residual glycosylation of α -DG resulting in a differently reduced binding affinity towards its physiological partners that cannot be evaluated with the available analytical techniques. This is consistent with previous reports that pointed out the lack of correlation between the glycosylation levels of α -DG and clinical severity of the pathology [57]. A deeper insight into the biochemical mechanism underlining primary dystroglycanopathies could be achieved accumulating structural information on LARGE and on the way it interacts with α -DG-Nt in order to correctly glycosylate the α -DG mucin-like domain, a goal that we will hopefully reach in the next future.

Supporting information

S1 Fig. Thermal denaturation assay using DSF with SYPRO dye. Comparison of averaged unfolding fluorescent curves with the respective error bars for WT (A), V72I (B), D109N (C) and T190M (D). The fitting Boltzmann function limited to the temperature region of interest (first transition) is shown.

(TIF)

S2 Fig. Guinier plots. Guinier plot representation of the SAXS data. The straight lines are the fitted data according to Guinier approximation to determine the radius of gyration and the scattering amplitude. A) V72I, B) D109N and C) T190M.

(TIF)

S3 Fig. Electrostatic potential maps of V72I (A) and WT (B). The electrostatic potential (in k_bT/e_c units) is mapped on solvent-accessible surface of the WT and V72I pathological mutant accessible surfaces. Negative potential is colored in blue, positive potential in red. Color scale varies between -2 and +2. Molecular models (stretch 105–116) are represented as ribbons with selected residues depicted as stick-and-ball.

(TIF)

S4 Fig. Experimental SAXS curves measured at different concentration: A) V72I, B) D109N and C) T190M. The zoomed regions of these graphs at low angles are presented on the right column.

(TIF)

S5 Fig. Comparison of the structural models of the WT and the three pathological mutants of α -DG-Nt. Typical CORAL models (green, red, blue and violet cartoon representations for the folded Ig-like and S6 domains, spheres for the restored missing fragments of WT (A), V72I (B), D109N (C) and T190M (D), respectively) overlaid to the respective CORAL models, where the domains have been fixed in the positions found in the crystal structures in yellow. (TIF)

S6 Fig. Distribution of the centers of masses of S6 domains: In the crystallographic models of the WT and the three mutants (represented by spheres) and in the respective CORAL models (represented by dots) respect to the Ig-like domain and its center of mass (represented by mesh).

(TIF)

S1 Table. Overall parameters calculated from SAXS experiments. SAXS data collection and processing information. Parameters derived from SAXS analysis are also reported.

(DOCX)

S2 Table. Additional SAXS structural parameters: Radius of gyration (R_g), maximum dimension (D_{max}), Porod volume (A^3), and MW (Da). D_{max} was obtained from the $p(r)$ distribution using GNOM; $I(0)$ (scattering intensity) was obtained from the scattering data by the Guinier analysis. Molecular mass (Mr) was estimated from comparison with $I(0)$ intensity of the standard BSA sample.

(DOCX)

S3 Table. Sequence alignments. Inter-specific alignments in the regions spanning the amino acid positions: A) 65–110 and B) 187–221, referring to the murine Ig-like domain of α -DG.

Accession codes for dystroglycan sequences: *Homo sapiens* (q14118), *Macaca mulatta* (f6ru72), *Pan troglodytes* (h2qml8), *Gorilla gorilla gorilla* (g3r897), *Callitrix jacchus* (f6trx4), *Otolemur garnettii* (h0xc26), *Papio anubis* (a0a096mt62), *Chlorocebus sabaues* (a0a0d9rr47), *Nomascus leucogenys* (g1r572), *Mus musculus* (q62165), *Rattus norvegicus* (f1m8k0), *Heterocephalus glaber* (a0a0p6jge5), *Oryctolagus cuniculus* (q28685), *Ictidomys tridecemlineatus* (i3n6c8), *Dipodomys ordii* (a0a1s3g6k3), *Felis catus* (b4xem8), *Canis lupus familiaris* (q9tsz6), *Sus scrofa* (q29243), *Bos taurus* (o18738), *Ovis aries* (w5pvz9), *Equus caballus* (f6x9u4), *Erinaceus europaeus* (a0a1s3wgq5), *Ailuropoda melanoleuca* (d2heu8), *Mustela putorius furo* (m3yqu5), *Loxodonta africana* (g3t6q9), *Sarcophilus harrisii* (g3wvr6), *Monodelphis domestica* (f7dbt8), *Gallus gallus* (a4var9), *Taeniopygia guttata* (h0z430), *Ficedula albicollis* (u3k2z0), *Anas platyrhynchos* (u3ibq2), *Meleagris gallopavo* (g1mw2), *Pelodiscus sinensis* (k7gfa6), *Anolis carolinensis* (g1kgb5), *Xenopus tropicalis* (f7ei21), *Xenopus laevis* (q7zx16), *Danio rerio* (q8jhu7), *Latimeria chalumnae* (h3b2q1), *Oreochromis niloticus* (i3kqf9), *Astyanax mexicanus* (w5l657), *Ictalurus punctatus* (w5ua28), *Aphyosemion striatum* (a0a1a7y3r0), *Nothobranchius furzeri* (a0a1a7zin9), *Nothobranchius rachovii* (a0a1a8plv8), *Nothobranchius kadleci* (a0a1a8dcc7), *Nothobranchius kuhntae* (a0a1a8jug2), *Nothobranchius pienaarri* (a0a1a8ms21), *Nothobranchius kortausae* (a0a1a8hgb2), *Poeciliopsis prolifica* (a0a0s7hfk7), *Salmon salar* (a0a1s3p849).

(DOCX)

Acknowledgments

We are grateful to Dr D. Lamba (Istituto di Cristallografia, CNR, Trieste, Italy) for the helpful discussions and for critically reading the manuscript.

We are grateful to the ELETTRA, to the EMBL Hamburg and to the ESRF staff for their help and assistance during the data collection.

Author Contributions

Conceptualization: Sonia Covaceuszach, Manuela Bozzi, Andrea Brancaccio, Alberto Cassetta.

Formal analysis: Sonia Covaceuszach, Manuela Bozzi, Maria Giulia Bigotti, Francesca Sciandra, Petr V. Konarev, Andrea Brancaccio, Alberto Cassetta.

Funding acquisition: Andrea Brancaccio.

Investigation: Sonia Covaceuszach, Manuela Bozzi, Alberto Cassetta.

Project administration: Andrea Brancaccio.

Visualization: Sonia Covaceuszach, Alberto Cassetta.

Writing – original draft: Sonia Covaceuszach, Manuela Bozzi, Andrea Brancaccio, Alberto Cassetta.

Writing – review & editing: Sonia Covaceuszach, Manuela Bozzi, Maria Giulia Bigotti, Francesca Sciandra, Petr V. Konarev, Andrea Brancaccio, Alberto Cassetta.

References

1. Ibraghimov-Beskrovnaya O, Ervasti JM, Leveille CJ, Slaughter CA, Sernett SW, Campbell KP. Primary structure of dystrophin-associated glycoproteins linking dystrophin to the extracellular matrix. *Nature* 1992; 355:696–702. <https://doi.org/10.1038/355696a0> PMID: 1741056
2. Barresi R, Campbell KP. Dystroglycan: from biosynthesis to pathogenesis of human disease. *J Cell Sci* 2006; 119:199–207. <https://doi.org/10.1242/jcs.02814> PMID: 16410545
3. Brancaccio A, Schulthess T, Gesemann M, Engel J. Electron microscopic evidence for a mucin-like region in chick muscle α -dystroglycan. *FEBS Lett*. 1995; 368:139–142. PMID: 7615068
4. Bozic D, Sciandra F, Lamba D, Brancaccio A. The structure of the N-terminal region of murine skeletal muscle α -dystroglycan discloses a modular architecture. *J Biol Chem*. 2004; 279:44812–44816. <https://doi.org/10.1074/jbc.C400353200> PMID: 15326183
5. De Rosa MC, Pirolli D, Bozzi M, Sciandra F, Giardina B, Brancaccio A. A second Ig-like domain identified in dystroglycan by molecular modelling and dynamics. *J Mol Graph Model*. 2011; 29:1015–1024. <https://doi.org/10.1016/j.jmgm.2011.04.008> PMID: 21605994
6. Bozzi M, Di Stasio E, Scaglione GL, Desiderio C, Martelli C, Giardina B, et al. Probing the stability of the “naked” mucin-like domain of human α -dystroglycan. *BMC Biochem*. 2013; 14:15. <https://doi.org/10.1186/1471-2091-14-15> PMID: 23815856
7. Yoshida-Moriguchi T, Campbell KP. Matriglycan: a novel polysaccharide that links dystroglycan to the basement membrane. *Glycobiology* 2015; 25:702–713. <https://doi.org/10.1093/glycob/cwv021> PMID: 25882296
8. Briggs DC, Yoshida-Moriguchi T, Zheng T, Venzke D, Anderson ME, Strazzulli A, et al. Structural basis of laminin binding to the LARGE glycans on dystroglycan. *Nat Chem Biol*. 2016; 12:810–814. <https://doi.org/10.1038/nchembio.2146> PMID: 27526028
9. Sciandra F, Bozzi M, Bigotti MG, Brancaccio A. The multiple affinities of α -dystroglycan. *Curr Protein Pept Sci*. 2013; 14:626–634. PMID: 24206164
10. Taniguchi-Ikeda M, Morioka I, Iijima K, Toda T. Mechanistic aspects of the formation of α -dystroglycan and therapeutic research for the treatment of α -dystroglycanopathy: A review. *Mol Aspects Med*. 2016; 51:115–124. <https://doi.org/10.1016/j.mam.2016.07.003> PMID: 27421908
11. Kanagawa M, Saito F, Kunz S, Yoshida-Moriguchi T, Barresi R, Kobayashi YM, et al. Molecular recognition by LARGE is essential for expression of functional dystroglycan. *Cell* 2004; 117:953–964. <https://doi.org/10.1016/j.cell.2004.06.003> PMID: 15210115
12. Hara Y, Balci-Hayta B, Yoshida-Moriguchi T, Kanagawa M, Beltrán-Valero de Bernabé D, Gündeşli H, et al. A dystroglycan mutation associated with limb-girdle muscular dystrophy. *N Engl J Med*. 2011; 364:939–946. <https://doi.org/10.1056/NEJMoa1006939> PMID: 21388311

13. Dong M, Noguchi S, Endo Y, Hayashi YK, Yoshida S, Nonaka I, et al. DAG1 mutations associated with asymptomatic hyperCKemia and hypoglycosylation of α -dystroglycan. *Neurology* 2015; 84:273–279. <https://doi.org/10.1212/WNL.0000000000001162> PMID: 25503980
14. Covaceuszach S, Bozzi M, Bigotti MG, Sciandra F, Konarev PV, Brancaccio A, Cassetta A. Structural flexibility of human α -dystroglycan. *FEBS Open Bio* 2017; 7:1064–1077. <https://doi.org/10.1002/2211-5463.12259> PMID: 28781947
15. Bozzi M, Cassetta A, Covaceuszach S, Bigotti MG, Bannister S, Hübner W, et al. The structure of the T190M mutant of murine α -dystroglycan at high resolution: insight into the molecular basis of a primary dystroglycanopathy. *PLoS ONE* 2015; 10:e0124277. <https://doi.org/10.1371/journal.pone.0124277> PMID: 25932631
16. Ericsson UB, Hallberg BM, Detitta GT, Dekker N, Nordlund P. Thermofluor-based high-throughput stability optimization of proteins for structural studies. *Anal Biochem.* 2006; 357:289–298. <https://doi.org/10.1016/j.ab.2006.07.027> PMID: 16962548
17. Schägger H, Jagow G von. Tricine-sodium dodecyl sulfate-polyacrylamide gel electrophoresis for the separation of proteins in the range from 1 to 100 kDa. *Anal Biochem.* 1987; 166:368–379. PMID: 2449095
18. Bernstorff S, Busetto E, Gramaccioni C, Lausi A, Olivi L, Zanini F, et al. The macromolecular crystallography beamline at ELETTRA. *Rev Sci Instrum.* 1995; 66:1661–1664.
19. Lausi A, Polentarutti M, Onesti S, Plaisier JR, Busetto E, Bais G, et al. Status of the crystallography beamlines at Elettra. *Eur Phys J Plus* 2015; 130:1–8.
20. Kabsch W. XDS. *Acta Crystallogr D Biol Crystallogr.* 2010; 66:125–132. <https://doi.org/10.1107/S0907444909047337> PMID: 20124692
21. McCoy AJ, Grosse-Kunstleve RW, Adams PD, Winn MD, Storoni LC, Read RJ. Phaser crystallographic software. *J Appl Crystallogr.* 2007; 40:658–674. <https://doi.org/10.1107/S0021889807021206> PMID: 19461840
22. Adams PD, Afonine PV, Bunkóczi G, Chen VB, Davis IW, Echols N, et al. *PHENIX*: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr D Biol Crystallogr.* 2010; 66:213–221. <https://doi.org/10.1107/S0907444909052925> PMID: 20124702
23. Emsley P, Lohkamp B, Scott WG, Cowtan K. Features and development of Coot. *Acta Crystallogr D Biol Crystallogr.* 2010; 66:486–501. <https://doi.org/10.1107/S0907444910007493> PMID: 20383002
24. Afonine PV, Grosse-Kunstleve RW, Echols N, Headd JJ, Moriarty NW, Mustyakimov M, et al. Towards automated crystallographic structure refinement with phenix.refine. *Acta Crystallogr D Biol Crystallogr.* 2012; 68:352–367. <https://doi.org/10.1107/S0907444912001308> PMID: 22505256
25. Chen VB, Arendall WB, Headd JJ, Keedy DA, Immormino RM, Kapral GJ, et al. *MolProbity*: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr D Biol Crystallogr.* 2010; 66:12–21. <https://doi.org/10.1107/S0907444909042073> PMID: 20057044
26. Collaborative Computational Project N° 4. The CCP4 suite: programs for protein crystallography. *Acta Crystallogr D Biol Crystallogr.* 1994; 50:760–3. <https://doi.org/10.1107/S0907444994003112> PMID: 15299374
27. Schrödinger LLC. The PyMOL Molecular Graphics System, Version 1.3r1. 2010
28. Pernot P, Round A, Barrett R, De Maria Antolinis A, Gobbo A, Gordon E, et al. Upgraded ESRF BM29 beamline for SAXS on macromolecules in solution. *J Synchrotron Radiat.* 2013; 20:660–664. <https://doi.org/10.1107/S0909049513010431> PMID: 23765312
29. Blanchet CE, Spilotros A, Schwemmer F, Graewert MA, Kikhney A, Jeffries CM, et al. Versatile sample environments and automation for biological solution X-ray scattering experiments at the P12 beamline (PETRA III, DESY). *J Appl Crystallogr.* 2015; 48:431–443. <https://doi.org/10.1107/S160057671500254X> PMID: 25844078
30. Konarev PV, Volkov VV, Sokolova AV, Koch MHJ, Svergun DI. *PRIMUS*: a Windows PC-based system for small-angle scattering data analysis. *J Appl Crystallogr.* 2003; 36:1277–1282.
31. Petoukhov MV, Franke D, Shkumatov AV, Tria G, Kikhney AG, Gajda M, et al. New developments in the ATSAS program package for small-angle scattering data analysis. *J Appl Crystallogr.* 2012; 45:342–350. <https://doi.org/10.1107/S0021889812007662> PMID: 25484842
32. Svergun DI. Determination of the regularization parameter in indirect-transform methods using perceptual criteria. *J Appl Crystallogr.* 1992; 25:495–503.
33. Porod G. *Small Angle X-Ray Scattering*, Academic Press; 1982. pp. 17–51.
34. Svergun DI. Restoring low resolution structure of biological macromolecules from solution scattering using simulated annealing. *Biophys J.* 1999; 76:2879–2886. [https://doi.org/10.1016/S0006-3495\(99\)77443-6](https://doi.org/10.1016/S0006-3495(99)77443-6) PMID: 10354416

35. Kozin MB, Svergun DI. Automated matching of high- and low-resolution structural models. *J Appl Crystallogr.* 2001; 34:33–41.
36. Volkov VV, Svergun DI. Uniqueness of *ab initio* shape determination in small-angle scattering. *J Appl Crystallogr.* 2003; 36:860–864.
37. Svergun D, Barberato C, Koch MHJ. CRY SOL—a program to evaluate X-ray solution scattering of biological macromolecules from atomic coordinates. *J Appl Crystallogr.* 1995; 28:768–773.
38. Bernadó P, Mylonas E, Petoukhov MV, Blackledge M, Svergun DI. Structural characterization of flexible proteins using small-angle X-ray scattering. *J Am Chem Soc.* 2007; 129:5656–5664. <https://doi.org/10.1021/ja069124n> PMID: 17411046
39. Tria G, Mertens HDT, Kachala M, Svergun DI. Advanced ensemble modelling of flexible macromolecules using X-ray solution scattering. *IUCrJ* 2015; 2:207–217. <https://doi.org/10.1107/S205225251500202X> PMID: 25866658
40. Thornton JM, Edwards MS, Taylor WR, Barlow DJ. Location of “continuous” antigenic determinants in the protruding regions of proteins. *EMBO J.* 1986; 5:409–413. PMID: 2423325
41. Phillips K, Peña AH de la. The combined use of the ThermoFluor assay and ThermoQ analytical software for the determination of protein stability and buffer optimization as an aid in protein crystallization. *Curr Protoc Mol Biol Ed Frederick M Ausubel Al*; 2011. pp.10.28.1–10.28.28.
42. Durand D, Vivès C, Cannella D, Pérez J, Pebay-Peyroula E, Vachette P, et al. NADPH oxidase activator p67phox behaves in solution as a multidomain protein with semi-flexible linkers. *J Struct Biol.* 2010; 169:45–53. <https://doi.org/10.1016/j.jsb.2009.08.009> PMID: 19723583
43. Kanagawa M, Toda T. α -Dystroglycanopathy. In: Takeda S, Miyagoe-Suzuki Y, Mori-Yoshimura M, editors. *Translational Research in Muscular Dystrophy*, Springer Japan; 2016. pp. 21–38.
44. Adams JC, Brancaccio A. The evolution of the dystroglycan complex, a major mediator of muscle integrity. *Biol Open* 2015; 4:1163–1179. <https://doi.org/10.1242/bio.012468> PMID: 26319583
45. Gottlieb E, Ciccone C, Darvish D, Naiem-Cohen S, Dalakas MC, Savelkoul PJ, et al. Single nucleotide polymorphisms in the dystroglycan gene do not correlate with disease severity in hereditary inclusion body myopathy. *Mol Genet Metab.* 2005; 86:244–249. <https://doi.org/10.1016/j.ymgme.2005.07.003> PMID: 16112887
46. Concolino P, Capoluongo E, Santonocito C, Vento G, Tana M, Romagnoli C, et al. Genetic analysis of the dystroglycan gene in bronchopulmonary dysplasia affected premature newborns. *Clin Chim Acta* 2007; 378:164–167. <https://doi.org/10.1016/j.cca.2006.11.016> PMID: 17196572
47. Saito F, Blank M, Schröder J, Manyá H, Shimizu T, Campbell KP, et al. Aberrant glycosylation of α -dystroglycan causes defective binding of laminin in the muscle of chicken muscular dystrophy. *FEBS Lett.* 2005; 579:2359–2363. <https://doi.org/10.1016/j.febslet.2005.03.033> PMID: 15848172
48. Parsons MJ, Campos I, Hirst EMA, Stemple DL. Removal of dystroglycan causes severe muscular dystrophy in zebrafish embryos. *Development* 2002; 129:3505–3512. PMID: 12091319
49. Gupta V, Kawahara G, Gundry SR, Chen AT, Lencer WI, Zhou Y, et al. The zebrafish *dag1* mutant: a novel genetic model for dystroglycanopathies. *Hum Mol Genet.* 2011; 20:1712–1725. <https://doi.org/10.1093/hmg/ddr047> PMID: 21296866
50. Ervasti JM, Roberds SL, Anderson RD, Sharp NJH, Kornegay JN, Campbell KP. α -Dystroglycan deficiency correlates with elevated serum creatine kinase and decreased muscle contraction tension in golden retriever muscular dystrophy. *FEBS Lett.* 1994; 350:173–176. PMID: 8070559
51. Martin PT, Shelton GD, Dickinson PJ, Sturges BK, Xu R, LeCouteur RA, et al. Muscular dystrophy associated with α -dystroglycan deficiency in Sphynx and Devon Rex cats. *Neuromuscul Disord.* 2008; 18:942–952. <https://doi.org/10.1016/j.nmd.2008.08.002> PMID: 18990577
52. Inamori K, Yoshida-Moriguchi T, Hara Y, Anderson ME, Yu L, Campbell KP. Dystroglycan function requires xylosyl- and glucuronyltransferase activities of LARGE. *Science* 2012; 335:93–96. <https://doi.org/10.1126/science.1214115> PMID: 22223806
53. Karshikoff A, Nilsson L, Ladenstein R. Rigidity versus flexibility: the dilemma of understanding protein thermal stability. *FEBS J.* 2015; 282:3899–3917. <https://doi.org/10.1111/febs.13343> PMID: 26074325
54. Sawle L, Ghosh K. How Do thermophilic proteins and proteomes withstand high temperature? *Biophys J.* 2011; 101:217–227. <https://doi.org/10.1016/j.bpj.2011.05.059> PMID: 21723832
55. Stadler AM, Garvey CJ, Bocahut A, Sacquin-Mora S, Digel I, Schneider GJ, et al. Thermal fluctuations of haemoglobin from different species: adaptation to temperature via conformational dynamics. *J. R. Soc. Interface* 2012; 9:2845–2855. <https://doi.org/10.1098/rsif.2012.0364> PMID: 22696485
56. Palmieri V, Bozzi M, Signorino G, Papi M, De Spirito M, Brancaccio A, et al. α -Dystroglycan hypoglycosylation affects cell migration by influencing β -dystroglycan membrane clustering and filopodia length: a multiscale confocal microscopy analysis. *Biochim Biophys Acta.* 2017; <https://doi.org/10.1016/j.bbdis.2017.05.025> PMID: 28572004

57. Jimenez-Mallebrera C, Torelli S, Feng L, Kim J, Godfrey C, Clement E, et al. A comparative study of α -dystroglycan glycosylation in dystroglycanopathies suggests that the hypoglycosylation of α -dystroglycan does not consistently correlate with clinical severity. *Brain Pathol.* 2009; 19:596–611. <https://doi.org/10.1111/j.1750-3639.2008.00198.x> PMID: 18691338