

The Structure of the T190M Mutant of Murine  $\alpha$ -Dystroglycan at High Resolution:  
Insight into the Molecular Basis of a Primary Dystroglycanopathy\*

**Manuela Bozzi<sup>1±</sup>, Alberto Cassetta<sup>2±#</sup>, Sonia Covaceuszach<sup>2</sup>, Maria Giulia Bigotti<sup>3</sup>, Saskia Bannister<sup>4</sup>,  
Wolfgang Hübner<sup>4</sup>, Francesca Sciandra<sup>5</sup>, Dorian Lamba<sup>2</sup> and Andrea Brancaccio<sup>3,5</sup>**

<sup>1</sup>Istituto di Biochimica e Biochimica Clinica, Università Cattolica del Sacro Cuore, Roma 00168

<sup>2</sup>Istituto di Cristallografia – CNR, Trieste Outstation, Trieste 34149

<sup>3</sup>School of Biochemistry, University of Bristol, Bristol BS8 1TD

<sup>4</sup>Department of Physics, University of Bielefeld, 33615 Bielefeld, Germany

<sup>5</sup>Istituto di Chimica del Riconoscimento Molecolare - CNR c/o Università Cattolica del Sacro Cuore, Roma 00168

<sup>±</sup>Both authors contributed equally to this work.

\*Running Title: *The structure of the  $\alpha$ -DG T190M mutant N-terminal domain*

To whom correspondence should be addressed: Andrea Brancaccio, Istituto di Chimica del Riconoscimento Molecolare (ICRM) - CNR c/o Università Cattolica del Sacro Cuore, Largo Francesco Vito 1, 00168 Roma, Italy, Tel: ++39 (06) 301-55135; Fax: ++39 (06) 305-3598, E-mail: [andrea.brancaccio@icrm.cnr.it](mailto:andrea.brancaccio@icrm.cnr.it) & Alberto Cassetta, Istituto di Cristallografia, CNR, Basovizza, SS14, km 163,5 Trieste 34012 Trieste, Italy, Tel: ++39 (040) 375-7520; Fax: ++39 (040) 922-1126, E-mail: [alberto.cassetta@ts.ic.cnr.it](mailto:alberto.cassetta@ts.ic.cnr.it)

**Keywords:**  $\alpha$ -Dystroglycan; muscular dystrophy; primary dystroglycanopathies; protein domain; protein denaturation; protein X-ray crystal structure; site-directed mutagenesis

**Background:** Mutations of  $\alpha$ -dystroglycan cause neuromuscular disorders known as primary dystroglycanopathies.

**Results:** The high-resolution crystal structure of the T190M mutant of the N-terminal domain of  $\alpha$ -dystroglycan has been determined suggesting a decreased structural flexibility.

**Conclusion:** The lower flexibility may account for the reduced glycosylation observed in T190M  $\alpha$ -dystroglycan.

**Significance:** The novel structural details suggest a molecular mechanism behind a severe form of muscular dystrophy.

## ABSTRACT

The severe dystroglycanopathy known as limb girdle disease is an autosomal recessive muscular dystrophy caused by the point mutation T192M in  $\alpha$ -dystroglycan. Functional expression analysis *in vitro* and *in vivo* indicated that the mutation was responsible for a decrease in posttranslational glycosylation of dystroglycan, eventually interfering with its extracellular-matrix receptor function and laminin binding in skeletal muscle and brain. The X-ray crystal structure of the missense variant T190M of the murine N-terminal domain of  $\alpha$ -dystroglycan (50-313) has been determined and showed an overall topology (Ig-like domain followed by a basket-shaped

domain reminiscent of the small subunit ribosomal protein S6) almost identical to that of the wild-type structure. An extended and highly flexible loop (encompassing residues 159-180) that could not be solved in the wild-type N-terminal domain could, in part, be traced in the T190M variant. The crystallographic analysis prompts to a solvent shells reorganization around Met190 that affects the interaction between the B1-B5 anti-parallel strands forming part of the floor of the basket-shaped domain, with likely repercussions on the folding stability of the protein domain(s) and on the overall molecular flexibility. Chemical denaturation and limited proteolysis experiments revealed a slightly decreased stability of the T190M variant with respect to its wild-type counterpart. This mutation may render the entire L-shaped protein architecture less flexible. The overall reduced flexibility and stability may affect the functional properties of  $\alpha$ -dystroglycan *via* negatively influencing its binding behavior to factors needed for dystroglycan maturation, and may lay the molecular basis of the T190M-driven primary dystroglycanopathy.

---

The dystroglycan (DG) complex is composed of two subunits,  $\alpha$  and  $\beta$ , that are formed from a unique precursor which undergoes a proteolytic event within the endoplasmic reticulum during the first steps of its maturation pathway (1,2). DG represents a widely expressed adhesion complex that plays a crucial role in offering stability to tissues being at the crossroad between cytoskeleton, plasma membrane and the surrounding extracellular matrix (3,4). The DG main role is to bind with high-affinity laminins that are central organizers of the molecular network behind specialized basement membranes surrounding skeletal muscle. Besides, DG may form a plethora of additional interactions with other binding partners sharing laminin globular (LG) domains within epithelia, endothelia and Schwann cells. Indeed, its binding affinity may largely vary based on the degree of glycosylation of its  $\alpha$  subunit (5).

DG has been particularly studied within skeletal muscle, a tissue where its  $\alpha$ -subunit is found to be highly glycosylated, since a number of neuromuscular diseases involving glycosyltransferases active in muscle have been identified during the last years (6). The commonly accepted molecular scenario implies that hypoglycosylated DG cannot bind the skeletal muscle laminin-2 (7) with the usual high affinity, which would lead to the reduction of sarcolemma stability at the basis of a wide number of severe, or later onset, muscular dystrophies, defined as secondary dystroglycanopathies (6).

The  $\alpha$ -subunit plays multiple important roles in DG, in that i) it is crucially involved in the maturation and posttranslational modification of the precursor; ii) it is highly modified with sugar chains especially in its central region (i.e. the mucin-like portion) from where some essential sugar moieties protrudes to bind laminin. The first is based on the noncovalent binding activity of  $\alpha$ -DG toward  $\beta$ -DG (8) while for the latter, it seems that the N-terminal domain may be central in recognizing and directing the activity of important modifying enzymes, such as the like-acetylglucosaminyltransferase (LARGE) that adds the repeating disaccharide block [- $\alpha$ 3-glucuronic acid (GlcA)- $\beta$ 3-xylose (Xyl)-] required for laminin-binding (9) to a phosphorylated mannose of  $\alpha$ -DG (10).

We have pioneered studies on  $\alpha$ -DG at the molecular level (11) having extensively characterized the structure of the N-terminal domain of  $\alpha$ -DG which is shaped in two subdomains (12). The first is a typical Ig-like domain, while the second one is similar to the S6 protein present in the small ribosomal subunit of *T. thermophilus* (13).

Although a plethora of interesting studies are currently unraveling the structure of the sugar blocks modifying the mucin-like domain of  $\alpha$ -DG as well as the exact residues of  $\alpha$ -DG involved in these modifications (14-17), no ultimate structural information is currently available on the dystroglycan/laminin complex at the molecular level. The challenges in obtaining crystals of the complex mainly arise from the highly heterogeneous nature of  $\alpha$ -DG glycosylation,

confirmed by the fact that the protein can be typically visualized in overlay assays or Western blots as a band with a broad smear appearance indicating a distribution of different molecular masses (1,3,4). In addition, the molecular mechanism behind the first case of primary dystroglycanopathy caused by the first disease-causing missense mutation found in a Turkish family in homozygous fashion (18) has yet to be clarified. Interestingly, the human T192M mutation (and its topological counterpart in mouse, T190M), lies within the strand B1 that together with the strands B2, B3, B4 and B5, arranged in an anti-parallel topology, defines the floor of the basket shaped S6-like domain (13). Theoretically, the Thr to Met semiconservative mutation is unlikely to cause *per se* a high degree of disorder/instability, and it should not represent a site of O-glycosylation.

Nonetheless, we believe that the structural analysis of the T190M mutant of  $\alpha$ -dystroglycan is a fundamental step towards the understanding at the molecular level of the mechanism leading to muscular dystrophy. Following this basic idea, we expressed the murine mutant in our recombinant expression system in *E.coli* and purified it for crystallization studies aimed at possibly solving its structure and comparing it with that of the wild-type protein previously reported (13).

## EXPERIMENTAL PROCEDURES

**Mutagenesis and Primers-** The murine  $\alpha$ -DG(50-313)R166H (hereinafter WT) DNA was cloned into a bacterial vector, pHis-Trx, for the expression of the protein as thioredoxin fusion product, also containing an N-terminal 6xHis tag and a thrombin cleavage site, as previously described (13). The point mutation T190M was introduced into the WT DNA construct using the QuikChange site-directed mutagenesis kit (Stratagene) and appropriate primers. Briefly, the mutation was inserted within the coding region of the WT N-terminal domain cloned in the vector pHis-Trx (13) for prokaryotic expression aimed at crystallographic analysis, and within the entire cDNA sequence of the murine DG cloned in the pEGFP-N1 plasmid for expression in eukaryotic cells (19), respectively, using the following

primers:

Forward: 5'-  
CCAGTGACTGTCCTTATGGTGATTCTGGAT  
GCT-3'

Reverse: 5'-  
AGCATCCAGAATCACCATAAGGACAGTCA  
CTGG-3'

The construct DG<sup>T190M</sup>-pEGFP-N1 allows to express dystroglycan with a Green Fluorescent Protein (GFP) fused at the C-terminus of  $\beta$ -DG. Moreover, a myc tag is present, inserted after Lys498, within the C-terminal domain of  $\alpha$ -DG (19). All constructs were verified by automated sequencing.

**Fusion protein expression and purification-** The recombinant  $\alpha$ -DG(50-313)R166H T190M (hereinafter T190M) fusion protein was expressed in *Escherichia coli* BL21(DE3) Codon Plus RIL strain and purified using nickel affinity chromatography. The fragment of interest was obtained upon thrombin cleavage. Further purification steps were carried out using anion exchange and gel filtration chromatography. Namely, after thrombin cleavage, the flow through of a HiTrap Chelating column (GE Healthcare) was applied onto a Hi-Trap Q HP column (GE Healthcare) pre-equilibrated with buffer A (25 mM Tris-HCl pH 7.5). T190M was eluted with a linear gradient of 0-0.5 M NaCl in buffer A. The fractions containing T190M were pooled, concentrated with Amicon Ultra 15 (Millipore) and loaded on Superdex 200 10/300 GL (GE Healthcare) pre-equilibrated with 25 mM Tris-HCl pH 7.5, 0.15M NaCl, at a flow rate of 0.4 mL/min: the core fractions of the peak were supplemented with 2.5% Glycerol and concentrated by Microcon GM10 (Millipore). The purity of the protein was confirmed by Tricine/SDS-PAGE (20).

The possibility that the methyl sulfide (thioether) of Met 190 would be oxidized in T190M, or in general that the recombinant peptide would harbour some additional undesired modifications, has been ruled out via mass spectrometry analysis exploiting an Orbitrap instrument (Thermo Scientific) (data not shown). The corresponding mass value obtained also confirmed the presence of a disulfide bridge

between Cys180 and Cys262.

*Crystallization, data collection, structure solution and refinement*- Crystals of T190M were grown by using the vapor diffusion hanging drop method, following the protocol used for WT (13). Drops were prepared by mixing 1  $\mu$ L of the protein solution (5.25 mg/mL in 25 mM Tris, 150 mM NaCl and 2.5 % glycerol; pH 7.5) with 1  $\mu$ L of the precipitant solution (0.8 M citrate buffer; pH 7.0) and equilibrated against the reservoir (1 mL) at 4  $^{\circ}$ C.

Data collection was carried out at the European Synchrotron Radiation Facility Grenoble (France) at the ID23-1 beamline, using a detector Pilatus-6M (Dectris) and the wavelength of 1.00  $\text{Å}$ . Data collection was carried out at -173  $^{\circ}$ C. Before being exposed to the X-ray beam, crystals were quickly dipped in cryoprotectant solution (25% ethylene glycol, 1.0 M citrate buffer, pH 7.0) and flash frozen in liquid nitrogen.

Indexing, integration and data reduction of the diffraction data were carried out by using a combination of XDS (21) and CCP4 (22) programs. Data reduction statistics are reported in Table I.

The structure solution of T190M was obtained by Patterson search methods with PHASER (23) using the crystal structure of WT as template (PDB id: 1U2C). Restrained refinement of the crystal structure was carried out by using *phenix.refine* (24). The improvement of the initial model underwent a protocol that included a rigid body fitting stage followed by simulated annealing, coordinates and individual B-factors refinement. The subsequent stages integrated a Translation-Libration-Screw (TLS) model parameterization before the individual B-factors refinement. Automatic refinement cycles were alternated with manual rebuilding sessions performed with COOT (25). Solvent molecules were identified by the automatic water-picking algorithm implemented in *phenix.refine* (24). The overall quality of these automatically picked solvent molecules were manually checked. Multiple conformations were introduced for selected side chains during the last cycles of refinement.

Protein stereochemistry was monitored

throughout the refinement process and during manual rebuilding with MolProbity (26). Statistics of the crystallographic refinement are reported in Table I.

Protein structures superposition was carried out using ProFit (Martin, A.C.R., <http://www.bioinf.org.uk/software/profit/>). The STRIDE web-server was used to assign the protein secondary structure (27). Interfaces between domains, solvent accessible areas and solvation energies were analyzed by using PISA as part of the CCP4 package (28). Water-protein interactions was analyzed using the WAP web-server (29).

Coordinates and structure factors have been deposited in the PDB, with accession code: 4WQI. Figures were prepared using PyMol (30).

*Fluorescence analysis*- Equilibrium fluorescence titrations were performed in 50 mM Tris HCl, pH 7.4, with spectra collected at each GdnHCl concentration (point by point, as to avoid photo-bleaching phenomena) at 25  $^{\circ}$ C in a 1-cm quartz cuvette using a fluorescence spectrophotometer Cary Eclipse (Agilent Technologies). The final concentration of protein (both WT and T190M) was 0.1  $\mu$ M, and for each point spectra were taken after 20 minutes of equilibration time (31). The excitation wavelength was 280 nm, and emission spectra were recorded between 300 and 400 nm. The fluorescence peak change as a function of guanidine concentration was fitted to the following phenomenological equation, that allowed to calculate the Gdn<sub>50%</sub> of the transition:

$$F_{peak} = a + \frac{(b - a)}{1 + \left(\frac{c}{[GdnHCl]}\right)^d}$$

where c represents the guanidine concentration at the midpoint of the curve (Gdn<sub>50%</sub>), and a, b and d are fitting parameters, with d denoting the slope of the curve; calculations were carried out using the Kaleidagraph software. All the reagents used were of high purity grade from Sigma.

*Limited proteolysis*- WT and T190M at 30  $\mu$ M in 50 mM Tris (pH 8) buffer were subjected to limited proteolysis at 37 °C by the addition of trypsin (Sigma) to a final concentration of 2  $\mu$ g/mL. The reaction was stopped after 1, 5, 10, 20, 40 and 60 min by adding SDS sample buffer to aliquots of the reaction mixture. The samples were analyzed by performing 15% Acrylamide SDS-PAGE and Coomassie staining. As a control the same time course was performed in the absence of the enzyme.

*SuperResolution Microscopy with 3D SIM*- Human osteosarcoma U2OS cells cultured in DMEM, 10% FBS, were seeded on #1.5H precision cover glass (Marienfeld-Superior). The cells were transfected with the corresponding pEGFP-N1 vectors expressing fusions to WT and T190M with polyethylenimine (PEI) precipitates based on a 3:1 ratio of PEI ( $\mu$ g) to total DNA ( $\mu$ g). 48h post transfection the cells were fixed in 4%PFA. Lipid staining was achieved after 30min incubation in a 1:2000 dilution of Cellmask Plasma Membrane Orange dye (Invitrogen). The cover glass was mounted in Vectashield containing DAPI (Vector Laboratories) for nuclear counter stain and sealed with nail polish. U2OS cells expressing WT-EGFP and T190M-EGFP were imaged on a 3D Structured Illumination Microscope (3D-SIM) OMX v.4 (GE Healthcare). The refraction index of the immersion oil was chosen to avoid spherical aberrations in the green emission channel. DAPI was excited with a 405 nm, EGFP with 488 nm and Cellmask Plasma Membrane Orange with 568 nm laser while sCMOS cameras recorded the corresponding emission in the following bandwidth of 436/31 nm, 528/48 nm and 609/37 nm. The 3D structured illumination images were reconstructed in the Softworx v6.1.1. software (GE Healthcare) at a wiener filter setting of 0.004 and using one week old recorded OTFs optimized for the 528 nm emission channel. For presentation here, the final images were adjusted linearly in intensities with the software package Fiji (32) to minimize any visible spherical aberrations in the blue and red emission channels. Raw data is available upon request.

## RESULTS

*Crystallization and crystal structure determination*- Well shaped crystals of T190M grew in 7-10 days as regular slabs that diffracted X-rays.

The crystal structure of T190M has been determined and refined up to a resolution of 1.6 Å. The final model includes residues in the ranges 58-163 and 179-303 which could be traced in the  $2F_o - F_c$  electron density map. The overall structure of T190M does not show remarkable differences with respect to WT. The RMSD between the T190M and WT structures is 0.758 Å (calculated on 268 C $\alpha$ ) which decreases to 0.297 Å when omitting from the RMSD calculation the 4 residues (157-160) being part of the flexible loop discussed in the next paragraph. As reported for WT (13), the structure of T190M consists of two domains, (Fig.1): the first one described as an Ig-like domain (residues 60-158) and the second one (residues 180-313) similar to the small subunit ribosomal protein S6. The spatial organization of the domains is conserved in the T190M crystal structure as well as the interaction surfaces between the two domains. A flexible loop, which links together the two domains, could be traced only at the N-terminal side (residues 159-163: His-Asn-Glu-Pro-Gln). It should be noted that the WT structure refers anyway to a protein carrying the additional single mutation R166H that was originally introduced to stabilize the recombinant protein during purification (13). The same mutation is also present in T190M, although this position falls into the region of the flexible loop that could not be traced either in WT or in T190M (see below).

The crystallographic analysis confirmed the substitution of Thr190 with Met, as showed in Fig.2 and, according to the refined model, the T190M mutation does not introduce significant conformational changes in the nearby residues. These findings altogether indicate that this mutation does not significantly alter the (either overall or local) structure of T190M with respect to that of WT.

Despite showing almost identical crystal structures, WT and T190M differ in the conformation of the flexible loop linking the two

domains (residues 159-179). Indeed the conformation of the residues 158-163 in T190M, that has been best modeled based on the Fourier maps (Fig.3) and by monitoring the  $R_{\text{free}}$ , differs with respect to the one being observed for residues 158-160 in WT (PDB id: 1U2C). It is worthy of note that the Fourier maps pinpoint to at least two alternative main-chain loop conformations (Fig.3, Fig.S1). However, attempts to model the loop with a conformation as observed in the WT structure did not result in improvement of either the Fourier maps or of the  $R_{\text{free}}$ , indicating that the WT conformation contributes with a rather low occupancy factor. The refined structure suggests that the ensemble of main-chain loop conformations has changed in T190M and that the most populated conformational state differs from that observed in WT. Moreover, in T190M the loop could be traced for three additional residues, up to Gln163, with respect to the WT structure, with residues 158-161 assuming a  $\beta$ -Turn structure (Fig.3, Fig.S2). This observation supports a different 159-179 loop organization in the T190M structure, at least in its N-terminal part. Furthermore, the better loop definition should not be merely taken as an effect of the higher resolution of the T190M crystal structure with respect to that of the WT (refined at 2.3 Å), since the backbone trace could not be extended in any of the other otherwise undefined regions of T190M such as the N- and C-terminal regions (residues 58 and 303) or the C-terminal part of the flexible loop. Being the WT and the T190M crystallization conditions almost identical and the crystal structures exhibiting the same packing features, it is reasonable to correlate such observed differences to the Thr vs Met mutation at position 190, which would result in an increased proximity of the loop to the side chain of the mutated residue.

Residue 190 is located at the center of the B1  $\beta$ -strand, at the bottom of the basket-like domain and is solvent exposed (13). The point mutation Thr to Met does not seem to introduce any significant change in the local structure. The Met residue nicely superpose on the position occupied by Thr in the WT (Fig.2), with the C $\beta$ -C $\gamma$  bond rotated by  $\approx 20^\circ$  with respect to the corresponding

bond (C $\beta$ -C $\gamma$ 2) in Thr, while the C $\gamma$ -S $\delta$ -C $\epsilon$  group is pointing towards the solvent region. Furthermore, in the T190M structure, the nearby residues geometries are not significantly affected by the point mutation.

Despite the limited effect of the Thr to Met mutation on the neighboring residues, the introduction of a bulkier and apolar group such as the methylthio group of the methionine causes a re-organization of the surrounding water molecules. An analysis made with PISA (28) shows that the solvent exposed surface is more than doubled when considering Thr in the WT ( $\approx 13 \text{ \AA}^2$ ) and Met in the T190M mutant ( $\approx 34 \text{ \AA}^2$ ). Furthermore, the solvation energy contribution estimated by PISA for each residue in the crystallographic models goes from -0.01 kcal/mole (Thr in WT) to 0.74 kcal/mole (Met in T190M), indicating a potentially perturbing effect of the methionine on the surrounding water molecules.

An analysis of the solvent structure in the crystallographic models shows that several waters surrounding residue 190 are conserved in WT and T190M. Their positions, although shifted by 0.4-1.0 Å, display the same interactions network. Nonetheless, some differences in the solvent organization are actually evident (as shown in the two panels of Fig.4). Namely, WAT304 (WT, PDB\_id: 1U2C) interacts with Thr190 OH in WT but it is absent in T190M (Fig 4a). The lack of WAT304 in T190M is rather interesting, as this water molecule bridges together the B1 and B5  $\beta$ -strands by interacting with the backbone of residues Val188 (O) and His296 (O), and with the side chains of residues Thr190 (O $\gamma$ 1) and His296 (N $\delta$ 1). Moreover, solvent structure changes were also found in a region occupied in T190M by Asn160, whereas the same space is populated by solvent waters in WT (Fig. 4b). Indeed, Asn160, which is part of the flexible loop discussed above, protrudes from the loop towards Met190 and interacts through O $\delta$ 1 with a WAT660 water molecule (T190M, PDB\_id: 4QWD), which in turn interacts with O $\gamma$  of Ser257 and with other solvent molecules. Such a water molecule is not present in the WT crystal structure.

It is quite apparent that the substitution of a

hydrophilic residue such as threonine with a bulkier and apolar one like methionine results in a solvent re-organization around the residue itself. Changes in the structure solvation sphere may have significant consequences on the overall stability of the protein domains or in locally altering the conformational ensemble distribution of the 159-179 loop, that in turn may also affect the reciprocal orientation of the two domains.

*Stability measurements-* Chemical unfolding experiments have been performed in order to establish the relative stabilities of WT and T190M in solution. Fluorescence emission has been used as a mean to record the exposure of the aromatic residues (mainly Trp) to the solvent, with emission maxima that range from  $\approx 310$  nm of the fully buried to  $\approx 350$  nm of the fully exposed tryptophan residues (33). The four tryptophan residues in the WT primary sequence are good fluorescent probes of the unfolding behavior of the domain and its mutant upon addition of a chemical denaturant such as Guanidine Hydrochloride (GdnHCl). Fig.5A shows the unfolding curves of WT and T190M as a function of GdnHCl concentration. In the absence of denaturant, the emission spectra of both WT and T190M are those of native-like proteins, with peaks in the spectral region of buried tryptophan residues; the fact that the mutant's peak is slightly red-shifted as compared to the wild-type might reflect an increased exposure of the mutant to the solvent. The plot of the emission spectra peaks as a function of GdnHCl reveals a cooperative unfolding behaviour for both WT and T190M, with transition midpoints (Gdn<sub>50%</sub>, i.e. the denaturant concentration inducing 50% of the total peak shift) of 1.45M and 1M, respectively, indicating that WT is slightly more stable than T190M.

To gain further insights into the stability of the two variants of  $\alpha$ -DG, limited proteolysis experiments have been performed. Both variants contain 19 predicted potential sites for trypsin cleavage (Peptide cutter tool at <http://web.expasy.org>) and therefore are likely to be completely degraded following trypsin digestion. T190M was rapidly cleaved by trypsin and completely degraded within 10 minutes of

digestion. WT instead was less susceptible to tryptic cleavage and showed a clear degradation only after 40 minutes of digestion (Fig. 5B) but was still present after 1 hour of digestion. Considering that neither of the two variants is subjected to degradation in the absence of the protease in the same experimental conditions (data not shown), T190M is clearly more susceptible to proteolysis than WT, confirming the latter to be more stable, in accordance with the results of the chemical denaturation experiments.

*SuperResolution Microscopy-* The Osteoclast cell line U2OS has been transfected both with the EGFP plasmids carrying the WT and T190M full-length dystroglycan variants (19), respectively and the cells have been analyzed by 3D structured illumination microscopy on an OMX v.4 system..

No relevant differences have been observed between WT (Fig. 6A) and T190M (Fig. 6B). In both variants,  $\alpha$ -DG undergoes its maturation cycle and it is extensively trafficked via the endoplasmic reticulum to the plasma membrane where it is nicely localized to filopodia and also to podosomes located at the cell periphery, suggesting no dramatic alterations as far as the distribution of  $\alpha$ -dystroglycan/ $\beta$ -dystroglycan is concerned.

## DISCUSSION

According to the high resolution X-ray crystal structure here presented, the overall structure of the N-terminal domain of  $\alpha$ -DG is not significantly affected by the T190M mutation, nor remarkable local conformational changes of the residues close to position 190 have been detected. At large, this finding seems to be fully supported by the fluorescence microscopy analysis, that does not record dramatic differences in the phenotype of T190M transfected osteoclasts with respect to their WT counterpart. The outcome of the fluorescence microscopy experiments (showing that at least  $\beta$ -dystroglycan has a typical localization pattern in T190M) is definitely indicating that the mutation is acting in quite a subtle and intriguing molecular regulatory fashion, or anyway that its effect would be much

less observable in an isolated cellular context than in skeletal muscle fibres (18).

The major effect of the T190M mutation seems to be related to the disruption by the bulkier and apolar methionine side chain of the surrounding solvent network structure. Indeed, it is well known that solvation may play a critical role on the stability and the conformational equilibrium of a macromolecule (34), as well as on the protein-ligands interactions (35). Furthermore, it has been reported that a single mutation, even at the surface of a protein, can affect the folding/unfolding free energies as a result of altering the solvent organization (36). In T190M the differences observed in the solvent organization around Met190 may account for the structural and thermodynamic differences observed with respect to WT. Indeed, T190M resolves in the formation of a new hydrogen bonding network involving Asn160 and the water molecule WAT660, which may influence the flexibility of the 159-179 loop. Moreover, the loss of stabilizing interactions between the B1 and B5  $\beta$ -strands due to the absence of WT WAT304, may account for the T190M lower stability revealed by the chemical denaturation and limited proteolysis experiments (see below).

Indeed, a particularly interesting aspect seems to be that in this case we have been able to trace part of the flexible loop (encompassing residues 160-180) that in the WT domain was even less structured, due to its high mobility and flexibility (13). Such a result may imply that the Thr to Met mutation has interesting consequences for the structural dynamics of the entire N-terminal domain. The enhanced rigidity of T190M may affect the overall dynamics and orientation of the two subdomains within the L-shaped  $\alpha$ -DG molecular architecture, eventually influencing its global conformation and its function.

This possibility seems to be corroborated by the evidence that in solution T190M has a lower stability than the WT, as suggested by the different melting points calculated on their unfolding curves. Although the T190M denaturation curve has a melting point at about 1 M GdnHCl, denoting a rather stable protein, the decrease in stability with respect to WT is small

but significant, as further confirmed by limited proteolysis experiments.

The exact molecular mechanism leading to hypoglycosylation of  $\alpha$ -DG carrying the T190M (T192M in human) mutation and causing a severe dystrophic phenotype, accompanied by some degree of intellectual impairment, is currently unknown (18). Via co-immunoprecipitation experiments, it was shown by Hara and colleagues that the mutated N-terminal domain of  $\alpha$ -DG is likely to exhibit a reduced affinity towards LARGE, that is responsible for adding a repeating disaccharide unit [ $\alpha$ 3-GlcA- $\beta$ 3-Xyl-], that is thought to be crucial for extracellular matrix protein binding, to O-mannosylated  $\alpha$ -DG (18). It is then possible that the T190M/T192M mutation is indeed provoking some degree of local perturbation that is affecting the conformation of the entire domain, or at least, of the two modular domains, influencing its binding to LARGE.

Kanagawa and colleagues have shown that the N-terminal domain of  $\alpha$ -DG is shed by furin, in a process that is likely to take place into the intracellular Golgi compartments in which the final steps of the DG maturation process are completed (37). Based on a possible alternative chronological order dictating the chaotic molecular events taking place in the Golgi, in which the furin enzymatic step would anticipate LARGE binding to  $\alpha$ -DG, it can be postulated that such an autonomous domain of  $\alpha$ -DG (12) may represent not only a recognition substrate but also an (allosteric?) activator for LARGE (38).

Interestingly, following the findings of a recent computational docking study, it was hypothesized that the T190M/T192M mutation might affect directly the binding of  $\alpha$ -DG to laminins (39). However, we believe that such a mechanism would be highly unlikely. Although we previously measured in the recombinant WT the presence of some residual binding activity towards laminin-1 (13), such an activity was displaying itself at a much lower affinity than that observed for full-length and glycosylated native  $\alpha$ -DG binding to laminin and agrin (5,11,40). In addition, such residual binding activity was apparently harboured by the first Ig-like domain of  $\alpha$ -dystroglycan and not by the S6-like domain

where the T190M/T192M mutations resides (13).

$\alpha$ -DG post-translational O-glycans decorations along the endoplasmic reticulum and modifications accomplished in the Golgi, are crucial for its ability to function as an extracellular matrix (EM) receptor and namely to bind to EM proteins such as laminin, perlecan, agrin and neurexin in the brain and pikachurin in the retina (5). Notably, the same are required for the binding of  $\alpha$ -DG to Old World arenaviruses (16). Recently, a  $\beta$ -1,4-glucuronyltransferase, designated B4GAT1, has been also identified to be involved in the initiation of the LARGE-dependent enzymatic process (41,42).

Although a few glycosylation sites have been proposed and/or localized within the N-terminal and C-terminal domains of  $\alpha$ -DG, most of the O-linked oligosaccharide chains protrude from Thr and Ser residues placed within its central mucin-like domain (2,43,44). A thorough biophysical characterization of the “naked” mucin-like domain of  $\alpha$ -DG clearly pointed to the dramatic role played by the glycosylation in its conformational stability (45).

Recently, the unveiled structural basis for the recognition of an O-glycan, GalNAc-Neu5Ac, and its attached mucin-like peptide GPATPAP by the paired Ig-like type 2 receptor  $\alpha$  (46) prompted us to envisage that the Ig-like modular architecture of the N-terminal  $\alpha$ -DG region may also function as a protein “hub” for the recognition of the oligosaccharides protruding from the neighboring  $\alpha$ -DG mucin-like domain. Furthermore, it may be a docking site for one of the two domains of the glycosyltransferase LARGE (47), belonging to the

GT-8 and GT-49 families respectively (48). Possibly, in solution, the effect of the T190M mutation, being next to the flexible hinge connecting the two subdomains, may “propagate” to the Ig-like domain by changing its relative orientation with respect to the S6-like domain also influencing its binding properties. It is worthy of note that the presence of a second Ig-like domain within the C-terminal region of  $\alpha$ -DG has been confirmed by bioinformatic analysis and molecular modeling; the domain has been biochemically characterized and shown to include the  $\beta$ -DG binding epitope (49).

Such interesting structural hypotheses, necessarily await further rounds of experimental/computational validations by means of techniques being able to explore the dynamical behaviour of  $\alpha$ -DG (NMR, Molecular Dynamics). Furthermore, molecular modeling and docking tools will be exploited to investigate the likelihood of  $\alpha$ -DG binding to LARGE or to other regulatory factors important for dystroglycan maturation. Our molecular analysis of the N-terminal T190M variant of  $\alpha$ -DG may pave the road to the elucidation of the molecular mechanism of a specific case of dystroglycanopathy, with important consequences for clinicians and for designing possible therapies, but in addition it may be also relevant from a cellular and molecular biology perspective, helping to improve the overall picture of the multiple structural aspects underlying the maturation and function of the dystroglycan complex.

## REFERENCES

1. Ibraghimov-Beskrovnaya, O., Ervasti, J.M., Leveille, C.J., Slaughter, C.A., Sernett, S.W., and Campbell, K.P. (1992) Primary structure of dystrophin-associated glycoproteins linking dystrophin to the extracellular matrix. *Nature* **355**, 696-702.
2. Barresi, R., and Campbell, K.P. (2006) Dystroglycan: from biosynthesis to pathogenesis of human disease. *J. Cell Sci.* **119**, 199-207.
3. Ervasti, J.M., and Campbell, K.P. (1991) Membrane organization of the dystrophin-glycoprotein complex. *Cell* **66**, 1121-1131.
4. Ervasti JM, and Campbell, KP. (1993) A role for the dystrophin-glycoprotein complex as a transmembrane linker between laminin and actin. *J. Cell Biol.* **122**, 809-823.
5. Sciandra, F., Bozzi, M., Bigotti, M.G., and Brancaccio, A. (2013) The multiple affinities of  $\alpha$ -dystroglycan. *Curr. Protein Pept. Sci.* **14**, 626-634.
6. Wells, L. (2013) The O-mannosylation pathway: glycosyltransferases and proteins implicated in congenital muscular dystrophy. *J. Biol Chem.* **288**, 6930-6935.
7. Michele, D.E., Barresi, R., Kanagawa, M., Saito, F., Cohn, R.D., Satz, J.S., Dollar, J., Nishino, I., Kelley, R.I., Somer, H., Straub, V., Mathews, K.D., Moore, S.A., and Campbell, K.P. (2002) Post-translational disruption of dystroglycan-ligand interactions in congenital muscular dystrophies. *Nature* **418**, 417-422.
8. Sciandra, F., Bozzi, M., Morlacchi, S., Galtieri, A., Giardina, B., and Brancaccio, A. (2009) Mutagenesis at the  $\alpha$ - $\beta$  interface impairs the cleavage of the dystroglycan precursor. *FEBS J.* **276**, 4933-4945.
9. Inamori, K., Yoshida-Moriguchi, T., Hara, Y., Anderson, M.E., Yu, L., and Campbell, K.P. (2012) Dystroglycan function requires xylosyl- and glucuronyltransferase activities of LARGE. *Science* **335**, 93-96.
10. Yoshida-Moriguchi, T., Willer, T., Anderson, M.E., Venzke, D., Whyte, T., Muntoni, F., Lee, H., Nelson, S.F., Yu, L., and Campbell, K.P. (2013) SGK196 is a glycosylation-specific O-mannose kinase required for dystroglycan function. *Science* **341**, 896-899.
11. Brancaccio, A., Schulthess, T., Gesemann, M., and Engel, J. (1995) Electron microscopic evidence for a mucin-like region in chick muscle  $\alpha$ -dystroglycan. *FEBS Lett.* **368**, 139-142.
12. Brancaccio, A., Schulthess, T., Gesemann, M., and Engel, J. (1997) The N-terminal region of  $\alpha$ -dystroglycan is an autonomous globular domain. *Eur. J. Biochem.* **246**, 166-172.
13. Bozic, D., Sciandra, F., Lamba, D., and Brancaccio, A. (2004) The structure of the N-terminal region of murine skeletal muscle  $\alpha$ -dystroglycan discloses a modular architecture. *J. Biol. Chem.* **279**, 44812-44816.
14. Stalnaker, S.H., Hashmi, S., Lim, J.M., Aoki, K., Porterfield, M., Gutierrez-Sanchez, G., Wheeler, J., Ervasti, J.M., Bergmann, C., Tiemeyer, M., and Wells, L. (2010) Site mapping and characterization of O-glycan structures on  $\alpha$ -dystroglycan isolated from rabbit skeletal muscle. *J. Biol. Chem.* **285**, 24882-24891.
15. Tran, D.T., Lim, J.M., Liu, M., Stalnaker, S.H., Wells, L., Ten Hagen, K.G., and Live, D. (2012) Glycosylation of  $\alpha$ -dystroglycan: O-mannosylation influences the subsequent addition of GalNAc by UDP-GalNAc polypeptide N-acetylgalactosaminyltransferases. *J. Biol. Chem.* **287**, 20967-20974.
16. Hara, Y., Kanagawa, M., Kunz, S., Yoshida-Moriguchi, T., Satz, J.S., Kobayashi, Y.M., Zhu, Z., Burden, S.J., Oldstone, M.B., and Campbell, K.P. (2011a) Like-acetylglucosaminyltransferase (LARGE)-dependent modification of dystroglycan at Thr-317/319 is required for laminin binding and arenavirus infection. *Proc. Natl. Acad. Sci. U.S.A.* **108**, 17426-17431.
17. Harrison, R., Hitchen, P.G., Panico, M., Morris, H.R., Mekhail, D., Pleass, R.J., Dell, A., Hewitt, J.E., and Haslam, S.M. (2012) Glycoproteomic characterization of recombinant mouse  $\alpha$ -dystroglycan. *Glycobiology* **22**, 662-675.
18. Hara, Y., Balci-Hayta, B., Yoshida-Moriguchi, T., Kanagawa, M., Beltrán-Valero de Bernabé, D., Gündeşli, H., Willer, T., Satz, J.S., Crawford, R.W., Burden, S.J., Kunz, S., Oldstone, M.B., Accardi, A., Talim, B., Muntoni, F., Topaloğlu, H., Dinçer, P., and Campbell, K.P. (2011b) A dystroglycan mutation associated with limb-girdle muscular dystrophy. *N. Engl. J. Med.* **364**, 939-946.

19. Morlacchi, S., Sciandra, F., Bigotti, M.G., Bozzi, M., Hübner, W., Galtieri, A., Giardina, B., and Brancaccio, A. (2012) Insertion of a myc-tag within  $\alpha$ -dystroglycan domains improves its biochemical and microscopic detection. *BMC Biochem.* **13**:14.
20. Schagger, H., and von Jagow, G. (1987) Tricine-sodium dodecyl sulfate-polyacrylamide gel electrophoresis for the separation of proteins in the range from 1 to 100 kDa. *Anal. Biochem.* **166**, 368-379.
21. Kabsch, W. (2010) XDS. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 125–132
22. Collaborative Computational Project, N. 4 (1994) The CCP4 suite: programs for protein crystallography. *Acta Crystallogr. D Biol. Crystallogr.* **50**, 760–763
23. McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C., and Read, R. J. (2007) Phaser crystallographic software. *J. Appl. Crystallogr.* **40**, 658–674
24. Afonine, P. V., Grosse-Kunstleve, R. W., Echols, N., Headd, J. J., Moriarty, N. W., Mustyakimov, M., Terwilliger, T. C., Urzhumtsev, A., Zwart, P. H., and Adams, P. D. (2012) Towards automated crystallographic structure refinement with phenix.refine. *Acta Crystallogr. D Biol. Crystallogr.* **68**, 352–367
25. Emsley, P., Lohkamp, B., Scott, W. G., and Cowtan, K. (2010) Features and development of Coot. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 486–501
26. Chen, V. B., Arendall, W. B., Headd, J. J., Keedy, D. A., Immormino, R. M., Kapral, G. J., Murray, L. W., Richardson, J. S., and Richardson, D. C. MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 12–21
27. Heinig, M., and Frishman, D. (2004) STRIDE: a web server for secondary structure assignment from known atomic coordinates of proteins. *Nucleic Acids Res* **32**, W500–502
28. Krissinel, E., and Henrick, K. (2007) Inference of Macromolecular Assemblies from Crystalline State. *J. Mol. Biol.* **372**, 774–797
29. Praveen, S., Ramesh, J., Sivasankari, P., Sowmiya, G., and Sekar, K. (2008) WAP (version 2.0): an updated computing and visualization server for water molecules. *J. Appl. Crystallogr.* **41**, 952–954
30. Schrödinger, L. (2010) The PyMOL Molecular Graphics System, Version 1.3r1.
31. Di Stasio, E., Bizzarri, P., Misiti, F., Pavoni, E., and Brancaccio, A. (2004) A fast and accurate procedure to collect and analyze unfolding fluorescence signal: the case of dystroglycan domains. *Biophys. Chem.* **107**, 197-211.
32. Schindelin, J., Arganda-Carreras, I., Frise, E., Kaynig, V., Longair, M., Pietzsch, T., Preibisch, S., Rueden, C., Saalfeld, S., Schmid, B., Tinevez, J., White, D.J., Hartenstein, V., Eliceiri, K., Tomancak, P. and Cardona, A. (2012) Fiji: an open-source platform for biological-image analysis. *Nat. Methods* **9**, 676-682
33. Lakowicz, J.R. Principles of fluorescence spectroscopy (2006) 3<sup>rd</sup> ed. Springer. ISBN 978-0-387-46312-4
34. Levy, Y., and Onuchic, J. N. (2006) Water Mediation in Protein Folding and Molecular Recognition. *Annu. Rev. Biophys. Biomol. Struct.* **35**, 389–415
35. Ladbury, J. E. (1996) Just add water! The effect of water on the specificity of protein-ligand binding sites and its potential application to drug design. *Chem. Biol.* **3**, 973–980
36. Covalt Jr, J. C., Roy, M., and Jennings, P. A. (2001) Core and surface mutations affect folding kinetics, stability and cooperativity in IL-1 $\beta$ : does alteration in buried water play a role? *J. Mol. Biol.* **307**, 657–669
37. Kanagawa, M., Saito, F., Kunz, S., Yoshida-Moriguchi, T., Barresi, R., Kobayashi, Y.M., Muschler, J., Dumanski, J.P., Michele, D.E., Oldstone, M.B., and Campbell, K.P. (2004) Molecular recognition by LARGE is essential for expression of functional dystroglycan. *Cell* **117**, 953-964
38. Brancaccio, A. (2013) Increased levels of expression of dystroglycan may protect the heart. *Neuromuscul. Disord.* **23**, 867-870
39. Bhattacharya, S., Das, A., Ghosh, S., Dasgupta, R., and Bagchi, A. (2014) Hypoglycosylation of dystroglycan due to T192M mutation: a molecular insight behind the fact. *Gene* **537**, 108-114
40. Gesemann, M., Brancaccio, A., Schumacher, B., Ruegg, M.A. (1998) Agrin is a high-affinity binding protein of dystroglycan in non-muscle tissue. *J. Biol. Chem.* **273**, 600-605
41. Willer, T., Inamori, K.I., Venzke, D., Harvey, C., Morgensen, G., Hara, Y., Beltrán Valero de Bernabé, D., Yu, L., Wright, K.M., Campbell, K.P. (2014) The glucuronyltransferase B4GAT1 is

- required for initiation of LARGE-mediated  $\alpha$ -dystroglycan functional glycosylation. *eLife* doi: 10.7554/eLife.03941.
42. Praissman, J.L., Live, D.H., Wang, S., Ramiah, A., Chinoy, Z.S., Boons, G.-J., Moremen, K.W., Wells, L. (2014) B4GAT1 is the priming enzyme for the LARGE-dependent functional glycosylation of  $\alpha$ -dystroglycan. *eLife* doi: 10.7554/eLife.03943.
  43. Liu, M., Borgert, A., Barany, G., Live, D. (2008) Conformational consequences of protein glycosylation: preparation of O-mannosyl serine and threonine building blocks, and their incorporation into glycopeptide sequences derived from  $\alpha$ -dystroglycan. *Biopolymers* **90**, 358-368
  44. Live, D., Wells, L., Boons, G.-J. (2013) Dissecting the molecular basis of the role of O-mannosylation pathway in disease:  $\alpha$ -dystroglycan and forms of muscular dystrophy. *ChemBioChem* **14**, 2392-2402
  45. Bozzi, M., Di Stasio E., Scaglione, G.L., Desiderio, C., Martelli, C., Giardina, B., Sciandra, F., Brancaccio A. (2013) Probing the stability of the “naked” mucin-like domain of human  $\alpha$ -dystroglycan. *BMC Biochemistry* **14**, 15
  46. Kuroki, K., Wang, J., Ose, T., Yamaguchi, M., Tabata, S., Maita, N., Nakamura, S., Kajikawa, M., Kogure, A., Satoh, T., Arase, H., Maenaka, K. (2014) Structural basis for simultaneous recognition of an O-glycan and its attached peptide of mucin family by immune receptor PILRa. *Proc. Natl. Acad. Sci. USA* **111**, 8877-8882
  47. Stanley, P. (2011) Golgi glycosylation. *Cold Spring Harb. Perspect. Biol.* **3**, a005199
  48. Lombard, V., Ramulu, H.G., Drula, E., Coutinho, P.M., Henrissat, B. (2013) The carbohydrate-active enzymes database (CAZy) in 2013. *Nucl. Acid Res.* **42**, D490-D495
  49. De Rosa, M.C., Pirolli, D., Bozzi, M., Sciandra F., Giardina, B., Brancaccio, A. (2011) A second Ig-like domain identified in dystroglycan by molecular modelling and dynamics. *J. Mol. Graph. Model.* **29**, 1015-1024

*Acknowledgments*—We thank Dr. Damir Bozic (University of Zurich) for thoughtful discussions and insights and Dr. Claudia Desiderio and Dr. Claudia Martelli (Catholic University of Rome) for the Orbitrap Mass Spectroscopy measurements.

## FOOTNOTES

\*This work was supported by a Wellcome Trust Career Re-entry fellowship to MGB.

<sup>‡</sup>Both authors contributed equally to this work.

<sup>1</sup>Istituto di Biochimica e Biochimica Clinica, Università Cattolica del Sacro Cuore, Roma 00168

<sup>2</sup>Istituto di Cristallografia – CNR, Trieste Outstation, Trieste 34149

<sup>3</sup>School of Biochemistry, University of Bristol, Bristol BS8 1TD

<sup>4</sup>Department of Physics, University of Bielefeld, 33615 Bielefeld, Germany

<sup>5</sup>Istituto di Chimica del Riconoscimento Molecolare - CNR c/o Università Cattolica del Sacro Cuore, Roma 00168

<sup>†</sup>To whom correspondence should be addressed: Andrea Brancaccio, Istituto di Chimica del Riconoscimento Molecolare - CNR c/o Università Cattolica del Sacro Cuore, Largo Francesco Vito 1, 00168 Roma, Italy, Tel: ++39 (06) 301-55135;

Fax: ++39 (06) 305-3598, E-mail: [andrea.brancaccio@icrm.cnr.it](mailto:andrea.brancaccio@icrm.cnr.it) & Alberto Cassetta, Istituto di Cristallografia, CNR, Area Science Park - Basovizza, S.S.14, km 163,5 Trieste 34149 Trieste, Italy, Tel: ++39 (040) 375-7520;

Fax: ++39 (040) 922-1126, E-mail: [alberto.cassetta@ts.ic.cnr.it](mailto:alberto.cassetta@ts.ic.cnr.it)

<sup>6</sup>The abbreviations used are: DG, dystroglycan, GdnHCl, guanidine hydrochloride, EM, extracellular matrix, LARGE, like-acetylglucosaminyltransferase, GFP, green fluorescent protein. SDS-PAGE, sodium dodecyl sulfate polyacrylamide gel electrophoresis.

## FIGURE LEGENDS

**FIGURE 1.** Cartoon depiction of the T190M crystal structure. The two domains are colored in violet (Ig-like domain) and yellow (small ribosomal S6 subunit domain), respectively; the B1 and B5  $\beta$ -strands are labeled for clarity. Spheres representation is employed for highlighting Met190. The loop regions connecting the structural elements are shown for both WT (colored in cyan) and T190M (colored in red).

**FIGURE 2.** Calculated  $\sigma_A$ -weighted  $2F_o-F_c$  map (contoured at  $1.0 \sigma$  and colored in light-blue) and the  $F_o-F_c$  map (contoured at  $\pm 3.0 \sigma$  and colored in green for positive values and in orange for negative values). The maps have been calculated following the Patterson search stage using the WT as a model (PDB id: 1U2C). The final model of T190M (colored in blue) is superposed to the initial Patterson search solution (colored in pink-salmon). Both models are represented as ribbon except residues at position 190 (Thr for the template structure and Met for the mutated final model), which are represented as sticks.

**FIGURE 3.** SA-omit  $\sigma_A$ -weighted  $2F_o-F_c$  map (contoured at  $1.0 \sigma$  and colored in grey) and the  $F_o-F_c$  map (contoured at  $3.0 \sigma$  and colored in green) overlaid with the loop region encompassing residues 156-163 as the T190M final model (PDB id 4WQI). Maps have been calculated using the T190M model, where residues 156-163 have been omitted from the calculation. The loop nicely fits the  $2F_o-F_c$  map which is indicative of the electron density of the molecule.

**FIGURE 4.** Details of the H-bond networks involving selected water molecules and residues. The structural elements of the N-terminal domain of  $\alpha$ -DG are represented as ribbon; residues involved in the interaction networks are represented as sticks. **A)** H-bond network involving WAT304 in WT. Met 190 (T190M mutant) is superimposed to Thr190 (WT) for comparison, with the B1 and B5  $\beta$ -strands labeled for clarity. **B)** H-bond network in the mutated T190M involving WAT660. Interatomic distances are reported in Å.

**FIGURE 5.** A. Equilibrium unfolding curves of WT and T190M. The red-shifts in the fluorescence peaks upon addition of GdnHCl indicate a progressive exposure of the tryptophan residues to the solvent: the data

fit to a single cooperative transition for both WT, closed circles, and T190M, open circles, with a difference in midpoints ( $Gdn_{50\%}$ ) of 0.45M, that reveals the mutant to be slightly less stable than the wild-type protein. Both experiments were performed in triplicate, and no significant deviation from the above curves was detected. B. Limited proteolysis of WT and T190M. Tryptic digestion time course of the two variants.

**FIGURE 6.** 3D Structured Illumination Microscopy images of U2OS cells expressing DG and DG<sup>T190M</sup> fused to EGFP. DG-EGFP (A) and DG<sup>T190M</sup>-EGFP (B) are both localized specifically at the plasma membrane and accumulate in the cytoplasm. The columns show the individual images for DAPI (blue), EGFP (green), Cellmask Plasma Membrane Orange (red) and the resulting RGB overlay. The first row of images is the z stack maximum intensity projection of the 3D recording of the cell. The two subsequent rows of images are at a specific z position in the stack (XY) and X position slice through the stack (XZ). The white bar represents 10 $\mu$ m.

**FIGURE S1.** SA-omit  $\sigma_A$ -weighted  $2F_o-F_c$  map (contoured at 1.0  $\sigma$  and colored in light-grey) and the  $F_o-F_c$  map (contoured at 3.0  $\sigma$  and colored in dark-grey) overlaid with the loop region encompassing residues 156-160 as in the WT model (PDB id 1U2C). Maps have been calculated by using the T190M model where residues 156-163 have been omitted from the calculation.

**FIGURE S2.** Detail of the superimposition of the WT and T190M mutant. The models are represented as ribbons, with WT colored in black and T190M colored in light-grey; Met 190 is depicted as a ball-and-stick model. The superimposition of the two crystal structures emphasize the different conformations of the respective loops encompassing residues 157-163.

**TABLE I: X-ray diffraction: data collection and model refinement statistics****Data Collection**

Space group	H 3
Unit-cell parameters (Å)	a=71.879, c=144.296
Molecules per asymmetric unit	1
Wavelength (Å)	1.00
Resolution (Å)	48.1 – 1.59 (1.65 – 1.59) <sup>a</sup>
Total observations	119373
Unique reflections	35504
R <sub>merge</sub> (%) <sup>b</sup>	4.3 (44.2)
<I/σ(I)>	15.28 (1.74)
Completeness (%)	95.0 (71.8)
Redundancy	3.4 (2.3)

**Refinement**

Number of reflections (work-set/test-set)	33720/ 1784
R <sub>work</sub> <sup>c</sup> / R <sub>free</sub> <sup>d</sup> (%)	14.61 / 16.32
Number of non-H atoms	
Protein	1741
Waters	194
Organic (ethylene glycol)	12
Ions (Cl <sup>-</sup> )	1
Average isotropic B factors (Å <sup>2</sup> )	32.9
Protein (main chain)	28.9
Protein (side chain)	35.8
Solvents	39.1
r.m.s. deviation	
Bond length (Å)	0.010
Angle (deg)	1.245
Ramachandran plot	
favored regions (%)	100
allowed regions (%)	0
disallowed regions (%)	0

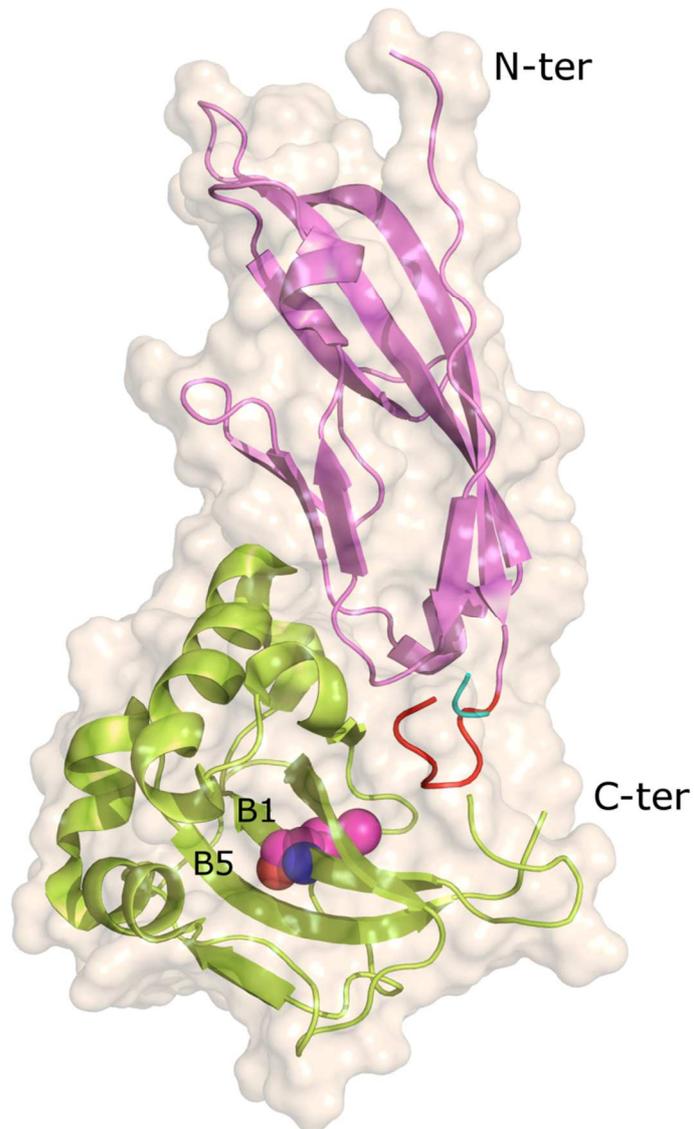
<sup>a</sup> Values in parenthesis are given for the highest resolution shell

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

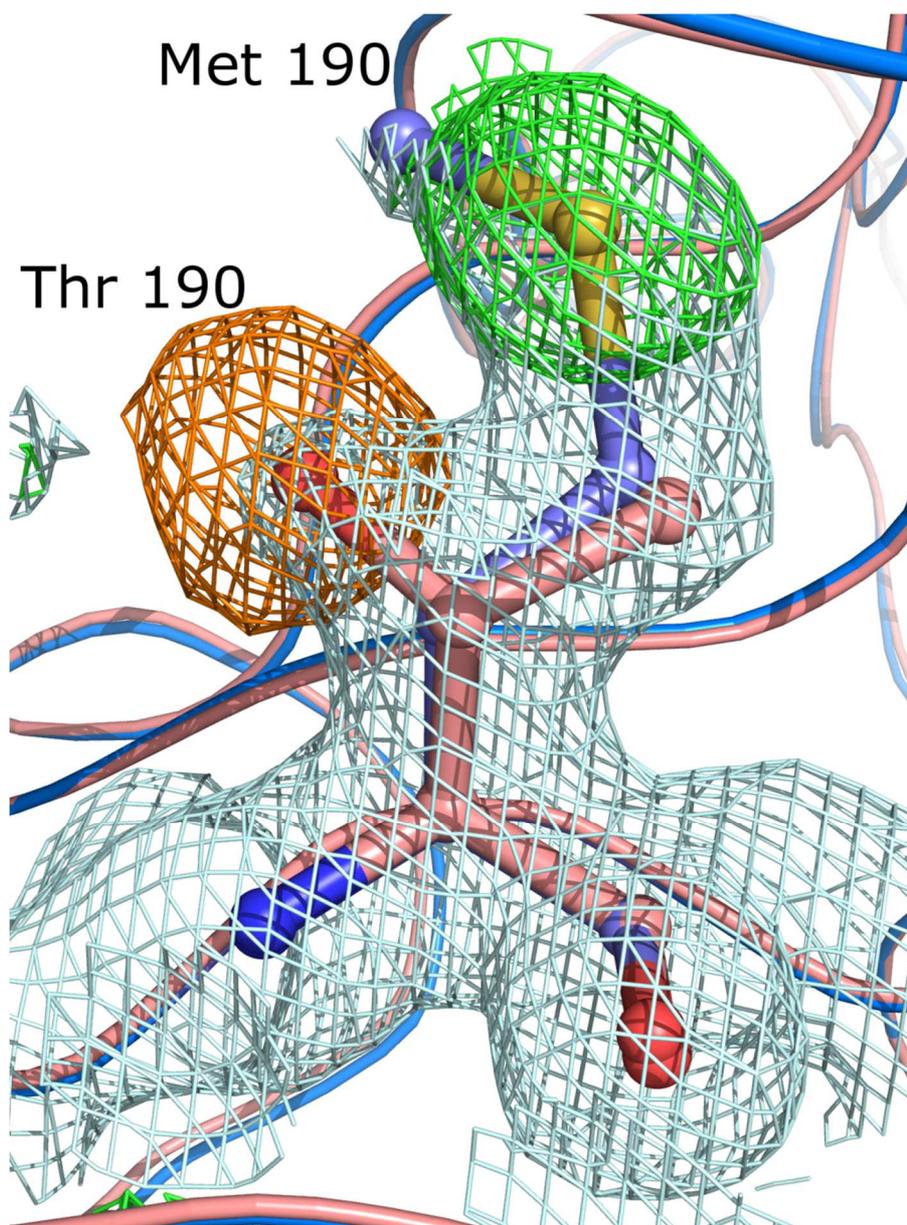
$$^c R_{\text{work}} = \frac{\sum_{\text{work-set}} |F_{\text{obs}} - F_{\text{cal}}|}{\sum_{\text{work-set}} F_{\text{obs}}}$$

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

**Figure 1**



**Figure 2**



**Figure 3**

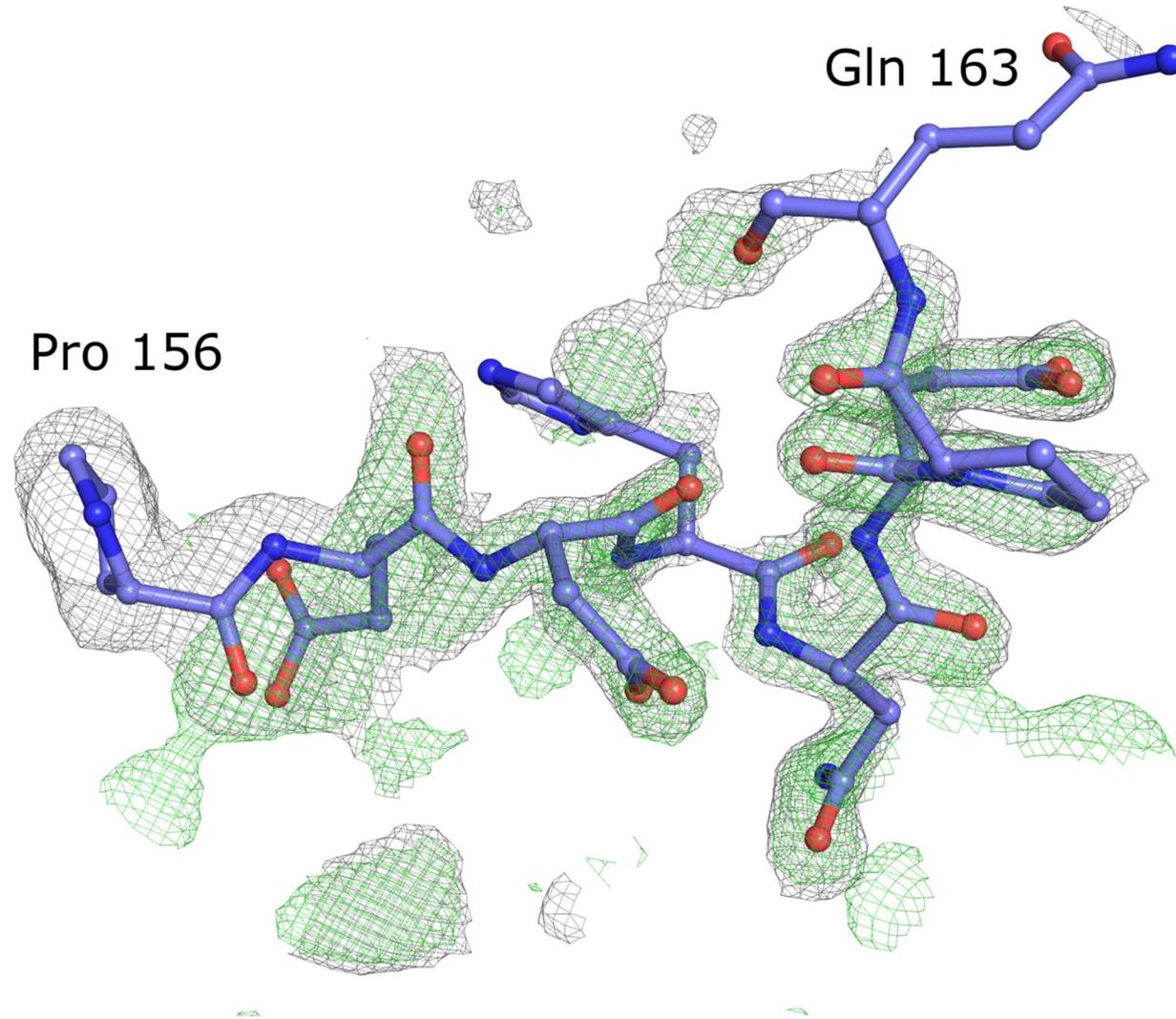


Figure 4

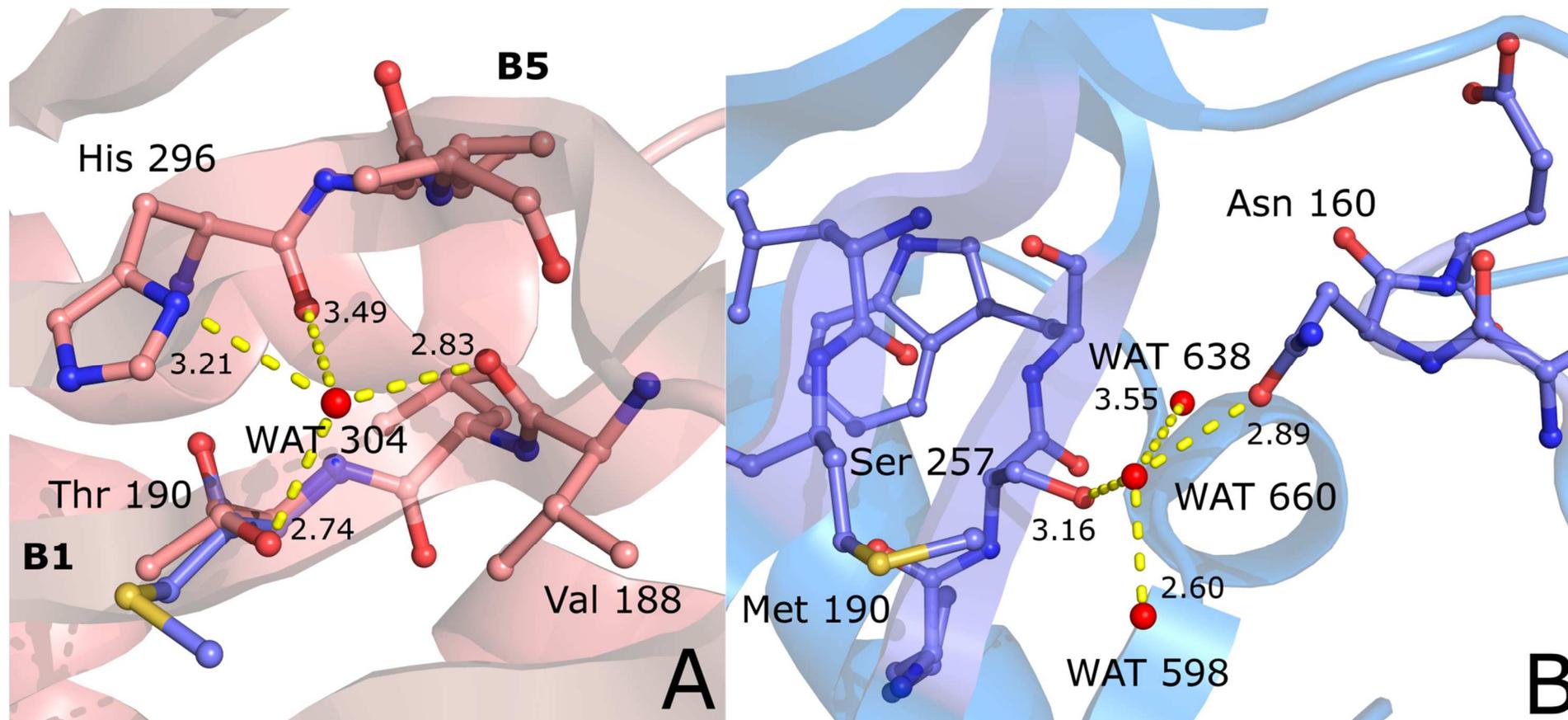
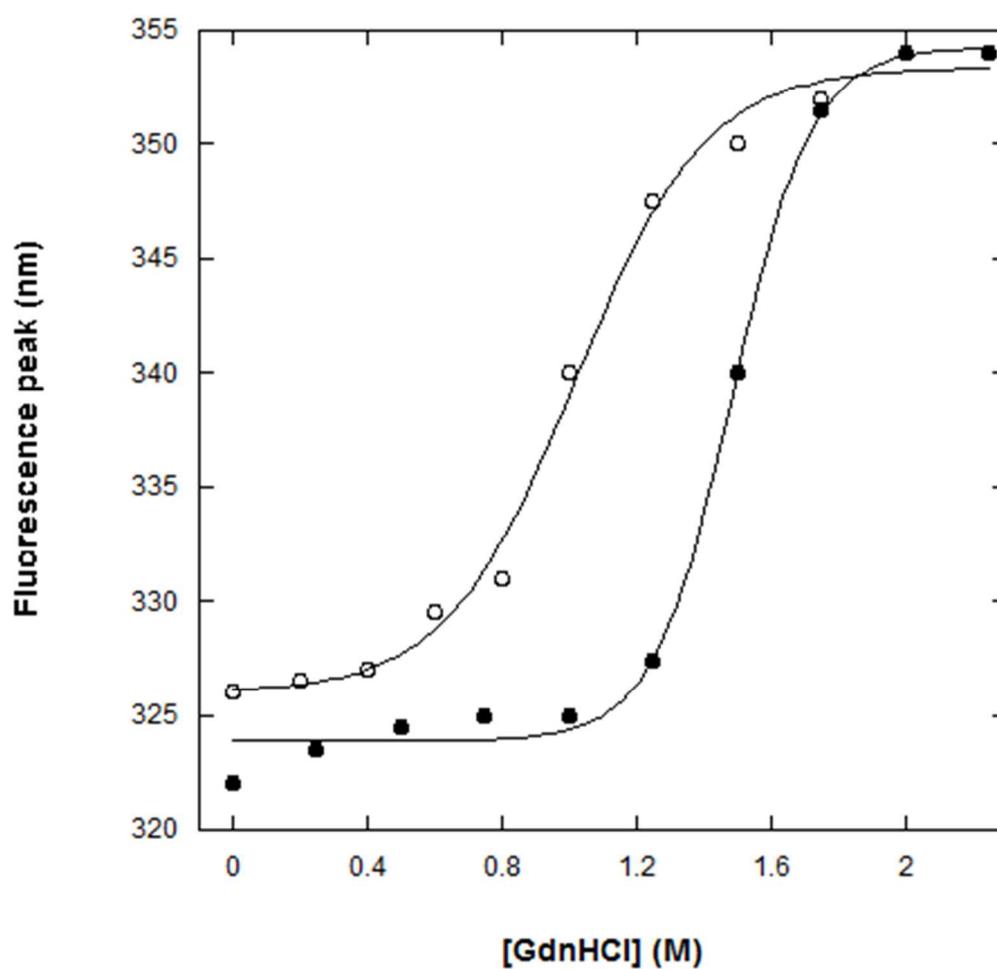


Figure 5a



**Figure 5b**

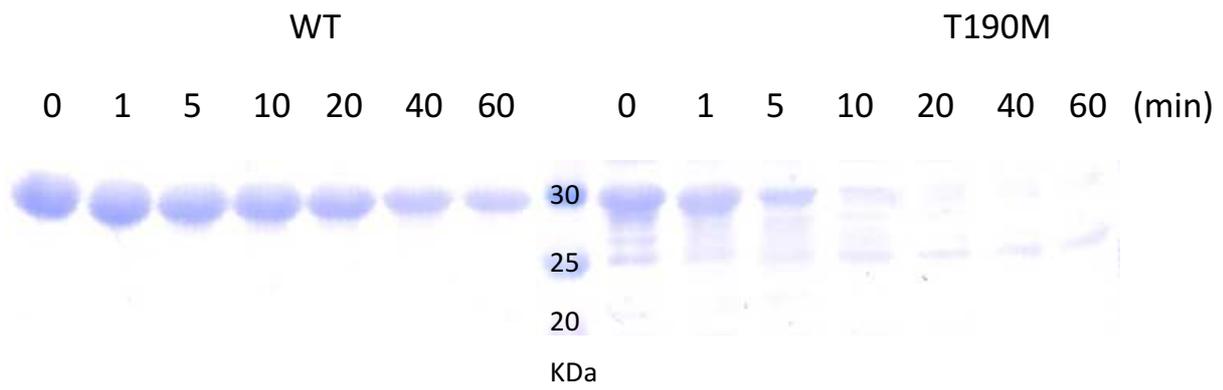
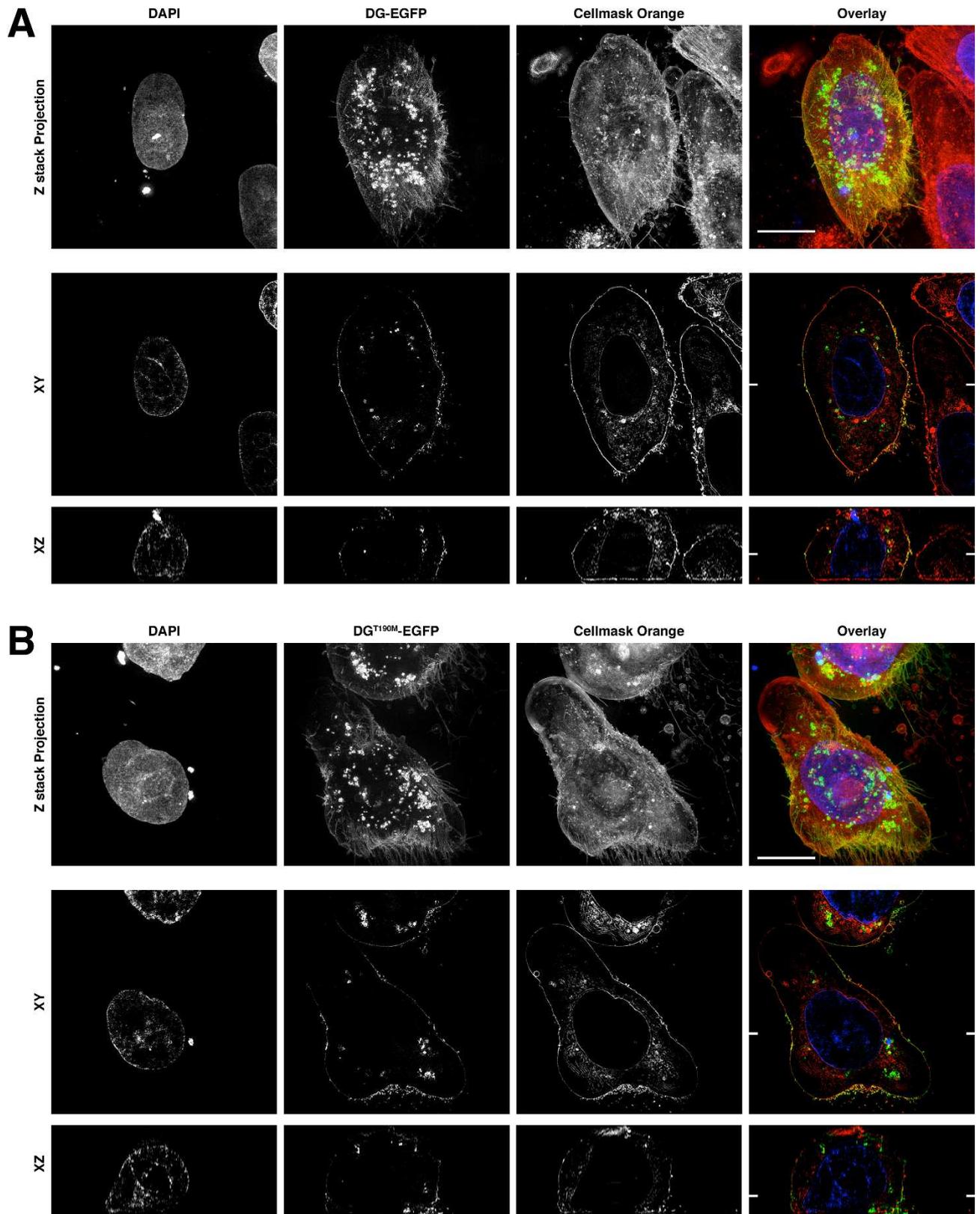
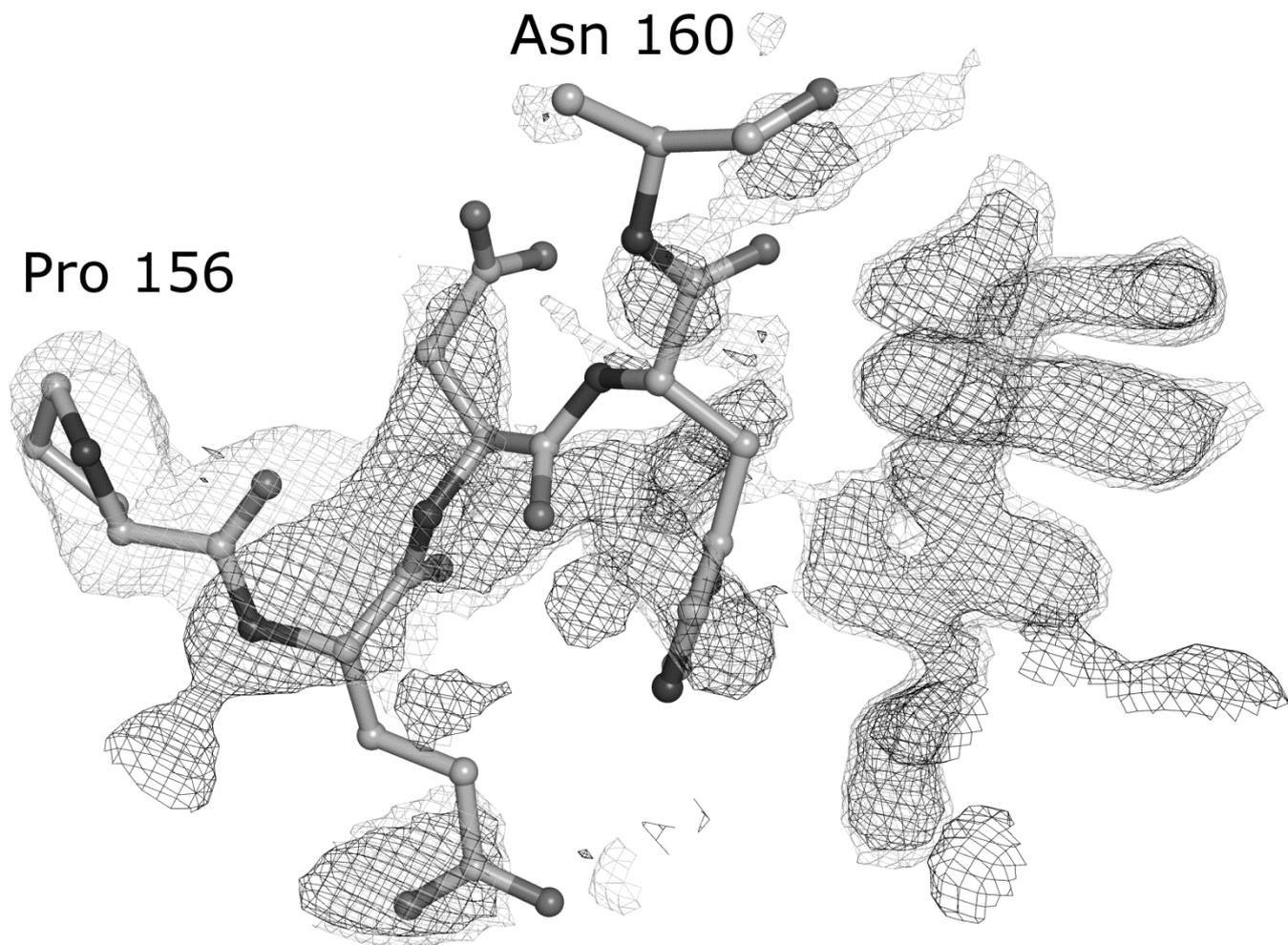


Figure 6



**Figure S1**



**Figure S2**

