JPP JOURNAL OF PROTEINS AND PROTEOMICS
4(2), December 2013, pp. 101-107

Communication

DP 71 AND BETA DYSTROGLYCAN INTERACTION: A MOLECULAR MODELING APPROACH TO UNDERSTAND DUCHENNE MUSCULAR DYSTROPHY

Simanti Bhattacharya, Amit Das, Rakhi Dasgupta and Angshuman Bagchi*

Department of Biochemistry and Biophysics, University of Kalyani, Nadia 741235, West Bengal, India

Abstract: Dp 71 is the most prevalent and widely expressed non muscle isoform of dystrophin (Dp) and its mutations are associated with Duchenne muscular dystrophy, a severe form of muscular disorder. Dp 71 deviates from the canonical Dp by means of its truncated N terminal which also has abolished certain amino acids that comprise WW domain in the canonical form. This WW domain is very crucial for Dp's interaction with partner proteins to establish a bridge between extra cellular matrices and cellular cytoskeleton. In our current study we have employed molecular modeling technique to understand the structural architecture of the N terminal region of Dp 71 and its deviation from the canonical form. We have further extended our studies to analyze the interaction probabilities between Dp 71 and β-DG applying molecular docking. Our studies for the first time have revealed that in spite of the underlying differences in terms of amino acids and structural organization, Dp 71 can interact with β-DG with its N terminal region which shares the similar molecular surface with the canonical form of Dp. These findings have opened up a platform to investigate the molecular interactions, spatio temporal orientations of the amino acids of Dp 71 and β-DG to understand the onset of DMD in much more greater detail.

Keywords: Duchenne muscular dystrophy; WW domain; dystrophin associated protein complex; homology modeling; molecular docking.

Introduction

Duchenne muscular dystrophy [OMIM: 310200] (DMD) is a common and X linked recessive lethal muscular dystrophy diagnosed as severe muscular degenerations associated with cognitive impairments (Moizard *et al.*, 2000; Daoud *et al.*, 2008). Genetic mapping analyses of the disease have identified mutation (s) at different regions in the largest gene, DMD gene (Ervasti, 2007). Dystrophin (Dp), a 427 kDa widely expressed protein, is the product of DMD gene. Dp along with dystroglycan (DG) forms the stem of dystrophin associated protein complex (DAPC) that acts as a bridge between extracellular matrix and cellular actin cytoskeleton (Ervasti, 2007; Henry and Campbell, 1996) (Figure 1A). Dp shares similarity with other actin binding proteins like α-actinin, spectrin etc in terms of its actin binding domain (ABD), spectrin repeat domains (Upper panel; Figure 1B) whereas its C terminal part is unique and this harbors its WW domain (Ervasti, 2007). WW domain is known the smallest known independently foldable protein architecture that is till date known to exist (Salah *et al.*, 2012). And it is well established that WW domain interacts with poly proline rich regions (PP region) through its conserved triple beta sheet structure (Salah *et al.*, 2012). Here also the WW domain of Dp (amino acid residues 3055-3088) has been reported to interact with PP region (amino acid residues 809-895) of beta subunit of DG (β-DG) to form a successful cascade for DAPC

Corresponding Author: **Rakhi Dasgupta, Angshuman Bagchi**
E-mail: rdgadg@gmail.com; angshuman_bagchi@yahoo.com

Received: November 10, 2013
Accepted: December 8, 2013
Published: December 14, 2013

mediated signal transduction. It is worthy to mention here that alternative splicing generates 10 isoforms for Dp. Among them isoform 4 (Dp4) has been considered as the canonical sequence and isoform 7 (Dp7or Dp71) is the highly expressed non-muscle dystrophin (Bolaños-Jiménez *et al.*, 2001). Mutations in DP71 expressing gene have been associated with two main non-muscular DMD phenotypes: retinal dysfunction and mental retardation (Tadayoni *et al*, 2012). Dp71 differs (Lower panel; Figure1B) from canonical sequence in terms of its most of the N terminal region is missing (1-3068 amino acid residues) and a modification at DP71 specific N terminal (3069-3075: KVPYYIN → MREQLKG). This reflects that DP71 lacks major residues of WW domain. But then also it forms a successful interaction with β-DG. These ambiguities have tempted us to investigate the impact of this truncated sequence on the WW domain structure of Dp71.

In our current study, we have employed homology modeling and molecular simulation techniques to understand the architecture and spatio temporal orientation of the N terminal of Dp 71 and have compared the models of Dp 4 and Dp 71 in terms of the WW domain. We have also docked the PP region of β-DG with Dp 71 to address the answer regarding the molecular interaction between β-DG and Dp 71.

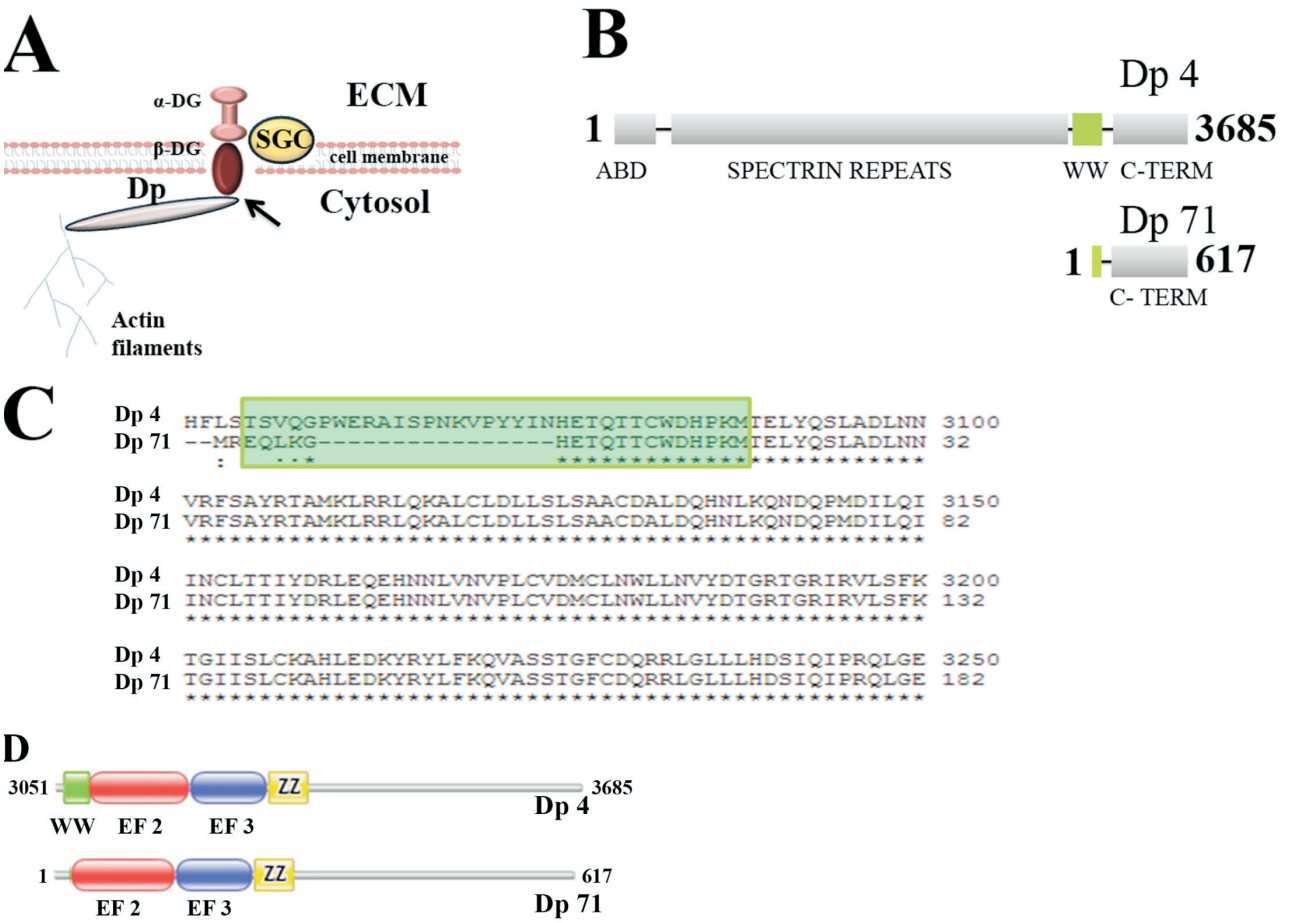


Figure 1: **Differences between Dp 4 and Dp 71 domain architectures. A. Simplified schematic representation of dystrophin associated protein complex (DAPC): Dp Dystrophin; α-DG alpha dystroglycan; β-DG beta dystroglycan; SGC sarcoglycans; extra cellular matrix ECM. Arrow indicates WW domain of Dp interacts with PP region of β-DG. B. Domain architecture of canonical Dp 4 having N terminal ABD- actin binding domain; 24 spectrin repeats and unique C terminus. WW domain for Dp 4 has been shown with green box. The majority of the N terminal portion is missing from Dp 71. C. A portion of multiple sequence alignment with total sequence reveals that except for its N terminal few amino acids, Dp 71 is identical with Dp 4. Green shaded portion marks the canonical WW domain for Dp 4. Most of the amino acids comprising WW domain are absent in Dp 71. Supplementary figure3 provides the detailed view of multiple sequence alignment. D. Pfam analyses with the amino acid stretches having maximum identity for the two sequences shows putative WW domain cannot be detected in case of Dp 71.**

Materials and Methods

At first, the total amino acid sequences for Dp 4 (UniProtKB: P11532) and Dp 71 (UniProtKB: P 11532-7) were aligned using ClustalW2 (*http://www.ebi.ac.uk/Tools/msa/clustalw2/*). The best matched portions were then used to search the domain organization in pfam (Punta *et al.*, 2012). Next, these respective sequences were employed as inputs to obtain the template to construct homology models for Dp 4 and Dp 71 in BLAST (Altschul *et al.*, 1990) against PDB (Berman, 2008) database. This returned A chain of the structure of a dystrophin WW domain fragment in complex with a beta dystroglycan peptide (pdb code: 1EG4) as the closest match template. Modbase (Pieper *et al.*, 2011) server was then used to build the models for Dp 4 and Dp 71 taking 1EG4 as template. All structures were validated using PROCHECK (Laskowski *et al.* 1993) and Verify3D (Eisenberg *et al.*, 1997). Structural validation report in the form of Ramanchandran plot (Ramachandran *et al.*, 1968) and ProSA web server [*https://prosa.services.came.sbg.ac.at/prosa.php*] for Dp 71 model have been provided in supplementary figure 1. Z dock server (Chen *et al.*, 2003) was employed to dock PP region of β-DG and N terminus of Dp 4 and Dp 71. A detailed account of energy change profile for the docked complex after minimization has been given in table 1. Analyses of resultant docked structure were carried out in discovery studio 2.5 and Protein Interaction calculator (Tina *et al.*, 2007).

Results and Discussions

Absence of the N terminal amino acids has affected WW domain architecture

Dp 71 the widely expressed and most important isoform of Dp has a large amino acids truncation at its N terminal. Multiple sequence alignment (Figure 1C) with total sequences of Dp 4 and Dp 71 has shown that apart from its N terminal rest of the amino acids residues of Dp 71 are identical with that of canonical isoform of Dp i.e. Dp 4. In Dp 4 amino acid residues 3055-3088 comprise putative WW domain (shaded green; Figure 1C). But this portion is largely missing in Dp 71 sequence. As mentioned earlier that this WW domain is very much essential for Dp to interact with the PP region of β-DG in order to maintain the signal flow. This triggers the question regarding the presence and architecture of WW domain in Dp 71. To address this query, we have used the region with maximum identity between the two sequences i.e. amino acid residues 3051-3685 of Dp 4 and 1-617 of Dp 71 respectively as input sequences for domain search in Pfam (*pfam.sanger.ac.uk*). The obtained result has clearly marked WW domain, EF hand 2, EF hand 3 and ZZ domain for Dp 4. While rest of the domains can successfully be predicted in case of Dp 71, surprisingly, WW domain cannot be predicted. Dp 71 is the first protein to be identified in embryonic stem cells (Tadayoni *et al.*, 2012) and this implies that the functioning of DAPC should have been strictly maintained. But here we are left with the finding that in case of Dp 71 WW domain cannot be detected as putative domain which is crucial for DAPC establishment. That is why we next have employed homology modeling to construct the models for both Dp 4 and Dp 71to visualize the structural orientations for both the sequences. Superimposition of the constructed models of Dp 4 and Dp 71 has shown the two structures to identical with root mean square deviation (RMSD) 0.485Å (Figure 2A). However, the conserved triple beta sheet architecture of WW domain of Dp 4 (arrow, left panel; Figure 2A) has not at all been preserved in case of Dp 71 (arrow, right panel; Figure 2A) reflecting the fact that in deed WW domain structure cannot be formed in Dp 71.

Similar surface architecture: answer to successful interaction

Our insight to the molecular structures has revealed the fact that Dp 71, the crucial protein during the early developmental stage (Tadayoni *et al.*, 2012), lacks its WW domain required for its interaction with PP region of β-DG. It is worthy to mention that molecular surface plays important role while considering protein-protein interactions. Our study with the surface architecture concerning the mismatched portion of the two models has provided a clue to the above mentioned puzzle. Figure 2B and C have clearly reflected that in spite of the prevalent differences at amino acid levels as well as triple beta sheet structure, the surface architecture is similar in both the cases with identical

Table 1
Energy change profile for minimized docked complexes. The resultant docked complexes have been subjected to minimization. The energy change profile has been represented

Name	*Forcefield*	*Initial Potential Energy (kcal/mol)*	*Potential Energy (kcal/mol)*	*Van der Waals Energy (kcal/mol)*	*Electrostatic Energy (kcal/mol)*	*Initial RMS Gradient (kcal/(mol x Angstrom)*	*Final RMS Gradient (kcal/(mol x Angstrom)*
Dp4complex.1	CHARMm	2122138.5838	-2863.20910	-220.18213	-3050.14967	27996.8040	0.09991
Dp7_complex.1	CHARMm	9062.68327	-1056.90781	-61.99305	-1205.73884	4458.75526	0.09972

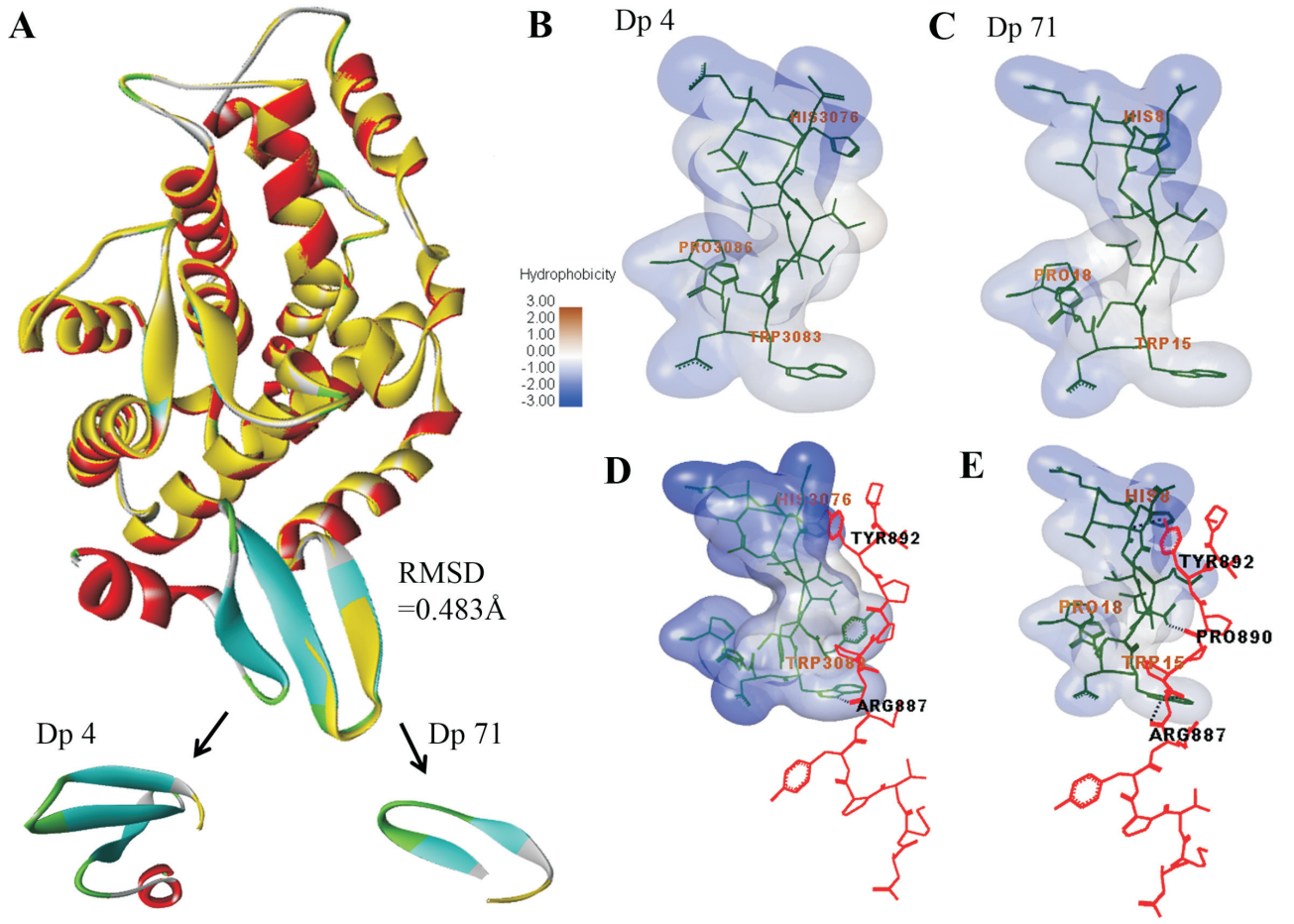


Figure 2: **Dp 71 can interact with PP region of β-DG. A. Superimposition of the modeled structures of Dp 4 and Dp 71 (yellow) reflects that these two structures are identical (RMSD=0.483Å) except for their N terminal regions, shown with arrow. B-C. In spite of the differences at their amino acid level, both Dp 4 and Dp 71 possess similar hydrophic surface architectures at their N terminal region. D-E. Molecular docking with PP region (red backbone) of β-DG has also been successfully occurred at the groove of Dp 4 and Dp 71 (green backbone) N terminus which is similar to the canonical Dp 4 structure.**

hydrophobic nature of these regions. And with this clue, we have proceeded to dock the PP region of β-DG with N terminus region of Dp 4 and Dp 71. We have analyzed the molecular interactions of the docked complexes (Figure 2D and 2E). Interestingly, we have found that PP region of β-DG has successfully interacted with the N terminus region of Dp 71 and the surface groove for this region of Dp 71 is very similar with that of Dp 4. We have also separately analyzed the inter molecular hydrogen bonds formed between PP of β-DG and WW domain of Dp in

case 1EG4. Interestingly, monitoring of possible inter molecular hydrogen bonds, formed in case of our docked model (black dashed lines; Figure 2D and 2E) have shared identity with that of 1EG4, strengthening further our finding. Moreover we have minimized the docked complexes both for Dp 4 and Dp 71 interacting with PP region of β-DG with explicit solvent simulation system. Thereafter analyses have revealed that the interactions remain unchanged in both the cases. Superimpositions of models minimized with solvent implicit and explicit reflect the models are indeed identical (supplementary figure 2). Studies with contact map with backbone residues (Figure 3) in both the cases have generated the detailed interaction pattern of WW domain of Dp with PP region of β-DG. In spite of lacking a proper WW domain, Dp 71 has shown more or less equal intermolecular interaction pattern (arrow, Figure 3). However, docking with a larger patch of C terminal portion of β-DG comprising its PP region will provide deeper insight in this matter. Our future work in this field will therefore be comprised of molecular simulations and dynamics with the modeled constructs to enlighten the interaction between Dp 71 and β-DG.

Conclusions

In summary, our present work has shown that Dp 71 can interact with PP region of β-DG, although it lacks essential amino acids to form WW domain triple beta sheet conserved architecture. The shortened N terminus and certain altered amino acid residue in the case of Dp 71 do not hamper the overall molecular surface of this region and the surface has similarity with that of canonical Dp 4 with identical surface hydrophibicity; and this in turn provides the platform for successful interaction between PP region of β-DG and Dp 71. This is so far the first report concerning the presence and structural impact of WW domain in Dp 71 and its interaction pattern with β-DG. Further analyses of the mutations in Dp 71 (Moizard *et al.*, 2000) leading to DMD will lead us to the molecular view of the disease onset.

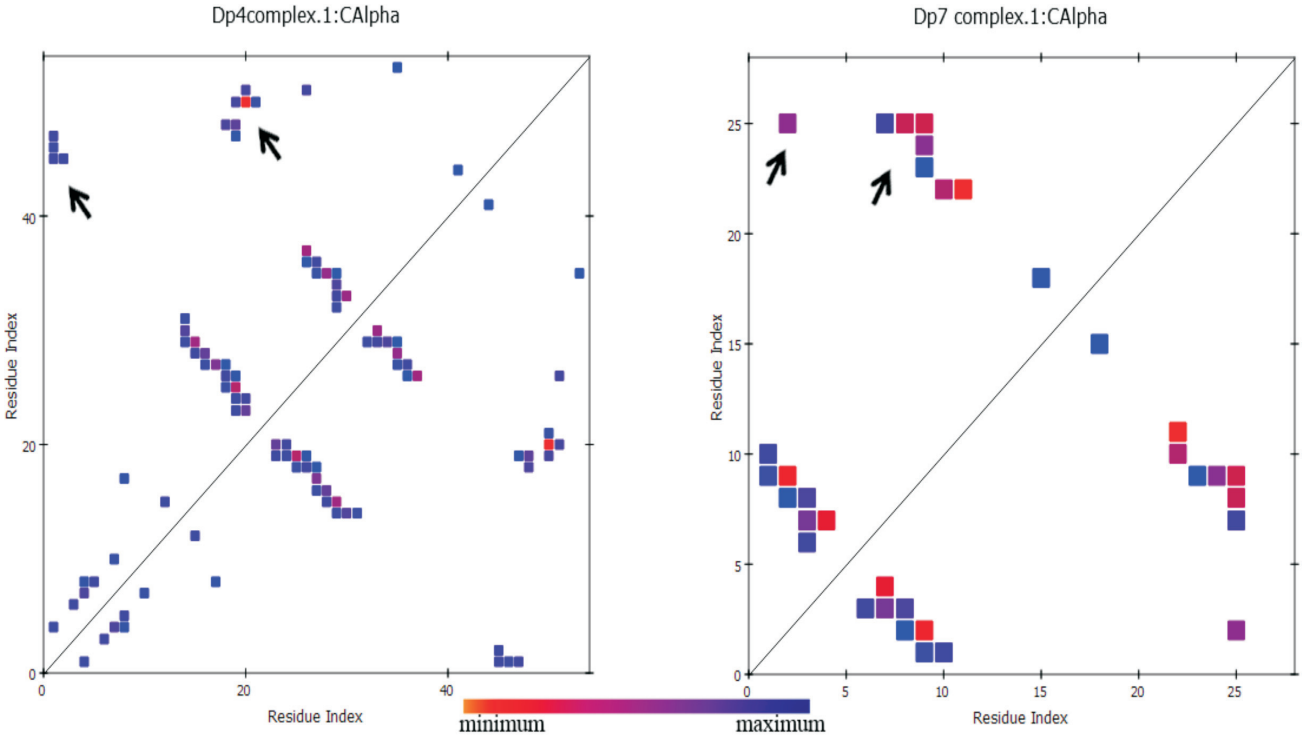


Figure 3: **Contact Map for the docked complex. Residue level contact map with backbone has been generated for the complexes separately. In the figure, left plot is for Dp 4 –PP complex and right plot is for Dp 71 –PP complex. Here blue color stands for maximum and red color stands for minimum interactions. Arrows in the figures indicate the inter protein interaction occurred between the backbone of two interacting peptide stretches.**

Acknowledgements

All the authors are thankful to Department of Biochemistry and Biophysics, University of Kalyani and Bioinformatics Infrastructure Facility of the department for providing the necessary instruments to carry out the reported study. The authors would like to thank the ongoing DST-PURSE program for financial assistance. SB and AD want to thank UGC (India) and CSIR (India) for their respective Ph.D fellowships.

Abbreviations

Dp, Dystrophin; DMD, Duchenne muscular dystrophy; ABD, Actin binding domain.

References

Altschul, S. F., Gish, W., Miller, W., Myers, E. W. and Lipman, D. J. (1990). Basic local alignment search tool. J. Mol. Biol. 215, 403-410.

Berman, H. M. (2008). The Protein Data Bank: a historical perspective, Acta Crystallogr A. 64, 88-95.

Bolaños-Jiménez, F., Bordais, A., Behra, M., Strähle, U., Sahel, J. and Rendón, A. (2001). Dystrophin and Dp71, two products of the DMD gene, show a different pattern of expression during embryonic development in zebrafish. Mech Dev. 102, 239-241.

Chen, R., Li, L. and Weng, Z. (2003). ZDOCK: an initial-stage protein-docking algorithm. Proteins 52, 80-87.

Daoud, F., Candelario-Martínez, A., Billard, J. M., Avital, A., Khelfaoui, M., Rozenvald, Y., Guegan, M., Mornet, D., Jaillard, D., Nudel, U., Chelly, J., Martínez-Rojas, D., Laroche, S., Yaffe, D. and Vaillend, C. (2008). Role of mental retardation-associated dystrophin-gene product Dp71 in excitatory synapse organization, synaptic plasticity and behavioral functions. PLoS One. 4, e6574.

Eisenberg, D., Lüthy, R. and Bowie, J. U. (1997). VERIFY3D: assessment of protein models with three-dimensional profiles. Methods Enzymol. 277, 396-404.

Ervasti, J. M. (2007). Dystrophin, its interactions with other proteins, and implications for muscular dystrophy. Biochim Biophys Acta. 1772, 108-117.

Henry, M. D. and Campbell, K. P. (1996). Dystroglycan: an extracellular matrix receptor linked to the cytoskeleton. Curr Opin Cell Biol. 8, 625-631.

Laskowski, R. A., MacArthur, M. W., Moss, D. S. and Thornton, J. M. (1993). PROCHECK - a program to check the stereochemical quality of protein structures. J. App. Cryst. 26, 283-291.

Moizard, M. P., Toutain, A., Fournier, D., Berret, F., Raynaud, M., Billard, C., Andres, C. and Moraine, C. (2000). Severe cognitive impairment in DMD: obvious clinical indication for Dp71 isoform point mutation screening. Eur J Hum Genet. 8, 552-556.

Pieper, U., Webb, B. M., Barkan, D. T., Schneidman-Duhovny, D., Schlessinger, A. and Braberg, H. (2011). ModBase, a database of annotated comparative protein structure models and associated resources. Nucleic Acids Res. 39, D465-74.

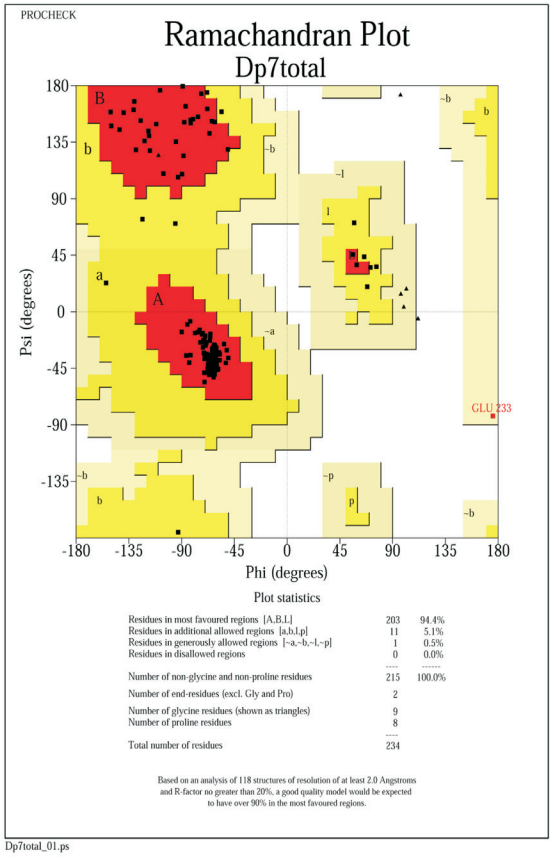
Punta, M., Coggill, P.C., Eberhardt, R.Y., Mistry, J., Tate, J., Boursnell, C., Pang, N., Forslund, K., Ceric, G., Clements, J., Heger, A., Holm, L., Sonnhammer, E.L., Eddy, S.R., Bateman, A. and Finn, R.D. (2012). The Pfam protein families database. Nucleic Acids Res. 40, D290-301.

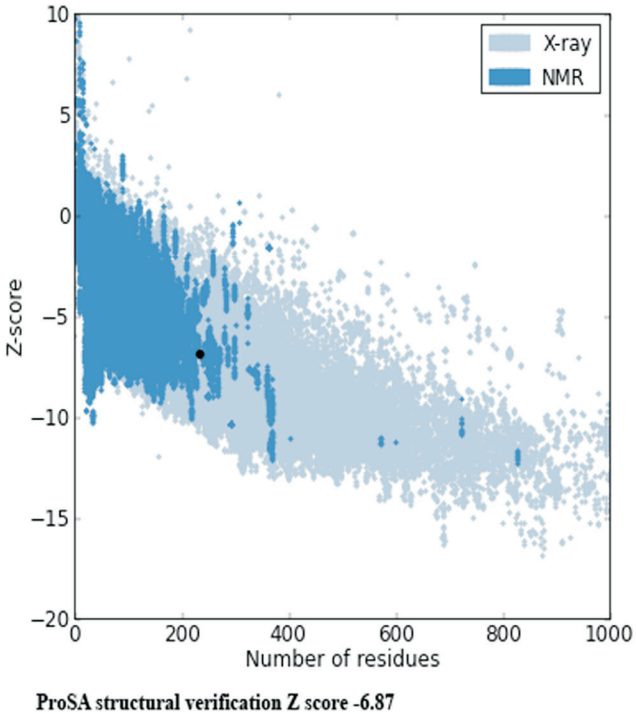
Ramachandran, G. N., Ramakrishnan, C and Sasisekharan, V. (1963). Stereochemistry of polypeptide chain configurations. J Mol Biol. 7, 95–96.

Salah, Z., Alian, A., and Aqeilan, R. I. (2012). WW domain-containing proteins: retrospectives and the future. Journal Front Biosci. (Landmark Ed). 17, 331-348.

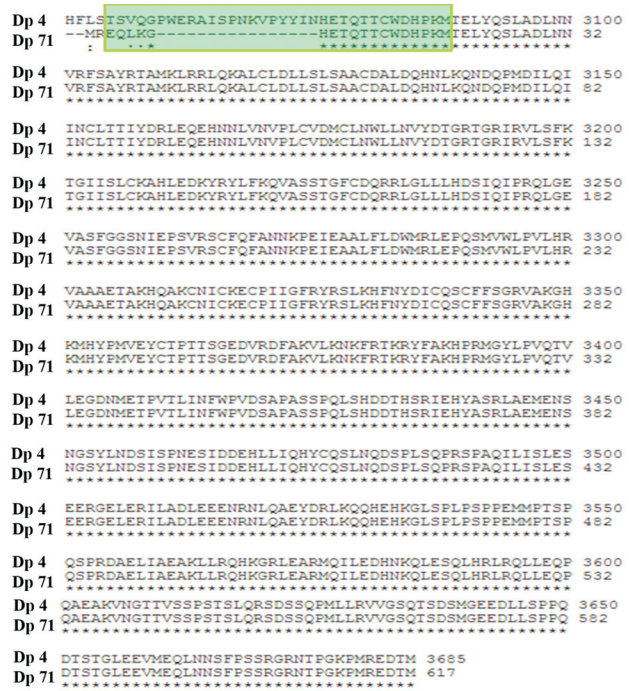
Tadayoni, R., Rendon, A., Soria-Jasso, L. E. and Cisneros, B. (2012). Dystrophin Dp71: the smallest but multifunctional product of the Duchenne muscular dystrophy gene. Mol Neurobiol. 45, 43-60.

Tina, K. G., Bhadra, R. and Srinivasan, N. (2007). PIC: Protein Interactions Calculator. Nucleic Acids Res. 35, W473–W476.

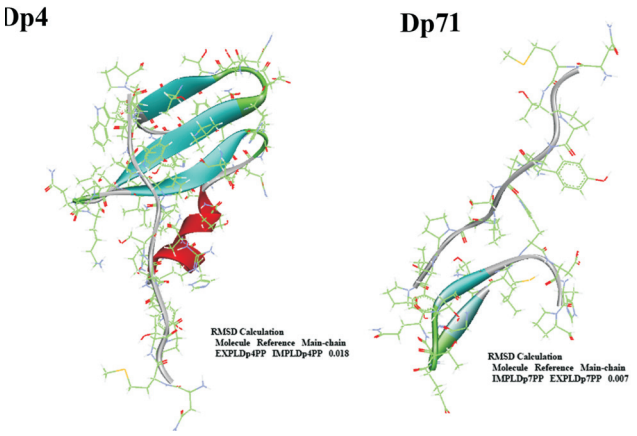




Supplementary Figure 1: **Structural validation report for Dp 71. Ramachandran plot showing there are no residues in the forbidden region as well as ProSA webserver analysis further validates the structure.**



Supplementary Figure 3. Sequence alignment of Dp 4 and Dp 71. Total sequences of Dp 4 and Dp 71 have aligned and the result reflects that except for its N terminal few residues, rest of the amino acids are identical. Green shaded portion indicates WW domain of Dp4 (UniProtKB: P11532.).



Supplementary Figure 2: **Validation of docking simulations. The docked complexes both for Dp 4 and Dp 71 interacting with PP region of â-DG have been minimized in two different conditions; one in explicit solvent system ann another with implicit solvent system. The resultant minimized structures have then been superimposed to track any structural change if present. Insignificant root mean square deviation (RMSD; 0.018Å and 0.007Å respectively) in the figures indicates that these structures are identical with respective molecular interactions intact. The backbone structure represents solvent implicit model whereas the ribbon structure represents solvent explicit model and the respective RMSDs have been mentioned in the figure.**