Indian Journal of Biochemistry & Biophysics
Vol. 63, November 2026, pp. 1357-1369
DOI: 10.56042/ijbb.v63i11.32006

Identification of differentially expressed genes and drug targets in duchenne muscular dystrophy using *in silico* approach with RNA-Seq analysis and molecular dynamics simulations

Mohandhass Venkatesan, Jino Blessy John Selesteen Charles*, Shruthi Chandrasekar & Shirley Lois Gundugollu

Department of Bioinformatics, Sri Ramachandra Faculty of Engineering and Technology,
Sri Ramachandra Institute of Higher Education and Research, Chennai-600 116, Tamil Nadu, India

Received 27 May 2026; *revised 30 September 2026*

Duchenne muscular dystrophy (DMD), is the most severe because mutation of dystrophin gene damages skeletal and cardiac muscles. Combining the results of RNA-Seq on 4 GEO data sets, we searched for DEGs related to DMD with the condition of logFC2.5 and $P< 0.05$. There were 10 hub genes in the PPI network: ACTN2, TNNT1, MYL4, TTN, TNNI2, MYH7, NEB, TPM1, CSRP3, and TCAP; it may play an important role in DMD. Molecular docking of ten phytochemical against these DMD hub proteins identified five complexes 4DB1-Mn-amppnp (-9.9 kcal/mol), 2X4F-BMN-195 (-9.5 kcal/mol), 2JT0-BMN-195 (-9.1 kcal/mol), 8EFH-NG25 (-9 kcal/mol), and 7KAA5-NG25 (-8.8 kcal/mol) with favorable binding energies. The top 5 complexes stability and dynamic behavior were checked further by RMSD, RMSF, H-bond, Radius of Gyration using MD for 100ns. Based on RNA-Seq, network analysis, Docking and MD simulation approach is providing the framework for further experimental investigation and insight in to prospective therapeutic molecules for the treatment of DMD.

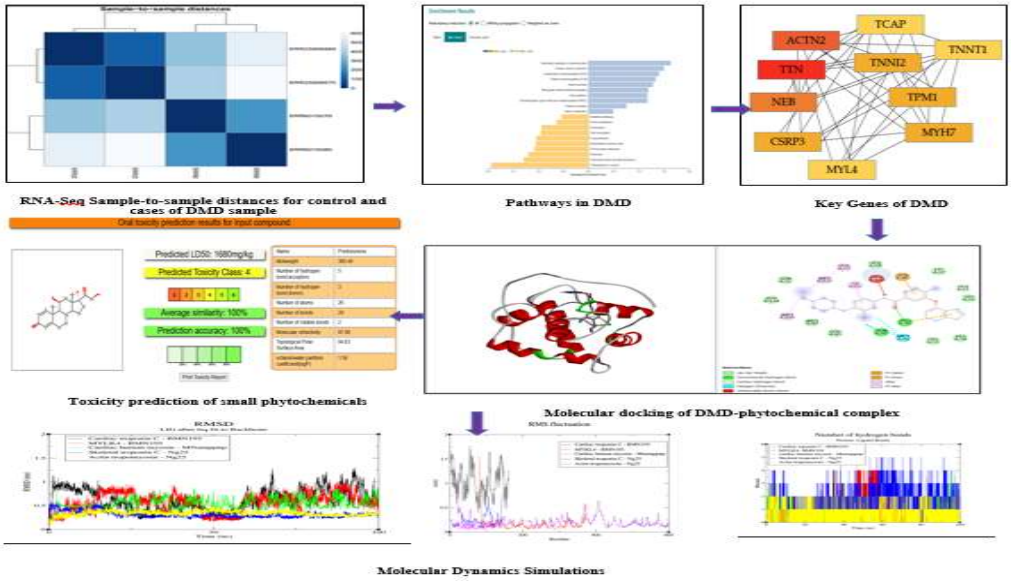
Keyword: Differentially expressed genes, Duchenne muscular dystrophy, Dystrophin gene, MD simulation, MYH7

Duchenne muscular dystrophy (DMD) is the most common and severe form of muscular dystrophy, which is an X-linked disorder caused by a mutation in the dystrophin gene, as initially documented over a century ago[1]. It is marked by progressive degeneration of skeletal muscles predominantly affecting males, with weakness typically manifesting in limb-girdle muscles around 5 years, followed by loss of independent ambulation between 8 and 12 years. Laboratory and clinical findings frequently include elevated creatine kinase levels, pseudohypertrophic calf muscles, reduced physical activity, and, in some cases, cognitive impairment[2]. Over 30 heterogenous disorders are classified as muscular dystrophies. Other relatively common muscular dystrophies include Becker muscular dystrophy, facioscapulohumeral muscular dystrophy, and myotonic dystrophy**,** whereas limb–girdle muscular dystrophy and congenital muscular dystrophy are themselves groups of several, usually very rare genetic disorders[3,4].

In DMD, dystrophin, a large 427 kDa cytoskeletal protein found on the cytoplasmic face of the sarcolemma in skeletal and cardiac myofibrils, is almost completely absent. The protein dystrophin functions by forming the dystrophin-glycoprotein complex, stabilising the muscle membrane against mechanical stress and enabling efficient transmission of contractile forces from the myofibrils to the surrounding extracellular matrix. When this protein is not present, the muscle membrane becomes fragile and susceptible to contraction-induced damage. The resulting membrane damage leads to increased intracellular Ca2+, which activates proteases that ultimately result in fiber death or necrosis[5,6]. Although regeneration of the myofibrils is seen after the damage for the first few years of life in the patients, this potential is lost as the patient advances in age. The necrotic fibers are replaced with fat and connective tissue, resulting in “pseudohypertrophy” of the calves[7].

Most of the treatments for MD have focused on DMD because of its frequency and severity. The pharmacological treatment of choice for DMD has historically been the corticosteroid prednisone, a potent anti-inflammatory that improves strength in mdx mice as well as in boys with DMD[8]. However, chronic exposure is accompanied by adverse effects,

*Correspondence:
E-mail: jinoblessy@sriramachandra.edu.in



Graphical abstract

including excessive weight gain, growth suppression, and osteoporosis. Deflazacort, an oxazoline derivative of prednisolone, an alternative with beneficial effects, causing less immunosuppression, reduced weight gain, reduced osteoporosis, thereby extending the period before loss of ambulation and scoliosis surgery in many patients[9]. For the subgroup of patients, approximately 10%, who harbor nonsense mutations in *DMD*, aminoglycoside antibiotics such as gentamicin have been evaluated as capable of partially suppressing premature stop codons and restoring production of functional dystrophin, with encouraging results in preclinical models and early human trials[10].

Pharmacological approaches, such as protease inhibitors, antioxidants, etc, although capable of slowing down the progress of DMD, can only do so up to a certain limit[11].

Therefore there is also a great deal of effort in the area of disease-modifying approaches, such as leveraging utrophin which is a dystrophin-like cytoskeletal protein that is ubiquitously expressed along the sarcolemma in foetal myogenesis and concentrated in the neuromuscular and myotendinous junctions of the adult[12]. In mdx mice, a higher level of utrophin expression could somewhat substitute the function of dystrophin, and simultaneously, knockout of dystrophin and utrophin lead to a dramatic progression of the myopathy, thus justifying up-regulation of utrophin as a therapeutic substitute[13,14]. Since utrophin is already synthesized endogenously in patients with DMD, it is unlikely that over-expression will trigger significant immune responses, and this strategy, combined with new gene-based approaches aimed at restoring dystrophin and medicines acting on downstream pathological processes, represents one of the most important concepts in the development of DMD therapies[15]. The present study applies RNA-seq analysis, molecular docking and dynamic simulations to identify the differentially expressed genes (DEGs) and interrogate their role in the key pathways related to DMD. These can serve as potential therapeutic targets. With all these bioinformatic workflows combined, an in-depth genetic analysis of DMD could be achieved and provide more insights into the disease mechanisms and new approaches toward therapeutic applications.

Materials and Methods

RNA Seq and pathway analysis

RNA-seq samples, including two cases (SRR25006670, SRR25006669) and two controls (SRR6015080, SRR6015076) were retrieved from NCBI GEO Dataset[16]. These samples were imported into the Galaxy server using the tool Faster Download and Extract Reads in FASTQ format from NCBI SRA[17]. Following this, the data quality of the files was assessed using the FastQC tool[18]. The files are subjected to the tool trimmomatic for adaptor removal and low-quality base trimming. Combined the outputs to a single report using MultiQC. Sequence aligned to reference human genome hg38 using HISAT2 tool,

output in the form of BAM file[19]. Quantified gene expression levels using feature counts. Used BAM file of mapped sequences and GTF file of gene annotations as input, output a mapped gene counts matrix. Performed differential expressed genes (DEGs) between case and control group using DESeq2: a statistical tool that uses negative binomial distribution model to estimates mean expression and variance of count data. Perform gene set enrichment analysis (GSEA) to the DEGs using WebGestalt and identify the significantly enriched categories, such asGOterms, KEGG pathways. The results illustrated possible biological pathways involved in DMD[20].

Protein-protein interaction network

The PPI interaction network was developed based on the STRING database[21] with the provided up-regulated and down-regulated genes. The network was built by uploading genes and download interaction map in the format of TSV, and import into Cytoscape. Cytoscape software was used to visualize relationships among genotypes, biological systems, and gene expression patterns[22]. The CytoHubba plugin was used to identify the top 10 hub genes based on the centrality measures and the distance between the genes in the network[23]. Using CytoHubba the top 10 hub genes were identified: ACTN2, TNNT1, MYL4, TTN, TNNI2, MYH7, NEB, TPM1, CSRP3 and TCAP. The proteins that are encoded by the hub genes were identified, and their PDB IDs are 5A38, 2JT0, 2X4F, 1NCT, 7KAA, 4DB1, 1ARK, 8EFH, 2CQ7 and 2F8V. Molecular docking was performed for DMD proteins against phytochemicals (Table 1).

Molecular Docking

The three-dimensional structure of the DMD target proteins 5A38 (A chain), 2JT0 (A chain), 2X4F (A chain), 1NCT (A chain), 7KAA (A chain), 4DB1 (A chain), 1ARK (A chain), 8EFH (B chain), 2CQ7 (A chain) and 2F8V (C chain) was retrieved from the Protein Data Bank (PDB)[24]. Ligand binding sites of the selected crystallographic structures were identified using PDBsum and the crystallographically resolved co-crystallized ligands for each structure[25]. Where a well-defined ligand-associated binding pocket was present in the corresponding crystal structure, a set of residues surrounding the co-crystallized ligand was used to define the docking region and site-specific docking was performed. Using this criteria, the proteins 2X4F, 7KAA, 4DB1 and 8EFH each displayed a well-defined ligand associated binding pocket and site-specific docking was performed. Conversely, the proteins 5A38, 2JT0, 1NCT, 1ARK, 2CQ7 and 2F8V did not have a crystallographically characterized ligand associated binding pocket present in the structural data available; therefore, the entire protein structure was used as the docking search space and blind docking was performed. This is consistent with previous blind-docking strategies where the presence of a binding pocket cannot be predicted and the entire protein structure is used in the docking search space[26,27]. Protein preparation was done using UCSF Chimaera, involving addition of polar hydrogen, charge assignment, removal of crystallographic water, protein chains and non-standard amino acids[28]. The molecular docking was performed using PyRx with AutoDock Vina[29]. A receptor grid box was defined around each protein's active site to constrain the ligand search space, with grid size and spacing adjusted to match the binding pocket geometry. In total, ten phytochemicals, Prednisolone, Prednisone, Mn-amppnp, Catalpol, Deflazacort, Vamorolone, Desacetyl deflazacort, Telethonin, Bmn-195 and Ng25, were selected based on previous studies on DMD and retrieved from the PubChem database[30-35].

Table 1 — Top 10 hub genes derived through PPI from Cytoscape using CytoHubba plugin

Hub Genes	Protein name	Protein symbol and chain	Active site
ACTN2	Alpha-actinin-2	5A38 A	No
TNNT1	Cardiac troponin-C	2JT0 A	No
MYL4	Myosin light chain 4	2X4F A	yes
TTN	Titin	1NCT A	No
TNNI2	Troponin I, fast skeletal muscle	7KAA A	Yes
MYH7	Myosin-7	4DB1 A	Yes
NEB	Nebulin	1ARK A	No
TPM1	Tropomyosin alpha-1 chain	8EFH B	Yes
CSRP3	Cysteine and glycine-rich protein 3	2CQ7 A	No
TCAP	Telethonin	2F8V C	No

Table 2 — Toxicity prediction of small molecules using Protox 2.0

Ligand name	Predicted LD50	Predicted Toxicity Class	Average similarity	Prediction accuracy
Prednisolone	1680 mg/kg	4	100%	100%
Prednisone	1680 mg/kg	4	84.98%	70.97%
Mn-amppnp	1190 mg/kg	4	100%	100%
Catalpol	2000 mg/kg	4	65.24%	68.07%
Deflazacort	5200 mg/kg	6	100%	100%
Vamorolone	4000 mg/kg	5	87.99%	70.97%
Desacetyl deflazacort	1190 mg/kg	4	100%	100%
Telethonin	1190 mg/kg	4	100%	100%
BMN-195	1600 mg/kg	4	65.35%	68.07%
NG25	440 mg/kg	4	51.69%	67.38%

The toxicity analysis was conducted using Protox 3.0 to predict the toxicity of the ten phytochemicals, and a toxicity report was prepared[36] (Table 2). The ligand preparation was done using energy minimization in PyRx and saved in PDBQT[37]. The prepared ligands were docked against the ten DMD target proteins, and the conformations were ranked based on the docking score and the relative mode. Molecular docking is a widely used technique in structure-based drug design because it effectively predicts how small molecule ligands bind and conform to their target binding sites[38].

Molecular dynamics simulations

The GROMACS software package is used to perform 100ns simulations for the Myosin-7-Mn-amppnp complex, Myosin light chain 4-Bmn-195 complex, Cardiac troponin-C-Bmn-195 complex, Tropomyosin-Ng25 complex, and Troponin I-Ng25 complex[39]. The force field parameters CHARMM36 are used for proteins and CGenFF for phytochemicals. The complex system was solvated by a dodecahedron box employing periodic boundary conditions, with an equal number of water molecules added from the TIP3P water model. The steepest descent algorithm of energy minimization was conducted to reduce unwanted steric conflicts in the Myosin-7-Mn-amppnp complex, Myosin light chain 4-Bmn-195 complex, Cardiac troponin-C-Bmn-195 complex, Tropomyosin-Ng25 complex, and Troponin I-Ng25 complex. The whole system was neutralized by introducing counter ions NA+ CL-. Subsequently, the system underwent relaxation under NVT conditions (constant number of atoms, volume, and temperature) at 300 K employing a Langevin thermostat, followed by NPT conditions (constant number of atoms, pressure, and temperature) utilizing a Berendsen barostat. Trajectory data were analyzed for RMSF, RMSD, hydrogen bond, and Radius of Gyration. Protein-ligand complex backbone RMSD, was calculated to track the deviation of structure and stability of ligand position throughout the trajectory. Root mean square fluctuation (RMSF) was calculated to determine the atomic fluctuations of residues in protein which would reveal differential conformational fluctuations. Radius of gyration (Rg) was analysed to track the overall compactness of the protein structure during the simulation. Protein ligand hydrogen-bond was also calculated which would demonstrate the consistency of non-covalent interaction between ligand and residue within the binding-site. The overall stability analysis of the complexes was carried out by the combined study of RMSD, RMSF, radius of gyration, and the pattern of stability of hydrogen-bond interactions over the 100 ns MD simulation time. The changes in parameters over time were studied to analyse the stability and stability pattern of the complexes. The 100 ns MD simulation timescale was also justified by previous protein ligand MD simulation work in which 100 ns trajectories were used to measure the stability and conformational changes of the complex using RMSD, RMSF, and radius of gyration and hydrogen-bond interactions[40,41].

Results

DMD datasets and Differentially expressed genes (DEGs)

The DESeq analysis on the four DMD samples (cases: SRR25006670, SRR25006669, and controls: SRR6015080, SRR6015076). Cluster analysis was done using a heatmap and principal component plot to analyze the expression patterns and clustering among the samples (Figs. 1 & 2). In the PCA plot, PC1 showed 93% variance, indicating a clear majority in the separation of cases and controls. The heat maps were used for assessing sample-to-sample distance across the cases and controls, where the dark blue regions indicate a shorter distance between the samples. The dispersion estimates plot revealed the variance in gene expression (Fig. 3). The differentially expressed genes were identified with a log2 fold change ≥ 2.5 and a significance threshold of $P< 0.05$. Genes both up-regulated and down-regulated were found in the group of DEGs. Based on the formula 25 significantly up-regulated and down-regulated genes were then retrieved.

Gene enrichment analysis using WebGestalt revealed altered biological pathways based on the normalized enrichment score. Several cardiac and muscle-related pathways were significantly up-regulated (FDR > 0.05), like the adrenergic signaling in cardiomyocytes, cardiac muscle contraction, hypertrophic cardiomyopathy (HCM), dilated cardiomyopathy (DCM), and arrhythmogenic right ventricular cardiomyopathy (ARVC), etc. The significantly down-regulated pathways (FDR > 0.05) included impairments in metabolic pathways, purine metabolism, endocytosis, biosynthesis in amino acids, and ECM-receptor interactions (Fig. 4).

Protein-Protein Interaction

The protein-protein interactions were constructed and analyzed between the DEGs from the STRING database, where the network was in the form of nodes and edges, where a node represents a protein from a single DEG and edge represents the interaction between the proteins (Fig. 5). Cytoscape was used to visualize the PPI network and conduct further analysis by altering the topological parameters. This led to the identification of the central hub genes, and it was further validated by using the MCODE plugin. The cytohubba plugin was used to infer the top 10 hub genes, and the analysis revealed the top 10 hub genes as ACTN2, TNNT1, MYL4, TTN, TNNI2, MYH7, NEB, TPM1, CSRP3, and TCAP (Table 1 & Fig. 6).

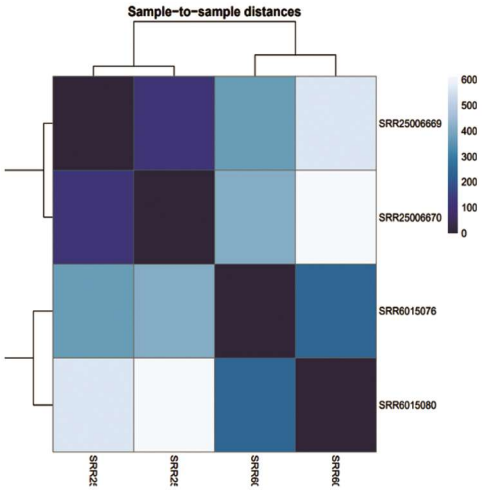


Fig. 1 — Sample-to-sample distances for control and cases of DMD sample

Phytochemical toxicity analysis

The toxicity of the molecules (Prednisolone, Prednisone, Mn-amppnp, Catalpol, Deflazacort, Vamorolone, Desacetyl deflazacort, Telethonin, Bmn-195 and Ng25) was assessed using Protox 3.0, where the compounds are classified into six classes according to their toxicity levels: class 1 (LD_{50} ≤ 5 mg/kg), class 2 (5–50 mg/kg), class 3 (50–300 mg/kg), class 4 (300–2000 mg/kg), class 5 (2000–5000 mg/kg) and class 6 (LD_{50} > 5000 mg/kg; non-toxic). The predicted LD_{50} values for the tested ligands ranged from 440 mg/kg to 5200 mg/kg, placing most molecules within classes 4–6 (Table 2). Prednisolone, Prednisone, Mn-amppnp, Catalpol, Desacetyl deflazacort, Telethonin, BMN-195 and NG25 were allocated to class 4, reflecting moderate toxicity but tolerable margins for further structural optimization. Deflazacort was assigned to class 6 with an LD_{50} of 5200 mg/kg, indicating a non-toxic profile, whereas Vamorolone was categorized as class 5 with an LD_{50} of 4000 mg/kg. Prednisolone, Prednisone, Mn-amppnp, Catalpol, Desacetyl deflazacort, Telethonin, BMN-195 and NG25 were allocated to class 4, reflecting moderate toxicity, and are permissible for further lead optimization and structural refinement under acceptable safety margins, and further research is required.

Molecular docking

Molecular docking was done between the proteins encoded by the top ten hub proteins, which are 5A38, 2JT0, 2X4F, 1NCT, 7KAA, 4DB1, 1ARK, 8EFH, 2CQ7, 2F8V, against the ten molecules Prednisolone, Prednisone, Mn-amppnp, Catalpol, Deflazacort, Vamorolone, Desacetyl deflazacort, Telethonin, Bmn-

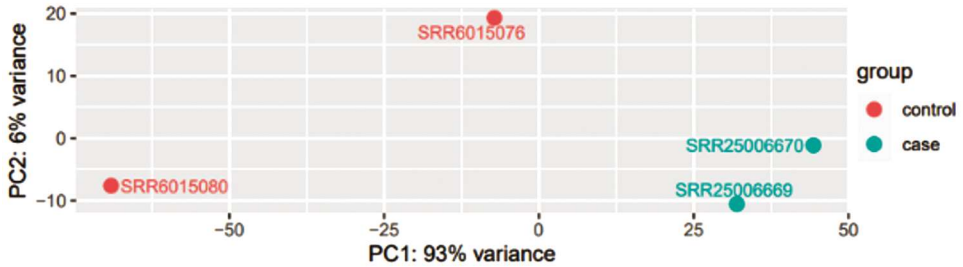


Fig. 2 — Principal component plot of the control vs case for up-regulated and down-regulated genes for the DMD sample

195 and Ng25. Five complexes with the favourable binding affinities, which is the lowest binding free energy scores, indicating stable complex formation were found to be: 4DB1-Mn-amppnp (Fig. 7), 2X4F-BMN-195 (Fig. 8), 2JT0-BMN-195 (Fig. 9), 8EFH-NG25 (Fig. 10), and 7KAA5-NG25 (Fig. 11) with scores (Kcal/mol) of -9.9, -9.5, -9.1, -9, and -8.8 (Tables 3 & 4).

Molecular dynamics simulation

The MD simulation provides information about the internal motions of the protein–ligand complex in a time-dependent manner. This protein-ligand complex was treated with solvent under flexible conditions. To validate the binding conformations and stability of the complexes 4DB1-Mn-amppnp (-9.9 kcal/mol), 2X4F-BMN-195 (-9.5 kcal/mol), 2JT0-BMN-195 (-9.1 kcal/mol), 8EFH-NG25 (-9 kcal/mol), and 7KAA5-NG25 (-8.8 kcal/mol), MD simulation was performed using the GROMACS. The conformations of the complexes obtained after MD simulation were found to be more stable throughout the 100ns of the simulation. The conformational stability of the complexes 4DB1-Mn-amppnp, 2X4F-BMN-195, 2JT0-BMN-195, 8EFH-NG25, and 7KAA5-NG25 was evaluated using Root Mean Square Deviation (RMSD), Root Mean Square Fluctuation (RMSF), Hydrogen bond analysis, and Radius of Gyration (Rg).

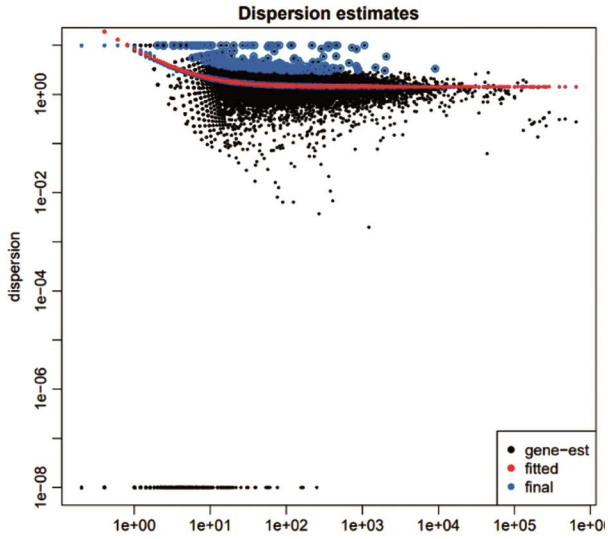


Fig. 3 — Dispersion estimates of gene-set, fitted, and final of DMD

RMSD

The RMSD plots various protein-ligand complexes over a 100 ns simulation period, illustrating the stability of the complexes over that time period. More

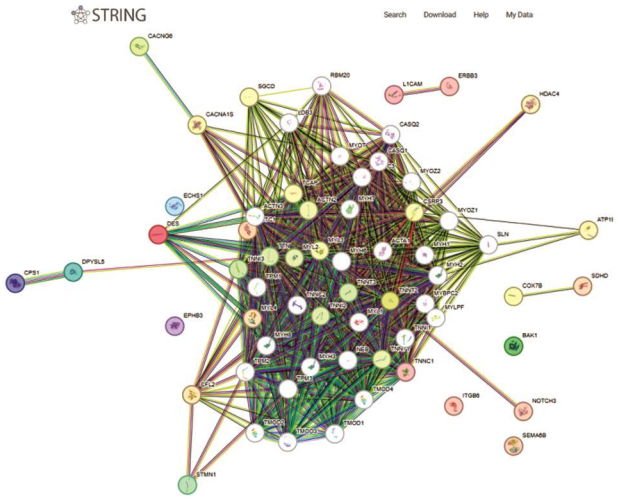


Fig. 5 — Protein-protein DMD interaction derived using StringDB

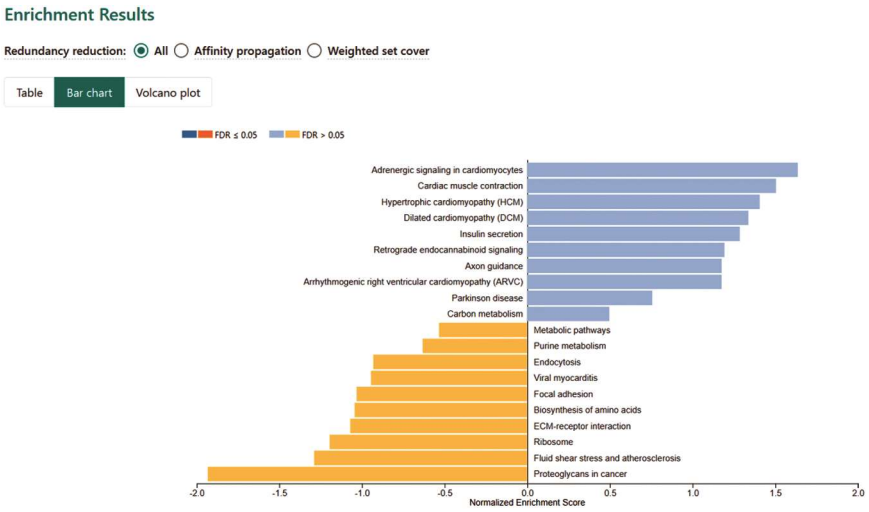


Fig. 4 — Enrichment analysis showing the altered biological pathways in DMD using WebGestalt

stable complex will have lower RMSD values. Cardiac troponin-C (2JT0)-BMN-195 complex (black) does not show any stability as RMSD values are fluctuating. Myosin light chain 4 (2X4F)-BMN-195 complex (red) is of moderate stability as RMSD values is less but in some cases it is showing the peak values. Cardiac human myosin (4DB1)-Mn-amppnp complex (green) is having low RMSD value and thus it seems to be stable enough. Skeletal troponin C (7KAA5)-NG25 complex (blue) is having the high stability with less RMSD fluctuation. Cardiac actin-tropomyosin (8EFH)-NG25 complex (yellow) show a lot of stability as its RMSD value is quite less and is steady in value (Fig. 12).

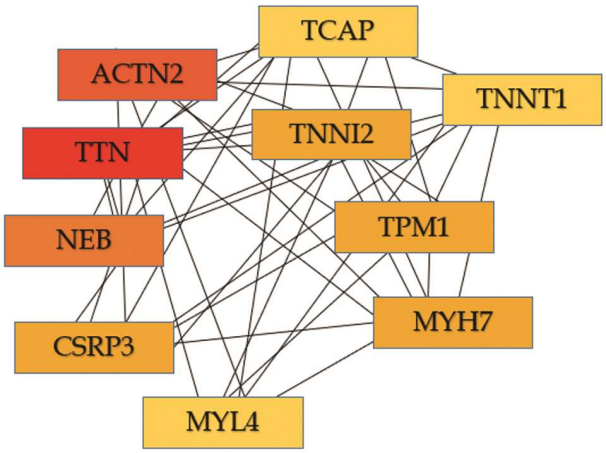


Fig. 6 — Top 10 hub genes of DMD identified using Cytoscape

RMSF

RMSF has also been plotted for each residue in all the protein-ligand complex. X-axis is represented by residue and Y-axis is represented by RMSF in nm. The high fluctuation corresponds to flexibility or instability in respective residues. In Cardiac troponin-C (2JT0)-BMN-195 complex (black) high fluctuations are found at the initial residues hence these residues are much more flexible. In Myosin light chain 4 (2X4F)-BMN-195 complex (red) moderate fluctuations can be seen and some peaks at some point. Cardiac human myosin (4DB1)-Mn-amppnp complex (magenta) residues are quite stable with slight increased fluctuations at some points. Skeletal troponin C (7KAA5)-NG25 complex (blue) residue has low fluctuation with stable residues. Actin-tropomyosin (8EFH)-NG25 complex (purple) has moderate fluctuation in throughout the simulation (Fig. 13).

Hydrogen bonds

The graph for the number of hydrogen bonds formed between the protein and ligand at various points of time (X-axis: Time in nanoseconds, Y-axis: number of hydrogen bonds) illustrates the stability and nature of the complex. More is the number of hydrogen bonds, stronger is the binding and vice-versa. In the graph, it is evident that the number of hydrogen bonds formed in Cardiac troponin-C (2JT0)-BMN-195 complex (black)

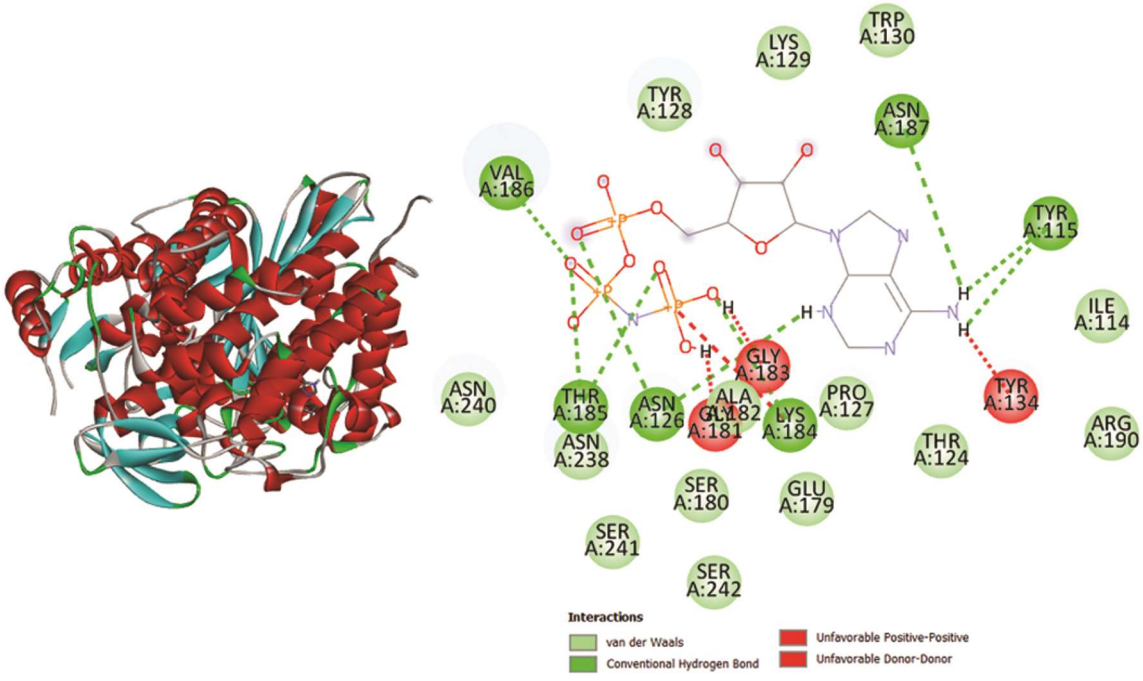


Fig. 7 — Docking results of 4DB1-Mn-amppnp complex

is oscillating indicating not a very strong bond. Similarly Myosin light chain 4 (2X4F)-BMN-195 complex (red) shows a considerable number of hydrogen bonds with moderate stability. Cardiac human myosin (4DB1)-Mn-amppnp complex (green) shows the highest and stable number of hydrogen bonds. Skeletal troponin C (7KAA5)-NG25 complex (blue) also has a consistently higher number of hydrogen bonds indicating stable interaction. Actin-tropomyosin (8EFH)-NG25 complex (yellow) shows less number of hydrogen bonds and a less stable complex (Fig. 14).

Radius of gyration

The graph shown here, plots the Radius of Gyration, the measure of structural compactness of a protein. If the Radius of gyration is high, it means that structure is a lot more spread out, if the radius is small it is more compact. Cardiac troponin-C (2JT0)-BMN-195 complex (black), has a steadily low radius of gyration indicating that it is structure is compact and

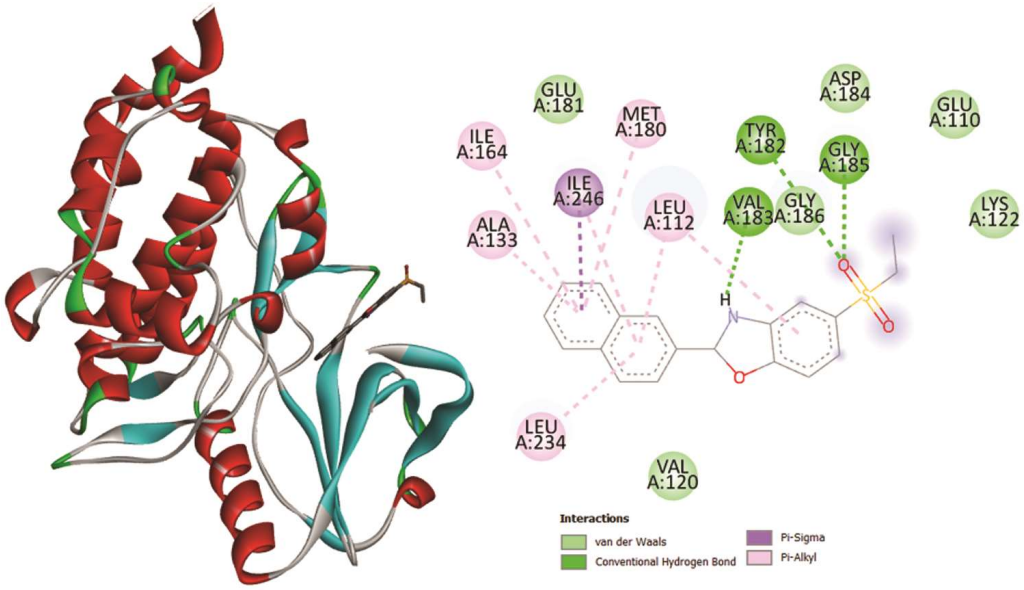


Fig. 8 — Docking results of 2X4F-BMN-195 complex

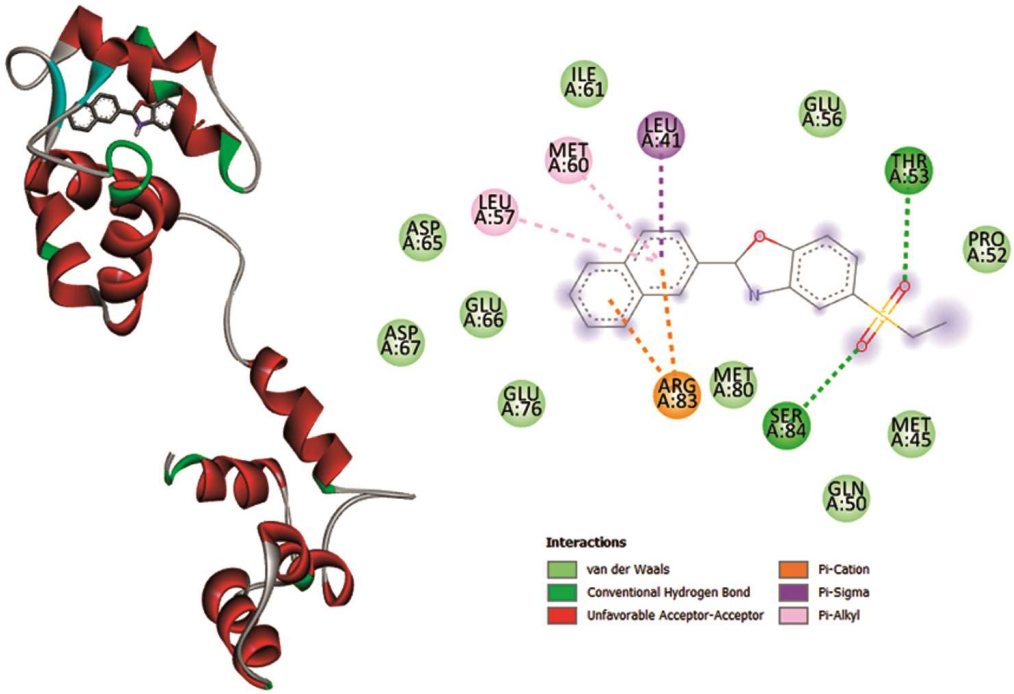


Fig. 9 — Docking results of 2JT0-BMN-195 complex

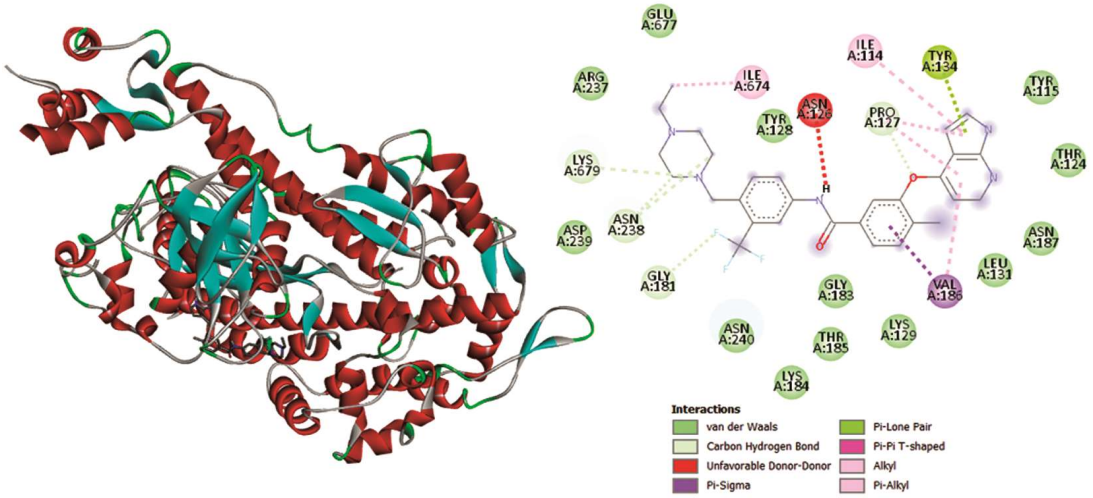


Fig. 10 — Docking results of 8EFH-NG25 complex

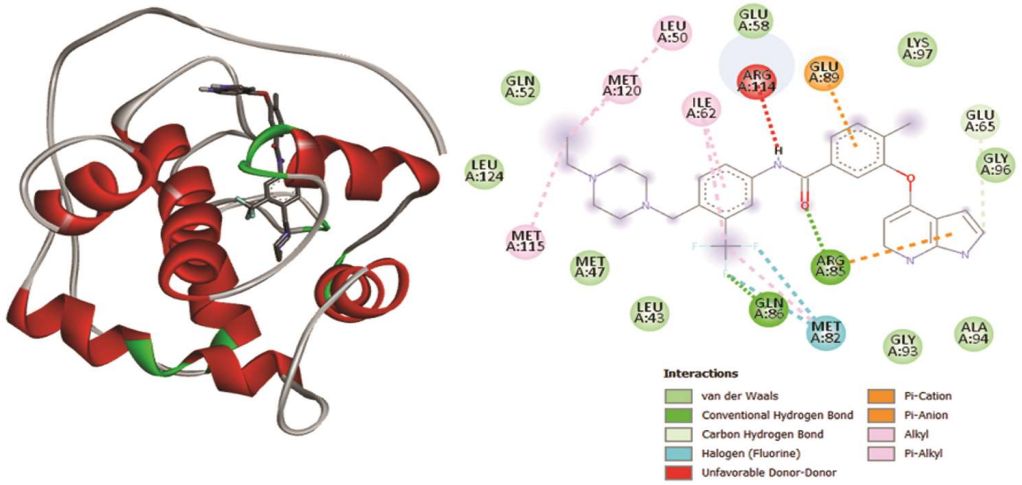


Fig. 11 — Docking results of 7KAA5-NG25 complex

Table 3 — Molecular docking of DMD target proteins with phytochemicals.

Protein and compound	1ark	1nct	2cq7	2f8v	2jt0	2x4f	4db1	5a38	7kaa	8efh
Prednisolone	-5.3	-6.2	-6.6	-5.3	-6.9	-6.5	-7.5	-6.7	-6.3	-7.1
Prednisone	-5.7	-6.3	-6.4	-5.6	-7.1	-6.8	-8	-7.1	-6.7	-7.2
Mn-amppnp	-5.2	-6.3	-6.6	-4.6	-7.5	-7.7	-9.9	-6.1	-7.1	-8
Catalpol	-5.4	-5.6	-5.6	-4.4	-6.3	-7.7	-8.4	-6.2	-6.6	-6.8
Deflazacort	-6.1	-6.2	-7	-5.7	-7.6	-6.9	-8.4	-7.1	-6.1	-7.8
Vamorolone	-5.9	-6.2	-6	-5	-7.5	-6.6	-7.6	-6.8	-6.6	-7.6
Desacetyl deflazacort	-6	-6.2	-6.5	-5.3	-7.4	-7	-8.5	-6.9	-5.9	-7.6
Telethonin	-5.9	-6.8	-7.3	-6	-8.8	-6.3	-8.5	-7.4	-7	-7.3
BMN-195	-6.3	-6.5	-6.7	-5.6	-9.1	-9.5	-8.7	-7.1	-7.5	-8.3
NG25	-6.4	-7	-6.7	-6.2	-7.8	-9.2	-8.8	-7.9	-8.8	-9

does not change.The Myosin Light Chain 4 (2X4F)-BMN-195 complex (red) has a low radius of gyration also, indicating that it too is quite compact. The Cardiac human myosin (4DB1)-Mn-amppnp complex (green) has a much higher radius and fluctuates a lot, indicating a less compact, more dynamic structure. The Skeletal troponin C (7KAA5)-NG25 complex (blue) and the Actin-tropomyosin (8EFH)-NG25 complex (yellow) both have a similar moderate radius of gyration (Fig. 15).

Table 4 — Top 5 DMD target proteins – phytochemicals complexes

Protein	Compound	Score kcal/mol
4db1	Mn-amppnp	-9.9
2x4f	BMN-195	-9.5
2jt0	BMN-195	-9.1
8efh	NG25	-9
7kaa5	NG25	-8.8

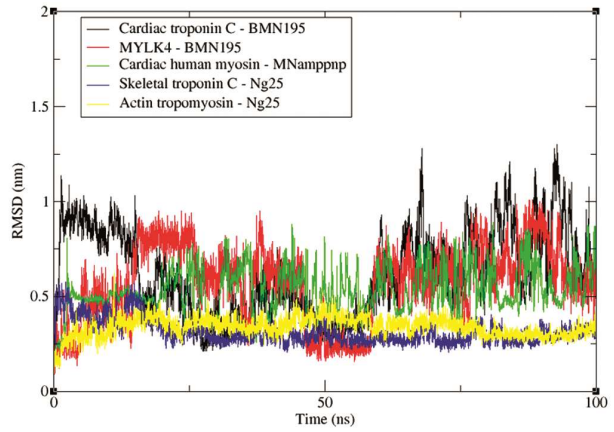


Fig. 12 — RMSD profile for the DMD-Phytochemical complex

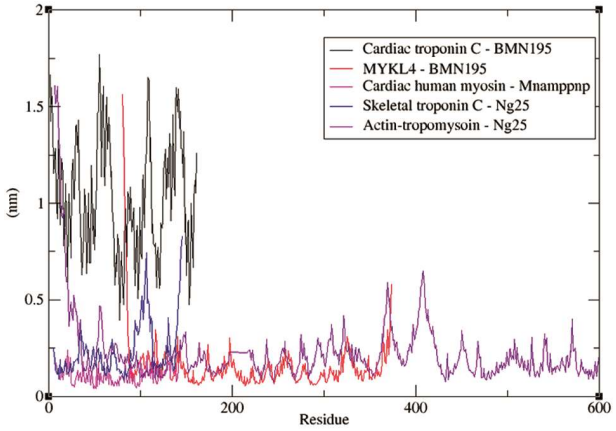


Fig. 13 — RMSF profile for the DMD-Phytochemical complex

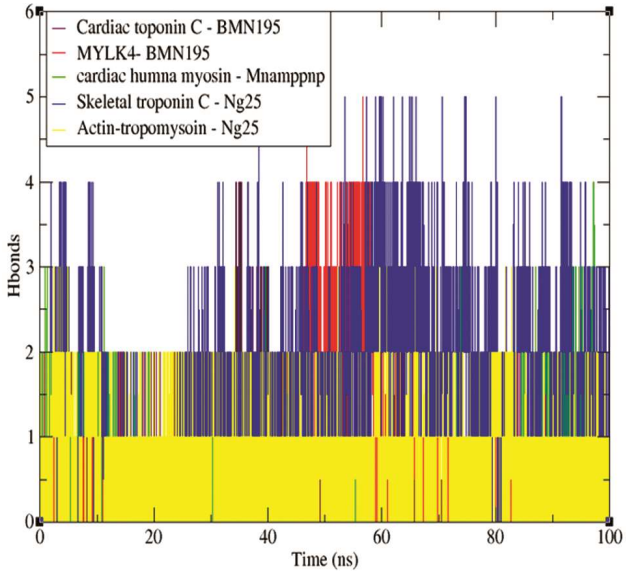


Fig. 14 — Hydrogen Bonds profile for the DMD-Phytochemical complex

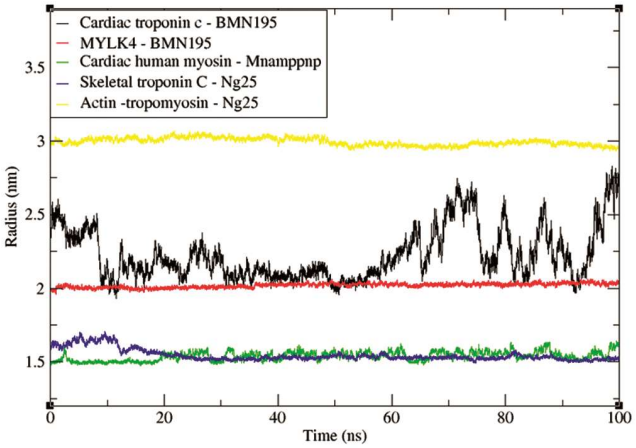


Fig. 15 — Radius of gyration (Rg) profile for the DMD-Phytochemical complex

Discussion

An integrative analysis of RNA-seq was implemented in this study, to understand molecular mechanism, therapeutic target for DMD, we integrated genomic information along with the analysis of RNA seq, molecular docking and molecular dynamics simulation. We selected DMD sample(cases: SRR25006670, SRR25006669 and controls: SRR6015080, SRR6015076) to analyze the RNA seq. PCA revealed the separate groups for the case and control sample, as expected, large transcriptional change in the dystrophic muscle was demonstrated. The high variance explanation of PC1(93%) from PCA indicated high robust of the gene expression change. The finding was consistent with earlier microarray and RNA seq based study. From the heatmaps, both case and control group demonstrated tightly clustered[42].

DEG analysis highlighted genes involved in cardiac and muscle-specific signaling pathways such as adrenergic signaling, cardiac muscle contraction and several cardiomyopathies were all up-regulated, indicating they are highly dysregulated in the DMD muscle and therefore it follows the disease phenotype in particular progressive cardiac dysfunction[43,44]. Down-regulated genes were identified as being involved in metabolic and biosynthetic pathways. The metabolic changes appear to be extensive and as

secondary downstream effects from the loss of dystrophin. Similar patterns of pathway enrichment have also been identified in separate studies, indicating a reproducible molecular signature[45].

Central regulators identified from PPI network construction and topology based hub analysis include ACTN2, TNNT1, MYL4, TTN, TNNI2, MYH7, NEB, TPM1, CSRP3 and TCAP, where many have already been reported to function in muscle contractility and sarcomere[46,47]. They are among the most enriched genes reported in network-centric DMD studies which further support their role in this disease and highlight their function as biomarkers and/or therapeutic targets due to their centrality and involvement with numerous genes[48].

Several small molecules such as Mn-amppnp, BMN-195, and NG25 were recognized for their capacity to create strong bonds with the top hub proteins, as demonstrated by the strong binding affinity calculations and supplemented by moderate toxicity profiles. In turn, these results may provide a framework to establish structure-based drug design strategies, along with a basis for *in silico* selection of candidates for repositioning drugs. The long term stability and dynamics of the identified protein-ligand complexes was confirmed through MD simulations and confirmed by the RMSD, RMSF, hydrogen bonds, and radius of gyration values that indicated some protein-ligand complexes remained strong over the 100 ns time period. Highly stable protein-ligand interactions such as that formed by 4DB1-Mn-amppnp and 7KAA5-NG25 might represent viable targets for future drug candidates targeting DMD[49]. This multi-faceted approach of DEG analysis, pathway enrichment, network generation and molecular modeling approaches provided a comprehensive view of the DMD molecular pathology layer by layer. The hub genes (ACTN2, TNNT1, MYL4, TTN, TNNI2, MYH7, NEB, TPM1, CSRP3, TCAP) along with prioritized binding sites represent rational candidates for future therapy development.

However, PCA and MM/GBSA analyses have not been included in the current study, and this should be considered as a limitation of the current MD study. Despite its lack of significance in the current study, PCA could have yielded further collective conformational information, and MM/GBSA could have yielded shallow energetic estimates of the protein-ligand interactions, although the current study prioritized trajectory-based conformational stability and interaction studies of the systems using RMSD, RMSF, radius of gyration, and hydrogen-bond interactions. These could be assessed in future studies to further characterize the nature of the complex.

Conclusion

Using next generation sequencing for this research enabled the detection of a diverse array of DMD gene mutations to support accurate clinical diagnosis. It may therefore be argued that integration of transcriptomics with network biology and structure-based design may aid in the discovery of effective DMD therapeutics. Based upon our DEGs and pathways enrichments, DMD represents a condition with significantly dysregulated signaling within both the cardiac and skeletal muscles, in addition to disordered metabolism and extracellular matrix pathways. Analyzing our predicted PPI network and identifying the respective hub genes points toward several sarcomeric and skeletal muscles regulatory proteins as potentially valuable drug targets in DMD (ACTN2, TNNT1, MYL4, TTN, TNNI2, MYH7, NEB, TPM1, CSRP3 and TCAP). Toxicity predictions suggest that the majority of screened molecules exhibit generally safe profiles and are suitable for subsequent modification. Furthermore, it is recommended that a more in depth analysis of our five most favorably binding protein-ligand complexes, selected through docking of all molecules to relevant protein targets with favorable binding energy, is performed with molecular dynamics simulation of each complexes over 100ns to determine structural stability and interaction. Majority of the predicted most favorable complexes in MD simulations show stable conformations and hydrogen bonds, acceptable compactness and therefore, we predict Mn-amppnp, BMN-195 and NG25 as potential leads to effectively target the essential proteins associated with DMD. These results should be confirmed through experimental investigation of the candidate lead molecules in suitable *in vitro* and *in vivo* model systems.

Acknowledgement

We thank the Department of Bioinformatics, Sri Ramachandra Faculty of Engineering and Technology, Sri Ramachandra Institute of Higher Education and Research for their computational facilities.

Conflict of interest

All authors declare no conflict of interest.

References

1 Duan D, Goemans N, Takeda S, Mercuri E & Aartsma-Rus A, Duchenne muscular dystrophy. *Nat Rev Dis Primers*, 7 (2021) 13.

2 Mercuri E, Bonnemann CG & Muntoni F, Muscular dystrophies. Lancet, 394 (2019) 2025.

3 Bozzi M, Diagnosis, pathogenesis and treatment of muscular dystrophy. *Biomedicines*, 13(8) (2025) 1820.

4 Hoang T & Dowdy RAE, A review of muscular dystrophies. *Anesth Prog*, 71 (2024) 44.

5 Dowling P, Gargan S, Murphy S, Zweyer M, Sabir H, Swandulla D & Ohlendieck K, The dystrophin node as integrator of cytoskeletal organization, lateral force transmission, fiber stability and cellular signaling in skeletal muscle. *Proteomes*, 9 (2021) 9.

6 Wilson DGS, Tinker A & Iskratsch T, The role of the dystrophin glycoprotein complex in muscle cell mechanotransduction. *Commun Biol*, 5 (2022) 1022.

7 Sarvutiene J, Ramanavicius A, Ramanavicius S & Prentice U, Advances in Duchenne muscular dystrophy: Diagnostic techniques and dystrophin domain insights. *Int J Mol Sci*, 26 (2025) 3579.

8 Komaki H, Duchenne muscular dystrophy: Evolving therapeutic strategies and multidimensional evaluation approaches. *Brain Dev*, 47 (2025) 104397.

9 Malaga M, Rodriguez-Calienes A, Chavez-Ecos FA, Huerta-Rosario A, Alvarado-Gamarra G, Cabanillas-Lazo M, Moran-Ballon P, Velásquez-Rimachi V, Martinez-Esteban P & Alva-Diaz C, Clinical practice guidelines for the diagnosis and management of Duchenne muscular dystrophy: A scoping review. *Front Neurol*, 14 (2024) 1260610.

10 Bushby K, Finkel R, Wong B, Barohn R, Campbell C, Comi GP, Connolly AM, Day JW, Flanigan KM, Goemans N, Jones KJ, Mercuri E, Quinlivan R, Renfroe JB, Russman B, Ryan MM, Tulinius M, Voit T, Moore SA, Lee Sweeney H, Abresch RT, Coleman KL, Eagle M, Florence J, Gappmaier E, Glanzman AM, Henricson E, Barth J, Elfring GL, Reha A, Spiegel RJ, O'donnell MW, Peltz SW & Mcdonald CM, Ataluren treatment of patients with nonsense mutation dystrophinopathy. *Muscle Nerve*, 50 (2014) 477.

11 Guiraud S & Davies KE, Pharmacological advances for treatment in Duchenne muscular dystrophy. *Curr Opin Pharmacol*, 34 (2017) 36.

12 Szwec S, Kapłucha Z, Chamberlain JS & Konieczny P, Dystrophin- and utrophin-based therapeutic approaches for treatment of Duchenne muscular dystrophy: A comparative review. *BioDrugs*, 38 (2024) 95.

13 Pons F, Robert A, Marini JF & Léger JJ, Does utrophin expression in muscles of mdx mice during postnatal development functionally compensate for dystrophin deficiency? *J Neurol Sci*, 122 (1994) 162.

14 Baban D & Davies KE, Microarray analysis of mdx mice expressing high levels of utrophin: Therapeutic implications for dystrophin deficiency. *Neuromuscul Disord*, 18 (2008) 239.

15 Wu R, Song Y, Wu S & Chen Y, Promising therapeutic approaches of utrophin replacing dystrophin in the treatment of Duchenne muscular dystrophy. *Fundam Res*, 2 (2022) 885.

16 Clough E, Barrett T, Wilhite SE, Ledoux P, Evangelista C, Kim IF, Tomashevsky M, Marshall KA, Phillippy KH, Sherman PM & Lee H, NCBI GEO: Archive for gene expression and epigenomics data sets: 23-year update. *Nucleic Acids Res*, 52 (2024) D138.

17 Leinonen R, Sugawara H, Shumway M & International Nucleotide Sequence Database Collaboration, The sequence read archive. *Nucleic Acids Res*, 39 (2010) D19.

18 González-Bodí S & Huerta-Cepas J, Reference-based gene expression analysis using Galaxy Server. *In Neural Stem Cells: Methods and Protocols*, (2025) 309.

19 Reddy RRS & Ramanujam MV, High throughput sequencing-based approaches for gene expression analysis. *In Gene Expression Analysis: Methods and Protocols*, (2018) 299.

20 Elizarraras JM, Liao Y, Shi Z, Zhu Q, Pico AR & Zhang B, WebGestalt 2024: Faster gene set analysis and new support for metabolomics and multi-omics. *Nucleic Acids Res*, 52 (2024) W415.

21 Szklarczyk D, Nastou K, Koutrouli M, Kirsch R, Mehryary F, Hachilif R, Hu D, Peluso ME, Huang Q, Fang T & Doncheva NT, The STRING database in 2025: Protein networks with directionality of regulation. *Nucleic Acids Res*, 53 (2025) D730.

22 Franz M, Lopes CT, Huck G, Dong Y, Sumer O & Bader GD, Cytoscape.js: A graph theory library for visualisation and analysis. *Bioinformatics*, 32 (2016) 309.

23 Chin CH, Chen SH, Wu HH, Ho CW, Ko MT & Lin CY, cytoHubba: Identifying hub objects and sub-networks from complex interactome. *BMC Syst Biol*, 8 (2014) S11.

24 Vallat B, Rose Y, Piehl DW, Duarte JM, Bittrich S, Bi C, Segura J, Zalevsky A, Sekharan MR, Webb BM & Bhikadiya C, RCSB Protein Data Bank: Delivering integrative structures alongside experimental structures and computed structure models. *Nucleic Acids Res*, 54 (2026) D489.

25 Laskowski RA, Jabłońska J, Pravda L, Vařeková RS & Thornton JM, 2018. PDBsum: Structural summaries of PDB entries. *Protein sci*, 27 (2018) 129.

26 Hetényi C & van der Spoel D, Efficient docking of peptides to proteins without prior knowledge of the binding site. *Protein Sci,* 11 (2002) 1729.

27 Ghersi D & Sanchez R, Improving accuracy and efficiency of blind protein-ligand docking by focusing on predicted binding sites. *Proteins*, 74 (2009) 417.

28 Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC & Ferrin TE, UCSF Chimera—A visualization system for exploratory research and analysis. *J Comput Chem*, 25 (2004) 1605.

29 Pawar RP & Rohane SH, Role of autodock vina in PyRx molecular docking. *Asian J Res Chem*, 14(2) (2021) 132.

30 Gulick AM, Bauer CB, Thoden JB & Rayment I, X-ray structures of the MgADP, MgATPγS, and MgAMPPNP complexes of the *Dictyostelium discoideum* myosin motor domain. *Biochemistry*, 36 (1997) 11619.

31 Bello L, Gordish-Dressman H, Morgenroth LP, Henricson EK, Duong T, Hoffman EP, Cnaan A, McDonald CM, Abresch RT, Cregan M & Goude E, Prednisone/prednisolone and deflazacort regimens in the CINRG Duchenne Natural History Study. *Neurology*, 85 (2015) 1048.

32 Griggs RC, Miller JP, Greenberg CR, Fehlings DL, Pestronk A, Mendell JR, Moxley RT 3rd, King W, Kissel JT, Cwik V, Vanasse M, Florence JM, Pandya S, Dubow JS & Meyer JM, Efficacy and safety of deflazacort vs prednisone and placebo for Duchenne muscular dystrophy. *Neurology*, 87 (2016) 2123.

33 Xu D, Zhao L, Jiang J, Li S, Sun Z, Huang X, Li C, Wang T, Sun L, Li X, Jiang Z & Zhang L, A potential therapeutic effect of catalpol in Duchenne muscular dystrophy revealed by binding with TAK1. *J Cachexia Sarcopenia Muscle*, 11 (2020) 1306.

34 Hoffman EP, Schwartz BD, Mengle-Gaw LJ, Smith EC, Castro D, Mah JK, McDonald CM, Kuntz NL, Finkel RS, Guglieri M, Bushby K, Tulinius M, Nevo Y, Ryan MM, Webster R, Smith AL, Morgenroth LP, Arrieta A, Shimony M, Siener C, Jaros M, Shale P, McCall JM, Nagaraju K, van den Anker J, Conklin LS, Cnaan A, Gordish-Dressman H, Damsker JM & Clemens PR, Vamorolone trial in Duchenne muscular dystrophy shows dose-related improvement of muscle function. *Neurology*, 93 (2019) e1312.

35 Kim S, Chen J, Cheng T, Gindulyte A, He J, He S, Li Q, Shoemaker BA, Thiessen PA, Yu B, Zaslavsky L, Zhang J & Bolton EE, PubChem 2019 update: Improved access to chemical data. *Nucleic Acids Res*, 47 (2019) D1102.

36 Banerjee P, Kemmler E, Dunkel M & Preissner R, ProTox 3.0: A webserver for the prediction of toxicity of chemicals. *Nucleic Acids Res*, 52 (2024) W513.

37 Dallakyan S & Olson AJ, Small-molecule library screening by docking with PyRx. *Methods Mol Biol*, (2014) 243.

38 Blessy JJ, Shanmugam NS, Veluraja K & Gromiha MM. Sialic acid analog as inhibitor for human coronavirus OC43-a study by molecular dynamics simulations. *J. Carbohydr. Chem*, 42, (2023) 197.

39 Pronk S, Páll S, Schulz R, Larsson P, Bjelkmar P, Apostolov R, Shirts MR, Smith JC, Kasson PM, Van Der Spoel D & Hess B, GROMACS 4.5: A high-throughput and highly parallel open source molecular simulation toolkit. *Bioinformatics*, 29 (2013) 845.

40 Liu T, Lyu C & Wang X, Resveratrol ameliorates sepsis-induced immune dysfunction by targeting GADD45A to maintain hematopoietic stem cell homeostasis: A predictive study integrating ensemble machine learning, single-cell transcriptomics, and molecular dynamics simulation. *Int. Immunopharmacol,* 176 (2026) 116470.

41 Tong Y, Chen R, Ling H, Shen Y, Tian H & Zeng L, Network pharmacology and molecular dynamics simulation elucidate the potential mechanism of Batatasin-III in *Bletilla striata* against ulcerative colitis. *Sci Rep*, 16 (2025) 1032.

42 Berglund AE, Welsh EA & Eschrich SA, Characteristics and validation techniques for PCA-based gene-expression signatures. *Int J Genomics*, (2017) 2354564.

43 Eisen B & Binah O, Modeling Duchenne muscular dystrophy cardiomyopathy with patients' induced pluripotent stem-cell-derived cardiomyocytes. *Int J Mol Sci*, 24 (2023) 8657.

44 Schultz TI, Raucci Jr FJ & Salloum FN, Cardiovascular disease in Duchenne muscular dystrophy: Overview and insight into novel therapeutic targets. *JACC Basic Transl Sci*, 7 (2022) 608.

45 Xu H, Cai X, Xu K & Wu Q, The metabolomic plasma profile of patients with Duchenne muscular dystrophy: Providing new evidence for its pathogenesis. *Orphanet J Rare Dis*, 18 (2023) 273.

46 Hershberger RE, Parks SB, Kushner JD, Li D, Ludwigsen S, Jakobs P, Nauman D, Burgess D, Partain J & Litt M, Coding sequence mutations identified in MYH7, TNNT2, SCN5A, CSRP3, LBD3 and TCAP from 313 patients with familial or idiopathic dilated cardiomyopathy. *Clin Transl Sci*, 1 (2008) 21.

47 Jolfayi AG, Kohansal E, Ghasemi S, Naderi N, Hesami M, Mozafary Bazargany M, Moghadam MH, Fazelifar AF, Maleki M & Kalayinia S, Exploring TTN variants as genetic insights into cardiomyopathy pathogenesis and potential emerging clues to molecular mechanisms in cardiomyopathies. *Sci Rep*, 14 (2024) 5313.

48 Ikelaar NA, Barnard AM, Eng SWM, Hosseini Vajargah S, Ha KCH, Kan HE, Vandenborne K, Niks EH, Walter GA & Spitali P, Large-scale serum protein biomarkers discovery associated with function and clinical milestones in Duchenne muscular dystrophy. *Nat Commun*, 16 (2025) 9073.

49 Heydemann A & Siemionow M, A brief review of Duchenne muscular dystrophy treatment options, with an emphasis on two novel strategies. *Biomedicines*, 11 (2023) 830.