



## OPEN ACCESS

## EDITED BY

Pukar Khanal,  
Emory University, United States

## REVIEWED BY

Bijendra Mandar,  
Sopan Pharmaceutical Pvt Ltd., Nepal  
Kunal Bhattacharya,  
Pratiksha Institute of Pharmaceutical Sciences  
(PIPS), India

## \*CORRESPONDENCE

Mohd Saeed,  
✉ mo.saeed@uoh.edu.sa

RECEIVED 02 March 2024

ACCEPTED 31 May 2024

PUBLISHED 02 October 2024

## CITATION

Saeed M, Haque A, Shoaib A and Danish Rizvi SM (2024), Exploring novel natural compound-based therapies for Duchenne muscular dystrophy management: insights from network pharmacology, QSAR modeling, molecular dynamics, and free energy calculations. *Front. Pharmacol.* 15:1395014. doi: 10.3389/fphar.2024.1395014

## COPYRIGHT

© 2024 Saeed, Haque, Shoaib and Danish Rizvi. This is an open-access article distributed under the terms of the [Creative Commons Attribution License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

# Exploring novel natural compound-based therapies for Duchenne muscular dystrophy management: insights from network pharmacology, QSAR modeling, molecular dynamics, and free energy calculations

Mohd Saeed<sup>1\*</sup>, Ashanul Haque<sup>2</sup>, Ambreen Shoaib<sup>3</sup> and Syed Mohd Danish Rizvi<sup>4</sup>

<sup>1</sup>Department of Biology, College of Sciences, University of Ha'il, Ha'il, Saudi Arabia, <sup>2</sup>Department of Chemistry, College of Sciences, University of Ha'il, Ha'il, Saudi Arabia, <sup>3</sup>Department of Clinical Pharmacy, College of Pharmacy, Jazan University, Jazan, Saudi Arabia, <sup>4</sup>Department of Pharmaceutics, College of Pharmacy, University of Ha'il, Ha'il, Saudi Arabia

Muscular dystrophies encompass a heterogeneous group of rare neuromuscular diseases characterized by progressive muscle degeneration and weakness. Among these, Duchenne muscular dystrophy (DMD) stands out as one of the most severe forms. The present study employs an integrative approach combining network pharmacology, quantitative structure-activity relationship (QSAR) modeling, molecular dynamics (MD) simulations, and free energy calculations to identify potential therapeutic targets and natural compounds for DMD. Upon analyzing the GSE38417 dataset, it was found that individuals with DMD exhibited 290 upregulated differentially expressed genes (DEGs) compared to healthy controls. By utilizing gene ontology (GO) and protein-protein interaction (PPI) network analysis, this study provides insights into the functional roles of the identified DEGs, identifying ten hub genes that play a critical role in the pathology of DMD. These key genes include DMD, TTN, PLEC, DTNA, PKP2, SLC24A, FBXO32, SNTA1, SMAD3, and NOS1. Furthermore, through the use of ligand-based pharmacophore modeling and virtual screening, three natural compounds were identified as potential inhibitors. Among these, compounds 3874518 and 12314417 have demonstrated significant promise as an inhibitor of the SMAD3 protein, a crucial factor in the fibrotic and inflammatory mechanisms associated with DMD. The therapeutic potential of the compounds was further supported by molecular dynamics simulation and Molecular Mechanics/Generalized Born Surface Area (MM/GBSA) analysis. These findings suggest that the compounds are viable candidates for experimental validation against DMD.

## KEYWORDS

Duchenne muscular dystrophy, network pharmacology, computational drug discovery, pharmacophore modeling, molecular dynamics simulation, QSAR

# 1 Introduction

Muscular dystrophies (MD) refer to a variety of inherited conditions that cause progressive weakening and degeneration of skeletal muscles. This group of disorders is known for its genetic and clinical diversity, with symptoms appearing from birth to late adulthood depending on the specific type. The most diverse forms of MD include limb-girdle, facioscapulohumeral, oculopharyngeal, emery-Dreifuss, distal, and Duchenne and Becker muscular dystrophies (Angelini, 2010). MD encompass a range of genetic disorders that can be X-linked, autosomal recessive, or autosomal dominant. These conditions manifest as muscular discomfort, weakness, and degeneration. Mutations in proteins within the sarcomere, nucleus, basement, or outer membrane of muscle cells and nonstructural enzymatic proteins are implicated in the pathogenesis of MD (Amato and Griggs, 2011). Different genetic deletions or mutations can lead to enzymatic or metabolic disorders, resulting in a variety of MDs (Gao and McNally, 2015; Allen et al., 2016). Variation in the X-linked gene DMD, where dystrophin is mutated frequently, causes Duchenne muscular dystrophy (DMD), which is the most severe form (Brinkmeyer-Langford and Kornegay, 2013). This condition affects around 1 in every 3,500 to 5,000 male infants born globally (Hoffman et al., 1987). Downregulation of protein expression during DMD or BMD is an outcome of mutations associated with Dystrophin gene (Le Rumeur, 2015). Limb-girdle muscular dystrophy (LGMD) is a progressive muscle weakness affecting the pelvic and shoulder girdles, primarily caused by mutations in proteins like myotilin, lamin, caveolin-3, calpain-3, dysferlin,  $\gamma$ -sarcoglycan, TCAP, TRIM32, FKRP, and titin. These mutations cause various subtypes of LGMD and are associated with other forms of muscular dystrophies, including distal and congenital dystrophies (Lovering et al., 2005).

Genetic profiling has revealed significant changes in gene expression in DMD. These changes offer insights into the disease's molecular basis and highlight potential therapeutic targets. Altered gene expression contributes to characteristic features of the disease. These altered gene expression leads to inflammation, fibrosis, and muscle degeneration. Studies have identified multiple genes associated with these biological processes, suggesting potential therapeutic interventions. Wang et al. (2013) found several hub genes associated with DMD, including C3AR1, TLR7, IRF8, and CD33, which are linked to immune and inflammation responses. Wang et al. (2021) identified proteins acting as hubs for DMD and BMD, finding 1,281 genes overexpressed and 189 downregulated. The informational evolution underscores the complexity of DMD and the need for ongoing research to understand its molecular intricacies.

Based on the immunomodulators of DMD pathology, various drug targets were used in phase I and preclinical trials (Malik et al., 2012; Tripodi et al., 2021). Inhibiting the SMAD3 signaling pathway to help reduce inflammation and fibrosis in DMD patients. SMAD3 is a protein with two domains, MH1 and MH2, crucial for TGF- $\beta$ -induced transcriptional activation. The MH2 domain interacts with transcriptional cofactors and the type I TGF- $\beta$  receptor. It binds to various proteins without a common sequence motif, making it essential for activating TGF- $\beta$  signals, as demonstrated by studies the domain of SMAD3 in transforming growth factor- $\beta$  signaling (Imoto et al., 2005a).

Halofuginone was tested in the mouse model to inhibit SMAD3 signaling, and it was observed that the inhibition of SMAD3 reduced fibrosis and improved muscle function (Tripodi et al., 2021). Osseni et al. (2022) found that tubastatin A can inhibit HDAC6, thereby enabling the acetylation of SMAD3. It also prevents nuclear translocation and Smad2/3 phosphorylation, reducing muscle atrophy and fibrosis by downregulating TGF- $\beta$  signaling through Smad3 acetylation. Overall, both studies showed that inhibiting SMAD3 restricts the progression of fibrosis improves muscle movement.

This study aims to focus on the SMAD3 gene, which is a key mediator in the TGF- $\beta$  signaling pathway in DMD. The gene expression data from the GEO database was used to identify differentially expressed genes and to construct a Protein-Protein Interaction network. Molecular docking was used to screen 2,569 natural compounds against SMAD3's MH2 domain, selecting three compounds with strong binding affinities. The study validated these compounds' potential as therapeutic agents through molecular dynamics simulations, principal component analysis (PCA), free energy landscape analysis (FEL analysis), and MM/GBSA calculations. This approach addresses a gap in current management strategies for DMD and sets the stage for developing more targeted and effective therapeutic options. The study's integrative approach combines high-throughput gene expression analysis, network-based bioinformatics, and advanced computational docking and simulation methods.

## 2 Methodology

### 2.1 Data collection, preprocessing and differential gene expression analysis

In order to find the gene expression data, a search was performed in the Gene Expression Omnibus (GEO) database using the keyword "muscular dystrophy" (Clough and Barrett, 2016). The data was filtered specifically for the species "*Homo sapiens*." The microarray dataset with the accession number GSE38417 was selected for further examination based on the data processing. The selected dataset contains the microarray datasets of healthy (control) and active MD patients. The microarray gene expression data from the selected groups were analysed using the GEO2R tool to determine the DEGs (Rezaei and Jabbari, 2022). The GEO2R software was used to obtain the *p*-value and false discovery rate (FDR) using the *T*-test and the Benjamini and Hochberg technique (Aubert et al., 2004). DEGs were selected by using a cut-off value of an absolute value of the  $\log(\text{fold change}) > 2$  and a significance threshold of  $p < 0.05$ . DEGs were classified as upregulated or downregulated according to  $\log\text{FC} \geq 2$  and  $\log\text{FC} \leq -2$ , respectively.

### 2.2 Protein-protein interaction network analysis, gene ontology and pathway analysis

This study integrated GEO for expression data, STRING for building interaction networks, and Cytoscape for network

visualization (Menche et al., 2015). In order to examine the protein-protein interaction (PPI) network of DEGs, the search Tool for the Retrieval of Interacting Genes' (STRING, version 12.0) (Szklarczyk et al., 2021) was utilised. The PPI network was visualised using the web-based software Cytoscope. The next step was to locate the hub genes in Cytoscape (Shannon et al., 2003) by using the MCC method and the CytoHubba plugin (Chin et al., 2014). Further, gene ontology was employed to evaluate the roles of the genes.

Here, the GO approach was used to analyze the molecular function of the gene. To perform GO analysis, the g:Profiler server (Reimand et al., 2007) was used, and the GOST tool was used to define the advanced parameters. Over-representation analysis (ORA), also known as gene set enrichment analysis, functional enrichment analysis, and other similar analyses, are all services provided by the GOST programme (Raudvere et al., 2019). This analysis is carried out on a list of input genes. Annotated genes were the primary ones chosen in this case under the advanced option. After that, g:SCS was selected as a significance threshold and set the user threshold to 0.5. When determining the  $p$ -values, the default parameter was used as the starting point, and a false discovery rate (FDR) threshold of  $p < 0.05$  was implemented in order to analyze the data. A statistically significant  $p$ -value of less than 0.05 was achieved by using the ShinyGO 0.80 programme to carry out the route analysis (Ge et al., 2020). The pathway database known as the Kyoto Encyclopaedia of Genes and Genomes (KEGG) was used by this algorithm in order to discover the genes that were implicated (Kanehisa and Goto, 2000). The KEGG pathway representation predominantly focuses the intricate relationship of gene products, primarily proteins, while additionally includes functional RNAs (Kanehisa and Goto, 2000). The FDR cut-off in the KEGG pathway has been set as 0.05. Consequently, relevant genes were acquired, and among them, additional selection was conducted based on the enrichment score (0.05).

## 2.3 Target preparation

The top hub genes were identified using information from the literature to analyze their functions. Based on this data, a possible target gene was selected. This selected gene was then queried in the Uniprot database (UniProt Consortium, 2015), and the results were refined by applying a filter for humans (*H. sapiens*). The data was then examined and further analyzed based on family and domain information. As a result, a specific domain was selected as the target protein for further investigation.

## 2.4 Molecular docking

### 2.4.1 Protein preparation

The target protein SMAD3 (PDB: 1MJS) was downloaded from the protein data bank (PDB) (UniProt Consortium, 2015). CASTp server (Tian et al., 2018) was used to identify the residues, and the CASTp server (Tian et al., 2018) was used associated with binding sites. The protein was prepared using the Autodock tool. The grid box with the following dimensions,  $18 \text{ \AA} \times 22 \text{ \AA} \times 26 \text{ \AA}$ , was generated to cover the binding site residues. The centre of the grid box was situated at  $28.39 \text{ \AA} \times 5.22 \text{ \AA} \times 4.37 \text{ \AA}$  along the x, y,

and z-axis Hydrogen atoms and Gasteiger charges are added, while water molecules and heteroatoms are removed before docking.

### 2.4.2 Ligand preparation

The compound library preparation process began with a search of the Selleckchem database for the desired natural substance. (<https://www.selleckchem.com/screening/natural-product-library.html>). The retrieved compounds were later converted into SMILES for PubChem (Kim et al., 2023) CID using the "PubChem Identifier Exchange Service." By removing duplicate entries from both conversion processes, distinct CIDs were obtained. Then, the CIDs corresponding to the 3D-SDF structures were acquired using the PubChem API. The other 2D-SDF structures were Collected and converted retrieved and converted to 3D-SDF. After that, the 3D-SDF files were optimized in size using the MMFF94 force field, and Open-babel was used to convert these to PDBQT (O'Boyle et al., 2011). Ligand was converted to PDB from SDF using an open-babel tool. Similarly, open- Babel was used to convert the PDB file to a PDBQT file. The addition of hydrogen was also done by the open-Babel programme.

### 2.4.3 Ligand-protein docking

The AutoDock 4.2 software (Morris et al., 2009) and AutoDock Vina 1.2.0 (Eberhardt et al., 2021) were used in order to carry out the protein-ligand docking process. AutoDock Vina is a validated and robust molecular docking tool favoured for its advanced scoring function, efficient optimization algorithms, and multithreading capabilities (Trott and Olson, 2010). These features enhance the accuracy and speed of docking simulations, which is crucial for the reliable prediction of binding energies and modes in protein-ligand interactions. The tool's development focuses on both empirical and knowledge-based potentials, which ensures improved prediction accuracy compared to its predecessors (Forli et al., 2016). Additionally, in the process of conducting docking-based virtual screening, the following characteristics were taken into consideration: twenty binding modes, ten exhaustiveness, and a maximum energy difference of four (kcal/mol).

### 2.4.4 Molecular dynamics (MD) simulation

The protein-ligand complex has been simulated employing the poses were used for a 100 ns molecular dynamics simulation using Gromacs 2022.4 (2022) (Bauer et al., 2022) (MD). The CHARMM36 force field parameter was used to establish the molecular topology (Huang and MacKerell, 2013). The CGneFF server was used to build topologies and force-field parameters for both the hit molecule and the control inhibitor (Vanommeslaeghe et al., 2010). The Ewald particle mesh method was used to determine the electrostatic force across a specific distance. The system was hydrated using the TIP3P model after being placed into a solvation box (cubic) with a distance of 1.0 nm from the wall (Harrach and Drossel, 2014). The neutralisation process was then carried out using  $\text{Na}^+$  and  $\text{Cl}^-$  ions. After fifty thousand cycles of the steepest descent (SD) method, the system was able to get rid of steric conflicts. Afterwards, the LINCS approach was employed to attain system stability and restrict the bonds (Hess et al., 1997). Furthermore, during a 100 ps simulation period in the NVT ensemble, the system's temperature was raised to 310 K using a 2 fs timestep. Furthermore, the system was exposed to continuous

pressure (NPT ensemble) at 310 K and 1 atmosphere for 1 ns. The simulation was run for 100 ns. The velocity-rescaling technique (Hess et al., 1997) was used to include temperature coupling, while the Parrinello-Rahman pressure method (Hess et al., 1997) was used to provide constant pressure was employed to maintain constant pressure. Hydrogen bonding, root mean square fluctuation (RMSF), and root mean square deviation (RMSD) were evaluated with the use of the GROMACS internal tool during a subsequent post-MD examination.

#### 2.4.5 Principal component analysis (PCA)

Prior to principal component analysis, the trajectory was cleaned by eliminating the periodic boundary condition. For computing the covariance matrix, the Gmx covar tool that is included in the GROMACS package was used. It is possible to utilise the covariance matrix in order to characterise the connection that exists between the atomic variations that are present in the protein-ligand complex. The eigenvalues and eigenvectors of the covariance matrix were obtained via the gmx anaeig function. The following step utilized the GROMACS software, specifically the "gmx anaproj" function, to obtain the PC coordinates for each frame.

#### 2.5 Free energy landscape (FEL)

The processes of biomolecule recognition, aggregation, and folding are some of the processes that can be better understood by examining the steady state, which is represented by the minima of the Free Energy Landscape (FEL), and the transient state, which is represented by the barriers of the FEL, in biological systems (Papaleo et al., 2009). The FEL was calculated by calculating the energy distribution according to Eq. 1a:

$$\Delta G(X) = -kBT \ln P(X) \quad (1a)$$

The variables  $X$ ,  $G$ ,  $kB$ ,  $T$ , and  $P(X)$  represent the reaction coordinate, Gibbs free energy, Boltzmann constant, absolute temperature, and probability distribution of the system along the reaction coordinate, respectively.

#### 2.6 Binding-free energy

When estimating the binding free energy of the protein-ligand complex, the GROMACS add-on tool gmx MM/PBSA was used as the information source (Valdes-Tresanco et al., 2021). The equations applied to compute the MM/GBSA are shown in Eqs 1–6.

$$\Delta G = G_{\text{complex}} - [G_{\text{receptor}} + G_{\text{ligand}}] \quad (1b)$$

$$\Delta G_{\text{binding}} = \Delta H - T\Delta S \quad (2)$$

$$\Delta H = \Delta G_{\text{GAS}} + \Delta G_{\text{SOLV}} \quad (3)$$

$$\Delta G_{\text{GAS}} = \Delta E_{\text{EL}} + \Delta E_{\text{VDWAALS}} \quad (4)$$

$$\Delta G_{\text{SOLV}} = \Delta E_{\text{GB}} + \Delta E_{\text{SURF}} \quad (5)$$

$$\Delta E_{\text{SURF}} = \gamma \cdot \text{SASA} \quad (6)$$

Here,  $\Delta G$  is defined as the variation in the protein-ligand formation's Gibbs free energy in Eq. 1b. In the solvent,  $G_{\text{complex}}$ ,  $G_{\text{receptor}}$  and  $G_{\text{ligand}}$  denote the total free energies of the protein-

ligand complex, free enzyme, and ligand, respectively.  $\Delta G_{\text{binding}}$  signifies the modification in the Gibbs free energy corresponding to the protein-ligand binding interaction, while  $\Delta H$  stands for the modification in enthalpy that incorporates both the gas-phase energy ( $\Delta G_{\text{GAS}}$ ) and the total solvation free energy ( $\Delta G_{\text{SOLV}}$ ). The binding free energy, shown as  $T\Delta S$ , is the cumulative sum of the change in entropy.  $\Delta E_{\text{EL}}$  represents a change in electrostatic energy, while  $\Delta E_{\text{VDWAALS}}$  signifies the change in van der Waals energy. These two components combine to yield  $\Delta G_{\text{GAS}}$ . Moreover,  $\Delta E_{\text{GB}}$  indicates the change in polar solvation energy brought on by polar group interaction, while  $\Delta E_{\text{SURF}}$  denotes a change in solvation-free energy brought on by the non-polar interaction. A variation of the solvent-accessible surface area (SASA) was made using the solvent surface tension parameter ( $\gamma$ ) in order to compute the solvent-accessible surface area (SURF) of the solvent.

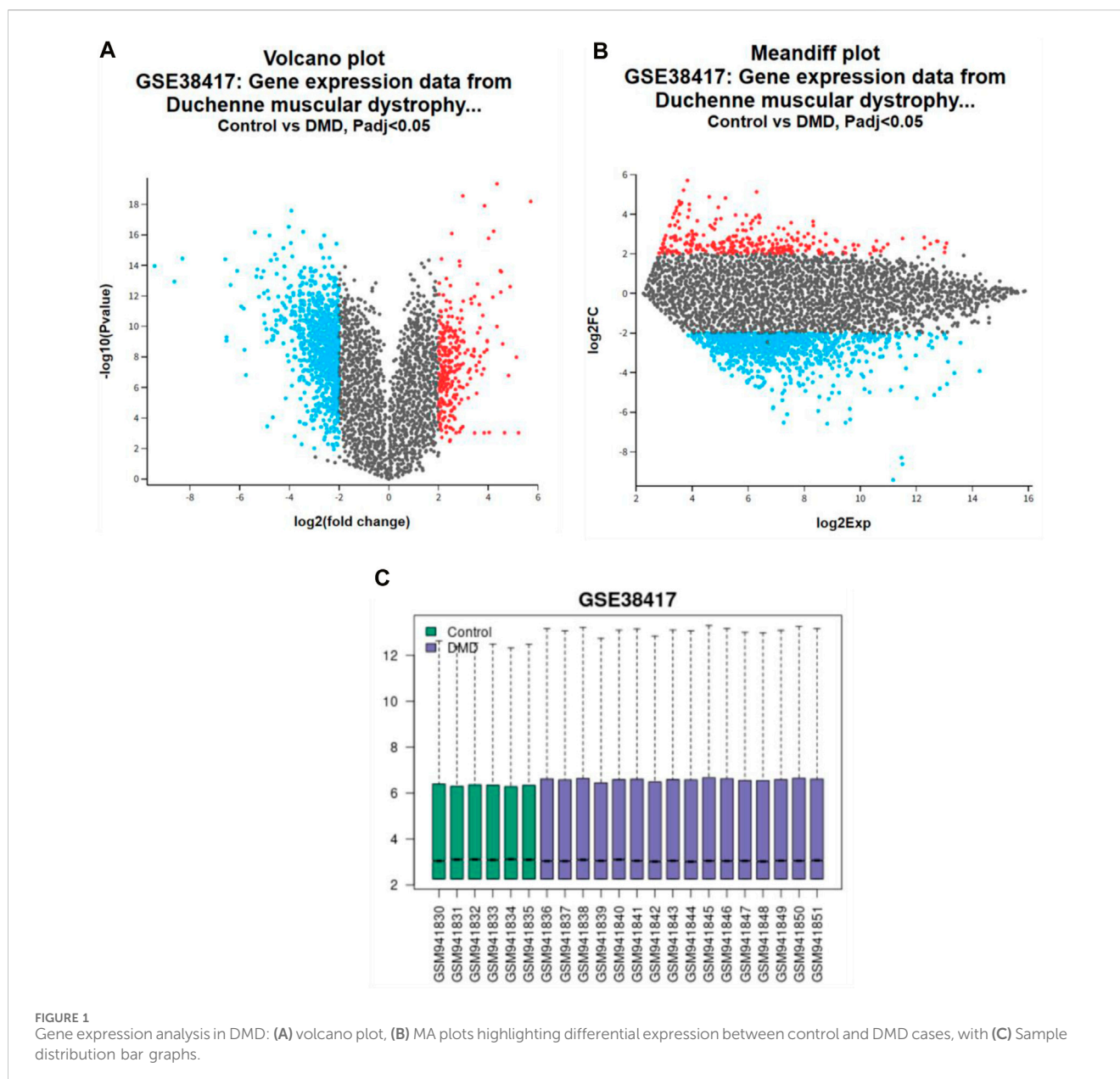
### 3 Results

#### 3.1 Healthy vs. active comparison for Duchenne muscular dystrophy (DMD)

Using microarray datasets from healthy persons and those with active DMD, the current research tried to discover the genes that are expressed differently between the two groups of people. The GSE38417 dataset was derived from the GEO database, which served as its primary source. The GEO2R approach was used to find DEGs in the dataset. A volcano plot is used to show the differential gene expression analysis results (Figure 1A). The x-axis displays the log<sub>2</sub> fold change between healthy and DMD conditions; each gene's expression level varies from condition to condition. The importance of the expression change is shown by negative log<sub>10</sub> of the  $p$ -value, displayed along the y-axis. The genes thought to be more differently expressed are those with greater and further to the left or right points. Following the requirements of an adjusted  $p$ -value ( $P_{\text{adj}}$ ) less than 0.05 and a logarithmic function (logFC) more than 2, genes that are substantially upregulated are represented by red dots, while genes that are significantly downregulated are represented by blue points. Here, we observed 290 upregulated and 977 downregulated DEGs. Figure 1B) illustrates a mean difference (MA) plot for gene expression data from DMD versus control samples. The y-axis shows the log<sub>2</sub>FC, log<sub>2</sub> Fold change and the x-axis shows the log<sub>2</sub>Exp, log<sub>2</sub> expression. The genes with red dots show a highly differentially expressed gene set, whereas genes with blue points may have modest expression levels but large fold changes.

Figure 1C shows the box plot that compares the distribution of gene expression levels between the healthy (green boxes) and the active DMD (blue boxes) for a series of samples on a logarithmic scale. The centre line in each box shows the median expression level, while the height of the box shows the interquartile range (the middle 50% of data points). There were 1,267 genes found to have differential expression between healthy and active DMD conditions, out of a total of 53,408, with an adjusted  $p$ -value less than 0.05. These plots were used to analyse and visualise the microarray dataset data and understand the biological significance of gene expression changes in diseases like Duchenne



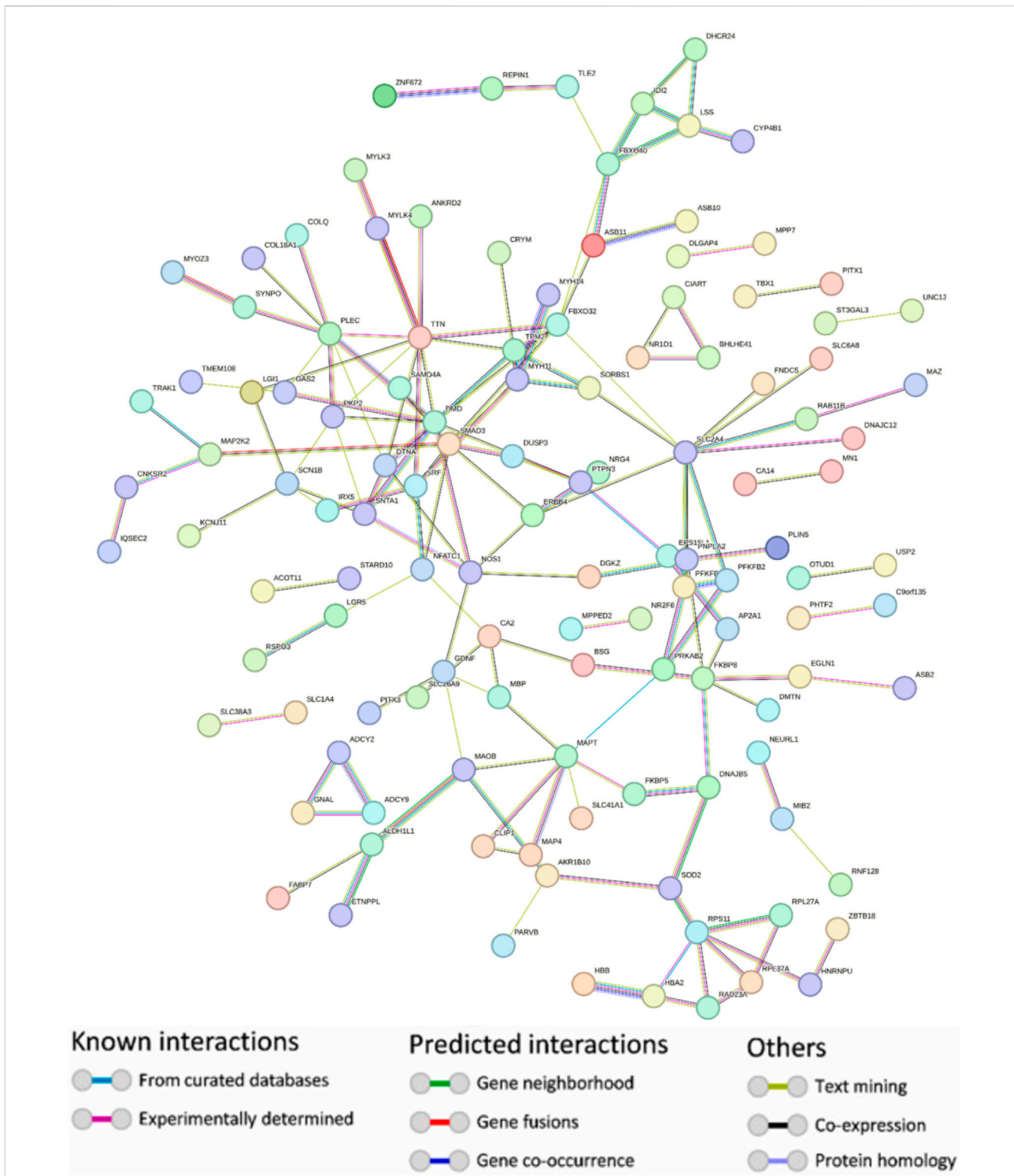


muscular dystrophy. There is a chance that the genes that are elevated (290 DEGs) are responsible for the production of proteins that play an important part in the development of DMD disease. Therefore, the purpose of the research was to examine the dataset GSE38417 in order to identify the DEGs that were elevated.

### 3.2 Protein-protein interaction

Using the STRING platform, the upregulated differentially expressed genes (DEGs) were analysed for protein-protein interactions (PPIs). The medium confidence level was set at 0.90, which is the lowest needed interaction score. Figure 2 illustrates the PPI, which represents the interaction network between DEGs. Disconnected nodes were ignored in the analysis as shown in the Figure 2. Multiple significant nodes and links were among the

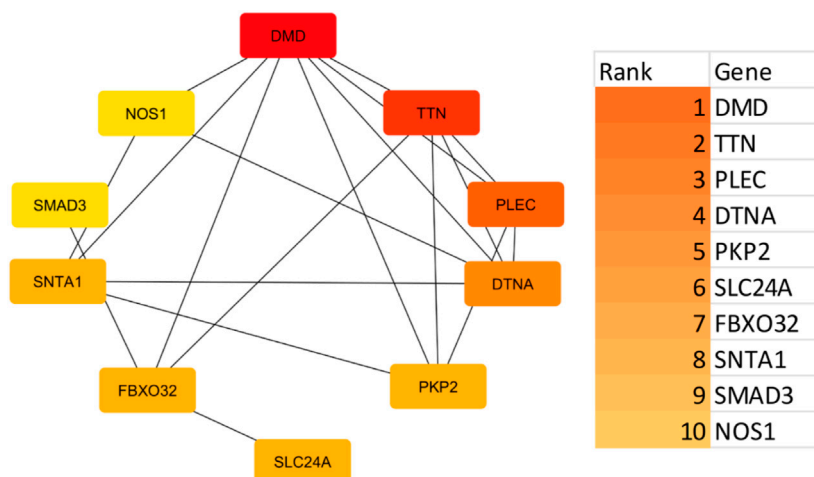
290 upregulated DEGs identified from GSE38417. TTN had the highest node degree, with a value of 11. The genes DMD and SLC2A4 exhibited a node degree of 10, which was the second highest among all genes. PLEC and SMAD3 followed this, which was 8 node degree. FBXO32 and MAPT exhibited significant node degree, with a value of 7. The genes were subjected to additional analysis using the Cytohubba plugin of Cytoscape. Cytoscape employed a colour gradient that spanned from red to yellow in order to represent the hubs, thereby signifying their varied degrees of importance within the network. Through this study, the most notable hubs are identified, which are subsequently used in constructing a PPI network. Figure 3 displayed the protein-protein interaction (PPI) network of the top 10 hubs, and Table 1 presented a ranking of the genes based on their scores. Both of these figures were shown in the same document. Figure 3 shows the use of Maximal Clique Centrality (MCC) values to



**FIGURE 2** Protein-protein interaction network depicting upregulated genes with hidden disconnected nodes. Node colours indicate different entity classifications, while edge colours represent the nature and source of interactions—ranging from experimentally determined (red) to predicted connections (green) and associations identified through text mining (blue) and co-expression (purple).

display the gene gradient. The top 10 hubs found by PPI network analysis were DMD, TTN, PLEC, DTNA, PKP2, SLC24A, FBXO32, SNTA1, SMAD3, and NOS1. The DMD gene exhibited the highest MMC score of 29, while the TTN gene

scored 23. These two genes, DMD and TTN, are the most significant differentially expressed genes (DEGs) that contribute to the development of muscular dystrophy. The PLEC gene scored 17, the DTNA gene exhibited a score of 12, and the PKP2, SLC24A,



**FIGURE 3**  
A key regulatory protein interaction network highlights central hub proteins as identified by cytohubba analysis. Gene interaction network with a gradient scale indicating Maximal Clique Centrality (MCC) values for each gene. The gradient bar on the right, ranging from dark orange to light yellow, visually represents the MCC values from highest to lowest, facilitating an understanding of the hierarchical importance of these genes in the network.

**TABLE 1** Top 10 hub proteins in network STRING network ranked by MCC method.

Rank	Name	Score	Gene symbol
1	ENSP00000354923	29	DMD
2	ENSP00000467141	23	TTN
3	ENSP00000323856	17	PLEC
4	ENSP00000470152	12	DTNA
5	ENSP00000070846	10	PKP2
5	ENSP00000320935	10	SLC24A
5	ENSP00000428205	10	FBXO32
5	ENSP00000217381	10	SNTA1
9	ENSP00000332973	9	SMAD3
9	ENSP00000477999	9	NOS1

FBXO32, and SNTA1 genes exhibited scores of 10. SMAD3 and NOS1 exhibited an MCC score of 9. The top 10 hubs were subsequently used for a literature analysis to examine the association between these genes and muscular dystrophy.

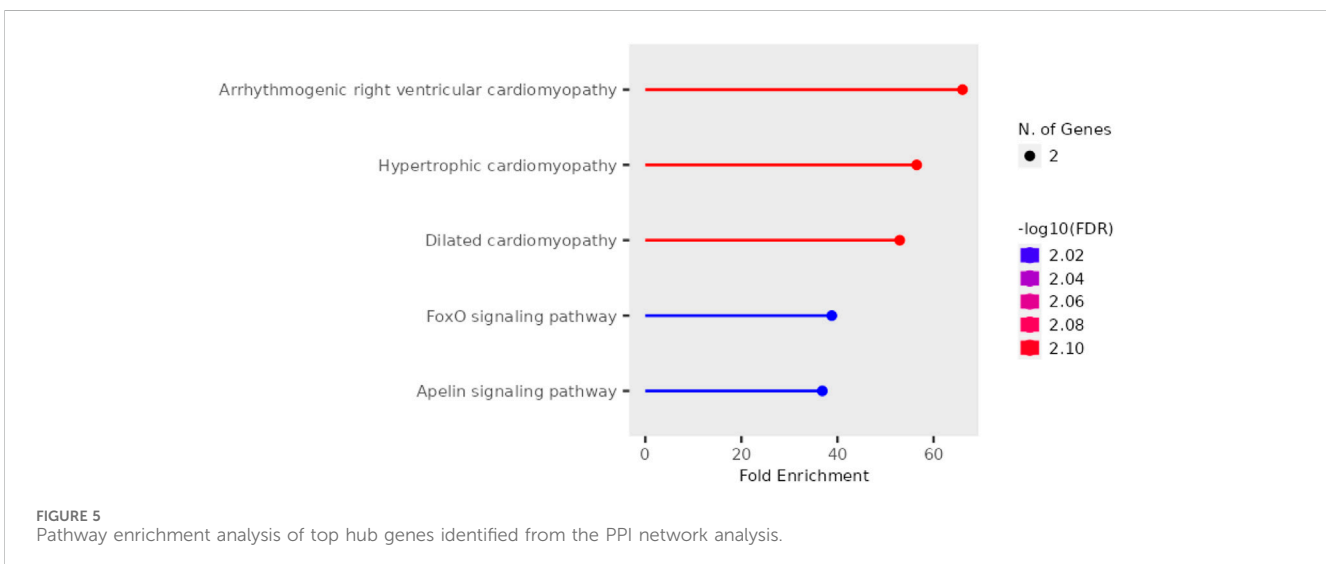
### 3.3 Gene ontology

Studies on the enrichment of pathways using Gene Ontology (GO) and the Kyoto Encyclopaedia of Genes and Genomes (KEGG) were carried out with the assistance of the g:Profiler tool. These efforts enhanced the understanding of the biological significance of the upregulated DEGs in the GSE38417 dataset. Examining the functional and route links between these DEGs helped researchers better understand their potential roles in biological systems. The user’s threshold was set to 0.5 in this case, which used the g:SCS threshold. The cellular components (CC), biological processes (BP),

and molecular functions (MF) domains were the primary areas of concentration for the GO analysis as shown in Figure 4. Sodium channel regulator activity, PDZ (PSD-95/Dlg/ZO-1) domain binding, dystroglycan binding, nitric-oxide synthase binding, and structural components of muscle were among the molecular processes linked to the DEGs. It was shown that the top ten hub proteins are linked to several biological processes, including those involving the cardiovascular system, striated muscle contraction, blood circulation, and cardiac muscle contraction. The top 10 hub proteins were shown to be connected with several cellular components, including the sarcolemma, Z disc, I band, syntrophin complex, and sarcomere. Duchenne muscular dystrophy (DMD) disrupts a complicated network of interactions and activities, as shown by the association between DEGs and these biological processes and cellular components. Hub proteins have recently been identified, which raises the possibility that they are therapeutic treatment targets due to their potential importance in disease progression. Gaining insight into these connections aids in understanding the underlying pathological processes of DMD and has the potential to identify efficient therapies.

### 3.4 Pathway analysis

Further, pathway enrichment analysis using the top 10 hub genes was performed. Figure 5 shows that each pathway has a horizontal line ending in a dot, with the dot’s position indicating the fold enrichment. The colour of the dot, which is a statistical metric that is used when testing several hypotheses, correlates to the  $-\log_{10}$  of the false discovery rate (FDR). In terms of statistical significance, a deeper shade indicates a greater enrichment than a lighter shade does. The size of the dot is a representation of the number of genes that are involved in the pathway; bigger dots will indicate that there are more genes involved. The arrhythmogenic right ventricular cardiomyopathy and dilated cardiomyopathy pathways exhibit the highest fold enrichment, accompanied by a



highly significant FDR (indicated by the red colour). Hypertrophic cardiomyopathy shows a slightly lower fold enrichment but still maintains a significant FDR. On the other hand, the FoxO signaling pathway and Apelin signaling pathway display lower fold

enrichments and less significant FDRs (indicated by the blue colour). The enrichment scores and the genes involved in the pathway are listed in Table 2. It was observed that DMD, TTN, and SMAD3 were involved in most of the pathways.



TABLE 2 Enrichment scores of key signalling pathways and their related genes.

Enrichment FDR	Fold enrichment	Pathway	Genes
0.007792883	66.03463203	Arrhythmogenic right ventricular cardiomyopathy	DMD, PKP2
0.007792883	56.4962963	Hypertrophic cardiomyopathy	DMD, TTN
0.007792883	52.96527778	Dilated cardiomyopathy	DMD, TTN
0.009610153	38.81424936	FoxO signalling pathway	FBXO32, SMAD3
0.009610153	36.84541063	Apelin signalling pathway	SMAD3, NOS1

### 3.5 SMAD3 a key target in Duchenne muscular dystrophy

The target identified by reviewing various literatures, the upregulation of specific genes in muscular dystrophies, including DMD, might involve intricate interactions between compensatory mechanisms, cellular stress responses, and pathogenic processes. The therapeutic genes involved in DMD were discussed here. The Dystrophin gene is generally not excessively expressed in DMD. However, mutations result in a lack or substantial decrease of dystrophin protein (Duan et al., 2021). Titin (TTN) could potentially be up-regulated as a compensatory mechanism in response to muscle injury. The function of titin in preserving muscle flexibility and structural integrity is critical. In order to preserve the structure and functionality of the muscles, the body may respond to damage to the muscle fibers by up-regulating TTN expression (Loescher et al., 2022). Plectin, or PLEC, serves as a link between the cell membrane and the cytoskeleton, especially in muscle cells. The overexpression of this gene may occur as a reaction to heightened mechanical strain and cellular instability in dystrophic muscles. The overexpression of DTNA may be a way for the dystrophin-associated protein complex to compensate for dystrophin's loss or malfunction by stabilising the membrane of the muscle cell (Wiche, 1998). The PKP2 gene, also known as Plakophilin 2, has a role in cell adhesion and may be upregulated in muscular dystrophies as a protective reaction to cellular stress and damage, in an effort to preserve cellular integrity (Papaleo et al., 2009). The SLC24A gene family is responsible for encoding sodium/potassium/calcium exchangers. Overexpression may be associated with disrupted ion homeostasis in dystrophic muscles, a prevalent characteristic in muscular diseases.

The FBXO32 gene, also known as Atrogin-1, plays a role in the development of muscular atrophy. Overexpression of this gene in muscular dystrophy may serve as an indicator of continuous muscle loss and atrophy, which are characteristic features of these disorders (Ghasemi et al., 2022). The SNTA1 gene, also known as Syntrophin Alpha 1, is a constituent of the dystrophin complex. Overexpression may arise as a compensatory strategy to stabilise the muscular membrane when functional dystrophin is lacking (Jimenez-Vazquez et al., 2022). NOS1, also known as Nitric Oxide Synthase 1, plays a role in the synthesis of nitric oxide, which serves as a signaling molecule. The overexpression of this gene may be associated with changes in blood circulation or inflammation in muscles affected by dystrophy (Buchwalow et al., 2006). SMAD3 overexpression, which is implicated in TGF-beta signaling, may be associated with fibrosis, a prevalent

consequence in muscular dystrophies (Liu et al., 1997). The upregulation of these genes is frequently a reaction to the underlying muscle dysfunction and not a primary instigator of the disease itself. These alterations can be involved in intricate feedback loops where the body endeavors to counterbalance or react to muscle injury, but they can also contribute to the advancement of disease in certain instances. Comprehending these alterations in gene expression is crucial for the development of precise treatments for muscle dystrophies.

SMAD3 overexpression has been associated with muscular dystrophies, according to research. The significance of this discovery lies in the fact that SMAD3 is an essential component of the TGF- $\beta$  signaling pathway, which plays a role in the processes of tissue remodeling, inflammation, and fibrosis. There are a total of two domains that make up SMAD3. The N-terminus is the location of one of the two, whereas the C-terminus is the location of the other. Because the MH2 domain engages in interactions with a multitude of transcriptional cofactors and the type I TGF- $\beta$  receptor (T $\beta$ R-I), it is essential for the activation of transcription by TGF- $\beta$ . One of the earlier studies indicated that the four particular lysine residues that make up the SMAD3 MH2 domain—Lys333, Lys341, Lys378, and Lys409—were essential to the operation of the domain (Imoto et al., 2005b). Therefore, the goal of targeting the MH2 domain of SMAD3 was to impede the fibrosis process associated with Duchenne muscular dystrophy (DMD).

The SMAD3 gene plays a critical role in DMD, primarily through its involvement in the TGF- $\beta$  signaling pathway, which is instrumental in processes such as tissue remodeling, inflammation, and fibrosis. These processes are central to the pathology of DMD, where progressive muscle degeneration is compounded by excessive fibrosis and chronic inflammation. The activation of SMAD3 in TGF- $\beta$  signaling leads to the upregulation of extracellular matrix proteins, contributing significantly to the muscle stiffness and fibrosis seen in DMD patients (Bernasconi et al., 1999; Qin et al., 2011). By modulating SMAD3 activity, there is potential to alter the course of tissue remodeling and inflammatory responses that exacerbate muscle damage (Roberts et al., 2003). Furthermore, the role of SMAD3 in DMD shares mechanistic similarities with its function in cancer metastasis, where it regulates cellular proliferation, migration, and invasion—processes similarly detrimental in DMD albeit through different pathological outcomes (Ismaeel et al., 2019). This crossover highlights SMAD3 as a versatile target in both oncological and muscular pathologies, underpinning its importance across different disease contexts. Targeting the TGF- $\beta$ /SMAD3 pathway in DMD could therefore mitigate fibrotic tissue development and reduce

TABLE 3 Average and top binding scores of the ligands, along with the number of hydrogen bonds (selected compounds are shown in Bold).

Ligands	Average binding energy (kcal/mol)	Top binding energy	Hydrogen bonds
<b>5,281,600</b>	-8.43	<b>-9.6</b>	<b>5</b>
72,950,887	-7.64	-9.6	1
<b>3,874,518</b>	-7.97	<b>-9.5</b>	<b>9</b>
<b>12,314,417</b>	-7.84	<b>-9.5</b>	<b>6</b>
4,485,132	-7.38	-9.4	4
4,463,283	-7.96	-9.3	6
437,080	-7.82	-9.3	6
5,281,847	-8.15	-9.2	3
4,441	-8.08	-9.2	6
328,441	-7.99	-9.2	0

inflammation, potentially preserving muscle function and improving patient outcomes (Flanders, 2004). Given this pivotal role, targeting SMAD3 presents a promising therapeutic strategy for managing DMD, as evidenced by research models showing reduced fibrosis and improved muscle functions upon modulation of this pathway (Zhou and Lu, 2010). These studies collectively advocate for further exploration of SMAD3 inhibitors as potential therapeutic agents in DMD treatment regimens.

### 3.6 Virtual screening

The MH2 domain of SMAD3 was tested against 2,569 natural compounds. The binding site residues of the MH2 domain of SMAD3 were GLN315, PRO317, ASN320, ALA328, ARG367, THR370, ILE371 and ARG372, as predicted by CASTp server. Top three hits were selected that showed a top binding energy > -9 kcal/mol (Table 3). The three compounds were 12,314,417 (Ojv-VI), 3,874,518 (Hederacoside C), and 5,281,600 (Amentoflavone). These compounds had binding values of -9.6 kcal/mol, -9.5 kcal/mol, and -9.5 kcal/mol, respectively.

Figure 6 shows the best-docked complexes had protein-ligand interactions. It was observed that 5,281,600 forms five hydrogen bonds with residues Pro87, Gln85, Glu152, Tyr153, and Thr136 (Figure 6A). 3,874,518 forms nine hydrogen bonds with the residues Tyr7, Glu166, Gln85, Arg154, Glu152, Gln91, Pro87, Ala98, and Arg142, (Figure 6B). 12,314,417 forms six hydrogen bonds with the residues Arg142, Thr140, Glu166, His168, Tyr133, and Thr136. The stability of interactions between proteins and ligands is dependent on hydrogen bonds. The identification of the individual residues involved in these interactions can offer valuable insights about the binding mechanism of the compound to the protein and its impact on the protein's function (Salentin et al., 2014). Therefore, in order to facilitate molecular-dynamic simulation, these three compounds were selected.

The 2D structure of the identified compounds are shown in the Supplementary Figure S1. The Tanimoto similarity of compounds 12,314,417, 3,874,518, and 5,281,600 with SIS3, the known inhibitor of SMAD3 phosphorylation, compared to study the fundamental differences in their mechanisms of action. The values in Supplementary Table S1 suggest that 5,281,600 has low to

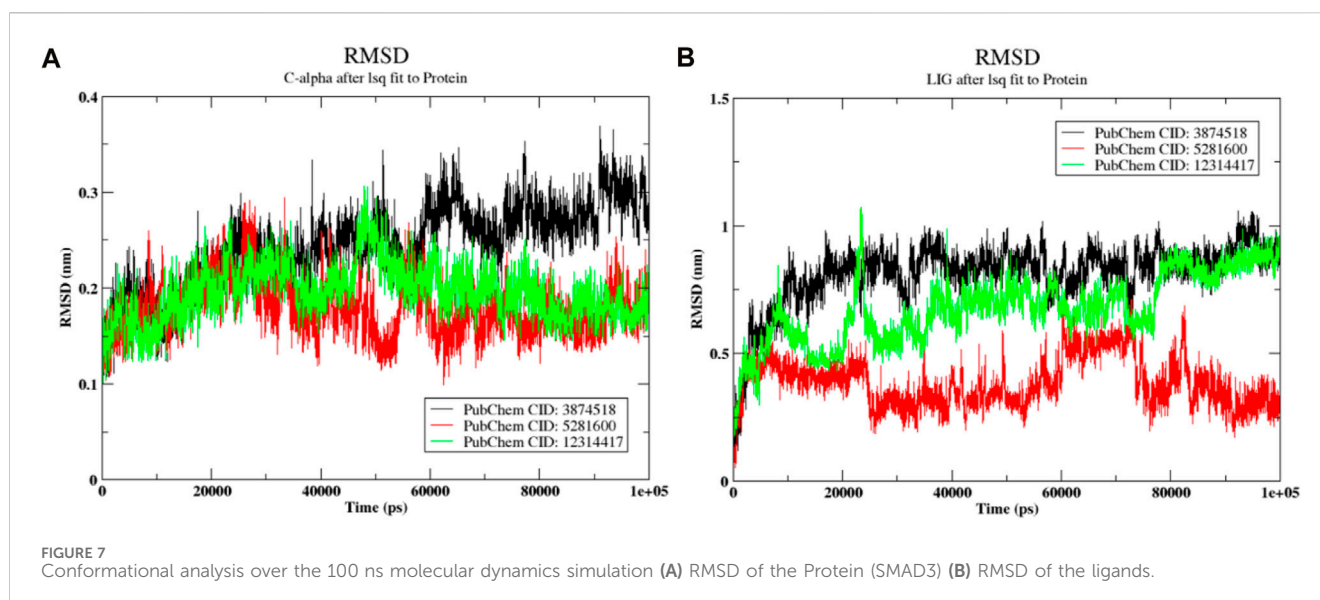
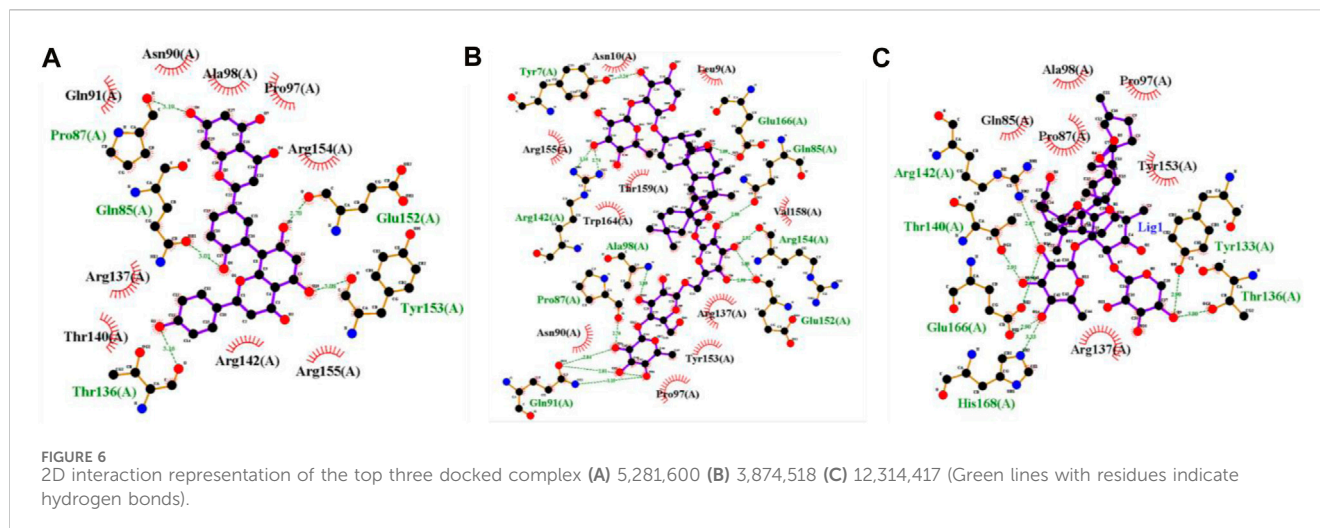
moderate structural resemblance to SIS3, while 3,874,518 and 12,314,417 show very little similarity. The Tanimoto scores indicate considerable structural differences from SIS3, suggesting that similar biological effects can be achieved through entirely different chemical structures and binding mechanisms.

### 3.7 Molecular dynamics simulation

The post-dynamics simulation study provides critical insights into the flexibility of protein-ligand complexes. During the 100 ns production run, only the best-docked complexes were analyzed. The Root Mean Square Deviation (RMSD) of both the protein and ligand, as shown in Figure 7, highlights the stability and conformational changes over time. These RMSD values indicate how closely the system maintains its initial structure, offering valuable information about the dynamic behavior and potential binding efficacy of the ligand within the active site of the protein.

#### 3.7.1 RMSD

The RMSD is a statistical measure that evaluates the degree toward which protein complexes deviate from their initial structure. This allows for the determination of the overall stability of the complexes. As can be seen in Figure 7A, the magnitude of the protein varied from 0.2 nm to 0.37 nm. The protein in the 3,874,518 complex first showed a variance of between 0.2 and 0.3 nm. After that, the more significant fluctuation rose to a range of 0.32–0.36 nm. With the exception of a single, more notable deviation at 45 ps by 0.3 nm, the protein molecule in the 5,281,600-ligand complex changed consistently between 0.2 nm and 0.25 nm. A stable conformation between 0.25 and 0.24 nm was also observed in the protein combination with 12,314,417. In general, the prediction of a stable conformation in the simulation was supported by the protein RMSD. Figure 7B displays the ligand's Root Mean Square Deviation (RMSD). At the initial stage, the ligand 3,874,518 exhibited significant fluctuation and reached 0.7 nm, followed by a subsequent period of consistent configuration. Similarly, during the early phase, 12,314,417 exhibited fluctuations with variations reached 0.7 nm and 1.2 nm. On the other side, after 70 ps, it arrived at a stable conformation with a



difference of 77 nm regarding the original docked stance. Meanwhile, the 5,281,600 showed deviation ranging from 0.5 nm to 0.4 nm and stability in the last 20 ps. Overall, in the last 20 ps all ligands showed stability.

### 3.7.2 Conformation analysis

Figures 8A,B,C show the positions of the compounds at 0, 90, and 100 ns during the molecular dynamics simulation, respectively. It was noted that the 3,874,518 showed a deviation from its starting state at 90 ns. However, from 90 to 100 ns, it maintained a similar conformation. Similarly, compounds 5,281,600 and 12,314,417 pose at 0, 90, and 100 ns, as shown in Figures 8D–J, showing a deviation from their starting states at 90 ns. However, from 90 to 100 ns, it maintained a stable conformation.

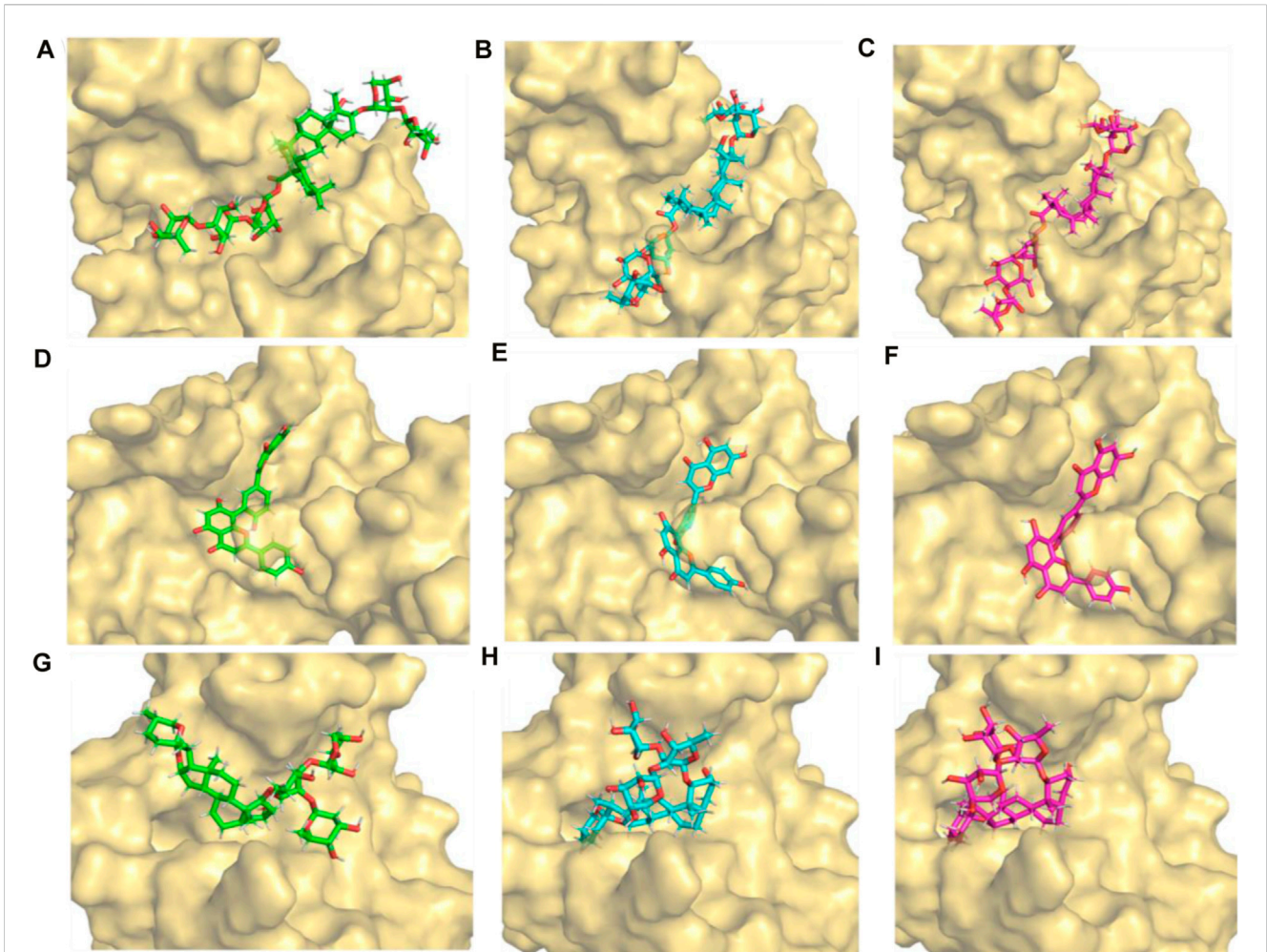
Despite some early spatial configurational deviations, all three compounds were able to achieve stable conformations by the conclusion of the simulation duration, as seen above. This suggests that, during the course of 100 ns, these compounds'

molecular motions go through an adaptation phase and then enter a stable state.

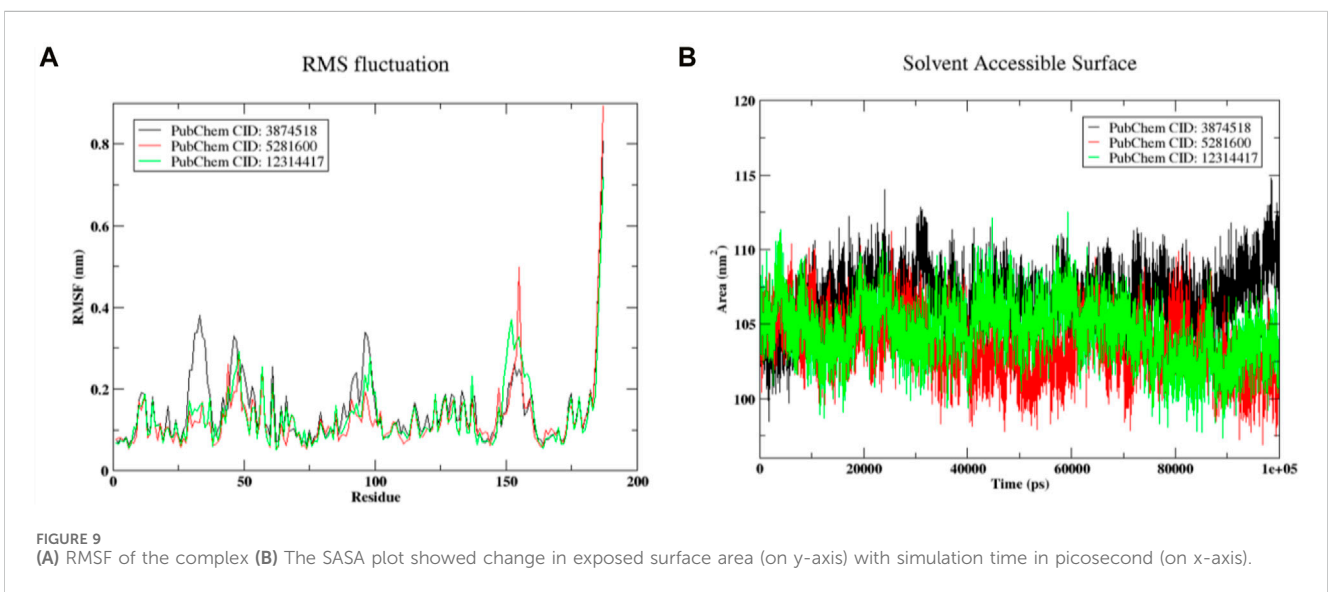
### 3.7.3 RMSF and SASA studies

Figure 9A displays the protein's root mean square fluctuation. In this case, it was shown that complexed protein residues with compound 3,874,518 showed larger fluctuations of 0.3 nm than residues with ASP31, PRO32, SER33, ASN34, and SER35. Protein in the complex of compound 5,281,600 showed a higher fluctuation of 0.3 nm, which appears in the residues VAL46 and ASN467. Similarly, protein complexed with compound 12,314,417 showed the maximum fluctuation of 0.3 nm that occurred in the residue HIS96 and PRO97. In general, the complexed form did not exhibit significant changes in protein residues. The SASA plot of SMAD3 is shown in Figure 9B when it is coupled to the ligands 3,874,518, 5,281,600, and 12,314,417. The 3,874,518 complex was shown to have a stable conformation at first, but in the last 20 ns, it diverged, suggesting changes in the solvent-exposed surface area. Similarly, in



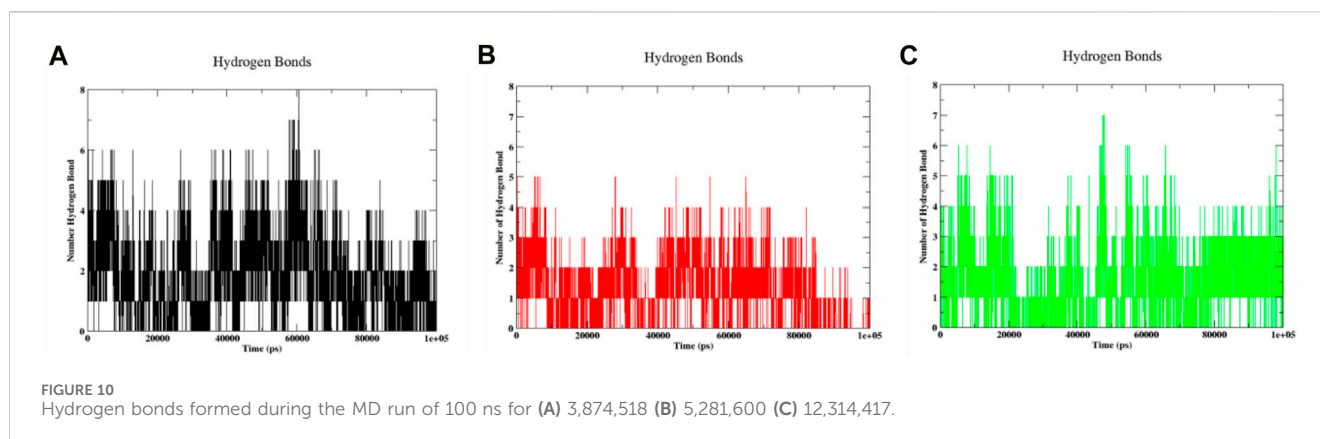


**FIGURE 8** Initial (0 ns), stable (90 ns), and final (100 ns) conformation of the protein-ligand complexes extracted from 100 ns molecular dynamics simulation (A–C) 3,874,518 (D–F) 5,281,600 and (G–I) 12,314,417.



**FIGURE 9** (A) RMSF of the complex (B) The SASA plot showed change in exposed surface area (on y-axis) with simulation time in picosecond (on x-axis).





the initial phase, 5,281,600 showed stable conformation, but at the last 20 ns, it deviated downward from 105 nm<sup>2</sup> to 110 nm<sup>2</sup> which indicated the extension of the molecule in the solvent. Additionally, 12,314,417 shows relatively stable SASA, suggesting a stable structure in the solvent during the simulation period. Change in SASA showed the conformational change, however, this change was in the acceptable range.

### 3.7.4 Hydrogen bonds analysis

Hydrogen bonds are crucial in molecular dynamics simulations as they significantly impact the structure, function, and interactions of biomolecules within a simulated environment. This facilitates the understanding of their behaviour within actual biological systems (Salo-Ahen et al., 2020). Within the context of the actual case, an estimation of the formation of hydrogen bonds was carried out during the molecular dynamics simulation of the protein-ligand pair. As the ligand 3,874,518 is attached to the protein, Figure 10A shows that it forms 6–8 hydrogen bonds during the 20–60 ns simulation and then progressively reduces to 3–6 hydrogen bonds after 60 ns until 100 ns. Figure 10B depicts the formation of a hydrogen bond when ligand 5,281,600 binds to the protein. It was found that the compound 5,281,600 produced 4–5 bonds within the first 70 ns, however, after that time period, the number of bonds rapidly decreased from 4 to 2. Figure 10C shows the synthesis of hydrogen resulting from the binding of the compound 12,314,417 to the protein. During the initial period of 20 ns, there was a downward trend of sloping down in the formation of hydrogen bonds. However, the value shifted from 20 to 40 ns to 2 to 5. At a time interval of 45 ns, it established a total of 7 hydrogen bonds. The bond formation steadily dropped from 45 ns to of the range of 6 to 4. Throughout the simulation, the amount of hydrogen bonds fluctuated, creating an uneven pattern. However, it is noticeable that detection of H-bonds is highly sensitive to the interatomic distance. These interatomic distances changed marginally during the simulation that can affect the detection of H-bonds. Moreover, the positive side of the observation shown in Figure 10 is that compounds always shown presence of minimum one hydrogen bond.

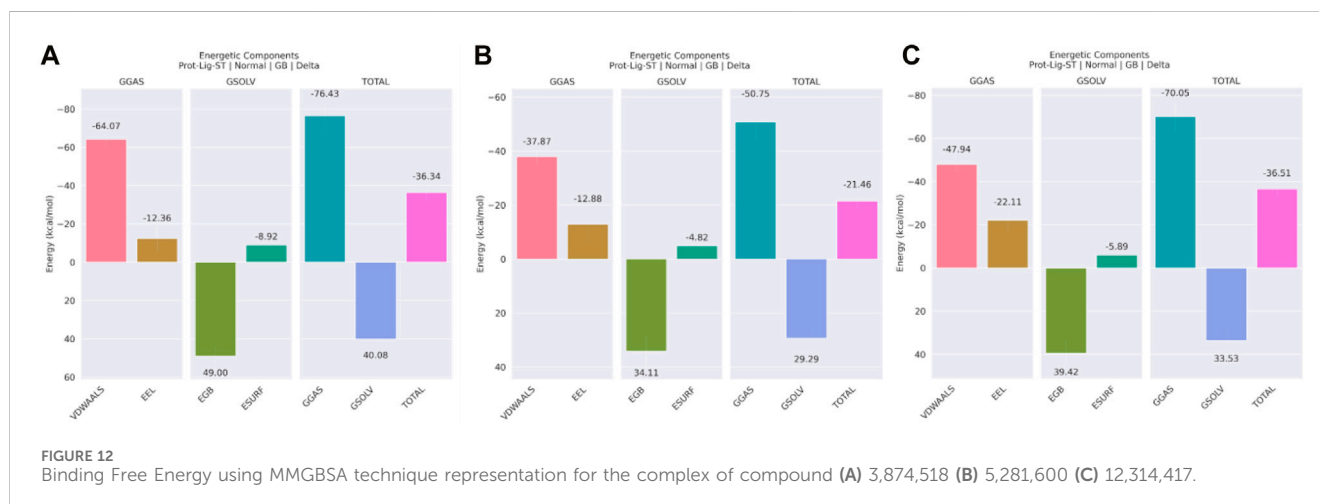
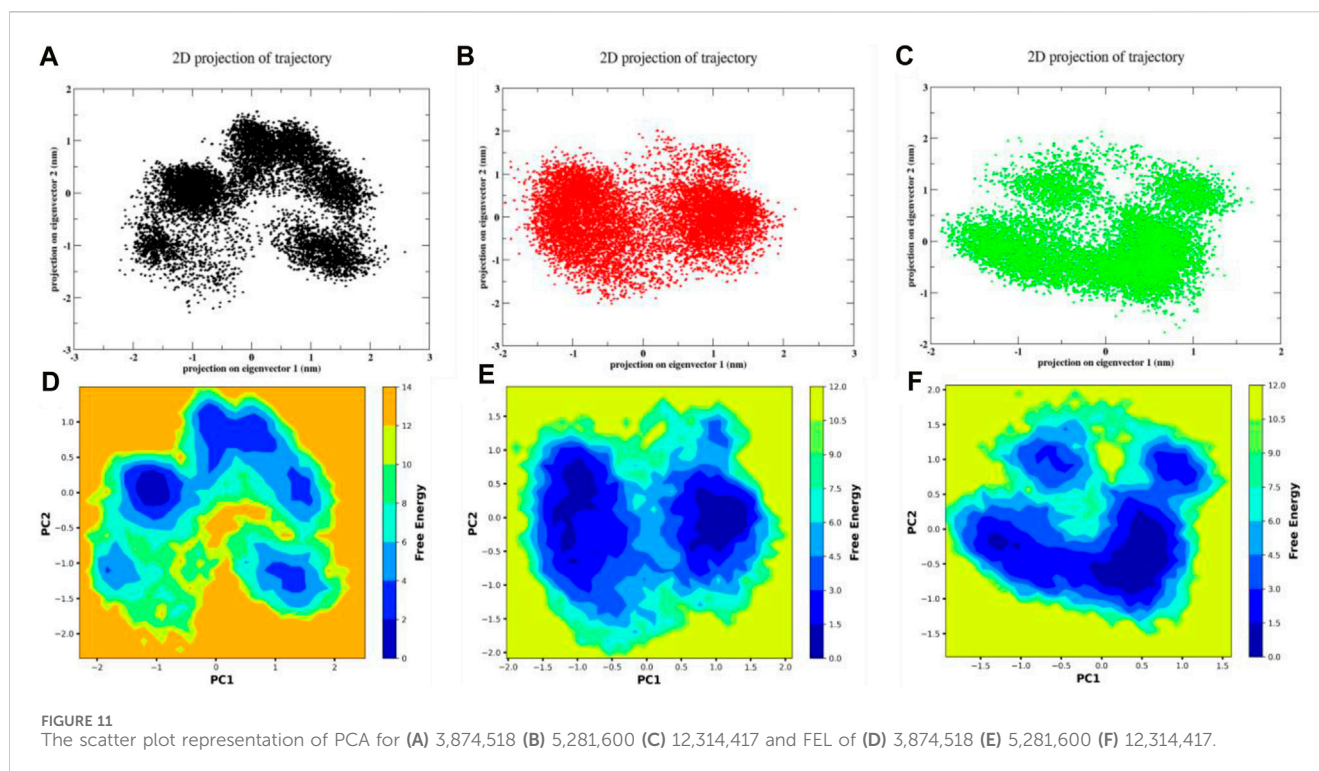
### 3.7.5 Principal component analysis (PCA) and free energy landscape (FEL)

A scatter plot is used to display the trajectory that was generated by the MD simulation. The Principal Component Analysis is used to provide this plot (PCA). Figure 11A showed the formation of four

distinct clusters by molecule 3,874,518 during the MD run, illustrating its transition from its initial to final state. Similarly, the 12,314,417 formed three distinct clusters, but these clusters exhibited a higher level of interconnection as shown in Figure 11C. This suggests that for this compound complex, there was a lower degree of transition compared to the 3,874,518. The 5,281,600 complex formed two closely grouped clusters as shown in Figure 11B, suggesting a higher level of stability compared to other molecules. Additionally, the free energy landscape (FEL) of the 3,874,518, 5,281,600, and 12,314,417, respectively, is shown in Figures 11(D–F). Figures 11(D–F) demonstrates the energy distribution across the conformational landscape. The plot shows regions of low free energy basin in blue, which correspond to more stable conformations or states that the protein-ligand complex frequently occupies. The surrounding areas in orange and yellow represent higher free energy states. The plots represented the stability states of several conformations for the aforementioned system that developed throughout the simulation run. The x-axis showed PC1, while the y-axis plotted PC2. These principal components served as reaction coordinates to show various conformations and their corresponding free energy. Here, in Figure 11C, the plot of 3,874,518 showed four energy basins with multiple energy barriers that suggested a stable conformations of the system separated in the space by barriers. Similarly, Figure 11E 12,314,417 showed three energy basins with multiple energy barriers this suggested system conformation during simulation. However, the FEL of 5,281,600 showed two energy basins with a single energy barrier, as observed in Figure 11D, which indicated stable conformation of the system. Overall, PCA and FEL results indicated that the 5,281,600 could be a strong binder of SMAD3. Overall, complex 5,281,600 and 12,314,417 showed more stable characteristics than 3,874,518 where the conformations are distributed in wider space with few minima. The total portion occupied by the blue color in the plots were larger for 5,281,600 and 12,314,417 that also suggests larger for 5,281,600 and 12,314,417, suggesting its higher chance of reaching minima.

### 3.7.6 Binding free energy

Several different energetic components were used to determine the binding-free energy of the protein-ligand combination.



According to the data shown in Figure 12A, the 3,874,518 complex, which is composed of the GGAS and GSOLV, has a total binding energy of  $-36.34$  kcal/mol overall. Moreover, in Figure 12B, total binding free energy was observed as  $-21.46$  kcal/mole for the 5,281,600 complex, and in Figure 12C, total binding energy of the 12,314,417 complex observed as  $-36.51$  kcal/mole. This suggested that the 12,314,417 complex showed strongest binding affinity with the protein molecule. Later, close observation on the different energetic components were made, van der Waals contributed most negatively ( $-64.07$ ,  $-37.87$ , and  $-47.94$  kcal/mol) in all three complexes. This showed that steric clashes were perfectly removed in all the complexes. Electrostatic interaction was also favorable for all the three complexes, in the 12,314,417 complex

the electrostatic interaction was best. The complex 3,874,518 and 12,314,417 showed very similar binding energy and emerged as the best ligand to bind with the target protein.

Additionally, the binding free energy of the compounds were calculated using the MM/PBSA technique. Supplementary Figure S2 showed the binding free energy of the complexes. 12,314,417 had the lowest binding free energy was 29.38 kcal/mol 3,874,518 also showed similar binding free energy with 27.81 kcal/mole 5,281,600 had binding free energy of 17.09 kcal/mol. The binding free energy trend was similar in both MM/GBSA and MM/PBSA technique. Moreover, binding free energy was calculated for comparative analysis of the three compounds. Among the three compounds, both 3,874,518 and 12,314,417 better binding free energy than 5,281,600.

### 3.8 Discussion

Recent studies on DMD have highlighted gene-targeted medicines, including exon skipping and gene editing, as effective methods to slow down the disease's advancement. The main goal of these therapies is to repair or substitute the function of the faulty dystrophin gene, which plays a crucial role in DMD pathogenesis (Chung Liang et al., 2022; Moorwood et al., 2011). Advancements in comprehending the interaction between the immune system and skeletal muscles have revealed new possibilities for therapeutic targeting. Studies have emphasised the significance of T-Cell profiling and transcriptome analysis of circulating immune cells in DMD, indicating their potential as new biomarkers or therapeutic targets. This immunological analysis could assist in identifying specific groups of cells that either worsen or can reduce muscle damage, allowing for a more customised therapeutic approach (Molinaro et al., 2024). Research on gene therapy and cell transplantation is ongoing. The procedures entail directly introducing accurate copies of the dystrophin gene or transplanting genetically repaired muscle stem cells into patients. Although these technologies show promise, they also pose notable technological and ethical obstacles such as delivery mechanisms, lasting effectiveness, and possible unintended consequences (Sun et al., 2020).

DMD is characterised by abnormal gene expression, and this work comprehensively evaluates those genes. The gene expression patterns of healthy individuals, as well as those with active DMD, are compared to achieve this. Specifically, the GSE38417 dataset from the Gene Expression Omnibus is used for the study (GEO). In a prior study, Lombardo et al. (2021) and Wu et al. (2022) used a dataset that was quite similar to GSE38417 in order to explore gene modules and their relationships with DMD. Additionally, 290 differentially expressed genes (DEGs) that were upregulated were discovered in the first analysis, demonstrating that substantial gene expression changes connected with DMD. The upregulation genes directly impact the diseased condition. Controlling these upregulated genes would bring down the physiopathology for a given disease, the investigation by Moorwood et al. (2011) revealed that utrophin gene upregulation may be used as a treatment strategy for DMD. Additionally, Kemaladewi et al. (2019) demonstrated that in pre-symptomatic MDC1A mice, up-regulating Lama1 via the CRISPR/dCas9-based approach averted muscle fibrosis and hindlimb paralysis when started early. According to Kemaladewi et al. (2019), it partially reversed the course of the illness and reduced dystrophic features when administered to symptomatic mice with muscle fibrosis and hind limb paralysis that were already evident. The enhanced DEGs were then constructed for PPI networks by using the STRING database and a stringent interaction confidence score threshold of 0.90. This was done in order to guarantee optimal analytical accuracy. It was recognised that the expression of particular genes in DMD might represent a compensatory response to cellular stress or a manifestation of pathogenic processes. Hence, a thorough literature study was used to identify gene targets strategically. The DMD gene typically exhibits low expression in DMD due to mutations that result in a marked reduction of dystrophin protein. Conversely, genes such as TTN, SMAD3 and PLEC may be upregulated in response to muscle injury or cellular instability. However, on the basis of enrichment scores, SMAD3 is identified as the most connected gene. Likewise, Multiple studies have demonstrated that TGF $\beta$ 1/SMAD3 directly interact with the muscular membrane and inhibits the progression of muscle weakening (Goodman et al., 2013; Zhang et al., 2019). Moreover,

Pathway enrichment studies offer a valuable understanding of the functional and pathway connections of these genes that are expressed differently (DEGs), highlighting their involvement in immune-related activities. Consequently, it becomes clear that it might be a great target for the therapy of muscular dystrophy, especially Duchenne muscular dystrophy. It was then tested against a library of naturally occurring chemicals. Natural compound PGC-1  $\alpha$  have shown effective results in DMD conditions by targeting PPAR $\gamma$  (Suntar et al., 2020). Another study demonstrated that Isolecanoric acid (ILA), a natural product, was used to explore its anti-inflammatory effect in DMD conditions (Matias-Valiente et al., 2024). Three compounds in the presented study were identified as the most promising candidates for targeting SMAD3 and its desired domain. The protein molecule in the complex state showed a minimum deviation from its native state, and RMSD was under 0.3 nm during the complete simulation. Studies have shown that under 0.3 nm, RMSD could be considered as minimum and acceptable deviation (Sharma et al., 2021). This confirmed that the binding of the ligand brought some change in the protein conformation, but these were considered stable. However, the RMSD of the ligand molecule was higher than that of the protein molecule. The deviation in the ligand molecule in the protein-bound state is acceptable at the early stage of the simulation, which also underline the limitation of the rigid docking algorithm.

The molecular interactions between molecules 3,874,518, 5,281,600, and 12,314,417, which target the MH2 domain of SMAD3, are essential for influencing the TGF- $\beta$  signaling pathway implicated in conditions such as DMD, which is distinguished by an excessive accumulation of fibrosis (Chen et al., 2014). These molecules exhibit a strong affinity for the MH2 domain and a low free energy of binding, indicating that they establish stable and energetically advantageous interactions. The stability of these compounds is of the utmost importance for their potential as therapeutic agents, as it indicates that they can consistently and effectively disrupt the normal functioning of SMAD3 within the TGF- $\beta$  pathway, which is well-known for its substantial contribution to the promotion of fibrosis (Budi et al., 2021). Based on the interaction between these molecules and critical residues in the MH2 domain, it is possible that they could alter the activity of SMAD3 by impeding its normal signaling functions. This could involve impeding essential phosphorylation processes that are required for the activation of genes associated with fibrosis. Consequently, these molecules may reduce the pathological signaling responsible for fibrotic tissue formation, a primary component of DMD pathology, by inhibiting SMAD3. This activity corresponds to the therapeutic objectives of managing and alleviating fibrosis in muscular dystrophies, presenting a potentially fruitful pathway for advancing treatments. Extensive research has been conducted previously on the role of the TGF- $\beta$ /SMAD signaling pathway in fibrosis across various diseases, underlining the therapeutic potential of modulating this pathway (Bujak and Frangogiannis, 2007; Schiro et al., 2011; Piersma et al., 2015).

The suggested chemical interactions of these drugs with SMAD3 present a potential therapeutic approach for treating disorders marked by excessive fibrosis, such as DMD. Additional empirical research, including thorough biochemical testing and clinical studies, is necessary to validate these results and guarantee the safety and effectiveness of these compounds for human usage. The future applications of targeting the SMAD3 pathway in fibrotic diseases such as DMD are promising and diverse. Precision medicine initiatives could leverage genetic

insights to tailor SMAD3 inhibitors to the unique genetic makeup of individual patients, potentially enhancing treatment efficacy and reducing side effects. Insights like these are crucial for creating specific treatments that can control these pathways. The study's discussion on prospective treatment strategies, like using natural chemicals that target specific genes or pathways critical in DMD, emphasises the study's translational potential. This component shows promise by indicating potential pathways for translating the findings of this study into clinical therapies. This thorough method improves the comprehension of DMD on a molecular scale and paves the way for future therapeutic advancements focused on more efficient treatment or control of the condition.

## 4 Conclusion

The present study was performed to identify hit molecules targeting the MH2 domain of SMAD3 that is involved in DMD. SMAD3 upregulation showed a significant impact on DMD. Screening of natural compounds found three most promising hit compounds: (a) 3,874,518 (b) 5,281,600 (c) 12,314,417. In the molecular dynamic study, 3,874,518 and 12,314,417 showed stable conformation after initial conformational jump. Combined, these two compounds a protein-ligand complex with a greater binding free energy. The fact that these chemicals remain stable at the protein's binding site suggests that they may impose some kind of inhibitory mechanism on the protein. Further, experimental assays are required to validate the computational findings.

## Data availability statement

The raw data supporting the conclusion of this article will be made available by the authors, without undue reservation.

## Author contributions

MS: Supervision, Writing–original draft, Conceptualization, Funding acquisition, Project administration, Writing–review and editing. AH: Supervision, Writing–original draft, Data curation,

Validation, Visualization. AS: Conceptualization, Methodology, Project administration, Visualization, Writing–review and editing. SD: Investigation, Methodology, Validation, Visualization, Writing–original draft.

## Funding

The author(s) declare that financial support was received for the research, authorship, and/or publication of this article.

## Acknowledgments

The authors extend their appreciation to the King Salman center For Disability Research for funding this work through Research Group no KSRG-2023-422.

## Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

## Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

## Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fphar.2024.1395014/full#supplementary-material>

## References

- Allen, D. G., Whitehead, N. P., and Froehner, S. C. (2016). Absence of dystrophin disrupts skeletal muscle signaling: roles of Ca<sup>2+</sup>, reactive oxygen species, and nitric oxide in the development of muscular dystrophy. *Physiol. Rev.* 96 (1), 253–305. doi:10.1152/physrev.00007.2015
- Amato, A. A., and Griggs, R. C. (2011). Overview of the muscular dystrophies. *Handb. Clin. Neurol.* 101, 1–9. doi:10.1016/B978-0-08-045031-5.00001-3
- Angelini, C. (2010). Chapter 31: muscular dystrophy. *Handb. Clin. Neurol.* 95, 477–488. doi:10.1016/S0072-9752(08)02131-3
- Aubert, J., Bar-Hen, A., Daudin, J. J., and Robin, S. (2004). Determination of the differentially expressed genes in microarray experiments using local FDR. *BMC Bioinforma.* 5, 125. doi:10.1186/1471-2105-5-125
- Bauer, P., Hess, B., and Lindahl, E. (2022). *GROMACS 2022.1 manual*. doi:10.5281/ZENODO.2022.6103568
- Bernasconi, P., Di Blasi, C., Mora, M., Morandi, L., Galbiati, S., Confalonieri, P., et al. (1999). Transforming growth factor-beta1 and fibrosis in congenital muscular dystrophies. *Neuromuscul. Disord.* 9 (1), 28–33. doi:10.1016/S0960-8966(98)00093-5
- Brinkmeyer-Langford, C., and Kornegay, J. N. (2013). Comparative genomics of X-linked muscular dystrophies: the golden retriever model. *Curr. Genomics* 14 (5), 330–342. doi:10.2174/13892029113149990004
- Buchwalow, I. B., Minin, E. A., Müller, F. U., Lewin, G., Samoilo, V. E., Schmitz, W., et al. (2006). Nitric oxide synthase in muscular dystrophies: a re-evaluation. *Acta Neuropathol.* 111 (6), 579–588. doi:10.1007/s00401-006-0069-5
- Budi, E. H., Schaub, J. R., Decaris, M., Turner, S., and Derynck, R. (2021). TGF- $\beta$  as a driver of fibrosis: physiological roles and therapeutic opportunities. *J. Pathol.* 254 (4), 358–373. doi:10.1002/path.5680
- Bujak, M., and Frangogiannis, N. G. (2007). The role of TGF-beta signaling in myocardial infarction and cardiac remodeling. *Cardiovasc Res.* 74 (2), 184–195. doi:10.1016/j.cardiores.2006.10.002
- Chen, J., Xia, Y., Lin, X., Feng, X. H., and Wang, Y. (2014). Smad3 signaling activates bone marrow-derived fibroblasts in renal fibrosis. *Lab. Invest.* 94 (5), 545–556. doi:10.1038/labinvest.2014.43
- Chin, C. H., Chen, S. H., Wu, H. H., Ho, C. W., Ko, M. T., and Lin, C. Y. (2014). cytoHubba: identifying hub objects and sub-networks from complex



- interactome. *BMC Syst. Biol.* 8 (Suppl. 4), S11. doi:10.1186/1752-0509-8-S4-S11
- Chung Liang, L., Sulaiman, N., and Yazid, M. D. (2022). A decade of progress in gene targeted therapeutic strategies in Duchenne muscular dystrophy: a systematic review. *Front. Bioeng. Biotechnol.* 10, 833833. doi:10.3389/fbioe.2022.833833
- Clough, E., and Barrett, T. (2016). The gene expression Omnibus database. *Methods Mol. Biol.* 1418, 93–110. doi:10.1007/978-1-4939-3578-9\_5
- Duan, D., Goemans, N., Takeda, S., Mercuri, E., and Aartsma-Rus, A. (2021). Duchenne muscular dystrophy. *Nat. Rev. Dis. Prim.* 7 (1), 13. doi:10.1038/s41572-021-00248-3
- Eberhardt, J., Santos-Martins, D., Tillack, A. F., and Forli, S. (2021). AutoDock Vina 1.2.0: new docking methods, expanded force field, and Python bindings. *J. Chem. Inf. Model* 61 (8), 3891–3898. doi:10.1021/acs.jcim.1c00203
- Flanders, K. C. (2004). Smad3 as a mediator of the fibrotic response. *Int. J. Exp. Pathol.* 85 (2), 47–64. doi:10.1111/j.0959-9673.2004.00377.x
- Forli, S., Huey, R., Pique, M. E., Sanner, M. F., Goodsell, D. S., and Olson, A. J. (2016). Computational protein-ligand docking and virtual drug screening with the AutoDock suite. *Nat. Protoc.* 11 (5), 905–919. doi:10.1038/nprot.2016.051
- Gao, Q. Q., and McNally, E. M. (2015). The dystrophin complex: structure, function, and implications for therapy. *Compr. Physiol.* 5 (3), 1223–1239. doi:10.1002/cphy.c140048
- Ge, S. X., Jung, D., and Yao, R. (2020). ShinyGO: a graphical gene-set enrichment tool for animals and plants. *Bioinformatics* 36 (8), 2628–2629. doi:10.1093/bioinformatics/btz931
- Ghasemi, S., Mahdavi, M., Maleki, M., Salahshourifar, I., and Kalayinia, S. (2022). A novel likely pathogenic variant in the FBXO32 gene associated with dilated cardiomyopathy according to whole-exome sequencing. *BMC Med. Genomics* 15 (1), 234. doi:10.1186/s12920-022-01388-5
- Goodman, C. A., McNally, R. M., Hoffmann, F. M., and Hornberger, T. A. (2013). Smad3 induces atrogen-1, inhibits mTOR and protein synthesis, and promotes muscle atrophy *in vivo*. *Mol. Endocrinol.* 27 (11), 1946–1957. doi:10.1210/me.2013-1194
- Harrach, M. F., and Drossel, B. (2014). Structure and dynamics of TIP3P, TIP4P, and TIP5P water near smooth and atomistic walls of different hydrophobicity. *J. Chem. Phys.* 140 (17), 174501. doi:10.1063/1.4872239
- Hess, B., Bekker, H., Berendsen, H. J. C., and Fraaije, J. G. E. M. (1997). LINCS: a linear constraint solver for molecular simulations. *J. Comput. Chem.* 18 (12), 1463–1472. doi:10.1002/(sici)1096-987x(199709)18:12<1463::aid-jcc4>3.3.co;2-1
- Hoffman, E. P., Brown, R. H., Jr., and Kunkel, L. M. (1987). Dystrophin: the protein product of the Duchenne muscular dystrophy locus. *Cell* 51 (6), 919–928. doi:10.1016/0092-8674(87)90579-4
- Huang, J., and MacKerell, A. D., Jr. (2013). CHARMM36 all-atom additive protein force field: validation based on comparison to NMR data. *J. Comput. Chem.* 34 (25), 2135–2145. doi:10.1002/jcc.23354
- Imoto, S., Sugiyama, K., Sekine, Y., and Matsuda, T. (2005a). Roles for lysine residues of the MH2 domain of Smad3 in transforming growth factor-beta signaling. *FEBS Lett.* 579 (13), 2853–2862. doi:10.1016/j.febslet.2005.04.023
- Imoto, S., Sugiyama, K., Sekine, Y., and Matsuda, T. (2005b). Roles for lysine residues of the MH2 domain of Smad3 in transforming growth factor-beta signaling. *FEBS Lett.* 579 (13), 2853–2862. doi:10.1016/j.febslet.2005.04.023
- Ismael, A., Kim, J. S., Kirk, J. S., Smith, R. S., Bohannon, W. T., and Koutakis, P. (2019). Role of transforming growth factor- $\beta$  in skeletal muscle fibrosis: a review. *Int. J. Mol. Sci.* 20 (10), 2446. doi:10.3390/ijms20102446
- Jimenez-Vazquez, E. N., Arad, M., Macias, A., Vera-Pedrosa, M. L., Cruz, F. M., Gutierrez, L. K., et al. (2022). SNTA1 gene rescues ion channel function and is antiarrhythmic in cardiomyocytes derived from induced pluripotent stem cells from muscular dystrophy patients. *Elife* 11, e76576. doi:10.7554/eLife.76576
- Kanehisa, M., and Goto, S. (2000). KEGG: kyoto encyclopedia of genes and genomes. *Nucleic Acids Res.* 28 (1), 27–30. doi:10.1093/nar/28.1.27
- Kemaladewi, D. U., Bassi, P. S., Erwood, S., Al-Basha, D., Gawlik, K. I., Lindsay, K., et al. (2019). A mutation-independent approach for muscular dystrophy via upregulation of a modifier gene. *Nature* 572 (7767), 125–130. doi:10.1038/s41586-019-1430-x
- Kim, S., Chen, J., Cheng, T., Gindulyte, A., He, J., He, S., et al. (2023). PubChem 2023 update. *Nucleic Acids Res.* 51 (D1), D1373–D1380. doi:10.1093/nar/gkac956
- Le Rumeur, E. (2015). Dystrophin and the two related genetic diseases, Duchenne and Becker muscular dystrophies. *Bosn. J. Basic Med. Sci.* 15 (3), 14–20. doi:10.17305/bjbm.2015.636
- Liu, X., Sun, Y., Constantinescu, S. N., Karam, E., Weinberg, R. A., and Lodish, H. F. (1997). Transforming growth factor beta-induced phosphorylation of Smad3 is required for growth inhibition and transcriptional induction in epithelial cells. *Proc. Natl. Acad. Sci. U. S. A.* 94 (20), 10669–10674. doi:10.1073/pnas.94.20.10669
- Loescher, C. M., Hobbach, A. J., and Linke, W. A. (2022). Titin (TTN): from molecule to modifications, mechanics, and medical significance. *Cardiovasc Res.* 118 (14), 2903–2918. doi:10.1093/cvr/cvab328
- Lombardo, S. D., Basile, M. S., Ciurleo, R., Bramanti, A., Arcidiacono, A., Mangano, K., et al. (2021). A network medicine approach for drug repurposing in Duchenne muscular dystrophy. *Genes (Basel)* 12 (4), 543. doi:10.3390/genes12040543
- Lovering, R. M., Porter, N. C., and Bloch, R. J. (2005). The muscular dystrophies: from genes to therapies. *Phys. Ther.* 85 (12), 1372–1388. doi:10.1093/ptj/85.12.1372
- Malik, V., Rodino-Klapac, L. R., and Mendell, J. R. (2012). Emerging drugs for Duchenne muscular dystrophy. *Expert Opin. Emerg. Drugs* 17 (2), 261–277. doi:10.1517/14728214.2012.691965
- Maratonak, R., Laio, A., and Parrinello, M. (2003). Predicting crystal structures: the Parrinello-Rahman method revisited. *Phys. Rev. Lett.* 90 (7), 075503. doi:10.1103/PhysRevLett.90.075503
- Matias-Valiente, L., Sanchez-Fernandez, C., Rodriguez-Outeiriño, L., Ramos, M. C., Diaz, C., Crespo, G., et al. (2024). Evaluation of pro-regenerative and anti-inflammatory effects of isolecanoric acid in the muscle: potential treatment of Duchenne Muscular Dystrophy. *Biomed. Pharmacother.* 170, 116056. doi:10.1016/j.biopha.2023.116056
- Menche, J., Sharma, A., Kitsak, M., Ghiassian, S. D., Vidal, M., Loscalzo, J., et al. (2015). Disease networks. Uncovering disease-disease relationships through the incomplete interactome. *Science* 347 (6224), 1257601. doi:10.1126/science.1257601
- Molinaro, M., Torrente, Y., Villa, C., and Farini, A. (2024). Advancing biomarker discovery and therapeutic targets in Duchenne muscular dystrophy: a comprehensive review. *Int. J. Mol. Sci.* 25 (1), 631. doi:10.3390/ijms25010631
- Moorwood, C., Lozynska, O., Suri, N., Napper, A. D., Diamond, S. L., and Khurana, T. S. (2011). Drug discovery for Duchenne muscular dystrophy via utrophin promoter activation screening. *PLoS One* 6 (10), e26169. doi:10.1371/journal.pone.0026169
- Morris, G. M., Huey, R., Lindstrom, W., Sanner, M. F., Belew, R. K., Goodsell, D. S., et al. (2009). AutoDock4 and AutoDockTools4: automated docking with selective receptor flexibility. *J. Comput. Chem.* 30 (16), 2785–2791. doi:10.1002/jcc.21256
- O'Boyle, N. M., Banck, M., James, C. A., Morley, C., Vandermeersch, T., and Hutchison, G. R. (2011). Open Babel: an open chemical toolbox. *J. Cheminform* 3, 33. doi:10.1186/1758-2946-3-33
- Osseni, A., Ravel-Chapuis, A., Belotti, E., Scionti, I., Gangloff, Y. G., Moncollin, V., et al. (2022). Pharmacological inhibition of HDAC6 improves muscle phenotypes in dystrophin-deficient mice by downregulating TGF- $\beta$  via Smad3 acetylation. *Nat. Commun.* 13 (1), 7108. doi:10.1038/s41467-022-34831-3
- Papaleo, E., Mereghetti, P., Fantucci, P., Grandori, R., and De Gioia, L. (2009). Free-energy landscape, principal component analysis, and structural clustering to identify representative conformations from molecular dynamics simulations: the myoglobin case. *J. Mol. Graph Model* 27 (8), 889–899. doi:10.1016/j.jmgm.2009.01.006
- Piersma, B., Bank, R. A., and Boersema, M. (2015). Signaling in fibrosis: TGF- $\beta$ , WNT5, and YAP/TAZ converge. *Front. Med. (Lausanne)* 2, 59. doi:10.3389/fmed.2015.00059
- Qin, W., Chung, A. C. K., Huang, X. R., Meng, X. M., Hui, D. S. C., Yu, C. M., et al. (2011). TGF- $\beta$ /Smad3 signaling promotes renal fibrosis by inhibiting miR-29. *J. Am. Soc. Nephrol.* 22 (8), 1462–1474. doi:10.1681/ASN.2010121308
- Raudvere, U., Kolberg, L., Kuzmin, I., Arak, T., Adler, P., Peterson, H., et al. (2019). g:Profiler: a web server for functional enrichment analysis and conversions of gene lists (2019 update). *Nucleic Acids Res.* 47 (W1), W191–W198. doi:10.1093/nar/gkz369
- Reimand, J., Kull, M., Peterson, H., Hansen, J., and Vilo, J. (2007). g:Profiler—a web-based toolset for functional profiling of gene lists from large-scale experiments. *Nucleic Acids Res.* 35 (Web Server issue), W193–W200. doi:10.1093/nar/gkm226
- Rezaei, N., and Jabbari, P. (2022). *Immunoinformatics of cancers: practical machine learning approaches using R*. Academic Press.
- Roberts, A. B., Russo, A., Felici, A., and Flanders, K. C. (2003). Smad3: a key player in pathogenic mechanisms dependent on TGF-beta. *Ann. N. Y. Acad. Sci.* 995, 1–10. doi:10.1111/j.1749-6632.2003.tb03205.x
- Salentin, S., Haupt, V. J., Daminelli, S., and Schroeder, M. (2014). Polypharmacology rescored: protein-ligand interaction profiles for remote binding site similarity assessment. *Prog. Biophys. Mol. Biol.* 116 (2–3), 174–186. doi:10.1016/j.pbiomolbio.2014.05.006
- Salo-Ahen, O. M., Alanko, I., Bhadane, R., Bonvin, A. M. J. J., Honorato, R. V., Hossain, S., et al. (2020). Molecular dynamics simulations in drug discovery and pharmaceutical development. *Processes* 9 (1), 71. doi:10.3390/pr9010071
- Schiro, M. M., Stauber, S. E., Peterson, T. L., Krueger, C., Darnell, S. J., Satyshur, K. A., et al. (2011). Mutations in protein-binding hot-spots on the hub protein Smad3 differentially affect its protein interactions and Smad3-regulated gene expression. *PLoS One* 6 (9), e25021. doi:10.1371/journal.pone.0025021
- Shannon, P., Markiel, A., Ozier, O., Baliga, N. S., Wang, J. T., Ramage, D., et al. (2003). Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res.* 13 (11), 2498–2504. doi:10.1101/gr.1239303
- Sharma, P., Joshi, T., Joshi, T., Chandra, S., and Tamta, S. (2021). Molecular dynamics simulation for screening phytochemicals as  $\alpha$ -amylase inhibitors from medicinal plants. *J. Biomol. Struct. Dyn.* 39 (17), 6524–6538. doi:10.1080/07391102.2020.1801507

- Sun, C., Shen, L., Zhang, Z., and Xie, X. (2020). Therapeutic strategies for Duchenne muscular dystrophy: an update. *Genes* (Basel) 11 (8), 837. doi:10.3390/genes11080837
- Suntar, I., Sureda, A., Belwal, T., Sanches Silva, A., Vacca, R. A., Tewari, D., et al. (2020). Natural products, PGC-1  $\alpha$ , and Duchenne muscular dystrophy. *Acta Pharm. Sin. B* 10 (5), 734–745. doi:10.1016/j.apsb.2020.01.001
- Szklarczyk, D., Gable, A. L., Nastou, K. C., Lyon, D., Kirsch, R., Pyysalo, S., et al. (2021). The STRING database in 2021: customizable protein-protein networks, and functional characterization of user-uploaded gene/measurement sets. *Nucleic Acids Res.* 49 (D1), D605–D612. doi:10.1093/nar/gkaa1074
- Tian, W., Chen, C., Lei, X., Zhao, J., and Liang, J. (2018). CASTp 3.0: computed atlas of surface topography of proteins. *Nucleic Acids Res.* 46 (W1), W363–W367. doi:10.1093/nar/gky473
- Tripodi, L., Villa, C., Molinaro, D., Torrente, Y., and Farini, A. (2021). The immune system in Duchenne muscular dystrophy pathogenesis. *Biomedicine* 9 (10), 1447. doi:10.3390/biomedicine9101447
- Trott, O., and Olson, A. J. (2010). AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *J. Comput. Chem.* 31 (2), 455–461. doi:10.1002/jcc.21334
- UniProt Consortium, (2015). UniProt: a hub for protein information. *Nucl. Acids Res.* 43 (Database issue), D204–D212. doi:10.1093/nar/gku989
- Valdes-Tresanco, M. S., Valiente, P. A., and Moreno, E. (2021). gmx\_MMPBSA: a new tool to perform end-state free energy calculations with GROMACS. *J. Chem. Theory Comput.* 17 (10), 6281–6291. doi:10.1021/acs.jctc.1c00645
- Vanommeslaeghe, K., Hatcher, E., Acharya, C., Kundu, S., Zhong, S., Shim, J., et al. (2010). CHARMM general force field: a force field for drug-like molecules compatible with the CHARMM all-atom additive biological force fields. *J. Comput. Chem.* 31 (4), 671–690. doi:10.1002/jcc.21367
- Wang, D., Wang, Z. J., Song, X. X., Pu, L. H., Li, X., and Wang, Y. (2013). Analysis of differentially expressed genes in various stages of Duchenne muscular dystrophy by using a network view. *Genet. Mol. Res.* 12 (4), 4480–4488. doi:10.4238/2013.October.10.13
- Wang, J., Fan, Q., Yu, T., and Zhang, Y. (2021). Identifying the hub genes for Duchenne muscular dystrophy and Becker muscular dystrophy by weighted correlation network analysis. *BMC Genom Data* 22 (1), 57. doi:10.1186/s12863-021-01014-w
- Wiche, G. (1998). Role of plectin in cytoskeleton organization and dynamics. *J. Cell Sci.* 111 (Pt 17), 2477–2486. doi:10.1242/jcs.111.17.2477
- Wu, X., Dong, N., Yu, L., Liu, M., Jiang, J., Tang, T., et al. (2022). Identification of immune-related features involved in Duchenne muscular dystrophy: a bidirectional transcriptome and proteome-driven analysis. *Front. Immunol.* 13, 1017423. doi:10.3389/fimmu.2022.1017423
- ww, P. D. B. c. (2019). Protein Data Bank: the single global archive for 3D macromolecular structure data. *Nucleic Acids Res.* 47 (D1), D520–D528. doi:10.1093/nar/gky949
- Zhang, P., He, J., Wang, F., Gong, J., Wang, L., Wu, Q., et al. (2019). Hemojuvelin is a novel suppressor for Duchenne muscular dystrophy and age-related muscle wasting. *J. Cachexia Sarcopenia Muscle* 10 (3), 557–573. doi:10.1002/jcsm.12414
- Zhou, L., and Lu, H. (2010). Targeting fibrosis in Duchenne muscular dystrophy. *J. Neuropathol. Exp. Neurol.* 69 (8), 771–776. doi:10.1097/NEN.0b013e3181e9a34b