Research Article

URF PUBLISHERS
connect with research world

Archives of Biotechnology and Pharmaceutical Research

https://urfpublishers.com/journal/biotech-pharma-research

Vol: 2 & Iss: 2

System-Level Transcriptomic and Structural Elucidation of NDUFA11 as Mitochondrial Nexus in Bovine Spongiform Encephalopathy

Nida[1], Ruqia Sartaj[1], Israr Hussain[1], Itazaz Ul Haq[2], Bilal Khan[3], Muhammad Rahiyab[1], Zahid Hussain[1], Syed Shujait Ali[1] and Arshad Iqbal[1]*

[1]Center for Biotechnology and Microbiology, University of Swat, KPK, Pakistan

[2]Sichuan Agricultural University, College of Life Science, Xinkang Road 46, Ya'an, 625014 P.R. China

[3]Center for Animal Sciences and Biodiversity, University of Swat, KPK, Pakistan

Citation: Nida, Sartaj R, Hussain I, et al. System-Level Transcriptomic and Structural Elucidation of NDUFA11 as Mitochondrial Nexus in Bovine Spongiform Encephalopathy. *Arch Biotech Pharma Res*, 2026;2(2):165-178.

Received: 28 April, 2026; **Accepted:** 12 May, 2026; **Published:** 15 May, 2026

***Corresponding author:** Arshad Iqbal, Center for Biotechnology and Microbiology, University of Swat, KPK, Pakistan, Email: arshad.iqbal@uswat.edu.pk

ABSTRACT

Bovine spongiform encephalopathy (BSE) is a fatal neurodegenerative disorder lacking well-defined molecular biomarkers and effective therapeutic targets. In this study, the gene expression dataset GSE69048 was retrieved from the Gene Expression Omnibus (GEO) database and analyzed using GEO2R to identify differentially expressed genes (DEGs). Protein–protein interaction (PPI) networks were constructed using STRING and visualized in Cytoscape to identify hub genes. The key gene, NDUFA11, was further subjected to three-dimensional structure prediction, refinement and validation. Active site analysis was performed using CASTp, followed by molecular docking against a library of compounds using EasyDock Vina. Pharmacokinetic and toxicity profiles were evaluated using SwissADME and ProTox-IIII. A total of 163 DEGs were identified, including 124 downregulated and 39 upregulated genes. Network analysis revealed NDUFA11 as the most significant hub gene associated with mitochondrial function. Structural modeling confirmed the stability and reliability of the predicted protein model. Molecular docking identified UCT1072M1 and Linagliptin as top candidate compounds, both exhibiting strong binding affinity (−9.4 kcal/mol) toward NDUFA11. ADMET analysis indicated favorable pharmacokinetic properties and acceptable toxicity profiles for both compounds. This study highlights NDUFA11 as a critical mitochondrial biomarker in BSE pathogenesis and proposes potential therapeutic candidates targeting this protein. These findings provide a foundation for future experimental validation and the development of targeted therapies for BSE.

Keywords: BSE, DEGs, NDUFA11, 3D structure modeling, Binding affinity, ADMET

1. Introduction

Bovine spongiform encephalopathy (BSE) is a fatal neurodegenerative disease in bovines and falls in the category of transmissible spongiform encephalopathies (TSEs) defined by the presence of disease-causing abnormal prion proteins (PrP) deposited in the bovine brain[1]. All initial cases of this disease were detected in the United Kingdom in 1987; however, retrospective analyses suggest that it may have been present as early as the 1980s, BSE has subsequently been reported in more than 25 countries worldwide, including nations across the Asian continent and North America[2]. The histopathological and clinical manifestation observed in affected bovines closely resembles those of transmissible spongiform encephalopathies (TSEs) reported in other species, such as scrapie in sheep and Creutzfeldt-Jakob disease in humans[3]. The etiology underlying all of these disorders is the pathological conversion in the normal cellular prion protein (PrP^C) into a misfolded, β-sheet rich conformation denoted as PrP^Sc, which accrues in the neural tissue and triggers neurodegeneration. Whereas inherited cases of prion disease have been linked to somatic mutation of the human prion protein gene (PRNP), no record of such mutation in cattle has been fully explored[4]. The etiology of the prion diseases relies on the fact that prion diseases are infectious in nature due to the capacity of the misfolded PrPSc to trigger conformational modifications in the normal prion proteins and thus induce the propagation of the malady into the central nervous system[5]. Although PrP^C and PrP^Sc share the same set of amino-acids, their functional and pathogenic behavior is determined by different conformations, 82-95 of PrP^C consists mainly of 82-95 amino-acids and PrP^Sc consists of predominantly 82-95 amino-acids[6].

Stanley Prusiner and his co-workers were the first to find that prions were the infectious agents causing TSEs and introduced the term prion to explain the nature of infectious agent and ultimately, he received the 1997 Nobel Prize in Physiology or Medicine[7]. The BSE is also characterized by various neurological and behavioral signs, clinically. These are sensitivity towards touch and sound, disturbance with motor, Anxiousness, ataxia, loss of weight and decrease production of milk, The characteristic clinical signs also include the loss of coordination and control over voluntary movements because of the dysfunction of the central nervous system progressively[8].

The bovine spongiform encephalopathy (BSE) has two subtypes, the atypical BSE that comprises:H-type (high molecular mass of non-glycosylated prion protein) and L type (low molecular weight of unglycosylated prion protein) or bovine amyloidotic spongiform encephalopathy (BASE). The most common by far is classical BSE that caused the extensive epidemic that arose in the 1980s in the United Kingdom. It has close linkage with the intake of tainted meat and bone meal (MBM) on the animals that are infected ruminants[9].

Conversely, the unusual BSE was first reported in 2004 in Italy (L-type) and France (H-types). They are believed to be random and mostly happen in aging cattle implying that they are spontaneous and do not originate in contaminated feed. It can be categorized as H-type and L-type on the basis of molecular mass of the unglycosylated form of the prion protein (PrP^Sc) that is generally found by Western blot method. One of the specificities of the L-type BSE is the presence of PrP amyloid plaques in the brain, which is not presented in classical BSE. Such a histopathological distinction confirms the adoption of the alternative nomenclature bovine amyloidotic spongiform encephalopathy to the L-type form. The cases of atypical BSE are not common, but they have been observed in different countries (Both in Europe and in the United States, Canada or Japan) causing concern about their worldwide distribution[10].

Although the information about the origin of classical BSE is interconnected with the improper feeding habits, the atypical forms of BSE do not seem to be produced under the influence of dieters, having arisen seemingly independently. However, a possible zoonotic risk with atypical BSE exists especially with the use of the bovine-derived products in animal feed and human foodstuffs[11].

Researchers initially became confused by the cause of bovine spongiform encephalopathy (BSE). Among the earliest theories, (outlined in 1992) was the possibility that the BSE epizootic in the United Kingdom might have originated in a prion disease of sheep, sheep scrapie, an endemic disease of sheep. Speculations to the effect that recurrent recycling of ruminant derived protein, such as cattle tissues and sheep tissue via meat and bone meal (MBM) as cattle feed, could had rendered possible the adjustment and enhancement of a scrapie-like agent, in the cattle, had been made[12].

There was no direct relation between scrapie and BSE as indicated by experimental result. Pathological and clinical characteristics of sheep scrapie differ with those of BSE in cattle[13]. When scrapie prions pass to cattle experimentally, the resulting transmissible spongiform encephalopathy (TSE) may have a different molecular and clinical phenotype as compared to classical BSE[14].

Recent advances in transcriptomics and systems biology have significantly improved the understanding of complex neurodegenerative diseases by enabling large-scale identification of dysregulated genes and molecular pathways. High-throughput gene expression profiling combined with network-based approaches has facilitated the discovery of key regulatory genes and biomarkers involved in disease progression. Moreover, structure-based drug design and molecular docking have emerged as powerful tools in neurodegeneration research, allowing the identification of potential therapeutic compounds through in silico screening and interaction analysis. Despite substantial progress in prion biology, the identification of mitochondrial-associated biomarkers and effective therapeutic targets in bovine spongiform encephalopathy (BSE) remains limited. Given the critical role of mitochondrial dysfunction in neurodegenerative disorders, there is a need to explore mitochondrial-linked genes that may contribute to BSE pathogenesis[15]. Therefore, this study integrates transcriptomic analysis and structural bioinformatics to identify key regulatory genes and potential therapeutic compounds targeting BSE, with a particular focus on the mitochondrial protein NDUFA11[16].

In the present study, we have adopted an integrative approach bioinformatics pipeline to Both Analyze the transcriptomics data of BSE patients to determine DEGs and Construct PPI networks and determine significant hub genes using Cytoscape and Cytohubba[17] and Predict the 3D structure and validated the stereochemical stability of NDUFA11, we also screened candidate small molecules through molecular docking and

Evaluated the pharmacokinetics and toxicity of the top molecules to drive viable therapeutic candidates as shown in **(Figure 1)**.

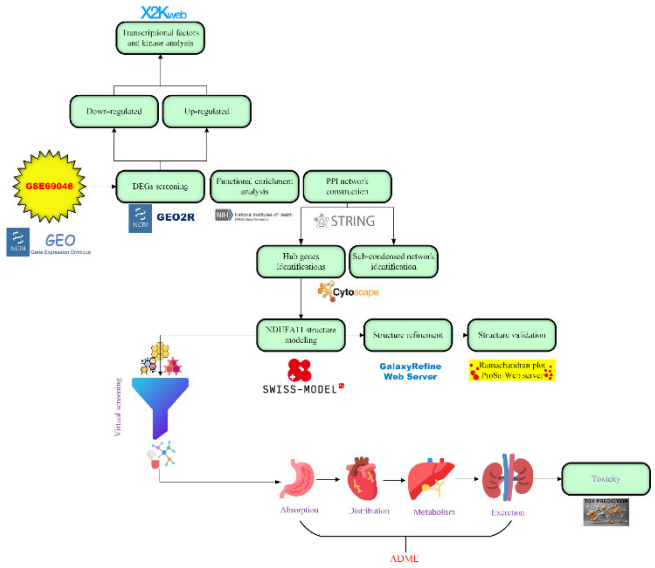

Figure 1: The schematic representation illustrates the stepwise methodology employed in this study

2. Materials and Methods

2.1. Dataset collection

The gene expression dataset GSE69048 was retrieved through NCBI gene expression omnibus (GEO) database (https://www.ncbi.nlm.nih.gov/geo/) that is a publicly available data store and distributed freely, containing comprehensive data of gene expression information submitted by research community[18]. The specified dataset consists of expression profiles that were obtained based on peripheral blood samples procured in cattle. It involves two control animals, non-infected and eight animals infected by bovine spongiform encephalopathy (BSE) that have been equally represented using H-type (n = 4) and L-type (n = 4) subtypes of BSE. The Biological Archive Group of the Animal and Plant Health Agency (APHA) in United Kingdom, generously gave these to us. The blood was sampled not once but twice at two points of disease progression i.e. before and after the disease i.e. preclinical and clinical respectively making a total of sixteen (eight at each stage) samples[19].

2.2. Preparing and analyzing DEGs data

GEO2R (http://www.ncbi.nlm.nih.gov/geo/geo2r/) is a web-based interactive toolkit that allows one to perform comparisons between the experimental conditions in GEO Series datasets. In order to determine differentially expressed genes (DEGs) between BSE-infected and control cattle, a publicly available microarray dataset GSE69048 was downloaded and analyzed by using GEO2R. The p-value required to find statistical significance was ≥ 0.05 wherein the log2 fold change (Log2FC) cutoff was set to 1. Genes with Log2FC > 1 were upregulated and the ones with Log2FC ≤ 1 were labeled downregulated[20]. A heatmap depicting the expression patterns of the identified differentially expressed genes was generated using the Morpheus web application (https://software.broadinstitute.org/morpheus/)[21].

2.3. Go pathway and functional enrichment analysis of DEGs

In genetic studies, large-scale gene expression data are essential and must be systematically analyzed and interpreted to identify underlying biological processes. Gene Ontology (GO) and functional enrichment analysis represent widely used approaches for this purpose. GO provides a structured, hierarchical framework for gene classification and is particularly useful for the annotation and interpretation of large gene lists[22]. Furthermore, Kyoto Encyclopedia of Genes and Genomes (KEGG) is a very complete database linking genes to currently described biochemical pathways, which facilitates the studies of molecular pathways and cellular networks[23].

The Database for Annotation, Visualization and Integrated Discovery (DAVID) (https://david.ncifcrf.gov/) was utilized to perform integrative functional annotation of differentially expressed genes. DAVID provides a comprehensive suite of tools for functional annotation and enrichment analysis. Gene Ontology (GO) enrichment analysis was conducted across three categories: biological processes (BP), cellular components (CC) and molecular functions (MF). In addition, relevant signaling and metabolic pathways were identified through Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment analysis[24]. The results of enrichment with p-values less than 0.05 were regarded as statistic. The biological relevance of upregulated and downregulated DEGs was illustrated by bubble plot constructed through SRplot software (https://www.bioinformatics.com.cn/en) presenting the enriched GO terms and KEGG pathways[25].

2.4. PPI (protein-protein interaction) network construction

Protein-protein interactions (PPI) networks give a detailed structure to display the physical and functional connections between proteins inside a cell or organism. PPI networks construction is core to unravel the molecular pathway in Bovine Spongiform Encephalopathy (BSE) of some fundamental cellular mechanisms. This was done in this paper using the STRING database (http://string.embl.de/) which is a powerful online database that constitutes both experimentally and computationally inferred protein interactions of a broad range (including both direct physical interactions as well as indirect functional associations[26]. The resulting PPI network were visualized via the Cytoscape software application and a confidence score of 0.4 and above cut-off was applied to confirm the consistency of interactions. Hub genes which were involved in pathogenesis of BSE were identified by the connectivity degree with the highest 10 connectivity degrees being considered as key regulatory nodes in the network.

2.5. Sub-condense network selection from PPI network

In computer science, network analysis, ways of identifying separate subnetworks in a larger network are essential in explaining the functional relationship between proteins. To this respect, Molecular Complex Detection (MCODE) plugin of Cytoscape was used to identify strongly interconnected parts of the PPI network. The default values of parameters, degree cutoff= 2, node score cutoff = 0.2, k-core = 2 and maximum depth 100 has been used[27].

2.6. Identification of hub genes

In genomics, there is a need to identify important genes, those performing crucial functions in biological networks, both to understand the mechanism of disease and to reveal therapeutic targets. Cytoscape is a platform that allows the assessment of hub genes in generated protein-protein interaction (PPI) networks via the use of a package called Cytohubba[28].The utility

relies upon eleven different topological algorithms to evaluate the usefulness of nodes and to pinpoint center hub genes in the network. In this respect, we used Cytohubba to find hub genes which could be some new targets in treating Bovine Spongiform Encephalopathy (BSE)[28].

2.7. Association between Transcription element and regulatory network

To clarify the connection between the regulatory factors and the gene expression, it is critical that transcription factor (TFs) and kinases that take part in these regulatory networks be identified and characterized. The relations between essential regulatory factors and the significant genes, therewith, were discovered based on the Expression2Kinases (X2K) platform (https://maayanlab.cloud/X2K/). The entire list of the differentially expressed genes (DEGs) was provided and certain gene markers were uploaded into the X2K tool to be analyzed. The Fisher exact test p-values allowed generating the top 10 most significant TFs and kinases. Next, a regulatory network was created and the resulting file with the denomination of "graphml" was visualized using Cytoscape. This control mechanism achieves adequate interconnection between edge nodes, which allows thorough examination of the network topology as it advances.

2.8. Structure prediction, refinement and validation of NDUFA11

Protein structure prediction, refinement and validation refer to the mechanization of more sophisticated computation tools to model and streamline the structural dictionaries of proteins in three-dimension (3D) forms justified through the sequences of their amino acids. In the prediction of the structure, we applied DMPfold2, a web-based tool that predicts the tertiary structure of single chains of proteins by relying on the amino acid sequence[29]. DUFA11 was downloaded as amino acids sequence in the UniProt database and uploaded in DMPfold2 to model it. The DMPfold2 has better accuracy and computational efficiencies than its past DMPfold version[29]. The resultant 3D model created by DMPfold2 was then improved with the help of GalaxyRefine web server (http://galaxy.seoklab.org/cgi-bin/submit.cgi?type=REFINE)[30]. GalaxyRefine uses molecular dynamics simulations to reconstruct side chains, repack them and relax the whole structure, this way improving the local quality of structures. In the CASP10 test, this approach had been identified as the refinement method available with the highest improvement. The quality of the refined model was measured with the help of several different scores such as GDT-HA, RMSD, the MolProbity score, the clash score and the Ramachandran plot statistics. The entire structural validation was carried out with MolProbity (http://molprobity.biochem.duke.edu/). MolProbity gathers knowledge-based and physics-based analysis with the most notable contribution being an all-atom contact analysis performed using Probe. The instrument copes with structural models of proteins in PDB form, which can be uploaded or found in the Nucleic Acid Database (NDB) or the Protein Data Bank (PDB) thru an appropriate designator[31]. MolProbity uses quality-filtered empirical Ramachandran distributions based on about 100,000 residues in the top 500 database as tested against a quality-filtered. distribution based on about 100,000 residues in the top 500 database and scores structural outliers in both the main and side chain[30].

2.9. Active sites prediction of NDUFA11

The CASTp (Computed Atlas of Surface Topography of Proteins) server was utilized to project any active loci of NDUFA11. CASTp is a web-based analyzer that is available at (http://sts.bioe.uic.edu/)[32], determines the topography of the protein surface. The topography topographical analysis of the protein structure in PDB format accompanied by the given radius of the probe was uploaded to the CASTp server to be thoroughly analyzed. CASTp identifies each atom which forms the surface pockets, internal cavity and channel in the protein structure in detail. It also computes the accurate area and volume of such features, the size of any openings that accompany them. Results of the CASTp analysis may be downloaded and the data displayed in PyMOL with the CASTpyMOL plugin or the UCSF Chimera software package.

2.10. Ligands selection and molecular docking analysis

Selection of ligands and molecular docking study is a very important part of drug target in drug discovery process, as this method can help to select possible drug candidates and player interaction specificity among them. In the present research, the natural compound library consisted of 3,368 ligands extracted out of the PubChem database[33]. The ligands were screened virtually against NDUFA11 protein by EasyDockVina, which is an online tool (high-throughput receptor-ligand docking based on Auto Dock Vina[34].

Docking workflow presupposed subsequent preparation of ligands and the target protein. The ligand and protein structures were translated into the PDBQT format according to the needs of Auto Dock Vina and then were uploaded to EasyDockVina server. Group box sizes required to mark out the binding site were obtained through PyRx and put into EasyDockVina to undertake the docking procedure[35]. 3,368 ligands were effectively docked with the molecule of NDUFA11[34]. The two ligands were chosen on the basis of their scores of affinities and the scoring of binding energies; the highest two ligands were chosen. Interaction plots to illustrate the protein ligand interactions were also generated using Discovery studio giving a clear understanding of the interactions between the ligands and the protein that hold the docked complexes.

2.11. In- silico ADMET analysis

In silico ADMET analysis The Absorption, Distribution, Metabolism, Excretion and Toxicity (In silico ADMET analysis) properties are those of a compound that are critical in determining its drug-likeness and pharmacokinetic profile, thus they can be predicted and scored in silico. In this research work, Online ADMET predictive tool SwissADME was used to carry out ADMET profiling of selected substances. The compounds were selected by searching their canonical-SMILES (Simplified Molecular Input Line Entry System) from PubChem database and dispatched to SwissADME service to analyze physicochemical properties of the compounds and physicochemical parameters[36]. ADMET characterization has become a strategic part of drug discovery and development because it allows selecting compounds with good pharmacological properties in the early stages before the project incurs the risk of a late-stage stumble.

2.12. Toxicity evaluation

Toxicity is a very important measure during drug design

since it assists in determining the safety characteristic of the drug of interest. In the present study, the toxicity of ligands of interest was modeled with ProTox-IIIII (https://tox.charite.de), a powerful online tool aimed at estimating toxicity of a compound, considering toxic structures[37]. ProTox-IIIII gives the reported value of estimated median lethal dose (LD 50) and may categorize the compounds into six classes of toxicity (Class I toxic, Class VI non-toxic). The toxicity endpoints and organ toxicity analysis were the main concerns of the group. This prediction assessment facilitates early detection of possibly dangerous compounds hence enhancing the efficiency and safety of drugs development.

3. Results

3.1. Data set collection

Analysis of the GSE69048 dataset was conducted using publicly available data retrieved from the Gene Expression Omnibus (GEO) database. This dataset was selected based on its relevance and the availability of comprehensive gene expression profiles. Detailed information regarding the characteristics of each sample, including experimental conditions, sample type and associated metadata, is summarized in **(Table 1)** to facilitate reproducibility and transparency of the analysis.

Table 1: GSE69048 dataset Sample information retrieved from the GEO database.

Group	Accession number	Title	Organism	Source name
Patient	GSM1691208	P1 blood_preclinical_H type	Bos taurus	Aberdeen angus
Patient	GSM1691209	P2 blood_preclinical_H type	Bos taurus	Aberdeen angus
Patient	GSM1691210	P7 blood_preclinical_H type	Bos taurus	Aberdeen angus
Patient	GSM1691211	P8 blood_preclinical_H type	Bos taurus	Aberdeen angus
Patient	GSM1691212	S1 blood_clinical_H type	Bos taurus	Aberdeen angus
Patient	GSM1691213	S2 blood_clinical_H type	Bos taurus	Aberdeen angus
Patient	GSM1691214	S7 blood_clinical_H type	Bos taurus	Aberdeen angus
Patient	GSM1691215	S8 blood_clinical_H type	Bos taurus	Aberdeen angus
Patient	GSM1691216	P4 blood_preclinical_L type	Bos taurus	Aberdeen angus
Patient	GSM1691217	P5 blood_preclinical_L type	Bos taurus	Aberdeen angus
Patient	GSM1691218	EP9 blood_preclinical_L type	Bos taurus	Aberdeen angus
Patient	GSM1691219	P10 blood_preclinical_L type	Bos taurus	Aberdeen angus
Patient	GSM1691220	S3 blood_clinical_L type	Bos taurus	Aberdeen angus
Patient	GSM1691221	S4 blood_clinical_L type	Bos taurus	Aberdeen angus
Patient	GSM1691222	S9 blood_clinical_L type	Bos taurus	Aberdeen angus
Patient	GSM1691223	S10 blood_clinical_L type	Bos taurus	Aberdeen angus
Control	GSM1691224	c.P3blood_noinoculated_control	Bos taurus	Aberdeen angus
Control	GSM1691225	c.P6blood_noinoculated_control	Bos taurus	Aberdeen angus
Control	GSM1691226	c.S6blood_noinoculated_control	Bos taurus	Aberdeen angus
Control	GSM1691227	c.9blood_no inoculated control	Bos taurus	Aberdeen angus
Control	GSM1691228	c.5blood_no inoculated control	Bos taurus	Aberdeen angus
Control	GSM1691229	c.2blood_not inoculated control	Bos taurus	Aberdeen angus
Control	GSM1691230	c.3blood_not inoculated control	Bos taurus	Aberdeen angus

3.2. Differentially expressed genes identification

The GEO2R tool was used in this study to determine the differentially expressed genes (DEGs) of BSE patients compared to healthy controls. The significant genes were identified with 163 significant genes using the cut off criteria of |log 2-fold change (log 2FC) | and p-value < 0.05. Out of these, 124 were downregulated whereas 39 upregulated. The need to visualize gene expression patterns required the creation of heat maps that depict top 20 upregulated and downregulated genes with the help of the tool Morpheus **(Figure 2)**. Also, the bar plot reflecting the distribution of DEGs was presented with the help of SR Plot, which is show in (**Figure 3**).

3.3. Enrichment analysis of DEGs using GO and KEGG Pathway

Functional annotation of the identified differentially expressed genes (DEGs) is essential for understanding their biological significance and for elucidating the underlying molecular mechanisms and pathways. To achieve this, Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment analyses were performed using the Database for Annotation, Visualization and Integrated Discovery (DAVID).

GO enrichment analysis of the upregulated DEGs revealed significant enrichment in several biological processes (BP), including negative regulation of cell population proliferation (GO:0008285), adherens junction assembly (GO:0034333), negative regulation of anoikis (GO:2000811), heterotypic cell–cell adhesion (GO:0034113) and protein localization to the cell surface (GO:0034394). Cellular component (CC) analysis indicated that these genes were primarily associated with the nucleoplasm (GO:0005654), cell–cell contact zone (GO:0044291) and focal adhesion (GO:0005925). Furthermore, molecular function (MF) analysis demonstrated enrichment in cell–cell adhesion mediator activity (GO:0098632), cysteine-

type endopeptidase activity (GO:0004197), protein binding (GO:0005515) and scavenger receptor activity (GO:0005044). KEGG pathway analysis further revealed that upregulated DEGs were significantly involved in leukocyte transendothelial migration (bta04670), cell adhesion molecules (CAMs) (bta04514) and tight junction (bta04530) pathways.

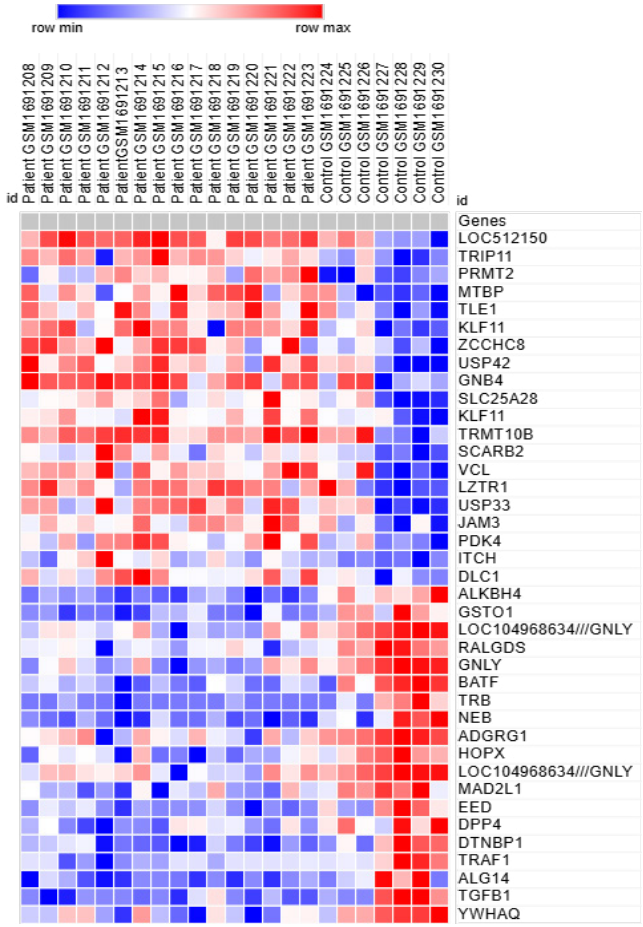

Figure 2: Heat map for the top 20 up and down-regulated gene constructed using the Morpheus tool.

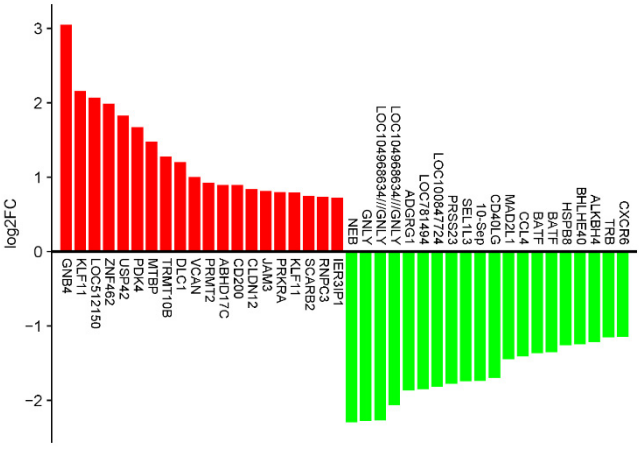

Figure 3: Red bars indicate up-regulated genes and green bars indicate down-regulated genes. The valves given on Y-axis are LogFC valves.

In contrast, GO enrichment analysis of the downregulated DEGs showed significant enrichment in biological processes such as mitochondrial respiratory chain complex I assembly (GO:0032981), protein export from the nucleus (GO:0006611), cytolysis of cells from other organisms (GO:0031640), protein export (GO:0009306) and phosphate-containing compound metabolic processes (GO:0006796). Cellular component analysis highlighted enrichment in mitochondrial respiratory chain complexes, including cytolytic granules (GO:0044194), mitochondrial respiratory chain complex IV (GO:0005751), mitochondrion (GO:0005739), cytoplasm (GO:0005737) and cytosol (GO:0005829). Molecular function analysis indicated enrichment in protein serine/threonine kinase activator activity (GO:0043539). KEGG pathway enrichment analysis demonstrated that downregulated DEGs were associated with several metabolic and disease-related pathways, including chemical carcinogenesis-reactive oxygen species (bta05208), oxidative phosphorylation (bta00190), diabetic cardiomyopathy (bta05415), Parkinson's disease (bta05012), non-alcoholic fatty liver disease (bta04932), pathways of neurodegeneration (bta05022) and Alzheimer's disease.

A comprehensive summary of enriched GO terms for upregulated and downregulated DEGs is presented in **(Table 2 and Table 3)**, respectively. Bubble plots illustrating GO term and KEGG pathway enrichment were generated using SRplot. **(Figures 4a and 4b)** represent GO terms and KEGG pathways for downregulated genes, respectively, whereas **(Figures 5a and 5b)** depict the corresponding enrichment results for upregulated genes.

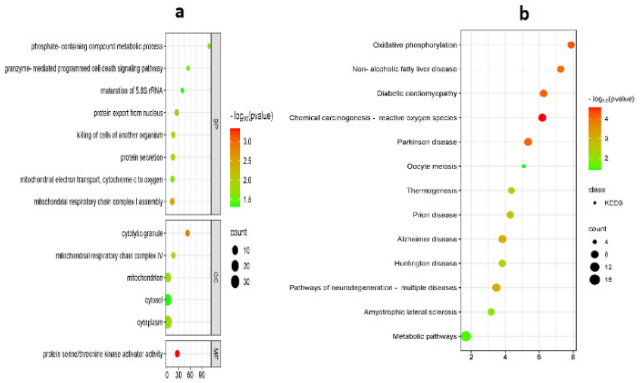

Figure 4: Functional annotation of downregulated genes. GO enrichment analysis, including BP, MF and CC, is shown in bubble plot (a), while KEGG pathway enrichment is presented in bubble plot (b). Bubble size indicates gene count and color represents statistical significance.

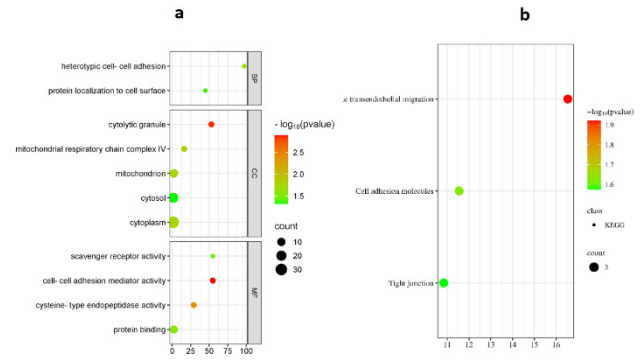

Figure 5: Functional annotation of upregulated genes. GO enrichment analysis, including BP, MF and CC, is shown in bubble plot (a), while KEGG pathway enrichment is presented in bubble plot (b). Bubble size represents gene count and color indicates statistical significance.

Table 2: Annotated result of GO terms and KEGG pathways for up - regulated genes.

Category	Term	Count	%	PValue	Fold Enrichment	FDR
GOTERM_BP_DIRECT	GO:0008285~negative regulation of cell population proliferation	4	10.52631579	0.002589	14.12502165	0.5488362
GOTERM_BP_DIRECT	GO:0034333~adherens junction assembly	2	5.263157895	0.013263	145.6642857	1
GOTERM_BP_DIRECT	GO:2000811~negative regulation of anoikis	2	5.263157895	0.014908	129.4793651	1
GOTERM_BP_DIRECT	GO:0034113~heterotypic cell-cell adhesion	2	5.263157895	0.01983	97.10952381	1
GOTERM_BP_DIRECT	GO:0034394~protein localization to cell surface	2	5.263157895	0.042482	44.81978022	1
GOTERM_CC_DIRECT	GO:0005654~nucleoplasm	8	21.05263158	0.015663	2.896466212	0.76889901
GOTERM_CC_DIRECT	GO:0044291~cell-cell contact zone	2	5.263157895	0.024709	77.86666667	0.76889901
GOTERM_CC_DIRECT	GO:0005925~focal adhesion	3	7.894736842	0.027138	11.37662338	0.76889901
GOTERM_MF_DIRECT	GO:0098632~cell-cell adhesion mediator activity	3	7.894736842	0.00128	54.85950413	0.11651459
GOTERM_MF_DIRECT	GO:0004197~cysteine-type endopeptidase activity	3	7.894736842	0.004454	29.19941349	0.20267591
GOTERM_MF_DIRECT	GO:0005515~protein binding	9	23.68421053	0.025346	2.371655419	0.76881706
GOTERM_MF_DIRECT	GO:0005044~scavenger receptor activity	2	5.263157895	0.03478	54.85950413	0.79124296
KEGG_PATHWAY	bta04670: Leukocyte transendothelial migration	3	7.894736842	0.012002	16.53728814	0.46152661
KEGG_PATHWAY	bta04514: Cell adhesion molecules	3	7.894736842	0.023675	11.54674556	0.46152661
KEGG_PATHWAY	bta04530: Tight junction	3	7.894736842	0.026627	10.84111111	0.46152661

Table 3: Annotated result of GO terms and KEGG pathways for down-regulated genes.

Category	Term	Count	%	Pvalue	Fold Enrichment	FDR
GOTERM_BP_DIRECT	GO:0032981~mitochondrial respiratory chain complex I assembly	4	3.669724771	0.003269	13.35057283	1
GOTERM_BP_DIRECT	GO:0006611~protein export from nucleus	3	2.752293578	0.006227	25.03232406	1
GOTERM_BP_DIRECT	GO:0031640~killing of cells of another organism	3	2.752293578	0.015032	15.87415672	1
GOTERM_BP_DIRECT	GO:0009306~protein secretion	3	2.752293578	0.016456	15.13582385	1
GOTERM_BP_DIRECT	GO:0006796~phosphate-containing compound metabolic process	2	1.834862385	0.018118	108.4734043	1
GOTERM_BP_DIRECT	GO:0006123~mitochondrial electron transport, cytochrome c to oxygen	3	2.752293578	0.018696	14.1487049	1
GOTERM_BP_DIRECT	GO:0140507~granzyme-mediated programmed cell death signaling pathway	2	1.834862385	0.035912	54.23670213	1
GOTERM_BP_DIRECT	GO:0000460~maturation of 5.8S rRNA	2	1.834862385	0.049048	39.44487427	1
GOTERM_CC_DIRECT	GO:0044194~cytolytic granule	3	2.752293578	0.001385	52.96078431	0.210491243
GOTERM_CC_DIRECT	GO:0005751~mitochondrial respiratory chain complex IV	3	2.752293578	0.014325	16.29562594	0.673228665
GOTERM_CC_DIRECT	GO:0005739~mitochondrion	11	10.09174312	0.017276	2.356192629	0.673228665
GOTERM_CC_DIRECT	GO:0005737~cytoplasm	30	27.52293578	0.017717	1.498890122	0.673228665
GOTERM_CC_DIRECT	GO:0005829~cytosol	20	18.34862385	0.047696	1.553102179	1
GOTERM_MF_DIRECT	GO:0043539~protein serine/threonine kinase activator activity	4	3.669724771	4.59E-04	26.20263158	0.078886726
KEGG_PATHWAY	bta05208: Chemical carcinogenesis - reactive oxygen species	10	9.174311927	2.53E-05	6.183143219	0.003397624
KEGG_PATHWAY	bta00190: Oxidative phosphorylation	8	7.339449541	5.67E-05	7.884444444	0.003397624
KEGG_PATHWAY	bta05415: Diabetic cardiomyopathy	9	8.256880734	7.50E-05	6.254487179	0.003397624
KEGG_PATHWAY	bta05012: Parkinson disease	10	9.174311927	7.84E-05	5.349232456	0.003397624
KEGG_PATHWAY	bta04932: Non-alcoholic fatty liver disease	8	7.339449541	9.49E-05	7.267783985	0.003397624

KEGG_PATHWAY	bta05022: Pathways of neurodegeneration - multiple diseases	11	10.09174312	9.18E-04	3.480123217	0.023604868
KEGG_PATHWAY	bta05010: Alzheimer disease	10	9.174311927	9.23E-04	3.835298742	0.023604868
KEGG_PATHWAY	bta05020: Prion disease	8	7.339449541	0.00222	4.293509351	0.049681035
KEGG_PATHWAY	bta05016: Huntington disease	8	7.339449541	0.004215	3.82627451	0.083829928
KEGG_PATHWAY	bta04714: Thermogenesis	7	6.422018349	0.004757	4.361366539	0.08514471
KEGG_PATHWAY	bta05014: Amyotrophic lateral sclerosis	8	7.339449541	0.011387	3.17300813	0.185294498
KEGG_PATHWAY	bta01100: Metabolic pathways	17	15.59633028	0.031247	1.683607389	0.466102623
KEGG_PATHWAY	bta04114: Oocyte meiosis	4	3.669724771	0.042126	5.081770833	0.580039977

3.4. PPI network construction

STRING database was used to study the functional and physical associations between the proteins coded by differentially expressed genes (DEGs). The cutoff value of the interaction confidence was adjusted to 0.4 to eliminate the insignificant protein-protein interactions. The interaction network produced was then built in STRING and then imported in Cytoscape to obtain a better visualisation **(Figure 6)**. The proteins in the network are denoted by nodes in the network with the edges being prediction or known interactions. The resultant network consisted of 145 nodes and 146 edges revealing that the number of interactions is considerably greatest than what would be possible with a random set of proteins of the same size. The analyses of network topology also indicated that the average node degree of the network was 2.01, the average local clustering coefficient was 0.385 and the p-value of PPI enrichment was highly significant (4.26 x 10 -6), meaning that there is strong functional connection among the proteins in question.

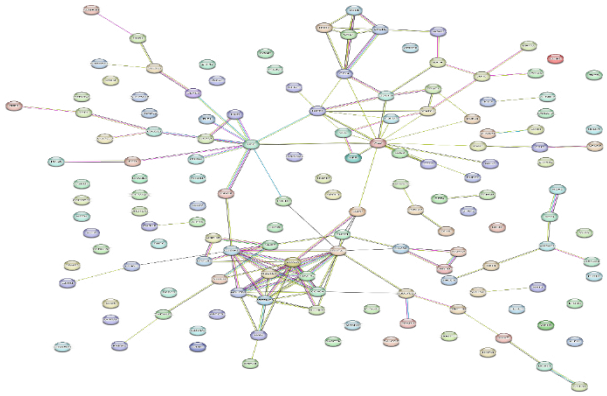

Figure 6: Protein–protein interaction (PPI) network visualization.

The network illustrates interactions among the identified proteins, where nodes represent proteins and edges indicate their interactions.

3.5. Sub condense network identification

The MCODE (Molecular Complex Detection) plugin in Cytoscape was used to identify clusters that denote neighborhoods of highly connected nodes in protein-protein interaction (PPI) network. Two large clusters of MCODE scores 7 and 2 were picked for the further analysis **(Figures 7a, 7b)**. Cluster 1 that was focused on NDUFA11 consisted of 8 nodes and 28 edges which implies densely connected subnetwork. Cluster 2 was a cluster created in CTSB and it had 4 nodes and 5 edges. Those clusters are probably functional molecular complexes and could be one of the factors in the development and progression of the Bovine Spongiform Encephalopathy (BSE)[38].

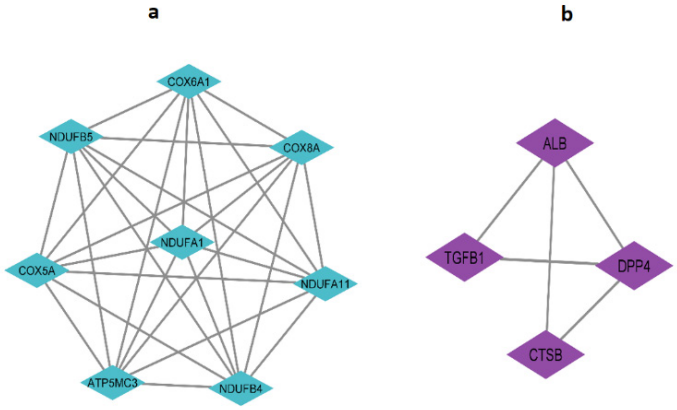

Figure 7: Subnetwork identified from the protein–protein interaction (PPI) network using the Cytoscape plugin MCODE.

3.6. Hub genes identification

It is important in the analysis of protein-protein interaction (PPI) networks because hub genes tend to play central regulatory functions in disease pathways. In the research, it used the CytoHubba plugin in the Cytoscape to demonstrate important hub genes in the developed PPI network. The top 10 hub genes were identified with the degree centrality approach that identified the genes with the highest connectivity score **(Figure 8)**. NDUFA11 was the most outstanding hub gene with the degree score of 13. ALB, COX5A, UBC, COX6A1, TGFB1, ATP5MC3, NDUFB5, COX8A and NDUFA1 are other hub genes and the network connectivity of them is big. These are probably genes that are very crucial in the molecular bases of Bovine Spongiform Encephalopathy (BSE). The detailed details of these hub genes are in shown **(Table 4)**.

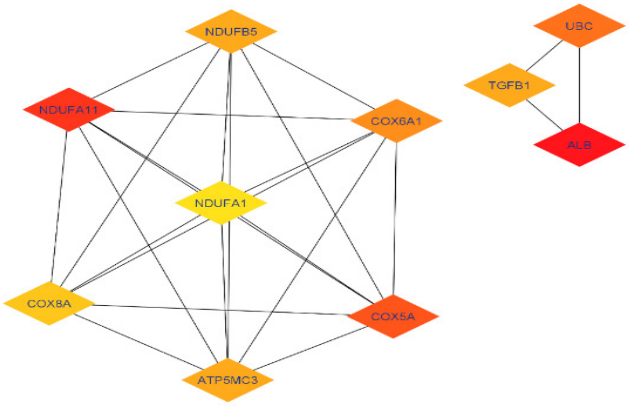

Figure 8: Network of the top 10 hub genes identified using the Cytoscape plugin CytoHubba.

Table 4: Top 10 Hub-genes identified in the resultant PPI Network.

Rank	Name	Score	P valve
1	ALB	15	8.52E-04
2	NDUFA11	13	1.03E-03
3	COX5A	12	1.52E-03
4	UBC	11	1.13E-03
5	COX6A1	10	1.44E-03
6	TGFB1	9	5.11E-04
6	ATP5MC3	9	5.11E-03
6	NDUFB5	9	1.35E-03
9	COX8A	8	5.23E-03
10	NDUFA1	7	5.91E-03

3.7. Transcription factors analysis

In this study, the major transcription factors (TFs) and protein kinases were related to the various DEGs, with an overall purpose of showing the importance of the regulation of molecular pathways playing a partial role in disease progression[39]. These regulatory factors together with intermediate proteins helped to create intricate transcriptional and signaling pathway. The transcriptional regulator prediction in the analysis started by integrating ChIP-seq-derived gene targets by using ChEA (ChIP-X Enrichment Analysis) database. This was followed by the mapping of the most pertinent TFs onto protein-protein interaction (PPI) network to determine their regulatory relationships. Regarding hypergeometric p-values, the most significant TFs were RUNX1, TAF1, ELF1, PML and MYC **(Figures 9a, 9b)**. Besides, the highest-ranking protein kinases were determined and placed on the PPI network as well. Kinases of the greatest regulatory potential according to hypergeometric p-values were CSNK2A1, CDK1, MAPK8, CDK2 and MAPK14 **(Figures 9c, 9d)**. These results indicate that the specified TFs and kinases can be used as possible upstream regulators/drug targets of Bovine Spongiform Encephalopathy (BSE).

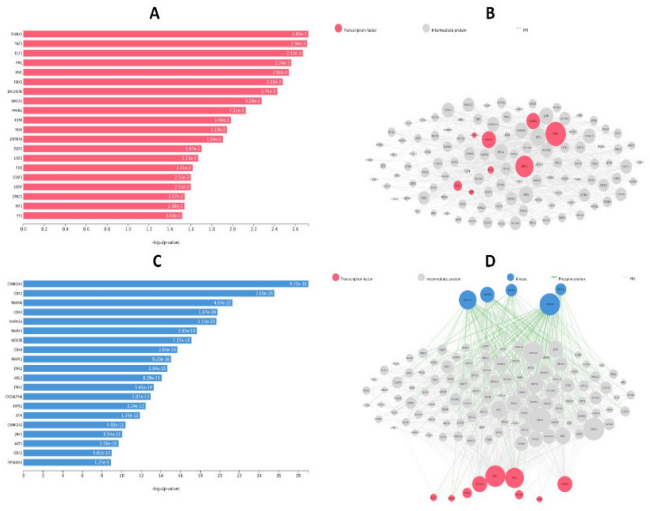

Figure 9: (a) and (b) depict transcription factors (TFs) identified from the analysis, while (c) and (d) represent the predicted protein kinases associated with the input gene set.

3.8. Structure prediction, refinement and validation of NDUFA11

The prediction of the three-dimensional (3D) structure of NDUFA11 was done with the help of DMPfold2 and the presented model was downloaded in PDB format to be used in further tests. Galaxy Refine web server was used to improve the accuracy of the initial model by refining, which yielded five refined models. Out of these, Model 1 was picked in virtue of its better-quality measures. The score of Global Distance Test - High Accuracy (GDT-HA) that defines similarity in structure among the refined and the original model was 0.9007, which defines great degree of similarity. The root means square deviation (RMSD) of changes between atoms positions was 0.578 A1, which is well below what would be accepted (01.2 A1) and it therefore indicates that the structure refined is stable. Further validation consisting of MolProbity additional validation parameters validated the quality of the model. The MolProbity score which together with the crystallographic resolution is a measure of model accuracy improved to 1.829 indicating that there will be fewer structural errors. The clash impact, which reflects the negative overlaps of atoms, reduced to 21.7, which illustrates the increased structural solidness[40]. Further, Ramachandran plot gave 98.6 percent of residues situated in energetically positive areas among which the conventionally accepted value is 85 percent and 0.0 percent of the poor rotamer outliers, which corresponded to the proper backbone and side-chain conformations.Additional verification by MolProbity indicated that 97.1 percent (135/139) of the residues were found in the preferred guide and 99.3 percent (138/139) in the allowed projections of the Ramachandran plot **(Figure 10)**, which confirms the fact that the refined model is of high structural quality in general.

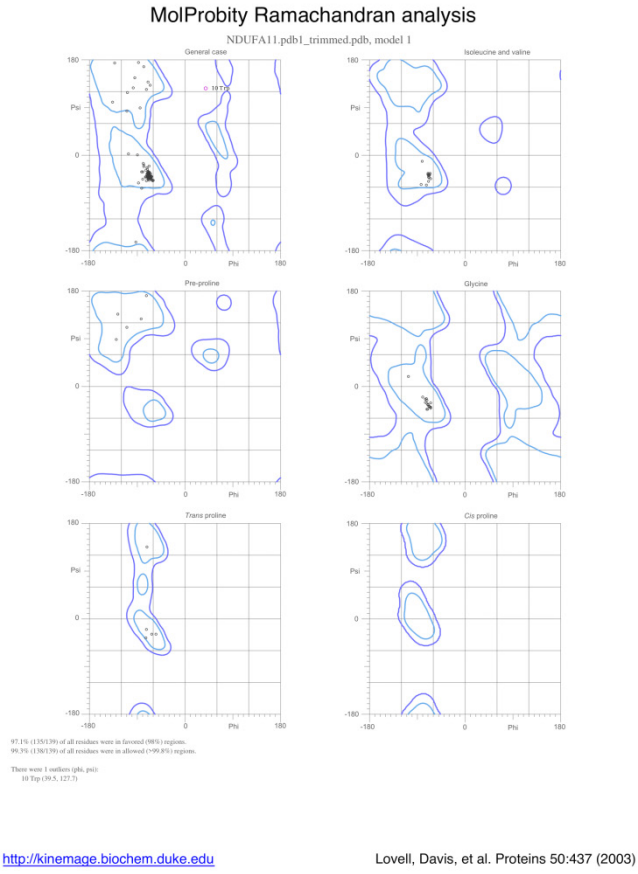

Figure 10: Ramachandran plot analysis of NDUFA11.

The plot illustrates the stereochemical quality of the predicted NDUFA11 protein structure, showing the distribution of amino acid residues in favored, allowed and disallowed regions of the phi (φ) and psi (ψ) torsion angles.

3.9. Active site evaluation of NDUFA11

Construction of an active site in NDUFA11 was carried out under default parameters in CASTp server and a probe radius of 1.4 AA. It was found that there are 3 different binding pockets and assessing them in terms of surface area and volume they can be chosen as ligand-binding sites in the further docking-related research. Of them, Pocket 1 was found to be the most accessible and presumably binding pocket because it was more exposed to the surface than Pocket 2 and Pocket 3.In fact, Pocket 1 had a surface area of 727.777 A 2 and the volume of 925.577 A 3. The residues LYS21, ALA24, THR25, ILE28, GLY29, ALA32, GLY33, VAL35, SER37, TYR39, THR56, TYR59, THR60, THR62, ALA63, ILE66, GLY67, PHE70, THR73, SER74, ALA78, LYS83, PRO84, ASN89, TYR90, GLY93, GLY94, GLY97,LEU101, ARG104, TYR117,MET118, THR121, ALA122, VAL125, LYS126, GLN129, GLN134, VAL135, PHE136, GLU138, PRO139 and VAL141.Comparatively, Pocket 2 and 3 were seen to have insignificantly smaller dimensions with surface area of 108.751 A 2 and 53.029 A 2 and volumes of 47.007 A and 14.762 A, respectively. Since it is much larger in size and with wide surface accessibilities, Pocket 1 has been picked as the main candidate to undergo molecular docking experiments[41].

3.10. Molecular docking analysis

In the case of molecular docking study, EasyDockVina was used to screen a library of 3,368 ligands against a target protein NDUFA11. Out of the five ligands the two best candidates as potential drug candidates were the molecules UCT1072M1 and Linagliptin which had a binding energy of -9.4 kcal/mol. Such values of the binding energies denote typical and positive interactions between NDUFA11 and ligands **(Figure 11a, 11b and Figure 12a, 12b)**.

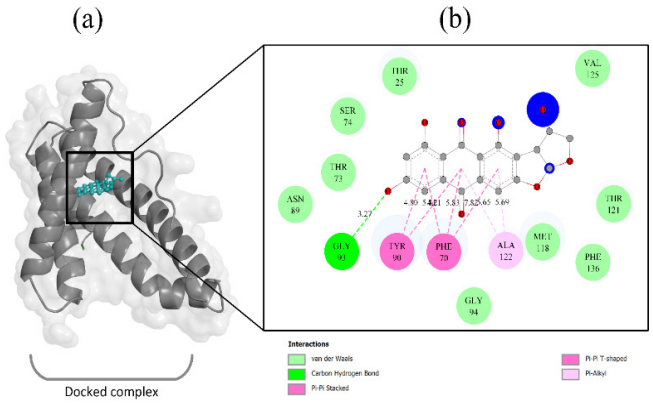

Figure 11: Docked complex visualization of NDUFA11 with UCT1072M1. The figure shows the molecular docking interaction between NDUFA11 and UCT1072M1, highlighting the binding pose and key interacting residues within the active site of the protein.

3.11. In- silico ADMET analysis of ligands

ADMET study is a sensitive and complex phase in drug discovery as well as development that involves assessment and review of properties; Absorption, Distribution, Metabolism, Excretion and Toxicity, which all collectively generate the pharmacokinetics of the candidate compounds. In this work, the properties of the selected ligands, pharmacokinetically, were evaluated with the help of such an online web tool as SwissADME that is proven to be trustworthy[42].

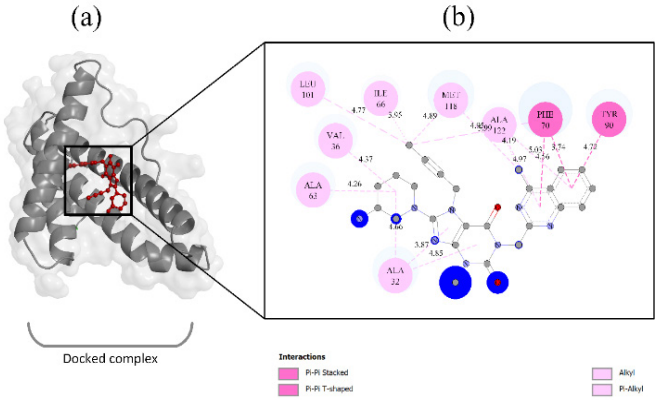

Figure 12: Docked complex of NDUFA11 with linagliptin visualized using molecular docking analysis.

In the case of UCT1072M1, ligand was found to have a molecular formula C18H12O8 with molecular weight of 356.28 g/mol. Its molar refractivity was 84.91 and the topological polar surface area (TPSA) was 133.52 A 2. This compound has 12 aromatic heavy atoms and no rotatable bond and has 26 heavy atoms. It possesses 8 hydrogen bonds acceptor and 4 hydrogen bond donors. The calculated lipophilicity values (log P) were as follows: iLOGP 1.69, XLOGP3 1.32, WLOGP 0.77, MLOGP 0.63, Silicos-IT Log P 1.33 and Consensus Log P 0.89. The solubility prediction in ESOL was at -3.22, which is equal to solubility of 2.14 10 -1 mg/mL, which could be classified as soluble. Notably, the main cytochrome P450 isoforms (CYP1A2, CYP2C19, CYP2C9, CYP2D6 and CYP3A4) were not inhibited by UCT 1072M1 and thus, it is less likely to produce adverse drug-drug interactions. The compound passes all five criteria in Lipinski rule of five or to be precise, it has zero violations, approving of good oral bioavailability and pharmacokinetic properties. The radar plot also proves that UCT1072M1 falls entirely within the ideal physicochemical space (pink zone), which is the indication of an agreeable ADMET profile **(Figure 13a)**. Further, BOILED-Egg diagram locates the ligand between white areas, which suggests that the ligand may have an increased probability of human intestinal absorption **(Figure 13b)**. On the same note, the ADMET analysis of Linagliptin provided a molecular formula of C 25 H 28 N 8 O 2 and an atomic mass of 472.54 g / mol. It also showed molar refractivity of 139.33 and TPSA 116.86 Å^2. The ligand has 4 rotatable bond, 35 heavy atoms (19 aromatic heavy atoms) and weighs 1138.1 Da. It has 6 hydrogen bond acceptor and 1 hydrogen bond donor. The reported values of lipophilicity were the iLOGP 3.73, XLOGP3 1.91, WLOGP 0.85, MLOGP 1.8, Silicos-IT Log P 1.56 and Consensus Log P 1.97. The ESOL value of 4.11 translate to a solubility of 3.66 X 10 -2 mg/mL making Linagliptin as moderately soluble. Like the UCT1072M1, Linagliptin did not show inhibition of the tested cytochrome p450 enzymes which reduces the risk of metabolic drug interaction. The adherence to Lipinski Rule of Five was perfect and there were no violations, which suggests good oral bioavailability, as well as acceptable pharmacokinetic characteristics. First, the radar plot has confirmed that Linagliptin was fully within the desirable ADMET area (pink) **(Figure 14a)**. In the BOILED-Egg diagram, Linagliptin was positioned in the white area confining that the compound had great potential of absorption in the intestines **(Figure 14b)**.

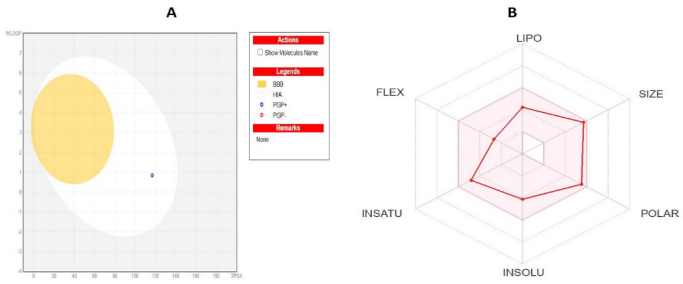

Figure 13: (a) shows the BOILED-Egg diagram, while Figure (b) presents the radar plot of UCT1072M1.

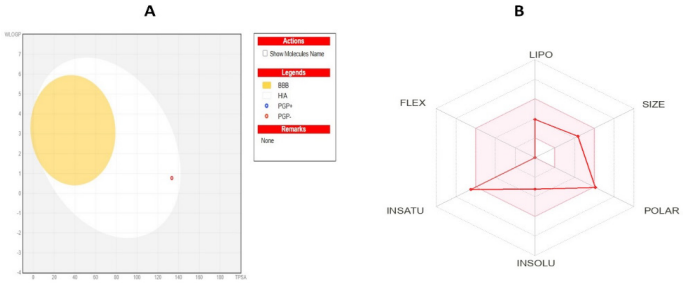

Figure 14: (a) shows the BOILED-Egg diagram, while Figure (b) presents the radar plot of linagliptin.

3.12. Toxicological profiling

Toxicity assessment remains an important element of the drug design process that aims at ensuring health hazards are minimal and that the environment is safe. In silico prediction of toxicity is a faster option compared to clinical trials performed on a traditional animal-based model that is also more cost-effective. Here, ProTox-IIII cyber tool was used to examine the toxicity of the considered ligands. ProTox-IIII predicts a wide range of toxicity end points which investigate immunotoxicity, mutagenicity, carcinogenicity, hepatotoxicity, acute toxicity, cytotoxicity, clinical toxicity and nutritional toxicity. Levels of toxicity fall into six groups; Class I toxicity is lethal; Class II and Class III are high toxicity; Class IV and Class V are moderate toxicity and Class VI is non-toxicity.

CT1072M1 had a prediction of the median lethal dose (LD50) of 3000 mg/kg, which belongs to Class V, which varies between moderate toxicity were listed in **(Table 5)**. The prediction showed a mean chemical similarity of 65.9 percent and the estimated expected success of 68.07 percent. On the same note, Linagliptin has also exhibited an LD50 of 684 mg / kg a figure that makes it fall under Class IV as well hinting at moderate toxicity were listed in **(Table 6)**. The similarity that matched with its prediction and accuracy were 57.56 and 67.38, respectively. These findings indicate that the two compounds have tolerable toxicity profile that can be developed into drugs in further stages[43].

Table 5: Toxicity evaluation of UCT1072M1.

Classification	Target	Shorthand	Prediction	Probability
Organ toxicity	Hepatotoxicity	Dili	Inactive	0.79
Organ toxicity	Neurotoxicity	neuro	Inactive	0.92
Organ toxicity	Nephrotoxicity	nephro	Active	0.58
Organ toxicity	Respiratory toxicity	respi	Active	0.73
Organ toxicity	Cardiotoxicity	cardio	Inactive	0.61
Toxicityend points	Carcinogenicity	carcino	Inactive	0.55
Toxicityend points	Immunotoxicity	immuno	Active	0.99
Toxicityend points	Mutagenicity	mutagen	Inactive	0.53
Toxicityend points	Cytotoxicity	cyto	Inactive	0.81
Toxicityend points	BBB-barrier	bbb	Active	0.61
Toxicityend points	Ecotoxicity	eco	Inactive	0.71
Toxicityend points	Clinical toxicity	clinical	Active	0.57
Toxicityend points	Nutritional toxicity	nutri	Active	0.50

Table 6: Toxicity assesment of Linagliptin.

Classification	Target	Shorthand	Prediction	Probability
Organ toxicity	Hepatotoxicity	dili	Inactive	0.70
Organ toxicity	Neurotoxicity	neuro	Active	0.86
Organ toxicity	Nephrotoxicity	nephro	Inactive	0.70
Organ toxicity	Respiratory toxicity	respi	Active	0.96
Organ toxicity	Cardiotoxicity	cardio	Inactive	0.91
Toxicity end points	Carcinogenicity	carcino	Inactive	0.58
Toxicity end points	Immunotoxicity	immuno	Inactive	0.88
Toxicity end points	Mutagenicity	mutagen	Inactive	0.51
Toxicity end points	Cytotoxicity	cyto	Inactive	0.60
Toxicity end points	BBB-barrier	bbb	Active	0.86
Toxicity end points	Ecotoxicity	eco	Inactive	0.61
Toxicity end points	Clinical toxicity	clinical	Active	0.71
Toxicity end points	Nutritional toxicity	nutri	Inactive	0.60

4. Discussion

Bovine spongiform encephalopathy (BSE) is a progressive neurodegenerative disease in cattle, worldwide. Though classical BSE cases that have mostly been related to the consumption of contaminated meat and bone meal have reduced significantly due to the strict input of international control measures, there has been an increase in the cases of atypical BSE[3]. Traditionally, BSE is diagnosed on the pathological alterations to the medulla oblongata which denotes the buildup of the abnormal proteins prions[2].

New genetic and bioinformatic methods as well as high-throughput methods of gene expression e.g. microarrays available in recent years have added new possibilities to the field of the prion diseases in solving the molecular mysteries behind them. We have examined the gene expression profile data (GSE69048) of the GEO database that represents the samples obtained in 16 BSE-infected and 7 normal animals. With the help of the GEO2R tool, we found 163 differentially expressed genes (DEGs) (124 downregulated genes and 39 upregulated genes). The findings complement the previous studies that, although being based on similar sample sizes of the H- and L- type of infected animals, indicated a limited number of statistically significant DEGs[44].

To get some idea about the functioning and pathways of these DEGs, we annotated them through Gene Ontology (GO) as well as KEGG pathway analysis through DAVID tool. The leading biological processes among the upregulated genes included such processes as negative regulation of cell population proliferation, adherens junction assembly, negative regulation of anoikis and protein localization to cell surface. Molecular functions were cysteine-type endopeptidase activity and scavenger receptor activity, whereas cellular components were nucleoplasm and focal adhesions sites. Mitochondrial processes were also enriched in downregulated genes, especially, mitochondrial respiratory chain complex I assembly, mitochondrial respiratory chain complex IV assembly and the carbon electron transport chain (cytochrome c to oxygen) activity as well as protein export and serine/threonine kinase activity[19].

The visualization of a protein, PPI, interaction network was generated in Cytoscape based on information held in the STRING database and contains 145 nodes and 146 edges. Under the CytoHubba plugin, we found ten major hub genes, i.e., ALB, NDUFA11, COX5A, UBC, COX6A1, TGFB1, ATP5MC3, NDUFB5, COX8A and NDUFA1. NDUFA11 was found to be the best downregulated hub gene among them and is of the most significant role in the assembly of the mitochondrial respiratory chain complex I. It is like the results of the former studies that the NDUFA11 gene is related to the stability of the mitochondrial complex I, the formation of super complexes and the formation of ATP. Interference of this gene may cause a threat to maintain the integrity of mitochondria and energy metabolism contributing to the pathogenesis of BSE[36].

Enrichment analysis of transcription factor and protein kinases was employed which showed a possible regulatory network causing the BSE pathology. RUNX1, TAF1, ELF1, PML and MYC have led to transcription factor involvement in the stress response process of the cell, immune response and neurodegeneration. An essential number of kinases such as CSNK2A1, CDK1, MAPK8, CDK2, as well as MAPK14 were linked to cell cycle, neuroinflammation and protein aggregation. These data provide information regarding upstream costs and can prove helpful in discovering preset biomarkers or therapeutic sites prior to the extent of prion build-up that could cause irreparable harm[36].

Because a 3D structure of NDUFA11 protein is not known with an experiment in the protein databases, we predicted the protein structure with the DMPfold algorithm and performed the refinement of its model on GalaxyRefine server. Of the five models developed, Model 1 had the best structural quality with GDT-HA score of 0.9007 meaning that the structural similarity between Model 1 and the native structure was high. The final model was also represented by a low RMSD of 0.578, a MolProbity score of 1.829 and Ramachandran plot 98.6%, which are good signs representing the model stability and stereochemical accuracy. There was little in the way of steric interference (clash score: 21.7) and MolProbity analysis had 97.1 percent of residues in the favorable regions[39,40].

Analysis of the refined NDUFA11 structure with CASTp, as an active site, revealing three possible binding pockets. Pocket 1 was identified as the most open and largest (the area 727.777 A 2; the volume 925.577 A 3) and so chosen as the major location of docking. The molecular docking has been carried out with the EasyDockVina tool that has strong flexibility and supports distributed and Python-enabled molecular docking[42]. The screening of a natural compound library (3,368 ligands) showed that two of them UCT1072M1 and Linagliptin became candidates due to high affinity of the ligand to NDUFA11 (binding energy: -9.4 kcal/mol in both cases)[43].

The SwissADME tool was used to perform ADMET profiling. UCT1072M1 (C 18H 12 O 8) exhibited attractive physicochemical features: molecular mass of 356.28 g/mol, TPSA = 133.52 A 2, neither of the criteria in the William Lipinski Rule of Five were violated and the water solubility of the compound was additional favorable characters. Notably, it did not inhibit all principal sleeping CYP450 enzymes, which decreased the chance of drug-to-drug interaction. Linagliptin (C 25 H 28 N 8 O 2) also passed Lipinski criteria, it had moderate solubility oral bioavailability was good and its ADME properties were favorable. The two compounds were found to have high rates of absorption in the human intestines according to BOILED-Egg model[37].

ProTox-III toxicity testing showed an LD 50 3000 mg/kg (toxicity group V: moderately toxic) for UCT1072M1 and an LD 50 684 mg/kg (toxicity group IV: moderately toxic) of the Linagliptin. Such levels of toxicity are reasonable when considering early drug discovery and they should be experimentally confirmed.

5. Conclusion

Based on our findings, NDUFA11 appears to be a significant biomolecular component involved in the pathogenesis of Bovine Spongiform Encephalopathy (BSE). As an essential subunit of mitochondrial complex, I, NDUFA11 plays a critical role in maintaining mitochondrial integrity and proper cellular function. Its downregulation may contribute to mitochondrial dysfunction, thereby facilitating disease progression in BSE.

Computational analyses identified UCT1072M1 and Linagliptin as potential ligands exhibiting strong binding affinity toward NDUFA11, suggesting their possible modulatory

effects on its activity. However, these findings require further experimental validation to confirm their biological relevance.

Moreover, the favorable ADMET profiles of these compounds highlight their potential as candidate therapeutic agents. Collectively, these results provide a promising foundation for future preclinical and clinical investigations aimed at developing effective therapeutic strategies against BSE.

6. Limitations

Despite the valuable insights provided by this in silico study, several limitations must be acknowledged. Firstly, the relatively small sample size may limit the generalizability of the findings. Future studies should incorporate larger and more diverse datasets to validate and extend the current results.

Furthermore, although this study identified potential drug candidates through computational modeling and molecular docking approaches, these compounds require experimental validation to confirm their biological activity, safety and therapeutic efficacy.

7. Funding

No funding was provided for this research.

8. Conflict of Interest Statement

The author, Itazaz Ul Haq, is a member of the editorial board of the Archives of Biotechnology and Pharmaceutical Research. To avoid any potential conflict of interest, the author had no involvement in the peer review process, editorial decision-making or handling of this manuscript. All review procedures were conducted independently in accordance with the journal's ethical guidelines.

9. Acknowledgement

We acknowledge that we used AI GPT in refining the language and text; however, the analysis, figures and conceptualization were performed without these AI methods.

10. Author Contributions

Nida and Ruqia Sartaj contributed equally to this work. Nida performed the system-level transcriptomic analysis and data interpretation. Ruqia Sartaj carried out the structural modeling, functional annotation and molecular analysis of NDUFA11. Israr Hussain assisted in methodology design, bioinformatics validation and manuscript drafting. Itazaz Ul Haq contributed to the statistical analysis and visualization of data. Bilal khan supported in literature review, dataset curation and technical editing. Muhammad Rahiyab and Zahid Hussain aided in computational validation and figure preparation. Syed Shujait Ali contributed to critical revision and intellectual input. Arshad Iqbal conceptualized and supervised the study, provided resources and finalized the manuscript.

11. References

1. Narayan KG, Sinha DK, Singh DK. Bovine Spongiform Encephalitis (BSE)/Mad Cow Disease, in: K. G. Narayan, et al. (Eds.), Veterinary Public Health & Epidemiology: Veterinary Public Health-Epidemiology-Zoonosis-One Health, Springer Nature Singapore, 2023: 235-247.
2. Alarcon P, Wall B, Barnes K, et al. Classical BSE in Great Britain: Review of its epidemic, risk factors, policy and impact. Food Control, 2023;146: 109490.
3. Haley NJ, Richt JA. Classical bovine spongiform encephalopathy and chronic wasting disease: two sides of the prion coin. Animal Diseases, 2023;3: 24.
4. Sikorska B, Liberski PP. Human Prion Diseases: From Kuru to Variant Creutzfeldt-Jakob Disease, in: J. R. Harris (Ed.), Protein Aggregation and Fibrillogenesis in Cerebral and Systemic Amyloid Disease, Springer Netherlands, Dordrecht, 2012: 457-496.
5. Legname G. Chapter 127 - Prions, in: Y.-W. Tang, et al. (Eds.), Molecular Medical Microbiology (Third Edition), Academic Press, 2024: 2577-2591.
6. Hirsch TZ, Martin-Lannerée S, Mouillet-Richard S. Chapter One - Functions of the Prion Protein, in: G. Legname and S. Vanni (Eds.), Progress in Molecular Biology and Translational Science, Academic Press, 2017: 1-34.
7. Aguzzi A, Weissmann C. Prion research: the next frontiers. Nature, 1997;389:795-798.
8. Casalone C, Hope J. Chapter 7 - Atypical and classic bovine spongiform encephalopathy, in: M. Pocchiari and J. Manson (Eds.), Handbook of Clinical Neurology, Elsevier, 2018: 121-134.
9. Dudas S, Czub S. Atypical BSE: current knowledge and knowledge gaps. Food Safety, 2017;5: 10-13.
10. Smith PG, Bradley R. Bovine spongiform encephalopathy (BSE) and its epidemiology. British medical bulletin, 2003;66: 185-198.
11. Ducrot C, Arnold M, De Koeijer A, et al. Review on the epidemiology and dynamics of BSE epidemics. Veterinary research, 2008;39: 1-18.
12. Bosque PJ. Bovine spongiform encephalopathy, chronic wasting disease, scrapie and the threat to humans from prion disease epizootics. Current Neurology and Neuroscience Reports, 2002;2: 488-495.
13. Hunter N. Scrapie and experimental BSE in sheep. British medical bulletin, 2003;66: 171-183.
14. Novakofski J, Brewer M, Mateus-Pinilla N, et al. Prion biology relevant to bovine spongiform encephalopathy. Journal of animal science, 2005;83: 1455-1476.
15. Hosseinkhani H. Nanomaterials in advanced medicine, 2019.
16. Hosseinkhani H. Biomedical engineering: materials, technology and applications John Wiley & Sons, 2022.
17. Domb AJ, Sharifzadeh G, Nahum V, et al. Safety evaluation of nanotechnology products. Pharmaceutics, 2021;13: 1615.
18. Clough E, Barrett T. The gene expression omnibus database, Statistical genomics: methods and protocols, Springer, 2016: 93-110.
19. Xerxa E, Barbisin M, Chieppa MN, et al. (2016) Whole Blood Gene Expression Profiling in Preclinical and Clinical Cattle Infected with Atypical Bovine Spongiform Encephalopathy. PLOS ONE, 2016;11: 0153425.
20. Dumas J, Gargano MA, Dancik GM. shinyGEO: a web-based application for analyzing gene expression omnibus datasets. Bioinformatics, 2016;32: 3679-3681.
21. Minguet EG, Segard S, Charavay C, et al. MORPHEUS, a webtool for transcription factor binding analysis using position weight matrices with dependency. PLoS One, 2015;10: 0135586.
22. Zhou T, Yao J, Liu Z. Gene ontology, enrichment analysis and pathway analysis. Bioinformatics in aquaculture: Principles and methods, 2017: 150-168.
23. Ogata H, Goto S, Sato K, et al. KEGG: Kyoto encyclopedia of genes and genomes. Nucleic acids research, 1999;27: 29-34.
24. Dennis Jr G, Sherman BT, Hosack DA, et al. DAVID: database for annotation, visualization and integrated discovery. Genome biology, 2003;4: 60.
25. Tang D, Chen M, Huang X, et al. SRplot: A free online platform for data visualization and graphing. PloS one, 2023;18: 0294236.

26. Shannon P, Markiel A, Ozier O, et al. Cytoscape: a software environment for integrated models of biomolecular interaction networks. Genome research, 2003;13: 2498-2504.
27. Chin C-H, Chen S-H, Wu H-H, et al. CytoHubba: identifying hub objects and sub-networks from complex interactome. BMC systems biology, 2014;8: 11.
28. Zhao P, Zhen H, Zhao H, et al. Identification of hub genes and potential molecular mechanisms related to radiotherapy sensitivity in rectal cancer based on multiple datasets. Journal of Translational Medicine, 2023;21:176.
29. Bryant P, Pozzati G, Elofsson A. Improved prediction of protein-protein interactions using AlphaFold2. Nature communications, 2022;13: 1265.
30. Heo L, Park H, Seok C. GalaxyRefine: Protein structure refinement driven by side-chain repacking. Nucleic acids research, 2013;41: 384-388.
31. Burley SK, Berman HM, Kleywegt GJ, et al. Protein Data Bank (PDB): the single global macromolecular structure archive. Protein crystallography: methods and protocols, 2017: 627-641.
32. Tian W, Chen C, Lei X, et al. CASTp 3.0: computed atlas of surface topography of proteins. Nucleic acids research, 2018;46: 363-367.
33. Kim S, Thiessen PA, Bolton EE, et al. PubChem substance and compound databases. Nucleic acids research, 2016;44: 1202-1213.
34. Minibaeva G, Ivanova A, Polishchuk P. EasyDock: customizable and scalable docking tool. Journal of Cheminformatics, 2023;15: 102.
35. Dallakyan S, Olson AJ. Small-molecule library screening by docking with PyRx, Chemical biology: methods and protocols, Springer, 2014: 243-250.
36. Daina A., Michielin O., Zoete V. SwissADME: a free web tool to evaluate pharmacokinetics, drug-likeness and medicinal chemistry friendliness of small molecules. Scientific reports, 2017;7: 42717.
37. Banerjee P, Eckert AO, Schrey AK, et al. ProTox-II: a webserver for the prediction of toxicity of chemicals. Nucleic acids research 2018;46: 257-263.
38. Sartaj R, Haq IU, Sartaj H, et al. Network pharmacology and molecular dynamics based elucidation of Mentha piperita phytochemicals in colorectal cancer therapy. In Silico Pharmacology, 2026;14: 59.
39. Hussain I, Haq IU, Rahiyab M, et al. Structure-guided drug repurposing and dynamics simulation reveal anti-viral candidates for Bourbon virus. In Silico Pharmacology, 2025;13: 155.
40. Hussain I, Rahiyab M, Iqbal A, et al. Structure-Based In Silico Discovery of Thymidine Kinase Inhibitors Targeting the Fatal Goatpox Virus: Integrating Multi-Library Screening and Molecular Dynamic Simulation. ChemistrySelect, 2025;10: 03462.
41. Owais, Hussain I, Rahiyab M, et al. Network pharmacology and molecular docking reveal PPARγ-directed antidiabetic phytochemical compounds of Momordica charantia. In Silico Pharmacology, 2026;14: 110.
42. Hussain I, Haq IU, Rahiyab M, et al. Integrative Network Pharmacology and Docking Analysis of Moringa oleifera in Alzheimer's Disease: Dual Targeting of MAPK1 and STAT3 by Active Phytochemical Compounds. Medinformatics, 2026.
43. Hussain I, Rahiyab M, Haq IU, et al. Identifying Significant Genes as Prognostic Biomarkers in Pancreatic Cancer Through Integrated Transcriptomic and in Silico Pharmacology Approaches. In Silico Research in Biomedicine, 2026: 100321.
44. Olech M. Conventional and state-of-the-art detection methods of bovine spongiform encephalopathy (BSE). International Journal of Molecular Sciences, 2023;24: 7135.
45. Bian X., Jiang H., Meng Y., et al. Regulation of gene expression by glycolytic and gluconeogenic enzymes. Trends in cell biology, 2022;32: 786-799.