IJPSR (2026), Volume 17, Issue 8 **(Research Article)**

INTERNATIONAL JOURNAL OF PHARMACEUTICAL SCIENCES AND RESEARCH

Received on 21 March 2026; received in revised form, 15 July 2026; accepted, 20 July 2026; published 01 August 2026

STRUCTURE BASED COMPUTATIONAL IDENTIFICATION OF POTENTIAL THERAPEUTIC LEADS TARGETING MSH3 PROTEIN

Adela Melari Lyngdoh [1], Amthul Saboora [1], Jedidiah Serena Abraham [1], Afreen Reesha [1], G. Gayathri [1], Rakshana Santosh G. [1], Erumalla Venkatanagaraju [2, * 3]

Department of Biotechnology [1], Mount Carmel College (Autonomous), Vasanth Nagar, Bengaluru - 560001, Karnataka, India.
Venkinformatics [2], Sidhi Vinayaka Layout, Kodigehalli, Vidyaranyapura, Bengaluru - 560097, Karnataka, India.
Department of Life Sciences [3], Indian Academy Degree College (Autonomous), Hennur Cross, Bengaluru - 560043, Karnataka, India.

Keywords:

MSH3 protein, ADMET prediction, Structure based drug design, Huntington's disease, Molecular docking

Correspondence to Author:
Dr. Erumalla Venkatanagaraju

Associate Professor,
Department of Life Sciences,
Indian Academy Degree College (Autonomous), Hennur Cross,
Bengaluru - 560043, Karnataka, India.

E-mail:venkatanagarajue@gmail.com

ABSTRACT: Huntington's disease is a progressive neurological condition characterized by motor dysfunction, cognitive decline, and psychiatric disturbances. Recent studies indicate that abnormalities in DNA mismatch repair mechanisms, particularly involving the MSH3 protein, contribute to the expansion of CAG repeats and the progression of the disease. In the current investigation, an integrated *in-silico* strategy was employed to identify small molecules capable of interacting with the MSH3 protein and to evaluate their pharmacological suitability for therapeutic development. Primarily, selected phytochemicals and FDA approved reference drugs used in Huntington's disease management, including tetrabenazine, deutetrabenazine, and valbenazine, were subjected to drug likeness and toxicity assessment using SwissADME and ProTox. Nine phytochemicals satisfied key ADMET criteria, demonstrating favorable gastrointestinal absorption, blood brain barrier permeability, and low predicted toxicity. Sequence similarity analysis using BLASTp confirmed the high conservation and correct identification of the MSH3 protein sequence, supporting its selection as a reliable computational target. Physicochemical analysis indicated that MSH3 possesses a slightly basic character, moderate conformational flexibility, and an overall hydrophilic nature. Structural validation using PROCHECK revealed that the experimentally determined structure (PDB ID: 3THX) exhibited superior stereochemical stability compared with models generated using trRosetta and AlphaFold. Molecular docking analysis demonstrated that tetrabenazine and deutetrabenazine showed the strongest binding affinity (-7.1 kcal/mol), whereas valbenazine showed -6.7 kcal/mol. Among the test ligands, valeric acid exhibited the most favorable interaction (-4.9 kcal/mol), followed by methyl isobutyl ketone (-4.2 kcal/mol) and isobutyric acid (-4.0 kcal/mol). These findings provide preliminary insights into ligandMSH3 interactions and may guide future structure based therapeutic development for Huntington's disease.

INTRODUCTION: Huntington's disease (HD) is a progressive autosomal dominant neurodegenerative disorder characterized by motor dysfunction, cognitive decline, and psychiatric disturbances [1].

QUICK RESPONSE CODE

DOI: 10.13040/IJPSR.0975-8232.17(8).2405-14

This article can be accessed online on www.ijpsr.com

DOI link: https://doi.org/10.13040/IJPSR.0975-8232.17(8).2405-14

The disease is primarily caused by abnormal expansion of Cytosine-Adenine-Guanine (CAG) trinucleotide repeats within the HTT gene, which leads to the production of mutant huntingtin protein and subsequent neuronal degeneration, particularly in the striatum and cerebral cortex [2].

Although several symptomatic treatments are available, including tetrabenazine, deutetrabenazine, and valbenazine, these drugs mainly alleviate motor symptoms and are often associated with adverse effects and

pharmacokinetic limitations [3, 4, 5]. Consequently, identifying new molecular targets and safer therapeutic candidates remains an important area of research in HD drug discovery.

Recent studies suggest that the expansion of CAG repeats is influenced not only by the mutant HTT gene but also by components of the DNA mismatch repair pathway [6]. Among these, the MSH3 protein has emerged as a critical factor involved in the recognition and processing of insertion-deletion loops during DNA replication and repair. Dysregulation of MSH3 has been linked to somatic expansion of CAG repeats, thereby contributing to disease severity and progression [7]. Therefore, targeting MSH3 or modulating its interactions may provide a promising strategy for controlling repeat instability associated with Huntington's disease.

In recent years, computational drug discovery approaches, including pharmacokinetic prediction, structural validation, and molecular docking, have significantly accelerated the identification of potential therapeutic candidates [8, 9]. *In-silico* screening enables rapid evaluation of compound libraries for drug-likeness, toxicity, and binding affinity with biological targets while reducing the time and cost associated with experimental screening [10].

In the present study, an integrated computational workflow was employed to investigate potential small molecules capable of interacting with the MSH3 protein. Selected phytochemicals and reference drugs were initially evaluated for pharmacokinetic and toxicity profiles using ADMET prediction tools. The structural reliability of MSH3 was assessed through physicochemical characterization and stereochemical validation before performing molecular docking studies. The interaction analysis aimed to identify key residues involved in ligand binding and to explore potential molecules that could serve as preliminary leads for future therapeutic development targeting pathways associated with Huntington's disease.

MATERIALS AND METHODS:

Mapping of Ligands and *In-silico* Drug Likeness Prediction: The selection of appropriate ligands is a critical step in the drug discovery and development pipeline, as it determines the success of subsequent computational and experimental evaluations. In the present study, phytochemical constituents from the following medicinal plants; *Ziziphus jujuba, Spondias dulcis, Moringa concanensis, Moringa pterygosperma, Moringa oleifera, Houttuynia cordata, Curcuma caesia, Glycosmis mauritiana, Oroxylum indicum, Rauwolfia serpentina, Bacopa monnieri, Abroma augusta, Aglaia odoratissima, Beta vulgaris, Viburnum opulus, Hibiscus sabdariffa, Cocos nucifera, Syzygium samarangense, Mucuna pruriens, Musa acuminate, Limnophila rugosa, Gossypium herbaceum and Psidium guajava* were selected for analysis. Phytochemical data for these plants were retrieved from the Indian Medicinal Plants Phytochemistry and Therapeutics 2.0 database [11]. The retrieved phytochemicals and their corresponding details are summarized in supplementary **Table 1.** To evaluate their drug likeness and molecular interactions, the corresponding three dimensional structures of the selected phytochemicals and FDA approved reference drugs, tetrabenazine, deutetrabenazine, and valbenazine, that are clinically prescribed for the management of Huntington's disease associated symptoms were obtained from the PubChem database [3, 4, 5, 12]. The drug likeness and toxicity of obtained ligands were predicted by SwissADME and ProTox 3.0 [13, 14]. The ADMET cleared ligands were further subjected to docking and interaction analysis.

Target Sequence Retrieval and Similarity Analysis: The target protein selected for the present *in-silico* investigation was MutS Homolog 3 (MSH3), a key component of the DNA mismatch repair pathway and increasingly implicated in the molecular pathology of Huntington's disease. The full-length amino acid sequence of human MSH3 was retrieved from the National Center for Biotechnology Information (NCBI) database with the accession number NP_002430.3.

To reduce the unintended interactions between ADMET screened ligands and essential cellular proteins, the amino acid sequence of the MSH3 protein was first evaluated for sequence similarity before proceeding with molecular docking. This assessment was carried out using the BLASTp program available through the NCBI platform [15].

The search was performed against the widely used curated protein repository Protein Data Bank (PDB), in order to identify possible homologous proteins. To ensure biologically relevant comparisons, the search was restricted to the Primates taxonomic group (taxid: 9443). The BLASTp analysis was conducted with a maximum target sequence limit of 5000, an E-value cutoff of 0.01, and a word size of 3. Sequence alignments were generated using the BLOSUM62 substitution matrix, with gap penalties defined as an existence cost of 11 and an extension cost of 1, together with conditional compositional score matrix adjustment. The resulting alignments were subsequently examined based on alignment score, percentage identity, and E-value, enabling the identification and evaluation of homologous protein sequences.

Physicochemical Characterization and Validation of Predicted and Experimental MSH3 Structures: Following target validation, the physicochemical characteristics of the MSH3 protein were analyzed using the ProtParam tool available on the ExPASy server [16].

Parameters such as molecular weight, theoretical isoelectric point (pI), amino acid and atomic composition, extinction coefficient, predicted half-life, instability index, aliphatic index, and grand average of hydropathicity (GRAVY) were computed to gain insights into the protein's stability, solubility, and structural behavior under physiological conditions. For three dimensional structure generation, the MSH3 amino acid sequence was modeled using two independent structure prediction platforms, trRosetta and AlphaFold, to ensure reliability and structural consistency [17, 18]. The predicted models were subsequently evaluated for stereochemical integrity using PROCHECK, which assesses backbone dihedral angles and overall geometric quality through Ramachandran plot analysis [19]. An experimentally resolved MSH3 structure retrieved from the Protein Data Bank (PDB ID: 3THX) was employed as a reference during validation to compare stereochemical parameters and ensure model accuracy. The protein structures that satisfied stereochemical quality criteria were considered structurally reliable and were subsequently utilized for molecular docking and protein ligand interaction analyses.

Molecular Docking and Interaction Analysis: Among the three dimensional structures evaluated, the experimentally resolved MSH3 crystal structure retrieved from the Protein Data Bank (PDB ID: 3THX) exhibited superior stereochemical quality and structural stability when compared with the predicted models generated using trRosetta and AlphaFold. Based on the PROCHECK validation metrics, this structure was therefore selected as the most reliable receptor model for subsequent molecular docking studies. Docking simulations were carried out to investigate the binding interactions of clinically approved reference compounds, namely valbenazine, tetrabenazine, and deutetrabenazine, along with selected test ligands, including 1-butanol, isoamyl alcohol, isobutyric acid, 2-pentanone, methyl isobutyl ketone, propionic acid, and valeric acid. These ligands were chosen based on their pharmacological relevance and physicochemical compatibility. Molecular docking was performed using the AutoDock module integrated within the PyRx virtual screening platform [20].

To ensure comprehensive sampling of the binding pocket, a grid box was defined with dimensions X=93.09 Å, Y=67.16 Å, and Z=135.30 Å, encompassing the predicted active site region of the MSH3 protein. Default docking parameters were maintained, and multiple conformations were generated for each ligand to identify energetically favorable binding poses. Following docking, the resulting proteinlig and complexes were subjected to detailed interaction analysis using BIOVIA Discovery Studio Visualizer [21]. Particular emphasis was placed on identifying conventional hydrogen bond contacts, which play critical roles in determining binding affinity and complex stability. The interaction profiles obtained provided insights into the molecular basis of ligand recognition and binding efficiency.

RESULTS AND DISCUSSION:

***In-silico* Drug Likeness and Toxicity Studies:** The *in-silico* evaluation of drug likeness and toxicity is a critical step in early stage drug discovery, as it helps prioritize compounds with favorable pharmacokinetic and safety profiles while minimizing late stage attrition. In the present investigation, the pharmacological suitability of selected phytochemicals and FDA approved

reference drugs was assessed using SwissADME and ProTox 3.0, focusing on absorption, distribution, metabolism, excretion, and toxicity (ADMET) parameters.

SwissADME analysis provides comprehensive insights into molecular features governing oral bioavailability, gastrointestinal absorption, bloodbrain barrier (BBB) permeability, metabolic stability, and drug likeness. Key parameters considered in the present study included Lipinski's rule of five compliance, topological polar surface area (TPSA), number of hydrogen bond donors and acceptors, rotatable bonds, P-glycoprotein (P-gp) substrate prediction, PAINS alerts, Brenk structural alerts, synthetic accessibility, and BBB permeation. These parameters collectively determine the pharmacokinetic feasibility of candidate molecules, particularly for central nervous system (CNS) disorders such as Huntington's disease, where BBB penetration is a critical requirement. Among the phytochemicals screened, nine compounds derived from *Ziziphus jujuba* successfully satisfied the predefined ADMET criteria, including favorable BBB permeability and high gastrointestinal absorption, indicating their potential suitability as CNS active agents. Notably, two of these compounds were structural isoforms, suggesting conserved physicochemical properties that may contribute to their effective pharmacokinetic behavior. The remaining phytochemicals from other plant sources failed to meet one or more critical ADMET parameters, primarily BBB permeability and metabolic stability, leading to their exclusion from subsequent docking studies.

An important observation from the SwissADME analysis was that the selected BBB permeable phytochemicals did not inhibit major cytochrome P450 isoenzymes, namely CYP1A2, CYP2C19, CYP2C9, CYP2D6, and CYP3A4. This finding is particularly significant, as inhibition of these enzymes is often associated with adverse drug interactions, metabolic instability, and toxicity risks. The absence of CYP inhibition suggests a lower probability of metabolic interference, thereby enhancing the therapeutic safety profile of these natural compounds. In contrast, the FDA approved reference drugs tetrabenazine, deutetrabenazine, and valbenazine, although demonstrating acceptable ADMET profiles in terms of BBB penetration, gastrointestinal absorption, Lipinski compliance, polar surface area, rotatable bonds, and synthetic accessibility, exhibited inhibitory effects against multiple CYP isoenzymes. This property may contribute to their known side effect profiles and complex pharmacokinetic interactions observed in clinical settings. Furthermore, these drugs possessed higher molecular weights and limited hydrogen bond donor capacity compared to the selected phytochemicals, which may influence binding flexibility and molecular recognition. Additionally, the toxicity prediction results indicated higher risk scores for the reference drugs relative to the plant derived leads.

Toxicological profiling using ProTox 3.0 further assessed critical safety endpoints, including organ toxicity, acute toxicity classes, stress response pathways, nuclear receptor signaling, molecular initiating events, and metabolic liability. The ADMET cleared phytochemicals demonstrated lower predicted toxicity, absence of hepatotoxicity and cardiotoxicity signals, and favorable metabolic profiles, highlighting their potential as safer therapeutic alternatives. In contrast, the FDA approved drugs showed comparatively higher toxicity risks, consistent with their documented adverse effects in long-term clinical usage. The corresponding ADMET details are summarized in supplementary **Table 1.**

Overall, the integrative *in-silico* screening approach enabled the identification of nine BBB permeable, metabolically stable, and low toxicity phytochemical leads from *Ziziphus jujuba*. These compounds exhibited superior pharmacokinetic and safety attributes compared to conventional FDA approved therapeutics. Consequently, they were selected for subsequent molecular docking and protein ligand interaction analyses to further explore their therapeutic potential in Huntington's disease.

Sequence Similarity Analysis and Target Validation Using BLASTp: The BLASTp analysis of the MSH3 amino acid sequence against curated protein databases restricted to *Homo sapiens* and the Primates taxonomic group produced sixteen significant alignments. The top ranking hits corresponded predominantly to the DNA mismatch repair protein MSH3, indicating

strong sequence conservation and confirming the accuracy of the retrieved query sequence. These matches displayed extremely high similarity, with query coverage reaching 100% and percentage identity values close to 100%, along with E values approaching zero, demonstrating highly significant alignments. Such results indicate that the analyzed sequence is highly conserved across the retrieved entries and corresponds to the expected human MSH3 protein. In addition to MSH3 entries, the BLAST output also identified homologous sequences belonging to other members of the MutS homolog family, particularly MSH6 and MSH2, which are known to participate in the DNA mismatch repair pathway. Although these proteins showed measurable similarity with the query sequence, their percentage identity values were considerably lower (approximately 27-28%) and their query coverage ranged between 62-77%, suggesting evolutionary relatedness but clear functional and structural distinction from MSH3. The presence of these homologous mismatch repair proteins reflects the conserved architecture of the MutS protein family involved in genome stability. However, the higher similarity scores observed for MSH3 sequences confirm the specificity of the query and support its suitability as the primary target for subsequent computational analyses. Overall, the BLASTp results validate the identity of the selected MSH3 sequence and provide confidence for its use in downstream structural modeling and molecular docking studies. A summary of the results is provided in **Table 1.**

TABLE 1: BLAST$_P$ BASED SEQUENCE SIMILARITY ANALYSIS OF THE MSH3 PROTEIN

S. no.	Protein	Chain	Organism	PDB Accession ID.	Query Coverage (%)	Identity (%)	E Value	Score
1	MSH3	B	*Homo sapiens*	8RZ7_B	100	100	0	2361
2	MSH3	B	*Homo sapiens*	8OLX_B	100	99.91	0.0	2359
3	MSH3	B	*Homo sapiens*	8R7V_B	100	99.74	0.0	2358
4	MSH3	B	*Homo sapiens*	8RAZ_B	100	99.65	0.0	2355
5	MSH3	B	*Homo sapiens*	8R7E_B	81	100.00	0.0	1905
6	MSH3	B	*Homo sapiens*	3THW_B	81	99.89	0.0	1904
7	MSH3	B	*Homo sapiens*	8RB1_B	81	99.78	0.0	1902
8	MSH3	B	*Homo sapiens*	8RAW_B	81	99.78	0.0	1900
9	MSH6	B	*Homo sapiens*	2O8B_B	77	27.09	2e-86	302
10	MSH6	B	*Homo sapiens*	8AG6_B	77	27.09	2e-85	304
11	MSH2	A	*Homo sapiens*	2O8E_A	62	27.85	5e-65	238
12	MSH2	A	*Homo sapiens*	2O8B_A	62	27.85	6e-65	238
13	MSH2	A	*Homo sapiens*	8R7C_A	62	27.85	6e-65	238
14	MSH2	A	*Homo sapiens*	8RB1_A	62	27.72	1e-64	237
15	MSH2	A	*Homo sapiens*	8RAZ_A	62	27.72	3e-64	236
16	MSH2	A	*Homo sapiens*	8RAX_A	62	27.72	6e-64	235

Physicochemical Analysis of MSH3 Protein: The physicochemical properties of the MSH3 protein were evaluated using ProtParam to assess its structural stability and biochemical behavior before molecular docking studies. The amino acid distribution **Fig. 1** demonstrated a balanced representation of hydrophobic, polar, and charged residues, supporting the multifunctional role of MSH3 in DNA mismatch repair. The calculated theoretical isoelectric point (pI) of 8.20 indicates that MSH3 is slightly basic in nature. At physiological pH (~7.4), the protein would carry a net positive charge, which may facilitate electrostatic interactions with negatively charged biomolecules such as DNA. This observation is consistent with the biological role of MSH3 in recognizing and binding to mismatched DNA regions. The presence of appreciable proportions of lysine and arginine residues further supports its potential for nucleic acid interaction and ligand binding through electrostatic stabilization. The instability index was calculated to be 47.95, suggesting that the protein may exhibit moderate instability under isolated *in-vitro* conditions. However, such values are not uncommon for large regulatory proteins and may reflect inherent conformational flexibility rather than structural weakness.

From a functional standpoint, moderate flexibility can be advantageous, particularly for proteins involved in molecular recognition and repair

processes, where conformational adaptability is required for substrate accommodation and complex formation. The grand average of hydropathicity (GRAVY) score of -0.324 indicates an overall hydrophilic character with localized hydrophobic regions. A negative GRAVY value suggests favorable solubility in aqueous environments, which aligns with the intracellular localization of MSH3. At the same time, the presence of hydrophobic residues such as leucine, isoleucine, and valine contributes to stable core formation and potential hydrophobic interaction sites within ligand binding pockets. Collectively, the slightly basic pI, moderate instability index, and negative GRAVY value indicate that MSH3 possesses a structurally adaptable yet functionally stable profile. These characteristics are particularly relevant for molecular docking studies, as they suggest the presence of accessible binding regions capable of supporting both hydrophobic contacts and hydrogen bond interactions. The physicochemical profile, therefore, supports the suitability of MSH3 as a reliable computational target in Huntington's disease related drug discovery investigations

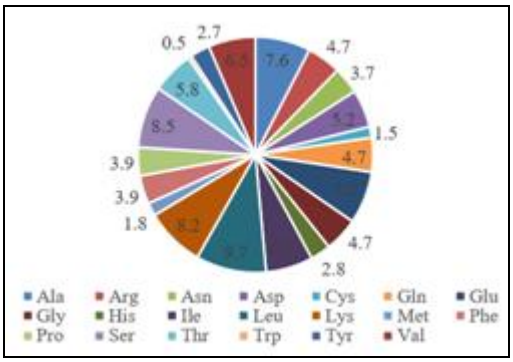


FIG. 1: DISTRIBUTION OF AMINO ACID RESIDUES IN THE MSH3 PROTEIN

Stereochemical Validation of Predicted and Experimental MSH3 Structures: To ensure the structural suitability of MSH3 for structure based drug design, all predicted models generated using trRosetta and AlphaFold, along with the experimentally resolved structure (PDB ID: 3THX), were subjected to stereochemical validation using PROCHECK. The Ramachandran plot analysis provided insights into backbone dihedral angle distribution and conformational stability. The experimentally resolved MSH3 structure (3THX) demonstrated 88.1% residues in the most favored regions, 11.4% in additionally allowed regions, 0.4% in generously allowed regions, and only 0.1% in disallowed regions. The Ramachandran plot revealed dense clustering of residues within the canonical α-helical (lower-left quadrant) and β-sheet (upper-left quadrant) regions, indicating energetically favorable φ (phi) and ψ (psi) angle distributions. The negligible proportion of residues in disallowed regions confirms minimal steric strain and high geometric precision **Fig. 2.**

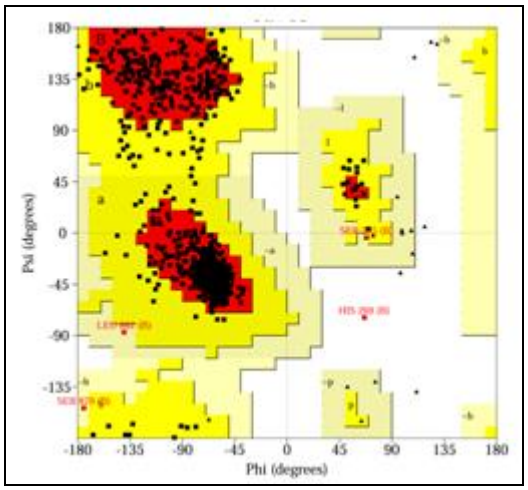


FIG. 2: RAMACHANDRAN PLOT ANALYSIS OF THE PDB MSH3 PROTEIN STRUCTURE

Among the trRosetta models, Model 4 exhibited the best stereochemical profile, with 93.0% residues in favored regions and only 0.2% in disallowed regions. All trRosetta models maintained more than 90% residues in favored regions, suggesting generally acceptable backbone conformations. However, slight increases in disallowed residues (up to 0.5%) and modest dispersion in allowed regions indicate minor deviations, particularly in loop or flexible domains. In contrast, the AlphaFold generated models showed comparatively reduced stereochemical quality. The percentage of residues in most favored regions ranged between 74.3% and 76.3%, while residues in disallowed regions ranged from 2.7% to 9.3%. The Ramachandran plots of these models displayed broader residue scattering outside the core α-helix and β-sheet clusters, with noticeable occupancy in generously allowed and disallowed regions. Such deviations may arise from modeling uncertainties in flexible segments or long disordered regions characteristic of DNA repair proteins.

Although trRosetta Model 4 numerically exceeded the PDB structure in favored region percentage, the

experimentally resolved 3THX structure demonstrated superior overall stereochemical balance, particularly with respect to nearzero disallowed residues and well-defined conformational clustering. Experimental crystal structures inherently reflect biologically stabilized conformations, providing higher reliability for active site geometry and binding pocket architecture. From a structure based drug discovery perspective, precise backbone geometry is critical, as even minor conformational distortions can influence docking accuracy, binding orientation, and interaction energy calculations. The higher proportion of residues in disallowed regions observed in AlphaFold models raises concerns regarding potential inaccuracies in binding site topology. Therefore, considering both quantitative Ramachandran statistics and qualitative plot distribution, the PDB derived MSH3 structure (3THX) was selected as the most stereochemically stable and biologically reliable model for subsequent molecular docking and interaction studies **Table 2.**

TABLE 2: COMPARATIVE STEREOCHEMICAL EVALUATION OF TRROSETTA, ALPHAFOLD AND PDB (3THX) MSH3 STRUCTURES

Models	Residues in Most Favored Regions		Residues in Additional Allowed Regions		Residues in Generously Allowed Regions		Residues in Disallowed Regions	
	Amino Acid Number	%	Amino Acid Number	%	Amino Acid Number	%	Amino Acid Number	%
trRosetta Model 1	950	91.5	82	7.9	2	0.2	4	0.4
trRosetta Model 2	942	90.8	93	9.0	1	0.1	2	0.2
trRosetta Model 3	947	91.2	85	8.2	3	0.3	3	0.3
trRosetta Model 4	965	93.0	70	6.7	1	0.1	2	0.2
trRosetta Model 5	953	91.8	74	7.1	6	0.6	5	0.5
PDB 3THX	686	88.1	89	11.4	3	0.4	1	0.1
AlphaFold Model 1	792	76.3	138	13.3	64	6.2	44	4.2
AlphaFold Model 2	786	75.7	123	11.8	94	9.1	35	3.4
AlphaFold Model 3	786	75.7	150	14.5	74	7.1	28	2.7
AlphaFold Model 4	778	75.0	98	9.4	65	6.3	97	9.3
AlphaFold Model 5	771	74.3	97	9.3	74	7.1	96	9.2

Overall, this comparative validation underscores the importance of rigorous structural assessment before docking simulations and reaffirms that experimentally determined structures remain the preferred choice for high confidence computational drug discovery workflows when available.

Molecular Docking and Interaction Analysis of MSH3 with Reference and Test Leads: To evaluate the binding potential of selected phytochemical derived compounds with the target protein MSH3, molecular docking studies were performed using the validated protein structure. Three FDA approved reference drugs used for the management of Huntington's disease symptoms, valbenazine, tetrabenazine, and deutetrabenazine, were included as control ligands. In addition, selected small molecules, including 1-butanol, isoamyl alcohol, isobutyric acid, 2-pentanone, methyl isobutyl ketone, propionic acid, and valeric acid, were docked to explore their interaction potential with the MSH3 binding region.

The predicted binding affinities and interacting amino acid residues are summarized in **Table 3.**

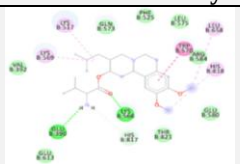
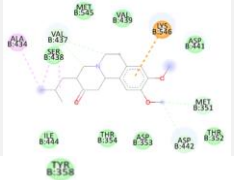
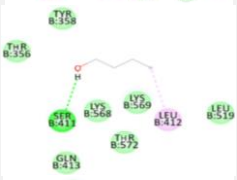
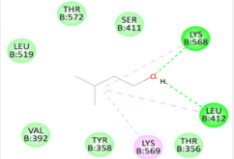
Among the reference compounds, tetrabenazine and deutetrabenazine exhibited the highest binding affinity values of -7.1 kcal/mol, indicating strong predicted interactions with the MSH3 protein. However, the interaction analysis did not reveal specific stabilizing amino acid contacts within the analyzed binding region, suggesting that the predicted binding may primarily involve hydrophobic interactions rather than strong hydrogen bonding. Valbenazine showed a slightly lower binding affinity (-6.7 kcal/mol) compared to the other reference drugs but demonstrated defined interactions with key residues Lys-566 and Gln-390. The presence of these interactions indicates potential stabilization of the ligand within the binding pocket through hydrogen bonding interactions, which may contribute to ligand retention despite a slightly lower docking score.

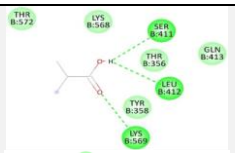
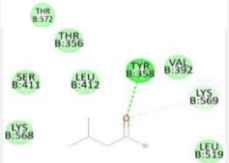
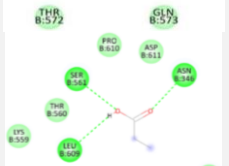
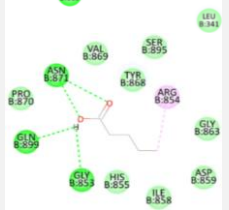
Among the test ligands, valeric acid displayed the most favorable binding energy (-4.9 kcal/mol) and formed interactions with residues Asn-871, Gln-899, and Gly-853, suggesting the formation of a stable ligandprotein complex. These interactions may contribute to enhanced binding stability through hydrogen bonding and polar interactions within the active region of the protein. Methyl isobutyl ketone also demonstrated relatively strong binding among the tested compounds with a docking score of -4.2 kcal/mol, interacting with Tyr-358, which may provide stabilization through hydrophobic contacts and possible π-related interactions. Similarly, isobutyric acid showed a binding affinity of -4.0 kcal/mol and formed multiple interactions involving Ser-411, Leu-412, and Lys-569, indicating that both polar and hydrophobic residues contribute to ligand binding.

Other compounds, including isoamyl alcohol (-3.9 kcal/mol), 2-pentanone (-3.8 kcal/mol), propionic acid (-3.4 kcal/mol), and 1-butanol (-3.3 kcal/mol), exhibited comparatively lower binding affinities. Nevertheless, these molecules still formed interactions with residues such as Ser-411, Lys-568, Leu-609, and Asn-346, suggesting that the MSH3 binding region accommodates small aliphatic molecules through a combination of hydrogen bonding interactions. Overall, the docking analysis demonstrated that the FDA approved drugs exhibited stronger binding affinities compared to the tested small molecules, which is expected given their optimized pharmacological structures.

However, certain test ligands, particularly valeric acid and isobutyric acid, showed moderate binding interactions with key amino acid residues of MSH3, indicating potential structural compatibility with the binding pocket. From a drug discovery perspective, the identification of interacting residues such as Ser-411, Lys-568/569, Gln-390, and Asn-871 suggests that these regions may play an important role in ligand recognition and stabilization within the MSH3 protein. These findings provide preliminary insights into the molecular interaction landscape of MSH3 and may guide the rational design or optimization of small molecules targeting pathways associated with Huntington's disease.

TABLE 3: COMPARATIVE MOLECULAR DOCKING AND 2D INTERACTION ANALYSIS OF STANDARD AND PHYTOCHEMICALS WITH THE MSH3 PROTEIN

Compound name	Binding affinity (kcal/mol)	Interacting amino acids	2D Interaction analysis
Valbenazine	-6.7	Lys-566 and Glu-390	
Tetrabenazine	-7.1	Nil	
Deutetrabenazine	-7.1	Nil	
1-Butanol	-3.3	Ser-411	
Isoamyl alcohol	-3.9	Lys-568 and Leu-412	

Isobutyric acid	-4.0	Ser-411, Leu-412 and Lys-569	
2-Pentanone	-3.8	Lys-568	
Methyl isobutyl ketone	-4.2	Tyr-358	
Propionic acid	-3.4	Ser-561, Asn-346 and Leu-609	
Valeric acid	-4.9	Asn-871, Gln-899 and Gly-853	

CONCLUSION: The present study employed an integrated *in-silico* drug discovery approach to explore potential small molecules capable of interacting with the DNA mismatch repair protein MSH3, a protein implicated in mechanisms associated with Huntington's disease progression. Initial pharmacokinetic and toxicity screening using computational ADMET tools identified several phytochemicals with favorable drug likeness, bloodbrain barrier permeability, and low predicted toxicity, highlighting their potential suitability for central nervous system applications. Compared with FDA approved reference drugs, the selected phytochemicals demonstrated lower predicted toxicity risks and minimal cytochrome P450 inhibition, suggesting improved metabolic safety profiles. Sequence similarity analysis confirmed the high conservation and accurate identity of the MSH3 protein, validating its selection as a reliable computational target. Physicochemical characterization further indicated that the protein possesses a slightly basic, moderately flexible, and hydrophilic nature, supporting its functional role in nucleic acid interaction and ligand recognition. Structural validation through Ramachandran plot analysis established that the experimentally resolved structure (PDB ID: 3THX) exhibited superior stereochemical stability compared with predicted models, making it the most suitable structure for docking analysis. Molecular docking revealed that while reference drugs showed stronger binding affinities, several tested small molecules demonstrated moderate but meaningful interactions with key residues of the MSH3 binding region. These findings provide preliminary insights into ligand-MSH3 interactions and may support future structure guided optimization of novel therapeutic candidates targeting molecular pathways involved in Huntington's disease.

ACKNOWLEDGMENTS: The authors would like to express gratitude to the Indian Academy Institutions management for providing the infrastructural facilities to carry out this work.

Supplementary Data: ADMET profile of standard and phytochemicals available as supplementary data **Table 1** (https://doi.org/10.5281/zenodo.19146640).

Financial Support: None

Ethics Statement: None

CONFLICT OF INTEREST: None

REFERENCES:

1. Christopher AR, Elizabeth HA, Edward JW, Douglas RL, Jeffrey DL, John HW, Rachael IS, Blair RL, Julie CS, Jane SP, Ralf R, Paul GU, Alice W, Russell LM and Sarah JT: Huntington disease: natural history, biomarkers and prospects for therapeutics. Nature Reviews Neurology 2014; 10: 204-16.
2. Joseph S and Tiago AM: Huntington's disease: latest frontiers in therapeutics. Current Neurology and Neuroscience Reports 2024; 24: 255-64.
3. Harshit G, Wil P, Cain S, Sathya K, Juyeon K, Adam K and Alan DK: Deutetrabenazine for the treatment of chorea associated with Huntington's disease. Health Psychol Res 2022; 10(5): 36040.
4. Daniel OC, Benjamin C, Lisa MDB, Eric W, Rajeev A, Sanjay G and David S: Indirect tolerability comparison of Deutetrabenazine and Tetrabenazine for Huntington disease. Journal of Clinical Movement Disorders 2017; 4(3): 1-11.
5. Christopher FOB, Roland J, Robert AH, Stewart AF, Joshua B, Daniel M and Julio CCG: NBI-98854, a selective monoamine transport inhibitor for the treatment of tardive dyskinesia: A randomized, double-blind, placebo controlled study. Mov Disord 2015; 30(12): 1681-87.
6. Ricardo MP, Ella D, Andrew K, Alejandro L, Edith L, Jason SC, Gagan BP, Caixia H, Kim H, Tammy G, Jolene RG, Paula EC, Guo ML, Christopher EP, Mark JD and Vanessa CW: Mismatch repair genes Mlh1 and Mlh3 modify CAG instability in Huntington's disease mice: genome-wide and candidate approaches. PLOS Genetics 2013; 9(10): e1003930.
7. Michael F, Vilija L, Marc C, Sarah C, Fernando M, Kitty L, Davina HM, Lesley J, Peter H, Darren GM and Sarah JT: MSH3 modifies somatic instability and disease severity in Huntington's and myotonic dystrophy type 1. Brain 2019; 142(7): 1876-86.
8. Gregory S, Sandeepkumar K, Jens M and Edward WL: Computational methods in drug discovery. Pharmacol Rev 2013; 66(1): 334-95.
9. Afza Kousar, Arshiya LF, Aliya F, Nousheen F and Erumalla V: *In-silico* molecular docking of *Salvadora persicametabolites* with myocarditis causative coxsackie virus B3 capsid proteins. International Journal of Pharmaceutical and Phytopharmacological Research 2025; 15(1): 4-12.
10. Bhargavi VH, Azra E, Manali NJ, Mallikarjun and Erumalla V: *In-silico* exploration of chyawanprash phytoconstituents binding affinity with key targets of Human metapneumovirus. International J of Pharma Sciences and Research 2026; 17(1): 315-32.
11. Vivek A, Karthikeyan M, Ajaya KS and Areejit S: IMPPAT 2.0: An Enhanced and Expanded Phytochemical Atlas of Indian Medicinal Plants. ACS Omega 2023; 8: 8827-45.
12. Sunghwan K, Jie C, Tiejun C, Asta G, Jia H, Siqian H, Qingliang L, Benjamin AS, Paul AT, Bo Y, Leonid Z, Jian Z and Evan EB: PubChem in 2021: new data content and improved web interfaces. Nucleic Acids Research 2021; 49(1): 1388-95.
13. Antoine D, Olivier M and Vincent Z: SwissADME: a free web tool to evaluate pharmacokinetics, drug-likeness and medicinal chemistry friendliness of small molecules. Scientific Reports 2017; 7: 42717.
14. Priyanka B, Emanuel K, Mathias D and Robert P: ProTox 3.0: a webserver for the prediction of toxicity of chemicals. Nucleic Acids Res 2024; 52(1): 513-20.
15. Mark J, Irena Z, Yan R, Yuri M, Scott M and Thomas LM: NCBI BLAST: a better web interface. Nucleic Acids Res 2008; 36: 5-9.
16. Gasteiger E, Hoogland C, Gattiker A, Duvaud S, Wilkins MR, Appel RD, Bairoch A and John MW: Protein Identification and Analysis Tools on the Expasy Server. The Proteomics Protocols Handbook. Humana Press 2005; 571-607.
17. Jianyi Y, Ivan A, Hahnbeom P and David B: Improved protein structure prediction using predicted inter residue orientations. Biophysics and Computational Biology 2020; 117(3): 1496-1503.
18. John J, Richard E, Alexander P, Tim G, Michael F, Olaf R, Kathryn T, Russ B, Augustin Z, Anna P, Alex B, Clemens M, Simon AAK, Andrew JB, Andrew C, Bernardino RP, Stanislav N, Rishub J, Jonas A, Trevor B, Stig P, David R, Ellen C, Michal Z, Martin S, Michalina P, Tamas B, Sebastian B, David S, Oriol V, Andrew WS, Koray K, Pushmeet K and Demis H: Highly accurate protein structure prediction with AlphaFold. Nature 2021; 596: 583-89.
19. Laskowski RA, MacArthur MW and Thornton JM: PROCHECK: validation of protein structure coordinates, in International Tables of Crystallography, Volume F. Crystallography of Biological Macromolecules, eds. Rossmann M G & Arnold E, Dordrecht, Kluwer Academic Publishers. The Netherlands 2001; 722-25.
20. Sargis D and Arthur JO: Small-Molecule Library Screening by Docking with PyRx. Chemical Biology 2014; 243-50.
21. BIOVIA. Discovery Studio Visualizer, v21. 1.0. 20298; Dassault Systèm: San Diego, CA, USA, 2021.

How to cite this article:
Lyngdoh AM, Saboora A, Abraham JS, Reesha A, Gayathri G, Santosh RG and Venkatanagaraju E: Structure based computational identification of potential therapeutic leads targeting MSH3 protein. Int J Pharm Sci & Res 2026; 17(8): 2405-14. doi: 10.13040/IJPSR.0975-8232.17(8).2405-14.