Biomedical Letters ISSN 2410-955X– An International Biannually Journal

Research article
2020 | Volume 6 | Issue 2 | Pages 138-148

ARTICLE INFO

Received
March 23, 2020
Revised
May 07, 2020
Accepted
June 15, 2020

***Corresponding Author**
Syeda Maheen Abid

E-mail
syedamaheenabid321@gmail.com

Keywords
Molecular docking
ALS
C9ORF72
Homology modeling
Virtual screening
AutoDock Vina

How to Cite
Abid SM, Ali M. Homology modeling, molecular docking and virtual screening to reveal potential inhibitor of ALS associated protein guanine nucleotide exchange C9ORF72. Biomedical Letters 2020; 6(2):138-148.

Scan QR code to see this publication on your mobile device.

Special Issue: Computational drug designing and molecular docking analyses

Open Access

Homology modeling, molecular docking and virtual screening to reveal potential inhibitor of ALS associated protein guanine nucleotide exchange C9ORF72

Syeda Maheen Abid[1*], Muhammad Ali[2]

[1]Department of Bioinformatics and Biotechnology, Government College University Faisalabad, Faisalabad, Pakistan
[2] District Head Quarter Hospital, Chiniot, Pakistan

Abstract

Repeated expansion of hexanucleotide in C9ORF72 encodes the protein Guanine Nucleotide Exchange considered as the main cause of Amyotrophic Lateral Sclerosis (ALS). The repeated expansion produces toxic products and autophagy deficits. Various *in silico* approaches were employed for structural 3D modeling and protein-protein molecular docking analyses of C9ORF72 followed by virtual screening. Homology modeling and threading approaches were applied to predict the 3D structures of C9ORF72 and 92.38% of quality factor was calculated by ERRAT evaluation tool. STRING database was utilized, and it was observed that SMCR8 has the ability to be the interacting partner of C9ORF72. Protein-protein molecular docking analyses of C9ORF72 with SMCR8 were performed and potential interacting residues were observed computationally. FDA library from ZINC database was utilized for virtual screening and comparative molecular docking analyses were performed by AutoDock Vina. It was proposed that the scrutinized compound ZINC131 have strong binding affinities and least binding energy of -11.3 kcal/mol. The suggested molecule may be used for further analyses in the drug discovery processes. The predicted 3D structure of C9ORF72 provides the structural insights for the better functional understanding of C9ORF72. Overall, the findings of present work may be helpful in designing the novel therapeutic targets against C9ORF72.

Introduction

Amyotrophic Lateral Sclerosis (ALS) was first described by Charcot in 1874 [1]. The cause of ALS/Lou Gehrig disease is not completely known and considered as spontaneously arise [2]. In ALS, upper motor neurons and lower motor neurons degenerate and die (motor neurons link between the brain and voluntary muscles) [3]. ALS affects the nerve cells in the brain and spinal cord. The brain lost the ability to initiate and control the movement of muscles [4]. The gradual decline in strength leads to paralysis of more and more muscles leads to death. The genetic cause of ALS includes mutation in C9ORF72, FUS, SOD1, VCP and TDP-43 [5]. Most of the cases for ALS are sporadic and about 25% of people were suffering for ALS due to family history [6]. The sporadic ALS are usually to the patients between the age of 55-65 years old and only 5% patients are <30 years old. Males and females are equally affected by ALS. Juvenile ALS (JALS) is a term used for patients below the age of 25 years [7].

Most of the time ALS is autosomal recessive while dominant inheritance linked with chromosome 9 [8]. ALS is a motor neuron disease characterized by muscle twitching [9], muscle stiffness and muscle weakness due to a decrease in muscle size. In most cases, patient lose the ability to speak, walk, breathe, swallow and hand movement [10, 11]. Some patients face difficulties in thinking and behavioral acts [12]. ALS has no cure yet however early diagnoses may help to treat and keep the muscle control little longer [13]. The death of ALS patient cause within 3 to 4 years due to respiratory failure [14]. ALS is related to parkinsonism and dementia [15]. There is no specific cure of ALS, moreover Riluzole have considerable relax to the patients [16]. Another drug approved by FDA for ALS is Edaravone [17].

C9ORF72 is localized on chromosome 9 [18]. Augmentation of GGGGCC hexanucleotide repeat extensively present in C9ORF72 considered as the common cause of ALS [19]. A repetition of hexanucleotides is considered as the genetic cause of almost 10% patients [20]. Guanine nucleotide exchange C9ORF72 is encoded by *C9ORF72* [21]. The structure of C9ORF72 is not known yet. *C9ORF72* is present on the short arm of the chromosome (p) in humans [22]. The length of the sequence is from 27,546,542 base pairs to 27,573,863 base pairs. The protein guanine nucleotide exchange C9ORF72 is present in most of the regions of the brain, presynaptic terminals and in the cytoplasm of neurons [23, 24]. In C9ORF72, there are fewer repeats of hexanucleotide GGGGCC as <30 normally however in patients of ALS, these repeats are in hundreds [25, 26]. These repeats decrease the autophagy regulator of protein C9ORF72 alters the expression leads to ALS. The lack of C9ORF72 might be the cause of disease. C9ORF72 emerged in most of the eukaryotes and have single copy of gene encoding *C9ORF72* [27]. The presence of four nucleotides of Guanine and two of Cytosine noncoding part repetition cause severe kind of mutational changes leads to ALS [28, 29]. In C9ORF72 hexanucleotide repetition can cause RNA toxicity through the confiscate and collection of RNA binding protein. The guanine nucleotide exchange protein has two isoforms; one has the sequence length of 481 amino acids while the other has 222 amino acids. Repeated expansion cause mutation in C9ORF72 leads to ALS. The neuronal function of C9ORF72 is unknown. C9ORF72 has structural homology with Differentially Expressed in Normal Neoplasia DENN [30].

During the last two decades, the number of known protein sequences has increased as compared to structures [31]. This unbalance between the protein sequence and structure has censoriously limited the ability to understand the molecular mechanism of proteins [32]. The structure formation rate of known protein is much slower as the structure prediction techniques (X-Ray Crystallography and NMR) are time consuming [33]. The development of structural bioinformatics helps to solve the biological macromolecules (DNA, RNA and protein) structural analyses [34]. There have been many achievements in computational drug designing and personalized medicine [35, 36]. Various possibilities are present to understand neurological disease which plays an important role in the medical field [37-39]. The focus of current work was to 3D structure prediction, evaluation and validation of Guanine Nucleotide Exchange C9ORF72 followed by protein-protein molecular docking and virtual screening.

Materials and Methods

The C9ORF72 have accession number Q96LT7 in Uniport Knowledge Base. In this work, 3D structure prediction, virtual screening and molecular docking analyses were performed.

The amino acid sequence (FASTA sequences) of Guanine nucleotide exchange C9ORF72 was retrieved from Uniport KB (http://www.uniprot.org/) [40]. The sequence was subjected to BLASTp for the selection of a suitable template against Protein Data Bank (PDB) [41, 42]. The automated program

MODELLER 9.20 [43] for homology modeling was used to predict the 3D structures of C9ORF72 by spatial restraints. The online tools including I-TASSER [44], RaptorX [45], CPHModel [46], HHpred [47], Phyre2 [48], SWISS-MODEL [49], MOD-WEB [50], Robetta [51], Sparks-X [52], 3D-Jigsaw [53] and ESyPred 3D [54] were also used to predict the protein structure. The 3D structures of C9ORF72 was visualized by UCSF Chimera 1.13.1 [55] and PyMOL [56]. UCSF Chimera also used to minimize selected structure. Rampage [57], Anolea [58], ProCheck [59] and ERRAT [60] evaluations tools were used to evaluate the quality of the model of protein structure. The produced Ramachandran plots for the assessment of predicted structures showed residues dispensation and also declared ϕ and ψ distribution of non-proline and non-glycine residues. For the differentiation of favorable and non-favorable regions, phi and psi angles were plotted against each other. These angles were used for the assessment of different regions. ERRAT evaluation tool was also used to calculate the quality factors of predicted structures [61].

To determine the functional interacting partner of target protein, STRING (Search tool for the retrieval of interacting genes/proteins) [62] and STITCH (Search Tool Interacting Chemical) [63] were used. The online server PatchDock [64, 65] was used for protein-protein molecular docking and FireDock [66] was used for the refinement and scoring of protein-protein docking solutions. Gramm-X was also employed for protein-protein docking analysis for the cross validation of the analyses. LigPlot [67, 68] was utilized to analyze the hydrophobic and electrostatic interaction and also used to generate schematic diagrams of protein-protein interactions.

PyRx [69] software was used to dock the small molecules with macromolecule and virtual screening. The blind docking was proceeded to analyze the protein and ligand interaction for confirmation and orientation. The FDA library of Zinc database [70] was retrieved for virtual screening against the target protein [71].

Results and Discussion

The study of neurology and structural bioinformatics are fields of exploring knowledge and providing an effective way for better understanding and development of different research approaches for the diagnosis, cure and detection of neuronal diseases including ALS. The ensemble genome browser was used to locate the exact position of C9ORF72 protein-coding gene in humans (**Fig. 1**).

Fig. 1: The presence of gene C9orf72 on the position of chromosome number 8.

It has been observed from extensive literature review that Guanine Nucleotide Exchange protein has no other reported member in family. The multiple sequence alignment (MSA) was performed for two isoforms of C9ORF72 carried out by Clustal omega [72]. An asteric(*) indicated positions which have single, fully conserved residues, colon indicated the similar residues and dot indicated the weakly similar residues. The positions with no dot indicated non-conserved residues (**Fig. 2**).

Coils, Protparam [73] tools were used to calculate the physiochemical properties of the receptor protein. The molecular weight of the protein based on average isotope masses of the amino acids was also studied. The theoretical PI was 5.82 depends on side chains determined the pH of the protein. The half-life of the protein was calculated 30 hours *in vitro*. The aliphatic index was -0.069, instability index was 50.54 and the number of positively charged residues were 50 as well as negatively charged residues with total number of an atom was also calculated 60 (**Table 1**) (**Fig. 3**).

The PONDR [74] tool was used to predict the ratio of natural disorders caused by the mutation in C9ORF72. The graph showed the composition of order and disorder. The center line was the threshold and peak above the threshold identified as disorder while the line below the threshold identified as an order of given protein C9ORF72 (**Fig. 4**).

Structure Prediction

The 3D structure of C9ORF72 was not reported by X-Ray crystallography and NMR yet. The comparative modeling and threading approaches

```
CLUSTAL O(1.2.4) multiple sequence alignment

sp|Q96LT7|CI072_HUMAN        MSTLCPPPSPAVAKTEIALSGKSPLLAATFAYWDNILGPRVRHIWAPKTEQVLLSDGEIT        60
sp|Q96LT7-2|CI072_HUMAN      MSTLCPPPSPAVAKTEIALSGKSPLLAATFAYWDNILGPRVRHIWAPKTEQVLLSDGEIT        60
                             ************************************************************

sp|Q96LT7|CI072_HUMAN        FLANHTLNGEILRNAESGAIDVKFFVLSEKGVIIVSLIFDGNWNGDRSTYGLSIILPQTE        120
sp|Q96LT7-2|CI072_HUMAN      FLANHTLNGEILRNAESGAIDVKFFVLSEKGVIIVSLIFDGNWNGDRSTYGLSIILPQTE        120
                             ************************************************************

sp|Q96LT7|CI072_HUMAN        LSFYLPLHRVCVDRLTHIIRKGRIWMHKERQENVQKIILEGTERMEDQGQSIIPMLTGEV        180
sp|Q96LT7-2|CI072_HUMAN      LSFYLPLHRVCVDRLTHIIRKGRIWMHKERQENVQKIILEGTERMEDQGQSIIPMLTGEV        180
                             ************************************************************

sp|Q96LT7|CI072_HUMAN        IPVMELLSSMKSHSVPEEIDIADTVLNDDDIGDSCHEGFLLNAISSHLQTCGCSVVVGSS        240
sp|Q96LT7-2|CI072_HUMAN      IPVMELLSSMKSHSVPEEIDIADTVLNDDDIGDSCHEGFLLK------------------        222
                             ****************************************:

sp|Q96LT7|CI072_HUMAN        AEKVNKIVRTLCLFLTPAERKCSRLCEAESSFKYESGLFVQGLLKDSTGSFVLPFRQVMY        300
sp|Q96LT7-2|CI072_HUMAN      ------------------------------------------------------------        222

sp|Q96LT7|CI072_HUMAN        APYPTTHIDVDVNTVKQMPPCHEHIYNQRRYMRSELTAFWRATSEEDMAQDTIIYTDESF        360
sp|Q96LT7-2|CI072_HUMAN      ------------------------------------------------------------        222

sp|Q96LT7|CI072_HUMAN        TPDLNIFQDVLHRDTLVKAFLDQVFQLKPGLSLRSTFLAQFLLVLHRKALTLIKYIEDDT        420
sp|Q96LT7-2|CI072_HUMAN      ------------------------------------------------------------        222

sp|Q96LT7|CI072_HUMAN        QKGKKPFKSLRNLKIDLDLTAEGDLNIIMALAEKIKPGLHSFIFGRPFYTSVQERDVLMT        480
sp|Q96LT7-2|CI072_HUMAN      ------------------------------------------------------------        222

sp|Q96LT7|CI072_HUMAN        F   481
sp|Q96LT7-2|CI072_HUMAN      -   222
```

Fig. 2: Alignment retrieved from the Clustal Omega of the related protein of mouse and bovine with a human which shows residues with * (identical) and: (somewhat similar).

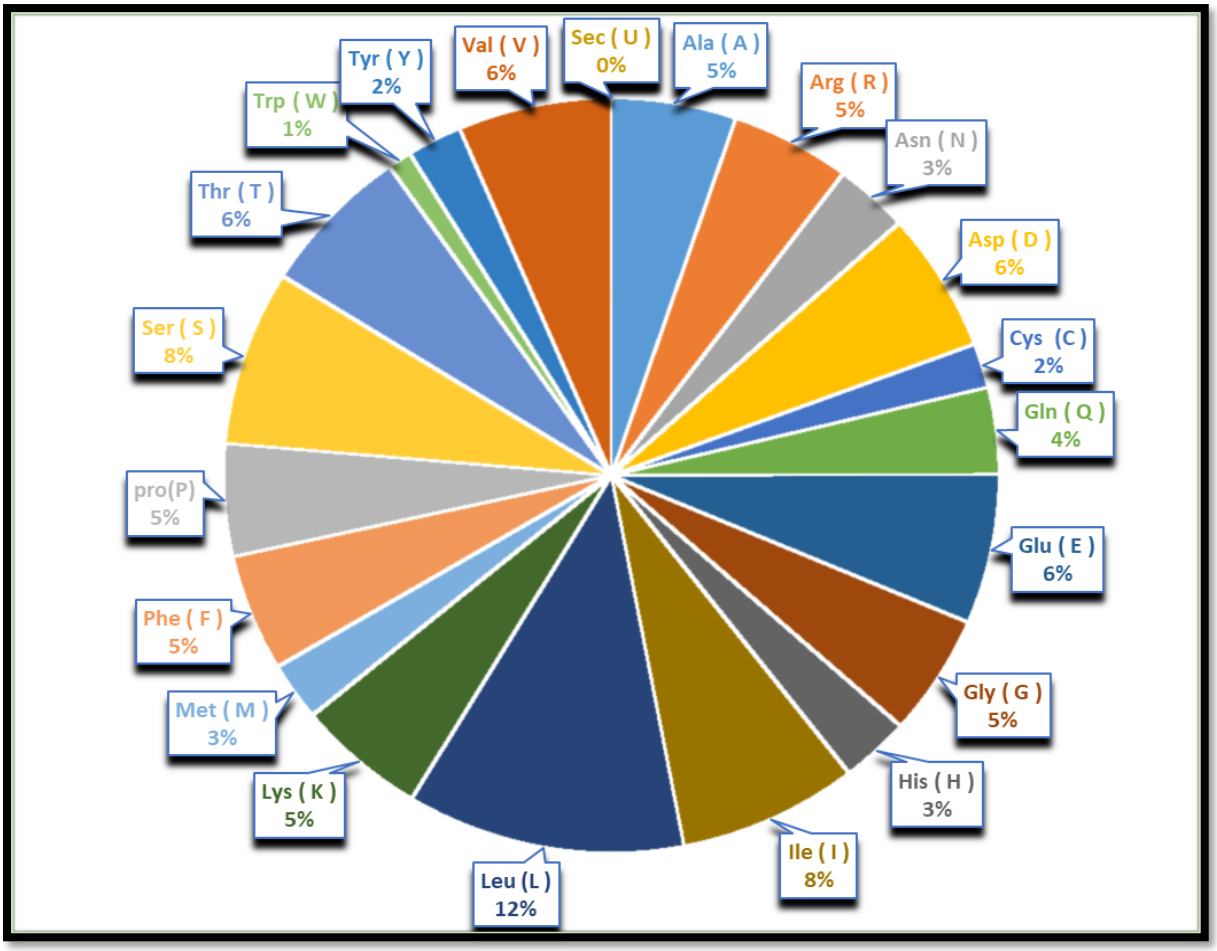

Fig. 3: Pie chart representation of amino acid composition of C9ORF72 and calculated percentage values.

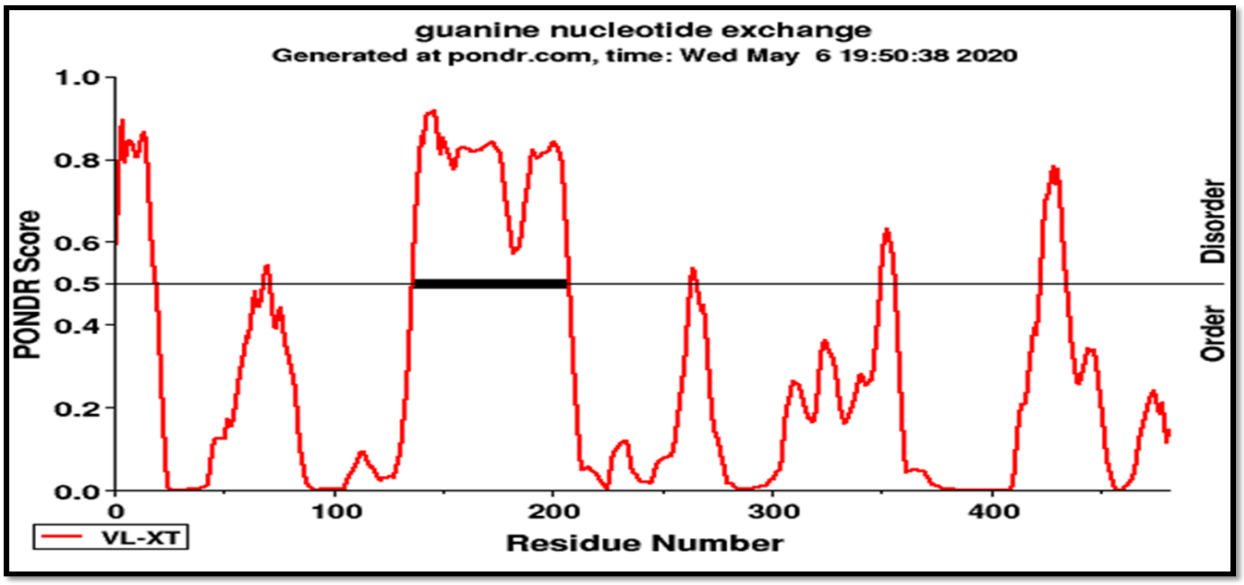

Fig. 4: Disorder residues of Guanine Nucleotide exchange C9ORF72.

were employed to predict the 3D structures. The BLASTp was used for the submission of protein sequence against PDB for the retrieval of suitable templates. Only one template has been appeared against the query sequence (**Table 2**).

Table 1: Calculated features of C9ORF72**.**

Features	Calculations
Aliphatic index	-0.069
Instability index	50.54
Total number of positive charge residues (Arg + Lys)	50
Total number of negative charge residues (Asp + Glu)	60
Total number of atoms	7683

All the 50 structures were evaluated on the bases of favored region, quality factor, allowed region and outlier regions. A comparative graph has been plotted to analyze the suitable structure among all the predicted structures. The suitable structure was selected from the plotted graph (**Fig. 5**).

There was variation in the quality factor values of the predicted structures and the selected structure of Guanine nucleotide exchange C9ORF72 has the overall quality factor of 92.38%. The Ramachandran plot was employed to evaluate the quality of the predicted structures and the selected structure has 99.60% value of the favored region, 0.40% of allowed region. Interestingly, no amino acid was observed in the outlier region. The energy minimization of the selected structure was performed to improve the stereochemical properties of the predicted structure. The minimization was performed by UCSF Chimera 1.13 with 1000 steepest decent and 1000 conjugate gradient runs (**Fig. 6**).

Protein-protein molecular docking analyses

Guanine Nucleotide Exchange expressed in many parts of the body specifically expressed in the brain. SMCR8, the interacting partner of C9ORF72 was analyzed and observed by employing STRING and STITCH databases for protein-protein molecular docking analyses. Comparative molecular docking analyses were done to evaluate the binding residues. The docked complexes of SMCR8 and C9ORF72 analyses predicted the interacting residues and analyzed on the basis of ACE [75] by utilizing PatchDock (Table 2). Numerous docked complexes were generated and then top 10 complexes with least ACE values were selected for further refinement through FireDock. The docked complexes were evaluated on the basis of least binding global energy. The molecular docking analyses suggested that SMCR8 and C9ORF72 have effective binding affinity [76]. The interacting residues were analyzed through UCSF Chimera (**Table 3**) (**Fig. 7**).

Table 2: BlastP against PDB.

Description	Max score	Total score	Query Coverage	E-Value	Identity	Accession
Cap-Associated protein CAF20	27.3	27.3	4%	7.84	5.83%	6FC3

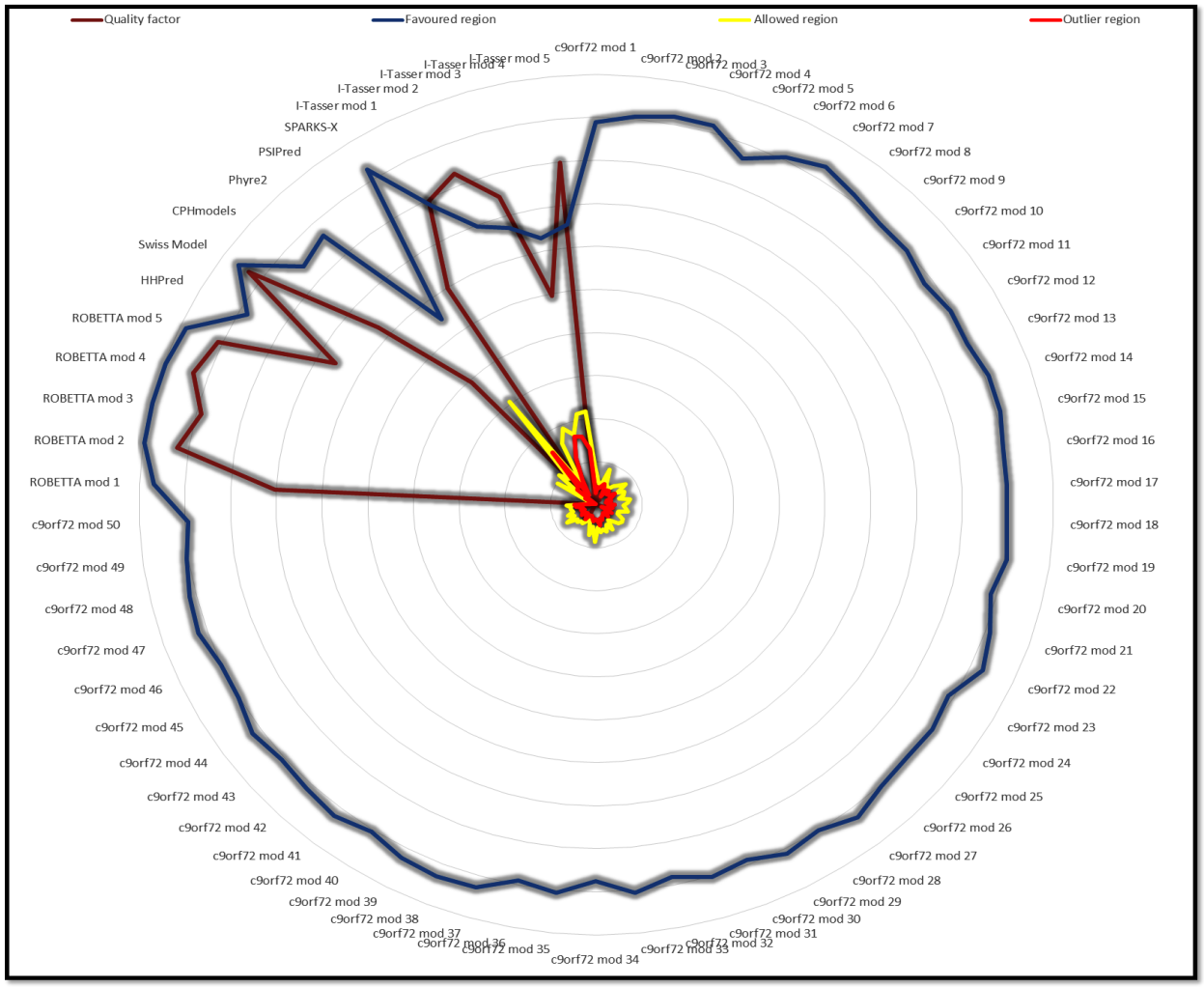

Fig. 5: Graph of quality factor, favored region, allowed region and outliers region of the C90RF72 structure prediction analyses evaluated through different software.

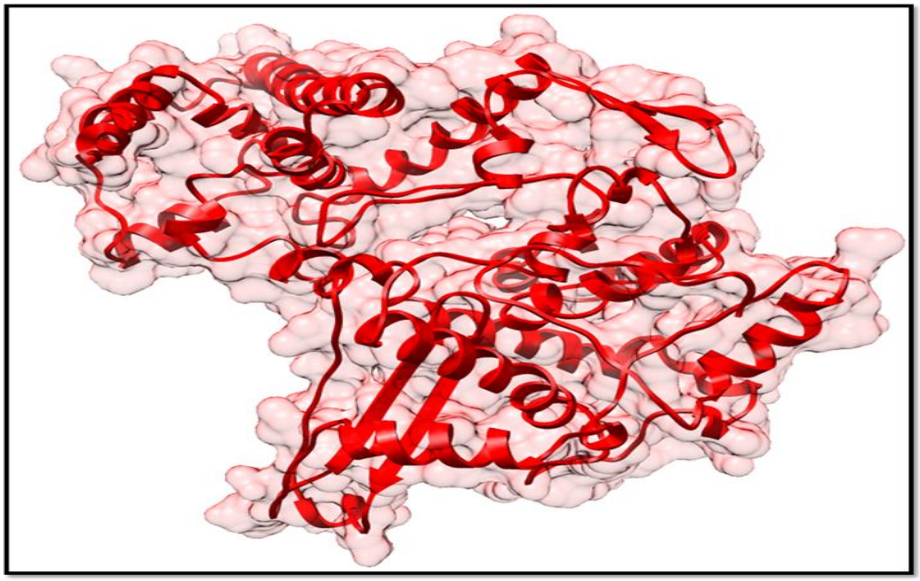

Fig. 6: 3D structure of ALS associated protein Guanine Nucleotide Exchange C9ORF72.

Table 3: Protein-protein interacting residues.

Interacting Protein	Interacting protein residues	Targeting Protein	Targeting protein residues
SMCR8	LYS 15, GLN 21, TYR 24, GLN17, LEU 18, ASN 20, ALA 19, ARG 43, ASN 23, TYR 39, LYS 42, GLU 36, GLY 35, SER 34, ARG 33, LYS 5, TYR 54	Guanine Nucleotide Exchange C9ORF72	PRO 467, THR 470, TYR 469, GLY 465, PHE 462, PHE 464, SER 361, VAL 384, PRO 389, LEU 402, SER 395, ALA 399, LEU 403, ARG 407, GLN 400, SER 392, ARG 329, TYR 326, LEU 391

Molecular docking analysis

The FDA library of ZINC databases was screened by using AutoDock Vina. After screening the FDA library of ZINC database, top ranked 4 compounds were observed. Comparative molecular docking analyses were done by utilizing AutoDock Vina (**Fig. 7**).

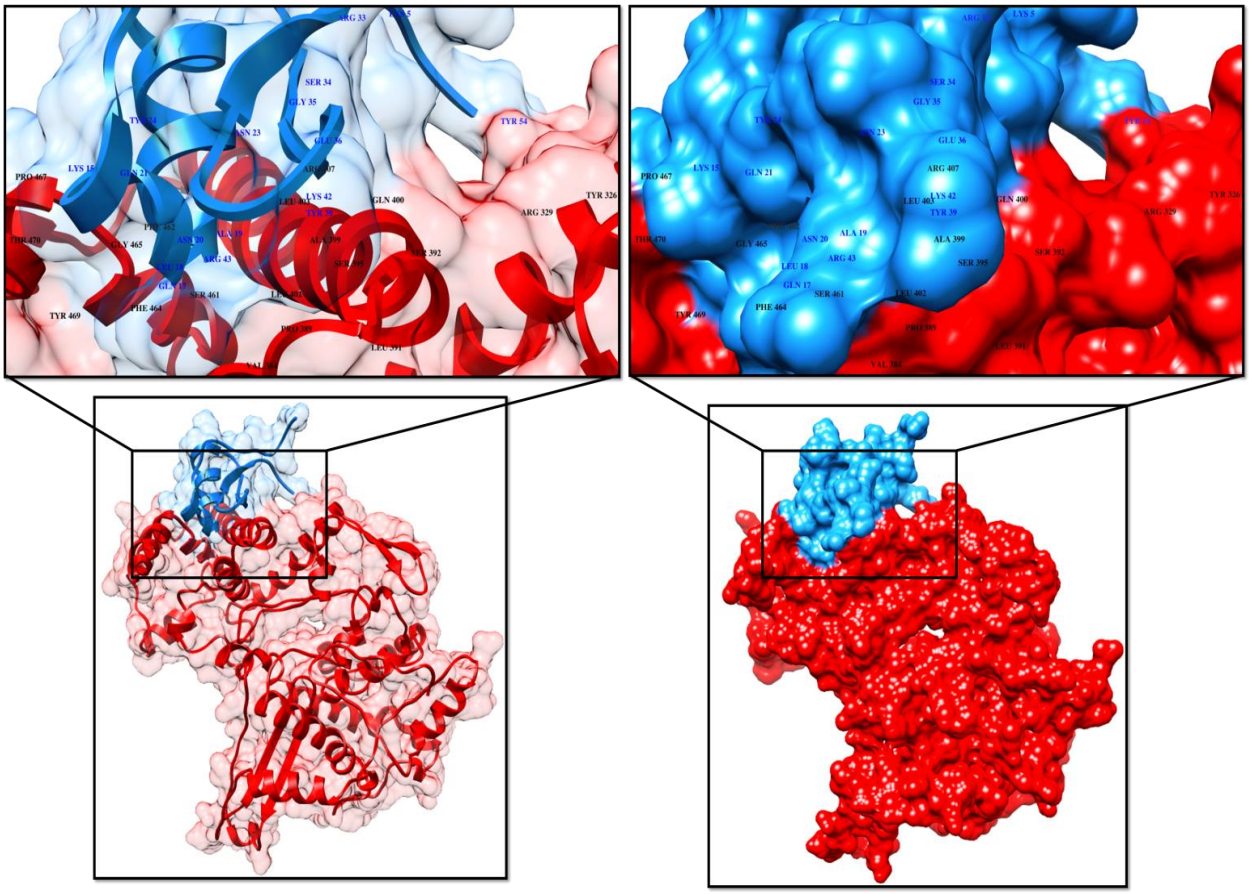

Fig. 7: Protein-Protein docking analysis and the interacting residues.

The generated docked complexes were ranked on the bases of least binding energy. The screened compounds showed similar binding domain. The selected compound may have potential against C9ORF72. The variation was observed in analyzed complexes having least binding energy. The compound ZINC131 showed least binding energy of -11.3 kcal/mol and 2D structure of the least binding affinity were develop from the ChemDraw Ultra 8.0. A plot of ligand-protein interactions was analyzed by employing UCSF Chimera (**Fig. 8**).

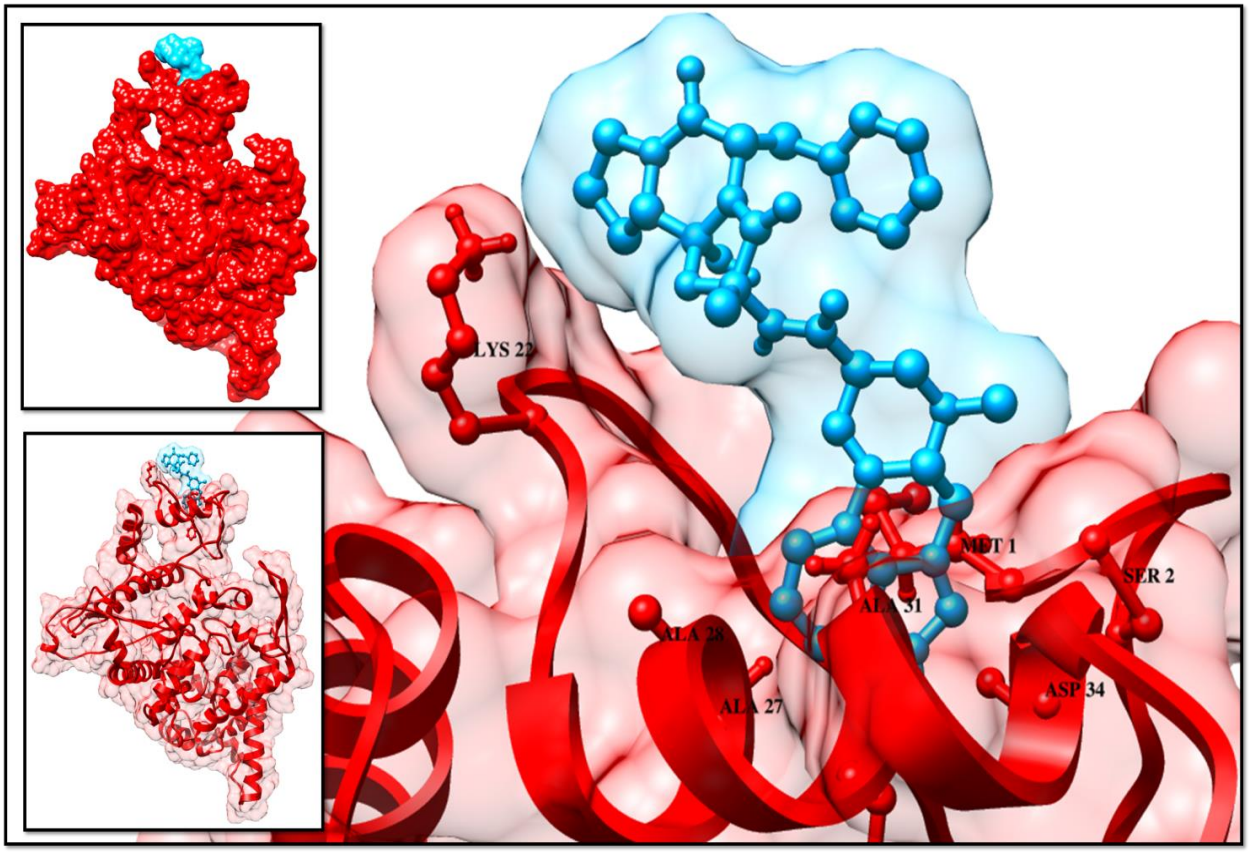

Fig. 8: The interactions of top ranked compound with C9ORF72. The residues analyzed from AutoDock Vina and UCSF Chimera.

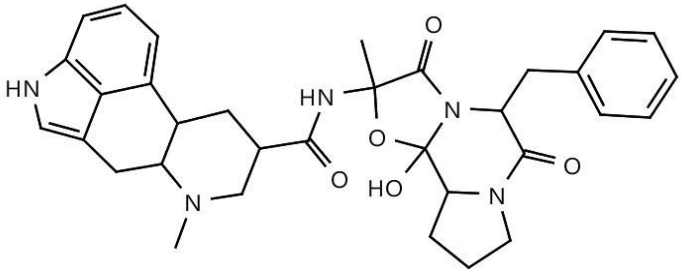

Fig. 9: 2D structure of least binding affinity compound from the molecular docking.

The function of protein depends on protein structures. The structural bioinformatics opens the way towards more progress in the analyses of protein function. The computational method of structure prediction is less time consuming [77]. The era of computational biology which is necessary for the prediction of function contributes well in the way of research.

Conclusion

By employing computational approaches and *in silico* analyses, the analyzed molecules showed binding residues in conserved region by AutoDock Vina. The *in silico* molecular docking analyses proposed that binding residues are significant to control the expression of C9ORF72. The observed results suggested that the selected molecule could be used for novel chemical compounds.

Acknowledgments

Authors are thankful to the Department of Bioinformatics and Biotechnology, Government College University Faisalabad to provide the research platform.

Conflict of interest

The authors declare no conflicts of interest.

References

[1] Rowland LP, Shneider NA. Amyotrophic lateral sclerosis. New England Journal of Medicine 2001;344(22):1688-1700.

[2] Kiernan MC, Vucic S, Cheah BC, Turner MR, Eisen A, Hardiman O, Burrell JR, Zoing MC. Amyotrophic lateral sclerosis. The lancet 2011;377(9769):942-955.

[3] Mitchell JD, Borasio GD. Amyotrophic lateral sclerosis. The lancet 2007;369(9578):2031-2041.

[4] Bede P, Iyer PM, Schuster C, Elamin M, Mclaughlin RL, Kenna K, Hardiman O. The selective anatomical vulnerability of ALS:'disease-defining'and 'disease-defying'brain regions. Amyotrophic Lateral Sclerosis and Frontotemporal Degeneration 2016;17(7-8):561-570.

[5] Shaw C, Quinn A, Daniel E. Amyotrophic lateral sclerosis/motor neurone disease. Palliative Care in Amyotrophic Lateral Sclerosis: From Diagnosis to Bereavement 2014:1.

[6] Webster CP, Smith EF, Bauer CS, Moller A, Hautbergue GM, Ferraiuolo L, Myszczynska MA, Higginbottom A, Walsh MJ, Whitworth AJ. The C9orf72 protein interacts with Rab1a and the ULK1 complex to regulate initiation of autophagy. The EMBO journal 2016;35(15):1656-1676.

[7] Matsumoto S, Kusaka H, Murakami N, Hashizume Y, Okazaki H, Hirano A. Basophilic inclusions in sporadic juvenile amyotrophic lateral sclerosis: an immunocytochemical and ultrastructural study. Acta neuropathologica 1992;83(6):579-583.

[8] Wijesekera LC, Leigh PN. Amyotrophic lateral sclerosis. Orphanet journal of rare diseases 2009;4(1):3.

[9] Sullivan PM, Zhou X, Robins AM, Paushter DH, Kim D, Smolka MB, Hu F. The ALS/FTLD associated protein C9orf72 associates with SMCR8 and WDR41 to regulate the autophagy-lysosome pathway. Acta neuropathologica communications 2016;4(1):51.

[10] Harrower T. The relationships between symptoms, disability, perceived health and quality of life in amyotrophic lateral sclerosis/motor neuron disease. 2019.

[11] Chiò A, Borghero G, Restagno G, Mora G, Drepper C, Traynor BJ, Sendtner M, Brunetti M, Ossola I, Calvo A. Clinical characteristics of patients with familial amyotrophic lateral sclerosis carrying the pathogenic GGGGCC hexanucleotide repeat expansion of C9ORF72. Brain 2012;135(3):784-793.

[12] Young CA, Ealing J, McDermott C, Williams T, Al-Chalabi A, Majeed T, Burke G, Pinto A, Dick D, Talbot K. The relationships between symptoms, disability, perceived health and quality of life in amyotrophic lateral sclerosis/motor neuron disease. Amyotrophic Lateral Sclerosis and Frontotemporal Degeneration 2019;20(5-6):317-327.

[13] van Es MA, Hardiman O, Chio A, Al-Chalabi A, Pasterkamp RJ, Veldink JH, Van den Berg LH. Amyotrophic lateral sclerosis. The Lancet 2017;390(10107):2084-2098.

[14] Pansarasa O, Gagliardi S, Sproviero D, Cereda C. RNA Metabolism and Therapeutics in Amyotrophic Lateral Sclerosis. Amyotrophic Lateral Sclerosis-Recent Advances and Therapeutic Challenges: IntechOpen; 2019.

[15] Zhu S, Wuolikainen A, Wu J, Öhman A, Wingsle G, Moritz T, Andersen PM, Forsgren L, Trupp M. Targeted Multiple Reaction Monitoring Analysis of CSF Identifies UCHL1 and GPNMB as Candidate Biomarkers for ALS. Journal of Molecular Neuroscience 2019;69(4):643-657.

[16] Fang T, Al Khleifat A, Meurgey J-H, Jones A, Leigh PN, Bensimon G, Al-Chalabi A. Stage at which riluzole treatment prolongs survival in patients with amyotrophic lateral sclerosis: a retrospective analysis of data from a dose-ranging study. The Lancet Neurology 2018;17(5):416-422.

[17] Rothstein JD. Edaravone: a new drug approved for ALS. Cell 2017;171(4):725.

[18] DeJesus-Hernandez M, Mackenzie IR, Boeve BF, Boxer AL, Baker M, Rutherford NJ, Nicholson AM, Finch NA, Flynn H, Adamson J. Expanded GGGGCC hexanucleotide repeat in noncoding region of C9ORF72 causes chromosome 9p-linked FTD and ALS. Neuron 2011;72(2):245-256.

[19] Waite AJ, Bäumer D, East S, Neal J, Morris HR, Ansorge O, Blake DJ. Reduced C9orf72 protein levels in frontal cortex of amyotrophic lateral sclerosis and frontotemporal degeneration brain with the C9ORF72 hexanucleotide repeat expansion. Neurobiology of aging 2014;35(7):1779. e5-1779. e13.

[20] Mori K, Weng S-M, Arzberger T, May S, Rentzsch K, Kremmer E, Schmid B, Kretzschmar HA, Cruts M, Van Broeckhoven C. The C9orf72 GGGGCC repeat is translated into aggregating dipeptide-repeat proteins in FTLD/ALS. Science 2013;339(6125):1335-1338.

[21] Zhu Q, Jiang J, Gendron TF, McAlonis-Downes M, Jiang L, Taylor A, Garcia SD, Dastidar SG, Rodriguez MJ, King P. Reduced C9ORF72 function exacerbates gain of toxicity from ALS/FTD-causing repeat expansion in C9orf72. Nature Neuroscience 2020:1-10.

[22] Wood H. A hexanucleotide repeat expansion in C9ORF72 links amyotrophic lateral sclerosis and frontotemporal dementia. Nature Reviews Neurology 2011;7(11):595-596.

[23] Laflamme C, McKeever P, Kumar R, Schwartz J, Kolahdouzan M, Chen CX-Q, You Z, Benaliouad F, Gileadi O, McBride HM. Implementation of an antibody validation procedure: Application to the major ALS/FTD disease gene C9ORF72. bioRxiv 2019:499350.

[24] Schludi MH, May S, Grässer FA, Rentzsch K, Kremmer E, Küpper C, Klopstock T, Alliance BBB, Arzberger T, Edbauer D. Distribution of dipeptide repeat proteins in cellular models and C9orf72 mutation cases suggests link to transcriptional silencing. Acta neuropathologica 2015;130(4):537-555.

[25] Vatsavayai SC, Nana AL, Yokoyama JS, Seeley WW. C9orf72-FTD/ALS pathogenesis: evidence from human neuropathological studies. Acta neuropathologica 2019;137(1):1-26.
[26] Cooper-Knock J, Bury JJ, Heath PR, Wyles M, Higginbottom A, Gelsthorpe C, Highley JR, Hautbergue G, Rattray M, Kirby J. C9ORF72 GGGGCC expanded repeats produce splicing dysregulation which correlates with disease severity in amyotrophic lateral sclerosis. PloS one 2015;10(5).
[27] Iyer S, Acharya KR, Subramanian V. A comparative bioinformatic analysis of C9orf72. PeerJ 2018;6:e4391.
[28] Fratta P, Mizielinska S, Nicoll AJ, Zloh M, Fisher EM, Parkinson G, Isaacs AM. C9orf72 hexanucleotide repeat associated with amyotrophic lateral sclerosis and frontotemporal dementia forms RNA G-quadruplexes. Scientific reports 2012;2:1016.
[29] Gómez-Tortosa E, Gallego J, Guerrero-López R, Marcos A, Gil-Neciga E, Sainz MJ, Díaz A, Franco-Macías E, Trujillo-Tiebas MJ, Ayuso C. C9ORF72 hexanucleotide expansions of 20–22 repeats are associated with frontotemporal deterioration. Neurology 2013;80(4):366-370.
[30] Ho WY, Tai YK, Chang J-C, Liang J, Tyan S-H, Chen S, Guan J-L, Zhou H, Shen H-M, Koo E. The ALS-FTD-linked gene product, C9orf72, regulates neuronal morphogenesis via autophagy. Autophagy 2019;15(5):827-842.
[31] Patel B, Singh V, Patel D. Structural Bioinformatics. Essentials of Bioinformatics, Volume I: Springer; 2019. p. 169-199.
[32] Chou K-C. Structural bioinformatics and its impact to biomedical science. Current medicinal chemistry 2004;11(16):2105-2134.
[33] Lee C, Wang Q. Bioinformatics analysis of alternative splicing. Briefings in bioinformatics 2005;6(1):23-33.
[34] Le Q, Torrisi M, Pollastri G. Deep learning methods in protein structure prediction. Computational and Structural Biotechnology Journal 2020.
[35] Kanakaveti V, Shanmugam A, Ramakrishnan C, Anoosha P, Sakthivel R, Rayala S, Gromiha MM. Computational approaches for identifying potential inhibitors on targeting protein interactions in drug discovery. Advances in protein chemistry and structural biology 2020;121:25.
[36] By Sheikh Arslan Sehgal AHM, Rana Adnan Tahir, Asif Mir. Quick Guideline for Computational Drug Design. Bentham Science Publisher 2018.
[37] Niland CN, Merry CR, Khalil AM. Emerging roles for long non-coding RNAs in cancer and neurological disorders. Frontiers in genetics 2012;3:25.
[38] Sehgal SA. Pharmacoinformatics and molecular docking studies reveal potential novel Proline Dehydrogenase (PRODH) compounds for Schizophrenia inhibition. Medicinal Chemistry Research 2017;26:314-326.
[39] Tahir RA, Wu H, Javed N, Khalique A, Khan SAF, Mir A, Ahmed MS, Barreto GE, Qing H, Ashraf GM, Sehgal SA. Pharmacoinformatics and molecular docking reveal potential drug candidates against Schizophrenia to target TAAR6. J Cell Physiol 2019;234(8):13263-13276.
[40] Boutet E, Lieberherr D, Tognolli M, Schneider M, Bairoch A. Uniprotkb/swiss-prot. Plant bioinformatics: Springer; 2007. p. 89-112.
[41] Berman H, Henrick K, Nakamura H. Announcing the worldwide protein data bank. Nature Structural & Molecular Biology 2003;10(12):980-980.
[42] Jacob A, Lancaster J, Buhler J, Harris B, Chamberlain RD. Mercury BLASTP: Accelerating protein sequence alignment. ACM Transactions on Reconfigurable Technology and Systems (TRETS) 2008;1(2):1-44.
[43] Eswar N, Webb B, Marti-Renom MA, Madhusudhan M, Eramian D, Shen My, Pieper U, Sali A. Comparative protein structure modeling using MODELLER. Current protocols in protein science 2007;50(1):2.9. 1-2.9. 31.
[44] Yang J, Yan R, Roy A, Xu D, Poisson J, Zhang Y. The I-TASSER Suite: protein structure and function prediction. Nature methods 2015;12(1):7.
[45] Peng J, Xu J. RaptorX: exploiting structure information for protein alignment by statistical inference. Proteins: Structure, Function, and Bioinformatics 2011;79(S10):161-171.
[46] Nielsen M, Lundegaard C, Lund O, Petersen TN. CPHmodels-3.0—remote homology modeling using structure-guided sequence profiles. Nucleic acids research 2010;38(suppl_2):W576-W581.
[47] Söding J, Biegert A, Lupas AN. The HHpred interactive server for protein homology detection and structure prediction. Nucleic acids research 2005;33(suppl_2):W244-W248.
[48] Kelley LA, Mezulis S, Yates CM, Wass MN, Sternberg MJ. The Phyre2 web portal for protein modeling, prediction and analysis. Nature protocols 2015;10(6):845.
[49] Schwede T, Kopp J, Guex N, Peitsch MC. SWISS-MODEL: an automated protein homology-modeling server. Nucleic acids research 2003;31(13):3381-3385.
[50] Pieper U, Eswar N, Braberg H, Madhusudhan MS, Davis FP, Stuart AC, Mirkovic N, Rossi A, Marti-Renom MA, Fiser A. MODBASE, a database of annotated comparative protein structure models, and associated resources. Nucleic acids research 2004;32(suppl_1):D217-D222.
[51] Kim DE, Chivian D, Baker D. Protein structure prediction and analysis using the Robetta server. Nucleic acids research 2004;32(suppl_2):W526-W531.
[52] Desai M, Chauhan JB. Predicting the functional and structural consequences of nsSNPs in human methionine synthase gene using computational tools. Systems biology in reproductive medicine 2019;65(4):288-300.
[53] Bates PA, Kelley LA, MacCallum RM, Sternberg MJ. Enhancement of protein modeling by human intervention in applying the automatic programs 3D-JIGSAW and 3D-PSSM. Proteins: Structure, Function, and Bioinformatics 2001;45(S5):39-46.
[54] Lambert C, Leonard N, De Bolle X, Depiereux E. ESyPred3D: Prediction of proteins 3D structures. Bioinformatics 2002;18(9):1250-1256.
[55] Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC, Ferrin TE. UCSF Chimera—a visualization system for exploratory research and

analysis. Journal of computational chemistry 2004;25(13):1605-1612.
[56] DeLano WL. PyMOL. 2002.
[57] Lovell SC, Davis IW, Arendall III WB, De Bakker PI, Word JM, Prisant MG, Richardson JS, Richardson DC. Structure validation by Cα geometry: ϕ, ψ and Cβ deviation. Proteins: Structure, Function, and Bioinformatics 2003;50(3):437-450.
[58] Melo F, Devos D, Depiereux E, Feytmans E, editors. ANOLEA: a www server to assess protein structures. Ismb; 1997.
[59] Laskowski RA, MacArthur MW, Moss DS, Thornton JM. PROCHECK: a program to check the stereochemical quality of protein structures. Journal of applied crystallography 1993;26(2):283-291.
[60] Colovos C, Yeates TO. Verification of protein structures: patterns of nonbonded atomic interactions. Protein science 1993;2(9):1511-1519.
[61] Dym O, Eisenberg D, Yeates T. ERRAT. 2012.
[62] Szklarczyk D, Franceschini A, Kuhn M, Simonovic M, Roth A, Minguez P, Doerks T, Stark M, Muller J, Bork P. The STRING database in 2011: functional interaction networks of proteins, globally integrated and scored. Nucleic acids research 2010;39(suppl_1):D561-D568.
[63] Szklarczyk D, Santos A, von Mering C, Jensen LJ, Bork P, Kuhn M. STITCH 5: augmenting protein–chemical interaction networks with tissue and affinity data. Nucleic acids research 2016;44(D1):D380-D384.
[64] Gong X, Wang P, Yang F, Chang S, Liu B, He H, Cao L, Xu X, Li C, Chen W. Protein–protein docking with binding site patch prediction and network-based terms enhanced combinatorial scoring. Proteins: Structure, Function, and Bioinformatics 2010;78(15):3150-3155.
[65] Schneidman-Duhovny D, Inbar Y, Nussinov R, Wolfson HJ. PatchDock and SymmDock: servers for rigid and symmetric docking. Nucleic acids research 2005;33(suppl_2):W363-W367.
[66] Mashiach E, Schneidman-Duhovny D, Andrusier N, Nussinov R, Wolfson HJ. FireDock: a web server for fast interaction refinement in molecular docking. Nucleic acids research 2008;36(suppl_2):W229-W232.
[67] Wallace AC, Laskowski RA, Thornton JM. LIGPLOT: a program to generate schematic diagrams of protein-ligand interactions. Protein engineering, design and selection 1995;8(2):127-134.
[68] Ding J, Arnold E. LIGPLOT. 2006.
[69] Dallakyan S, Olson AJ. Small-molecule library screening by docking with PyRx. Chemical biology: Springer; 2015. p. 243-250.
[70] Irwin JJ, Shoichet BK. ZINC− a free database of commercially available compounds for virtual screening. Journal of chemical information and modeling 2005;45(1):177-182.
[71] Tahir RA, Hassan F, Kareem A, Iftikhar U, Sehgal SA. Ligand-Based Pharmacophore Modeling and Virtual Screening to Discover Novel CYP1A1 Inhibitors. Curr Top Med Chem 2019;19(30):2782-2794.
[72] Sievers F, Higgins DG. Clustal omega. Current protocols in bioinformatics 2014;48(1):3.13. 1-3.13. 16.
[73] ProtParam E. ExPASy-ProtParam tool. 2017.
[74] Xue B, Dunbrack RL, Williams RW, Dunker AK, Uversky VN. PONDR-FIT: a meta-predictor of intrinsically disordered amino acids. Biochimica et Biophysica Acta (BBA)-Proteins and Proteomics 2010;1804(4):996-1010.
[75] Sehgal SA. Pharmacoinformatics, Adaptive Evolution, and Elucidation of Six Novel Compounds for Schizophrenia Treatment by Targeting DAOA (G72) Isoforms. Biomed Res Int 2017;2017:5925714.
[76] Tahir RA, Wu H, Rizwan MA, Jafar TH, Saleem S, Sehgal SA. Immunoinformatics and molecular docking studies reveal potential epitope-based peptide vaccine against DENV-NS3 protein. J Theor Biol 2018;459:162-170.
[77] Tahir RA, Bashir A, Yousaf MN, Ahmed A, Dali Y, Khan S, Sehgal SA. In Silico identification of angiotensin-converting enzyme inhibitory peptides from MRJP1. PLoS One 2020;15(2):e0228265.