History of Medicine:Vol.11 No.1(2025):35-49
https://doi.org/10.48047/HM.11.1.2025.35-49

Structural and Functional Annotation of Hypothetical Protein of Clostridium botulinum—Potential Drug Target

Hafiza Nida Shehzadi1, Neelum Shehzadi2, Mareena Kanwal3, Maryam Bukhari 4, Tehreem Shabbir1, Zaib Un Nisa5, *Saba Abbas1

[1]School of Medical Lab Technology, Minhaj University Lahore,

[2]Department of Basic and Applied Chemistry, University of Central Punjab Lahore,

[3]Department of Biotechnology, International Islamic University Islamabad,

[4]Department of Microbiology, University of Central Punjab Lahore

[5]School of Chemistry, Minjah University Lahore.

Correspondence;

*Saba Abbas

Email: sabaabbas786786@gmail.com

Abstract

Clostridium botulinum is a gram positive and obligate anaerobic bacterium have ability to produce a protein containing neurotoxicity and food poisoning characteristics. C. botulinum release a neurotoxin, which is a poison that attack the nervous system. It also causing food poisoning when food containing toxins are ingested. Using in-silico approach hypothetical protein KON12272.1 of Clostridium Botulinum is characterized. Using different servers and tools, protein physiochemical properties, disulfide bond, subcellular localization and, molecular interaction of protein-ligand complex were predicted. This protein homology was predicted and characterized to be more matching with tetratricopeptide repeated proteins and also show some matching with putative PEP- CTERM system TRP-repeat lipoprotein and also have virulence and pathogenic functions. Sub-cellular localization had confirmed that such virulent protein is lies in extracellular of the cell. Due to its pathogenic and disease causing factor, protein 3D structure got from I-TASSER, was docking against 6 ligands which acquired from PubChem, and ChEMBL. These ligands could be good candidates in future to fight against the food poisoning if the resistance to available drug had observed. Protein is virulent and so it was confirmed by docking which showed excellent molecular interaction with the protein. The best docking score were predicted against 4-Dechloro-4-(4- chlorophenyl) Loperamide was -9.7 KJ/Mol score. Results after molecular docking and ADMET showed the best drugs i.e. 4-Dechloro-4-(4-chlorophenyl) Loperamide, Cipro (Ciprofloxacin), 2,6-dimethyl-4-(2,3,4,5-tetrahydro-1H-1-benzazepin-3-yl) thiomorpholi. In future, these findings will be used to design in-vitro research methodology. The drugs candidate can be used to fight against the food poisoning.

Keywords: Clostridium botulinum, Tetratricopeptide repeat protein, Pathogens

Introduction

Clostridium botulinum is a gram positive, rod-shaped, anaerobic, and spore producing bacterium. Tetratricopeptide repeats (TRP) is a hypothetical protein which is found in Clostridium Botulinum. It produces toxins in food [1]. “Botulism” is a severe form of food poisoning results when we ingest these toxins. C. Botulinum is associated with canned food [2]. C. botulinum is widely distributed as a saprophyte in soil animal manure vegetables and sea mugs.

The degenerate sequence of tetratricopeptide repeats (TPR) that consist of 34 amino acids which occurs in different proteins in tandem repeats [3]. The tetratricopeptide repeat (TPR) structural motif was originally identified in yeasts [4]. TPR repeats leads to their identification in a large number of proteins in a broad variety of species. Each unit TPR motif consists of 2 helices, generally designated A-helix and B-helix. Being an all- helical domain, the TPR domain bears close structural similarity with other alpha-helical domains. In fact, in the Structural Classification of Proteins (SCOP), it is a member of the “TPR-like” superfamily, which in turn belongs to the “alpha-alpha superhelix” fold in the “all alpha protein” class [5, 6].

TPR-containing proteins were found to be significant participants in many diverse processes in eukaryotic cells, including peroxisomal targeting and import synaptic vesicle fusion and mitochondrial and chloroplast import [7]. TRPs are essential for many bacterial pathways, such as bio-mineralization of iron oxides in magnetotactic bacteria, outer membrane assembly, and pathogenesis. Mutations in TPR-containing proteins have been associated with a variety of human diseases, such as Leber's congenital amaurosis and chronic granulomatous disease [8].

TRP involve in various pathogenic functions like phagolysosomal maturation blockage, attachment with host cells and transfer the virulence factors into host cells [9]. In case of bacterial pathogenesis, Class II chaperons of the type III secretion system (TTSS) is widely studied TRP- containing proteins and required for the recognition [10]. It also stabilizes the two hydrophobic proteins known as translocators shown on the eukaryotic cell membrane as a pore. Translocators help the bacterium specific effectors enter into host cell [11]. In various pathogens pili act as virulence factor and involve in many functions such as locomotion, attachment, translocation of DNA, and biofilm formation [12].

TRP-containing proteins are essential for bacterial virulence mechanisms, which focus on the importance of TRP motifs. TRP involves assembling multiprotein complexes and protein- protein interaction [13]. TRP- containing proteins are involved in various biological processes, like transcriptional control, protein folding, neurogenesis, regulation of cell cycle, and transport of peroxisomal and mitochondrial protein [14].

Our main objective is to better understand the hypothetical protein and further drug targets by predicting its structure and functions via in-silico analysis.

Material and Methods

Sequence Retrieval

For sequence retrieval, NCBI (https://www.ncbi.nlm.nih.gov/) database was used in order to collect the FASTA sequence of hypothetical protein under accession number KON12272.1 of Clostridium botulinum with having 1094 amino acids residues. UniProt (https://www.uniprot.org/) database is used for functional annotation under accession no. ACA56520.1 [15].

Sequence Analysis

Sequence similarity was predicted by using Basic Local Alignment Search Tool (BLAST) (https://blast.ncbi.nlm.nih.gov/Blast.cgi) against our protein sequence. Homology predicted by using NCBI and BLASTp for protein database by comparing sequences of query and subject [16].

Conserved Domain Database and Family identification

Conserved domains was detected by four tools using NCBI CD-Search (https://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi), while the other three tools i.e. Interproscan (https://www.ebi.ac.uk/interpro/search/sequence/), pfam (https://pfam.xfam.org/), Cath3D (https://www.cathdb.info/search) were used for protein functional analysis that classified them into families [15].

Predicting Physiochemical properties

For predicting physical and chemical properties of interest protein, ExPASy Protparam (https://web.expasy.org/protparam/) was used. The detection of amino composition molecular -weight, atomic composition, theoretical pl, half-life, total number of positively and negatively charged residues, and also instability index also by using ExPASy Protparam [17].
Analyzing Subcellular Localization

Subcellular localization was done to find the cellular functions of protein by using Cello (http://cello.life.nctu.edu.tw/) . Subcellular localization means that a protein can be present in inner or outer membrane, periplasm, cytoplasm or in extracellular space. Cello also was used to predict the extracellular partitioned sequences, peptides and also composition of amino acid [18].

Predicting Disulfide Bonds

In order to find the protein folding in protein tertiary structure, disulfide bonds prediction was done with the help of design 2 tool (http://cptweb.cpt.wayne.edu/DbD2/) [19].

Predicting Secondary and Tertiary Structure

For predicting protein secondary structure, Psipred (http://bioinf.cs.ucl.ac.uk/psipred/) was used which inestegated about the position of alpha helix and coils number. For getting 3D structure of protein, I-Tasser was used. For this purpose, hypothetical protein with having 1094 residues by ID S 691817 was submitted at (https://seq2fun.dcmb.med.umich.edu//I-TASSER/output/S691817/) by using I-Tasser [20].

Predicting three dimensional structure refinement and validation

For protein 3D structure refinement, Galaxy- Refine (https://galaxy.seoklab.org/cgi- bin/submit.cgi?type=REFINE) was used. For the validation of protein 3D structure, ERRAT (https://saves.mbi.ucla.edu/) and Ramachandran Plot was used [21, 22].

Predicting Protein Network Map

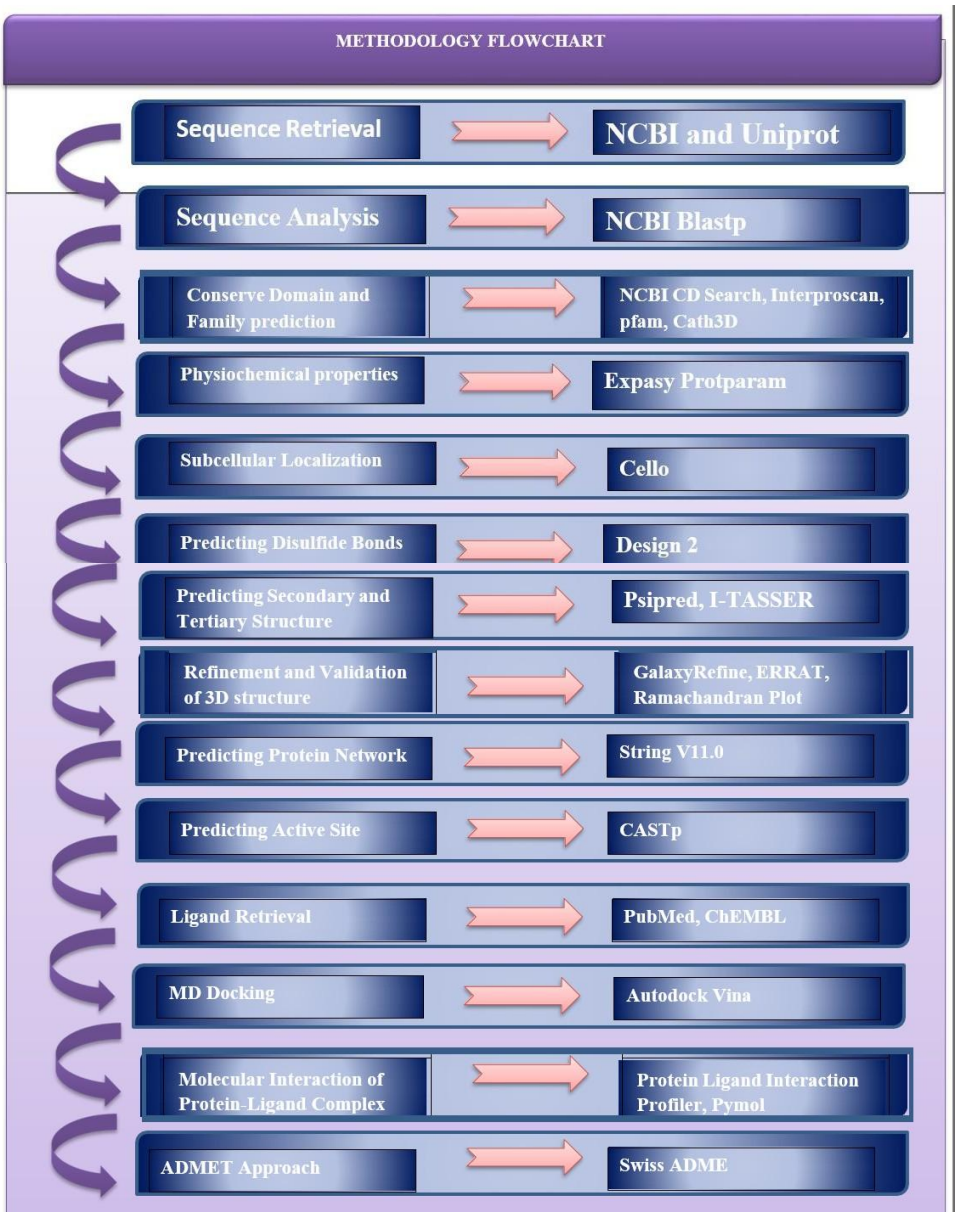
Protein-protein interaction was predicted using STRING (https://string-db.org/). Protein-protein interaction helps to predict the target protein function and also molecules drug ability [23].

MD Docking

Ligand retrieval from using PubChem (https://pubchem.ncbi.nlm.nih.gov/) and ChEMBL (https://www.ebi.ac.uk/chembl/). For predicting active site of the protein for docking, CASTp (Computed Atlas of Surface Topography of proteins) was used (http:/sts.bioe.uic.ed u/castp/index.html.2011). The docking of protein ligand predicted by molecular modeling simulation software i.e. Autodoc Vina [24]. Protein Ligand Interaction Profiler (PLIP) (https://plip- tool.biotec.tu-dresden.de/plip-web/plip/index) and Pymol was also used for predicting the molecular interaction of protein-ligand complex [25].

ADMET Approach

w3aqAThe most efficient tool for drug deigning SwissADME (http://www.swissadme.ch/) was used for predicted pharmacokinetics of small molecules. It expresses molar solubility, absorption level, hepatotoxicity and plasma protein binding level about drug [26].



Results

Sequence Retrieval

The FASTA sequence of hypothetical protein under accession number KON12272.1 of Clostridium botulinum with having 1094 amino acids residues is obtained from NCBI.

Sequence Analysis

NCBI BLASTp is used in order to analyze the protein sequence homology. Our protein sequence matched with tetratricopeptide repeat protein with having querry coverage of 97% and per. Identity 99.91%.

Conserve Domain Database and Family identification

By using NCBI CD search and Cath3D, we predict the conserve domain database and they show that our protein belongs to tetratricopeptide repeat protein. While the Pfam and interproscan illustrate the two protein matching and also identify family. The first one was putative PEP- CTERM system TRP-repeat lipoprotein and other was tetratricopeptide repeat. These two tools show more matching with tetratricopeptide repeat proteins.

Physiochemical properties

Expasy ProtParam is used in order to find out the physiochemical properties of our protein with accession no. KON12272. By using this tool, we predicted the number of amino acids, molecular weight, theoretical pIh, Total number of negatively charged residues (Asp + Glu), total no. of positively charged residues (Arg + Lys), total number of atoms, aliphatic index, grand average of hydropathicity (GRAVY) are 1094, 130771.73, 4.74, 229, 155, 18139, 78.17, -0.791
respectively. Following table shows the results of Expasy ProtParam.

Table # 01 Expasy ProtParam results

Parameters	Values
No. of Amino Acids	1094
Molecular wt.	130771.73
Theoretical pI	4.74
Total no. of negatively charged residues (Asp + Glu)	229
Total no. of positively charges residues (Arg + Lys)	155
Atomic Composition	Carbon (C) 5908 Hydrogen (H) 8915 Nitrogen (N) 1441 Oxygen (O) 1835 Sulfur (S) 40
Formula	$C_{5908}H_{8915}N_{1441}O_{1835}S_{40}$
Total no. of atoms	18139
Aliphatic index	78.17
Grand average of hydropathicity (GRAVY)	-0.791

Table 2: CELLO results

Analysis Report: SVM	LOCALIZATION	RELIABILITY
Amino Acid Comp.	Cytoplasmic Extracellular	0.647
N-peptide Comp.		0.592
Partitioned seq. Comp.		0.742
Physico-chemical Comp.		0.350
Neighboring seq. Comp.	Cytoplasmic Extracellular . Cytoplasmic	0.751

Prediction of Disulfide Bond

Design 2 tool was used in order to predict the disulfide bond. Conformation of protein occur when bond is formed at the cysteine residue. Dynamics can be changed. Design 2 tool was used to increase the stability and protein functional characteristics. Fig 1 shown the disulfide bonds by yellow color. A total of 60 cysteine bonds for protein turning were analyzed shown in Fig 2.

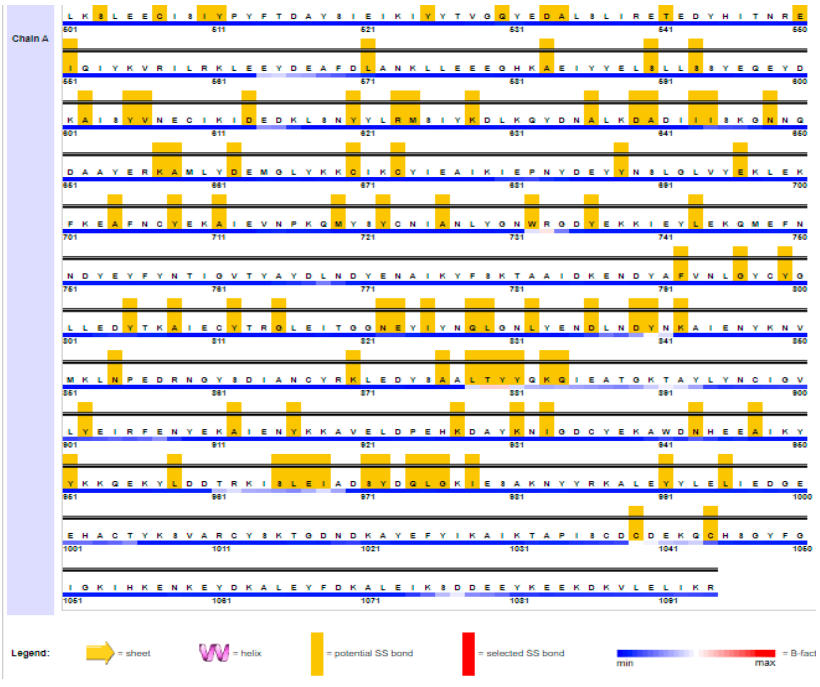


Figure 1: 2Disulfide illustration of all 60 disulfide bonds

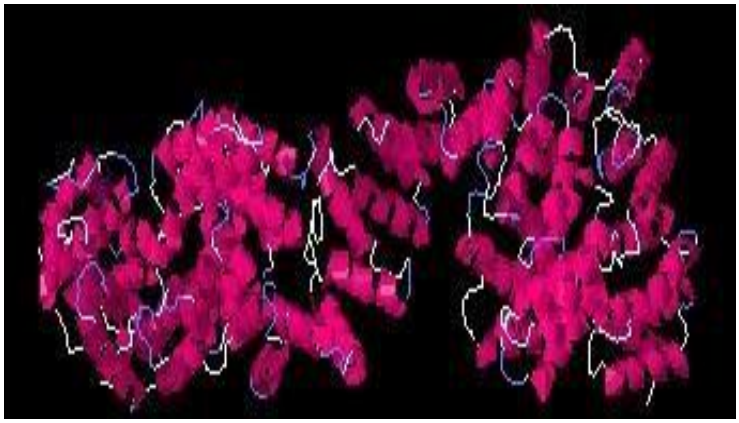


Fig. 2: Disulfund for Hypothetical protein—3D illustration

Prediction of Secondary and Tertiary Structure

For predicting protein secondary structure, PsiPred server was used which predicted that secondary structure has more alpha-helix than beta strands shown in Fig. 3.

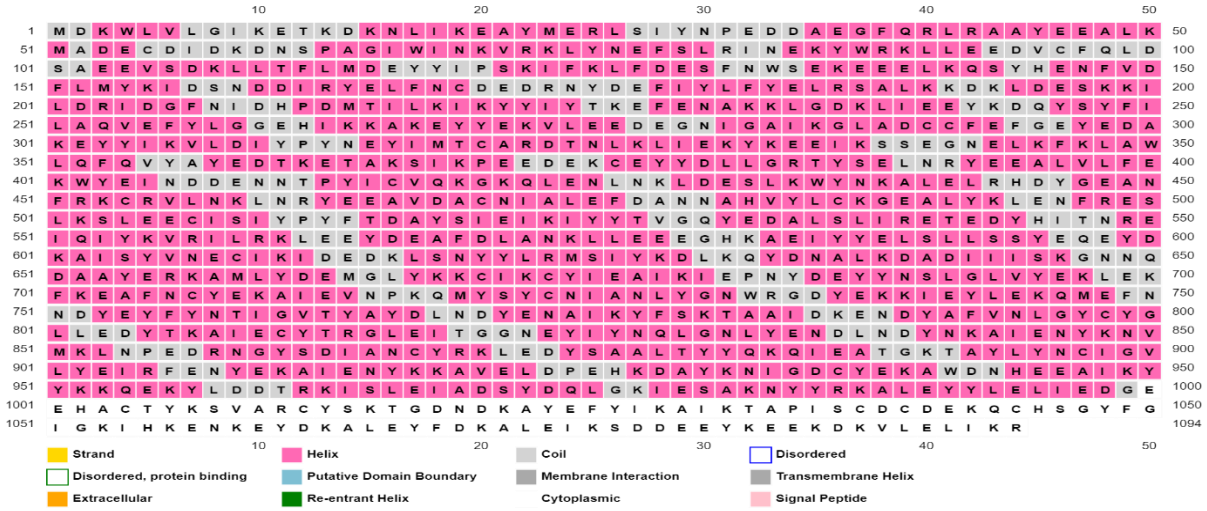


Fig 3: PsiPred 2D structure prediction

For predicting that protein is either Intracellular or Extracellular, Memstat was used. Memstat results illustrated that hypothetical protein is extracellular shown in Fig 4 (A) and Fig 4 (B).

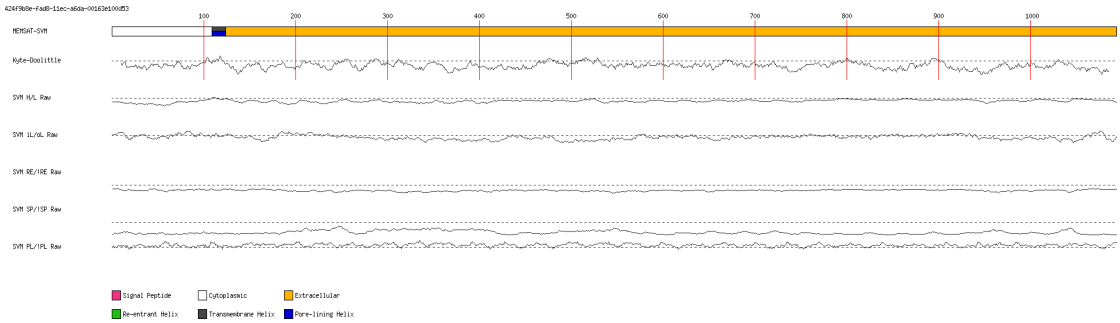


Fig 4 (A) Memstat graphical illustration

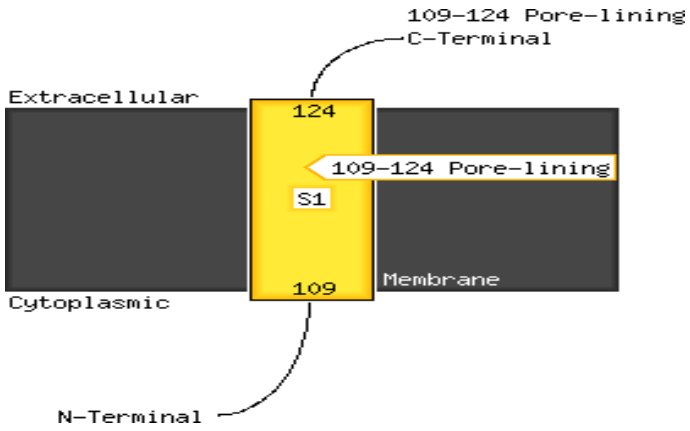


Fig 4 (B): Memstat

For predicting 3D structure, I-TASSER was used, whose results showed five best structure with TM score 0.71 ± 0.12, predicted RSMD was 9.2 ± 4.6 Å, along which the structure with C score
0.06 was selected. It helps to analyzed protein as well as also predicted homology with tetratricopeptide repeat proteins and gave us the 3D model.

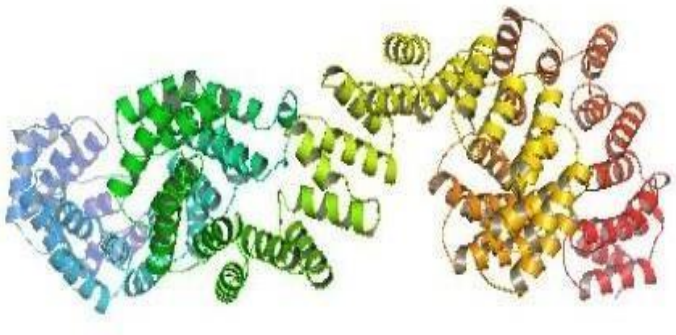


Fig 5: 3D structure of HP KON12272.1— I-TASSER Prediction

Protein 3D Structure Validation

For validating the 3D structure of protein, Ramachandran plot was used. Yellow color showed allowed region, the red color shows the most allowed region, while the rest regions illustrate the disallowed region shown in Fig. 6. Plot statistics showed the residues in most favored region contain 79 %, while additional allowed regions contains 14.7 %. Residues in generously allowed regions contains
4.0 %, while the disallowed regions have 1.8 % residues. Glycine and proline residues are 40 and 13 respectively.

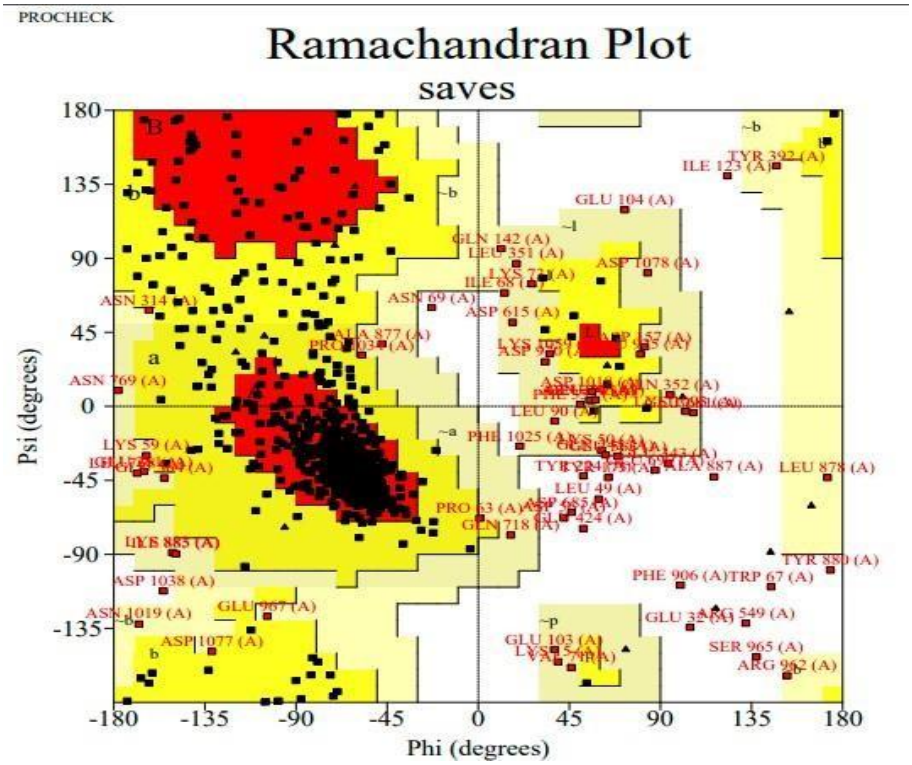


Fig 6: Ramachandran Plot

ERRAT was used to validate the 3D structure of protein which predicted that the protein non bonded interaction in the form of error. ERRAT result showed 92.081 overall quality factor as shown in Fig 7.

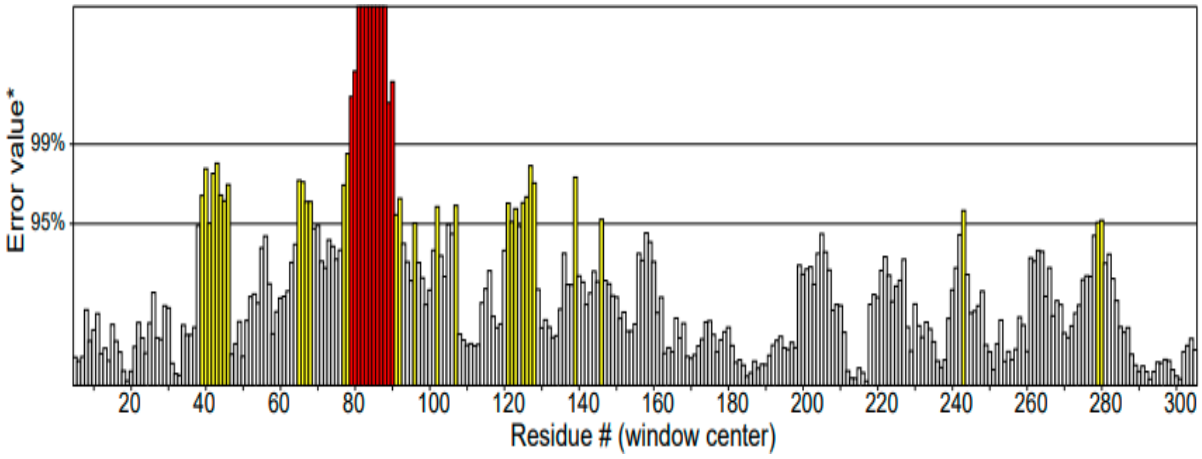


Fig 7: ERRAT graphical representation

Refinement of protein 3D structure

GalaxyRefine server was used for refinement of protein 3D structure. Below table# 03 showed refine scores i.e.

Table # 03: Refinement scores by GalaxyRefine Web server

Model	GDT-HA	RMSD	MolProbity	Clash score	Poor rotamers	Rama favored
Initial	1.0000	0.000	3.438	16.1	20.6	80.3
MODEL 1	0.9383	0.450	2.238	21.0	1.0	93.7
MODEL 2	0.9333	0.467	2.250	19.8	1.1	93.6
MODEL 3	0.9390	0.449	2.344	21.6	1.4	94.0
MODEL 4	0.9378	0.455	2.290	21.0	1.2	93.9
MODEL 5	0.9381	0.458	2.192	20.5	1.0	94.4

Protein Network Map

For predicting protein-protein interactions, String V11.0 was used. Metabolic process pathways illustrated by this tool and showed by interacting protein with other proteins in cell, body and processes and form a network. Network showed 11 nodes and 37 edges by String are as shown in Fig 8.

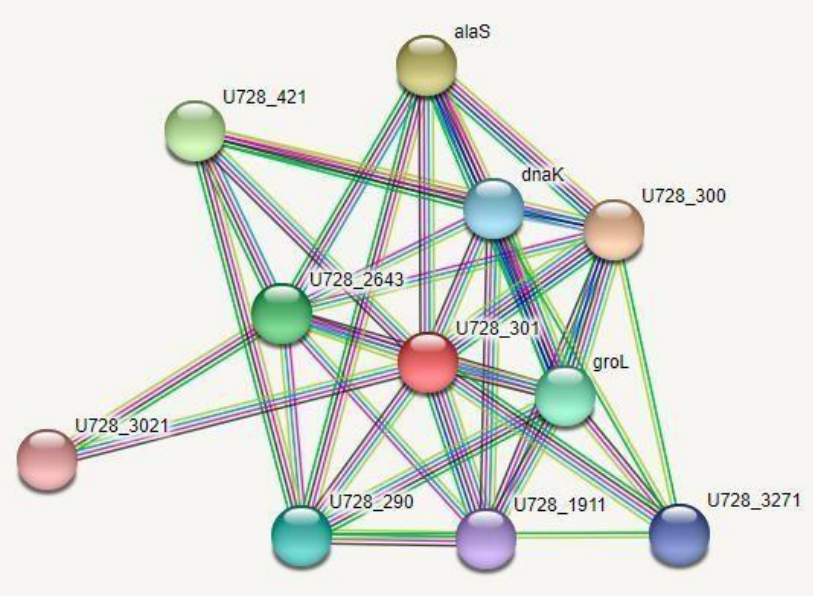


Fig 8. String V11.0 shows Protein Network Map

MD DOCKING

1. Ligand Retrieval

For ligand retrieval, PubChem and ChEMBL was used according to the targeted protein. Following table shown PubChem CID and ChEMBL IDs.

Table # 4: Ligands

No.	Name	PubChem CID	ChEMBL ID
1.	Loperamide hydrochloride	71420	CHEMBL1707
2.	4-Dechloro-4-(4-chlorophenyl) Loperamide	71315245	CHEMBL503879
3.	2,6-dimethyl-4-(2,3,4,5-tetrahydro-1H-1-benzazepin-3-yl) thiomorpholi	103314161	CHEMBL1773903
4	Ethyl 5-pentyl-2-thioxo-1,3-dithiole-4-carboxylate	387186199	CHEMBL8659
5.	Cipro (Ciprofloxacin)	2764	CHEMBL8
6.	N-Desmethyl Loperamide-d3	46781171	CHEMBL1459

2. Active Site Prediction

The active site of protein was predicted by using CASTp tool. 3D structure of hypothetical protein taken fro TASSER. CASTp predicted in below figure 9 showed that 127 residues in pocket of active site and also involved in ligand interaction.

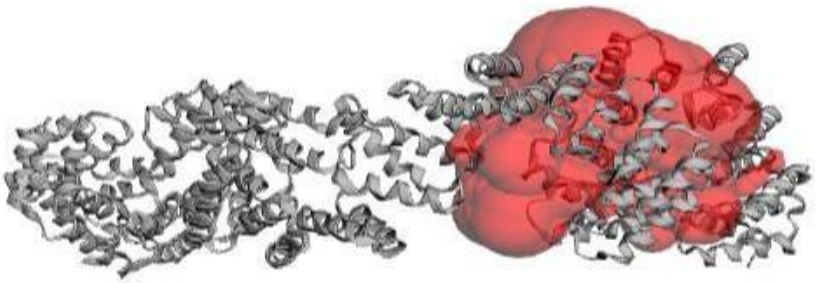


Fig 9: Active site prediction by CASTp

3. Energies Prediction

For predicting energies, Autodoc vina tools used for docking of ligand with the protein. Grid box was set according to active site predicted by CASTp. Below table # 5 shows that the results of Autodoc vina was the best dock score of Hypothetical Protein with 4-Dechloro-4- (4-chlorophenyl) Loperamide having -9.7 KJ/Mol energy. Other best score were 2,6-dimethyl-4- (2,3,4,5- tetrahydro-1H-1-benzazepin-3-y1) thiomorpholi and Cipro (Ciprofloxacin).

No.	Ligands	Affinity (kcal/mol)
1.	Loperamide hydrochloride	-4.8
2.	4-Dechloro-4-(4-chlorophenyl) Loperamide	-9.7
3.	2,6-dimethyl-4-(2,3,4,5-tetrahydro-1H-1-benzazepin-3-yl) thiomorpholi	-7.3
4	Ethyl 5-pentyl-2-thioxo-1,3-dithiole-4-carboxylate	-5.5
5.	Cipro (Ciprofloxacin)	-7.6
6.	N-Desmethyl Loperamide-d3	-4.7

4. Molecular interaction of protein-ligand complex

Protein Ligand Interaction Profiler (PLIP) and Pymol were used for predicted the molecular interaction of protein-ligand complex. Fig. 10 (A) shows protein ligand interaction while Fig. 10 (B) shows the Pymol result by using Protein Ligand Interaction Profiler (PLIP)

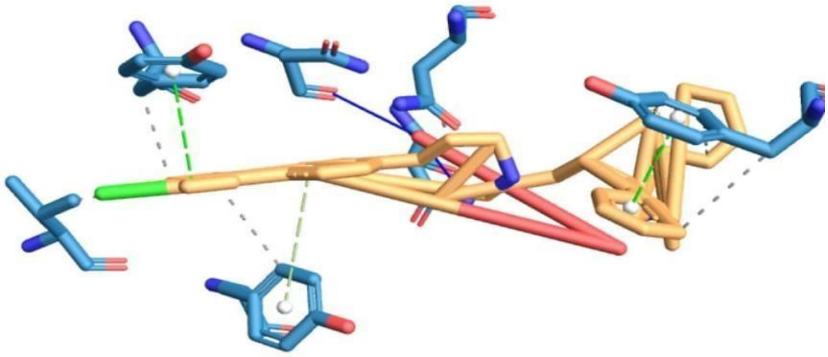


Fig # 10 (A): PLIP results showing molecular interaction of protein-ligand complex

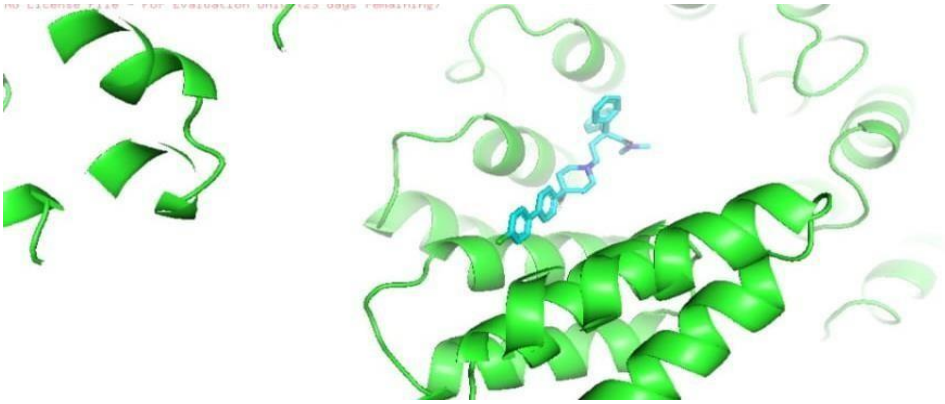


Fig. # 10 (B): Pymol results

5. ADMET Approach

In order to perform ADMET approach on ligand, SwissADME were used. Boiled- Egg diagram taken by using SwissADME which illustrate that which drug crosses the Gastrointestinal tract (PGP) and Blood-Brain Barrier (BBB). Boiled-Egg diagram shown that 3 drugs cross BBB. Out of 5 drugs, 4 drugs don't cross the PDP+ and 1 drug cross the PGP-.

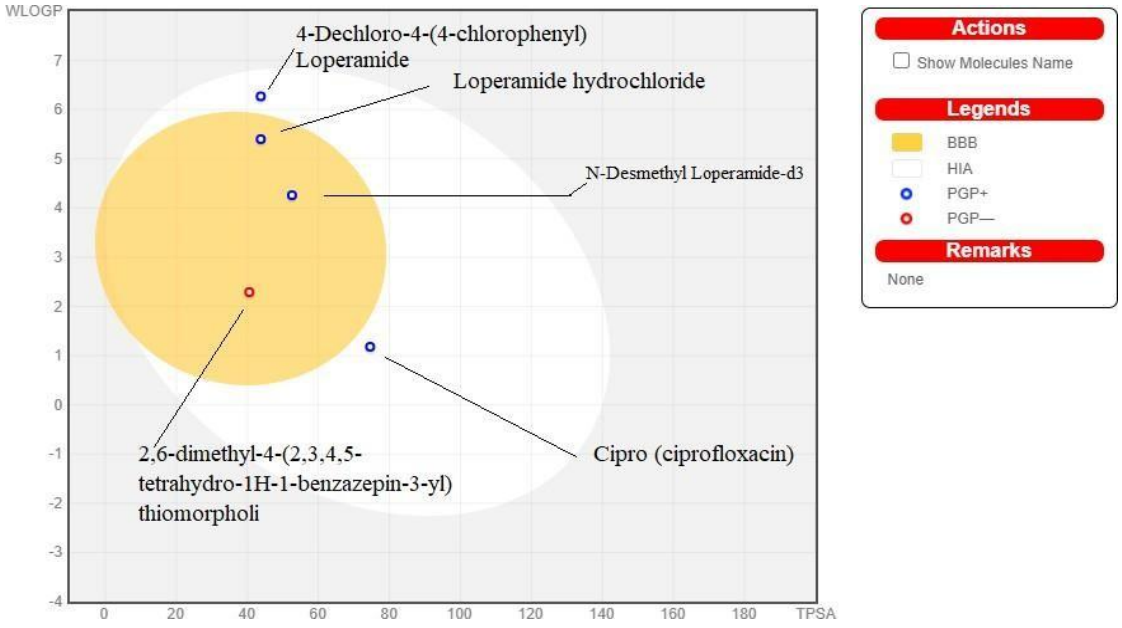


Fig. # 13: SwissADME showing Boiled-Egg

Discussions

TRP-containing proteins of bacterial pathogens are directly involved in pathogenic and virulence functions indicated by many reports in recent years [27]. In-silico analysis of hypothetical protein KON12272.1 and its physiochemical properties illustrated that this hypothetical protein showing negative GRAVY point and is hydrophilic. Through subcellular localization, it predicted that our hypothetical protein is extracellular. Virulent proteins lie in the extracellular region of many pathogenic bacteria such as Clostridium, mycobacterium, and streptococcus etc. Family prediction and homology was done and it shows more matching with tetratricopeptide repeat protein and some matching with putative PEP-CTERM system TRP-repeat lipoprotein. The protein associated with these families have pathogen and virulence characteristics.

Lot of protein folding involved in the protein tertiary structure shown by lot of cysteine binding that were predicted and confirmed by disulfide bond analysis. A total 60 cystiene bonds were predicted that this protein is cycteine-rich and very stable. For better docking results, there is a need of refinement of 2D and 3D structures. I-TASSER results illustrated that this protein is pathogenic and have virulence characters, which is consistent with previous studies that TRP- containing proteins play a critical role in pathogenicity and causing disease [27, 28].

In comparison with other studies, the results of this study are consistent with the idea that TRP- containing proteins plays a critical role in pathogenicity and causing disease. For example, a study by [29] showed that the TRP-containing protein, InvH, is involved in the regulation of flagellar gene expression in Salmonella typhimurium. Similarly, a study by [30] showed that the TRP- containing protein, YfiH, is involved in the regulation of flagellar gene expression in Escherichia coli.

Till now its pathway is unknown and in-vitro work is required to find its pathway, but the fate of protein functions had understand by protein network mapping. Structural and functional annotation of hypothetical protein has revealed that it is a pathogenic protein and has virulence factor and is predicted through protein interaction mapping and docking. All ligands related to drugs against food poisoning and antimicrobials are got dock against it. In the future, they may be used as the best drug candidate against food poisoning. These all results directed us towards that tetratricopeptide repeat protein plays a critical role in pathogenicity and causing disease [27]. Molecular interaction of protein ligand complex and ADME approach were used to predict the best drug candidate among these. All the results are variable, but these drugs can be used against food poisoning in the future to fight against the disease if drug resistance has seemed in Clostridium botulinum.

Conclusion

Structural and functional annotation of Hypothetical protein KON12272.1 illustrated that it is cysteine-rich and stable protein. 2D and 3D structure confirms that this protein is hydrophilic and extracellular. Pathway and cell fate of this protein showed that tetratricopeptide repeat protein is a virulent protein and involved in food poisoning. The best drug after molecular docking and ADME approach are reported as 4-Dechloro-4-(4-chlorophenyl) Loperamide, Cipro (Ciprofloxacin), 2,6-dimethyl-4-(2,3,4,5-tetrahydro-1H-1-benzazepin-3-yl) thiomorpholi. In the future, these drugs will be used to fight against food poisoning caused by Clostridium botulinum.

References:

1) Cerveny, L., Straskova, A., Dankova, V., Hartlova, A., Ceckova, M., Staud, F., & Stulik, J. (2013). Tetratricopeptide repeat motifs in the world of bacterial pathogens: role in virulence mechanisms. Infection and immunity, 81(3), 629- 635..

2) Perez-Riba A, Itzhaki LS. (2019). The tetratricopeptide-repeat motif is a versatile platform that enables diverse modes of molecular recognition. Curr Opin Struct Biol.; 54:43-49.

3) Barik, S. (2019). Protein tetratricopeptide repeat and the companion non- tetratricopeptide repeat helices: Bioinformatic analysis of interhelical interactions. Bioinformatics and Biology Insights, 13, 11.

4) Zeytuni, N., & Zarivach, R. (2012). Structural and functional discussion of the tetra-trico-peptide repeat, a protein interaction module. Structure, 20(3), 397- 405.

5) Main ER, Stott K, Jackson SE, Regan L. (2015). Local and long-range stability in tandemly arrayed tetratricopeptide repeats. Proc Natl Acad Sci USA; 102:5721- 5726.

6) Cauerhff, A., & Galigniana, M. D. (2018). Structural characteristics of the TPR protein-Hsp90 interaction: A new target in biotechnology. Bentham Science 1, 73-173.

7) Natalie Zeytuni, Raz Zarivach,. (2013). Structural and Functional Discussion of the Tetra-TricoPeptide Repeat, a Protein Interaction Module, Structure, 20(3).

8) Albert Perez-Riba, Laura S Itzhaki, (2019). The tetratricopeptide-repeat motif is a versatile platform that enables diverse modes of molecular recognition. Current Opinion in Structural Biology, 54 43-49.

9) Bröms JE, Edqvist PJ, Forsberg A, Francis MS. 2006. Tetratricopeptide repeats are essential for PcrH chaperone function in Pseudomonas aeruginosa type III secretion. FEMS Microbiol. 25, 57–66.

10) Christie, P. J. (2019). The rich tapestry of bacterial protein translocation systems. The protein journal, 38, 389-408.

11) Mueller CA, Broz P, and Cornelis GR. 2008. The type III secretion system tip complex and translocon. Mol. Microbiol. 68, 1085–1095.

12) Salomonsson EN, Forslund A-L, and Forsberg A. 2011. Type IV pili in Francisella—a virulence trait in an intracellular pathogen. Front. Microbiol. 2, 29-35.

13) Graham, J. B., Canniff, N. P., & Hebert, D. N. (2019). TPR-containing proteins control protein organization and homeostasis for the endoplasmic reticulum. Critical reviews in biochemistry and molecular biology, 54(2), 103- 118.

14) Trudova, E. (2024). A Taste of Bioinformatics: Exploring the Pathogenetics of Clostridium botulinum.

15) Bhardwaj, T., Haque, S., & Somvanshi, P. (2018). In silico identification of molecular mimics involved in the pathogenesis of Clostridium botulinum ATCC 3502 strain. Microbial pathogenesis, 121, 238-244.

16) Tan, C., Zhu, F., Xiao, Y., Wu, Y., Meng, X., Liu, S., ... & Wu, A. (2022).Immunoinformatics approach toward the introduction of a novel multi-epitope vaccine against Clostridium difficile. Frontiers in Immunology, 13, 887061.

17) Bhardwaj, T., Haque, S., & Somvanshi, P. (2019). Comparative assessment of the therapeutic drug targets of C. botulinum ATCC 3502 and C. difficile str. 630 using in silico subtractive proteomics approach. Journal of Cellular Biochemistry, 120(9), 16160-16184.

18) Kesheri, M., Kanchan, S., Chowdhury, S., & Sinha, R. P. (2015). Secondary and tertiary structure prediction of proteins: a bioinformatic approach. Complex System Modelling and ControlThrough Intelligent Soft Computations, 541- 569.

19) Motamedi, H., Ari, M. M., Shahlaei, M., Moradi, S., Farhadikia, P., Alvandi, A., & Abiri, R. (2023). Designing multi-epitope vaccine against important colorectal cancer (CRC) associated pathogens based on immunoinformatics approach. BMC bioinformatics, 24(1), 65.

20) Irajie, C., Mohkam, M., Vakili, B., & Nezafat, N. (2021). Computational Elucidation of Phylogenetic, Functional and Structural Features of Methioninase from Pseudomonas, Escherichia, Clostridium and Citrobacter Strains. Recent Patents on Biotechnology, 15(4), 286-301.

21) Devarakonda, Y., Rajratna, A. D., Ray, A., & Syal, K. (2024). Novel edible multi-epitope vaccine construct against Enterococcus faecalis. The Nucleus, 1- 19.

22) Roja, B., Saranya, S., Prathiviraj, R., & Chellapandi, P. (2024). Functional prediction and assignment of Clostridium botulinum type A1 operome: A quest for prioritizing drug targets. Medicine in Omics, 12, 100040.

23) Kumar, P., Mohanan, A. G., & AK, A. K. (2020). Homology Modelling, Phylogenetic Analysis, and Molecular Docking of Glutamine Aminotransferase, gatD from Clostridium botulinum.

24) Davies, A. H., McGlashan, J., Posner, M. G., Roberts, A. K., Shone, C. C., & Acharya, K. R. (2016). Functional significance of active site residues in the enzymatic component of the Clostridium difficile binary toxin. Biochemistry and Biophysics Reports, 8, 55-61.

25) Azad, I., Khan, T., Ahmad, N., Khan, A. R., & Akhter, Y. (2023). Updates on drug designing approach through computational strategies: a review. Future Science OA, 9(5), FSO862.

26) Shang, D., Han, X., Du, W., Kou, Z., & Jiang, F. (2021). Trp-containing antibacterial peptides impair quorum sensing and biofilm development in multidrug-resistant Pseudomonas aeruginosa and exhibit synergistic effects with antibiotics. Frontiers in microbiology, 12, 611009.

27) Pathirana, R. D., et al. (2012). The tetratricopeptide repeat protein, YfiH, is involved in the regulation of flagellar gene expression in Escherichia coli. Journal of Bacteriology, 194(10), 2511-2522.

28) Christersson, L. A., et al. (2016). Tetratricopeptide repeat proteins in bacterial pathogenesis. Journal of Bacteriology, 198(10), 1431-1441.

29) Futterer, K., et al. (2009). The structure of the TPR domain of protein phosphatase 5. Journal of Molecular Biology, 386(2), 431-443.

30) Mueller, C. A., et al. (2003). The tetratricopeptide repeat domain of the Salmonella typhimurium InvH protein. Journal of Bacteriology, 185(10), 2943- 2952.