Acta Scientific Microbiology (ISSN: 2581-3226)

Volume 9 Issue 1 January 2026

Research Article

Unravelling the Functional Potential of a Hypothetical Protein from *Elizabethkingia meningoseptica*: An *In silico* Analysis

Hitesh Thakare[1]*, Rakesh Thakare[2], Yajuved N Selokar[1] and Dayanand Gogle[1]

[1]*Department of Molecular Biology and Genetic Engineering, Rashtrasant Tukadoji Maharaj Nagpur University, Nagpur, Maharashtra, India*

[2]*Department of Microbiology, Yashwantrao Chavhan College Lakhandur, Bhandara, Maharashtra, India*

***Corresponding Author:** Hitesh Thakare, Department of Molecular Biology and Genetic Engineering, Rashtrasant Tukadoji Maharaj Nagpur University, Nagpur, Maharashtra, India.

DOI: 10.31080/ASMI.2026.09.1569

Received: November 25, 2025

Published: December 27, 2025

Abstract

Globally, nosocomial infections are primarily caused by the gram-negative bacillus bacteria *Elizabethkingia meningoseptica*. The elucidation of *Elizabethkingia meningoseptica's* genome sequence has improved our understanding of the pathogenicity and lifestyle of this clinically significant organism. The study sought to provide an overview of a hypothetical protein that may be crucial to *Elizabethkingia meningoseptica* pathogenicity and virulence, including information on its expected structure, likely biological roles, and significance in this particular environment. *Elizabethkingia meningoseptica's* putative protein, which contains 179 amino acid residues, was selected for study and modelling using a variety of bioinformatics tools and databases in this work. Investigations of the primary and secondary structures illustrate that AQX04507.1 is a stable hydrophilic protein containing a sizable number of α-helices. According to subcellular localization predictions made by PSORTb, SOSUI server, and CELLO, the protein is cytoplasmic. Functional genomics studies using NCBI-CDD, Pfam, and InterProScan suggested that the putative protein might contain the putative TPR-repeat lipoprotein domain of the PEP-CTERM system. Using the template structure of Parabacteroides merdae ATCC 43184, which is the crystal structure of a tetratricopeptide repeat (PARMER_03812) (PDB ID: 4R7S), an X-ray diffraction model with 99.81% sequence identity with the hypothetical protein, the HHpred server was used to create its 3D structure in the homology modeling method. Following energy minimization, the created protein model was found to be dependable and of acceptable quality based on a number of quality assessments and validation variables. The hypothetical protein AQX04507.1 of *Elizabethkingia meningoseptica* has been thoroughly characterized and functionally annotated in the current study. Additional experimental validation is necessary to ascertain the true function of AQX04507.1 and to validate its potential as a therapeutic target.

Keywords: *Elizabethkingia meningoseptica*; Hypothetical Protein; Homology Modelling; Tetratricopeptide Repeat (TPR) Lipoprotein; UCSF Chimera

Introduction

Elizabethkingia meningoseptica is a gram-negative, saprophytic, nonfermentative, narrow, slightly curved, and indole-positive bacteria. Previously, it was referred to as *Flavobacterium meningosepticum* and *Chryseobacterium meningosepticum*. In 1959, an American bacteriologist named Elizabeth O. King discovered this species, and it was subsequently identified as the culprit behind neonatal septicaemia [1,2]. Even as he turned into analysing unclassified bacteria related to meningitis in infants, he named the organism that he recovered *Flavobacterium* ("the yellow bacillus") *meningosepticum* ("associated with meningitis and sepsis"). In 1994, It was reclassified as *Chryseobacterium meningosepticum* and placed inside the *Chryseobacterium* genus. *Elizabethkingia* is a common microbe that is widely distributed throughout the environment [3-6].

Chryseobacterium meningosepticum causes ailment predominantly in premature newborns and infants. Meningitis and bacteria are maximum not unusual medical presentation. However, C. meningosepticum remains an extraordinary pathogen in cases of bacterial meningitis in adults and kids [2,7,8]. *Elizabethkingia meningoseptica* is mainly disbursed in soil, plants, water, frogs, foodstuffs, and fishes. it's far a health cen-

ter-obtained pathogen pronounced in water assets, disinfectants, and medical instruments in hospitals and can be extracted from the sputum of patients with cystic fibrosis [9]. *E. meningoseptica,* a newly identified hospital-acquired pathogen, affects patients with compromised or healthy immune systems and exhibits resistance to multiple drugs. A case series demonstrated a significant rise in the annual occurrence of *E. meningoseptica* bacteremia between 2002 and 2006 (from 6.8–13.1 to 26.6–39.9 per 100,000 admissions; P = 0.006). Based on epidemiological reports, Taiwan has emerged as the central location for *E. meningoseptica* infections in the past ten years [4,9-17].

The strain of *E. meningoseptica* showed a high level of resistance to 16 out of 13 antibiotics, which made it a strong multi-drug resistant pathogen [18,19]. The management of *E. meningoseptica* infections should rely on the susceptibility test results, specifically the minimum inhibitory concentration (MIC). Nevertheless, different strains of *E. meningoseptica* have been found to exhibit resistance to multiple medications, particularly lactams. This resistance is linked to the presence of various types of lactamases, including class A extended-spectrum lactamases and class B metallo-lactamases (MBLs). The Virulence Factor Database (VF) (Table 2) has identified 766 prevalent virulence factors for *E. meningoseptica*. These factors encompass a wide range of genes that contribute to the synthesis of lipo-oligosaccharides, capsule polysaccharide, catalases, proteases, peroxidase, a two-component regulatory system, superoxide dismutase, heat shock protein, and various other elements associated with virulence [20–22].

Prematurity is a significant contributing factor to the risk of *E. meningosepticum* infection. This condition has been linked to various serious infections, including Meningitis, endocarditis, cellulitis, dialysis-associated peritonitis, septic arthritis, ocular infections, sinusitis, wound infection, epididymitis, and abdominal infections. These infections primarily affect severely immunocompromised patients, such as those with end-stage hepatic and renal disease, extensive burns, acquired immune deficiency syndrome, as well as individuals with community-acquired necrotizing *fasciitis, pneumonia, and bacteremia* [15,23,24].

Patients diagnosed with *E. meningoseptica* bacteremia face a bleak outlook, and the situation is exacerbated by the administration of inappropriate antibiotics. It is worth noting that all these patients were already being treated with colistin for infection caused by multidrug resistant organisms before contracting *E. meningoseptica*. Furthermore, the organism displayed resistance to the majority of antimicrobial agents that were tested, and even developed resistance to additional treatments during the course of treatment. The establishment of Clinical and Laboratory Standards Institute (CLSI) breakpoints for this particular organism is yet to be determined, which poses a significant challenge for both microbiologists and clinicians in selecting the appropriate antibiotic. Additionally, the mortality observed in the aforementioned patients may also be influenced by comorbid conditions and underlying diseases [25].

Antibiotics are used to treat bacterial infections. Microorganisms such as bacteria, fungi, or viruses develop antibiotic resistance over time as a defence mechanism to adapt to their surroundings and shield themselves from antibiotic therapy [26]. It is a severe threat to global health that is responsible for frightening rates of sickness and mortality in both humans and animals [27]. The emergence of antibiotic resistance is a direct result of bacterial evolution, as bacteria adapt to the ever-changing environmental conditions. Therefore, it is crucial to develop new strategies to effectively combat antibiotic-resistant pathogens. Biological preparations have the potential to enhance the host's immune response against any type of infection [28]. The only method to more effectively control bacterial infections resistant to antibiotics is to have a healthy supply of vaccines. To successfully produce a vaccine, potential candidates must meet certain criteria, such as being highly antigenic, conserved, and non-homologous to the host's typical fora or host [29].

The *E. meningoseptica* genome sequence comprises 4,038,467 base pairs (bp) and has a G+C content of 36.37%. Numerous genes responsible for encoding proteins linked to virulence, disease, and defense were detected within the genome. Additionally, we uncovered protein-coding genes associated with capsule formation, potentially boosting the bacteria's capacity to induce infection, evade phagocytosis, and adhere to medical device surfaces. An initial analysis of the genome revealed the presence of genes in *E. meningoseptica* that provide resistance to β-lactams. Additionally, this bacterium possesses hemolysin genes and a collection of genes responsible for heme uptake and utilization, indicating its capability to potentially induce bloodstream infections [30].

The study utilized the hypothetical protein (AQX04507.1) from *E. meningoseptica* due to its unknown structural characteristics, (Figure 5) yet known core amino acid sequence. The objective was to analyze the physiochemical and secondary structural features of the potential *E. meningoseptica* protein (AQX04507.1), develop its initial three-dimensional (3D) model via homology modelling, and carry out functional and comparative genomics investigations us-

ing Basic Local Alignment Search Tool for proteins (BLASTp) and multiple sequence alignment (MSA) analysis. This research aims to improve our comprehension of the functional roles played by *Elizabethkingia* members, thus providing valuable insights into potential therapeutic targets.

Materials and Methods

Procedures For the filtration and selection of a specific hypothetical protein

The proteomic information for *Elizabethkingia meningoseptica* was obtained from the Kyoto Encyclopedia of Genes and Genomes (KEGG) (https://www.genome.jp/kegg/) [57] database, with around 296 proteins selected for further *in silico* analysis after excluding those with amino acid sequences shorter than 50 residues due to compromised folding characteristics. All 296 Hypothetical proteins are longer than 50 residue amino acid sequences were analysed with ProtParam tool for protein physicochemical characteristics. Out of these, approximately 70 HP were discarded from further analysis as they showed unstable characteristics based on ProtParam tool results. Out of the remaining stable (226) HP, HP that consistently showed the same result with all four tools were selected for Protein Domain and Motif Prediction. Subcellular Location Analysis The remaining stable HP were subjected to subcellular location analysis with CELLO V2.5, PSoRTb, SOSUI and PSLpred. Subcellular localization and subsequent screening steps resulted in the rejection of 137 HP. In the next step, homology modelling was used to predict the 3D structure of the HP. HP with less than 80 % sequence similarity to a published protein structure is discarded for further analysis. We were able to narrow the HP number down to 29. The predicted models were tested for protein quality and accuracy. The number of predicted HPs was successfully reduced down as several models failed the quality assessment and were excluded from further selection. From this narrow selection, clinical relevance and function were carefully considered for all predicted HPs and the best performing HP (with 179 amino acids) was selected for our manuscript as the representative. The general workflow for this screening is shown in Figure 1.

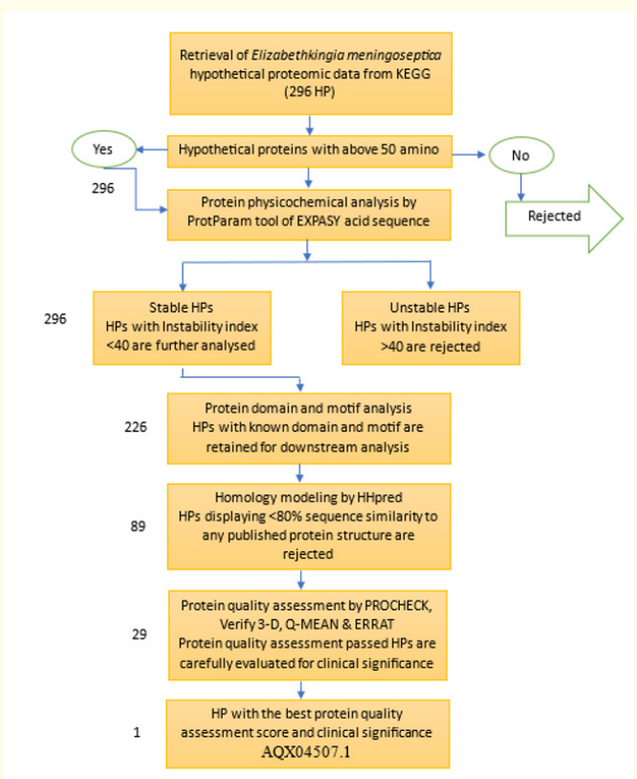

Figure 1: Workflow of the filtration and selection process of the hypothetical protein (AQX04507.1). Number of HP at each step of the filtration process are indicated on the left side.

Physicochemical properties analysis

The physicochemical properties of HP were characterized with the help of the ExPasy ProtParam (https://web.expasy.org/protparam/) tool [31,32]. Parameters such as Molecular weight, Aliphatic Index (AI), Extinction Coefficient, Amino Acid Composition, GRAVY (Grand Average of Hydropathy), Isoelectric Point (PI), and Estimated Half-Life were analysed.

Prediction of protein subcellular localisation

The putative subcellular localization of the Hypothetical proteins (AQX04507.1) is as follows: CELLO vs.2.5 predicts subcellular localization based on two-Level support vector prediction (SVM). The subcellular localization predicted in CELLO is correlated with the result in PSoRTb (https://www.psort.org/psortb/) [33], SOSUI (https://harrier.nagahama-ibio.ac.jp/sosui/mobile/) [34], and PSLpred (https://webs.iiitd.edu.in/raghava/pslpred/submit.html) [35]. SOSUI distinguishes between soluble and transmembrane proteins based on the average hydrophobic activity of protein. PSORTb and PSLpred predict subcellular localization for prokaryotic protein based on various parameters such as evolutionary information PSI-BLAST, amino acid, dipeptide, and physicochemical properties.

Identification of protein domain and motif

NCBI Conserved Domain Search (NCBI CD-Search) (https://structure.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi) [36], Protein families database (Pfam34.0) (http://pfam.xfam.org/) [37], and InterProScan5 (http://www.ebi.ac.uk/Tools/services/web/toolform.ebi?tool=iprscan5 &sequence=uniprot:KPYMHUMAN) [32,38] had been used for domain analysis of AQX04507.1. We applied the Conserved Domain Database (CDD) through Reverse Position-Specific BLAST (RPS-BLAST) and the InterProscan tool for our analyses. Pfam, a protein family database, hired hidden Markov models (HMMs) to generate annotations and multiple sequence alignments. To perceive the protein sequence motif, we hired the MOTIF search (https://www.genome.jp/tools/motif/) online InterProscan tool. Pfam is a protein family database that makes use of hidden Markov models (HMMs) with the intention to generate annotations and multiple sequence alignments. To decide the protein sequence motif, MOTIF Search (https://www.genome.jp/tools/motif/) online tool became used [39].

Protein family and phylogenetic tree analysis

In order to recognise the homologs of the HP (AQX04507.1), a protein-BLAST (BLASTp) (https://blast.ncbi.nlm.nih.gov/Blast.cgi?PAGE=Proteins) [40] from NCBI (National Centre for Biotechnology Information) towards the non-redundant database with default parameters changed into performed. This method is primarily based totally at the Local alignment of protein series to discover similar proteins. Multiple sequence alignment was carried out using CLC Sequence Viewer version 8 (https://clc-sequence-viewer.software.informer.com/8.0/), which was also used to create a phylogenetic tree for a particular subset of sequence [32].

Secondary structure prediction

2D Structure of the AQX04507.1 protein was determined using SOPMA (self-optimized prediction method with alignment) (https://npsa-prabi.ibcp.fr/cgibin/npsa_automat.pl?page=/NPSA/npsa_sopma.html) [41] and PSI-PRED (position-specific iterated – BLAST) (http://bioinf.cs. ucl.ac.uk/psipred/) [42]. The PSI-PRED result and the SOPMA analysis result were correlated.

Homology modelling

The 3D structure of AQX04507.1 was ascertained using the HHpred server (https://toolkit.tuebingen.mpg.de/tools/hhpred) [43], with the performance of this determination being based on the pairwise comparison profile of hidden Markov models (HMMs). A variety of databases, including the PDB, SCOP, Pfam, SMART, COGs, and CDD, can be searched using the HHpred server. Aspects of the target-template alignment have been used to project the quality of each detected template. For homology modeling, we selected the template protein of a Parabacteroides merdae ATCC 43184 which is crystal structure of a tetratricopeptide repeat protein (PARMER_03812) (PDB ID: 4R7S) with 99.81% sequence identity to our hypothetical protein. The 3D model structure was visualized using UCSF Chimera 1.17 [44].

Quality assessment

Several programs from the ExPASy server of SWISSMODEL Workspace, including PROCHECK (https://www.ebi.ac.uk/thornton-srv/software/PROCHECK/) [45], Verify 3D (https://servicesn.mbi.ucla.edu/Verify3D/) [46], ERRAT [47], and Qualitative Model Energy Analysis (QMEAN) (https://swissmodel.expasy.org/qmean/) [48], were used to evaluate the structural aspects of the protein model. In addition to measuring torsion angles, surface areas, bond angles, and atomic distances, PROCHECK also generates a Ramachandran plot [46]. The model evaluation was checked for errors in localized regions, stereo chemical characteristics like bond lengths and angles, and the overall fold/structure accuracy. Verify 3D uses structural class assignment (alpha, beta, loop, polar, nonpolar, etc.) based on the location and environment of an atomic model (3D) to determine whether it is compatible with its own

amino acid sequence (1D) [46]. The composite scoring function known as QMEAN, or Qualitative Model Energy Analysis, describes the key geometrical features of protein structures. "Evaluation of Protein Structure by Ramachandran Plot Assessment" is what ERRAT stands for. One measure for evaluating the precision and Caliber of protein models is the ERRAT score. The ERRAT score assesses how well the model fits known protein structures by calculating the statistical significance of the difference between predicted and expected atomic interactions [32,47]. These analyses offer insightful information about the precision and quality of the protein models, guaranteeing their dependability for additional study and interpretation.

Energy minimization of the model structure

Through the use of the YASARA energy minimization server the 3D structure of the hypothetical protein AQX04507.1 was improved [49]. The model structure was uploaded in order to carry out the protein energy minimization of the three-dimensional protein PDB file. By offering a more accurate and stable three-dimensional representation of the target protein, the server reduces the amount of energy needed (AQX04507.1).

Active site analysis

Utilizing the Computed Atlas of Surface Topography of Proteins (CASTp) (http://sts.bioe.uic.edu/castp/) server, the ligand binding sites of the putative protein were identified. AQX04507.1. CASTp provides a thorough, quantitative, and detailed analysis of a protein's topographical features. The regions and important residues of proteins that interact with ligands, as well as the interior site of the three-dimensional structure, are predicted by CASTp to contain active pockets. It has consequently developed into a crucial tool for predicting protein regions and important residues that interact with ligands [32,50]. PyMOL software was also used to display the CASTp result [51].

Results

Analysis of physicochemical properties and subcellular localization

ExPASy's ProtParam server was used to analyze the theoretical physiochemical characteristics of the hypothetical protein AQX04507.1 (Table 1). It was estimated that the protein would have a molecular weight of 20810.39 Daltons, an isoelectric point (PI) of 9.25, and 179 amino acids. It is commonly known that proteins are classified as stable if their instability index is less than 40 and unstable if it is greater than 40 [52]. The instability index for the hypothetical protein under analysis, AQX04507.1, was found to be 35.68, meaning that it is within the stable range. A protein that is both hydrophilic and soluble is indicated by a negative grand average of hydropathicity (GRAVY) index of -0.491. Leucine (23) and Alanine (16) were determined to be the next most abundant amino acid residues, after Lysine (24). Histidine was determined to be the lowest (2). There were 32 positively charged residues (arginine+lysine) and 23 negatively charged residues (aspartic acid+glutamic acid) in the sequence. 2959 atoms make up the atomic makeup, and the protein's molecular formula is $C_{931}H_{1502}N_{248}O_{266}S_{12}$. The location of a protein within a cell has a major impact on its function. Since distinct cellular sites correspond to distinct activities, it would be advantageous to predict the subcellular distribution of unknown proteins. The research of illness mechanisms and the creation of novel medications may benefit from this knowledge [30,53]. It was expected that the subcellular location of our query protein (AQX04507.1) would be cytoplasmic. AQX04507.1's subcellular position was examined by CELLO, and PSORTb v3.2.0, SOSUIGramN, and PSLpred server verified the location.

Table 1: Physiochemical properties of the AQX04507.1 using ProtParam.

Descriptions	Value
Number of amino acids	179
Molecular weight	20810.39 KDa
Theoretical pI	9.25
Total number of negatively charged residues	23
Total number of positively charged residues	32
Ext. coefcient	18130 $M^{-1}cm^{-1}$
Instability index	35.68
Aliphatic index	88.94
Grand average of hydropathicity (GRAVY)	-0.491

Secondary structure prediction of AQX04507.1.

The SOPMA and PSI-PRED servers were utilized to look into the AQX04507.1's secondary structure. The protein's alpha helix, beta turn, extended strand, and random coil proportions were found to be 68.16%, 3.91%, 4.47%, and 23.46%, respectively, according to the SOPMA secondary prediction server analysis. Comparable outcomes were also noted in the PSI-PRED tool (Figure 2).

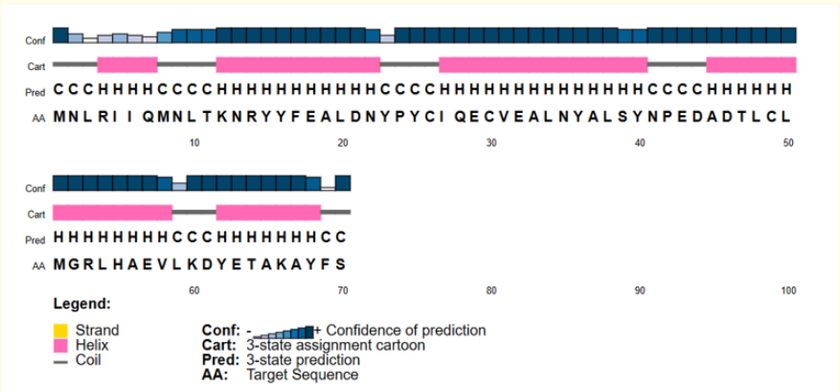

Figure 2: Protein secondary structure prediction of the (AQX04507.1) use the server PSI-PRED. There are four distinct portions in this graphical illustration. Bars of various heights make up the initial portion. The relationship between the bar's height and confidence score is linear. In the second part, the colour pink, yellow, and Gray stand for the alpha helix, beta sheets, or strands, and coils, respectively. A certain beta sheet and a specific alpha helix are joined by the coil. The secondary structure of a protein is represented alphabetically in the third section by the letters E, H, and C, which stand for beta sheets, alpha helices, and coils, respectively. In the last part, the amino acid arrangement is listed alphabetically.

Prediction of protein family by domain and motif analysis

The annotation tools NCBI-CD Search, Pfam, and InterProScan were utilized to determine the conserved domains and possible function of AQX04507.1. The query protein is predicted to belong to the tetratricopeptide repeat (TPR) lipo super family by the specific hit examined by the conserved domain (CD) search tool (putative PEP-CTERM system TPR-repeat lipoprotein). The tetratricopeptide repeat (TPR) protein in this family has an E-value of 7.41e-03 for amino acid residues 1–157 in our protein sequence. The results of the CD search analysis were found to be similar to those of the InterProscan and Pfam domain searching tools. With an E-value of 6.4e-13, InterProscan spans amino acid residues 1-157. The tetratricopeptide repeat (TPR) superfamily is predicted by the Pfam tool to span 1–157 amino acid residues, with an E-value of 6.4e–13. With an E-value of 2e-05, the MOTIF server predicted the presence of putative PEP-CTERM system TPR-repeat lipoprotein at positions 1 - 154 amino acid residues.

Comparative genomics analysis of AQX04507.1 using multiple sequence alignment and phylogeny

Tetratricopeptide repeat (TPR) lipo super family protein homology (up to 100% sequence similarity) with other known types from different *Elizabethkingia meningoseptica* was found using the BLASTp search against the non-redundant database (Table 2). For multiple sequence alignment (MSA), a total of ten chosen protein sequences and the target sequence were obtained from BLASTp analysis. To view the conserved and dissimilar residues among the homologs, MSA was performed using the BLASTp (Figure 3). A phylogenic tree was constructed using the same data (Figure 4). The target protein appears to share an ancestor with the WP260259498.1 and WP026149195.1, as well as two other proteins from MDE5438179.1 and OPC19382.1. The line segment with the number (0.022) indicates the amount of genetic change, and the scale bar estimates sequence divergence.

Table 2: Identification of homologs of AQX04507.1 through protein BLASTp search analysis.

Description	Scientific Name	Max Score	Total Score	Querry cover	E- Value	Per. Ident	Acc. Len	Accession
BBD33_04245 [*Elizabethkingia meningoseptica*]	*Elizabethkingia meningoseptica*	362	362	100%	7e-126	100.00	179	AQX04507.1
hypothetical protein BES09_10680 [*Elizabethkingia meningoseptica*]	*Elizabethkingia meningoseptica*	357	357	100%	9e-124	98.32	179	ODM53283.1
hypothetical protein [*Elizabethkingia meningoseptica*]	*Elizabethkingia meningoseptica*	350	350	100%	3e-121	96.09	179	EJK5328553.1
hypothetical protein [*Elizabethkingia meningoseptica*]	*Elizabethkingia meningoseptica*	348	348	96%	2e-120	100.00	172	WP_026149195.1
hypothetical protein [*Elizabethkingia meningoseptica*]	*Elizabethkingia meningoseptica*	342	342	96%	5e-118	98.26	172	WP_260259498.1
hypothetical protein [*Elizabethkingia meningoseptica*]	*Elizabethkingia meningoseptica*	340	340	96%	1e-117	97.09	172	WP_260257295.1
hypothetical protein [*Elizabethkingia meningoseptica*]	*Elizabethkingia meningoseptica*	336	336	96%	1e-115	95.93	172	WP_271151726.1
hypothetical protein [*Elizabethkingia meningoseptica*]	*Elizabethkingia meningoseptica*	313	313	96%	1e-106	95.35	172	WP_336664706.1
hypothetical protein BAX95_11915 [*Elizabethkingia meningoseptica*]	*Elizabethkingia meningoseptica*	311	311	87%	7e-106	96.18	179	OPC19382.1
hypothetical protein [*Elizabethkingia meningoseptica*]	*Elizabethkingia meningoseptica*	308	308	87%	1e-104	95.54	179	MDE5438179.1

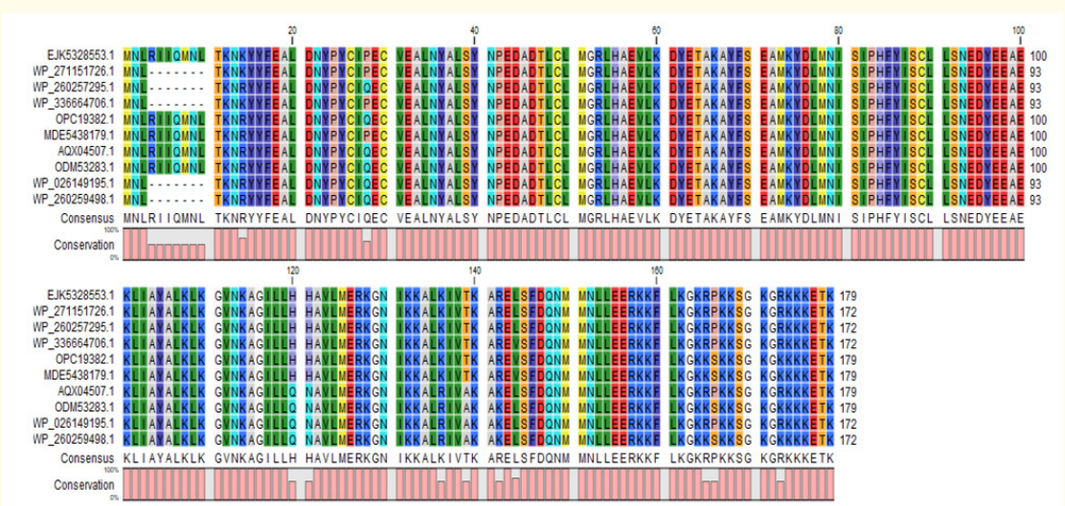

Figure 3: MSA analysis among the different types of hypothetical proteins with the AQX04507.1. MSA indicates multiple sequence alignment generated by CLS Sequence viewer version 8.

Threedimensional structure determination and model quality assessment

The query sequence was uploaded to the HHpred server in order to predict the structure and detect protein homology [43]. The template structure of Parabacteroides merdae ATCC 43184 which is crystal structure of a tetratricopeptide repeat (PARMER_03812) (PDB ID: 4R7S) protein was used to determine the 3D structure of AQX04507.1. This structure showed 98.81% identity with AQX04507.1 in the HHPred server. USCF Chimera 1.16 viewed the 3D model, which is displayed in (Figure 5). We employed QMEAN, ERRAT, PROCHECK, and Verify 3D to assess the quality of our 3D model. The program PROCHECK uses various evaluation metrics, such as surface area, bond angle, torsion angle, and atomic distance analysis, as well as the creation of a Ramachandran plot, to validate

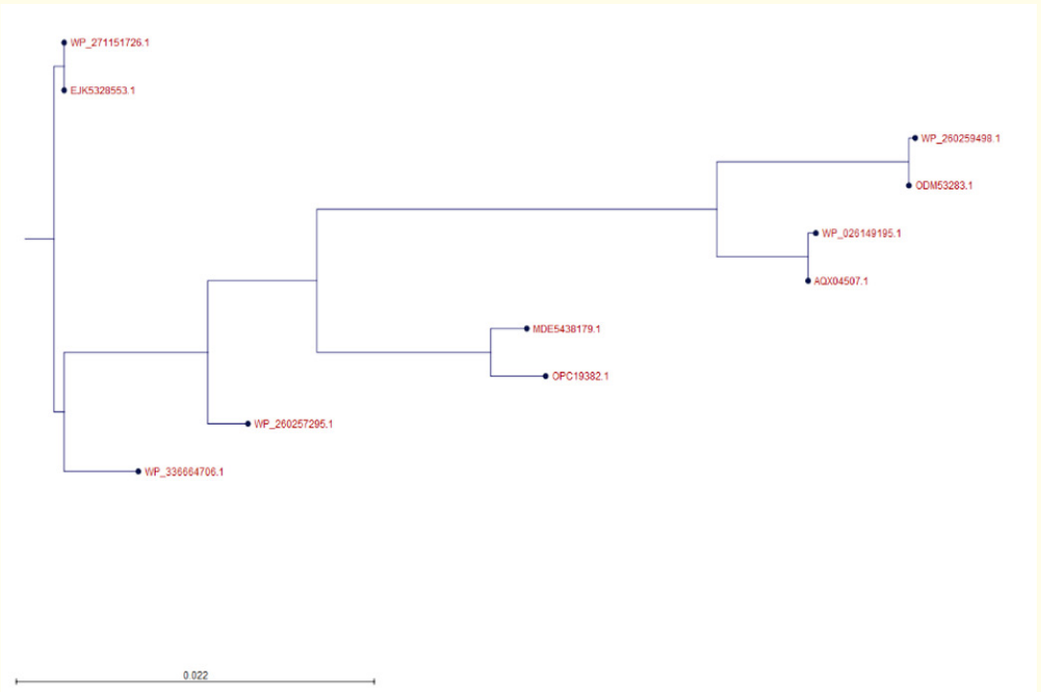

Figure 4: Phylogenetic tree illustrating the relationship of AQX04507.1 with closely related proteins generated by using CLS Sequence viewer version 8.

protein models. When evaluating the accuracy and structural integrity of protein models, these computations are essential [45]. The model was deemed dependable and of high quality based on the PROCHECK results, which showed that the most favoured region in the "Ramachandran plot" contained 91.6% of amino acid residues, with 8.4% and 0.0% of residues in additional allowed and generously allowed regions, respectively (Table 3 and Figure 6A). By analyzing the statistics of non-bonded interactions between different atom types based on characteristic atomic interactions, ERRAT was used to evaluate the model's reliability. With an overall quality factor of 99.5816, the template was determined to have a good, high-resolution structure. 97.17% of residues had an averaged 3D (atomic model)–1D (amino acid) score ≥0.1, which indicates that these structures were excellent and compatible, according to the Verify 3D tool. The QMEAN tool classified the model, with a QMEAN value of 0.87, into the dark grey zone. Since the QMEAN score threshold value, which goes from 0 to 1, is within the acceptable range, this score is regarded as good (Figure 6B).

Table 3: Ramachandran plot statistics of the hypothetical protein (AQX04507.1).

Statistics	Number of AA residues	Percentage (%)
Residues in the most favored regions (A, B, L)	217	91.6%
Residues in the additional allowed regions (a, b, l, p)	20	8.4%
Residues in the generously allowed regions (~ a, ~ b, ~ l, ~ p)	0	0.0%
Residues in disallowed regions	0	0.0%
Number of non-glycine and non-proline residues	237	100.0%
Number of end-residues (excl. Gly and Pro)	1	
Number of glycine residues (shown as triangles)	13	
Number of proline residues	5	
Total number of residues	256	

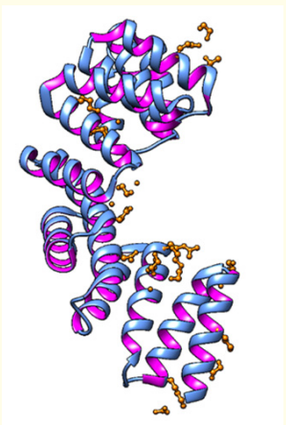

Figure 5: Predicted three-dimensional structure of the AQX04507.1 (visualized by UCSF Chimera 1.17).

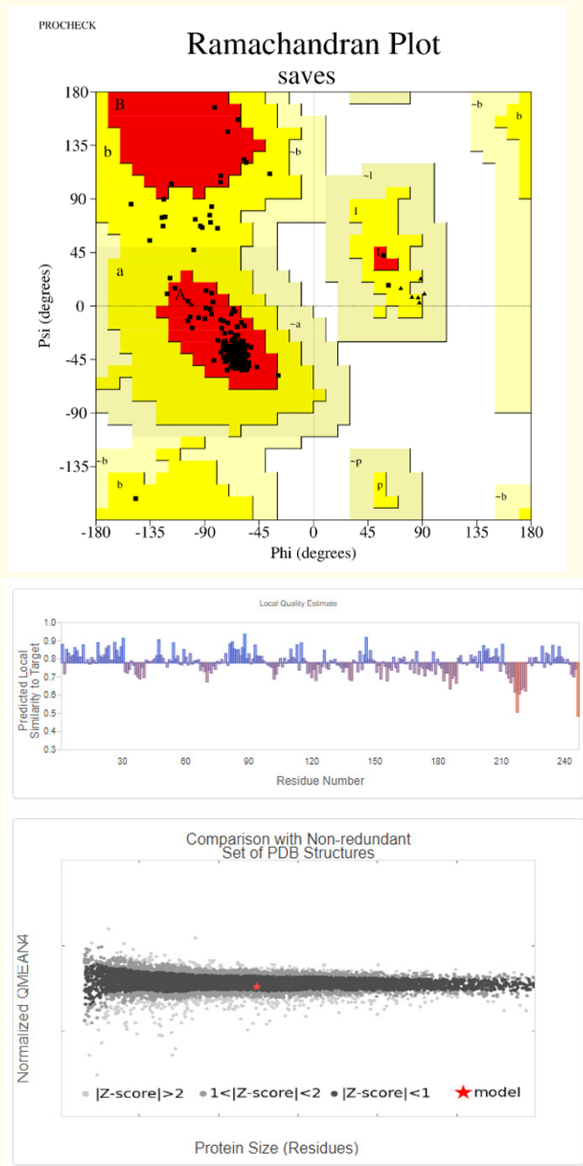

Figure 6: Evaluate the quality of the model. a Ramachandran plot of the PROCHECK server-validated model structure. Here, the most favoured areas (A, B, and L) were covered by 91.6% amino acid residues. B graphical depiction of the model structure's QMEAN outcome. In this case, the expected model's Z score was 0.87, indicating a strong match between the model's structure and an experimental structure of a comparable size.

Active site determination

A crucial step in the development of a medication or inhibitor is the identification and characterization of active site residues. The model structure's active site evaluation and the identification of the amino acid residues in the active site were conducted using the CASTp server. In one of the largest pockets, the top active sites of the model protein were identified by utilizing the area of 368.407 and the volume of 683.233 amino acids. The model protein's active residues are displayed in (Table 4 and Figure 7) based on CASTp prediction.

Table 4: CASTp analysis result: Active site of amino acid residues. Here, A.A, amino acid; SeqID, position of AA in protein sequence.

A.A	SeqID	A.A	SeqID	A.A	SeqID	A.A	SeqID
THR	104	ILE	138	ASP	203	VAL	264
GLU	105	GLN	139	THR	228		
LEU	106	LEU	165	ALA	229		
GLY	107	ILE	169	ASN	230		
THR	109	LYS	172	LEU	231		
TYR	116	PHE	184	LEU	234		
TYR	131	ILE	196	GLY	259		
CYS	132	LEU	197	TYR	260		
LEU	135	GLU	199	ASN	261		
LEU	136	GLY	200	GLU	263		

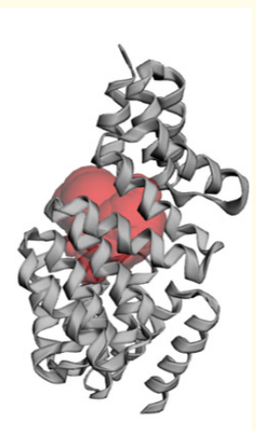

Figure 7: Determination of active site of YP_498675.1 using the CASTp server. The largest active site was found in the areas with 368.407 and the volume of 683.233 amino acids.

Energy minimization result

YASARA force field minimizer was used to minimize the energy of the three-dimensional structure of the predicted protein. After energy minimization, the energy was down from -137789.6 kJ/mol to -176410.7 kJ/mol. The final score changed from -0.05 to -1.28 after energy minimization, indicating a more stable structure.

Discussion

Elizabethkingia is not a typical component of human flora and is widely distributed in nature [1]. After 48 hours of incubation at 37° C in enriched media, the bacteria—which is Gram-negative, non-motile, non-fermenting, beta-hemolytic, oxidase and catalase positive—forms biofilms and grows into yellow-colored colonies [1,54].

Emerging as a significant opportunistic bacterium, *E. meningoseptica* mostly lives in nosocomial environments. It is unknown which treatment plan works best and what variables are linked to poor results [55].

The characterization of *Elizabethkingia meningoseptica* Hypothetical Protein AQX04507.1 can help with the development of efficient therapeutics, the formulation of disease control plans, and our understanding of the metabolic regulations of bacteria. In order to characterize the hypothetical protein AQX04507.1 of *Elizabethkingia meningoseptica* from structural and functional aspects, a variety of computational resources were used in this study. The protein's physiochemical properties analysis showed that it has a theoretical PI of 9.25, a molecular weight of 20810.39, a grand average of hydrophilicity (GRAVY) score of -0.491, and a sequence of 281 amino acids (Table 1). The HP extinction coefficient at 280 nm, which ranges from 18,130 to 17,880 M−1 cm−1, is computed using the Protparam tool. For the quantitative analysis of protein interactions, including those with ligands and other proteins, this coefficient is useful [49]. To analyze the hypothetical protein AQX04507.1, we used CELLO, a subcellular location prediction tool, in our study. The CELLO results showed that the protein primarily localizes in the cytoplasm, which is in complete agreement with the PSoRTb, SOSUI, and PSLpred findings. Additionally, the ProtParam GRAVY index (−0.491) supported this observation by indicating that the protein is hydrophilic. The CELLO prediction further supports our findings, since hydrophilic proteins are typically located in the cytoplasmic compartment in cellular environments. The prevalence of extended strand, beta turn, alpha helix, and random coil is shown by secondary structure analysis of the protein. The tetratricopeptide repeats (TPR) lipoprotein motif is predicted to be present in the hypothetical protein under analysis (AQX04507.1) in this work. This protein family occurs in strictly within a subset of Gram-negative bacterial species with the proposed PEP-CTERM/exosortase system, analogous to the LPXTG/sortase system common in Gram-positive bacteria. This protein occurs in a species if and only if a transmembrane histidine kinase (TIGR02916) and a DNA-binding response regulator (TIGR02915) also occur. The present of tetratricopeptide repeats (TPR) suggests protein-protein interaction, possibly for the regulation of PEP-CTERM protein expression, since many PEP-CTERM proteins in these genomes are preceded by a proposed DNA binding site for the response regulator.

To completely understand the unique function of the hypothetical protein AQX04507.1 and its connection to *Elizabethkingia meningoseptica* infections, more investigation is required. putative PEP-CTERM system TPR-repeat lipoprotein is necessary for the virulence and survival of pathogenic bacteria like Elizabethkingia meningoseptica. Tetratricopeptide repeat, Plant specific mitochondrial import receptor subunit TOM20, Anaphase-promoting complex, cyclosome, subunit 3, MIT (microtubule interacting and transport) domain of the hypothetical protein AQX04507.1. The precise role and importance of this protein are still being investigated, but its relationship to bacterial sorting and control adds to the intriguing complexity of microbial systems. AQX04507.1 demonstrated up to 100% sequence similarity with other hypothetical protein of Elizabethkingia meningoseptica, according to further analysis by protein BLASTp against the non-redundant database (Table 2).

The results of BLASTp and protein domain analyses unequivocally show that the hypothetical protein AQX04507.1 may play a significant functional role in *Elizabethkingia meningoseptica* cellular metabolism. Comprehending the interactions, functions, and localization of proteins requires an understanding of their three-dimensional structure. Homology modelling is the most commonly used technique for protein structure prediction. In this work, we used homology modelling to suggest the first three-dimensional structure of a *Elizabethkingia meningoseptica* protein known as AQX04507.1. Further investigation into drug design and protein interactions will be made possible by this predicted structure, which will provide insightful information about the structure and function of the protein [56].

The AQX04507.1's tertiary structure was created using the HHpred server, and evaluation tools such as Verify 3D, PROCHECK, ERRAT, and QMEAN were used to gauge the model's quality. The most favoured area in the Ramachandran plot was estimated to be covered by 91.6% of the amino acid residues in the model 3D structure, indicating that the model quality is valid (Figure 6A). The QMEAN4 server result (Figure 6B) revealed that the expected model's Z score was 0.87, indicating a high-quality model as well. Following the YASARA energy minimization procedure, the hypothetical protein AQX04507.1's 3D structure stabilized. An essential first step in the development of a medication or inhibitor is the CASTP server's prediction of active-site residues.

To find out more about the protein's enzymatic activity, binding characteristics, and potential roles in a variety of biological processes, these active site residues can be located and studied. A database server called CASTp is able to recognize and describe different regions on proteins. It is capable of identifying these regions' borders, estimating their sizes, and analysing their dimensions [49]. These areas include both internal voids within the protein structure and pockets on the protein's surface. The main active sites of the protein model were identified with high accuracy using CASTp; their sizes ranged from 368.407 (area) to 683.233 (volume). Active sites with a solvent-accessible (SA) surface area of 368.407 and a volume of 683.233 amino acids were identified as the largest pock-

et in the CASTp analysis. All things considered, the CASTp analysis contributes to our growing knowledge of the protein's structure–function relationship and makes it possible to investigate the specific molecular pathways at play.

Conclusion

We investigated the *in silico* structural and functional annotation of a novel hypothetical *Elizabethkingia meningoseptica* protein, designated AQX04507.1. Through a thorough computational analysis, we were able to learn important things about the properties and functions of this protein. Potential functional domains (putative PEP-CTERM system TPR-repeat lipoprotein) and motifs (Tetratricopeptide repeat) were uncovered by functional annotation, offering hints about their potential biological roles. Through a detailed analysis of AQX04507.1, we have added to our understanding of its possible role in essential cellular interactions and processes. By identifying potential targets for the development of a medication or vaccination against this pathogenic bacterium, our research contributes to our understanding of the genetic and proteomic profile of *Elizabethkingia meningoseptica*. Although this work is an important first step toward understanding the functional significance of AQX04507.1, additional experimental validation and functional characterization are necessary to validate the expected structural and functional characteristics. However, our thorough *in silico* analysis provides a strong basis for further investigation, providing insightful information about the possible functions of AQX04507.1 and its implications regarding the field of *Elizabethkingia meningoseptica* physiology and pathogenesis.

Bibliography

1. KING EO. "Studies on a group of previously unclassified bacteria associated with meningitis in infants". *American Journal of Clinical Pathology* 31.3 (1959): 241-247.

2. Idrees M., *et al*. "Core-Proteomics-Based Annotation of Antigenic Targets and Reverse-Vaccinology-Assisted Design of Ensemble Immunogen against the Emerging Nosocomial Infection-Causing Bacterium Elizabethkingia meningoseptica". *International Journal of Environmental Research and Public Health* 19.1 (2021).

3. Cone LA., *et al*. "Osteomyelitis due to Chryseobacterium (Flavobacterium) meningosepticum". *Antimicrobe and Infectious Disease Newsletter* (1998).

4. Lin PY., *et al*. "Clinical and microbiological analysis of bloodstream infections caused by Chryseobacterium meningosepticum in nonneonatal patients". *Journal of Clinical Microbiology* (2004).

5. Schreckenberger PC., *et al*. "Acinetobacter, Achromobacter, Chryseobacterium, Moraxella, and other nonfermentative gram-negative rods". *Manuel of Clinical Microbiology* (2003).

6. Ozkalay N., *et al*. "Community-acquired meningitis and sepsis caused by Chryseobacterium meningosepticum in a patient diagnosed with thalassemia major". *Journal of Clinical Microbiology* 44.8 (2006): 3037-3039.

7. Bloch KC., *et al*. "Chryseobacterium meningosepticum: an emerging pathogen among immunocompromised adults: report of 6 cases and literature review". *Medicine* (1997).

8. Sztajnbok J and EJT. "Community-acquired Chryseobacterium meningosepticum pneumonia and sepsis in a previously healthy child". *Journal of Infectious* (1998).

9. Hsu MS., *et al*. "Clinical features, antimicrobial susceptibilities, and outcomes of Elizabethkingia meningoseptica (Chryseobacterium meningosepticum) bacteremia at a medical center in Taiwan". *European Journal of Clinical Microbiology and Infectious Diseases* (2006).

10. Jean S., *et al*. "Elizabethkingia meningoseptica: An important emerging pathogen causing health care associated infections". *Journal of Hospital Infection* (2014).

11. Tuon FF., *et al*. "Chryseobacterium meningosepticum as a cause of cellulitis and sepsis in an immunocompetent patient". *Journal of Medical Microbiology* (2007).

12. Lin YT., *et al*. "Clinical and microbiological analysis of Elizabethkingia meningoseptica bacteremia in adult patients in Taiwan". *Journal of Infectious* (2009).

13. Chiu CH., *et al*. "Atypical Chryseobacterium meningosepticum and meningitis and sepsis in newborns and the immunocompromised, Taiwan". *Emergency Infectious Diseases* (2000).

14. Steinberg JP and BEM. "Other gram-negative and gram-variable bacilli". Mandell, Douglas, and Bennett's Principles and Practice of Infectious Diseases (2015).

15. Güngör S., *et al*. "Chryseobacterium meningosepticum outbreak in a neonatal ward". *Infection Control and Hospital Epidemiology* (2003).

16. Hoque S., *et al*. "Chryseobacterium (Flavobacterium) meningosepticum outbreak associated with colonization of water taps in a neonatal intensive care unit". *Journal of Infectious* (2001).

17. Hsueh PR., *et al*. "Susceptibilities of Chryseobacterium indologenes and Chryseobacterium meningosepticum to cefepime and cefpirome". *Journal of Clinical Microbiology* (1997).

18. Chen S., *et al*. "Comparative genomic analyses reveal diverse virulence factors and antimicrobial resistance mechanisms in clinical Elizabethkingia meningoseptica strains". *PLoS ONE* (2019).

19. Lin JN., *et al*. "Elizabethkingia infections in humans: From genomics to clinics". *Microorganisms* (2019).

20. Chen CH., *et al*. "Bacteremia caused by Elizabethkingia meningoseptica in a mechanically ventilated patient successfully treated with imipenem-cilastatin and ciprofloxacin". *The Revista do Instituto de Medicina Tropical de São Paulo* (2017).

21. Chen S., *et al*. "Comparative genomic analyses reveal diverse virulence factors and antimicrobial resistance mechanisms in clinical Elizabethkingia meningoseptica strains". *PLoS One* (2019).

22. Meza B., *et al*. "A novel design of a multi-antigenic, multistage and multi-epitope vaccine against Helicobacter pylori: An in-silico approach". *Infection, Genetics and Evolution* (2017).

23. Adachi A MTSTYATNIY. "Chryseobacterium meningosepticum septicemia in a recipient of allogeneic cord blood transplantation". *Scandinavian Journal of Infectious Diseases* (2004).

24. Kirby JT SHWTJRN. "Antimicrobial susceptibility and epidemiology of a worldwide collection of Chryseobacterium spp: Report from the SENTRY antimicrobial surveillance program". *Journal of Clinical Microbiology* (2001).

25. Ratnamani MS RR. "Elizabethkingia meningoseptica: Emerging nosocomial pathogen in bedside hemodialysis patients". *Indian Journal of Critical Care Medicine* (2013).

26. Hutchings M., *et al*. "Antibiotics: past, present and future". *Current Opinion on Microbiology* (2019).

27. Negahdaripour M., *et al*. "Structural vaccinology considerations for in silico designing of a multi-epitope vaccine". *Infection Genetics and Evolution* (2018).

28. Covián C., *et al*. "BCG-induced cross-protection and development of trained immunity: Implication for vaccine design". *Frontiers in Immunology* (2019).

29. Okafor CN RAMII. "Bacillus Calmette Guerin". Treasure Island (FL): StatPearls Publishing (2024).

30. Chen S., *et al*. "Genome Sequence of Elizabethkingia meningoseptica EM1, Isolated from a Patient with a Bloodstream Infection". *Genome Announcement* 4.5 (2016).

31. Wilkins MR., *et al*. "Protein Identification and Analysis Tools in the ExPASy Server". In: 2-D Proteome Analysis Protocols. New Jersey: Humana Press 531-552.

32. Chakma V., *et al*. "In silico analysis of a novel hypothetical protein (YP_498675.1) from Staphylococcus aureus unravels the protein of tryptophan synthase beta superfamily (Try-synth-beta_ II)". *Journal of Genetic Engineering and Biotechnology* 21.1 (2023): 135.

33. Gardy JL., *et al*. "PSORTb v.2.0: Expanded prediction of bacterial protein subcellular localization and insights gained from comparative proteome analysis". *Bioinformatics* 21.5 (2005): 617-23.

34. Hirokawa T., *et al*. "SOSUI: classification and secondary structure prediction system for membrane proteins". *Bioinformatics* 14.4 (1998): 378-379.

35. Bhasin M., *et al*. "PSLpred: prediction of subcellular localization of bacterial proteins". *Bioinformatics* 21.10 (2005): 2522-4.

36. Marchler-Bauer A., *et al*. "CDD: a conserved domain database for interactive domain family analysis". *Nucleic Acids Research* 35 (2007): D237-240.

37. Bateman A., *et al*. "@Pfam@The Pfam protein families database". *Nucleic Acids Research* 32 (2004): D138-141.

38. Quevillon E., *et al*. "InterProScan: protein domains identifier". *Nucleic Acids Research* 33 (2005): W116-120.

39. Smoly IY., *et al*. "MotifNet: a web-server for network motif analysis". *Bioinformatics* 33.12 (2017): 1907-1909.

40. Altschul SF., *et al*. "Basic local alignment search tool". *Journal of Molecular Biology* 215.3 (1990): 403-410.

41. Geourjon C and Deléage G. "SOPMA: significant improvements in protein secondary structure prediction by consensus prediction from multiple alignments". *Bioinformatics* 11.6 (1995): 681-684.

42. Buchan DWA and Jones DT. "The PSIPRED Protein Analysis Workbench: 20 years on". *Nucleic Acids Research* 47 (2019): W402-407.

43. Soding J., *et al*. "The HHpred interactive server for protein homology detection and structure prediction". *Nucleic Acids Research* 33 (2005): W244-248.

44. Huang CC., *et al*. "Enhancing UCSF Chimera through web services". *Nucleic Acids Research* 42.W1 (2014): W478-484.

45. Laskowski RA., *et al*. "PROCHECK: a program to check the stereochemical quality of protein structures". *Journal of Applied Crystallography* 26.2 (1993): 283-291.

46. Eisenberg D., *et al*. "VERIFY3D: Assessment of protein models with three-dimensional profiles". In (1997): 396-404.

47. Colovos C and Yeates TO. "Verification of protein structures: Patterns of nonbonded atomic interactions". *Protein Science* 2.9 (1993): 1511-1519.

48. Benkert P., *et al*. "Toward the estimation of the absolute quality of individual protein structure models". *Bioinformatics* 27.3 (2011): 343-350.

49. Land H and Humble MS. "YASARA: A Tool to Obtain Structural Guidance in Biocatalytic Investigations". In (2018): 43-67.

50. Tian W., *et al*. "CASTp 3.0: computed atlas of surface topography of proteins". *Nucleic Acids Research* 46.W1 (2018): W363-367.

51. Rigsby RE and Parker AB. "Using the PyMOL application to reinforce visual understanding of protein structure". *Biochemistry and Molecular Biology Education* 44.5 (2016): 433-437.

52. Naveed M., *et al*. "Structural and functional annotation of hypothetical proteins of human adenovirus: prioritizing the novel drug targets". *BMC Research Notes* 10.1 (2016): 706.

53. Okafor CN., *et al*. "Bacillus Calmette Guerin". (2024).

54. Dias M. "Case Series: Elizabethkingia Meningosepticum". *Journal of Clinical and Diagnostic Research* (2012).

55. Umair A and Nasir N. "Clinical features and outcomes of critically ill patients with Elizabethkingia meningoseptica: an emerging pathogen". *Acute and Critical Care* 36.3 (2021): 256-61.

56. Sobia Idrees., *et al*. "In silico Sequence Analysis, Homology Modeling and Function Annotation of Ocimum basilicum Hypothetical Protein G1CT28_OCIBA". *International Journal Bioautomation* 16.2 (2012): 111-118.

57. Kanehisa M., *et al*. "KEGG: biological systems database as a model of the real world". *Nucleic Acids Research* 53 (2025): D672-D677.