IOSR Journal Of Pharmacy And Biological Sciences (IOSR-JPBS)
e-ISSN:2278-3008, p-ISSN:2319-7676. Volume 16, Issue 1 Ser. IV (Jan. – Feb. 2021), PP 21-27
www.Iosrjournals.Org

Computational Annotation of a Hypothetical Protein from *Pseudomonas aeruginosa* SE5458 Strain Reveals a Bacterial Antitoxin Protein, Tsi6

Md. Fazley Rabbi[1]*, Aradhan Sarkar[1]
[1]*Department of Biotechnology and Genetic Engineering, Noakhali Science and Technology University, Noakhali-3814, Bangladesh.*
**Corresponding author*

Abstract
Background: *Multidrug resistant Pseudomonas aeruginosa is a gram-negative bacterium that causes opportunistic infections in human. The bacterial genome of SE5458 contains many hypothetical proteins yet to be annotated. Annotation of these proteins might reveal new insights about this pathogenic organism; hence a hypothetical protein WP_003111794.1 was selected in this study for comprehensive characterization by computational approach.*
Materials and Methods: *Several computational tools were used in this study for characterization of the target protein. Physicochemical properties were estimated using ProtParam tool. CELLO, LocTree3 and PSLpred server was used for subcellular localization prediction. Functional analysis was performed using NCBI-CDD, InterProScan5 and Pfam server. Secondary and tertiary structure (3D) of the protein was determined by SOPMA and MODELLER server respectively. Moreover, quality of the 3D structure was validated by several structure assessment tools. The active site of the protein was analyzed by CASTp server. Finally, protein-protein interaction network was built using STRING database.*
Results: *The 94 amino acid containing protein was found as a stable protein. The protein was predicted to be a Tsi6 family protein which neutralizes the toxic activities of integral membrane toxin Tse6 of P. aeruginosa. Tse6 inhibits growth of target bacterial cells by degrading essential NAD+ and NADP+; therefore, influences microbial niche formation. 3D structure of the protein was successfully determined which passed all the quality assessment tools. Active site analysis revealed several key interacting residues. Protein-protein interaction analysis by STRING database revealed several interacting partners. Interestingly, two of them were found to be part (like Tse6) of type VI secretion system further supporting our annotation.*
Conclusion: *The protein was found to be a Tsi6 family protein that neutralizes the toxic activities of Tse6. We encourage further research and experimentations to validate our findings.*

Keywords: *Computational annotation, hypothetical protein, Pseudomonas aeruginosa, antitoxin protein, Tsi6*

Date of Submission: 29-01-2021 Date of Acceptance: 12-02-2021

I. Introduction

Pseudomonas aeruginosa is a Gram-negative, rod-shaped, opportunistic bacterium that causes disease in different species ranging from plants to animals, including humans [1]. The bacterium is commonly found in soil, water, skin flora, medical equipment etc [2]. It generally affects the immunocompromised people typically infecting the airway, urinary tract, blood, burns, and wounds [3, 4, 5]. *P. aeruginosa* infections have become a critical issue responsible for a total of 51,000 healthcare infections in the USA per year [6, 7, 8]. It has been designated as an ESKAPE pathogen (*Enterococcus faecium, Staphylococcus aureus, Klebsiella pneumoniae, Acinetobacter baumannii, Pseudomonas aeruginosa,* and *Enterobacter* species), a group of pathogens with a high rate of antibiotic resistance that are responsible for the majority of nosocomial infections [9]. Due to widespread resistance to many common first-line antibiotics, people are left with limited treatment options for *P. aeruginosa* infections. Hence, extensive research endeavors for the discovery of novel antibiotics and drugs against the bacterium is essential. However, the bacterial genome contains many hypothetical proteins (HPs) whose functions are yet to be discovered. Hypothetical proteins (HPs) are those proteins of which primary sequences are known but there exists no experimental evidence about their function *in vivo* [10]. Exploring the structural and functional features of these HPs may lead us to identify markers and potential therapeutic targets [10]. Computational annotation of hypothetical proteins may help in determining three-dimensional (3D) structures which may reveal new domains and motifs, metabolic pathways and protein-protein interaction networks etc. [11, 12, 13].

Pseudomonas aeruginosa SE5458 strain was isolated from respiratory sample of human from China in 2011. The 7.15 Mb long genome contains 6,646 proteins, of which 1,034 proteins are hypothetical (as of 20 January, 2021). Various bioinformatics resources have previously been used to annotate the functions of hypothetical proteins in different pathogenic microorganisms [14, 15, 16, 17, 18]. A hypothetical protein WP_003111794.1 of *P. aeruginosa* SE5458 strain was chosen in this study for comprehensive structural and functional analysis using computational approach. The protein was selected for the study as its primary sequence is available but structural and functional information is yet to be elucidated.

II. Materials and Methods

Primary sequence retrieval: There are 5,662 genomes of *Pseudomonas aeruginosa* available in NCBI database (https://www.ncbi.nlm.nih.gov/) [19]. We browsed the NCBI protein database and searched for hypothetical proteins of *Pseudomonas aeruginosa* SE5458. A hypothetical protein (Accession No.: WP_003111794.1) consisting of 94 amino acids was chosen for detailed structural and functional exploration. The sequence was stored as FASTA format for subsequent analysis.

Determination of physicochemical properties: Physiochemical properties of the protein were analyzed using ProtParam (http://web.expasy.org/protparam/) [20] tool of ExPASy server. This tool utilizes primary sequence of the protein to estimate physical and chemical properties. Various parameters including the amino acid composition, molecular weight, extinction coefficient, estimated half-life, theoretical pI, instability index, aliphatic index and grand average of hydropathicity (GRAVY) were determined.

Subcellular localization prediction: Identifying protein's cellular location is crucial step toward understanding its cellular function. CELLO (http://cello.life.nctu.edu.tw/) [21], LocTree3 (https://rostlab.org/services/loctree3/) [22] and PSLpred (http://crdd.osdd.net/raghava/pslpred/) [23] tools were used to predict subcellular localization of the hypothetical protein WP_003111794.1.

Functional annotation: NCBI Conserved Domain Search Service (CD Search) (https://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi), Pfam (https://pfam.xfam.org/) and InterProScan5 (http://www.ebi.ac.uk/Tools/services/web/ toolform.ebi?tool=iprscan5) was used to annotate the protein based on domain analysis. CD Search compares a query protein sequence by performing RPS-BLAST (Reverse Position-Specific BLAST) against position-specific score matrices resulting from conserved domain alignments present in the Conserved Domain Database (CDD) [24]. Pfam is a protein family database that includes annotations and multiple sequence alignments generated using Hidden Markov models (HMMs) [25]. InterProScan5 provides functional analysis of proteins by classifying them into families and predicting domains and important sites [26].

Structure determination: Secondary structure of the protein was predicted by using the self optimized prediction method with alignment (SOPMA) (https://npsa-prabi.ibcp.fr/cgi-bin/npsa_automat.pl?page=/NPSA/npsa_sopma.html) [27] and PSIPRED (http://bioinf.cs.ucl.ac.uk/psipred/) [28] servers.

MODELLER through HHpred tools of the Max Planck Institute for Developmental Biology [29] was used to develop the three dimensional (3D) structure of the protein based on homology modeling. The server performs BLASTp search to fetch templates for each protein sequence given in FASTA format. The template protein 4ZV0_B (PDB ID: 4ZV0) was selected to initiate modeling. This is an X-ray diffraction model of Tse6/Tsi6 complex from *Pseudomonas aeruginosa* PAO1. BIOVIA Discovery Studio Visualizer (version 20.1.0.19295) was used to view the 3D structure.

Structure quality assessment: PROCHECK [30], Verify3D [31] and ERRAT [32] programs through structure validation server SAVES version 6.0 (https://saves.mbi.ucla.edu/) was used to assess the quality of the developed 3D structure. QMEAN4 value was determined by QMEAN program of ExPASy server of SWISS-MODEL Workspace (https://swissmodel.expasy.org/qmean/) [33]. The value is used to assess whether the developed structure is in good agreement with the experimental structures of similar size. The Z score for template and target protein was also determined by ProSA-web server [34].

Active site prediction: Computed Atlas of Surface Topography of proteins (CASTp) (http://sts.bioe.uic.edu/castp/) server was used to determine the active site of the protein. The developed 3D structure in PDB format was uploaded in the server. CASTp locates all pockets and voids on a protein structure and measures the volume and area of each pocket and void analytically [35].

Comparative genomics approach: A BLASTp search (https://blast.ncbi.nlm.nih.gov/Blast.cgi?PAGE=Proteins) [36] between *Homo sapiens* proteome and our target protein was performed to test if our target protein has any human resemblance. Protein sequence was given as FASTA format with default parameters.

Protein-protein interaction network: STRING (https://string-db.org/cgi/input.pl) database was used to build protein-protein interaction network (PPI) of the target hypothetical protein. It is a database of known and

predicted protein-protein interactions which include both direct (physical) and indirect (functional) associations derived from computational prediction, knowledge transfer and curation of other primary databases [37].

III. Results

Physicochemical properties: Various physicochemical properties of the hypothetical protein WP_003111794.1 were estimated by ProtParam tool (Table 1). The protein contains 94 amino acid residues, possess a molecular weight of 10465.23, theoretical pI of 8.89 and grand average of hydropathicity (GRAVY) of -0.018. The instability index (II) of the target protein was found to be 36.06 indicating the protein as stable one.

Table 1: Physicochemical properties of the hypothetical protein

Properties	Value
No. of amino acids	94
Molecular weight	10465.23
Theoretical pI	8.89
Negatively charged residues (Asp + Glu)	12
Positively charged residues (Arg + Lys)	14
Extinction coefficients (M-1 cm-1)	4470
Estimated half-life (in vitro)	30 hours
Instability index (II)	36.06
Aliphatic index	111.06
Grand average of hydropathicity (GRAVY)	-0.018

Subcellular localization: The subcellular location of our target hypothetical protein was predicted by three different tools like CELLO, LocTree3 and PSLpred. The protein was predicted as cytoplasmic protein by all the tools with high confidence. Cellular location of a hypothetical protein can give us insights about cellular function which can also be utilized to design a therapeutic against a desired protein [38].

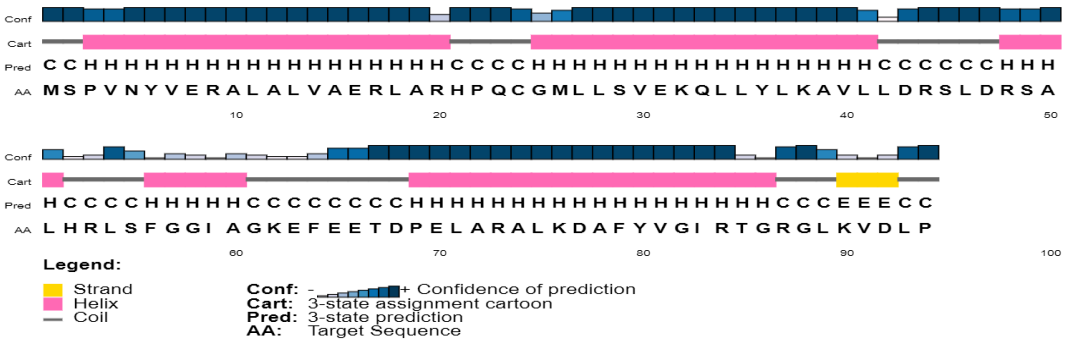

Figure 1: Secondary structure of the target protein obtained from PSIPRED server.

Functional annotation result: Several web-based annotation tools were used to identify conserved domains of our target protein. According to predictions made by NCBI-CD Search, Pfam and InterProScan5, the target protein was suggested as Tsi6 family proteins with high confidence. NCBI-CDD server predicted the Tsi6 domain at 2-84 amino acid residues with an E-value (expected value) of 9.88e-34. Pfam also predicted the Tsi6 domain at same position with an E-value of 2.4e-28. The result was also consistent (3-83 amino acids position) with InterProScan 5 with an E-value of 1.4e-28.

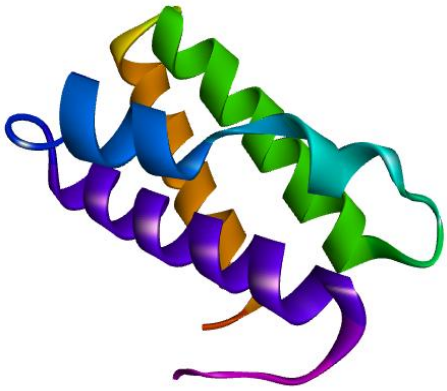

Figure 2: Three dimensional structure of the target protein developed through MODELLER (visualized by BIOVIA Discovery Studio Visualizer).

Table 2: Ramachandran plot statistics of the target protein

Ramachandran Plot Statistics	Number of a.a residues	Percentage (%)
Residues in most favored regions [A, B, L]	80	97.6%
Residues in additional allowed regions [a, b, l, p]	2	2.4%
Residues in 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	82	100%
Number of end-residues (excl. Gly and Pro)	1	
Number of glycine residues (shown as triangles)	7	
Number of proline residues	4	
Total number of residues	94	

Secondary and tertiary structure analysis: The secondary structure of the target protein was primarily predicted by the SOPMA server. Alpha helix predominates (64.89%) in secondary structure types followed by random coil (17.02%), extended strand (10.64%) and beta turn (7.45%). Similar results were also observed in PSIPRED prediction (Figure 1). Tertiary structure (3D) of the protein was obtained from MODELLER server using the template 4ZV0_B (PDB ID: 4ZV0). The template protein is an X-ray diffraction model of Tse6/Tsi6 complex from *Pseudomonas aeruginosa* PAO1 further strengthening our annotation. The 3D structure derived from MODELLER is shown in Figure 2.

3D structure quality assessment result: According to PROCHECK, 97.6% amino acid residues fell within the most favored region in "Ramachandran plot" (Table 2 and Figure 3A). The model structure successfully passed the Verify 3D server where 98.94% of the residues have averaged 3D-1D score >= 0.2. QMEAN4 value generated through QMEAN program was -0.83 which indicates that the structure is compliant with the experimental structures of similar size (Figure 3B). ERRAT also predicted the protein structure to be of good quality with a quality factor of 94.186.

The Z score generated from ProSA was -5.51 for the template (Figure 4A) and -5.45 for the model (Figure 4B) proposing homology between them. It implies that the input structure is within the range of scores usually found for native proteins of similar size [34].

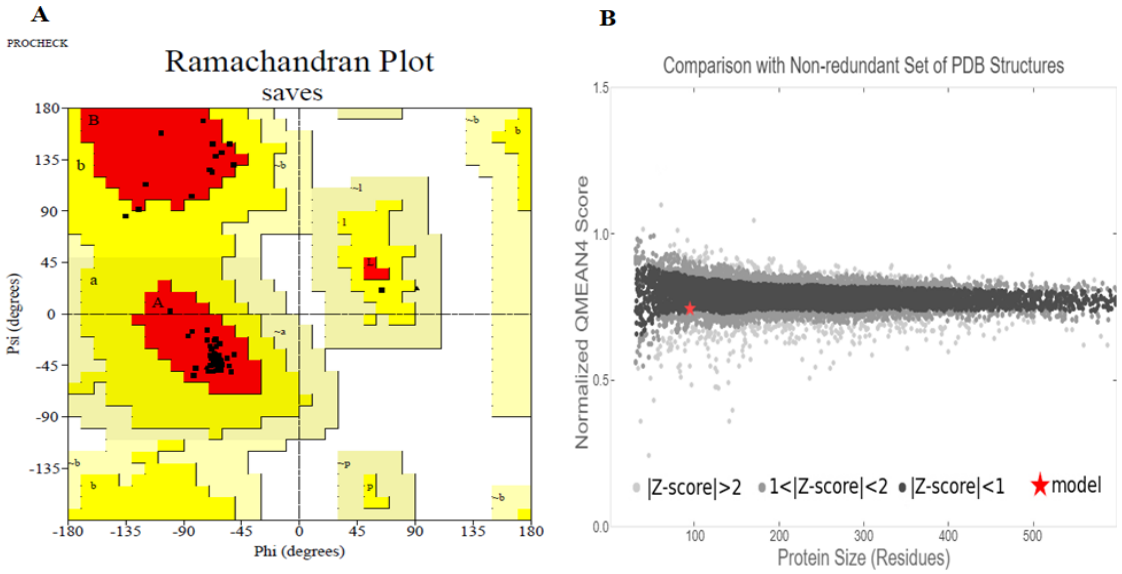

Figure 3: Quality assessment of the developed 3D structure. (A) Ramachandran plot of the target protein. (B) Graphical output of QMEAN result.

Active site analysis: The active site of the developed 3D structure was determined using the CASTp server (Figure 5). The best active site was found to be a pocket with solvent-accessible (SA) surface area of 53.311 and volume of 31.999. Key active residues predicted from CASTp are Arg[17], Arg[20], His[21], Cys[24], Leu[27], Phe[64], Thr[67] and Asp[68]. Identification of active site amino acids is the key step toward the design of a drug or an inhibitor.

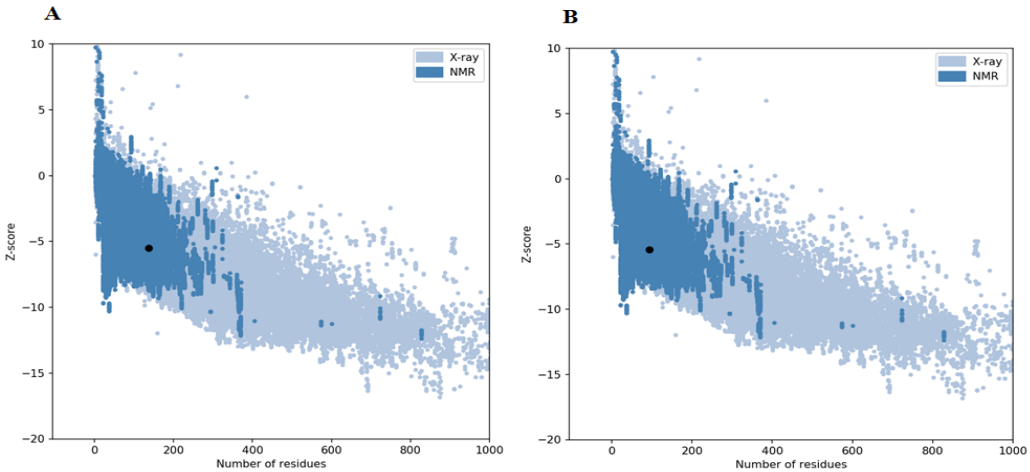

Figure 4: Z scores of the proteins determined by ProSA server (A: template protein, B: target protein).

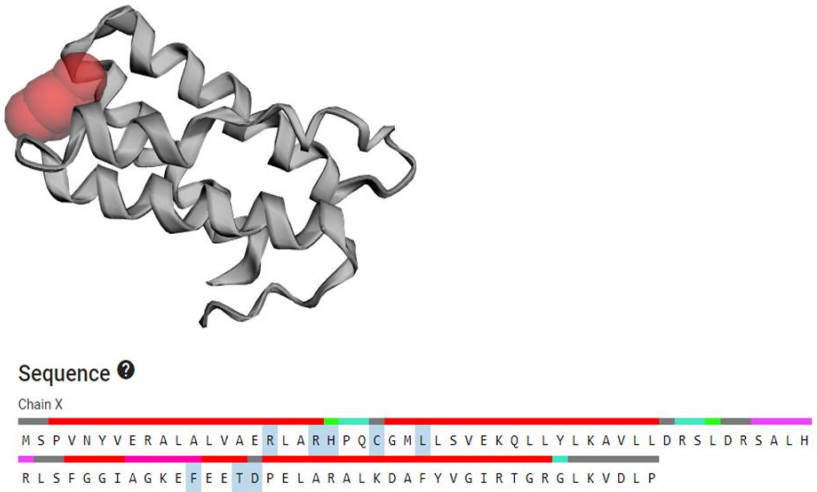

Figure 5: Active site of the protein predicted by CASTp server. The red sphere indicates the active site (upper) and the highlighted residues indicate amino acids of that active site (below).

Comparative genomics: A BLASTp search against the human proteome did not reveal any homology to any known human protein. Hence, the protein was identified as a unique protein of *Pseudomonas aeruginosa*. Bacterial proteins that are non-homologous to human proteins might be potential candidates in terms of developing drug or vaccine.

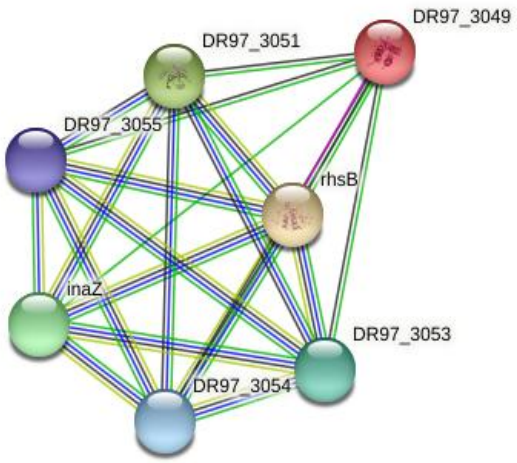

Figure 6: Protein-protein interaction network from STRING database (DR97_3049 denotes target protein).

Protein-protein interaction: Different interacting partners of our target protein were predicted using STRING database (Figure 6). The interacting proteins included RhsB, InaZ, DR97_3051, DR97_3053, DR97_3054 and DR97_3055. RhsB is a PAAR motif-containing protein located in cell membrane. Ice nucleation protein, InaZ is a member protein of Type VI secretion system (T6SS). Rests of the interacting partners are uncharacterized proteins of *Pseudomonas aeruginosa.*

IV. Discussion

About 15% of the proteins of *Pseudomonas aeruginosa* SE5458 strain are hypothetical. Annotation of these proteins is essential in order to better understand its pathogenicity and disease progression. In this study, we chose the hypothetical protein WP_003111794.1 for characterization by various bioinformatics resources. The 94 amino acids containing protein was found as a stable protein with a molecular weight of 10465.23, theoretical pI of 8.89 and grand average of hydropathicity (GRAVY) of -0.018 (Table 1). The protein was predicted as a cytoplasmic protein by CELLO, LocTree3 and PSLpred tools. Secondary structure types include alpha helix (64.89%), random coil (17.02%), extended strand (10.64%) and beta turn (7.45%) (Figure 1). The protein was found to be a Tsi6 superfamily protein according to consistent prediction made by NCBI CD search, Pfam and InterProScan 5 tools with high confidence. Tsi6 neutralize the toxic activities of Tse6, a T6S integral membrane toxin from *Pseudomonas aeruginosa*. Tse6 influences the composition of microbial communities through the delivery of toxins between neighboring bacterial cells. It acts on target cells by degrading the universally essential dinucleotides NAD+ and NADP+. Tsi6 interacts with the putative NAD+ binding pocket of Tse6, suggesting its importance for the toxic activity of Tse6 [39, 40]. We also developed 3D structure of the protein through MODELLER server (Figure 2) which successfully passed several quality assessment tools like PROCHECK, Verify 3D, QMEAN and ERRAT. The active site and interacting amino acids were determined through CASTp (Figure 5). From STRING database, several interacting partners of our target protein were identified. Among them, InaZ and RhsB are parts (like Tse6) of type VI secretion system [41, 42] further supporting our annotation. Finally, the protein was found to be a unique protein of *P. aeruginosa* non-homologous to human through comparative genomics approach. Further experimental validations are needed to confirm our findings. We underscore the importance of continued research on this crucial protein to understand bacterial pathogenesis and to decipher its role in interbacterial competition.

V. Conclusion

The hypothetical protein was successfully annotated from both structural and functional aspects. It was found to be a Tsi6 family protein that neutralizes the toxic activities of Tse6.

References

[1]. Bassetti M, Vena A, Croxatto A, Righi E, Guery B. How to manage Pseudomonas aeruginosa infections. *Drugs in context* 2018, 7: 212527.

[2]. Lyczak JB, Cannon CL, Pier GB. Establishment of Pseudomonas aeruginosa infection: lessons from a versatile opportunist1*Address for correspondence: Channing Laboratory, 181 Longwood Avenue, Boston, MA 02115, USA. *Microbes and Infection* 2000, 2(9): 1051-1060.

[3]. National Nosocomial Infections Surveillance (NNIS) System Report, data summary from January 1992 through June 2004, issued October 2004. *American journal of infection control* 2004, 32(8): 470-485.

[4]. Nathwani D, Raman G, Sulham K, Gavaghan M, Menon V. Clinical and economic consequences of hospital-acquired resistant and multidrug-resistant Pseudomonas aeruginosa infections: a systematic review and meta-analysis. *Antimicrobial resistance and infection control* 2014, 3(1): 32.

[5]. Trubiano JA, Padiglione AA. Nosocomial infections in the intensive care unit. *Anaesthesia & Intensive Care Medicine* 2015, 16(12): 598-602.

[6]. Fujii A, Seki M, Higashiguchi M, Tachibana I, Kumanogoh A, Tomono K. Community-acquired, hospital-acquired, and healthcare-associated pneumonia caused by Pseudomonas aeruginosa. *Respiratory medicine case reports* 2014, 12: 30-33.

[7]. Sligl WI, Dragan T, Smith SW. Nosocomial Gram-negative bacteremia in intensive care: epidemiology, antimicrobial susceptibilities, and outcomes. *International journal of infectious diseases : IJID : official publication of the International Society for Infectious Diseases* 2015, 37: 129-134.

[8]. Öncül O, Öksüz S, Acar A, Ülkür E, Turhan V, Uygur F, *et al.* Nosocomial infection characteristics in a burn intensive care unit: analysis of an eleven-year active surveillance. *Burns : journal of the International Society for Burn Injuries* 2014, 40(5): 835-841.

[9]. Rice LB. Federal funding for the study of antimicrobial resistance in nosocomial pathogens: no ESKAPE. *The Journal of infectious diseases* 2008, 197(8): 1079-1081.

[10]. Lubec G, Afjehi-Sadat L, Yang JW, John JP. Searching for hypothetical proteins: theory and practice based upon original data and literature. *Progress in neurobiology* 2005, 77(1-2): 90-127.

[11]. Nimrod G, Schushan M, Steinberg DM, Ben-Tal N. Detection of functionally important regions in "hypothetical proteins" of known structure. *Structure (London, England : 1993)* 2008, 16(12): 1755-1763.

[12]. Pellegrini M, Marcotte EM, Thompson MJ, Eisenberg D, Yeates TO. Assigning protein functions by comparative genome analysis: protein phylogenetic profiles. *Proceedings of the National Academy of Sciences of the United States of America* 1999, 96(8): 4285-4288.

[13]. Idrees S NS, Kanwal S, Ehsan B, Yousaf A, Nadeem S,, MI R. In silico sequence analysis, homology modeling and function annotation of Ocimum basilicum hypothetical protein G1CT28_OCIBA. *Int J Bioautomation* 2012, 16(2): 111-118.

[14]. Naqvi AA, Anjum F, Khan FI, Islam A, Ahmad F, Hassan MI. Sequence Analysis of Hypothetical Proteins from Helicobacter pylori 26695 to Identify Potential Virulence Factors. *Genomics & informatics* 2016, 14(3): 125-135.
[15]. Shahbaaz M, Hassan MI, Ahmad F. Functional annotation of conserved hypothetical proteins from Haemophilus influenzae Rd KW20. *PloS one* 2013, 8(12): e84263.
[16]. Yang Z, Zeng X, Tsui SK. Investigating function roles of hypothetical proteins encoded by the Mycobacterium tuberculosis H37Rv genome. *BMC genomics* 2019, 20(1): 394.
[17]. Islam MS, Shahik SM, Sohel M, Patwary NI, Hasan MA. In Silico Structural and Functional Annotation of Hypothetical Proteins of Vibrio cholerae O139. *Genomics & informatics* 2015, 13(2): 53-59.
[18]. Ferdous N, Reza MN, Emon MTH, Islam MS, Mohiuddin AKM, Hossain MU. Molecular characterization and functional annotation of a hypothetical protein (SCO0618) of Streptomyces coelicolor A3(2). *Genomics & informatics* 2020, 18(3): e28.
[19]. Benson DA, Karsch-Mizrachi I, Lipman DJ, Ostell J, Rapp BA, Wheeler DL. GenBank. *Nucleic acids research* 2002, 30(1): 17-20.
[20]. Gasteiger E. HC, Gattiker A., Duvaud S., Wilkins M.R., Appel R.D., Bairoch A. Protein Identification and Analysis Tools on the ExPASy Server. *(In) John M Walker (ed): The Proteomics Protocols Handbook, Humana Press* 2005: pp. 571-607.
[21]. Yu CS, Chen YC, Lu CH, Hwang JK. Prediction of protein subcellular localization. *Proteins* 2006, 64(3): 643-651.
[22]. Goldberg T, Hecht M, Hamp T, Karl T, Yachdav G, Ahmed N, *et al.* LocTree3 prediction of localization. *Nucleic acids research* 2014, 42(Web Server issue): W350-355.
[23]. Bhasin M, Garg A, Raghava GP. PSLpred: prediction of subcellular localization of bacterial proteins. *Bioinformatics (Oxford, England)* 2005, 21(10): 2522-2524.
[24]. Lu S, Wang J, Chitsaz F, Derbyshire MK, Geer RC, Gonzales NR, *et al.* CDD/SPARCLE: the conserved domain database in 2020. *Nucleic acids research* 2020, 48(D1): D265-d268.
[25]. El-Gebali S, Mistry J, Bateman A, Eddy SR, Luciani A, Potter SC, *et al.* The Pfam protein families database in 2019. *Nucleic acids research* 2019, 47(D1): D427-d432.
[26]. Madeira F, Park YM, Lee J, Buso N, Gur T, Madhusoodanan N, *et al.* The EMBL-EBI search and sequence analysis tools APIs in 2019. *Nucleic acids research* 2019, 47(W1): W636-w641.
[27]. Combet C, Blanchet C, Geourjon C, Deléage G. NPS@: network protein sequence analysis. *Trends in biochemical sciences* 2000, 25(3): 147-150.
[28]. Buchan DWA, Jones DT. The PSIPRED Protein Analysis Workbench: 20 years on. *Nucleic acids research* 2019, 47(W1): W402-w407.
[29]. Zimmermann L, Stephens A, Nam SZ, Rau D, Kübler J, Lozajic M, *et al.* A Completely Reimplemented MPI Bioinformatics Toolkit with a New HHpred Server at its Core. *Journal of molecular biology* 2018, 430(15): 2237-2243.
[30]. Laskowski R A MMW, Moss D S, Thornton J M. PROCHECK - a program to check the stereochemical quality of protein structures. *J App Cryst* 1993, 26: 283-291.
[31]. Eisenberg D, Lüthy R, Bowie JU. VERIFY3D: assessment of protein models with three-dimensional profiles. *Methods in enzymology* 1997, 277: 396-404.
[32]. Colovos C, Yeates TO. Verification of protein structures: patterns of nonbonded atomic interactions. *Protein science : a publication of the Protein Society* 1993, 2(9): 1511-1519.
[33]. Benkert P, Biasini M, Schwede T. Toward the estimation of the absolute quality of individual protein structure models. *Bioinformatics (Oxford, England)* 2011, 27(3): 343-350.
[34]. Wiederstein M, Sippl MJ. ProSA-web: interactive web service for the recognition of errors in three-dimensional structures of proteins. *Nucleic acids research* 2007, 35(Web Server issue): W407-410.
[35]. Tian W, Chen C, Lei X, Zhao J, Liang J. CASTp 3.0: computed atlas of surface topography of proteins. *Nucleic acids research* 2018, 46(W1): W363-w367.
[36]. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. Basic local alignment search tool. *Journal of molecular biology* 1990, 215(3): 403-410.
[37]. Szklarczyk D, Gable AL, Lyon D, Junge A, Wyder S, Huerta-Cepas J, *et al.* STRING v11: protein-protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. *Nucleic acids research* 2019, 47(D1): D607-d613.
[38]. Wang J, Sung WK, Krishnan A, Li KB. Protein subcellular localization prediction for Gram-negative bacteria using amino acid subalphabets and a combination of multiple support vector machines. *BMC bioinformatics* 2005, 6: 174.
[39]. Whitney JC, Quentin D, Sawai S, LeRoux M, Harding BN, Ledvina HE, *et al.* An interbacterial NAD(P)(+) glycohydrolase toxin requires elongation factor Tu for delivery to target cells. *Cell* 2015, 163(3): 607-619.
[40]. Tang JY, Bullen NP, Ahmad S, Whitney JC. Diverse NADase effector families mediate interbacterial antagonism via the type VI secretion system. *The Journal of biological chemistry* 2018, 293(5): 1504-1514.
[41]. Koskiniemi S, Lamoureux JG, Nikolakakis KC, t'Kint de Roodenbeke C, Kaplan MD, Low DA, *et al.* Rhs proteins from diverse bacteria mediate intercellular competition. *Proc Natl Acad Sci U S A* 2013, 110(17): 7032-7037.
[42]. Quentin D, Ahmad S, Shanthamoorthy P, Mougous JD, Whitney JC, Raunser S. Mechanism of loading and translocation of type VI secretion system effector Tse6. *Nat Microbiol* 2018, 3(10): 1142-1152.

Md. Fazley Rabbi, et. al. "Computational Annotation of a Hypothetical Protein from *Pseudomonas aeruginosa* SE5458 Strain Reveals a Bacterial Antitoxin Protein, Tsi6." *IOSR Journal of Pharmacy and Biological Sciences (IOSR-JPBS)*, 16(1), (2021): pp. 21-27.