International Journal of Pharmacology

ISSN 1811-7775

science alert

ansinet
Asian Network for Scientific Information

OPEN ACCESS International Journal of Pharmacology

ISSN 1811-7775
DOI: 10.3923/ijp.2021.262.270

Research Article
Functional and Structural Annotation of a Hypothetical Protein (PA2373) from *Pseudomonas aeruginosa* PA01

[1]Alariqi Reem, [2]Rokayya Sami, [3]Marwa. Y.F. Koko, [4]Abeer Essam Noman, [5]Yousif A. Algabri, [6]Raj Aditya Kumar, [2]Ebtihal Khojah and [1]Zhao-Hua Zhong

[1]Department of Microbiology, Harbin Medical University, 157 Baojian Road, Harbin 150081, China
[2]Department of Food Science and Nutrition, College of Sciences, Taif University, P.O. 11099, Taif 21944, Saudi Arabia
[3]Department of Food Science, Grease and Vegetable Protein, Northeast Agriculture University, China
[4]College of Food Science and Technology, HuaZhong Agricultural University, Wuhan 430070, China
[5]Department of Biomedical Engineering, School of Control Science and Engineering, Shandong University, Jinan, Shandong, 250061, China
[6]Department of Physiology, Harbin Medical University,157 Baojian Road, Harbin 150081China

Abstract

Background and Objective: A comprehensive study of *Pseudomonas aeruginosa* proteomics found that 25% of proteins are retained as hypothetical proteins whose functions have not yet been ultimately determined. In this study, it's attempted to assign a particular role to one such hypothetical protein PA2373 for which no experimental knowledge is currently available. **Materials and Methods:** To achieve this, the newest versions of bioinformatics tools was applied that provide various information, such as regarding amino acid sequences, protein families, sequence-function relationships and motifs. To identify homologous proteins, sequence similarities were searched using accessible bioinformatics databases. **Results:** Obtained results showed that PA2373 has two functional domains, including VI_Rhs_Vgr and 5 superfamilies. Functional annotation showed that PA2373 could be a Vgr protein and type VI secretion system. In addition, protein-protein interactions of selected hypothetical proteins show that certain functional partners play a significant role in pathogen survival. **Conclusion:** This study's findings may help better understand the mechanisms of virulence, drug resistance and pathogenesis of *P. aeruginosa* infections and to innovate the treatment strategies subsequently.

Key words: Hypothetical protein, VgrG1, T6SSs, virulence, *P. aeruginosa*, PA2373

Citation: Reem, A., R. Sami, M.Y.F. Koko, A.E. Noman, Y.A. Algabri, R.A. Kumar, E. Khojah and Z.H. Zhong, 2021. Functional and structural annotation of a hypothetical protein (PA2373) from *Pseudomonas aeruginosa* PA01. Int. J. Pharmacol., 17: 262-270.

Corresponding Author: Rokayya Sami, Department of Food Science and Nutrition, College of Sciences, Taif University, P.O. 11099, Taif 21944, Saudi Arabia
Zhao-Hua Zhong, Department of Microbiology, Harbin Medical University, 157 Baojian Road, Harbin 150081, China

Copyright: © 2021 Alariqi Reem *et al.* This is an open access article distributed under the terms of the creative commons attribution License, which permits unrestricted use, distribution and reproduction in any medium, provided the original author and source are credited.

Competing Interest: The authors have declared that no competing interest exists.

Data Availability: All relevant data are within the paper and its supporting information files.

INTRODUCTION

Secretion systems of proteins are essential for bacteria to grow and survive in various environments. In some cases, pathogenic bacteria use secretion systems to control host immunity and create a replica space[1]. In eukaryotic cells, this secretory strategy entails the translocation of altered molecules from the inside to the outside of the cell, as well as the transport of effector molecules such as toxins, enzymes or proteins. Therefore, secretion is a vital part of the pathogenesis and virulence of bacteria[2]. There are different types of bacterial secretion systems; the type VI secretion system (T6SS) is the sixth and the most recently known bacterial secretion system[3]. T6SS has been identified in *P. aeruginosa* and included in antibacterial competition and virulence[4]. Virulence factors are microorganism-secreted molecules that allow them to colonize the host, suppress the host's immune responses and cause illness. Thus, understanding the molecular mechanisms of bacterial virulence plays a vital role in bacterial pathogenesis.

T6SS is very wide, with a maximum of 21 proteins encoded in an adjacent cluster of genes. Furthermore, it has three major components of Multimeric: the baseplate (TssEFGK), transmembrane core (TssJLM), tail and sheath complex (Hcp, VgrG and TssBC)[5]. Intriguingly, T6SS shares structural homology with phage tails. Moreover, VgrG1 is structurally like the T4 phage (gp27)3-(gp5)3 puncturing. The compound of the spike is organized as a needle-like β-helix and is sited at the tip of the machinery of the shooting phage[6]. The dual hemolysin co-regulated protein 1 and valine-glycine repeat protein G1 (Hcp1/VgrG1) approach has been suggested to act as a spike wherein the long tail constitutes Hcp1 and the penetrating tip is VgrG1. Similar to the tail and tip used by phages to inject the host cell with their DNA, together, each of these proteins crosses the target membrane[7]. Furthermore, several VgrG proteins contain C-terminal extensions which have an effector function, such as the VgrG-1 actin cross-linking domain (ACD). VgrG-1 or its ACD domain is translocated to cross-link cell actin into the host cell cytosol following the diffusion of bacteria into macrophage cell lines, affecting the subsequent phagocytic process[8]. *P. aeruginosa* is a gram-negative bacillus pathogen associated with Cystic Fibrosis (CF) and affects immunocompromised patients[9]. Typically, *P. aeruginosa* genomes have sizes of approximately 6 and 7 Mb and it has the lowest AT content (33%) by 5,570 Open Reading Frames (ORFs). This genome helps *P. aeruginosa* survive and withstand the effects of several antimicrobial substances in different environments[10]. Interestingly, several bacterium proteins are measured as hypothetical proteins because their biological and structural roles remain unknown. Therefore, bioinformatics methods could play a key role in predicting and studying the structure, biological functions and protein-protein interactions of these hypothetical proteins. Accordingly, this study aimed to assign the hypothetical protein PA2373 structural and biological functions using *in silico* analysis. Subcellular localization, physiochemical properties and secondary structure were predicted; additionally, conserved domain, family and superfamily were identified and protein-protein interactions were analyzed. Homology modelling techniques were used to produce a quality model of PA2373.

MATERIALS AND METHODS

Study area: The current work was carried out at the pathogen biology Department, Harbin Medical University, from November, 2020 to April, 2021.

Sequence assembly: The hypothetical protein (PA2373) sequence information was obtained through the NCBI database. In a FASTA format, the sequence was taken and then presented for *in silico* characterization to several prediction servers (Table 1). Similarity searches were carried out with the NCBI Protein Database and Swiss Prot to identify proteins that may have structural similarities to uncharacterized proteins using the "BLAST p tool of the BLAST software.

Table 1: Tools used for the in-silico characterization of hypothetical protein PA2373

Server name	Purpose
BLAST Prot BLAST MUSCLE	Similarity search multiple sequence alignment
ProtParam PSORTb	Physicochemical characterization
CELLO SOSUIGram N TMHMM CCTOP	Topology prediction
Motif	Motif discovery
Pfam	Family relationship identification
Superfamily	Superfamily search
InterPro	Functional classification
PSIPRED	Secondary structure prediction
HHpred	Tertiary structure prediction
STRING	Interaction network analysis
PROCHECK ERRAT	Structure verification

Alignment of sequences and study of phylogeny: Multiple sequence alignments were achieved using the EBI MUSCLE server[11]. Phylogenetic analysis was performed using Phylogeny.fr[12].

Analysis of physicochemical properties: Physical and chemical features such as amino acid composition, predictable half-life, the total number of negatively charged residues (Asp +Glu), the total number of positively charged residues (Arg+Lys), aliphatic index, instability index and the grand average of hydropathicity (GRAVY) predictions were determined using the ProtParam tool of ExPASy[13].

Subcellular localization: The CELLO was used to predict the subcellular localization[14]. Upshots were also verified through subcellular localization predictions gained from SOSUIGramN[15] and PSLpred[16] TMHMM[17] and HMMTOP[18].

Conserved domain and family identification: The conserved domains on the Conserved Domains Database (CDD), available on the NCBI webpage has been used for research[19]. Two domain search methods, InterProScan[20] and Pfam[21] were also used to confirm the outcomes. The protein motif search was performed using the MOTIF search tool (Genome Net, the Institute for Chemical Research, Kyoto University, Japan)[22] and SUPERFAMILY[23]. The COILS server was used to classify the coiled-coil conformation into the protein[24]. The PFP-FunDSeqE server was used to recognize protein folding patterns[25]. The virulence activity of HPs was predicted using VICMpred[26].

Prediction of secondary structure: A PSI blast-based secondary structure prediction (PSIPRED) was implemented to investigate the structure of the protein. The Self-Optimized Prediction Method with Alignment (SOPMA), an online tool was also applied to predict the secondary protein structure[27].

Prediction of the three-dimensional structure: The HHpred server predicted the three-dimensional structure[28]. At The Max Planck Institute for Developmental Biology, Tubingen. The process used the pairwise comparison profile of Hidden Markov Models (HMMs). For advanced precision, the 3D structure was predicted based on the best score template. The homology-modelled protein validated by the expected three-dimensional model evaluated using PROCHECK, Ramachandran plots[29], ERRAT[30] and ProSA web[31]. The ProB is the server that used the homology-modelled hypothetical protein to classify structurally related protein-binding sites[32]. The homology-modelled protein was superimposed with the UCSF Chimera 1.10[33].

RESULTS AND DISCUSSION

Sequence similarity data: Homology with the other VgrG proteins were identified in the BLAST (BLASTp search) against non-redundant protein sequence which showed homology with another type VI secretion system tip protein VgrG proteins in the same genus: *Pseudomonas*, *Pseudomonas aeruginosa*, *Pseudomonas fluorescens* with sequence identity (100 and 99.8%) respectively through score 1367 and E-value 0.0. Table 2. Furthermore, Table 3 showed the homology with Type VI secretion system in the different genus of bacteria, *Aeromonas hydrophila*, *Dickeya dadantii*, *Vibrio cholerae* 01and *Vibrio cholerae* 0395, with sequence identity (34.56, 34.40, 37.11, 36.93, 33.99 and 35.34%) respectively through scores (348,342, 341, 316 and 345) and E value Ie-109,

Table 2: Similar protein obtained from non-redundant UniProt KB/swiss prot sequences

Protein ID	Protein name	Organism	Identity (%)	Score	E-value
WP_003114515.1	Type VI secretion system tip protein VgrG	*Pseudomonas*	100.0	1367	0.0
MBI9132181.1	Type VI secretion system tip protein VgrG	*Pseudomonas aeruginosa*	99.8	1367	0.0
MBG7402737.1	Type VI secretion system tip protein VgrG	*Pseudomonas aeruginosa*	99.8	1367	0.0
WP_134302382.1	Type VI secretion system tip protein VgrG	*Pseudomonas aeruginosa*	99.8	1367	0.0
WP_114230990.1	Type VI secretion system tip protein VgrG protein VgrG	*Pseudomonas aeruginosa*	99.8	1367	0.0
WP_003122692.1	VgrG	*Pseudomonas fluorescens*	99.8	1367	0.0

Table 3: Similar protein obtained from UniProt database

Sequence ID	Protein name	Organism	Identity (%)	Score	E-value
K7WKL8.1	Type VI secretion system spike protein VgrG1	*Aeromonas hydrophila*	34.56	348	1e-109
A0KJB0.1	Type VI secretion system spike protein VgrG1	*Aeromonas hydrophila*	34.40	348	6e-109
E0SAL0.1	Putative type VI secretion system protein VgrGA	*Dickeya dadantii*	37.11	342	8e-108
E0SIS4.1	Putative type VI secretion system protein VgrGB	*Dickeya dadantii*	36.93	341	1e-107
Q9KS45.1	Actin cross-linking toxin VgrG1	*Vibrio cholerae* 01	33.99	316	1e-93
A0A0H3AIG7.1	Actin cross-linking toxin VgrG1	*Vibrio cholerae* O395	35.34	345	8e-90

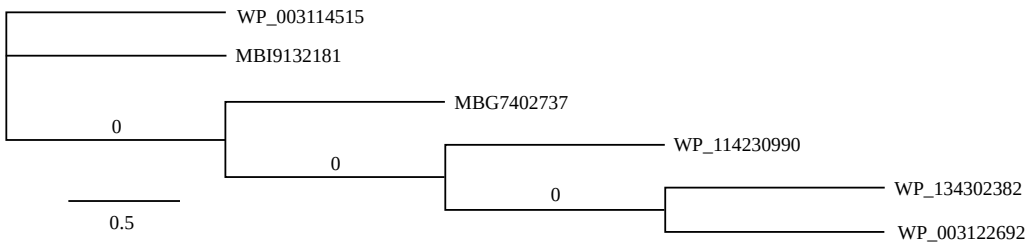

Fig. 1: Phylogenetic trees of different VgrG proteins

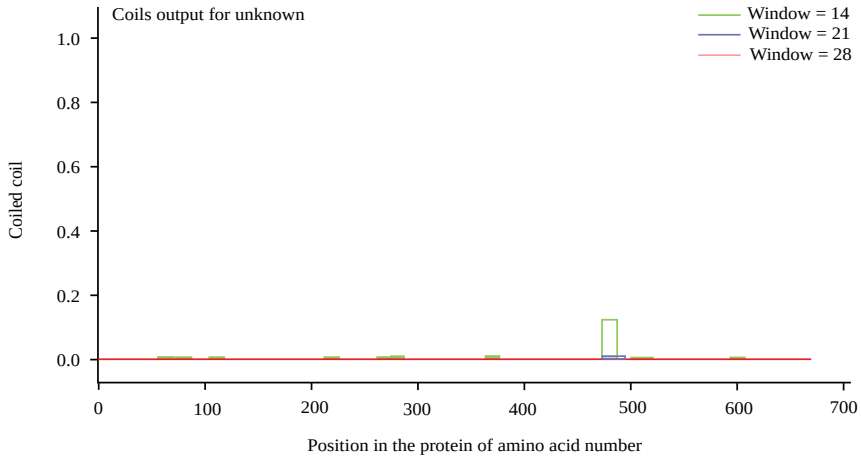

Fig. 2: Coil depicts the heptads corresponding to the residue windows 14 (green), 21(blue) and 28 (red)

61-109, 8e-108, Ie-107, Ie-93 and 8e-90) in the Swiss Prot databases. Hypothetical protein (PA2373) FASTA sequences and annotated homologous proteins were systematized. To validate the homology between the proteins, a phylogenetic analysis was also performed. Based on the alignment and BLAST results, a phylogenetic tree was constructed, which provided a similar concept of protein, as shown in Fig. 1.

Physicochemical features: The protein contained 668 amino acids and the molecular weight measured was 7877.95 Da and the theoretical pI was 5.207, representing that the protein was negatively charged. Among the most abundant charged residues, the whole number of positively charged residues (Arg+Lys) was 12 and the whole number of negatively charged residues was 13. The determined index of instability classifies the protein as stable. The measured protein half-life was 7.2 "hrs" in mammalian reticulocytes (*in vitro*), >20 hrs in yeast (*in vivo*) and >10 hrs in *Escherichia coli* (*in vivo*). The aliphatic index was 71.39, which indicates the stability of the protein in a high-temperature range. The GRAVY value was -0.335. A negative GRAVY value shows that the protein is nonpolar. The molecular protein formula was defined as $C_{356}H_{539}N_{95}O_{102}S_3$.

Subcellular localization: The hypothetical protein (2373) subcellular localization was expected to be in the extracellular space. The absence of transmembrane helices predicted by THMM and HMMTOP besides enhances the final result of extracellular protein[15-18].

Functional annotation and virulence prediction: The conserved domain search tool showed two domains in this hypothetical protein sequence, VI_Rhs_Vgr (accession no TIGR03361) and 5 superfamilies (accession no cl33691). The Pfam server predicted the Phage_GPD at 29–353 amino acid residues with an e-value of 4.9e-82, phage base V at 357-583 with an e-value of 3.1e-13 and Gp5 C at 560-646 amino acid residues with an e-value of 0.0031. The Inter-ProScan server predicted the type VI secretion system, RhsGE-related Vgr protein family subset. Phage GPD family subset, Phage_GPDPhage_base_V, Gp5_C, and DUF3540 domains were also predicted using the MOTIF server. Fold pattern predicted by the PFP-FunDSeqE tool shown the presence of conA-like lectin/glucanases; fold inside the protein sequence. The x-axis of the diagram shows the position in the protein of amino acid number (starting at the N-terminus) and the y-axis represents the coiled-coil whereas 'Window' refers to the width of the amino acid in Fig. 2.

Superfamily analyses identified the presence of phage tail proteins and the phage fibre protein superfamily. To predict the bacterial virulence factor of PA2373, we used the VICMpred bioinformatics tool. PA2373 includes a virulent element that may offer clues for future drug design as a treatment objective. The characteristics of HP PA2373 from *Pseudomonas aeruginosa* PA01 were investigated and studied for the first time. The recent finding was agreed with Wiehlmann L *et al.*[34] who made the report that HPs are attractive therapeutic targets since they are involved in important pathways such as (cell wall integrity e and maintaining genome).

Analysis of the secondary structure: The study using the secondary structure prediction SOPMA tool showed the alpha helix, beta-turn, extended strand and random protein coil proportions as 20.96, 6.8, 28.1 and 44.01%, respectively.

Analysis of the three-dimensional structure: 3D structure prediction was accomplished using the HHpred server with the highest scoring template for 100 percent identification (PDB ID: 4UHV), Fig. 3 and 4. The UHV is a trimetric fold crystal structure that separates two major fragments assembly of the protein, the head and the spike bound by a neck centred on a cord[29-31]. This unique molecular structure is identical to the T4 bacteriophage (gp27)3-(gp5)3 complex puncturing-device machinery used by the virus to inject its DNA into the host cell[34]. The energy minimization server Modrefiner was used for the refined prediction of homology models. For both the energy aspects and geometries, the ProSA web helped evaluate the energy-minimized structures by accessing the Z-score protein structure, representing the overall quality of the model, Fig. 4. It was helpful to verify whether the input structure was within the range of scores typically seen for similar-sized native proteins. The overall model quality Z-score was -6.57, suggesting a similarity between the template and query structure. The Ramachandran plot revealed that 71.9% of the residues were located in the most favourable regions, where the distribution of ϕ and ψ angle in the model within the limits are shown as illustrated in Fig. 5. The ERRAT server predicted an overall model quality factor of 58.0813, which indicates a good model. The homology-modelled protein was superimposed with the UCSF Chimera 1.10 is presented in Fig. 6.

Analysis of protein-protein interaction: The STRING 10.0 tool was used to establish potential protein functional interaction networks[35,36]. The partners of functional listed with scores

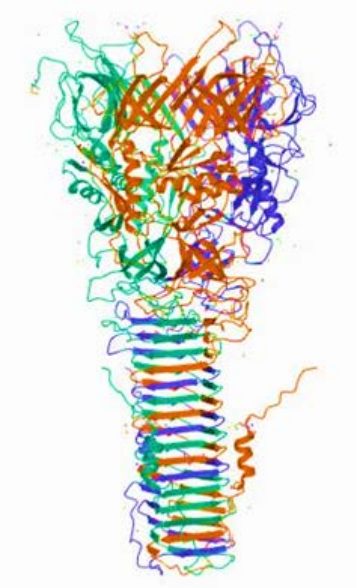

Fig. 3: Predicted three-dimensional hypothetical protein structure

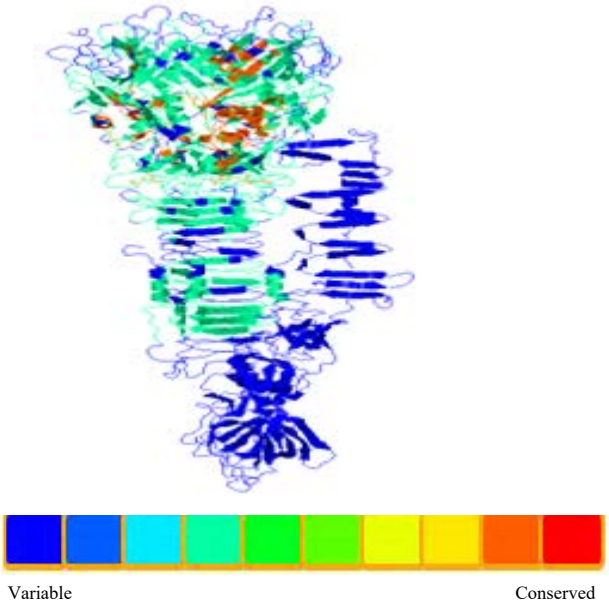

Fig. 4: Predicted functional regions of selected hypothetical protein by using ProBis server
Highly conserved residues are shown in red and the least conserved regions are shown in blue

were; PA2369 (0.938), PA2370 (0.934), PA2367 (0.928), PA2366 (0.915), ClpV1 (0.910), PA2365 (0.909), PA2371 (0.908), PA0088 (0.900), IcmF1 (0.900), PA1660 (0.899). PA2369, PA2370, PA2367, PA2365, PA0088 and PA1660 are hypothetical

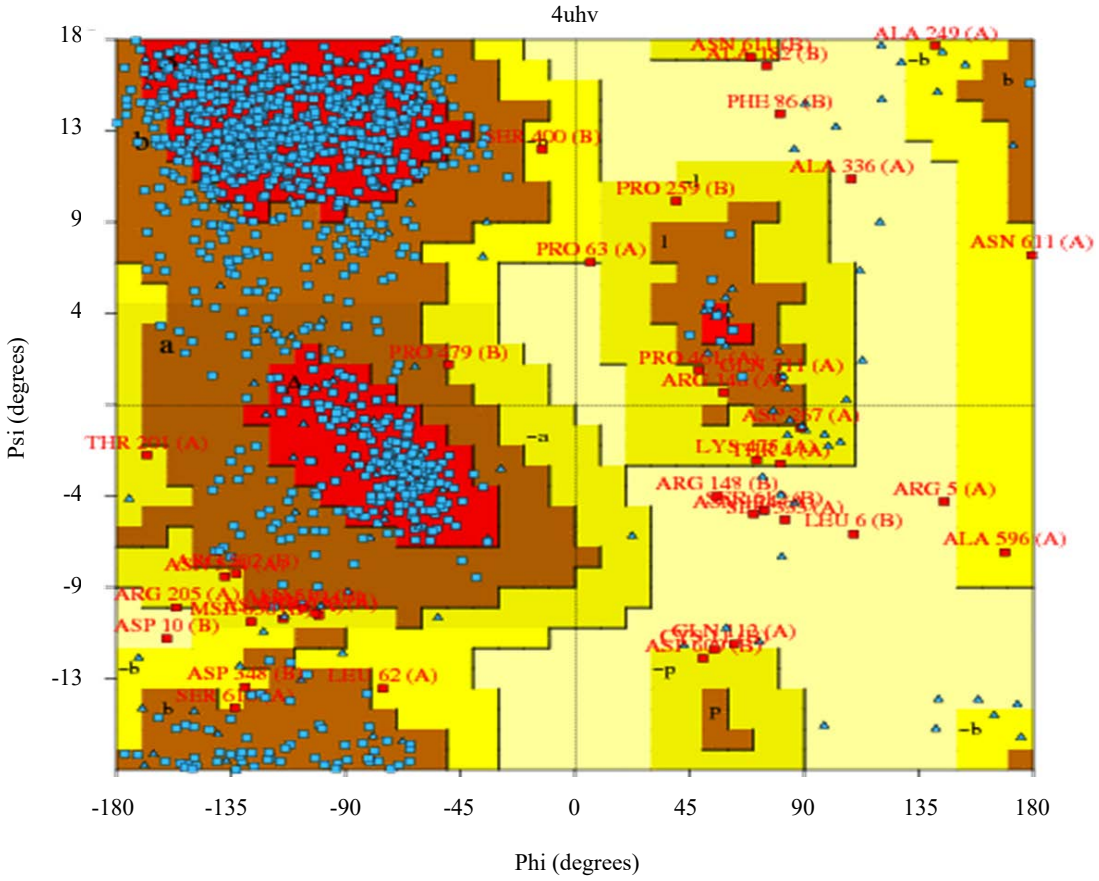

Fig. 5: Ramachandran plot of modelled structure validated by PROCHECK program

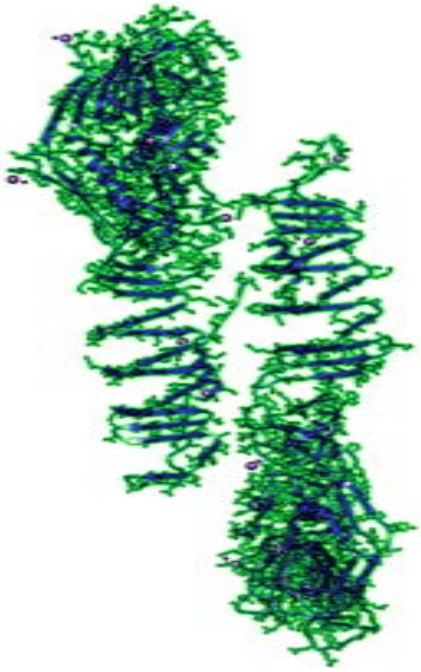

Fig. 6: Homology modelled proteins are superimposed with the template of (PDB ID: 4uhv) by using UCSF chimera-1.10

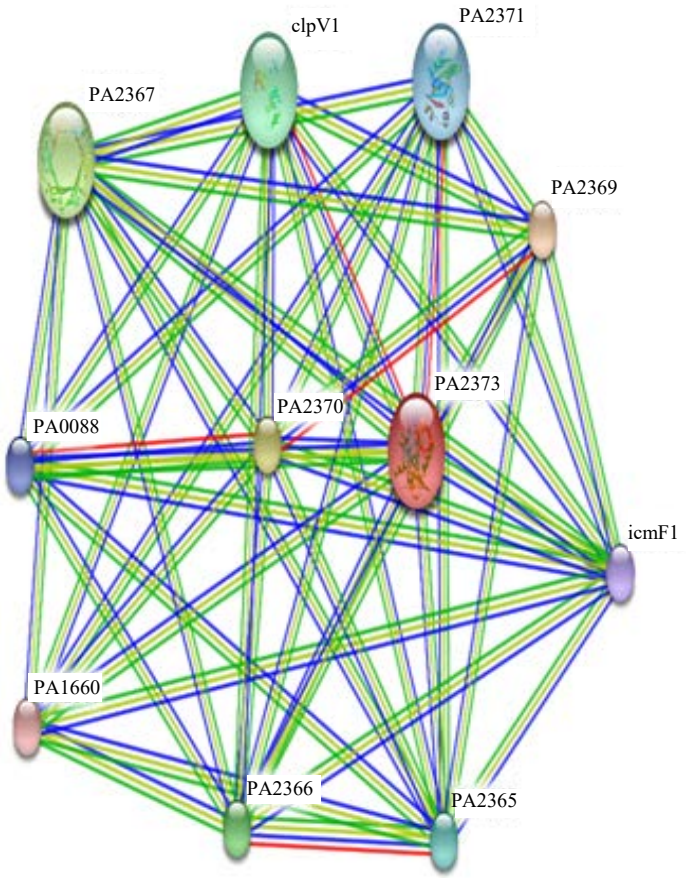

Fig. 7: A string network analysis of the hypothetical protein is identified as PA2373

proteins. PA2366 is an uricase, ClpV1 is a protein, which is required for secretion of hcp1 probably by providing an energy source for its translocation, PA2371 is a ClpA/B-type protease and IcmF1 as shown in Fig. 7.

CONCLUSION

The detection of protein functions is crucial to the understanding of biological processes. Therefore, this study aimed to define the biological and structural role of PA2373, a hypothetical protein from *P. aeruginosa*, via *in silico* analysis. The hypothetical protein that was identified revealed some characteristics, including extracellular protein, Vgr protein and type VI secretion system, containing two domains. These hypothetical protein characteristics provide basic information about *P. aeruginosa*. The findings of this study will help determine mechanisms of virulence, drug resistance and pathogenesis of *P. aeruginosa* infections through extended *in vitro* experiments to innovate treatment strategies for these infections subsequently.

SIGNIFICANCE STATEMENT

The current research is aimed at describing the structure and function annotation of HPs which have an antibiotic-resistant activity that can help find new targets to improve *P. aeruginosa* treatment and investigation. Hypothetical proteins in *P. aeruginosa* have been predicted using several structural and functional annotation informatics servers.

ACKNOWLEDGMENT

We thank our professors and support from the Department of Microbiology, Harbin Medical University, Harbin, China. Taif University Researchers Supporting Project Number (TURSP-2020/140), Taif University, Taif, Saudi Arabia.

REFERENCES

1. Green, E.R. and J. Mecsas, 2016. Bacterial Secretion Systems: An Overview. In: Virulence Mechanisms of Bacterial Pathogens, Kudva, I.T., N.A. Cornick, P.J. Plummer, Q. Zhang, T.L. Nicholson, J.P. Bannantine and B.H. Bellaire (Eds.)., American Society of Microbiology, pp: 215-239.
2. Cossart, P. and P.J. Sansonetti, 2004. Bacterial invasion: The paradigms of enteroinvasive pathogens. Science, 304: 242-248.
3. Mougous, J.D., M.E. Cuff, S. Raunser, A. Shen and M. Zhou *et al*., 2006. A virulence locus of *Pseudomonas aeruginosa* encodes a protein secretion apparatus. Science, 312: 1526-1530.
4. Kapitein, N. and A. Mogk, 2013. Deadly syringes: Type VI secretion system activities in pathogenicity and interbacterial competition. Curr. Opin. Microbiol., 16: 52-58.
5. Chen, L., Y. Zou, P. She and Y. Wu, 2015. Composition, function and regulation of T6SS in *Pseudomonas aeruginosa*. Microbiol. Res., 172: 19-25.
6. Wettstadt, S., T.E. Wood, S. Fecht and A. Filloux, 2019. Delivery of the *Pseudomonas aeruginosa* phospholipase effectors PldA and PldB in a VgrG- and H2-t6SS-dependent manner. Front. Microbiol., Vol. 10. 10.3389/fmicb.2019.01718.
7. Hood, R.D., P. Singh, F. Hsu, T. Güvener and M.A. Carl *et al*., 2010. A type VI secretion system of *Pseudomonas aeruginosa* targets a toxin to bacteria. Cell Host Microbe., 21: 25-37.
8. Cherrak, Y., N. Flaugnatti, E. Durand, L. Journet and E. Cascales, 2019. Structure and activity of the type VI secretion system. Microbiol. Spectr., Vol. 7. 10.1128/microbiolspec.PSIB-0031-2019.
9. Lorè, N.I., C. Cigana, I.D. Fino, C. Riva and M. Juhas *et al*., 2012. Cystic fibrosis-niche adaptation of *Pseudomonas aeruginosa* reduces virulence in multiple infection hosts. PLoS ONE, Vol. 7. 10.1371/journal.pone.0035648.
10. Stover, C.K., X.Q. Pham, A.L. Erwin, S.D. Mizoguchi and P. Warrener *et al*., 2000. Complete genome sequence of *Pseudomonas aeruginosa* PAO1, an opportunistic pathogen. Nature, 406: 959-964.
11. Madeira, F., Y.M. Park, J. Lee, N. Buso and T. Gur *et al*., 2019. The EMBL-EBI search and sequence analysis tools APIs in 2019. Nucleic Acids Res., 47: W636-W641.
12. Dereeper, A., V. Guignon, G. Blanc, S. Audic and S. Buffet *et al*., 2008. Phylogeny.fr: Robust phylogenetic analysis for the non-specialist. Nucleic Acids Res., 36: W465-W469.
13. Gasteiger, E., A. Gattiker, C. Hoogland, I. Ivanyi, R.D. Appel and A. Bairoch, 2003. ExPASy: The proteomics server for in-depth protein knowledge and analysis. Nucleic Acids Res., 31: 3784-3788.
14. Yu, C.S., Y.C. Chen, C.H. Lu and J.K. Hwang, 2006. Prediction of protein subcellular localization. Proteins: Struct. Funct., Bioinf., 64: 643-651.
15. Imai, K., N. Asakawa, T. Tsuji, F. Akazawa, A. Ino, M. Sonoyama and S. Mitaku, 2008. SOSUI-gramN: High performance prediction for sub-cellular localization of proteins in gram-negative bacteria. Bioinformation, 2: 417-421.
16. Guo, J., Y. Lin and X. Liu, 2006. GNBSL: A new integrative system to predict the subcellular location for gram-negative bacteria proteins. Proteomics, 6: 5099-5105.
17. Moller, S., M.D.R. Croning and R. Apweiler, 2001. Evaluation of methods for the prediction of membrane spanning regions. Bioinformatics, 17: 646-653.
18. Tusnady, G.E. and I. Simon, 2001. The HMMTOP transmembrane topology prediction server. Bioinformatics, 17: 849-850.
19. Marchler-Bauer, A., J.B. Anderson, P.F. Cherukuri, C. DeWeese-Scott and L.Y. Geer et al., 2005. CDD: A conserved domain database for protein classification. Nucleic Acids Res., 33: D192-D196.
20. Mulder, N. and R. Apweiler, 2007. InterPro and InterProScan. In: Comparative Genomics, Bergman, N.H. (Ed.)., Humana Press, pp: 59-70.
21. Finn, R.D., A. Bateman, J. Clements, P. Coggill and R.Y. Eberhardt *et al*., 2014. Pfam: The protein families database. Nucleic Acids Res., 42: D222-D230.
22. Kanehisa, M., S. Goto, S. Kawashima and A. Nakaya, 2002. The KEG databases at genomenet. Nucleic Acids Res., 30: 42-46.
23. Wilson, D., M. Madera, C. Vogel, C. Chothia and J. Gough, 2007. The SUPERFAMILY database in 2007: Families and functions. Nucleic Acids Res., 35: D308-D313.
24. Ludwiczak, J., A. Winski, K. Szczepaniak, V. Alva and S. Dunin-Horkawicz, 2019. DeepCoil-a fast and accurate prediction of coiled-coil domains in protein sequences. Bioinformatics, 35: 2790-2795.
25. Shen, H.B. and K.C. Chou, 2009. Predicting protein fold pattern with functional domain and sequential evolution information. J. Theor. Biol., 256: 441-446.
26. Saha, S. and G.P.S. Raghava, 2006. VICMpred: An SVM-based method for the prediction of functional proteins of gram-negative bacteria using amino acid patterns and composition. Genomics Proteomics Bioinf., 4: 42-47.
27. Combet, C., C. Blanchet, C. Geourjon and G. Deleage, 2000. NPS@: Network protein sequence analysis. Trends Biochem. Sci., 25: 147-150.
28. Zimmermann, L., A. Stephens, S.Z. Nam, D. Rau and J. Kübler *et al*., 2018. A completely reimplemented MPI bioinformatics toolkit with a new HHpred server at its core. J. Mol. Biol., 430: 2237-2243.
29. Carugo, O. and K. Djinović-Carugo, 2013. Half a century of ramachandran plots. Acta Crystallogr. Sect. D Biol. Crystallogr., 69: 1333-1341.
30. Nagibina, G.S., V.V. Marchenkov, K.A. Glukhova, T.N. Melnik and B.S. Melnik, 2020. Verification of the stabilized protein design based on the prediction of intrinsically disordered regions: Ribosomal proteins L1. Biochemistry, (Moscow), 85: 90-98.

31. Wiederstein, M. and M.J. Sippl, 2007. ProSA-web: Interactive web service for the recognition of errors in three-dimensional structures of proteins. Nucleic Acids Res., 35: W407-W410.
32. Konc, J. and D. Janežič, 2014. ProBiS-ligands: A web server for prediction of ligands by examination of protein binding sites. Nucleic Acids Res., 42: W215-W220.
33. Pettersen, E.F., T.D. Goddard, C.C. Huang, G.S. Couch, D.M. Greenblatt, E.C. Meng and T.E. Ferrin, 2004. UCSF chimera-A visualization system for exploratory research and analysis. J. Comput. Chem., 25: 1605-1612.
34. Wiehlmann, L., A. Munder, T. Adams, M. Juhas, H. Kolmar, P. Salunkhe and B. Tümmler, 2007. Functional genomics of *Pseudomonas aeruginosa* to identify habitat-specific determinants of pathogenicity. Int. J. Med. Microbiol., 297: 615-623.
35. Spínola-Amilibia, M., I. Davó-Siguero, F.M. Ruiz, E. Santillana, F.J. Medrano and A. Romero, 2016. The structure of VgrG1 from *Pseudomonas aeruginosa*, the needle tip of the bacterial type VI secretion system. Acta Crystallogr. Sect. D Struct. Biol., 72: 22-33.
36. Szklarczyk, D., A. Franceschini, S. Wyder, K. Forslund and D. Heller *et al*., 2014. STRING v10: Protein-protein interaction networks, integrated over the tree of life. Nucl. Acids Res., 43: D447-D452.