Original Article :

In Silico Structural and Functional Annotation of Helicobacter Pylori (Strain Hpag1) Hypothetical Protein Hpag1_0137

Mohammed Alaidarous

1Department of Medical Laboratory Sciences, College of Applied Medical Sciences, Majmaah University, Al-Majmaah 11952, Saudi Arabia

Received on 21. 11.2022, accepted for publication on 30.1.2023, dx.doi.org/10.5455/mjhs.2023.02.002

Corresponding Author:

Mohammed Alaidarous

Department of Medical Laboratory Sciences, College of Applied Medical Sciences, Majmaah University, Al-Majmaah 11952, Saudi Arabia, E-mail- m.alaidarous@mu.edu.sa

OPEN ACCESS Distributed under Creative Commons CC BY-NC-ND 4.0

Abstract

Background and Aims:

Helicobacter pylori (H. pylori) is a Gram-negative pathogenic bacterium that causes gastritis in humans. The exact molecular mechanism of most of the non-annotated hypothetical proteins involved in the virulence of *H. pylori* is yet to be identified. This paper aims to characterize the structure and function of the non-annotated hypothetical protein (HP) HPAG1_0137 from the *H. pylori* bacterium (strain HPAG1) in silic..

Methods:

We predicted several properties of HPAG1_0137, such as its physicochemical characteristics, secondary and tertiary structures, and functional annotation using different computational tools, followed by structural validation and quality assessment. We also performed virulence, protein crystallization, and active site predictions.

Results:

The study found that HPAG1_0137 has 242 amino acids and an acidic pI and is thermally stable, water-soluble, and distributed in the cytoplasm. Structurally, alpha-helices and random coils are this protein's most prominent structural components. We found structural models to be highly native and of decent quality.

Conclusions:

This study will facilitate a better understanding of the HPAG1_0137 protein molecular mechanism of action. In addition, it will provide the basis for additional proteomic, genomic data, and structural and pharmacological research for functional and therapeutic potential identification.

Keywords:

Helicobacter pylori, hypothetical protein, HPAG1_0137 protein, homology modeling, in silico characterization

Introduction:

The existence of hypothetical proteins (HPs) has been predicted. However, the HPs' in vivo functions have not yet been entirely determined.[1] In most genomes, HPs mostly cover around half of the protein-coding areas. Despite their functions not being thoroughly characterized, they

may be essential for providing complete genomic and proteomic information.[2] Our understanding of protein mosaics is still incomplete. However, appropriate structural and functional annotations for a specific genome's HPs may result in the discovery of novel structures and functions and assist in producing a list of different protein pathways and cascades.[3] More profound knowledge of protein-protein interactions and networks in various living forms, including plants and microorganisms, may unravel the structural and functional mysteries of these HPs.[4] Several HPs may serve as diagnostic markers and pharmacological targets for drug design discoveries.[5]

In recent years, numerous HPs have been found in the genomes of multiple creatures. However, complete genome annotations have not yet been attained because of drawbacks including the high cost and time needed for performing experiments.[6] In addition, it is challenging to study the large number of HPs in a genome. Hence, bioinformatics methods using various databases and algorithms to estimate protein functions are a promising alternative to laboratory-based approaches. Combined experimental results are the base of these algorithms and databases that are effectively used to assign the functional and structural annotations of HPs.[3,7-8]

Here, we aim to expand the virulent protein annotation and genome sequence information by analyzing the physicochemical characteristics, secondary structural features, generated three-dimensional (3D) protein models, comparative genomics analyses, and computational functional characteristics of the *H. pylori* HPAG1_0137 hypothetical protein. We selected this HP because the amino acid sequence of the protein is available; however, its structural details are unknown to our current knowledge. The outcomes of this work will highlight our understanding of the mechanisms of pathogenesis and furnish the base for identifying novel therapeutic targets against *H. pylori*.

Materials and Methods

Sequence retrieval

We retrieved the amino acid sequence of *H. pylori* (strain HPAG1) HPAG1_0137 from the Uniprot database (http://www.uniprot.org) using the primary accession number Q1CV18 and entry name Q1CV18_HELPH.

Physiochemical analysis

We analyzed the primary structures of HPAG1_0137 using the ExPASy-ProtParam tool.9 The examined physicochemical characteristics included the amino acid composition (AA), molecular weight (Mwt), theoretical isoelectric point (pI),

instability index (II), number of positive residues (+R), number of negative residues (-R), extinction coefficient at 280 nm (EC), aliphatic index (AI), half-life (in hrs), and grand average of hydropathy (GRAVY).

Secondary structure prediction

To predict secondary structural components, we used the default parameters of three web servers: SOPMA[10], GOR IV[11], and XtalPred-RF.[12] The servers use available x-ray crystallographic protein structures as templates to analyze the relative frequencies for each amino acid in the helixes, sheets, and coils to predict the percentages of secondary structural components (helixes, sheets, and coils).

Homology modeling

We constructed a 3D structural model of HPAG1_0137 using the homology modeling servers SWISS-MODEL[13] and Phyre2[14] with the complete amino acid sequence of the protein in FASTA format, again employing the available default options. We checked the initial structural model for errors in the 3D structure using the ERRAT[15] and Verify3D[16] programs included in the structural analysis and verification SAVES server (http://nihserver.mbi.ucla.edu/SAVES/).

Model quality and verification assessment

We verified the 3D models using the Ramachandran plot[17] and ProSA-web[18] servers. Generally, the Ramachandran plots for the best models reveal a high percentage of residues positioned in the most favored allowed regions and a lower percentage of residues positioned in the outlier or disallowed regions. The model quality assessment server ProSA-web is a z-scoring system generated using protein structures from the Protein Data Bank (PDB) based on the z-scoring system. We visualized the models using the PyMOL molecular graphics system (Version 2.0; Schrödinger LLC, New York, USA).

Protein crystallization prediction using the XtalPred-RF server

We used the online server XtalPred-RF[12] to predict protein crystallization. The server combines individual crystallization probabilities calculated for several protein features into a single crystallization score. Then, the protein is assigned to one of five crystallization classes based on this score. Based on the sequence length, isoelectric point, predicted structural disorder, GRAVY index, instability index, predicted coiled secondary structure, predicted coiled-coil structure and insertion score, hydrophobicity, side-chain entropy of surface residues, predicted surface ruggedness, and amino acid composition of the predicted protein surface, each class represents a different crystallization success

rate.

Active Site and Ligand Prediction

We predicted the active site of the protein using the PrankWeb server.[19] PrankWeb is a machine-learning web server that predicts ligand-binding sites based on a protein's combined sequence, 3D structures, immediate visual analysis, and binding pocket lists. It works directly with multidomain protein structures.

Subcellular localization and virulence prediction

We predicted the subcellular localization of HPAG1_0137 using CELLO version 2.5[20,21] and PSORTb version 3.0.2.[22] For virulence prediction, we used the MP3 server (http://metagenomics. iiserb. ac.in/ mp3/).[23] It uses support vector machines (SVMs) or hidden Markov models (HMMs) utilizing algorithms to predict the pathogenesis of a query protein.

Functional annotation

We analyzed the presence of conserved domains in the *H. pylori* HPAG1_0137 protein based on a sequence similarity search of closely orthologous family members. We used three bioinformatics tools and databases, including the Protein Families Database (Pfam)[24], InterProScan[25], and the NCBI Conserved Domains Database (NCBI-CDD)[26] for this purpose.

We employed the Pfam (https://pfam.xfam.org/)[24] and NCBI-CDD search servers[26] to predict the likely domain or family of the query protein. This can provide insight into a query protein's possible role and interactions by considering the functions and structures of similar proteins.

Comparative genomics analysis

We performed a comparative genomics analysis of HPAG1_0137 by conducting a BLASTp search (https://blast.ncbi.nlm.nih.gov/Blast.cgi?PAGE=Proteins) of homo sapiens proteomes and HPAG1_0137 to test the resemblance of HPAG1_0137 to humans. We filtered hits based on the expectation value (e-value), using the inclusion threshold set at 0.005 and a minimum bit score of 100.

Results and Discussion

Physicochemical characteristics

Although there is a large amount of genomic and proteomic data generated from the low-cost sequencing technologies, the number of research studies on HPs is scarce.[27] Characterizing HPs helps better understand bacterial metabolic pathways, disease progression, disease control strategies, and drug development.[27] Here, we used the ExPASy-ProtParam tool[9] to analyze the physicochemical properties of HPAG1_0137 HP protein (Table 1). Of the 20 amino acids, leucine and lysine had the highest percentages, and tryptophan

had the lowest (Table 1). The protein primary sequence was computed to contain 242 amino acids (Table 2). The molecular weight was estimated to be 27.493 kDa (Table 2). The computed pI value of the HPAG1_0137 protein was 6.53 (Table 2). The calculated pI value is practical when purifying a protein of interest, and it helps in screening for a suitable buffering system, ensuring the stability of the purified protein.[9,28] Table 2 shows the total number of +R residues (Arg + Lys) and -R residues (Asp + Glu). There were 32 negatively charged residues and 31 positively charged residues. Estimating the protein's charged residue composition may facilitate understating the nature of the protein's intercellular protein-protein interactions.[28,29] The extinction coefficient (EC) at a wavelength of 280 nm was 22,390 concerning the concentrations of Trp, Tyr, and cystine amino acids. Knowing the EC value of a protein can help researchers optimize the purification protocols for their protein of interest.[28] The AI for the protein is estimated to be 87.40. The protein is predicted to be hydrophobic in nature and thermally stable. This comes to a high hydrophobic amino acid composition, including Met, Gly, Leu, Ala, Phe, Pro, Val, Trp and Ile.[28] The instability index (II) value is computed to be 40.36. Studies showed that proteins are estimated to be unstable in a test tube when the computed II values are above 40 and vice versa.[9,28]

Table 1: Numbers and percentages of amino acids present in HPAG1_0137 estimated using the ExPASy-ProtParam tool.

Amino Acid	Number	%
Ala	14	5.8
Arg	7	2.9
Asn	13	5.4
Asp	11	4.5
Cys	9	3.7
Gln	6	2.5
Glu	21	8.7
Gly	13	5.4
His	5	2.1
Ile	14	5.8
Leu	24	9.9
Lys	24	9.9
Met	9	3.7
Phe	11	4.5
Pro	6	2.5
Ser	18	7.4
Thr	8	3.3
Trp	1	0.4
Tyr	11	4.5
Val	17	7.0

The ExpPASy-ProtParam tool predicts the time taken for half the amount of a protein in a cell to be degraded post-synthesis (known as protein's half-life).[30] It predicts the half-life of proteins in yeast, humans, Escherichia coli, and other similar organisms.[9,28] Here, the half-life of the protein is predicted to be more than 20 hrs in yeast (in vivo), 30 hrs in mammalian reticulocytes (in vitro), and > 10 hrs in Escherichia coli (in vivo). A standard bacterial protein has a half-life of 20 hrs.[31] Hence, further in vivo, and in vitro investigations are needed to

determine the exact half-life. The protein's GRAVY value is computed to be –0.200. It was reported that when the GRAVY value is high, this reflects the protein's hydrophilic nature and the possibility of better interaction with water molecules.[9,28]

Table 2: Physiochemical characteristics of HPAG1_0137 predicted using the ExPASy-ProtParam tool.

Parameters	Value
Number of amino acids	242
Mwt	27.493 kDa
pI	6.53
Number of -R residues (Asp + Glu) Number of +R residues (Arg + Lys)	32 31
Extinction coefficient	22,390 (Abs 0.1 % (= 1 g/l) 0.814, assuming all pairs of Cys residues formed cystines) 21890 (Abs 0.1 % (= 1 g/l) 0.796, assuming all Cys residues were reduced)
Half-life	30 hrs (in vitro - mammalian reticulocytes) > 20 hrs (in vivo - yeast) > 10 hrs (in vivo - Escherichia coli).
II	40.36
AI	87.40
GRAVY	-0.200

Secondary structural prediction

We use the SOPMA[10], GOR IV[11], and XtalPred-RF[12] servers to predict the HPAG1_0137 secondary structure. The alpha helix was the most predominant (SOPMA = 50.83%, GOR IV = 36.36%, XtalPred – RF = 58%), followed by random coil (SOPMA = 32.23%, GOR IV = 40.91%, XtalPred – RF = 30%) and beta-sheet (SOPMA = 16.95%, GOR IV = 22.73%, XtalPred – RF = 12%; Table 3).

Table 3: Protein secondary structural prediction for HPAG1_0137 using the SOPMA, GOR IV, and XtalPred-RF servers.

Secondary structure	%	Secondary structure prediction server
Alpha helix	50.83	SOPMA
	36.36	GOR IV
	58	XtalPred-RF
Beta Sheet	16.95	SOPMA
	22.73	GOR IV
	12	XtalPred-RF
Random coil	32.23	SOPMA
	40.91	GOR IV
	30	XtalPred-RF

Homology modeling, quality assessment, and structural validation

The 3D structure of HPAG1_0137 has not yet been determined. In this study, we constructed full-length HPAG1_0137 protein models using two homology modeling servers: Phyre2[14] and SWISS-MODEL.[13] The servers produced similar HPAG1_0137 protein models (Figure 1) and predicted globular structures containing many alpha helices. Intriguingly, this was consistent with the secondary structure prediction for this protein, which was predicted to have a higher alpha helix content than other secondary structures (see Table 3). To validate the constructed HPAG1_0137 3D models, the Ramachandran plot server was employed.[17] The server uses Ramachandran plots to present the backbone conformations of proteins based on phi/psi distribution in the model and the position of the non-Gly residue in the disallowed regions.[17] The generated scores estimate the quality of the protein model by comparing the model to similar-sized experimentally identified reference structures deposited in the PDB. The Ramachandran plots of the HPAG1_0137 models revealed that most of the residues were positioned in the favored and allowed regions. The Ramachandran plot calculation for the HPAG1_0137 SWISS-MODEL is 94.118% residues in the favored regions, 3.167% residues in the allowed region and 2.715% residues in the disallowed region. The Ramachandran plot calculation for the HPAG1_0137 Phyre2 model is 95.475% residues in the favored regions, 3.167% residues in the allowed region and 1.357% residues in the disallowed region. This indicates the decent quality of the models generated using the two modeling servers. We validated the structural models of HPAG1_0137 using the ProSA-web server.[18] A ProSA-web score (called the z-score) determines the modeled protein's degree of nativeness.[18] ProSA is used to confirm the quality of a modeled protein structure, where significant negative z-scores indicate nativeness. However, positive z-scores indicating erroneous or problematic models.[18] Here, the z-scores for the two SWISS-MODEL and Phyre2 HPAG1_0137 models were highly negative, -6.87 and -6.12, respectively. Indicating reasonable quality models with a high degree of nativeness. More structural studies on bacterial proteins are needed to be performed to advance computational structural modeling tools. This will produce quality protein models that can highlight the molecular functional mechanisms bacterial proteins.

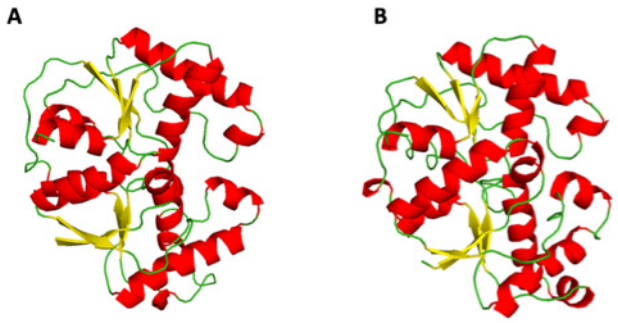


Figure 1: (A) Schematic representation of the HPAG1_0137 model generated by the SWISS_MODEL server. (B) Schematic representation of the HPAG1_0137 model generated by the Phyre2 server. The Pymol tool was used to produce the figure.

Protein crystallization prediction

Obtaining diffraction-quality protein crystals is vital for producing high-resolution x-ray protein structures.[32,33] We used the online server XtalPred-RF[12] to predict protein crystallization. The lower the score, the more promising a protein is for crystallization, and the protein in this study was unpromising for crystallization (Random forest crystallization class = 10). Hence, the protein should be subjected to experimental laboratory crystallization trials to confirm the prediction. Using truncated versions of proteins may result in promising soluble proteins and better crystallization behavior.[34]

Subcellular localization and virulence prediction

The unknown proteins' subcellular localization prediction is performed to provide information about their cellular functions. Both servers; CELLO[20,21] and PSORTb[22], predicted from subcellular localization that the query protein was cytoplasmic (Table 4). The MP3 server[23] predicted that the protein was pathogenic. This data can help in understanding disease mechanisms and establish drug discovery investigations.[35]

Table 4: Subcellular localization scores were predicted using the CELLO and PSORTb servers.

Server	Result	
CELLO	Localization scores	
	Cytoplasmic	3.766
	Periplasmic	0.728
	Inner membrane	0.047
	Outer membrane	0.116
	Extracellular	0.343
PSORTb	Localization scores	
	Cytoplasmic	2.11
	Cytoplasmic membrane	7.88
	Periplasmic	0.00
	Outer membrane.	0.00
	Extracellular	0.00

Functional annotation and comparative genomic analysis

The function of HPAG1_0137 has not yet been confirmed. Here, we used three web tools including InterProScan[25], Pfam[24], and NCBI-CDD[26] to investigate conserved domains and potential functions of the HPAG1_0137 protein. HPAG1_0137 was

indicated to contain cysteine-rich domains. NCBI-CDD predicted that this protein was a FeS-binding protein. FeS-binding proteins are found in all living organisms and are essential for their cellular processes, including gene regulation, RNA modification, DNA repair and replication, central metabolism, and respiration.[36]

We further applied a comparative genomics approach to characterize HPAG1_0137. The BLASTp (https://blast.ncbi.nlm.nih.gov/Blast.cgi?PAGE=Proteins) search against human proteomes was performed to identify whether HPAG1_0137 protein had any human homologs. HPAG1_0137 protein was identified as a unique protein of *H. pylori*. It has no homology to any human proteins. Previous studies showed that targeting bacterial proteins that are non-homologous to human proteins for pharmacological drug discoveries is an effective and safe alternative in terms of producing side effects.[7]

Conclusion

The present study aims to generate three-dimensional structure models and investigate possible functions of *H. pylori* hypothetical protein HPAG1_0137. Homology modeling generated good quality structural models for the protein. Then those models were refined using several structural assessment methods, and the outcomes were suitable subjects for further structural and functional studies. Here, using prediction tools, we found that HPAG1_0137 is a stable cytoplasmic protein with a probable FeS-binding nature. HPAG1_0137 protein was found to be essential for *H. pylori* virulence action. The protein is non-homologous to human proteins; hence, it may be used as a potential candidate for drug design studies. The outcome of the study will be the basis of future drug discovery research aims to characterize putative drug targets against *H. pylori* and other clinically significant pathogens.

Conflicts of interest

The author declares no conflicts of interest.

References

1. Lobb B, Tremblay BJ, Moreno-Hagelsieb G, Doxey AC. An assessment of genome annotation coverage across the bacterial tree of life. Microb Genom. 2020;6(3).
2. Zerbino DR, Frankish A, Flicek P. Progress, Challenges, and Surprises in Annotating the Human Genome. Annu Rev Genomics Hum Genet. 2020;21:55-79.
3. Kader MA, Ahammed A, Khan MS, Ashik SAA, Islam MS, Hossain MU. Hypothetical protein predicted to be tumor suppressor: a protein functional analysis. Genomics Inform.

2022;20(1):e6.

4. Acharya D, Dutta TK. Elucidating the network features and evolutionary attributes of intra- and interspecific protein-protein interactions between human and pathogenic bacteria. Sci Rep. 2021;11(1):190.
5. Kootery KP, Sarojini S. Structural and functional characterization of a hypothetical protein in the RD7 region in clinical isolates of Mycobacterium tuberculosis - an in silico approach to candidate vaccines. J Genet Eng Biotechnol. 2022;20(1):55.
6. Galperin MY, Koonin EV. From complete genome sequence to 'complete' understanding? Trends Biotechnol. 2010;28(8):398-406.
7. Ashrafi H, Siraji MI, Showva NN, Hossain MM, Hossan T, Hasan MA, et al. Structure to function analysis with antigenic characterization of a hypothetical protein,HPAG1_0576 from Helicobacter pylori HPAG1. Bioinformation. 2019;15(7):456-66.
8. Segelke BW. Functional Annotation from Structural Homology. Methods Mol Biol. 2022;2349:215-57.
9. Gasteiger E, Hoogland C, Gattiker A, Duvaud Se, Wilkins MR, Appel RD, et al. Protein Identification and Analysis Tools on the ExPASy Server. In: Walker JM, editor. The Proteomics Protocols Handbook. Totowa, NJ: Humana Press; 2005. p. 571-607.
10. Geourjon C, Deléage G. SOPMA: significant improvements in protein secondary structure prediction by consensus prediction from multiple alignments. Comput Appl Biosci. 1995;11(6):681-4.
11. Kouza M, Faraggi E, Kolinski A, Kloczkowski A. The GOR Method of Protein Secondary Structure Prediction and Its Application as a Protein Aggregation Prediction Tool. Methods Mol Biol. 2017;1484:7-24.
12. Slabinski L, Jaroszewski L, Rychlewski L, Wilson IA, Lesley SA, Godzik A. XtalPred: a web server for prediction of protein crystallizability. Bioinformatics. 2007;23(24):3403-5.
13. Waterhouse A, Bertoni M, Bienert S, Studer G, Tauriello G, Gumienny R, et al. SWISS-MODEL: homology modelling of protein structures and complexes. Nucleic Acids Res. 2018;46(W1):W296-W303.
14. Kelley LA, Mezulis S, Yates CM, Wass MN, Sternberg MJ. The Phyre2 web portal for protein modeling, prediction and analysis. Nat Protoc. 2015;10(6):845-58.

15. Colovos C, Yeates TO. Verification of protein structures: patterns of non-bonded atomic interactions. Protein Sci.1993;2(9):1511-9.

16. Eisenberg D, Lüthy R, Bowie JU. VERIFY3D: assessment of protein models with three-dimensional profiles. Methods Enzymol. 1997;277:396-404.

17. Anderson RJ, Weng Z, Campbell RK, Jiang X. Main-chain conformational tendencies of amino acids. Proteins. 2005;60(4):679-89.

18. Wiederstein M, Sippl MJ. ProSA-web: interactive web service for the recognition of errors in three-dimensional structures of proteins. Nucleic Acids Res. 2007;35(Web Server issue):W407-10.

19. Jendele L, Krivak R, Skoda P, Novotny M, Hoksza D. PrankWeb: a web server for ligand binding site prediction and visualization. Nucleic Acids Res. 2019;47(W1):W345-W9.

20. Yu CS, Chen YC, Lu CH, Hwang JK. Prediction of protein subcellular localization. Proteins. 2006;64(3):643-51.

21. Yu CS, Lin CJ, Hwang JK. Predicting subcellular localization of proteins for Gram-negative bacteria by support vector machines based on n-peptide compositions. Protein Sci. 2004;13(5):1402-6.

22. Yu NY, Wagner JR, Laird MR, Melli G, Rey S, Lo R, et al. PSORTb 3.0: improved protein subcellular localization prediction with refined localization subcategories and predictive capabilities for all prokaryotes. Bioinformatics. 2010;26(13):1608-15.

23. Gupta A, Kapil R, Dhakan DB, Sharma VK. MP3: a software tool for the prediction of pathogenic proteins in genomic and metagenomic data. PLoS One. 2014;9(4):e93907.

24. Mistry J, Chuguransky S, Williams L, Qureshi M, Salazar GA, Sonnhammer ELL, et al. Pfam: The protein families database in 2021. Nucleic Acids Res. 2021;49(D1):D412-D9.

25. Blum M, Chang HY, Chuguransky S, Grego T, Kandasaamy S, Mitchell A, et al. The InterPro protein families and domains database: 20 years on. Nucleic Acids Res. 2021;49(D1):D344-D54.

26. Lu S, Wang J, Chitsaz F, Derbyshire MK, Geer RC, Gonzales NR, et al. CDD/SPARCLE: the conserved domain database in 2020. Nucleic Acids Res. 2020;48(D1):D265-D8.

27. Rabbi MF, Akter SA, Hasan MJ, Amin A. In Silico Characterization of a Hypothetical Protein from. Bioinform Biol In-

sights.2021;15:11779322211011140.

28. Alaidarous M. In silico structural homology modeling and characterization of multiple N-terminal domains of selected bacterial Tcps. PeerJ. 2020;8:e10143.

29. Bhagavan NV, Ha C-E. Essentials of medical biochemistry: with clinical cases. 1st ed. Amsterdam ; Boston: Elsevier/Academic Press; 2011. xvii, 581 p. p.

30. Bojkowska K, Santoni de Sio F, Barde I, Offner S, Verp S, Heinis C, et al. Measuring in vivo protein half-life. Chem Biol. 2011;18(6):805-15.

31. Moran MA, Satinsky B, Gifford SM, Luo H, Rivers A, Chan LK, et al. Sizing up metatranscriptomics. ISME J. 2013;7(2):237-43.

32. Holcomb J, Spellmon N, Zhang Y, Doughan M, Li C, Yang Z. Protein crystallization: Eluding the bottleneck of X-ray crystallography. AIMS Biophys.2017;4(4):557-75.

33. Jia Q, Zeng H, Zhang J, Gao S, Xiao N, Tang J, et al. The Crystal Structure of the. Insects. 2021;12(7).

34. Ullah MO, Ve T, Dkhar J, Alaidarous M, Ericsson DJ, Sweet MJ, et al. Crystallization and X-ray diffraction analysis of the N-terminal domain of the Toll-like receptor signalling adaptor protein TRIF/TICAM-1. Acta Crystallographica Section F-Structural Biology and Crystallization Communications.2013;69:766-70.

35. Li FM, Gao XW. Predicting Gram-Positive Bacterial Protein Subcellular Location by Using Combined Features. Biomed Res Int. 2020;2020:9701734.

36. Py B, Barras F. Building Fe-S proteins: bacterial strategies. Nat Rev Microbiol. 2010;8(6):436-46.