Khwaja Yunus Ali University Journal

Publisher homepage: www.kyau.edu.bd

OPEN ACCESS

ISSN: 2791-3759 (Online), 2521-3121 (Print)

Journal homepage:www.journal.kyau.edu.bd

Research Article

Structure prediction and functional annotation of an uncharacterized protein bcrivmbc845_02328 of *Bacillus cereus*: An *in-silico* approach

Md Foyzur Rahman[1, *], Joyanto Mahato[2], Aoulia Yeasmin[3], Mohammad Shahangir Biswas[1], Rubait Hasan[1], Jamiatul Husna Shathi[1], Munna Kumar Podder[1], Md Shahidur Rahman[1]

[1]Department of Biochemistry and Biotechnology, Khwaja Yunus Ali University, Sirajganj, Bangladesh
[2]Department of Biochemistry and Molecular Biology, Rajshahi University, Rajshahi, Bangladesh;
[3]Department of Botany, National University, Bangladesh

***Correspondence:** foyzur.bcbt@kyau.edu.bd (Md Foyzur Rahman**,** Department of Biochemistry and Biotechnology, School of Biomedical Science, Khwaja Yunus Ali University, Sirajganj 6751, Bangladesh

Abstract

Characterization as well as prediction of secondary and tertiary structure of hypothetical proteins from their amino acid sequences uploaded in databases by in silico approach are the critical issues in computational biology. Bacillus cereus, a foodborne pathogen that causes gastrointestinal illness, possesses a wide range of proteins, many of which are still uncharacterized. Hence, the current study was conducted to reveal the physicochemical characteristics and structures of an uncharacterized protein BCRIVMBC845_02328 of B. cereus. Following the common flowchart of characterizing a hypothetical protein, several sophisticated computerized tools e.g., ExPASy Protparam, CD Search, SOPMA, PSIPRED, HHpred, etc. were employed to discover the functions and structures of BCRIVMBC845_02328. After delineating the secondary and tertiary structures of the protein, some quality evaluating tools e.g., PROCHECK, ProSA-web etc. were performed to assess the structures and later the active site was identified also by CASTp v.3.0. The protein contains more positively charged residues than negatively charged residues and a high aliphatic index value which make the protein more stable. The 2D and 3D structures modelled by several bioinformatics tools ensured that the protein has a domain in the N-terminal which indicates that it is a functional protein. Moreover, active site was found in the protein where ligand could bind. The study was aimed to unveil the features and structures of an uncharacterized protein BCRIVMBC845_02328 of B. cereus which can be a therapeutic target for development of drugs against the bacterium.

Keywords: BCRIVMBC845_02328**,** hypothetical protein, functional annotation, *Bacillus cereus*, bioinformatics.

1. Introduction

Bacillus cereus, a gram-positive, anaerobic bacterium, is associated with 2-5 % of foodborne diseases and is widely distributed in nature (air, soil and water) (Paul *et al*., 2021). It was first isolated from the cow-shed air. In 1887, Frankland isolated it from the cow-shed air and characterized as a highly motile bacterial species (Frankland and Frankland, 1887). It is mesophile in character and grows in the temperature range between 10-50°C. So, this organism can easily be spread from

the nature to the food products and causes health problems to mankind (Lindbäck *et al*., 2004; Ngamwongsatit *et al*., 2008). Since, Hauge's successful experimental evidence for *B. cereus* as a diarrheal agent in 1950s, complex protein toxin (Nhe and Hbl), monomeric protein toxin (Cytk), exoproteins and proteases or phospholipases activities of *B. cereus* have been discussed as diarrheal agents for food poisoning (Granum *et al*., 2008; Doll *et al*., 2013). More distinctly, *B. cereus* by expressing its pathogenic characteristics (by releasing toxins), plays a significant role in the development of both emetic and diarrheic forms of food poisoning (mostly mild and self-limited) (Bottone, 2010). A single heat-stable toxin released by vegetable cells during growth in food causes emetic syndrome, which manifests with nausea, vomiting, and diarrhea following ingestion. However, heat-labile enterotoxin is generated by B. cereus during its vegetative development in the small intestine, and it is responsible for the diarrheal syndrome that causes abdominal discomfort and watery diarrhea (Schoeni and Wong, 2005; Carroll *et al*., 2019). In addition to being known to cause infections in the eye and the body as a whole, this bacterium has recently come to light as a potential causal agent of wound infections. Additionally, B. cereus is responsible for a wide range of non-GI illnesses, including bacteremia, debilitating endophthalmitis, septicemia, meningitis, cellulites, abscess development, endocarditis, osteomyelitis, upper and lower urinary tract infection, pulmonary infection, and more (Glasset *et al*., 2018). Toxins from B. cereus may be very dangerous, even fatal, for preterm infants, addicts, and those with impaired immune systems. (McDowell *et al*., 2021).

Due to the diverse growing advances in computational biology, this field has attracted researchers to examine the disease-causing proteins and biomarker identification related works. Computational biology offers a wide range of resources and approaches for predicting known and unknown protein structure, homology modeling, active site identification, protein-protein interaction analysis, gene expression analysis, Pfam domain and motif searching, conserved domain identification etc. (Gong *et al*., 2019; Saikat *et al*., 2020). Moreover, an in-silico study can allow us to predict and evaluate the secondary and tertiary structure of the proteins, their functional annotation, and their roles in pharmacology for drug design against many diseases (Azevedo, 2011; Mills *et al*., 2015). These might be useful for delving further into the infected process and finding new pathways that *B. cereus* uses to cause disease, and discovering new potential drugs against this disease-causing organism.

The uncharacterized protein BCRIVMBC845_02328 is found in *B. cereus* and the function of the protein is unknown. Furthermore, neither the secondary nor the tertiary structure of this protein has been reported yet. The present study will evaluate the physiochemical properties of the protein's physiochemical characteristics, functional annotation, secondary and tertiary structure modeling and active site prediction by using various computational methods and tools.

2. Materials and Methods

2.1. Sequence Retrieval

The FASTA file containing the amino acid sequence of BCRIVMBC845 02328 was acquired from the NCBI database (https://www.ncbi.nlm.nih.gov/) with the accession ID of SCV19836.1.

2.2. Physiochemical Characterization

To characterize the physical and chemical parameters such as molecular weight, amino acid composition, the instability index, the alphabetic index, the measurement of hydrophobicity or hydrophilicity of protein, the ProtParam method was used in the ExPASy server (https://web.expasy.org/protparam/) (Gasteiger *et al*., 2005). Moreover, the theoretical isoelectric point (pI) of BCRIVMBC845_02328 was evaluated by the SMS v.2.0 server (https://www.bioinformatics.org.sms2/) (Stothard, 2000).

2.3. Sub-Cellular Localization Analysis

CELLO v.2.5, a sub-cellular localization prediction tool, was used to determine where in the cell BCRIVMBC845 02328 would be found (https://cello.life.nctu.edu.tw/) (Yu *et al*., 2006).

2.4. Functional Annotation Prediction

The CDD search tool in NCBI (https://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi) (Lu *et al*., 2020) was used to identify the conserved domain in BCRIVMBC845_02328. Protein motif search was carried out using Motif tool in Genome Net

server (https://www.genome.jp/tools/motif/) (Finn *et al*., 2014).

2.5. Secondary Structure Prediction

The secondary structure of BCRIVMBC845_02328 was predicted by the PSI-blast based secondary structure prediction tool (Jones, 1999). SOPMA framework was used for the element prediction of BCRIVMBC845_02328 (Geourjon and Deleage, 1995).

2.6. Tertiary Structure Prediction and Validation

The tertiary structures of BCRIVMBC845_02328 was predicted by the Modeller (Webb and Sali, 2016) tool following the HHpred tool (Zimmermann *et al*., 2018), Phyre2 (Kelley *et al*., 2015), and the Swiss-Model server (Amold *et al*., 2006). Experiments were performed to evaluate the projected structures' structural quality. After that, the Ramachandran plot was evaluated by PROCHECK (Laskowski *et al*., 1993) to validate the predicted tertiary structure of BCRIVMBC845_02328 protein.

2.7. Active Site Prediction

BCRIVMBC845 02328's active site was predicted using CASTp v.3.0 (http://sts.bioe.uic.edu/castp/), an online tool for identifying, outlining, and quantifying concave surface areas on tertiary structures of proteins.

3. Results

3.1. Sequence Retrieval

According to the NCBI database, the BCRIVMBC845_02328 protein of *Bacillus cereus* consists of a total of 560 amino acids. The following **Table 1** represents the further information on the uncharacterized protein BCRIVMBC845_02328:

Table 1. Protein information retrieved from the NCBI database.

Title	Information
Locus ID	*SCV19836*
Definition	Uncharacterized protein BCRIVMBC845_02328
Amino acid sequence	560 aa
Organism name	*Bacillus cereus*

3.2. Physical and Chemical Properties

The protein BCRIVMBC845_02328 contains a 560 aa long sequence with a molecular weight of 64746.99 Da and a pI of 8.16. The total number of negatively charged residues (Asp + Glu) is 71 and the total number of positively charged residues (Arg + Lys) is 74. The extinction coefficient (all pairs of Cys residues from cystines) is 87015 and the extinction coefficient (all Cys residues are reduced) is 86640. The protein has an estimated half-life of 30 hours (mammalian reticulocytes, in vitro), >20 hours (yeast, in vivo) and >10 hours (*Escherichia coli*, in vivo). Moreover, the instability index of the protein is computed to be 32.71, aliphatic index is 92.04 and the grand average of hydropathicity (GRAVY) is -0.422 (**Table 2 & 3**).

Table 2. Physicochemical properties of the uncharacterized protein BCRIVMBC845_02328.

Name of the properties	Value
Molecular weight	64746.99 Da
Theoretical pI	8.16
Total no. of negatively charged residues (Asp + Glu)	71
Total no. of positively charged residues (Arg + Lys)	74
Formula	$C_{2926}H_{4579}N_{765}O_{864}S_{14}$
Total no. of atoms	9148
Extinction coefficient (all pairs of Cys residues from cystines)	87015
Extinction coefficient (all Cys residues are reduced)	86640
Estimated half life	a) 30 hours in mammalian, *in vitro* b) >20 hours, *in vivo* c) >10 hours, *E. coli*, in vivo

Instability index (II)	32.71
Aliphatic index	92.04
Grand average of hydropathicity (GRAVY)	-0.422

Table 3. Amino acid composition of the uncharacterized protein BCRIVMBC845_02328.

Name of the mino acid	No. of amino acids	Percentage
Ala (A)	23	4.1 %
Arg (R)	20	3.6 %
Asn (N)	37	6.6 %
Asp (D)	32	5.7 %
Cys (C)	7	1.2 %
Gln (Q)	22	3.9 %
Glu (E)	39	7.0 %
Gly (G)	31	5.5 %
His (H)	13	2.3 %
Ile (I)	46	8.2 %
Leu (L)	52	9.3 %
Lys (K)	54	9.6 %
Met (M)	7	1.2 %
Phe (F)	20	3.6 %
Pro (P)	11	2.0 %
Ser (S)	31	5.5 %
Thr (T)	35	6.2 %
Trp (W)	6	1.1 %
Tyr (Y)	36	6.4 %
Val (V)	38	6.8 %
Pyl (O)	0	0.0 %
Sec (U)	0	0.0 %

3.3. Subcellular Localization

CELLO v.2.5, a subcellular localization predictor tool was used to predict the subcellular location of the uncharacterized protein BCRIVMBC845_02328 in *Bacillus cereus*. It predicted the localization of amino acid composition, N-peptide composition, partitioned sequence composition, physio-chemical composition, neighboring sequence composition with reliability values of 0.641, 0.959, 0.760, 0.345 and 0.653, respectively. The results also show the CELLO prediction of cytoplasmic, extracellular, membrane and cell wall with a reliability value of 3.335, 1.117, 0.527 and 0.021, respectively (**Table 4**).

Table 4. Subcellular localization of BCRIVMBC845_02328.

Support Vector Machine (SVM)	Localization	Reliability value
Amino acid comp.	Cytoplasmic	0.641
N-peptide comp.	Cytoplasmic	0.959
Partitioned seq. comp.	Cytoplasmic	0.760
Physichemical comp.	Extracellular	0.345
Neighboring comp.	Cytoplasmic	0.653
CELLO Prediction	Cytoplasmic	3.335*
	Extracellular	1.117
	Membrane	0.527
	Cellwall	0.021

*CELLO v.2.5 predicted the subcellular location of the protein as cytoplasmic.

3.4. Functional Annotation Prediction

The CDD search tool in NCBI website predicted a conserved domain as DUF2075 domain-containing protein with a GIY-YIG nuclease domain in N-terminus (GIY-YIG_COG3410, accession no. cd10439) (**Fig. 1**). The function of the GIY-YIG nuclease domain is unknown.

Fig. 1. Functional annotation of the uncharacterized protein BCRIVMBC845_02328.

3.5. Secondary Structure Prediction

Using a protein sequence of 560 amino acids, SOPMA projected that 34.46 percent of the residues would be in random coils, compared to 41.79% in alpha-helix, 15.89% in an extended strand, and 7.86% in a beta turn **(Table 5)**. Higher confidence in the prediction of the protein's helix, strand, or coil structure is shown by the PSIPRED result.

Table 5. Secondary structure prediction result obtained from SOMPA tool.

Elements	Values (%)
Alpha helix (Hh)	41.79 %
3_{10} helix (Gg)	0.00 %
Pi helix (Ii)	0.00 %
Beta bridge (Bb)	0.00 %
Extended strand (Ee)	15.89 %
Beta turn (Tt)	7.86 %
Beta region (Ss)	0.00 %
Random coil (Cc)	34.46 %

3.6. Tertiary Structure Prediction and Validation

The uncharacterized protein BCRIVMBC845 02328 was modelled using the Modeller tool, with the tertiary structure predicted by using the best-fitting template. Based on these metrics and others, we settled on a 476-amino-acid (aa) target template with a probability rate of 0.8%, an E value of 10.15 percent, a sequence similarity (SS) score of 1.0 percent, a col percentage (Cols) score of 40.7 percent. The modeled tertiary structure of the protein is depicted in **Fig. 2** and is maintained in the Protein Data Bank (PDB). The predicted tertiary structure of the protein was evaluated using Ramachandran plot analysis by PROCHECK (by modeller). 87.6% of the total residues (169) were located in the most preferred regions (A, B, L) as shown by the Ramachandran plot (**Fig. 3**), while 6.2% (12) were located in the additional allowed regions (a, b, l, p), 4.7% (9) were located in the generously allowed regions (a, b, l, p), and 1.6% (3) were located in the disallowed regions.

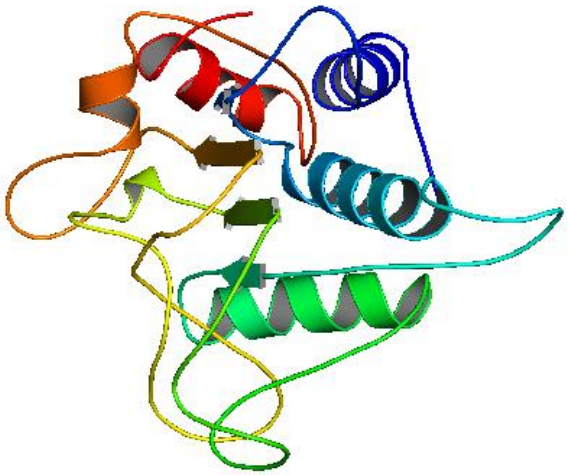

Fig. 2. Tertiary structure of the uncharacterized protein BCRIVMBC845_02328 predicted by Modeller.

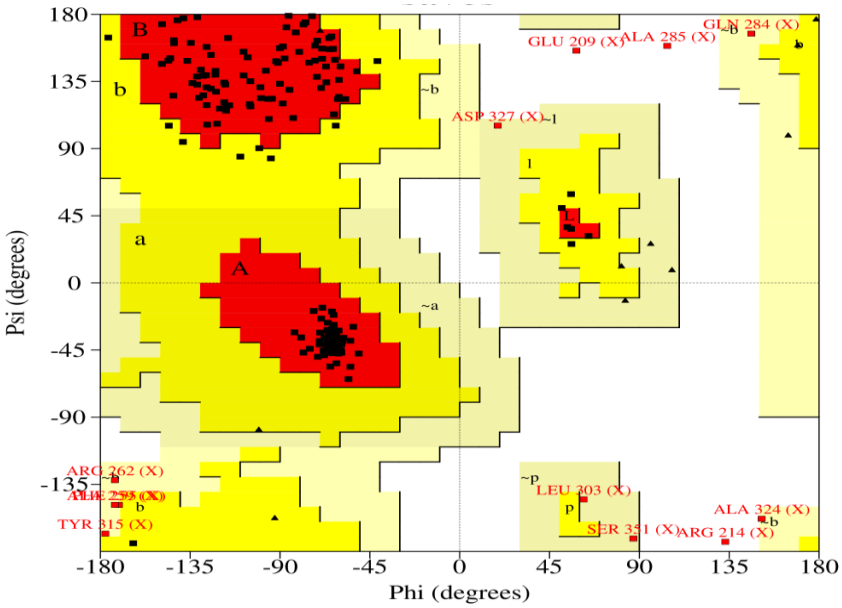

Fig. 3. Ramachandran plot analysis of the predicted tertiary protein structure.

There were 193 (100%), and neither glycine nor proline were present. Moreover, there were a total of 6 proline residues, 2 glycine residues, and 2 end residues (excluding Gly and Pro).

Table 6. Ramachandran plot reports of the uncharacterized protein BCRIVMBC845_02328.

Characteristics	No. of residues	Percentage
Residues in the most favored regions [A, B, L]	169	87.6 %
Residues in the additional allowed regions [a, b, l, p]	12	6.2 %
Residues in the generously allowed regions [~a, ~b, ~l, ~p]	9	4.7 %
Residues in the in disallowed regions	3	1.6 %
No. of non-glycine and non-proline residues	193	100 %
No. of end residues (excl. Gly and Pro)	2	1.03%
No. of glycine residues (shown as triangles)	10	5.18%
No. of proline residues	6	3.1%

3.7. Active Site Prediction

The active site of the uncharacterized protein BCRIVMBC845_02328 of *Bacillus cereus* was predicted by using the CASTp v.3.0 online tool (**Fig. 4**).

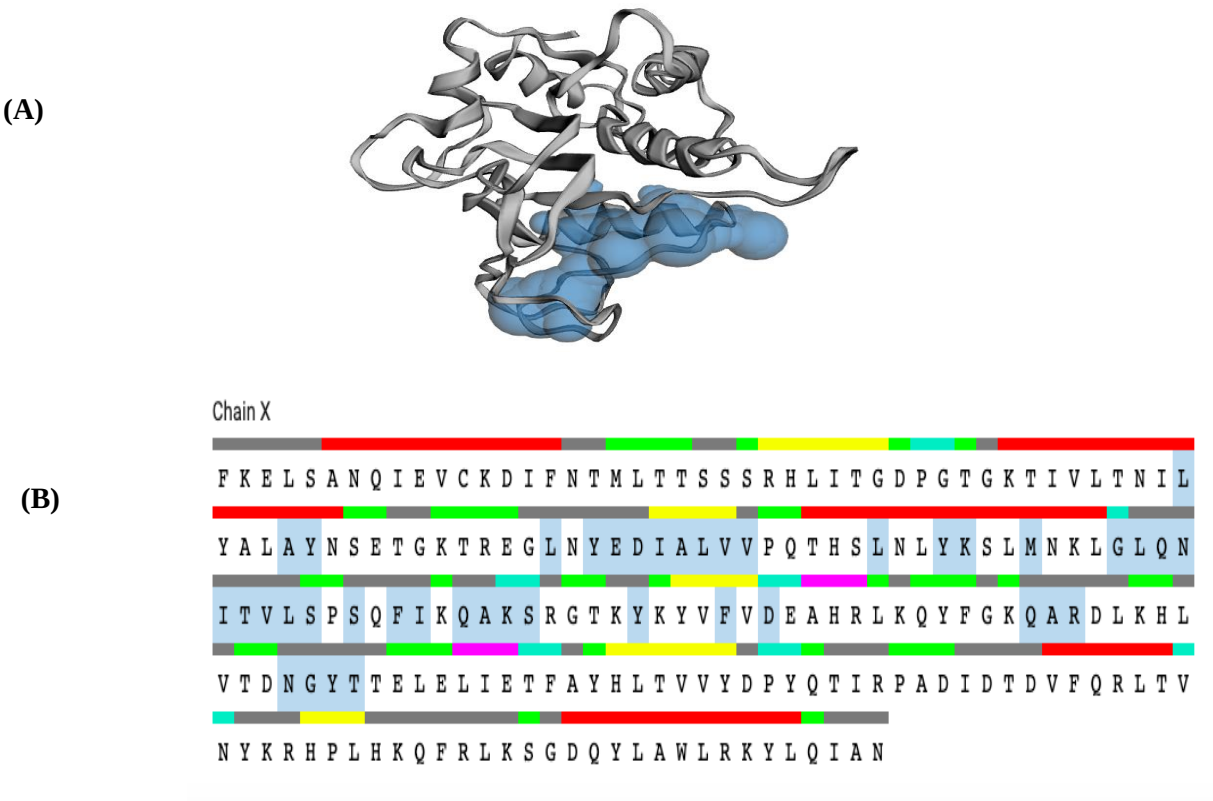

Fig. 4. Active site prediction of the uncharacterized protein BCRIVMBC845_02328; (A) Blue color indicates the active sites of the protein, (B) Amino acid residues in the active sites of the protein.

4. Discussion

The amino acid (aa) FASTA format sequence of the hypothetical protein BCRIVMBC845_02328 of *Bacillus cereus* was retrieved from the NCBI and employed as a query sequence to ascertain its physico-chemical properties where the theoretical isoelectric point (8.16) indicates the basic nature of the protein (**Table 2**). In addition, the protein had an instability index of 32.71 (40), which indicated its stability. The amino acid content (**Table 3**) and secondary structure components (**Table 5**) reveal the characteristics of the protein BCRIVMBC845_02328. The protein's subcellular location analysis report included information on its function in which the protein was predicted by CELLO v.3.0 to be located in the cytoplasm (**Table 4**). The GIY-YIG nuclease domain is a conserved domain identified by the CD Search tool as being present in a protein that contains the DUF2075 domain but whose function is unclear. By Ramachandran plot analysis, PROCHECK, and other methods, the modeled tertiary structures generated from the Modeller tool were confirmed (**Table 6**). 87.6% of amino acid residues were discovered in the locations that the Modeller preferred. The forbidden sections had 3 residues, however the extra and generously permitted regions contained 6.2% and 4.7% of residues, respectively.

5. Conclusion

Characterization of a protein using sophisticated bioinformatics tools is another novel task as like as other systems biology works. In our study we aimed at to reveal the physicochemical characteristics, structures and functions of a hypothetical protein BCRIVMBC845_02328 of *Bacillus cereus*. The 560 amino acid containing protein contains more positively charged residues and a high aliphatic index value and a low instability index value which make the protein more temperature stable. The secondary structure modelled by several bioinformatics tools ensured that

the proteins had domain in it which indicated it was a functional protein and tertiary structure prediction showed the protein had a fine 3D structure validated by various servers. Moreover, active site was found in the protein where ligand could bind. Further study of the protein is needed to find novel therapeutic drug for the gastrointestinal illness treatment through targeting the protein.

6. Acknowledgement
We are grateful to the Khwaja Yunus Ali University, Enayetpur, Sirajganj, Bangladesh authority as this study was sponsored by them.

7. Conflict of Interests
The authors declare no conflicts of interests.

8. Ethical Statement
Not Applicable.

9. Funding
This study was funded by the Research Grant Committee, Khwaja Yunus Ali University, Sirajganj, Bangladesh.

10. Authors Contributions
Research concept- MFR; Research design-MFR, JM, AY, MSB; Supervision-JM, MSB; Data analysis and Interpretation- MFR, RH, JHS, MKP, MSR; Writing article- MFR, JM; Article editing- MFR, AY, MSR; Final approval- All authors.

11. References

1. Arnold, K., Bordoli, L., Kopp, J., & Schwede, T. (2006). The SWISS-MODEL workspace: a web-based environment for protein structure homology modelling. *Bioinformatics*, *22*(2), 195-201. https://doi.org/10.1093/bioinformatics/bti770.
2. Bottone, E. J. (2010). Bacillus cereus, a volatile human pathogen. *Clinical microbiology reviews*, *23*(2), 382-398.
3. Carroll, L. M., Wiedmann, M., Mukherjee, M., Nicholas, D. C., Mingle, L. A., Dumas, N. B.,m& Kovac, J. (2019). Characterization of emetic and diarrheal Bacillus cereus strains from a 2016 foodborne outbreak using whole-genome sequencing: addressing the microbiological, epidemiological, and bioinformatic challenges. *Frontiers in microbiology*, 144. https://doi.org/10.3389/fmicb.2019.00144.
4. Doll, V. M., Ehling-Schulz, M., & Vogelmann, R. (2013). Concerted action of sphingomyelinase and non-hemolytic enterotoxin in pathogenic Bacillus cereus. *PLoS One*, *8*(4), e61404. https:/doi.org/10.1371/journal.pone.0061404
5. F de Azevedo, W. (2011). Molecular dynamics simulations of protein targets identified in *Mycobacterium tuberculosis. Current medicinal chemistry*, *18*(9), 1353-1366. https://doi.org/10.2174/092986711795029519.
6. Finn, R. D., Bateman, A., Clements, J., Coggill, P., Eberhardt, R. Y., Eddy, S. R., & Punta, M. (2014). Pfam: the protein families database. *Nucleic acids research*, *42*(D1), D222-D230. https://doi.org/10.1093/nar/gkt1223.
7. Frankland, G. C., & Frankland, P. F. (1887). XI. Studies on some new micro-organisms obtained from air. *Philosophical Transactions of the Royal Society of London (B.)*, (178), 257-287. https://doi.org/10.1098/rstb.1887.0011.
8. Gasteiger, E., Hoogland, C., Gattiker, A., Wilkins, M. R., Appel, R. D., & Bairoch, A. (2005). Protein identification and analysis tools on the ExPASy server. *The proteomics protocols handbook*, 571-607. doi.org/10.1385/1-59259-890-0:571.
9. Geourjon, C., & Deleage, G. (1995). SOPMA: significant improvements in protein secondary structure prediction by consensus prediction from multiple alignments. *Bioinformatics*, *11*(6), 681-684. https://doi.org/10.1093/bioinformatics/11.6.681.
10. Glasset, B., Herbin, S., Granier, S. A., Cavalié, L., Lafeuille, E., Guérin, C., & Ramarao, N. (2018). Bacillus cereus, a serious cause of nosocomial infections: epidemiologic and genetic survey. *PloS one*, *13*(5), e0194346.
11. Gong, J., Chen, Y., Pu, F., Sun, P., He, F., Zhang, L.,& Wang, H. (2019). Understanding membrane protein drug targets in computational perspective. *Current Drug Targets*, *20*(5), 551-564. https://doi.org/10.2174/1389450120666181204164721.
12. Granum, P., Arnesen, L. S., & Fagerlund, A. (2008). From Soil to Gut: *Bacillus cereus* and its food poisoning toxins. *Fems Microbiology Reviews*, *32*(4), 579-606.
13. Jones, D. T. (1999). Protein secondary structure prediction based on position-specific scoring

matrices. *Journal of molecular biology*, *292*(2), 195-202. https://doi.org/10.1006/jmbi.1999.3091

14. Kelley, L. A., Mezulis, S., Yates, C. M., Wass, M. N., & Sternberg, M. J. (2015). The Phyre2 web portal for protein modeling, prediction and analysis. *Nature protocols*, *10*(6), 845-858. https://doi.org/10.1038/nprot.2015.05
15. Laskowski, R. A., MacArthur, M. W., Moss, D. S., & Thornton, J. M. (1993). PROCHECK: a program to check the stereochemical quality of protein structures. *Journal of applied crystallography*, *26*(2), 283-291. https://doi.org/10.1107/S0021889892009944
16. Lindbäck, T., Fagerlund, A., Rødland, M. S., & Granum, P. E. (2004). Characterization of the Bacillus cereus Nhe enterotoxin. *Microbiology*, *150*(12), 3959-3967.https://doi.org/10.1099/mic.0.27359-0
17. Lu, S., Wang, J., Chitsaz, F., Derbyshire, M. K., Geer, R. C., Gonzales, N. R., & Marchler-Bauer, A. (2020). CDD/SPARCLE: the conserved domain database in 2020. *Nucleic acids research*, *48*(D1), D265-D268. https://doi.org/10.1093/nar/gkz991
18. McDowell, R. H., Sands, E. M., & Friedman, H. (2021). Bacillus cereus. In *StatPearls*. StatPearls Publishing. https://www.ncbi.nlm.nih.gov/books/NBK459121/.
19. Mills, C. L., Beuning, P. J., & Ondrechen, M. J. (2015). Biochemical functional predictions for protein structures of unknown or uncertain function. *Computational and Structural Biotechnology Journal*, *13*, 182-191. https://doi.org/10.1016/j.csbj.2015.02.003
20. Ngamwongsatit, P., Buasri, W., Pianariyanon, P., Pulsrikarn, C., Ohba, M., Assavanig, A., & Panbangred, W. (2008). Broad distribution of enterotoxin genes (hblCDA, nheABC, cytK, and entFM) among Bacillus thuringiensis and Bacillus cereus as shown by novel primers. *International journal of food microbiology*, *121*(3), 352-356. https://doi.org/10.1016/j.ijfoodmicro.2007.11.013
21. Paul, S. I., Rahman, M. M., Salam, M. A., Khan, M. A. R., & Islam, M. T. (2021). Identification of marine sponge-associated bacteria of the Saint Martin's Island of the Bay of Bengal emphasizing on the prevention of motile Aeromonas septicemia in Labeo rohita. *Aquaculture*, *545*, 737156. https://doi.org/10.1016/j.aquaculture.2021.737156.
22. Saikat, A. S. M., Islam, R., Mahmud, S., Imran, M., Sayeed, A., Alam, M. S., & Uddin, M. (2020). Structural and functional.
23. Schoeni, J. L., & Wong, A. C. (2005). *Bacillus cereus* food poisoning and its toxins. *Journal of food protection*, *68*(3), 636-648. https://doi.org/10.4315/0362-028X-68.3.636.
24. Stothard, P. (2000). The sequence manipulation suite: JavaScript programs for analyzing and formatting protein and DNA sequences. *Biotechniques*, *28*(6), 1102-1104. https://doi.org/10.2144/00286ir01.
25. Webb, B., & Sali, A. (2016). Comparative protein structure modeling using MODELLER. *Current protocols in bioinformatics*, *54*(1), 5-6. https://doi.org/10.1002/cpps.20.
26. Yu, C. S., Chen, Y. C., Lu, C. H., & Hwang, J. K. (2006). Prediction of protein subcellular localization. *Proteins: Structure, Function and Bioinformatics*, *64*(3), 643-651. https://doi.org/10.1002/prot.21018.
27. Zimmermann, L., Stephens, A., Nam, S. Z., Rau, D., Kübler, J., Lozajic, M., & Alva, V. (2018). A completely reimplemented MPI bioinformatics toolkit with a new HHpred server at its core. *Journal of molecular biology*,*430*(15), 2237-2243. https://doi.org/10.1016/j.jmb.2017.12.007.

Citation: Rahman, M.F., Mahato, J., Yeasmin,A., Biswas, M.S., Hasan, R., Shathi, J.H., Podder, M.K.,& Rahman, M.S.(2022). Structure prediction and functional annotation of an uncharacterized protein bcrivmbc845_02328 of *Bacillus cereus*: An *in-silico* approach, *Khwaja Yunus Ali Uni. J.5(2),136-144*