IOCD 2025 Conference

The 1st International Online Conference on Diseases

09–11 September 2025 | Online

MDPI

Computational approaches for structure-based functional annotation of an uncharacterized protein (Q182S9) of Clostridium difficile

Md. Tahosin Hossain, Abu Saim Mohammad Saikat

Gopalganj Science and Technology University, Gopalganj 8100, Bangladesh

INTRODUCTION & AIM

Clostridium difficile, sometimes known as C. Difficile, is a Gram-positive, spore-forming, anaerobic, bacillus that is extensively found in the environment and in both human and animal digestive tracts.C. Difficile infections have increased in frequency and severity over the last ten years,making it one of the most widespread nosocomial infections globally.Severe infectious colitis brought on by Clostridium difficile causes a great deal of morbidity and mortality throughout the globe.Two important virulence factors produced by C. Difficile during infection, which are toxin A and toxin B.Experiments utilizing pure toxins have shown that Toxin A alone can produce symptoms of a C. Difficile infection. On the other hand, toxin B cannot cause symptoms unless it is combined with toxin A or there has already been damage to the gut mucosa.

OBJECTIVES

1. The aim of this project is future medicines and immunizations that can target this protein to prevent microbial infections and the protein's ability to infect cells and prevent further morbidity and mortality.

2. To investigate the pathogenic mechanisms of Clostridium difficile, with particular emphasis on the roles of toxins A and B, spore germination, and antimicrobial resistance in disease progression.

3. To assess the clinical and epidemiological impact of C. difficile infections (CDI), including risk factors such as antibiotic use, in order to develop strategies for prevention and management.

METHOD

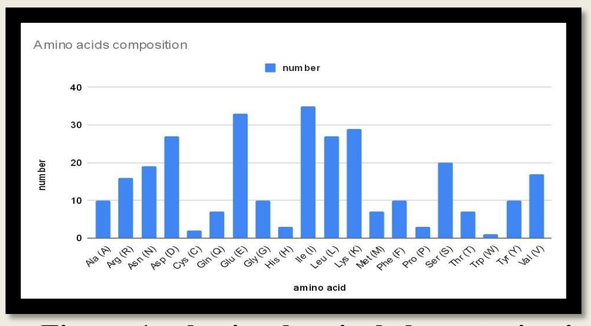
1.Sequence retrieval
The protein sequence of Clostridium difficile was collected in FASTA format from uniport protein database(accession date, september 7,2024).

2. physicochemical characterization:
The selected protein's physicochemical parameters were determined using the PortParam database.

3. Functional annotation
The conserved domain was identified using the NCBI's CD search tool.

4. structure prediction and validation
Secondary structural components were identified using the SOPMA server.
Moreover, the secondary structure of the chosen protein was predicted and modeled by the PSIPRED tool/server.
The AlphaFold and SWISS-MODEL services were used to predict the tertiary structure of the selected protein. ProSa-web servers and UCLA Saves software later verified the structure.

RESULTS & DISCUSSION



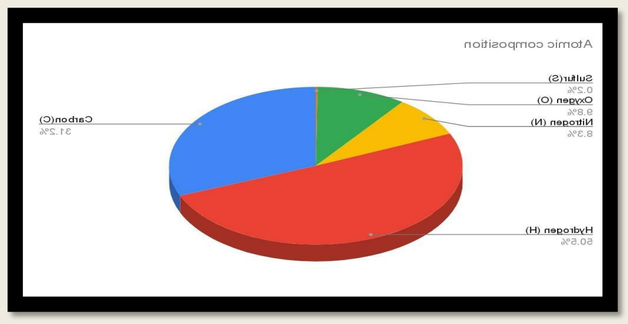
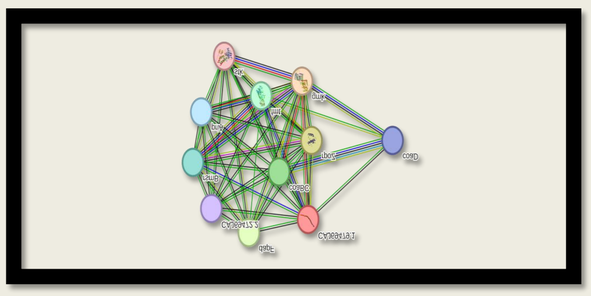
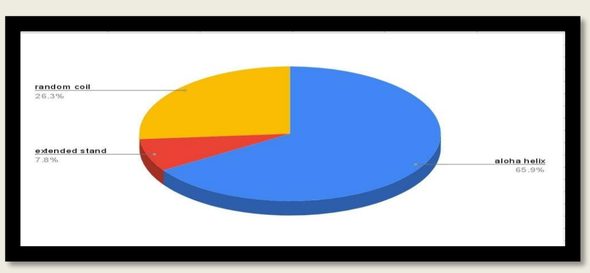


Figure-1: physicochemical characterization by portParam database



Accession	Name	Ontology
GO:0004521	RNA endonuclease activity	Molecular function
GO:0030435	Sporulation resulting in formation of a cellular spore	Biological process

Figure -2: protein's PPI network and the conserved domain were identified by The string (version 12.0) and NCBI's CD search tool.



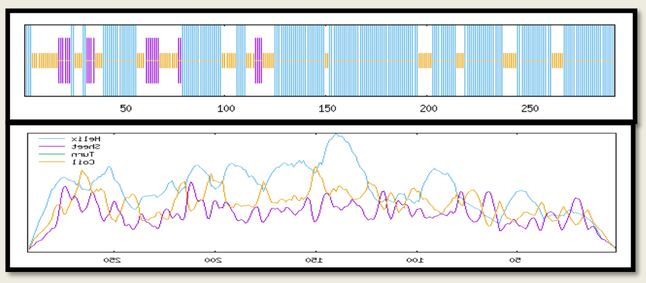


Figure-3: Secondary structural components were identified using the SOPMA server and the secondary structure of the chosen protein was predicted and modeled by the PSIPRED tool/server

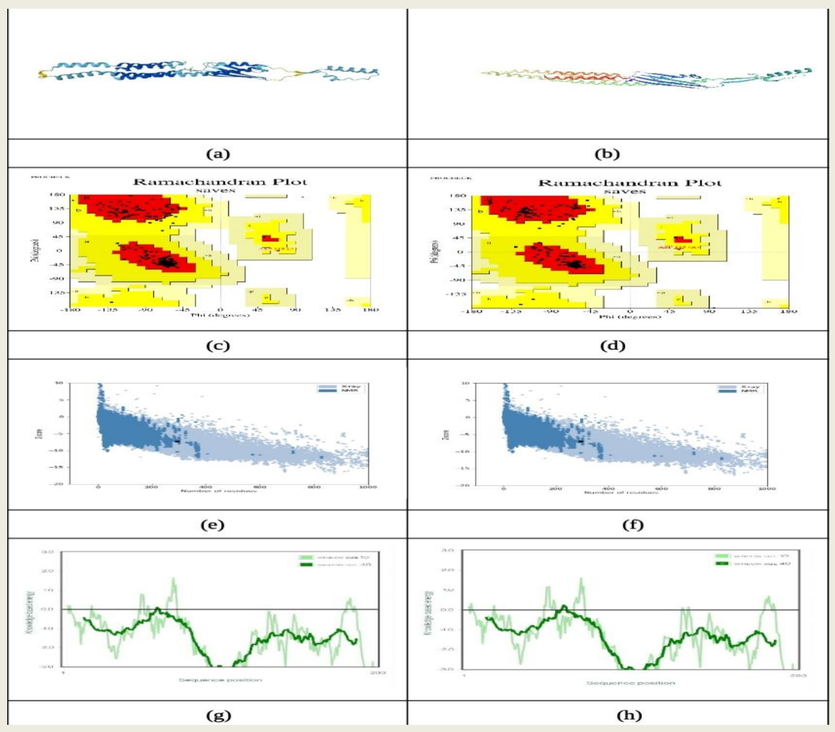


Figure-4: shows the predicted tertiary structure of the selected protein using (a)AlphaFold and (b) SWISS-MODEL. Ramachandran Plot of the chosen protein predicted by (c) AlphaFold and (d) SWISS-MODEL. The Z-scores of the modeled tertiary proteins were obtained from the ProSA-web for (e)AlphaFold (Z-Score: -7.28) and (f) SWISS-MODEL (Z-Score: -7.04). The local model quality analysis for (g) AlphaFold and (h) SWISS-MODE shows Z-Scores of -7.28 and -7.04, respectively.

Characteristics	AlphaFold	Swiss-Model
Residues in most favored regions	266 (95.7%)	267 (96.0%)
Residues in additional allowed regions	11(4.0%)	10 (3.6%)
Residues in generously allowed regions	1 (0.4%)	1 (0.4%)
Residues in disallowed regions	0 (0.0%)	0 (0.0%)
Number of glycine residues	10	10
Number of proline residues	3	3

Figure-5:Comparison Between AlphaFold and SWISS-MODEL based on Ramachandran plot study

CONCLUSION

This study analyzed an uncharacterized Clostridium difficile protein and found it to be hydrophilic, thermostable, and mainly alpha-helical in structure. Domain analysis suggested possible roles in RNA degradation and glucose transport. Among the predicted structures, SWISS-MODEL was the most accurate. These findings provide useful insights for developing future therapies and precaution against C. difficile.

REFERENCES

[1] J. Czepiel *et al.*, "Clostridium difficile infection: review," *Eur J Clin Microbiol Infect Dis*, vol. 38, no. 7, pp. 1211–1221, Jul. 2019, doi: 10.1007/S10096-019-03539-6.
[2] K. E. Burke and J. T. Lamont, "Clostridium difficile infection: a worldwide disease," *Gut Liver*, vol. 8, no. 1, pp. 1–6, Jan. 2014, doi: 10.5009/GNL.2014.8.1.1.
[3] H. Q. Qu and Z. D. Jiang, "Clostridium difficile infection in diabetes," *Diabetes Res Clin Pract*, vol. 105, no. 3, pp. 285–294, 2014, doi: 10.1016/J.DIABRES.2014.06.002.
[4] R. Shrestha and J. A. Sorg, "Terbium chloride influences Clostridium difficile spore germination," *Anaerobe*, vol. 58, pp. 80–88, Aug. 2019, doi: 10.1016/J.ANAEROBE.2019.03.016.
[5] K. Imwattana *et al.*, "Antimicrobial resistance in Clostridium difficile ribotype 017," *Expert Rev Anti Infect Ther*, vol. 18, no. 1, pp. 17–25, Jan. 2020, doi: 10.1080/14787210.2020.1701436.