Cole-Puli International Conference Publishing Group

Science, Engineering and Technology Proceedings

Article

2025 International Conference on Natural Sciences, Agricultural Economics, Biomedicine and Sustainable Development (AEBSD 2025)

Analyzing Protein Structures and Functions in Tube A & B from Borduria

Xinjue Zhou [1,*]

[1] University of Melbourne, Victoria, Australia
* Correspondence: Xinjue Zhou, University of Melbourne, Victoria, Australia

Abstract: During a routine retrieval operation at a remote research station, a field technician collected two unidentified laboratory sample tubes, labeled A and B, for subsequent scientific evaluation. These samples were believed to originate from an ongoing experimental project whose details had not been fully documented, and their analysis was expected to clarify the project's technical focus and research direction. While the technician was returning with the materials, Tube B was accidentally damaged, resulting in the loss of part of its contents and a brief exposure incident. The technician later developed acute symptoms and was unable to provide further information about the circumstances surrounding the spill. This event increased the urgency of examining the remaining Tube A to determine the nature, stability, and potential laboratory relevance of the preserved sample. The incident also underscored the importance of traceable sample management, standardized handling procedures, and careful cross-disciplinary analysis when dealing with unknown research materials. The following study provides a systematic evaluation of the specimen in Tube A, aiming to identify structural features, possible biological origins, and any functional properties that may be relevant to ongoing scientific work.

Keywords: protein structure analysis; biochemical exposure; hazardous compounds; experimental sample characterization; intelligence investigation; cross-disciplinary laboratory evaluation

Received: 07 November 2025
Revised: 21 November 2025
Accepted: 18 December 2025
Published: 20 December 2025

1. Introduction

During a routine retrieval task conducted at a remote research facility, a field technician collected two experimental sample tubes, labeled A and B, for subsequent laboratory assessment. Although the specific nature of the samples was not clearly documented at the time, the materials were believed to contain information relevant to an ongoing scientific project. In the period leading up to this incident, several research programs in the region had been undergoing rapid development, prompting growing interest in understanding their technical orientation and potential applications. Within this broader context, the acquisition of the samples was expected to provide early insights into the experimental directions and methodological advances underlying the associated project.

According to internal records, the technician encountered an unexpected accident while returning from the facility. During this event, Tube B was damaged, resulting in the loss of part of its contents before they could be properly secured. This spill created significant difficulties in reconstructing the original properties of the sample. Shortly after the incident, the technician experienced severe injuries and was unable to provide further details. This development increased the urgency of safeguarding Tube A for immediate

laboratory examination and raised concerns about the need for improved procedures for handling and transporting sensitive research materials.

Following the technician's death, the organization conducted a systematic review of the circumstances under which the samples were collected. The aim was to clarify whether the accident was isolated or whether it reflected broader challenges associated with field retrieval tasks. At the same time, scientific teams initiated controlled assessments of Tube A using standardized containment procedures to prevent additional material loss and to ensure the safety of personnel working with the sample. Initial analyses focused on basic physical properties, stability under different conditions, and compliance with established laboratory safety requirements. Even preliminary findings were regarded as valuable for understanding the sequence of events surrounding the mishandled sample.

The incident also prompted broader internal discussions about the importance of rigorous research documentation, transparent sample tracking, and consistent safety protocols. Although the organization had traditionally concentrated on routine project monitoring, the appearance of an undocumented experimental material highlighted the need for strengthened management practices. This case demonstrated that laboratory products, developmental materials, and intermediate research samples may have implications that extend beyond their immediate scientific context. Consequently, defining the purpose, composition, and potential applications of Tube A became essential—not only for reconstructing the conditions under which the accident occurred, but also for improving oversight mechanisms and ensuring responsible research practices.

By placing the loss of Tube B and the preservation of Tube A within this broader analytical perspective, this chapter establishes the foundation for a more detailed examination of the remaining sample. It also clarifies why Tube A continues to hold scientific value, despite the challenges created by the accidental spill. The discussion that follows builds on this introductory framework and aims to interpret the technical, procedural, and contextual aspects surrounding the retrieval of the samples while maintaining a focus on safety, neutrality, and academic rigor.

2. Method

2.1. Basic Local Alignment Search Tool (BLAST)

BLAST is a computational method used to compare a query nucleotide or protein sequence against a reference database in order to determine potential matches and infer the biological identity of the detected sequence [1]. In practical applications, BLAST assists in identifying related genes, homologous proteins, or conserved functional regions by evaluating the similarity between the query and known entries. Two key parameters-score and E value-play a central role in interpreting the results. The score represents the cumulative value derived from substitution matches and gap penalties, with higher scores generally indicating a closer similarity between the sequences being compared. The E value, or Expectation Value, reflects the number of alignments that are statistically expected to achieve a similar or better score by chance. A lower E value suggests a more reliable similarity, as it indicates that the observed alignment is unlikely to be random [1]. These parameters jointly guide the evaluation of sequence relevance and help determine whether a detected sequence may share structural or functional attributes with known biological molecules.

2.2. Position-specific Iterative Basic Local Alignment (PSI-BLAST)

PSI-BLAST is an extended version of BLAST designed for deeper protein sequence analysis, enabling the detection of remote homologs through iterative refinement of a position-specific scoring matrix [2]. During each iteration, PSI-BLAST evaluates the sequences that meet the predefined E value threshold and incorporates them into an evolving profile. This process continues until no new sequences above the threshold are identified, thereby improving sensitivity in detecting conserved motifs or weakly similar

protein families. By progressively updating the scoring matrix, PSI-BLAST allows the analysis to capture subtle evolutionary signals that standard BLAST searches may miss. As a result, it is particularly useful for exploring protein domains, classifying sequences into broader families, and revealing structural or functional relationships that are not immediately apparent from direct pairwise comparison.

2.3. Protein Data Bank (PDB)

The Protein Data Bank (PDB) provides comprehensive information on experimentally determined protein structures, including 3D conformation, sequence details, and associated biochemical data [3]. In this study, PDB was used to visualize the three-dimensional architecture of the detected protein, allowing a clearer understanding of its spatial organization, potential binding regions, and structural motifs. Through the structural models available in PDB, it becomes possible to assess how certain regions of the protein may contribute to stability, catalytic function, or molecular interactions. Such structural interpretation complements sequence-based analysis and supports more accurate functional prediction.

PROSITE is a database designed to identify protein families, conserved domains, and functional sites by comparing the detected peptide sequence to documented structural and functional motifs [4]. It was applied in this work to infer potential biological functions of the protein by examining characteristic patterns, such as active sites, binding residues, or signature motifs. By matching these patterns to established domain profiles, PROSITE enables a more informed interpretation of the protein's functional properties and assists in linking sequence features to biological activity.

3. Result

3.1. Tube A

In Tube A, sequence analysis revealed four distinct protein fragments. Their best-aligned matches and corresponding species are summarized in Table 1. These results offer an initial overview of the possible biological origins of each detected protein and provide the basis for subsequent structural and functional assessments.

Table 1. The best matching protein and the original species of Protein1-4.

Name	Best match protein	Species	Sequence ID
Protein 1	glycoprotein 1, partial	Mammarenavirus machupoense	AAS77647.1
Protein 2	tetanus toxin precursor, partial	Clostridium sp. RKD	CAJ28911.1
Protein 3	Chain A, MATRIX PROTEIN VP40	Ebola virus sp.	1H2C_A
Protein 4	Chain X, DNA topoisomerase 1	Variola virus	2H7G_X

Protein 4 was selected as the primary focus for further investigation because its alignment results exhibited high identity values and consistent matches across multiple related sequences. As shown in Table 2, BLAST analysis identified the five most closely related proteins, all of which share substantial sequence similarity with Protein 4.

Table 2. The 5 proteins most closely related to Protein 4 in the results of the BLAST research.

Name of protein	No. identical amino acid	No. similar amino acid	%identity	%similarity	Score	E value	Sequence ID
Chain X, DNA topoisomerase 1 [Variola virus]	49	49	100%	100%	108 bits	2e-26	2H7G_X
Chain X, DNA topoisomerase 1 [Variola virus]	49	49	100%	100%	107 bits	2e-26	2H7F_X
DNA topoisomerase type I [Variola virus]	49	49	100%	100%	107 bits	2e-26	NP_042133.1
CPXV115 protein [Cowpox virus]	47	48	96%	97%	102 bits	3e-26	ATB55156.1
Chain A, DNA TOPOISOMERASE I [Vaccinia virus WR]	47	48	96%	97%	100 bits	3e-26	1VCC_A

As indicated in Table 2, the top two matches correspond to Chain X of DNA topoisomerase 1, each presenting an E value of 2e-26. The first entry exhibited a score of 108 bits, which is slightly higher than the 107 bits of the second entry. Given these metrics, the most credible inference is that Protein 4 belongs to DNA topoisomerase 1 and corresponds to the sequence identified as 2H7G. This conclusion is supported by the complete identity observed in the aligned amino acids and the extremely low E value, suggesting that the detected fragment is highly unlikely to arise from unrelated proteins.

By entering the inferred sequence ID into the PDB database, the three-dimensional structure of the corresponding noncovalent topoisomerase-DNA complex was obtained, as shown in Figure 1. This structural visualization provides clearer insight into the spatial configuration of the protein, including the arrangement of α-helices, β-sheets, and the bound DNA double helix. The structural information also assists in understanding functional aspects of the protein, such as DNA binding regions and catalytic sites.

Figure 1. 3D view of noncovalent binding topoisomerase-DNA complex (Accession No. 2H7G). The purple part is DNA double helix. The red part is the α helices structure of the protein. The yellow part is the β sheets structure of the protein.

3.2. Tube B

Sequence analysis of Tube B was conducted using PSI-BLAST with three iterations. As shown in Table 3 and Table 4, each iteration produced alignments with statistically significant E values, and results gradually converged toward a consistent protein family.

Table 3. The sequence that significant alignment with E-values better than threshold in PSI-BLAST iteration 1.

Name	Genus/species	Identities	Similarity	Score	E value
Chain A, Alpha-conotoxin Vc1A	Synthetic construct	16/16 (100%)	16/16 (100%)	60.4 bits	5e-10
Chain A, Alpha-conotoxin Vc1A	Conus victoriae	16/16 (100%)	16/16 (100%)	60.4 bits	9e-10
RecName: Full = Alpha-conotoxin Vc1a; Short = Alpha-Vc1a; AltName: Full = ACV1; AltName: Full = Vc1.1; Flags: Precursor	Conus victoriae	17/17 (100%)	17/17 (100%)	63.8 bits	9e-10
Chain A, Alpha-conotoxin Vc1A	Conus victoriae	14/15 (93%)	14/15 (93%)	51.5 bits	2e-6
Chain A, Alpha-conotoxin Vc1a	Conus victoriae	14/16 (88%)	14/16 (87%)	51.1 bits	2e-6
Chain A, Alpha-conotoxin Vc1A	Conus victoriae	14/16 (88%)	14/16 (87%)	51.1 bits	2e-6
Chain A, Alpha-conotoxin Vc1A	Conus victoriae	14/16 (88%)	14/16 (87%)	51.5 bits	3e-6
RecName: Full = Alpha-conotoxin LvIA; Short = Alpha-CTx LvIA; Flags: Precursor	Conus lividus	14/17 (82%)	14/17 (82%)	45.2 bits	0.002
Chain X, Alpha-conotoxin Vc1A	Conus victoriae	13/14 (93%)	13/14 (92%)	43.9 bits	0.002

Table 4. The top 5 sequences that significant alignment with E-values better than threshold in PSI-BLAST iteration 3.

Name	Genus/species	Identities	Similarity	Score	E value
RecName: Full = Alpha-conotoxin LvIA; Short = Alpha-CTx LvIA; Flags: Precursor	Conus lividus	14/18 (78%)	14/18 (77%)	52.5 bits	5e-6
RecName: Full = Alpha-conotoxin Vc1a; Short = Alpha-Vc1a; AltName: Full = ACV1; AltName: Full = Vc1.1; Flags: Precursor	Conus victoriae	17/18 (94%)	17/18 (94%)	51.2 bits	4e-5
alpha-conopeptide precursor Bt1.4	Conus betulinus	11/19 (58%)	12/19 (63%)	50.8 bits	5e-5
conotoxin	Conus betulinus	11/19 (58%)	12/19 (63%)	50.8 bits	6e-5
Chain A, Alpha-conotoxin-like	Conus lividus	13/16 (81%)	13/16 (81%)	47.8 bits	6e-5

During the iterative process, noticeable differences were observed between the sets of alignments obtained at each stage. In the third iteration, none of the top five sequences reached 100% identity; however, the highest identity recorded was 94% for alpha-

conotoxin Vc1a from *Conus victoriae*. Although this sequence did not yield the highest score or lowest E value, its high identity suggests a strong resemblance. Meanwhile, alpha-conotoxin LvIA from *Conus lividus* achieved a better overall score and lower E value, indicating that it may represent a more statistically robust match. Some other sequences achieved 100% identity but exhibited lower scores and higher E values, which prevented them from ranking within the top five [5]. This variation reflects how PSI-BLAST prioritizes statistical confidence in addition to identity alone.

To complement the iterative alignment, the detected sequence was further analyzed using PROSITE. The scan identified the motif "CCSDPRCNYDHPEIC" as a signature pattern characteristic of the alpha-conotoxin family, in which all four cysteine residues participate in disulfide bond formation [6]. This structural pattern is essential for the functional stability and biological activity commonly associated with this peptide group. The PROSITE result corresponds well with the family classification indicated by PSI-BLAST, thereby reinforcing the conclusion.

Taken together, the results from PSI-BLAST and PROSITE indicate that the substance contained in Tube B can be reliably classified as a member of the alpha-conotoxin family. The consistent detection of characteristic motifs and the convergence of sequence alignments across iterations support this determination [7]. This classification also suggests that the protein likely possesses biochemical properties typical of short, disulfide-rich peptides with defined structural loops and potential biological activity.

4. Discussion

4.1. Tube A

According to the previous analysis, the best alignment DNA topoisomerase 1 comes from variola virus, a highly transmissible virus that causes a deadly disease smallpox with a 20-30% mortality rate. DNA topoisomerase is an important type of enzyme for viruses. It functions for reducing the topological stress during some critical events such as transcription and replication when the double helix is underwinding or overwinding [8]. A specific target sequence, 5'-CCCTT-3', is required to start a sequence-specific recognition between the enzyme and the DNA double helix [9]. After the event is complete, the enzyme leaving group sugars undergo a conformational change that promotes DNA ligation [10]. Considering the importance of this enzyme for viral survival, topoisomerase-DNA complexes are considered as targets for disease treatment [11]. Asp168 is the critical regulator of the active site to balance among DNA cleavage, religation, and product release, providing a guide to antiviral drug design [12].

The four proteins in tube A are known to derive from different viruses. At the same time, it has been informed of the presence of immunostimulatory complexes (ISCOM) in Tube A. ISCOMs are particles with multiple copies of antigen and are constructed for antigen presentation in immune system, so it is often used in the manufacture of vaccines [13]. Based on the above information, it can be inferred that ISCOM is present to deliver these proteins into the cell to induce viral infection or immune response.

Given that the four proteins come from four different pathogens, the results of the analysis support a reasonable suspicion that they are trying to develop complex viruses. But given the role of ISCOM in vaccine development, it is also possible that they are working on vaccine development to target multi-viral infections.

4.2. Tube B

Analysis showed that the protein in tube B belonged to alpha conotoxin (α-CTx), a small peptide with a length of 12 to 20 amino acid residues.[13] Studies have shown that this family can selectively block the nicotinic receptors, which leads to paralysis or even involves the diaphragm [14].

The previously known information included that the source of tube B was a laboratory that housed many exotic sea creatures. Before her death, the spy had

experienced numb, stiff lips, blurred vision, and paralysis of her legs. These known conditions are consistent with α-CTx poisoning. Her autopsy showed that she had sample powder on her injured left hand, suggesting that she had been poisoned from the wound. Post-mortem tests showed that her organs were healthy, so she should have died of asphyxiation caused by diaphragm paralysis.

Since the alignments only aligned part of the sequence to be detected, and none of the top five alignments reached 100% identity, it is assumed that α-CTx in tube B was lab-synthesized rather than natural.

5. Conclusion

In summary, the analyses conducted on the samples from Tube A and Tube B indicate that the research facility involved in this investigation is engaged in studies related to viral proteins and bioactive peptide components. The identification of several viral protein fragments in Tube A, together with the characterization of an alpha-conotoxin family peptide in Tube B, suggests that the laboratory is working with materials that possess notable biochemical activity. Considering the known physiological effects of alpha conotoxins, including their potential to disrupt normal neuromuscular transmission, it is plausible that exposure to the leaked contents of Tube B could have caused acute respiratory difficulty leading to the spy's death. Although this possibility cannot be confirmed solely from sequence data, it remains consistent with the biochemical nature of the detected peptide.

Regarding the purpose of the research activities, two general possibilities may be considered from a technical standpoint. One possibility is that the facility is conducting studies on complex viral components and naturally occurring peptide toxins to explore their structural and functional characteristics. Another possibility is that these efforts are directed toward developing vaccines, inhibitors, or detoxification agents aimed at reducing the health risks associated with such biological substances. Given the uncertainty surrounding the research intentions, it would be prudent to promote further investigation of the known protein components and to consider the early development of corresponding protective measures. Preparing preventive technologies, such as targeted antidotes or immunological tools, may help mitigate potential risks and improve the capacity to respond to future unforeseen situations.

References

1. J. Fassler, and P. Cooper, "BLAST glossary," *BLAST® Help*, 2011.
2. M. Bhagwat, and L. Aravind, "Psi-blast tutorial," In *Comparative genomics*, 2007, pp. 177-186. doi: 10.1385/1-59745-514-8:177
3. P. W. Rose, C. Bi, W. F. Bluhm, C. H. Christie, D. Dimitropoulos, S. Dutta, and P. E. Bourne, "The RCSB Protein Data Bank: new resources for research and education," *Nucleic acids research*, vol. 41, no. D1, pp. D475-D482, 2012. doi: 10.1093/nar/gks1200
4. N. Hulo, A. Bairoch, V. Bulliard, L. Cerutti, B. A. Cuche, E. De Castro, and C. J. Sigrist, "The 20 years of PROSITE," *Nucleic acids research*, vol. 36, no. suppl_1, pp. D245-D249, 2007.
5. R. J. Tuieng, S. H. Cartmell, C. C. Kirwan, A. Eckersley, and M. J. Sherratt, "X-Ray Exposure Induces Structural Changes in Human Breast Proteins," *International Journal of Molecular Sciences*, vol. 26, no. 12, p. 5696, 2025. doi: 10.3390/ijms26125696
6. E. Kamau, and A. Grove, "Fluoroquinolone-dependent DNA supercoiling by vaccinia topoisomerase I," *Journal of molecular biology*, vol. 342, no. 2, pp. 479-487, 2004. doi: 10.1016/j.jmb.2004.06.082
7. D. Sehnal, S. Bittrich, M. Deshpande, R. Svobodová, K. Berka, V. Bazgier, and A. S. Rose, "Mol* Viewer: modern web app for 3D visualization and analysis of large biomolecular structures," *Nucleic acids research*, vol. 49, no. W1, pp. W431-W437, 2021. doi: 10.1093/nar/gkab314
8. K. Perry, Y. Hwang, F. D. Bushman, and G. D. Van Duyne, "Structural basis for specificity in the poxvirus topoisomerase," *Molecular cell*, vol. 23, no. 3, pp. 343-354, 2006. doi: 10.1016/j.molcel.2006.06.015
9. B. G. Anderson, and J. T. Stivers, "Variola type IB DNA topoisomerase: DNA binding and supercoil unwinding using engineered DNA minicircles," *Biochemistry*, vol. 53, no. 26, pp. 4302-4315, 2014. doi: 10.1021/bi500571q
10. K. Perry, Y. Hwang, F. D. Bushman, and G. D. Van Duyne, "Insights from the structure of a smallpox virus topoisomerase-DNA transition state mimic," *Structure*, vol. 18, no. 1, pp. 127-137, 2010. doi: 10.1016/j.str.2009.10.020

11. Y. Hwang, N. Minkah, K. Perry, G. D. Van Duyne, and F. D. Bushman, "Regulation of catalysis by the smallpox virus topoisomerase," *Journal of Biological Chemistry*, vol. 281, no. 49, pp. 38052-38060, 2006. doi: 10.1074/jbc.m608858200
12. B. Morein, K. Lövgren, B. Rönnberg, A. Sjölander, and M. Villacrés-Eriksson, "Immunostimulating complexes: clinical potential in vaccine development," *Clinical Immunotherapeutics*, vol. 3, no. 6, pp. 461-475, 1995.
13. F. Rahimzadeh, L. M. Khanli, P. Salehpoor, F. Golabi, and S. PourBahrami, "Unveiling the evolution of policies for enhancing protein structure predictions: A comprehensive analysis," *Computers in Biology and Medicine*, vol. 179, p. 108815, 2024.
14. S. Kapil, S. Hendriksen, and J. S. Cooper, "Cone snail toxicity," 2017.