International Journal of Agricultural Technology 2026 Vol. 22(4):1633-1648
DOI: https://doi.org/10.63369/ijat.2026.22.4.1633-1648
Available online http://www.ijat-aatsea.com
ISSN 2630-0192 (Online)

# In silico structural analysis, classification, and functional annotation of uncharacterized protein from *Corbicula fluminea*

**Barangan, J. B. M.*** **and Jacob, J. K. S.**

Department of Biological Sciences, College of Arts and Sciences, Isabela State University – Echague, Isabela, Philippines 3309.

Barangan, J. B. M. and Jacob, J. K. S. (2026). In silico structural analysis, classification, and functional annotation of uncharacterized protein from *Corbicula fluminea*. International Journal of Agricultural Technology 22(4):1633-1648.

**Abstract** Heat shock protein 70 (HSP70) chaperones are essential molecular machines that facilitate the folding, assembly, and stabilization of numerous client proteins through ATP-dependent cycles of substrate binding and release. An uncharacterized protein sequence from *Corbicula fluminea* was initially annotated to contain domains with potential biotechnological relevance. This study employed in silico approaches using online tools for protein characterization including BLASTp, ExPASy, I-TASSER, EzMol, DeepLoc, SMART, and MEGA 12 software to determine its structural characteristics, classification, and potential functional role. Results revealed the presence of a conserved Pfam HSP70 domain (positions 6–612) and a Pfam MreB_Mbl domain (positions 116–384), along with a low-complexity region (positions 615–646). Subcellular localization prediction indicated a cytoplasmic distribution. Multiple sequence alignment and phylogenetic reconstruction using the Maximum Likelihood method showed that the HSP70 sequence from *C. fluminea* is closely related to known HSP70 sequences from related species, confirming its taxonomic placement within HSP70 protein family. Collectively, these results confirmed the protein as a member of HSP70 family and provide foundational insights into its structural and functional characteristics, contributing to the understanding of molecular chaperone biology in bivalves and informing future applications in biotechnology.

**Keywords:** *Corbicula fluminea*, HSP70, In silico analysis, Molecular chaperone, Protein annotation

## Introduction

The Asian clam (*Corbicula fluminea*) is an ecologically and economically significant freshwater bivalve that is widely distributed in Asia, Europe, and North America (Zhang *et al.*, 2021). *C. fluminea* serves as a crucial food resource and a bioindicator for assessing environmental health; however, it also poses ecological concerns due to its high adaptability and invasiveness—traits that can disrupt native biodiversity and alter ecological functions (Okawa *et al.*, 2016;

* **Corresponding Author:** Barangan, J. B. M..; **Email:** jinobryle.m.barangan@isu.edu.ph

Sidhu, 2023). As such, effective management and deeper biological perspectives on this species rely on molecular methods—particularly those that focus on genetic regulation and stress-response mechanisms (Mashjel and Al-Taher, 2025; Zhang *et al.*, 2021).

Heat shock proteins (HSPs) are a molecular system that underlies stress resilience in aquatic invertebrates, such as the HSP70 family of proteins. These stable proteins are essential for protecting cell integrity against environmental, chemical, and thermal stress, which involves roles in physical folding, repair, and protein degradation (Chen *et al.*, 2013). In *C. fluminea*, HSP70 expression is significantly activated in response to various abiotic conditions, including temperature changes, pollutant exposure, and osmotic pressure, making HSP70 genes reliable molecular biomarkers for monitoring different environmental parameters (Falfushynska *et al.*, 2016; Valenzuela-Castillo *et al.*, 2019). Empirical studies indicate that enhanced HSP70 expression not only confers improved stress tolerance but also supports the species and its invasiveness capacity by enabling rapid physiological adaptation to diverse and changing habitats (Sidhu, 2023).

Despite the pivotal role of HSP70 in environmental adaptation, most existing research focus on the gene family as a whole or examines generalized gene expression profiles, often overlooking the diversity and specificity of individual mRNA isoforms (Xie, 2017). A substantial gap remains in resolving structural, functional, and evolutionary nuances of HSP70 mRNA variants, including those specific to *Corbicula fluminea*. Advances in de novo transcriptome assembly and recent genomic analyses have emphasized the genetic complexity of many aquatic species and highlighted the potential for molecular diversity to influence phenotype and adaptability (Chen *et al.*, 2013; Zhang *et al.*, 2021). However, comprehensive *in silico* characterization of HSP70 isoforms—particularly those derived from *C. fluminea*—remains limited and underexplored.

Thus, this study aimed to characterize an unannotated protein sequence from *Corbicula fluminea* through *in silico* analysis, with specific focus on its physicochemical properties, conserved domains, functional classification, and predicted subcellular localization.

## Materials and methods

### *Time and place of the study*

This study was conducted at the Protein Chemistry and Biosensor Laboratory of Isabela State University, Echague, Isabela, Philippines from June 2025 to August 2025.

### *In silico structural analysis, classification, and functional annotation*

The study adapted several methodological components from Bitacura and Santos (2022), particularly those related in mRNA retrieval and translation, physicochemical characterization, domain classification and protein function identification, structure prediction, subcellular localization, and protein evolution analysis.

### *mRNA retrieval and confirmation of mRNA translation*

The mRNA FASTA sequence of *Corbicula fluminea* (Accession No. KJ461738.1) was submitted to the ExPASy Translate Tool and translated, selecting the compact output format (M, –) without spaces (Gasteiger *et al.*, 2005). The longest open reading frame (ORF) beginning with a start codon—deemed to be the most suitable one—was selected among the possible ORFs and aligned with a known annotated protein from *C. fluminea* (Accession No. AHY03302.1) using the Clustal Omega server to confirm proper translation and identity (Sievers and Higgins, 2017).

### *Physicochemical characterization*

To determine the physicochemical characteristics of the uncharacterized protein from *C. fluminea*, this study utilized the ProtParam tool available on the Expert Protein Analysis System (ExPASy) server of Gasteiger *et al.* (2005). Several physicochemical parameters were calculated such as molecular weight, theoretical isoelectric point (pI), instability index, aliphatic index, and the grand average of hydropathicity (GRAVY).

### *Domain architecture and protein classification*

The Simple Molecular Architecture Research Tool (SMART) was used to identify the conserved domains within the uncharacterized protein of *C. fluminea* (Letunic *et al.*, 2020). This tool enables the detection of genetic domains and annotation of domain structures across diverse protein families. Moreover, a comparison was made between the identified domain arrangements and known proteins to infer potential functional classifications.

### *Structure prediction*

To predict the three-dimensional (3D) structure of the uncharacterized *C. fluminea* protein, the study utilized Iterative Threading ASSEmbly Refinement

(I-TASSER) tool developed by Yang and Zhang (2015). This tool is a template-based structure prediction tool that constructs full-length atomic models through iterative simulations based on multiple structural templates. Subsequently, the generated PDB structure files were analyzed and visualized using online tool EzMol v2.1 (Reynolds *et al.*, 2018).

### *Subcellular localization*

The FASTA sequence of the uncharacterized protein from *C. fluminea* was submitted to the DeepLoc-1.0 server of Armenteros *et al.* (2017) to predict its subcellular localization. This online tool utilizes a Neural Network algorithm trained on annotated proteins with subcellular localization evidence from UniProt, and thus uses only the sequence information to infer localization.

### *Protein evolution analysis*

This study conducted ortholog analysis to assess the evolutionary relationship of the uncharacterized protein from *C. fluminea* and orthologs from other species within order Venerida. BLASTp (Johnson *et al.*, 2008) was used in searching protein orthologs. Thereafter, multiple sequence alignment was performed using ClustaW, and phylogenetic tree was constructed through the MEGA 12 software (Kumar *et al.*, 2024). This study employed the Maximum Parsimony approach, gaps and missing data were treated as all sites, and with 1,000 bootstraps replicates.

## Results

### *Translated mRNA*

Several open reading frames (ORFs) were identified in both the forward and reverse strands of the translated mRNA expressed by the uncharacterized protein, annotated as heat shock protein 70 (HSP70) from *C. lumina.* The longest ORF was located in the 5'–3' Frame 2 (green box, Figure 1B). An alignment of the amino acid residues of this ORF with the uncharacterized protein sequence confirmed complete residue identity (Figure 1A).

```
VIRT-1756796:5'3'      MSKAPAIGIDLGTTYSCVGVFQHGKVEIIANDQGNRTTPSYVAFTDTERLIGDAAKNQVA     60
AHY03302.1             MSKAPAIGIDLGTTYSCVGVFQHGKVEIIANDQGNRTTPSYVAFTDTERLIGDAAKNQVA     60
                       ************************************************************

VIRT-1756796:5'3'      MNPTNTIFDAKRLIGRKFDDASVQSDMKHWPFEVISEGGKPKLRVEYKGEQKTFFPEEIS     120
AHY03302.1             MNPTNTIFDAKRLIGRKFDDASVQSDMKHWPFEVISEGGKPKLRVEYKGEQKTFFPEEIS     120
                       ************************************************************

VIRT-1756796:5'3'      SMVLNKMKETAEAFLGKTVTNAVVTVPAYFNDSQRQATKDAGTISGLNVLRIINEPTAAA     180
AHY03302.1             SMVLNKMKETAEAFLGKTVTNAVVTVPAYFNDSQRQATKDAGTISGLNVLRIINEPTAAA     180
                       ************************************************************

VIRT-1756796:5'3'      IAYGLDKKVGGERNVLIFDLGGGTFDVSILTIEDGIFEVKSTSGDTHLGGEDFDNRMVTH     240
AHY03302.1             IAYGLDKKVGGERNVLIFDLGGGTFDVSILTIEDGIFEVKSTSGDTHLGGEDFDNRMVTH     240
                       ************************************************************

VIRT-1756796:5'3'      FVQEFKRKHKKDMSDNKRAVRRLRTACERAKRTLSSSTQASIEIDSLYEGIDYYTSITRA     300
AHY03302.1             FVQEFKRKHKKDMSDNKRAVRRLRTACERAKRTLSSSTQASIEIDSLYEGIDYYTSITRA     300
                       ************************************************************

VIRT-1756796:5'3'      RFEELNADLFRGTLEPVEKALRDAKMDKPGVNEIVLVGGSTRIPKIQKLLQDFFNGKELN     360
AHY03302.1             RFEELNADLFRGTLEPVEKALRDAKMDKPGVNEIVLVGGSTRIPKIQKLLQDFFNGKELN     360
                       ************************************************************

VIRT-1756796:5'3'      KSINPDEAVAYGAAVQAAILHGDKSEEVQDLLLLDVTPLSLGIETAGGVMTSLIKRNTTI     420
AHY03302.1             KSINPDEAVAYGAAVQAAILHGDKSEEVQDLLLLDVTPLSLGIETAGGVMTSLIKRNTTI     420
                       ************************************************************

VIRT-1756796:5'3'      PTKQTQTFTTYSDNQPGVLIQVYEGERAMTKDNNLLGKFELTGIPPAPRGVPQIEVTFDI     480
AHY03302.1             PTKQTQTFTTYSDNQPGVLIQVYEGERAMTKDNNLLGKFELTGIPPAPRGVPQIEVTFDI     480
                       ************************************************************

VIRT-1756796:5'3'      DANGILNVTACDKSTGKENKITITNDKGRLSKEDIDRMVNDAEKYKAEDEKQKDRISAKN     540
AHY03302.1             DANGILNVTACDKSTGKENKITITNDKGRLSKEDIDRMVNDAEKYKAEDEKQKDRISAKN     540
                       ************************************************************

VIRT-1756796:5'3'      HLESYSFNMKSTVEDENLKDKLSEDDKKKILDKCNEVIHWLDTNQLADKEEYEHKQKELE     600
AHY03302.1             HLESYSFNMKSTVEDENLKDKLSEDDKKKILDKCNEVIHWLDTNQLADKEEYEHKQKELE     600
                       ************************************************************

VIRT-1756796:5'3'      GVCNPIITKLYQSTGGAPGGMPGGMPNFGGAGGAGAAHGGGSGGGPTIEEVD     652
AHY03302.1             GVCNPIITKLYQSTGGAPGGMPGGMPNFGGAGGAGAAHGGGSGGGPTIEEVD     652
                       ***************************************************
```

**Figure 1 A.** Alignment of the ORF with the uncharacterized sequence (Accession No. AHY03302.1)


**Figure 1 B.** ORFs predicted by ExPASy Translate: The green box highlights the longest ORF

### *Physicochemical characteristics*

The physicochemical characteristics of the uncharacterized protein from *C. fluminea* were predicted using the ProtParam tool available on the ExPASy server. The calculated parameters are presented in Table 1.

**Table 1.** Physicochemical characteristics of the uncharacterized protein from *C. fluminea*

| Physicochemical Characteristics | Values |
|---|---|
| Number of amino acids | 652 |
| Molecular weight | 71401.47 |
| Theoretical pI | 5.40 |
| | |
| Amino acid composition | |
| Ala (A) | 47 |
| Arg (R) | 27 |
| Asn (N) | 34 |
| Asp (D) | 48 |
| Cys (C) | 5 |
| Gln (Q) | 22 |
| Glu (E) | 50 |
| Gly (G) | 60 |
| His (H) | 10 |
| Ile (I) | 43 |
| Leu (L) | 46 |
| Lys (K) | 56 |
| Met (M) | 14 |
| Phe (F) | 24 |
| Pro (P) | 24 |
| Ser (S) | 35 |
| Thr (T) | 50 |
| Trp (W) | 2 |
| Tyr (Y) | 15 |
| Val (V) | 40 |
| Pyl (O) | 0 |
| Sec (U) | 0 |
| | |
| Total number of negatively charged residues (Asp + Glu) | 98 |
| Total number of positively charged residues (Arg + Lys) | 83 |
| | |
| Atomic composition | |
| Carbon (C) | 3124 |
| Hydrogen (H) | 5007 |
| Nitrogen (N) | 867 |
| Oxygen (O) | 1005 |
| Sulfur (S) | 19 |

| | |
|---|---|
| Formula | $C_{3124}H_{5007}N_{867}O_{1005}S_{19}$ |
| Total number of atoms | 10022 |
| | |
| Extinction coefficients: | |
| Assuming all pairs of Cys residues form cystines | 33600 |
| Assuming all Cys residues are reduced | 33350 |
| | |
| Estimated half-life | |
| o (Mammalian reticulocytes, in vitro) | 30 hours |
| o (Yeast, in vivo) | >20 hours |
| o (*Escherichia coli*, in vivo) | >10 hours |
| | |
| Instability index | 35.88 |
| | |
| Aliphatic index | 78.24 |
| Grand average of hydropathicity (GRAVY) | -0.507 |

The uncharacterized protein consists of 652 amino acids, with a calculated molecular weight of 71, 401.47 Da. Its theoretical isoelectric point (pI) measured 5.40. The uncharacterized protein contains comparable numbers of positively and negatively charged residues. The instability index was calculated 35.88, while the aliphatic index was 78.24. Lastly, the grand average of hydropathicity (GRAVY) value was –0.507.

### *Domain architecture and protein classification*

The conserved domain architecture of the uncharacterized protein of *C. fluminea* was analyzed using the SMART database to facilitate functional classification.

Domain analysis revealed the presence of two major overlapping regions: a Pfam HSP70 domain (residues 6–612) and a Pfam MreB_Mbl domain (residues 116–384). A low-complexity region was also located at the C-terminus between residues 615–646. These domains were identified and shown in Figure 2.

### *Predicted structure*

Homology modelling with I-TASSER generated ten threading templates (Figure 3) and 10 structural analogs (Figure 4).

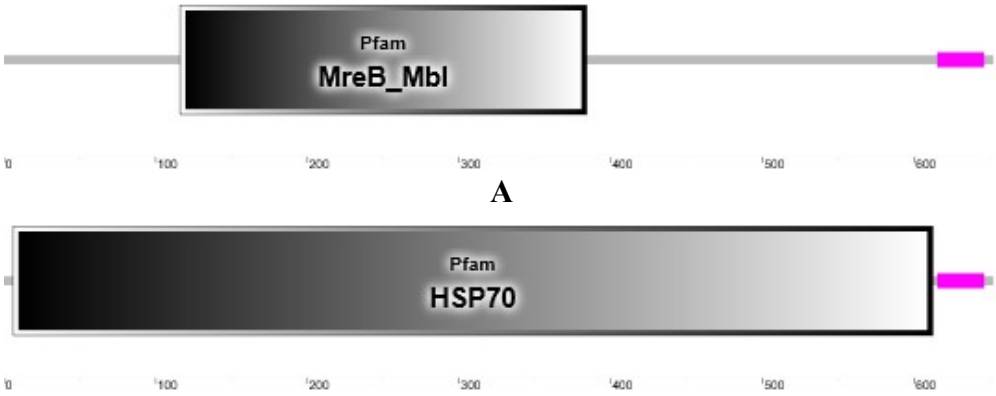


**Figure 2.** Domain architecture of the uncharacterized protein form *C. fluminea:* (A) Representation showing Pfam MreB_Mbl and low complexity regions. (B) Representation showing Pfam HSP70 domain

| Rank | PDB Hit | Iden1 | Iden2 | Cov | Norm. Z-score | Download Align. |
|---|---|---|---|---|---|---|
| 1 | 5e84A | 0.65 | 0.60 | 0.92 | 5.28 | Download |
| 2 | 7n1rA | 0.65 | 0.60 | 0.92 | 5.22 | Download |
| 3 | 5e84 | 0.65 | 0.60 | 0.92 | 2.79 | Download |
| 4 | 5tky | 0.58 | 0.54 | 0.91 | 2.25 | Download |
| 5 | 7n1rA | 0.65 | 0.60 | 0.92 | 6.73 | Download |
| 6 | 5e84 | 0.65 | 0.60 | 0.92 | 3.13 | Download |
| 7 | 6asyA | 0.65 | 0.60 | 0.92 | 1.10 | Download |
| 8 | 7n1rA | 0.65 | 0.60 | 0.92 | 6.16 | Download |
| 9 | 2khoA | 0.51 | 0.47 | 0.91 | 7.91 | Download |
| 10 | 5tkyA | 0.58 | 0.54 | 0.91 | 9.39 | Download |

**Figure 3.** Threading templates used by I-TASSER

Although Template 10 exhibited the highest Z-score, Template 1 was selected as the optimal model due to its superior overall alignment coverage, consensus support, and structural relevance, consistent with I-TASSER's ranking criteria. The structural analogs identified in the Protein Data Bank (PDB) confirmed that the closest analog is 5e84A (TM-score = 0.920) (Figure 4).

**Top 10 Identified stuctural analogs in PDB**

| Click to view | Rank | PDB Hit | TM-score | RMSD[a] | IDEN[a] | Cov | Alignment |
|---|---|---|---|---|---|---|---|
| ○ | 1 | 5e84A | 0.920 | 0.55 | 0.648 | 0.923 | Download |
| ○ | 2 | 5tkyA | 0.856 | 2.58 | 0.571 | 0.909 | Download |
| ○ | 3 | 7ko2A | 0.845 | 2.58 | 0.477 | 0.900 | Download |
| ○ | 4 | 2qxlB | 0.811 | 3.68 | 0.262 | 0.919 | Download |
| ○ | 5 | 7b7zB | 0.777 | 1.68 | 0.688 | 0.797 | Download |
| ○ | 6 | 4j8fA | 0.558 | 3.89 | 0.695 | 0.636 | Download |
| ○ | 7 | 6gfaA | 0.555 | 1.42 | 0.369 | 0.569 | Download |
| ○ | 8 | 7x3kB | 0.539 | 3.32 | 0.248 | 0.598 | Download |
| ○ | 9 | 1yuwA | 0.534 | 4.21 | 0.682 | 0.626 | Download |
| ○ | 10 | 3fe1C | 0.529 | 2.85 | 0.802 | 0.580 | Download |

**Figure 4.** Top 10 structural analogs from PDB

The final top model (Model 1) exhibited the highest confidence, with a C-score of 0.45, an estimated TM-score of 0.77 ± 0.10, and an RMSD of 6.9 ± 4.6 Å (Figure 5A). The closest analog was visualized using EzMol for comparison (Figure 5B).

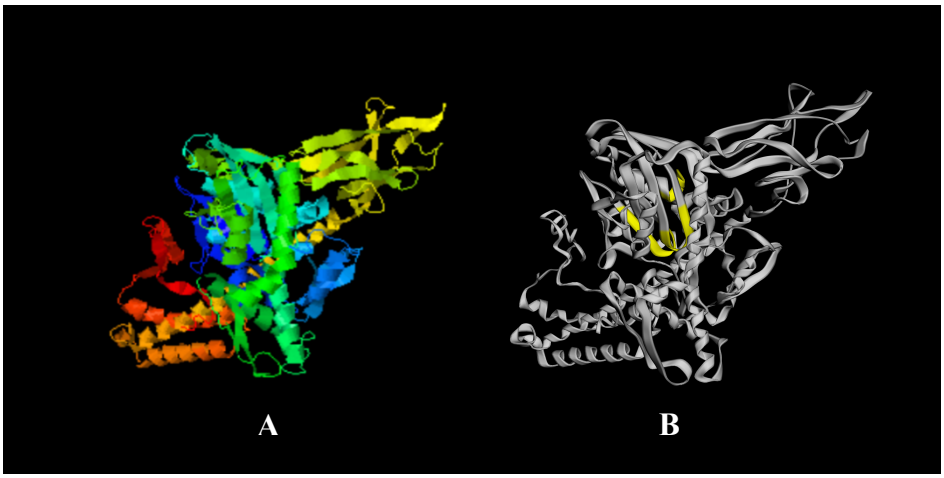


**Figure 5.** Tertiary structural characterization of the uncharacterized *C. fluminea* protein: (A) Full-length 3D model predicted by I-TASSER; (B) Functional mapping via EzMol, highlighting the conserved ATP-binding pocket in yellow within the N-terminal Nucleotide Binding Domain (NBD)

## *Subcellular localization*

DeepLoc analysis predicted the uncharacterized protein as cytoplasmic, with a high probability (0.8946) (Figure 6).

**Prediction: Cytoplasm, Soluble**

| Localization | Cytoplasm | Peroxisome | Mitochondrion | Plastid | Lysosome/Vacuole | Nucleus | Endoplasmic reticulum | Cell membrane | Golgi apparatus | Extracellular |
|---|---|---|---|---|---|---|---|---|---|---|
| Likelihood | 0.8946 | 0.06 | 0.0136 | 0.0109 | 0.0067 | 0.0051 | 0.0051 | 0.0023 | 0.0014 | 0.0004 |

| Type | Soluble | Membrane |
|---|---|---|
| Likelihood | 0.9845 | 0.0155 |

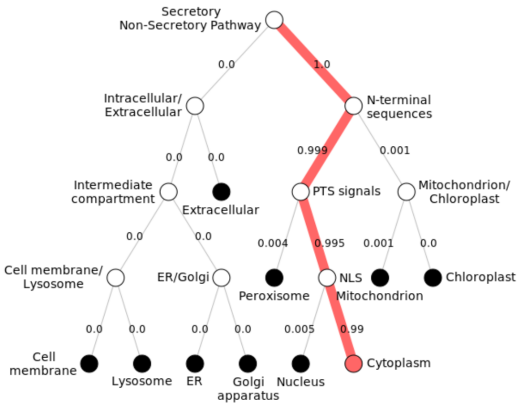


**Figure 6.** Subcellular localization prediction using DeepLoc

## *Protein evolution*

Phylogenetic reconstruction using Maximum Likelihood analysis with 1,000 bootstrap replicates placed the uncharacterized protein within HSP70 clade of *C. fluminea* (Figure 7).

Strong bootstrap support (100%) confirmed its close evolutionary relationship with known HSP70 sequences from the order Venerida. This suggests that the uncharacterized protein is a heat shock protein both structural and functional annotations.

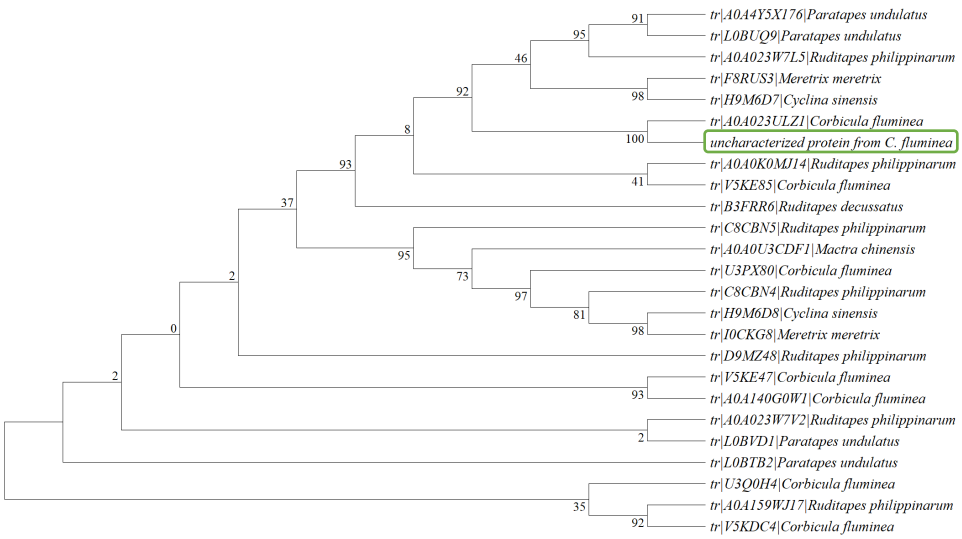


**Figure 7.** Phylogenetic tree showing evolutionary placement of the uncharacterized protein with orthologous HSP70 sequences in the order Venerida

## Discussion

The *in-silico* analysis performed in this study provides strong evidence that the uncharacterized protein from *Corbicula fluminea* belongs to the heat shock 70 protein (HSP70) family. Translation and alignment confirmed its derivation from the HSP70 mRNA clone.

The physicochemical characterization of the uncharacterized protein revealed features consistent with a stable, soluble, and thermostable protein. The molecular weight of 71401.47 kDa features are characteristics of heat shock proteins (HSPs) proteins, as these highly conserved molecular chaperones typically weighs ~70 kDa that assist in protein folding, stabilization, and protection under stress (Craig, 1989; Hu *et al.*, 2022).

The uncharacterized protein from *C. fluminea* also exhibits a calculated isoelectric point (pI) of 5.40 and a negative GRAVY score of -0.507. These additional physicochemical parameters indicate an acidic, highly hydrophilic protein. Proteins with an acidic isoelectric point remain negatively charged at physiological pH, generating electrostatic repulsion that minimizes non-specific interactions and reduces the likelihood of aggregation, particularly under stress conditions where proteins are prone to misfolding. Furthermore, the instability index of 35.88 classifies the protein as structurally stable (<40), while the negative GRAVY score confirms its highly solubility in aqueous cytoplasm.

These traits are characteristics of the HSP70 family, which must remain soluble and functional to act as "molecular chaperones" during environmental stressors such as thermal fluctuations or heavy metal exposure, as previously documented in bivalve species (Xie, 2017).

Domain analysis using SMART identified a conserved Pfam HSP70 domain (6–612), with an overlapping Pfam MreB_Mbl domain (116–384) (Letunic *et al.*, 2020). This structural arrangement reflects the structural homology of the HSP70 nucleotide-binding domain (NBD) to the actin-like ATPase superfamily. The presence of the MreB-like fold represents an evolutionary conserved structural framework that enables ATP binding and hydrolysis to drive the conformational changes essential for the protein folding cycle (Modrzejewska and Zdanowska, 2024; Wentink *et al.*, 2025). The C-terminal low-complexity region (residues 615–646) is likely intrinsically disordered and may serve as a flexible interaction interface that could potentially facilitate association with co-chaperones, although this remains to be experimentally validated (Wentink *et al.*, 2025). Overall, this domain architecture is consistent with the standard HSP70 organization, in which the N-terminal nucleotide-binding domain (NBD) mediates ATP binding and hydrolysis, while the C-terminal substrate-binding domain (SBD) facilitates client protein interaction (Hu *et al.*, 2024; Wentink *et al.*, 2025).

Consistent with this, the generated top model by the I-TASSER exhibited a high structural reliability, with a TM-score of 0.77 and a C-score of 0.45. A TM-score > 0.5 serves as a quantitative benchmark for a correct global fold (Yang and Zhang, 2015). The structural similarity to the analog 5e84A (TM-score 0.920) suggests the deep evolutionary conservation of the HSP70 fold. The functional mapping using EzMol localized a conserved ATP-binding pocket within the N-terminal domain. This spatial conservation is consistent with the canonical function of HSP70 proteins, in which ATP binding and hydrolysis drive conformational changes necessary for chaperon activity (Modrzejewska and Zdanowska, 2024; Wentink *et al.*, 2025). Such structural features are important for maintaining proteostasis, particularly under stress conditions that promote protein misfolding (Modrzejewska and Zdanowska, 2024). In this context, the conservation of ATP-dependent chaperone function may contribute to the physiological resilience of *C. fluminea* in diverse freshwater environments, but such a relationship remains to be validated.

The phylogenetic reconstruction placed the sequence within the HSP70 clade od the order Venerida with 100% bootstrap support, which indicates an evolutionary affiliation with known HSP0 orthologs. This placement, together with the DeepLoc-predicted cytoplasmic localization probability of 0.8946, is consistent the protein being a cytosolic HSP70. The degree of conservation

observed across taxa suggests that this protein is a part of a broadly conserved stress-response system among bivalves. Given the established role of HSP70 proteins in maintaining proteostasis under stress conditions, such conservation may contribute to physiological resilience in fluctuating environments. In aquatic organisms, HSP70 are widely recognized as molecular biomarkers of environmental stress due to their inducible expression under diverse stress conditions such as temperature fluctuations, pollution and pathogenic exposure (Chahouri *et al.*, 2023; Jeyachandran *et al.*, 2023; Xu *et al.*, 2022). However, species-specific functional validation is required.

## Acknowledgements

The first author expresses sincere gratitude to his parents for their constant support and encouragement throughout the course of this study. Deep appreciation is also extended to the National Research Council of the Philippines (NRCP) Capability Advancement through Scientific Capacity Development (CASCaDe) Program support. The authors likewise acknowledge the Department of Biological Sciences at Isabela State University for providing the resources and academic environment that made this work possible.

## Conflicts of interest

The authors declare no conflict of interest.

## References

Armenteros, J. J. A., Sønderby, C. K., Sønderby, S. K., Nielsen, H. and Winther, O. (2017). DeepLoc: Prediction of protein subcellular localization using deep learning. *Bioinformatics*, 33:3387-3395. https://doi.org/10.1093/bioinformatics/btx431

Bitacura, J. G. and Santos, M. D. (2022). In silico structural analysis, classification, and functional annotation of an uncharacterized protein from an aquatic fungus *Lindgomyces ingoldianus*. Genetics of Aquatic Organisms, 7. https://doi.org/10.4194/ga527

Chahouri, A., Yacoubi, B., Moukrim, A. and Banaoui, A. (2023). Bivalve molluscs as bioindicators of multiple stressors in the marine environment: Recent advances. Continental Shelf Research, 264:105056. https://doi.org/10.1016/j.csr.2023.105056

Chen, H., Zha, J., Liang, X., Bu, J., Wang, M. and Wang, Z. (2013). Sequencing and de novo assembly of the Asian clam (*Corbicula fluminea*) transcriptome using the Illumina GAIIx method. PLoS ONE, 8:e79516. https://doi.org/10.1371/journal.pone.0079516

Craig, E. A. (1989). Essential roles of 70 kDa heat inducible proteins. BioEssays, 11:48-52. https://doi.org/10.1002/bies.950110203

Falfushynska, H. I., Phan, T. and Sokolova, I. M. (2016). Long-term acclimation to different thermal regimes affects molecular responses to heat stress in a freshwater clam *Corbicula fluminea*. Scientific Reports, 6. https://doi.org/10.1038/srep39476

Gasteiger, E., Hoogland, C., Gattiker, A., Duvaud, S., Wilkins, M. R., Appel, R. D. and Bairoch, A. (2005). Protein identification and analysis tools on the ExPASY server. In J. M. Walker (Ed.), The proteomics protocols handbook. Humana Press, pp.571-607 https://doi.org/10.1385/1-59259-890-0:571

Hu, B., Liu, G., Zhao, K. and Zhang, G. (2024). Diversity of extracellular HSP70 in cancer: Advancing from a molecular biomarker to a novel therapeutic target. Frontiers in Oncology, 14:1388999. https://doi.org/10.3389/fonc.2024.1388999

Hu, C., Yang, J., Qi, Z., Wu, H., Wang, B., Zou, F., Mei, H., Liu, J., Wang, W. and Liu, Q. (2022). Heat Shock proteins: Biological functions, pathological roles, and therapeutic opportunities. MedComm, 3:e161. https://doi.org/10.1002/mco2.161

Jeyachandran, S., Chellapandian, H., Park, K. and Kwak, I. (2023). A review on the involvement of heat shock proteins (extrinsic chaperones) in response to stress conditions in aquatic organisms. Antioxidants, 12:1444. https://doi.org/10.3390/antiox12071444

Johnson, M., Zaretskaya, I., Raytselis, Y., Merezhuk, Y., McGinnis, S. and Madden, T. L. (2008). NCBI BLAST: A better web interface. Nucleic Acids Research, 36(Web Server), W5-W9. https://doi.org/10.1093/nar/gkn201

Kumar, S., Stecher, G., Suleski, M., Sanderford, M., Sharma, S. and Tamura, K. (2024). MEGA12: Molecular evolutionary genetic analysis version 12 for adaptive and green computing. Molecular Biology and Evolution, 41. https://doi.org/10.1093/molbev/msae263

Letunic, I., Khedkar, S. and Bork, P. (2020). SMART: Recent updates, new developments and status in 2020. Nucleic Acids Research, 49:D458-D460. https://doi.org/10.1093/nar/gkaa937

Mashjel, Z. and Al-Taher, Q. (2025). Molecular identification of the bivalves *Corbicula fluminea* and *Sinanodonta woodiana* in Thi-Qar Province, Iraq, using mitochondrial COX-I Gene Sequencing. Egyptian Journal of Aquatic Biology and Fisheries, 29:2347-2356. https://doi.org/10.21608/ejabf.2025.422893

Modrzejewska, M. and Zdanowska, O. (2024). The role of heat shock protein 70 (HSP70) in the pathogenesis of ocular diseases—Current literature review. Journal of Clinical Medicine, 13:3851. https://doi.org/10.3390/jcm13133851

Okawa, T., Kurita, Y., Kanno, K., Koyama, A. and Onikura, N. (2016). Molecular analysis of the distributions of the invasive Asian clam, *Corbicula fluminea* (O.F. Müller, 1774), and threatened native clam, *C. leana* Prime, 1867, on Kyushu Island, Japan. BioInvasions Records, 5:25-29. https://doi.org/10.3391/bir.2016.5.1.05

Reynolds, C. R., Islam, S. A. and Sternberg, M. J. (2018). EZMOL: A web server wizard for the rapid visualization and image production of protein and nucleic acid structures. Journal of Molecular Biology, 430:2244-2248. https://doi.org/10.1016/j.jmb.2018.01.013

Sidhu, A. (2023). Biological factors and its effect on HSP70 in *Corbicula fluminea*. American Journal of Biomedical Science & Research, 19:673-682. https://doi.org/10.34297/ajbsr.2023.19.002636

Sievers, F. and Higgins, D. G. (2017). Clustal Omega for making accurate alignments of many protein sequences. Protein Science, 27:135-145. https://doi.org/10.1002/pro.3290

Valenzuela-Castillo, A. V., Sanchez-Paz, A. S., Castro-Longoria, R. C., Lopez-Torres, M. L. and Grijalva-Chon, J. G. (2019). Hsp70 function and polymorphism, its implications for mollusk aquaculture: A review. Latin American Journal of Aquatic Research, 47:224-231. https://doi.org/10.3856/vol47-issue2-fulltext-2

Wentink, A., Rosenzweig, R., Kampinga, H. and Bukau, B. (2025). Mechanisms and regulation of the Hsp70 chaperone network. Nature Reviews Molecular Cell Biology, 27:110-128. https://doi.org/10.1038/s41580-025-00890-9

Xie, Y. (2017). Molecular characterization of the HSP70 and HSP90 genes in Asian clam (*Corbicula fluminea*) and their expression analysis during heavy metal exposure. Gene Reports, 7:18-24. https://doi.org/10.1016/j.genrep.2017.01.002

Xu, Y., Liang, J., He, G., Liu, X., Yang, K., Masanja, F., Deng, Y. and Zhao, L. (2022). Responses of pearl oysters to marine heatwaves as indicated by HSP70. Frontiers in Marine Science, 9. https://doi.org/10.3389/fmars.2022.847585

Yang, J. and Zhang, Y. (2015). I-TASSER server: New development for protein structure and function predictions. Nucleic Acids Research, 43:W174-W181. https://doi.org/10.1093/nar/gkv342

Zhang, T., Yin, J., Tang, S., Li, D., Gu, X., Zhang, S., Suo, W., Liu, X., Liu, Y., Jiang, Q., Zhao, M., Yin, Y. and Pan, J. (2021). Dissecting the chromosome-level genome of the Asian clam (*Corbicula fluminea*). Scientific Reports, 11. https://doi.org/10.1038/s41598-021-94545-2

(Received: 15 August 2025, Revised: 20 June 2026, Accepted: 3 July 2026)