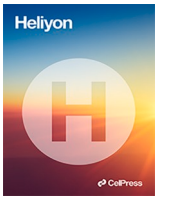
Heliyon 10 (2024) e32748

Contents lists available at ScienceDirect

Heliyon

journal homepage: www.cell.com/heliyon



Research article

In silico characterization and homology modeling of Nile tilapia (*Oreochromis niloticus*) Hsp70cBi and Hsp70cBc proteins

Geraldine B. Dayrit [a,b,c,*], Normela Patricia F. Burigsay [c], Emmanuel M. Vera Cruz [d], Mudjekeewis D. Santos [a,d,e]

[a] *The Graduate School, University of Santo Tomas, España Boulevard, Manila, 1015, Philippines*
[b] *College of Public Health, University of the Philippines Manila, Ermita, Manila, 1000, Philippines*
[c] *ONE ARM, Department of Medical Microbiology, College of Public Health, University of the Philippines Manila, Ermita, Manila, 1000, Philippines*
[d] *Freshwater Aquaculture Center, Central Luzon State University, Science City of Muñoz, Nueva Ecija, Philippines*
[e] *National Fisheries Research and Development Institute, Genetic Fingerprinting Laboratory, 101 Mother Ignacia Ave., South Triangle, Quezon City, 1101, Philippines*

A R T I C L E I N F O

Keywords:
Hsp70
Oreochromis niloticus
In silico characterization
Homology modeling

A B S T R A C T

The molecular chaperone heat shock proteins 70 (Hsp70) play a pivotal role in preserving cellular integrity and managing stress. This study extensively examined two Hsp70 proteins, *On*-Hsp70cBi, inducible, and *On*-Hsp70cBc, constitutively expressed, in Nile tilapia (*Oreochromis niloticus*) utilizing in silico analysis, homology modeling, and functional annotation. Employing the SWISS-MODEL program for homology modeling, the proposed models underwent thorough reliability assessment via ProSA, Verify 3D, PROVE, ERRAT, and Ramachandran plot analyses. Key features of On-Hsp70cBi and On-Hsp70cBc included amino acid lengths (640 and 645) and molecular weights (70,233.48 and 70,773.17 Da). Moreover, theoretical isoelectric points (pI = 5.63 and 5.28), indicated their acidic nature. Counts of negatively and positively charged residues (95 and 86; 95 and 81) revealed neutrality, while instability index (II) values of 35.27 (*On*-Hsp70cBi) and 38.85 (*On*-Hsp70cBc) suggested stability. Aliphatic index (AI) values were notably high for both proteins (84.58 and 82.85), indicating stability across a broad temperature range. Domain architecture analysis showed both proteins to contain an MreB/Mbl domain. Protein-protein interaction analysis identified the co-chaperone Stip1 as a primary functional partner. Comparative modeling yielded highly reliable 3D models, showcasing structural similarity to known proteins and predicted binding sites. Additionally, both proteins are primarily localized in the cytoplasm. Functional analysis predicted an AMP-PNP binding site for *On*-Hsp70cBi and an ATP binding site for *On*-Hsp70cBc. These findings deepened our understanding of Hsp70cBc and Hsp70cBi in Nile tilapia, underscoring their significance in fish physiology and warranting further investigation, thus advancing our knowledge of these proteins' roles in cellular processes and stress responses, potentially impacting fish health and resilience.

* Corresponding author. Department of Medical Microbiology, College of Public Health, 3rd F Lara Hall, Pedro Gil St., Ermita, Manila, 1000, Philippines.
E-mail address: gbdayrit@up.edu.ph (G.B. Dayrit).

https://doi.org/10.1016/j.heliyon.2024.e32748
Received 16 February 2024; Received in revised form 4 June 2024; Accepted 7 June 2024
Available online 8 June 2024

1. Introduction

Nile tilapia (*Oreochromis niloticus*) is a vital fish species in the Philippines, significantly contributing to its aquaculture industry. However, the escalating challenges posed by climate change, particularly heat stress, have impeded its production. Ectothermic by nature, fish are exceptionally susceptible to fluctuations in temperature and other environmental stressors [1]. Amid these challenges, the protein Hsp70 (Heat shock protein 70) emerges as a key player in temperature adaptation, exerting crucial roles in cellular thermotolerance [2]. Hsp70 aids in the folding and unfolding of proteins, facilitating their proper conformation even under stressful conditions [3]. Additionally, it acts as a molecular chaperone, safeguarding proteins from denaturation and assisting in their refolding [4]. Notably, Hsp70 is pivotal in preventing cellular damage induced by heat stress, thereby promoting the survival of fish under adverse environmental conditions. Literature further underscores the significance of Hsp70 in facilitating cellular homeostasis and stress responses in fish species, including tilapia [5]. Understanding the intricate mechanisms underlying the functions of Hsp70 is imperative for elucidating the adaptive strategies of tilapia in the face of environmental challenges.

This study employs computational techniques to characterize two distinct isoforms of Hsp70 in Nile tilapia: Hsp70cBi (inducible) and Hsc70cBc (constitutive). By analyzing both the structural and functional characteristics of these proteins, our research aims to gain a thorough understanding of their properties. Specifically, we aim to elucidate the specific distinctions between their physicochemical traits, structural arrangements, subcellular localization, and interactions with other proteins. Achieving these research objectives will provide valuable insights into the adaptive mechanisms of tilapia and shed light on the fundamental molecular pathways that regulate their cellular functions and responses to stress.

Furthermore, to the best of our knowledge, no in silico three-dimensional (3D) models of proteins from the *O. niloticus* family have been published. Therefore, insights gained from this study may have applications beyond tilapia and extend to other fish species, offering potential benefits in fisheries and aquatic ecosystems.

In the forthcoming sections of this paper, we will present the methodologies employed for our in-depth analysis of tilapia's Hsp70cBi and Hsc70cBc proteins. This study integrates physicochemical characterization, domain architecture analysis, comparative homology modeling, structural analysis, and function prediction to provide a comprehensive overview of these proteins.

2. Materials and methods

2.1. Sequence retrieval

The mRNA sequences of *On*-Hsp70cBi (Accession No.: XM_019357557.1) and *On*-Hsp70cBc (Accession No.: WLP61716.1 were uploaded to Expasy's Translate Server using compact formatting with no spaces on both the forward and reverse strands [6]. The obtained protein sequences were aligned with each other using MUSCLE, available on the EMBL-EBI server [7].

2.2. Phylogenetic analysis

Amino acid sequences of Hsp70cBi and Hsp70cBc from various fish species and representatives of previously described Hsp70 family members were downloaded from NCBI database (Table 1) [8]. The sequences were screened and selected by BLASTp search for homolog proteins to *On*-Hsp70cBi and *On*-Hsp70cBc. All sequences were aligned and a maximum likelihood (ML) phylogenetic tree was constructed using MEGA X software [9]. The best fit amino acid substitution model was selected according to the lowest BIC score, and phylogenetic analysis was performed with LG + G model and 1000 replications.

Table 1
Amino acid sequences used for multiple alignment with *On*-Hsp70cBi and *On*-Hsp70cBc.

Phylum	Family	Species	Protein	GenBank accession no	Lineage
Chordata	Cichlidae	*Oreochromis niloticus*	Hsp70i	XP_019213102	
Chordata	Cichlidae	*Oreochromis niloticus*	Hsp70c	WLP61716.1	
Chordata	Cichlidae	*Oreochromis aureus*	Hsp70i	ACJ04036.1	
Chordata	Cichlidae	*Oreochromis mossambicus*	Hsp70i	CAA04673.1	
Chordata	Cichlidae	*Simochromis diagramma*	Hsp70 i(LOC120733556)	XP_063328606.1	
Chordata	Cichlidae	*Pelmatolapia mariae*	Hsp70c	XP_063344520.1	
Chordata	Cichlidae	*Simochromis diagramma*	Hsp70c	XP_039885390.1	
Chordata	Cichlidae	*Astatotilapia calliptera*	Hsp70c	XP_026040332.1	
Arthropoda	Insecta	*Drosophila melanogaster*	Hsp70i	AAG26887.1	lineage A
Mollusca	Gastropoda	*Lottia gigantea*	hypothetical protein LOTGIDRAFT_181897	XP_009051716.1	lineage A
Rotifera	Monogononta	*Brachionus ibericus*	Hsp70i	ADR79281.1	lineage A
Chordata	Mammalia	*Bos tuarus*	Hsp70i	AAA73914.1	lineage B
Nematoda	Chromadorea	*Caenorhabditis elegans*	Hsp70i	AAA28078.1	lineage B
Mollusca	Bivalvia	*Crassostrea gigas*	Hsp70ci	BAD15287.1	lineage B
Nematoda	Chromadorea	*Caenorhabditis elegans*	BiP, heat shock protein 3	AAA28074.1	ER

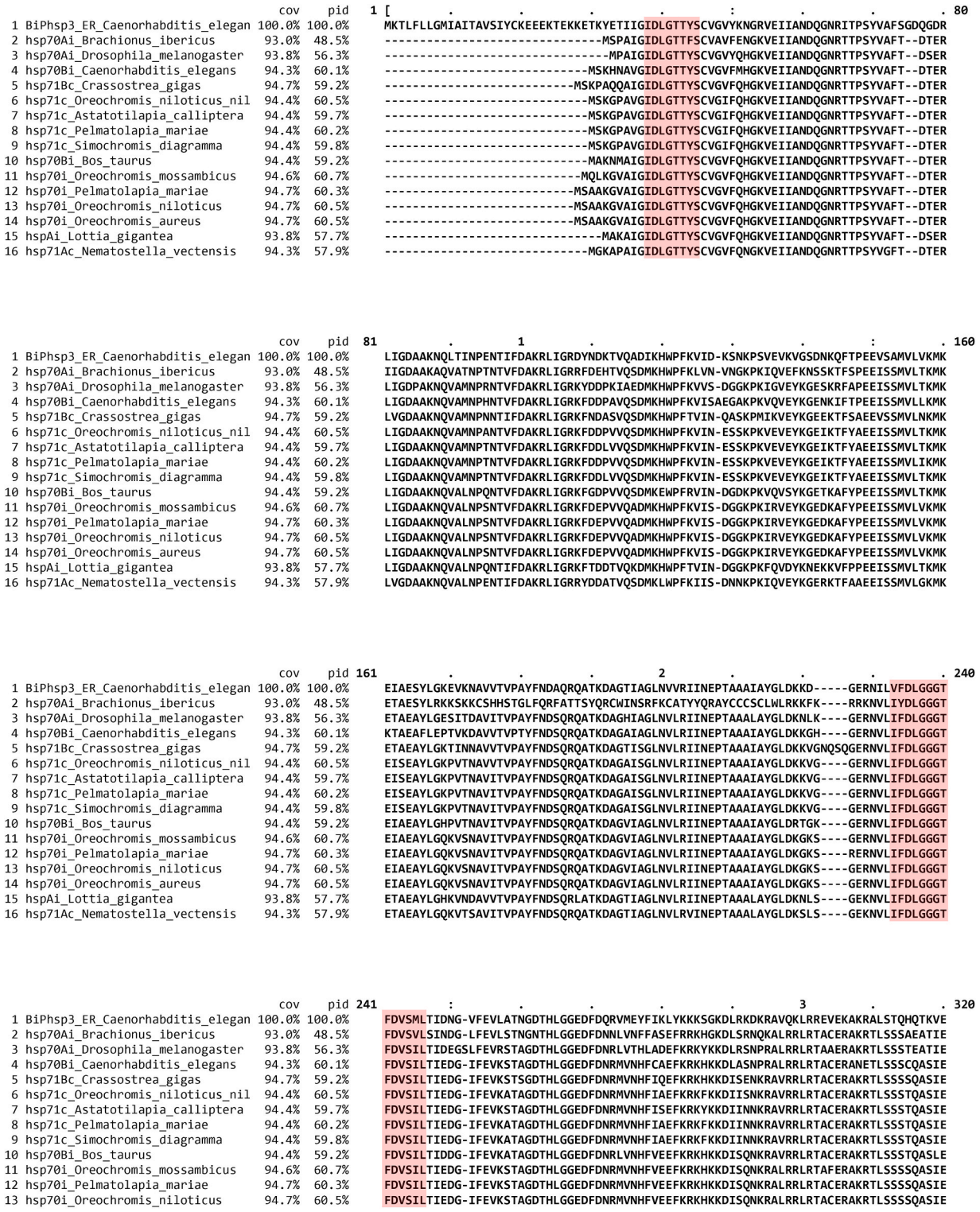


Fig. 1. Alignment of *On*-Hsp70cBi, *On*-Hsp70cBc, and other homolog proteins.

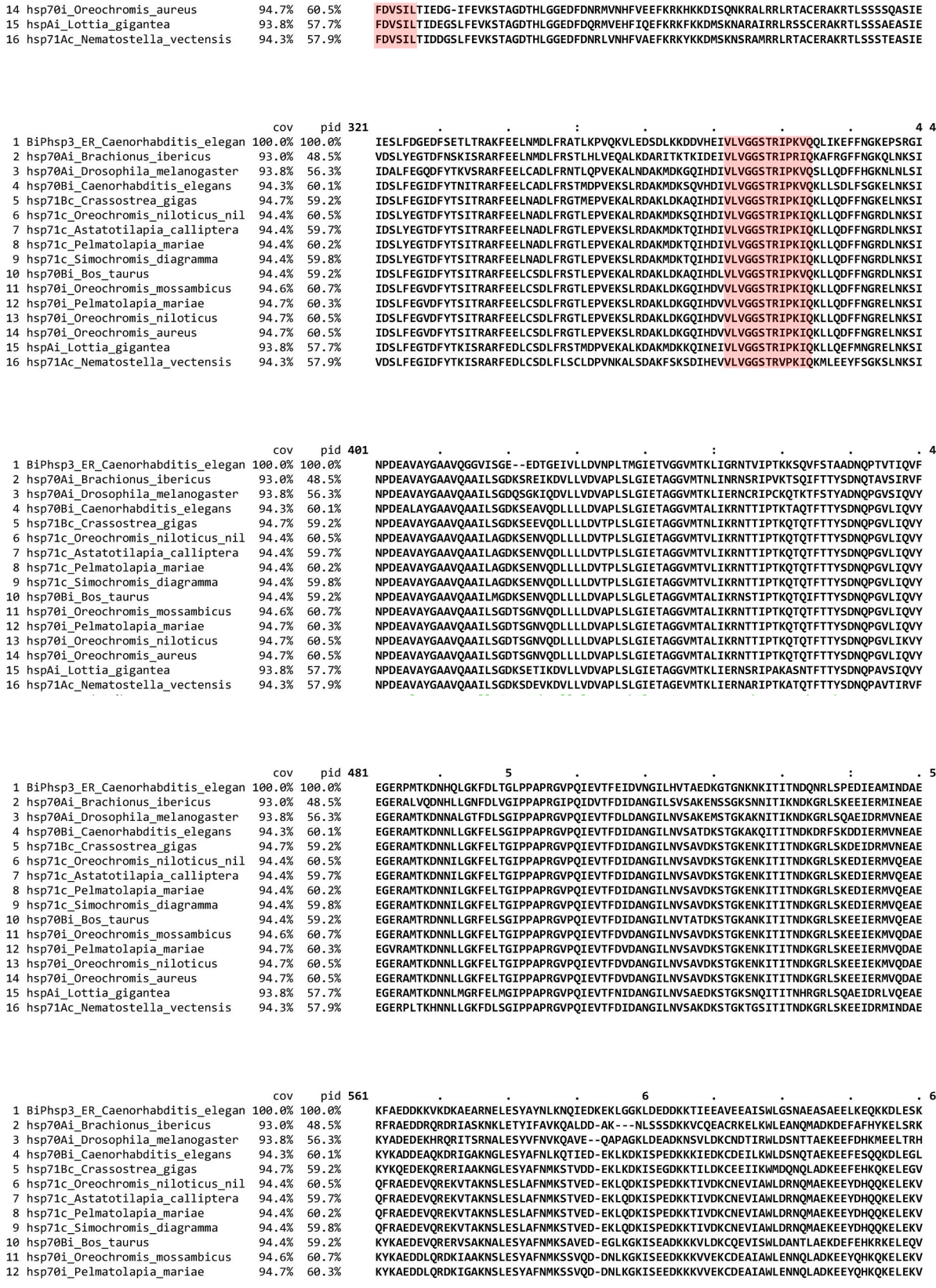


Fig. 1. (*continued*).

```
 13 hsp70i_Oreochromis_niloticus      94.7%  60.5%     KYKAEDDLQRDKIAAKNSLESYAFNMKSSVQD-DNLKGKISEEDKKKVVEKCDEAIAWLENNQLADKEEYQHKQKELEKV
 14 hsp70i_Oreochromis_aureus         94.7%  60.5%     KYKAEDDLQRDKIAAKNSLESYAFNMKSSVQD-DNLKGKISEEDKKKVVEKCDETIAWLENNQLADKEEYQHKQKELEKV
 15 hspAi_Lottia_gigantea             93.8%  57.7%     KFKDEDNKQKEKMAARNQLESYIYSVKQTV---DAAGDKLAAEDKETAKKACENGLNWLEGNALADKDEYEYKMQEIQKV
 16 hsp71Ac_Nematostella_vectensis    94.3%  57.9%     KYKSEDEAQREKIAARNRLESYAFGVKSAISE-PSLEGKLSQSDKDTVKNKVEEVLNWLEKNSLAEKEEFEEQEKELQRV
```

```
                                     cov    pid 641         :         .         .         .         7] 701
  1 BiPhsp3_ER_Caenorhabditis_elegan 100.0% 100.0%     VQPIVSKLYK---------------D------------AGAGER--RPQKRDLDDKDEL
  2 hsp70Ai_Brachionus_ibericus       93.0%  48.5%     CMPLMKKIHS--G--------------------------ERNG-----PTVEEVE-----
  3 hsp70Ai_Drosophila_melanogaster   93.8%  56.3%     CSPIMTKMHQ-QGAGAAGGPGANCGQQ-------------AGGFGGYSGPTVEEVD-----
  4 hsp70Bi_Caenorhabditis_elegans    94.3%  60.1%     AKPDLSKLYQSAG-GAPP--GAAPGG--------------AAGGAG--GPTIEEVD-----
  5 hsp71Bc_Crassostrea_gigas         94.7%  59.2%     CNPIITKLYQASG-GAPG--GGMPGGMPNFGGGAPGG--GAPGGGSGGGPTIEEVD-----
  6 hsp71c_Oreochromis_niloticus_nil  94.4%  60.5%     CNPIISKLYQ--G-GMP---GGMPGGMP---GGFPG----GAAGGSSSGPTIEEVD-----
  7 hsp71c_Astatotilapia_calliptera   94.4%  59.7%     CNPIISKLYQ--G-GMP---GAMPGGMP---GGMPGGFPGGAGGGSSSGPTIEEVD-----
  8 hsp71c_Pelmatolapia_mariae        94.4%  60.2%     CNPIISKLYQ--G-GMP---GGMPGGMP---GGFPG----GAGGGSSSGPTIEEVD-----
  9 hsp71c_Simochromis_diagramma      94.4%  59.8%     CNPIISKLYQ--G-GMP---GGMPGGMP---GGMPGGFPDGAGGGSSSGPTIEEVD-----
 10 hsp70Bi_Bos_taurus                94.4%  59.2%     CNPIISRLYQ--GAGGPG--AGGFGAQ-------------GPKGGSGSGPTIEEVD-----
 11 hsp70i_Oreochromis_mossambicus    94.6%  60.7%     CNPIISKLYQ--G-GMPT--GATCGEQ-------------ARAGSQ--GPTIEEVD-----
 12 hsp70i_Pelmatolapia_mariae        94.7%  60.3%     CNPIISKLYQ--G-GMPT--GATCGEQ-------------ARAGSQ--GPTIEEVD-----
 13 hsp70i_Oreochromis_niloticus      94.7%  60.5%     CNPIISKLYQ--G-GKPT--GATCGEQ-------------ARAGSQ--GPTIEEVD-----
 14 hsp70i_Oreochromis_aureus         94.7%  60.5%     CNPIISKLYQ--G-GMPT--GATCGEQ-------------ARAGSQ--GPTIEEVD-----
 15 hspAi_Lottia_gigantea             93.8%  57.7%     CSPIMTKMHQGGG-G-----QQQAGGQ-------------QSAGSDGSGPTVEEVD-----
 16 hsp71Ac_Nematostella_vectensis    94.3%  57.9%     CSPIMAKVH---G-GSGG--TQMPGQT-------------AGHGQSAGGPTIEEVD-----
```

Fig. 1. (*continued*).

2.3. Physiochemical characterization

The amino acid sequences of Hsp70cBi and Hsp70cBc from *O. niloticus* were submitted to Expasy's ProtParam Server to determine their physicochemical properties [6]. A protein sequence or Swiss-Prot or TrEMBL entry can be used to compute physical and chemical parameters for the protein.

2.4. Domain architecture analysis

The conserved domains within Hsp70cBi and Hsp70cBc proteins were identified through domain architecture analysis. Utilizing the Simple Molecular Architecture Research Tool (SMART) [10], this analysis focused on outlier homologues, PFAM domains, signal peptides, and internal repeats in genomic mode. This approach facilitated the identification and annotation of genetically mobile domains, allowing for the analysis of domain architectures. Each domain's presence across distinct species, functional categories, three-dimensional structures, and critical residues essential for function were thoroughly documented. To ensure accuracy, every domain identified in a non-repetitive protein database, along with corresponding search criteria and taxonomic details, was meticulously recorded in a relational database system [11].

2.5. Comparative homology modeling

The SWISS-MODEL service was used to model the proteins' homology by aligning the input target protein with existing templates [12]. Sequence identity and GMQE were used to choose the best template for building the 3D model [13]. GalaxyRefine was then used to refine protein structure through side-chain reconstruction, then side-chain repacking, and finally, relaxation of the entire structure via molecular dynamics simulations [14]. Various analyses, including PROCHECK's Ramachandran plot analysis, ERRAT, PROVE, Verify3D, and ProSA (all accessible from the SAVES server at http://nihserver.mbi.ucla.edu), were used to evaluate the stereochemical quality and correctness of the predicted models [15]. Lastly, structural analysis was conducted, and Swiss PDB Viewer was used to create model representations.

2.6. Subcellular localization

The DeepLoc-1.0 server was used to predict the proteins' subcellular position [16]. This server's Neural Networks algorithm was trained on eukaryotic proteins from UniProt that have experimental evidence supporting their subcellular localization. Note that this prediction process bases all conclusions entirely on information from the protein's sequence.

2.7. Structural similarity and functional annotation

The TM-align method was used to perform a global structural match on the COFACTOR web server [17,18]. The TM-score, a measure of the overall structural similarity, was calculated as a result of this comparison. A TM-score of 1.0 indicates a perfect match between two structures [19]. TM-scores range from 0.0 to 1.0. Scores below 0.17, on the other hand, indicate that the proteins were selected at random and are unrelated. A score higher than 0.5 denotes that the structures have a fold that is usually similar [20]. The I-TASSER suite was used to annotate the analysis on ligand-binding locations, gene ontology, and enzyme commission [21].

3. Results

3.1. Sequence characterization

The sequence alignment depicted in Fig. 1 illustrates the comparison between *On*-Hsp70cBi and *On*-Hsp70cBc, revealing an 85.24 % identity between these two heat shock proteins as determined by MUSCLE analysis. Three signature motifs (IDLGTTYS, IFDLGGGTFDVSIL, and VVLVGGSTRIPRIQK) were identified in the HspP70 amino acid sequences, representing highly conserved regions present across all Hsp70 family members [22]. The cytoplasmic Hsp70 C-terminal region of EEVD was present in all proteins except the outgroup protein *C. elegans*, which is known to be a part of the endoplasmic reticulum (ER) lineage.

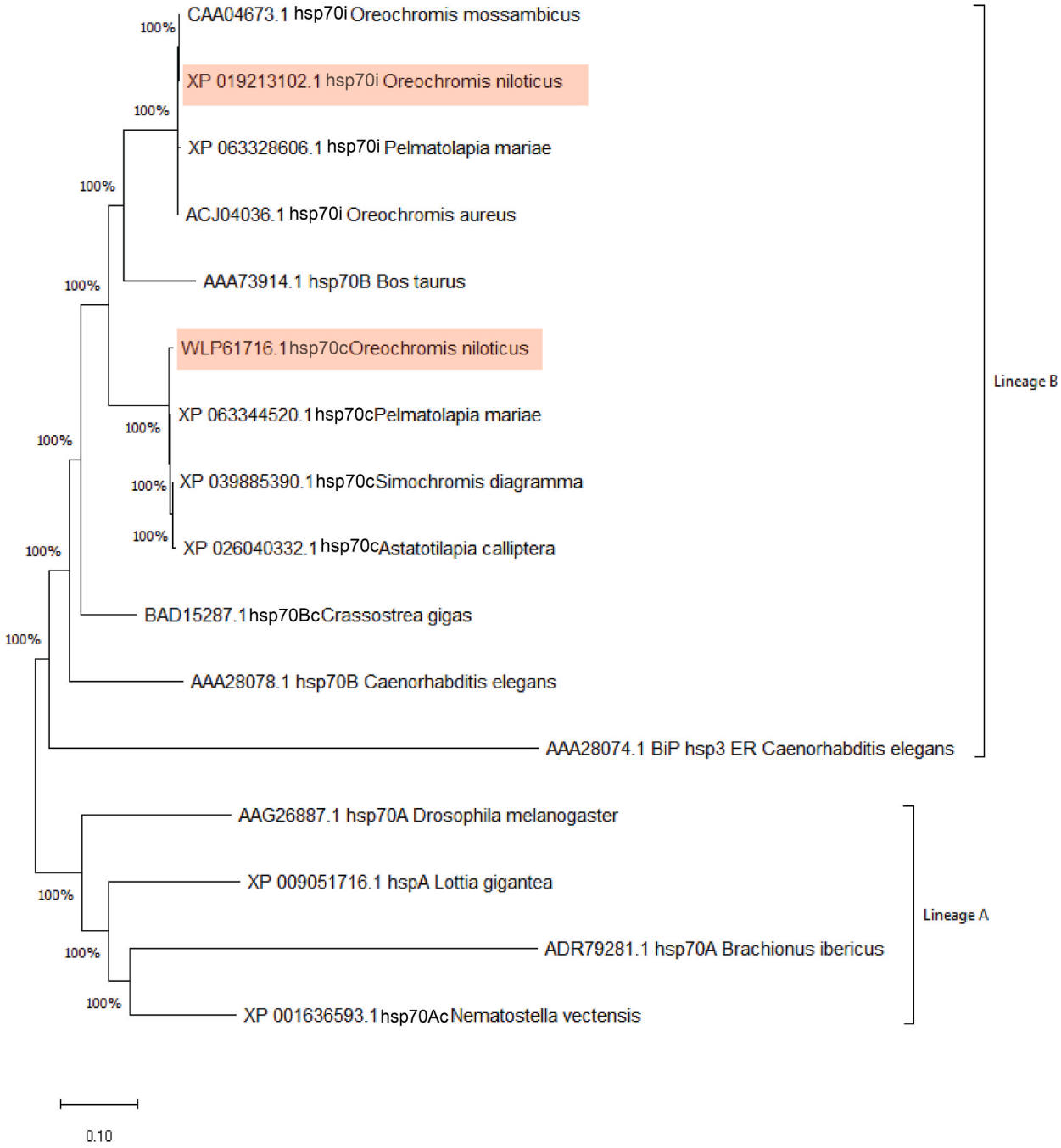


Fig. 2. Phylogenetic tree of *On*-Hsp70cBi, *On*-Hsp70cBc, and other homolog proteins.

3.2. Phylogenetic analysis

In the phylogenetic analysis as shown in Fig. 2, inducible proteins from various species were observed to cluster together, as did constitutively expressed proteins, forming distinct clades. Both proteins of interest, *On*-Hsp70cBi and *On*-Hsp70cBc, belonged to the cytosolic lineage B but did not form a monophyletic clade, supporting data from previous studies [9].

3.3. Physiochemical characteristics

The key characteristics of the *On*-Hsp70cBi and *On*-Hsp70cBc protein are summarized in Table 2, including amino acid lengths (640 and 645), molecular weights (70,233.48 and 70,773.17 Da), theoretical isoelectric points (pI = 5.63 and 5.28), and the total counts of negatively and positively charged residues (95 and 86; 95 and 81). Among the novel findings, the computed pI value revealed the acidic nature of the proteins, signifying their suitability for purification using isoelectric focusing on a polyacrylamide gel (pI < 7). The calculated extinction coefficient (EC), directly linked to the cysteine, tryptophan, and tyrosine content, at 280 nm absorbance, amounted to 33725/32110 (assuming all pairs of cysteine residues form cysteines) and 33350/31860 (assuming all cysteine residues are reduced) M−1 cm−1 for *On*-Hsp70cBi and *On*-Hsp70cBc proteins, respectively. Furthermore, the instability index (II) values were 35.27 (*On*-Hsp70cBi) and 38.85 (*On*-Hsp70cBc), indicating the stability of both proteins (II < 40). The aliphatic index (AI) exhibited high values for both proteins, reaching 84.58 (*On*-Hsp70cBi) and 82.85 (*On*-Hsp70cBc), underscoring their stability across a broad temperature range.

The estimated half-life values were consistent for both proteins: exceeding 30 h in mammalian reticulocytes (in vitro), over 20 h in yeast (in vivo), and more than 10 h in *Escherichia coli* (in vivo). Furthermore, the grand average hydropathicity (GRAVY) values of *On*-Hsp70cBi and *On*-Hsp70cBc were found to be −0.440 and −0.404, respectively, confirming that these proteins are more likely to be globular and hydrophilic rather than membranous and hydrophobic. An analysis of amino acid composition revealed that *On*-Hsp70cBi had elevated levels of alanine (8.3 %), lysine (9.1 %), and glycine (8.0 %), while *On*-Hsp70cBc contained higher proportions of alanine (7.8 %), glutamic acid (7.8 %), lysine (8.1 %), and glycine (8.2 %).

Table 2
Physicochemical properties of *On-Hsp70cBi* and *On-Hsp70cBc*.

Property	On-Hsp70cBi	On-Hsp70cBc
Number of amino acids	645	640
Molecular weight	70,773.17 Da	70,233.48 Da
Theoretical pI	5.28	5.63
Amino Acid Composition (% of total amino acids)		
Ala (A)	50 (7.8 %)	53 (8.3 %)
Arg (R)	29 (4.5 %)	28 (4.4 %)
Asn (N)	32 (5.0 %)	31 (4.8 %)
Asp (D)	45 (7.0 %)	46 (7.2 %)
Cys (C)	4 (0.6 %)	6 (0.9 %)
Gln (Q)	25 (3.9 %)	27 (4.2 %)
Glu (E)	50 (7.8 %)	49 (7.7 %)
Gly (G)	53 (8.2 %)	51 (8.0 %)
His (H)	6 (0.9 %)	7 (1.1 %)
Ile (I)	48 (7.4 %)	44 (6.9 %)
Leu (L)	42 (6.5 %)	47 (7.3 %)
Lys (K)	52 (8.1 %)	58 (9.1 %)
Met (M)	15 (2.3 %)	9 (1.4 %)
Phe (F)	25 (3.9 %)	23 (3.6 %)
Pro (P)	26 (4.0 %)	21 (3.3 %)
Ser (S)	36 (5.6 %)	38 (5.9 %)
Thr (T)	45 (7.0 %)	39 (6.1 %)
Trp (W)	2 (0.3 %)	2 (0.3 %)
Tyr (Y)	14 (2.2 %)	15 (2.3 %)
Val (V)	46 (7.1 %)	46 (7.2 %)
Pyl (O)	0 (0.0 %)	0 (0.0 %)
Sec (U)	0 (0.0 %)	0 (0.0 %)
(B)	0 (0.0 %)	0 (0.0 %)
(Z)	0 (0.0 %)	0 (0.0 %)
(X)	0 (0.0 %)	0 (0.0 %)
Total number of negatively charged residues (Asp + Glu)	95	95
Total number of positively charged residues (Arg + Lys)	81	86
Extinction coefficients		
Ext. coefficient	32,110	33,725
Abs 0.1 % (=1 g/l)	0.454 (Cys residues form cystines)	0.480 (Cys residues form cystines)
Ext. coefficient	31,860	33,350
Abs 0.1 % (=1 g/l)	0.450 (Cys residues are reduced)	0.475 (Cys residues are reduced)

3.4. Domain architecture analysis

Fig. 3 shows *On*-Hsp70cBi and *On*-Hsp70cBc containing an MreB/Mbl domain at positions 119 to 383 and 108 to 382, respectively. The only difference detected between *On*-Hsp70cBi and *On*-Hsp70cBc was the low complexity region (LCR) along positions 613 to 639 on *On*-Hsp70cBc. LCRs, which are extremely common in eukaryotic proteins, are amino acid sequences containing repeats of one amino acid or brief amino acid motifs [23].

3.5. Comparative homology modeling

The Hsp 70A Ciliary C1 central pair apparatus from *Chlamydomonas reinhardtii* (PDB ID: 7sqc) was selected as the most suitable template for homology modeling of *On*-Hsp70cBi and *On*-Hsp70cBc exhibited in Fig. 4, showing a GMQE score of 0.82 and percent sequence identities of 74.10 % and 74.65 %, respectively.

The quality of the models was assessed using various validation tools summarized in Table 3, while Fig. 5 reveals their corresponding plots. The Ramachandran plot analysis revealed excellent quality, with 95.50 % of residues falling within the most favored regions for *On*-Hsp70cBi and 95.40 % for *On*-Hsp70cBc. The overall G-factor values for *On*-Hsp70cBi (0.02) and *On*-Hsp70cBc (0.20) were well within the acceptable range, indicating the reliability of the models.

Verify3D analysis showed an average 3D-1D score of $\geq$0.1 for 86.26 % of On-Hsp70cBi and 80.63 % of *On*-Hsp70cBc residues, supporting the validity of the predicted models. The ERRAT quality factor values for *On*-Hsp70cBi (92.1466 %) and *On*-Hsp70cBc (94.6274 %) exceeded the commonly accepted threshold of 95 % for high-resolution structures.

Additional assessments included Z-scores, with values of −11.76 for *On*-Hsp70cBi and −11.96 for *On-Hsp70cBc*, falling within the expected range for native proteins of comparable size. The plot of residue energies revealed predominantly negative values, indicating a lack of problematic or erroneous segments in the input structures.

3.6. Subcellular localization

The DeepLoc-1.0 server suggested a predominant presence within the cytoplasm for both proteins, as shown in Table 4. Additionally, studies have identified EEVD as the motif for the cytoplasmic Hsp70 C-terminal region, a finding depicted in Fig. 6 for both sequences. For *On*-Hsp70cBi, the next most plausible subcellular locations included the nucleus, cell membrane, and extracellular space, while for *On*-Hsp70cBc, these included the extracellular space, lysosome, and nucleus. It is noteworthy, however, that the localizations ranked as the next most probable options did not meet the server's predefined threshold for confidence.

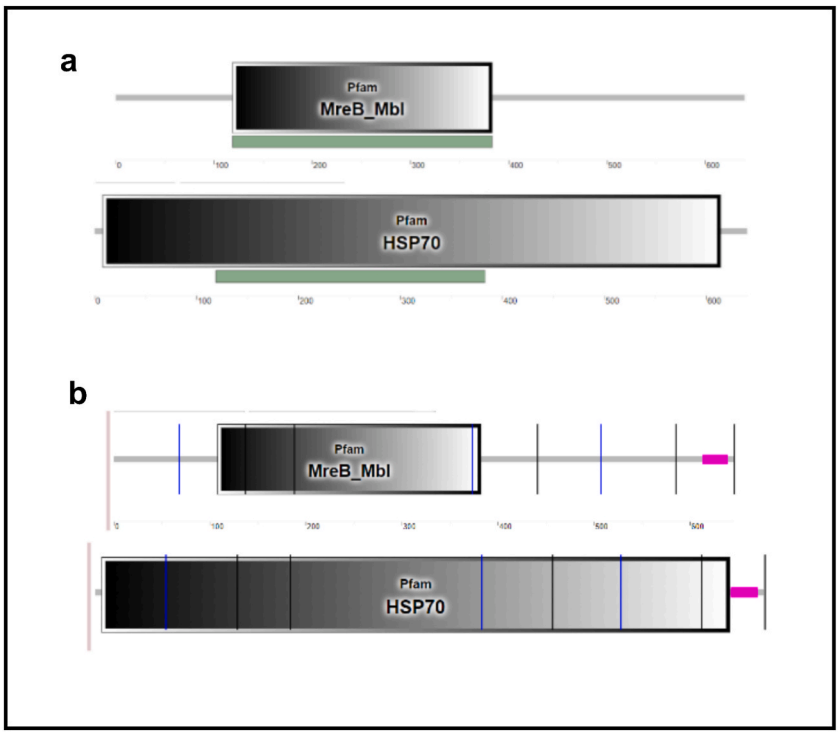


Fig. 3. Domain architecture of **(a)** *On*-Hsp70cBi and **(b)** *On*-Hsp70cBc.

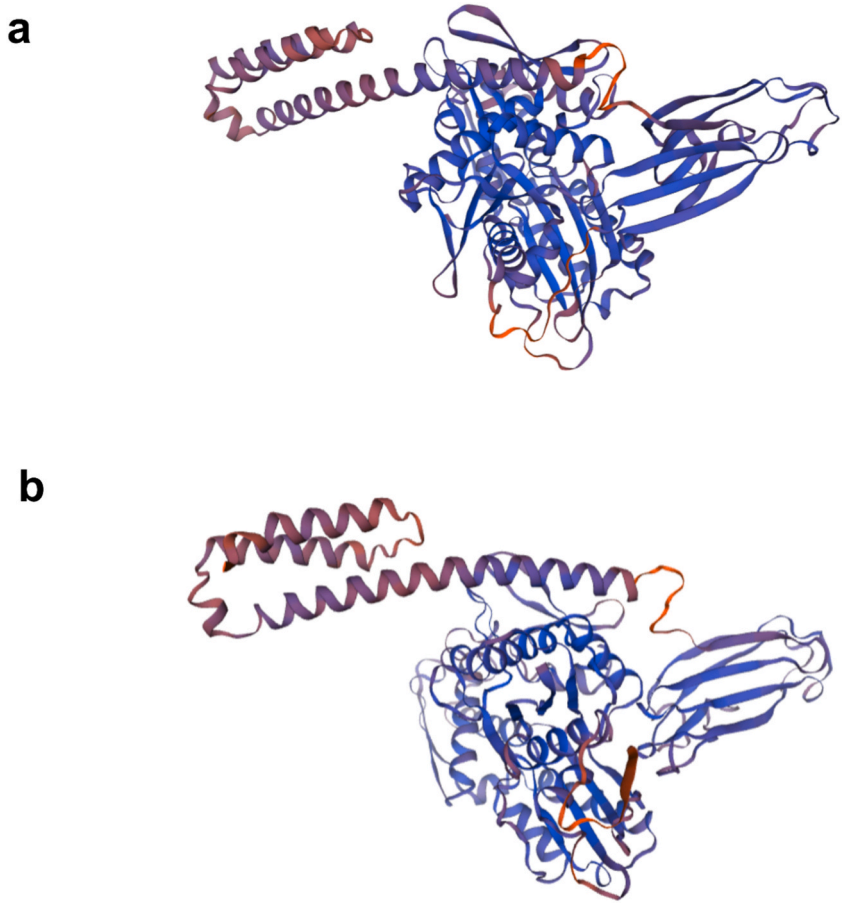


Fig. 4. Predicted models of **(a)** *On*-Hsp70cBi and **(b)** *On*-Hsp70cBc.

Table 3
Assessment of the predicted three-dimensional structures of *On-Hsp70cBi* and *On-Hsp70cBc* proteins.

Validation Index	*On-Hsp70cBi*	*On-Hsp70cBc*
Ramachandran plot		
Residues in most favored regions	95.5 %	95.4
Residues in additional allowed regions	3.9 %	4.2 %
Residues in generously allowed regions	0.4 %	0.0 %
Residues in disallowed regions	0.2 %	0.4 %
Overall G-factor	0.20	0.20
ProSA Z-Score	−11.76	−11.96
ERRAT	92.1466	94.6274

3.7. Protein-protein interaction analysis

Table 5 revealed STRING query results indicating that stress-induced phosphoprotein 1 (stip1) stands out as the functional partner with the highest confidence score for both *On*-Hsp70cBi (0.833) and *On*-Hsp70cBc proteins (0.909). The common functional partners of *On*-Hsp70cBi and *On*-Hsp70cBc, which include GAK, DNAJA4, DNAJC2, LOC100708200, LOC100703595, and Dnaja2, play a variety of roles in cellular processes. Interestingly, *On*-Hsp70cBi's unique functional partner LOC100711553 is a DnaJ heat shock protein family (Hsp40) member B1a, which binds to unfolded proteins, while *On*-Hsp70cBc's bag3 is a BCL2-associated athanogene 3, that is not well studied yet in fish but is in tetrapods and humans, wherein it plays a role in chaperone binding.

3.8. Molecular docking

Molecular docking was conducted for both proteins with ATP as their ligand (Fig. 7). Remarkably, among the ten docking models generated for *On*-Hsp70cBi, eight exhibited highly probable binding, evidenced by confidence scores exceeding 0.70, while the

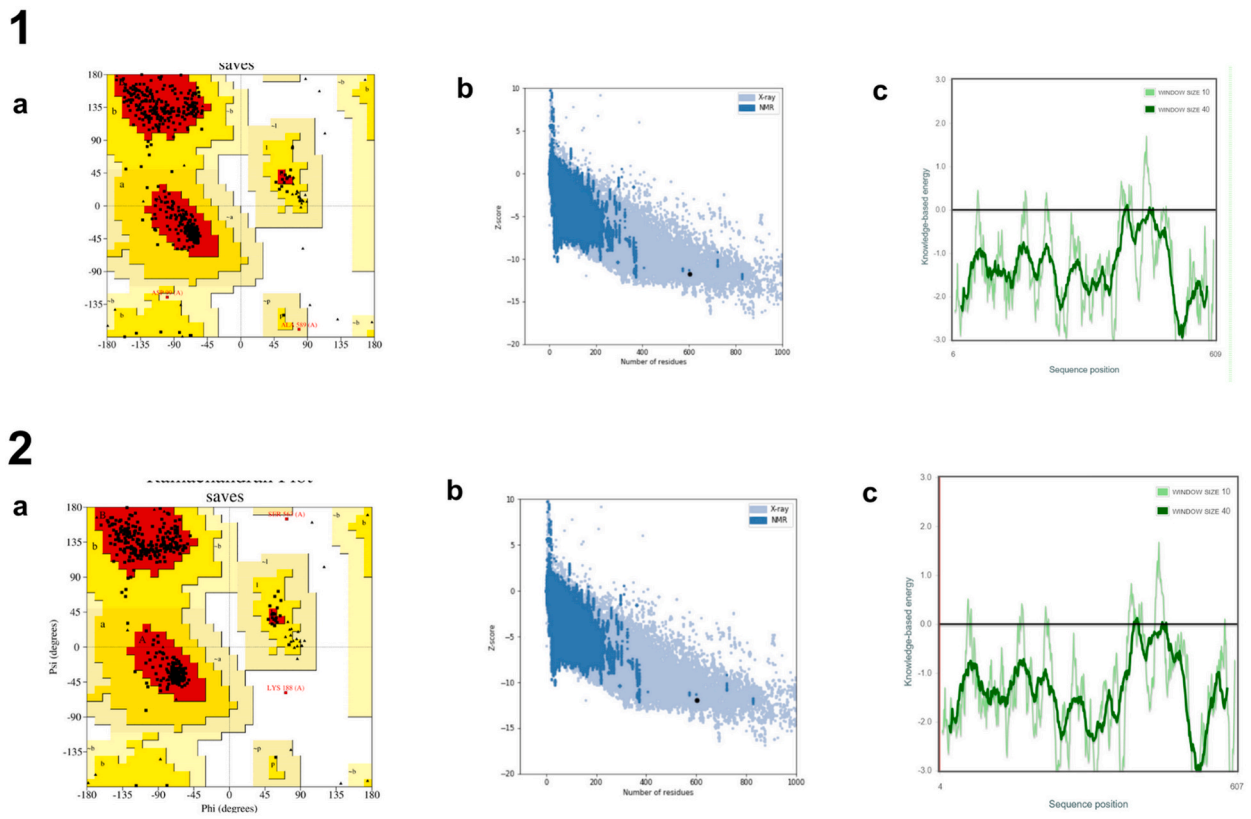


Fig. 5. Validation indices for **(1)** *On*-Hsp70cBi and **(2)** *On*-Hsp70cBc. **(a)** Ramachandran plot analysis shows where residues are located, emphasizing those that are in the desirable (red), permissible (yellow), liberally allowed (light yellow), and forbidden (white) categories. **(b)** Plotted alongside Z-scores from several experimentally determined protein chains in the Protein Data Bank is the model's Z-score, which is shown as a black dot. The light-blue and dark-blue colors represent various sources, such as X-ray and NMR, respectively. **(c)** The ERRAT program is used to evaluate the overall model quality. It uses two horizontal lines on the error axis (representing 95 % and 99 % confidence levels) to show places where structural components can be reliably rejected.

remaining two models demonstrated potential binding. On the other hand, for *On*-Hsp70cBc, seven out of the ten models highlighted robust binding, while three models displayed possible binding. This outcome accentuates the comparable binding affinities observed between the two molecular chaperones and ATP, shedding light on their shared molecular recognition mechanisms.

3.9. Structure similarity analysis

TM-scores >0.5 in Table 6 show that the folding of the two proteins is comparable. Notably, *On*-Hsp70cBi, *On*-Hsp70cBc, and the *Bos taurus* Hsc70 protein (PDB ID: 1yuwA) have a TM-score of 0.99.

3.10. Function prediction

Results from the COFACTOR server showed that an Adenylyl Imidodiphosphate (AMP-PNP) binding site was predicted for *On*-Hsp70cBi based on human Hsp70 (PDB ID: 2e8aA, C-score = 0.83). As with previous findings, *On*-Hsp70cBc has an ATP binding site based on structural similarity with bovine Hsc70 (PDB ID: 1kaxA, C-score = 0.85). Furthermore, COFACTOR predicted that the *On*-Hsp70cBi and *On*-Hsp70cBc proteins have phosphate ion (PO4)-binding sites according to their structural similarity to other proteins with known PO4-binding sites, such as human Hsp70 proteins 3jxuA (C-score = 0.42) and 3gdqA (C-score = 0.42), respectively.

Furthermore, in Table 7, I-TASSER's Enzyme Commission analysis predicted that both *On*-Hsp70cBi and *On*-Hsp70cBc are structurally similar to the isozymes rhamnulokinase, glucokinase, hexokinase, and xylulokinase, all of which can transfer an inorganic phosphate group from ATP to a substrate. Predicted Gene Ontology (GO) Terms also showed that heterocyclic compound binding, such as ATP binding, is a molecular function of both proteins.

4. Discussion

Sequence alignment provides strong support for the claim made in the literature that Hsp70cBi and Hsp70cBc have an 85 %

Table 4
Probability results for the subcellular localizations of *On-Hsp70cBi* and *On-Hsp70cBc*.

Localization	Cytoplasm	Nucleus	Extracellular	Cell membrane	Mitochondrion	Plastid	Endoplasmic reticulum	Lysosome/Vacuole	Golgi apparatus	Peroxisome
On-Hsp70cBi	0.7637	0.3319	0.3223	0.3284	0.2132	0.0305	0.1050	0.2913	0.2607	0.1437
On-Hsp70cBc	0.7788	0.3006	0.3759	0.2958	0.1947	0.0247	0.1397	0.3111	0.2534	0.0871

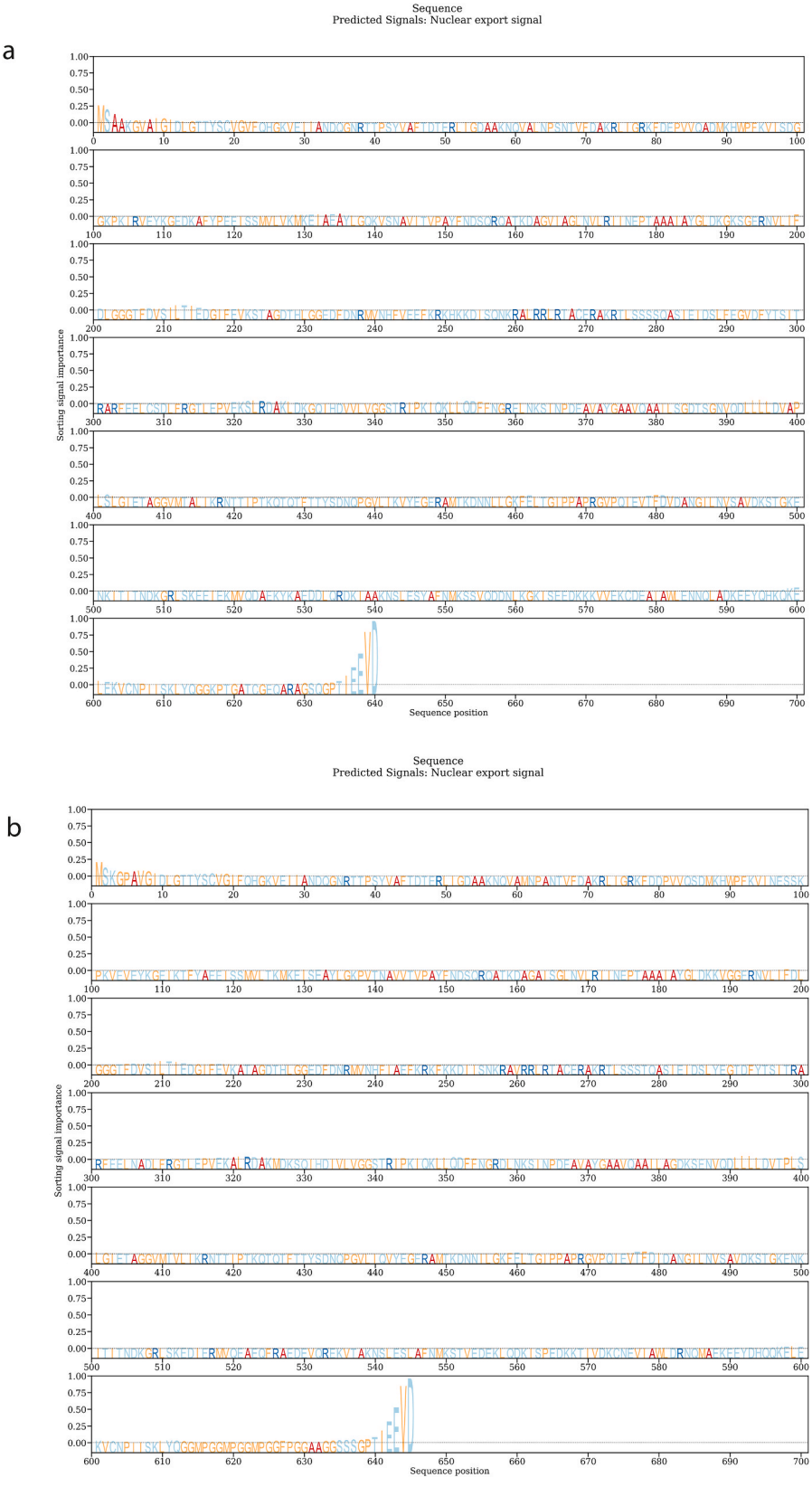


Fig. 6. Sorting signal importance plot of **(a)** *On*-Hsp70cBi and **(b)** *On*-Hsp70cBc.

Table 5
Predicted functional partners of *On*-Hsp70 and *On-Hsp70cBc*.

	Functional Partner	Description	Score
On-Hsp70cBi	stip1	Stress-induced phosphoprotein 1	0.883
	Hsp90ab1	Heat shock protein 90, alpha (cytosolic), class B member 1	0.873
	Gak	Cyclin G-associated kinase	0.863
	DNAJA4	DnaJ heat shock protein family (Hsp40) member A4	0.840
	dnajc2	DnaJ (Hsp40) homolog, subfamily C, member 2	0.840
	LOC100708200	DnaJ heat shock protein family (Hsp40) member A1	0.840
	LOC100703595	DnaJ heat shock protein family (Hsp40) member A2b	0.838
	LOC100701982	Heat shock protein 90, alpha (cytosolic), class A member 1, tandem duplicate 1	0.837
	Dnaja2	DnaJ heat shock protein family (Hsp40) member A2	0.833
	LOC100711553	DnaJ heat shock protein family (Hsp40) member Bla	0.816
On-Hsp70cBc	stip1	Stress-induced phosphoprotein 1	0.909
	LOC100701982	Heat shock protein 90, alpha (cytosolic), class A member 1, tandem duplicate 1	0.883
	bag3	BCL2 associated athanogene 3	0.876
	gak	Cyclin G-associated kinase	0.863
	Hsp90ab1	Heat shock protein 90, alpha (cytosolic), class B member 1	0.858
	dnajc2	DnaJ (Hsp40) homolog, subfamily C, member 2	0.840
	DNAJA4	DnaJ heat shock protein family (Hsp40) member A4	0.839
	LOC100708200	DnaJ heat shock protein family (Hsp40) member A1	0.839
	LOC100703595	DnaJ heat shock protein family (Hsp40) member A2b	0.837
	Dnaja2	DnaJ heat shock protein family (Hsp40) member A2	0.832

sequence identity and are thought to play similar roles in cells [24]. The present study's high degree of sequence similarity underlines the evolutionary relationship and functional similarity of these two proteins. This similarity in sequence identity provides more evidence that *On*-Hsp70cBi and *On*-Hsp70cBc have related and overlapping functions in cellular processes.

Their domains show MreB and Mbl proteins from bacteria. MreB is a protein that contributes to the construction of the bacterial cytoskeleton and is known to determine the rod shape in bacteria [25]. MreB/Mbl-coding genes are exclusive to elongated bacteria and not coccoid forms [26]. According to a theory, the eukaryotic cytoskeleton's components, tubulin and actin, may have originated from prokaryotic precursor proteins linked to the modern bacterial proteins FtsZ and MreB/Mbl [27]. The chaperone Hsp70, the actin subfamily, and sugar kinases are all members of the broader superfamily that contains actin, highlighting an evolutionary link between the found domain and Hsp70.

Moreover, the extensive tools for validation confirm the high quality and dependability of the suggested 3D models for *On*-Hsp70cBi and *On*-Hsp70cBc. Both models satisfy strict structural requirements, as shown by the positive results of the Ramachandran plot analysis, G-factor calculation, Verify3D analysis, ERRAT quality factor evaluation, Z-scores, and residue energies plot. These results demonstrate the homology models' potential as accurate representations of the corresponding heat shock proteins and provide confidence in their applicability for additional research. The models' agreement with multiple validation metrics and their alignment with experimental data support the validity of the structural predictions and open the door to further in-depth studies of the functional properties and interactions of *On*-Hsp70cBi and *On*-Hsp70cBc.

The localization patterns observed for *On*-Hsp70cBi and *On*-Hsp70cBc are consistent with existing literature. A study by Rai et al. [28] supports this notion, as it was also found that human *On*-Hsp70cBi and *On*-Hsp70cBc were seen primarily in the cytoplasm, supporting multiple metabolic processes.

Further supporting evidence comes from Balogi et al. [29], who noted that Hsp70 is typically found in the cytoplasm but can translocate to the nucleus under stress conditions. The protein's presence in the endosomal-lysosomal system, on the cell membrane, and in the extracellular environment during pathophysiological states adds complexity to its subcellular distribution.

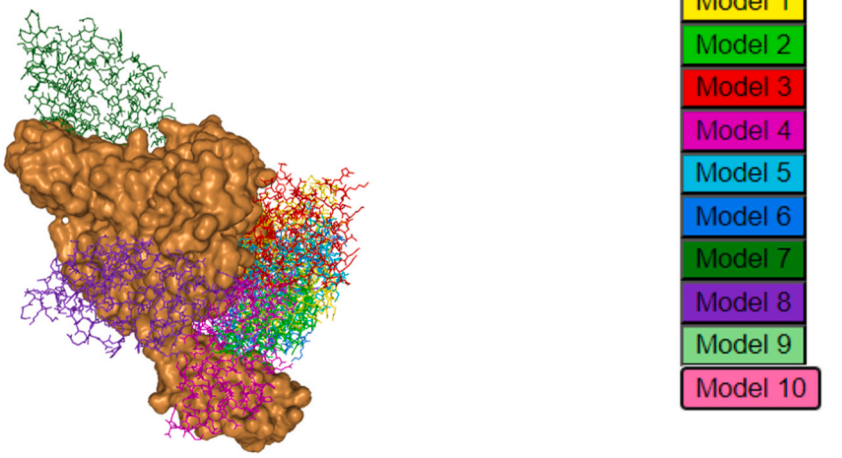
On-Hsp70cBc's findings align with the study of Bonam et al. [30], confirming its presence in the nucleus and external exosomes. Additionally, its localization in lysosomes is supported by its involvement in chaperone-mediated autophagy, as described by Tekirdag and Cuervo [31]. These diverse localizations underscore the multifunctional roles of Hsp70cBc in cellular processes beyond the cytoplasm, emphasizing its involvement in various cellular compartments and pathways.

The STRING analysis provided valuable insights into the potential functional partners of *On*-Hsp70cBi and *On*-Hsp70cBc. Stip1, also known as stress-induced Phosphoprotein 1, is a well-known co-chaperone of heat shock proteins. This protein plays a pivotal role in facilitating protein folding and stability, particularly in times of cellular stress. Beyond its chaperone function, stip1 has been implicated in promoting cell proliferation and is even suggested to function as an oncogenic factor [32].

The common functional partners of both proteins include GAK, Dnaja4, Dnajc2, Loc100708200, Loc100703595, and Dnaja2, which have various roles in cellular processes. GAK is associated with cyclin G and participates in cell cycle regulation and proliferation signaling. The Dnaja4 is a member of the DnaJ heat shock protein family and contributes to cellular stress responses and protein homeostasis. The Dnaja2, another member of the DnaJ family, assists heat shock proteins in protein folding. Loc100708200 and Loc100703595, both members of the Hsp40 class, serve as co-chaperones to ensure proper protein folding. Dnaja2, also an Hsp40 member, plays a vital role in chaperone machinery by assisting in protein folding and stability. Collectively, these proteins contribute to the cellular stress response, maintenance of protein homeostasis, and the proper functioning of proteins in the cell, especially in challenging conditions.

For unique partners, Loc100711553, in humans, is mostly found in the cytoplasm and nucleus. It interacts with Hsp70cBi and can

A

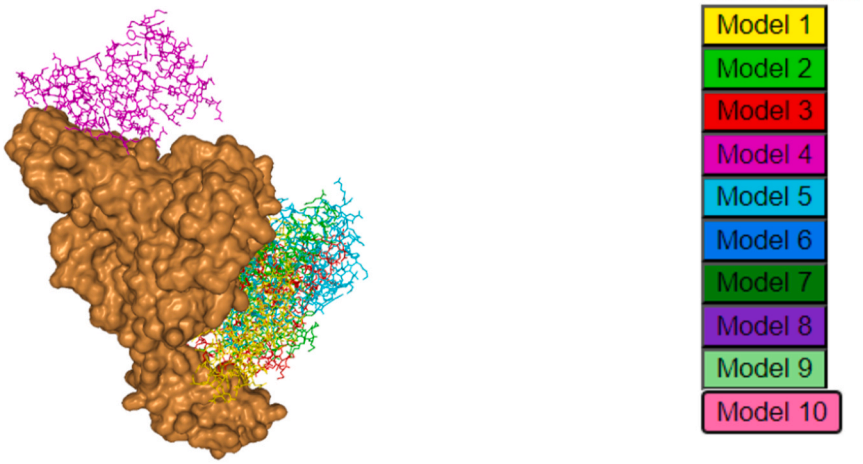


Summary of the Top 10 Models

Rank	1	2	3	4	5	6	7	8	9	10
Docking Score	-217.27	-214.77	-206.54	-202.47	-202.27	-197.86	-197.85	-196.27	-186.76	-185.43
Confidence Score	0.7934	0.7851	0.7560	0.7407	0.7399	0.7226	0.7225	0.7161	0.6759	0.6701
Ligand rmsd (Å)	519.32	525.38	509.20	510.41	522.94	530.55	497.42	485.88	496.02	480.32
Interface residues	model_1	model_2	model_3	model_4	model_5	model_6	model_7	model_8	model_9	model_10

Note: The models are ranked according to the docking scores. Please click help for the explanations of evaluation metrics

B



Summary of the Top 10 Models

Rank	1	2	3	4	5	6	7	8	9	10
Docking Score	-217.54	-214.80	-200.60	-198.51	-196.35	-196.29	-185.06	-185.05	-182.45	-180.27
Confidence Score	0.7943	0.7852	0.7334	0.7252	0.7165	0.7162	0.6685	0.6684	0.6568	0.6469
Ligand rmsd (Å)	515.82	527.48	526.87	509.30	520.80	516.96	507.93	524.15	527.48	514.83
Interface residues	model_1	model_2	model_3	model_4	model_5	model_6	model_7	model_8	model_9	model_10

Note: The models are ranked according to the docking scores. Please click help for the explanations of evaluation metrics

Fig. 7. Top ten models showing binding interactions of **(A)** *On*-Hsp70cBi and **(B)** *On*-Hsp70cBc with ATP as the ligand.

Table 6
Top five identified structural analogs in the Protein Data Bank (PDB) library.

	Protein	Species	PDB Hit	TM - score	RMSDa	IDENa	Cov.
On-Hsp70cBi	Hsc	*Bos taurus*	1yuwA	0.99	1.24	0.892	1
	Hsp	*Homo sapiens. Rattus norvegicus*	4j8fA	0.69	1.95	0.672	0.715
	Hsp	*Homo sapiens*	3iucc	0.68	1.09	0.684	0.688
	RAC	*Saccharomyces cerevisiae*	7 × 3kB	0.67	3.42	0.271	0.752
	BiP	*Cricetulus griseus*	7a4uA	0.67	2.27	0.507	0.701
On-Hsc 70	Hsc	*Bos taurus*	1yuwA	0.99	1.22	0.923	1
	Hsp	*Homo sapiens. Rattus norvegicus*	4j8fA	0.69	1.94	0.656	0.715
	Hsp	*Homo sapiens*	3iuc C	0.68	1.09	0.695	0.688
	RAC	*Saccharomyces cerevisiae*	7 × 3kB	0.67	3.42	0.264	0.752
	BiP	*Cricetulus griseus*	7a4uA	0.67	2.27	0.513	0.701

Table 7
Enzyme Commission (EC) predictions for *On*-Hsp70 and *On-Hsp70cBc* proteins.

	Enzyme	C-score	PDBHit	TM-score	RMSD	IDEN	Cov.	EC Number	Predicted Active Site Residue
On-Hsp70cBi	rhamnulokinase	0.060	2uytA	0.398	5.94	0.081	0.544	2.7.1.5	203,340,369
	Hexokinase glucokinase	0.060	1v4tA	0.325	4.87	0.060	0.411	2.7.1.22.7.1.1	NA
	xylulokinase	0.060	2nlxA	0.393	4.46	0.119	0.482	2.7.1.17	14,18
	glucokinase	0.060	1q18A	0.354	4.90	0.133	0.454	2.7.1.2	NA
	hexokinase	0.060	3b8aX	0.441	4.88	0.111	0.558	2.7.1.1	NA
On-Hsp70cBc	rhamnulokinase	0.060	2uytA	0.399	5.96	0.092	0.544	2.7.1.5	201,338,367
	Hexokinase glucokinase	0.060	1v4tA	0.325	4.87	0.053	0.411	2.7.1.2.2.7.1.1	NA
	xylulokinase	0.060	2nlxA	0.393	4.46	0.110	0.480	2.7.1.17	12
	glucokinase	0.060	1q18A	0.354	4.91	0.125	0.454	2.7.1.2	NA
	hexokinase	0.060	3b8aX	0.441	4.89	0.111	0.558	2.7.1.1	NA

stimulate its ATPase activity and association between Hsp70cBc and HIP. Bag3 (BAG family molecular chaperone regulator 3) is found in cytoplasm, membrane, neuron projection, nucleus and plasma membrane and functions as co-chaperone for Hsp70cBi and Hsp70cBc. It acts as a nucleotide-exchange factor (NEF), promoting the release of ADP from the Hsp70cBi and Hsp70cBc proteins, thereby triggering client/substrate protein release.

The molecular docking results show that *On*-Hsp70cBi and *On*- Hsp70cBc have remarkably similar ATP binding affinities. Most docking models have high confidence scores, indicating strong and probable binding interactions between ATP and chaperones. The ATP-binding sites of *On*-Hsp70cBi and *On*-Hsp70cBc appear to have some functional similarities based on their shared molecular recognition mechanism.

Intriguing questions concerning the structural and functional similarities between these two heat shock proteins are also brought up by their similar binding affinities. Comprehending the subtleties of their associations with ATP can aid in clarifying their functions in cellular procedures, like the folding and unfolding of proteins dependent on ATP. These docking results lay the groundwork for additional research into the residues and molecular characteristics that influence the binding affinity of *On*-Hsp70cBi and *On-Hsp70cBc* with ATP, offering valuable insights into their functional mechanisms and potential regulatory pathways.

The two proteins, *On*-Hsp70cBi and *On*-Hsp70cBc, fold similarly as their TM-scores are more than 0.5. According to Sharma and Masison's work [33], the study's results, especially the TM-scores higher than 0.66 (Table 5), demonstrate how highly conserved both proteins are across species, including humans, yeast, and rodents.

One interesting finding is that the *Bos taurus* Hsp70cBc protein (PDB ID: 1yuwA), *On*-Hsp70cBi, and *On*-Hsp70cBc all have TM-scores of 0.99, which indicates almost exact structural similarities. The same Cov. Scores bolster their homology even more. The three-dimensional structures of these proteins appear to be more similar than their linear amino acid sequences, based on minor variations in the RMSD and IDEN scores. The complex folding patterns that support the overall structural conservation seen may be the source of this subtle variation.

The challenges caused by the lack of fish Hsp70cBi/Hsp70cBc templates in the Protein Data Bank (PDB) are also acknowledged in the discussion. The inability to compare the primary and secondary structures of *On*-Hsp70cBi and *On*-Hsp70cBc with other proteins may have resulted from this absence. These restrictions on the availability of templates highlight the need for care when interpreting some structural metrics and highlight the significance of further structural studies using templates specific to fish to improve our comprehension of the structural properties of these heat shock proteins.

Results from the COFACTOR server provided insights into potential ligand binding sites for *On*-Hsp70cBi and *On*-Hsp70cBc. The prediction of an AMP-PNP binding site for *On*-Hsp70cBi suggests the presence of an ATP binding site. AMP-PNP, an ATP analog, is commonly used to study ATPases and their roles in cellular processes, further supporting the likelihood of an ATP binding site in *On*-Hsp70cBi. Similarly, structural similarity with bovine Hsc70 confirms the presence of an ATP binding site in *On*-Hsp70cBc. Additionally, COFACTOR predicted phosphate ion (PO4)-binding sites for both proteins, reinforcing their structural similarity to proteins

with known PO4-binding sites.

Further analysis by I-TASSER revealed structural similarities between *On*-Hsp70cBi and *On*-Hsp70cBc and isozymes such as rhamnulokinase, glucokinase, hexokinase, and xylulokinase. These enzymes are known for transferring inorganic phosphate groups from ATP to substrates. Predicted Gene Ontology (GO) Terms indicated the molecular function of heterocyclic compound binding, including ATP binding, for both proteins.

These findings align with previously described operational mechanisms, where Hsp70s in their ATP-bound state rapidly capture and release their substrates, whereas Hsp70s in their ADP-bound state firmly grasp them. Hsp70s conduct their chaperone role by transitioning between the ATP and ADP-bound states. Nevertheless, the analysis lacked the sensitivity to differentiate between the roles of the constitutive and inducible isoforms. This limitation is not a significant hindrance since it has been suggested that the functional disparity primarily pertains to regulating interactions between the substrate-binding domain (SBD) and substrates rather than the inherent properties of the two ATPase domains.

With all these analyses, we have observed that *On*-Hsp70cBi and *On*-Hsp70cBc share a high degree of similarities. Further, the Hsp70 protein is the umbrella of molecular chaperones and folding catalysts that aid a wide range of protein folding activities within the cell, which is coupled with the action of other chaperones like the Hsp70cBc [34]. These proteins and other chaperones constitute a complex network of folding machines that are needed to maintain cellular protein homeostasis. To enhance our understanding of these important proteins, it is recommended to conduct a structure-based functional analysis detailing functionally significant residues, motifs, or pockets as well as a more thorough analysis of each protein structure level. Moreover, the researchers also suggest integrating computational predictions with experimental evidence for further functional investigations.

5. Conclusions

To gain a deeper understanding of the structural and functional biology of inducible and constitutive proteins in Nile tilapia, various computational tools were employed. These tools were used to analyze the physical and chemical properties, create accurate models of both proteins and predict their functions by comparing their structural features to those of other known proteins. Notably, the study highlighted that the generated 3D models from the study are highly reliable. Significant structural resemblance to Hsp70cBi/Hsp70cBc proteins from various animal species that are taxonomically distant also supports the idea of a remarkably high degree of evolutionary conservation within this protein family. Also, the unique presence of BAG3 as a functional partner of *On-Hsp70cBc* protein revealed additional potential diversity in the regulation of Hsc70/Hsp70cBi chaperones. Functional annotation based on structural similarity provides reliable indirect evidence of substantial functional conservation of these two proteins in tilapia fish. However, it is worth noting that this method may not be sensitive enough to differentiate between the two isoforms.

In summary, even though the assignment of gene functions based on protein structure similarity is currently limited by the availability of protein structures in the PDB, it shows promise as a rapid, cost-effective, and reasonably reliable approach for assigning gene functions in fish on a large scale.

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Data availability statement

Data included in article/supp. material/referenced in article. Data associated with this study has not been deposited into a publicly available repository other than this.

CRediT authorship contribution statement

Geraldine B. Dayrit: Writing – review & editing, Writing – original draft, Visualization, Validation, Software, Methodology, Formal analysis, Data curation, Conceptualization. **Normela Patricia F. Burigsay:** Writing – original draft, Visualization, Software, Methodology, Data curation. **Emmanuel M. Vera Cruz:** Writing – review & editing, Validation, Methodology. **Mudjekeewis D. Santos:** Writing – review & editing, Validation, Supervision, Project administration, Methodology, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

The authors would like to acknowledge the extended help and expertise of Mr. Jayzon G. Bitacura.

References

[1] H. Volkoff, I. Rønnestad, Effects of temperature on feeding and digestive processes in fish, Temperature 7 (2020) 307–320, https://doi.org/10.1080/23328940.2020.1765950.

[2] R.J. Roberts, C. Agius, C. Saliba, P. Bossier, Y.Y. Sung, Heat shock proteins (chaperones) in fish and shellfish and their potential role in relation to fish health: a review, J. Fish. Dis. 33 (2010) 789–801, https://doi.org/10.1111/j.1365-2761.2010.01183.x.

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

[4] N.T. Tran, I. Jakovlić, W.-M. Wang, In silico characterisation, homology modelling and structure-based functional annotation of blunt snout bream (*Megalobrama amblycephala*) Hsp70 and Hsc70 proteins, J. Anim. Sci. Technol. 57 (2015) 44, https://doi.org/10.1186/s40781-015-0077-x.

[5] C. Zhang, K. Lu, J. Wang, Q. qian, X. Yuan, C. Pu, Molecular cloning, expression HSP70 and its response to bacterial challenge and heat stress in Microptenus salmoides, Fish Physiol. Biochem. 46 (2020) 2389–2402, https://doi.org/10.1007/s10695-020-00883-9.

[6] E. Gasteiger, C. Hoogland, A. Gattiker, S. Duvaud, M.R. Wilkins, R.D. Appel, A. Bairoch, Protein identification and analysis tools on the ExPASy server, in: J. M. Walker (Ed.), Proteomics Protoc. Handb., Humana Press, Totowa, NJ, 2005, pp. 571–607, https://doi.org/10.1385/1-59259-890-0:571.

[7] H. McWilliam, W. Li, M. Uludag, S. Squizzato, Y.M. Park, N. Buso, A.P. Cowley, R. Lopez, Analysis tool web services from the EMBL-EBI, Nucleic Acids Res. 41 (2013) W597–W600, https://doi.org/10.1093/nar/gkt376.

[8] E. Yu, T. Yoshinaga, F.L. Jalufka, H. Ehsan, D.B. Mark Welch, G. Kaneko, The complex evolution of the metazoan HSP70 gene family, Sci. Rep. 11 (1) (2021) 1–10, https://doi.org/10.1038/s41598-021-97192-9.

[9] S. Kumar, G. Stecher, M. Li, C. Knyaz, K. Tamura, Mega X: molecular evolutionary genetics analysis across computing platforms, Mol. Biol. Evol. 35 (6) (2018) 1547–1549, https://doi.org/10.1093/molbev/msy096.

[10] I. Letunic, S. Khedkar, P. Bork, SMART: recent updates, new developments and status in 2020, Nucleic Acids Res. 49 (2021) D458–D460, https://doi.org/10.1093/nar/gkaa937.

[11] J. Schultz, R.R. Copley, T. Doerks, C.P. Ponting, P. Bork, SMART: a web-based tool for the study of genetically mobile domains, Nucleic Acids Res. 28 (2000) 231–234, https://doi.org/10.1093/nar/28.1.231.

[12] K. Arnold, L. Bordoli, J. Kopp, T. Schwede, The SWISS-MODEL workspace: a web-based environment for protein structure homology modelling, Bioinformatics 22 (2006) 195–201, https://doi.org/10.1093/bioinformatics/bti770.

[13] A. Fiser, Template-based protein structure modeling, Methods Mol. Biol. Clifton NJ 673 (2010) 73–94, https://doi.org/10.1007/978-1-60761-842-3_6.

[14] L. Heo, H. Park, C. Seok, GalaxyRefine: protein structure refinement driven by side-chain repacking, Nucleic Acids Res. 41 (2013) W384–W388, https://doi.org/10.1093/nar/gkt458.

[15] M.J. Sippl, Recognition of errors in three-dimensional structures of proteins, Proteins: Struct., Funct., Bioinf. 17 (1993) 355–362, https://doi.org/10.1002/prot.340170404.

[16] J.J. Almagro Armenteros, C.K. Sønderby, S.K. Sønderby, H. Nielsen, O. Winther, DeepLoc: prediction of protein subcellular localization using deep learning, Bioinformatics 33 (2017) 3387–3395, https://doi.org/10.1093/bioinformatics/btx431.

[17] A. Roy, J. Yang, Y. Zhang, COFACTOR: an accurate comparative algorithm for structure-based protein function annotation, Nucleic Acids Res. 40 (2012) W471–W477, https://doi.org/10.1093/nar/gks372.

[18] C. Zhang, P.L. Freddolino, Y. Zhang, COFACTOR: improved protein function prediction by combining structure, sequence and protein–protein interaction information, Nucleic Acids Res. 45 (2017) W291–W299, https://doi.org/10.1093/nar/gkx366.

[19] Y. Zhang, J. Skolnick, TM-align: a protein structure alignment algorithm based on the TM-score, Nucleic Acids Res. 33 (2005) 2302–2309, https://doi.org/10.1093/nar/gki524.

[20] Y. Zhang, I-TASSER server for protein 3D structure prediction, BMC Bioinf. 9 (2008) 40, https://doi.org/10.1186/1471-2105-9-40.

[21] J. Yang, R. Yan, A. Roy, D. Xu, J. Poisson, Y. Zhang, The I-TASSER Suite: protein structure and function prediction, Nat. Methods 12 (2015) 7–8, https://doi.org/10.1038/nmeth.3213.

[22] H.M. Song, X.D. Mu, D.E. Gu, D. Luo, Y.X. Yang, M. Xu, J.R. Luo, J.E. Zhang, Y.C. Hu, Molecular characteristics of the HSP70 gene and its differential expression in female and male golden apple snails (Pomacea canaliculata) under temperature stimulation, Cell Stress Chaperones 19 (4) (2014) 579–589, https://doi.org/10.1007/s12192-013-0485-0.

[23] K. Kastano, P. Mier, M.A. Andrade-Navarro, The role of low complexity regions in protein interaction modes: an illustration in huntingtin, Int. J. Mol. Sci. 22 (2021) 1727, https://doi.org/10.3390/ijms22041727.

[24] M. Kabani, C.N. Martineau, Multiple Hsp70 Isoforms in the Eukaryotic Cytosol: Mere Redundancy or Functional Specificity?, Curr. Genom. 9 (n.d.) 338–348. https://www.eurekaselect.com/article/12297 (accessed January 16, 2024).

[25] D. Muñoz-Espín, G. Serrano-Heras, M. Salas, Chapter 10 - role of host factors in bacteriophage φ29 DNA replication, in: M. Łobocka, W.T. Szybalski (Eds.), Adv. Virus Res., Academic Press, 2012, pp. 351–383, https://doi.org/10.1016/B978-0-12-394621-8.00020-0.

[26] A. Zapun, T. Vernet, M.G. Pinho, The different shapes of cocci, FEMS Microbiol. Rev. 32 (2008) 345–360, https://doi.org/10.1111/j.1574-6976.2007.00098.x.

[27] F. Mayer, Cytoskeletons in prokaryotes, Cell Biol. Int. 27 (2003) 429–438, https://doi.org/10.1016/S1065-6995(03)00035-0.

[28] R. Rai, A.L. Kennedy, Z.R. Isingizwe, P. Javadian, D.M. Benbrook, Similarities and differences of Hsp70, hsc70, Grp78 and mortalin as cancer biomarkers and drug targets, Cells 10 (2021) 2996, https://doi.org/10.3390/cells10112996.

[29] Z. Balogi, G. Multhoff, T.K. Jensen, E. Lloyd-Evans, T. Yamashima, M. Jäättelä, J.L. Harwood, L. Vígh, Hsp70 interactions with membrane lipids regulate cellular functions in health and disease, Prog. Lipid Res. 74 (2019) 18–30, https://doi.org/10.1016/j.plipres.2019.01.004.

[30] S.R. Bonam, M. Ruff, S. Muller, HSPA8/HSC70 in immune disorders: a molecular rheostat that adjusts chaperone-mediated autophagy substrates, Cells 8 (2019) 849, https://doi.org/10.3390/cells8080849.

[31] K. Tekirdag, A.M. Cuervo, Chaperone-mediated autophagy and endosomal microautophagy: jointed by a chaperone, J. Biol. Chem. 293 (2018) 5414–5424, https://doi.org/10.1074/jbc.R117.818237.

[32] C.-L. Tsai, A.-S. Chao, S.-M. Jung, C.-Y. Lin, A. Chao, T.-H. Wang, Stress-induced phosphoprotein 1 acts as a scaffold protein for glycogen synthase kinase-3 beta-mediated phosphorylation of lysine-specific demethylase 1, Oncogenesis 7 (2018) 1–17, https://doi.org/10.1038/s41389-018-0040-z.

[33] D. Sharma, D.C. Masison, Hsp70 Structure, Function, Regulation and Influence on Yeast Prions, Protein Pept. Letture. 16 (n.d.) 571–581. https://www.eurekaselect.com/article/29549 (accessed January 16, 2024).

[34] M.P. Mayer, B. Bukau, Hsp70 chaperones: cellular functions and molecular mechanism, Cell. Mol. Life Sci. 62 (2005) 670, https://doi.org/10.1007/s00018-004-4464-6.