IJPSR (2022), Volume 13, Issue 10 (Research Article)

INTERNATIONAL JOURNAL OF PHARMACEUTICAL SCIENCES AND RESEARCH

Received on 25 February 2022; received in revised form, 15 June 2022; accepted 20 June 2022; published 01 October 2022

IN-SILICO HOMOLOGY MODELING OF CRUSTIN PROTEIN IN *PORTUNUS TRITUBERCULATUS* CRAB SPECIES

S. Moganapriya and K. Shoba *

Department of Biochemistry and Biotechnology, D. K. M. College for Women (Autonomous), Vellore - 632001, Tamil Nadu, India.

Keywords:

Crustin, *Portunus trituberculatus*, NCBI, Protparam and helix turn helix tool and cph, Discovery studio software

Correspondence to Author:
Dr. K. Shoba

Assistant Professor,
Department of Biochemistry and Biotechnology, D. K. M. College for Women (Autonomous), Vellore - 632001, Tamilnadu, India.

E-mail: danishoba@gmail.com

ABSTRACT: The aim of the present *In-silico* studies is to find out the potential protein target and find out the 3D structure of crustin antimicrobial protein in *Portunus trituberculatus*. Crustin are small cationic proteins assumed to be antimicrobial effectors against mainly Gram-positive bacteria. The swimming crab *Portunus trituberculatus* is a commercially important crab species in East Asia countries. Gonadal development is a physiological process of great significance to reproduction and commercial seed production for *P. trituberculatus.* The nucleotide and protein sequence of crustins protein from *P. trituberculatus* is retrieved from the NCBI database in Fasta format. Further, the primary and functional analysis of crustin protein was done through protparam and helix turn helix tools. The retrieved protein sequence was applied to an advanced protein modeling server Cph to deliver the 3D structure. After modeling, the protein's three Dimensional structure was viewed with Discovery studio software's help. Our method enables the production of new antimicrobial peptides as potential next-generation antibiotics.

INTRODUCTION: Crabs are decapod crustaceans of the *Brachyura infraorder*, which usually have a very small protruding "tail" (abdomen) (Greek: β) (Romanized: Brachys = short,) usually concealed completely under the thorax [1]. They live in all the deep sea, in freshwater, and on land, typically covered with a dense exoskeleton with a single pair of pincers. Southern Europe's river crab (Lenten crab, *Potamon fluviatile*) is an example of the numerous freshwater crabs in much of the colder regions of the country [2].

QUICK RESPONSE CODE

DOI: 10.13040/IJPSR.0975-8232.13(10).4239-43

This article can be accessed online on www.ijpsr.com

DOI link: http://dx.doi.org/10.13040/IJPSR.0975-8232.13(10).4239-43

As a rule, crabs breathe gills, which lie in a pair of cavities under the sides of the carapace, but in real land crabs the cavities are widened and modified to serve as lungs for breathing oxygen [3, 4]. The thick exoskeleton primarily consists of finely mineralized chitin. *Portunus trituberculatus* (Crustacea: Decapoda: Brachyura), commonly known as the swimming crab, is widely distributed in the coastal waters of Korea, Japan, China, and Southeast Asia.

This species inhabits estuaries and coastal waters, which belong to typical euryhaline crab species. In China, it is a major edible crab species and one of the most important fishery resources and its production has now reached 92, 907 tons 2011 [5]. At present, commercial crab farming largely depends on wild seed stock and the commercial characteristics (growth rate, flesh quality and

disease resistance) of the cultured stocks have also declined after many years of culturing and wild populations of *Portunus trituberculatus* have dramatically declined for the last decades due to over-exploitation and the deterioration of environmental conditions in China. *Portunus trituberculatus* AMPs are often described as small, cationic, amphipathic molecules encoded by a single gene, as is the case in many other animals [6]. However, this classification has recently been updated to include less common AMPs such as anionic peptides. These multifunctional proteins play a substantial role in other cellular processes and antimicrobial protein-derived fragments [7].

Surprisingly, many gene-encoded AMPs in *Portunus trituberculatus* are made up of distinct structural domains [8]. Each of these domains has unique characteristics observed in different AMP groups, such as the presence of cysteine residues that form disulfide bonds or the overrepresentation of particular amino acids. In addition to their involvement in innate immunity [9], it has been postulated that these chimeric peptides might behave as multifunctional proteins in other physiological systems (Bachère *et al.,* 2004). Based on amino acid makeup and structure, we divided the antimicrobial peptide families discovered in crab species into four major categories: (1) single-domain linear-helical AMPs and peptides enriched in specific amino acids, (2) single-domain peptides containing cysteine residues engaged in disulfide bonds, (3) multi-domain or chimeric AMPs and (4) unconventional AMPs, which included multi-functional proteins and protein-derived fragments with antimicrobial activity [10].

Crustins are cationic antibacterial polypeptides with many domains (7-14 kDa) and a single whey acidic protein (WAP) domain at the C-terminus (Smith *et al.,* 2008). The first crustin member discovered was an 11.5-kDa protein isolated from the granular hemocytes of the beach crab *C. maenas*, which has a strong antibacterial effect against Gram-positive marine or salt-tolerant bacteria (Relf *et al.,* 1999). Later, Bartlett *et al*. (2002) proposed the term crustin to describe transcripts found in two *Penaeid shrimp* species (L. vannamei and L. setiferus) with significant sequence resemblance to the crab 11.5-kDa protein (carcinin) (Brockton *et al.,* 2007).

EST-based approaches have identified over 50 crustins and crustin-like sequences in many *Portunus trituberculatus* species [11], including crayfish, shrimp, freshwater prawns, crabs and lobsters, and non-decapod *Portunus trituberculatus* such as amphipods (Smith *et al.,* 2008).

MATERIALS AND METHODS:

Protein Sequence Retrieval System: The protein crustin sequence was retrieved from NCBI in FASTA format to perform protein modeling studies.

Primary and Functional Prediction: The retrieved protein sequence was applied to protparam and advanced HTH (Helix –Trun-Helix) motif sequence regions to find out the primary protein sequence analysis and functional part of the sequences [12].

Protein Modeling and Visualization: The protein sequence was applied to an advanced protein modeling server Cph server to deliver the 3D structure [13]. After modeling, the 3 Dimensional protein was viewed with the help of Discovery studio software.

RESULTS:

Protein: >AFU61590.1 crustin 2 *Portunus trituberculatus*

MQNRVVCLMVVMAVAMSVASASSCSTFCK
DPYLNIQGEYVCCDKNPGTCPERDECPPLAQ
EDVRQGIRFCHYDPECHPNEKCCFDICIKQKV
CKLADP

Nucleotide: >JQ728435.1:86-382 *Portunus trituberculatus* crustin 2 mRNA, complete cds

ATGCAGAACCGAGTCGTGTGCCTGATGGTG
GTGATGGCGGTGGCTATGTCCGTTGCCAGC
GCCTCCTCCTGCAGCACCTTTTGCAAGGAC
CCGTACCTTAACATCCAGGGAGAGTACGTG
TGTTGTGACAAAAATCCCGGCACATGCCCC
GAACGAGATGAGTGTCCCCCGCTCGCGCAA
GAGGATGTTCGCCAAGGAATCAGGTTTTGC
CACTACGATCCAGAGTGTCACCCAAATGAG
AAATGCTGCTTTGATATCTGCATCAAACAG
AAAGTGTGCAAACTTGCCGATCCCTAA

Primary Analysis –Protparam:

User-provided Sequence: 10 20 30 40 50 60
MQNRVVCLMV VMAVAMSVAS
ASSCSTFCKD PYLNIQGEYV CCDKNPGTCP

ERDECPPLAQ 70 80 90 EDVRQGIRFC HYDPECHPNE KCCFDICIKQ KVCKLADP

Number of Amino Acids: 98

Molecular Weight: 10969.74

Theoretical pI: 5.23

Amino Acid Composition: Ala (A) 6 6.1% Arg (R) 4 4.1% Asn (N) 4 4.1% Asp (D) 7 7.1% Cys (C) 13 13.3% Gln (Q) 5 5.1% Glu (E) 6 6.1% Gly (G) 3 3.1% His (H) 2 2.0% Ile (I) 4 4.1% Leu (L) 4 4.1% Lys (K) 6 6.1% Met (M) 4 4.1% Phe (F) 3 3.1% Pro (P) 8 8.2% Ser (S) 5 5.1% Thr (T) 2 2.0% Trp (W) 0 0.0% Tyr (Y) 3 3.1% Val (V) 9 9.2% Pyl (O) 0 0.0% Sec (U) 0 0.0% (B) 0 0.0% (Z) 0 0.0% (X) 0 0.0%

Total Number of Negatively Charged Residues (Asp + Glu): 13

Total Number of Positively Charged Residues (Arg + Lys): 10

Atomic Composition: Carbon C 464 Hydrogen H 735 Nitrogen N 129 Oxygen O 144 Sulfur S 17

Formula: $C_{464}H_{735}N_{129}O_{144}S_{17}$

Total Number of Atoms: 1489

Extinction Coefficients: This protein does not contain any Trp residues. Experience shows that this could result in more than 10% error in the computed extinction coefficient. Extinction coefficients are in units of M^{-1} cm^{-1}, at 280 nm measured in water. Ext. coefficient 5220

Abs 0.1% (=1 g/l) 0.476, assuming all pairs of Cys residues form cystines

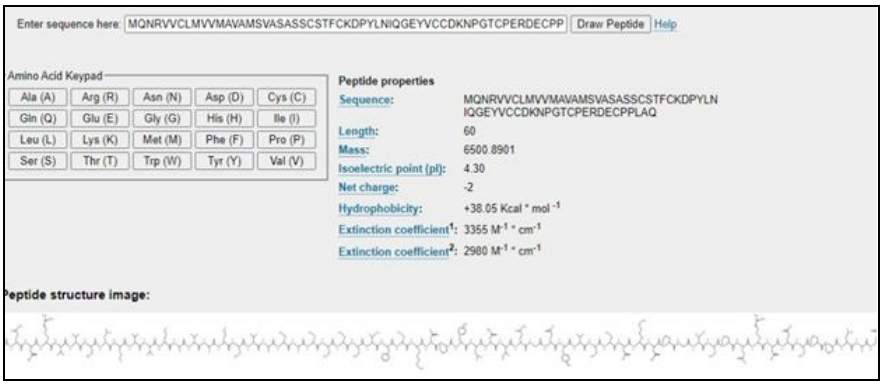
Ext. coefficient 4470 Abs 0.1% (=1 g/l) 0.407, assuming all Cys residues are reduced

Estimated Half-life: The sequence's N-terminal is considered M (Met). The estimated half-life is: 30 h (mammalian reticulocytes *in vitro*). >20 h (yeast, *in-vivo*). >10 h (Escherichia coli, *in-vivo*).

Instability Index: The instability index (II) is computed to be 63.95. This classifies the protein as unstable.

Aliphatic Index: 64.59

Grand Average of Hydropathicity (GRAVY): -0.182 Pepdraw.



Helix Turn – Helix:
Motif Prediction Server:

FIG. 1: PRIMARY AND MOTIF ANALYSIS

Protein Modelling: The Identified motif sequence was applied into the Protein Modelling Server. (The amino acids Protein sequence was converted into 3D structure). The predicted 3 Dimensional structure was viewed with the help of advanced molecular visualization tools such as Discovery Studio Software.

Discovery Studio Software 3D Structure of Protein Crustin: Hydrophobicyellow1ala, val, phe, pro, met, ile, leupolarpinkser, thr, tyr, his, cys, cyss, asn, gln, trp, glycharged (+) blue lys, arg charged (-) redasp, glu.

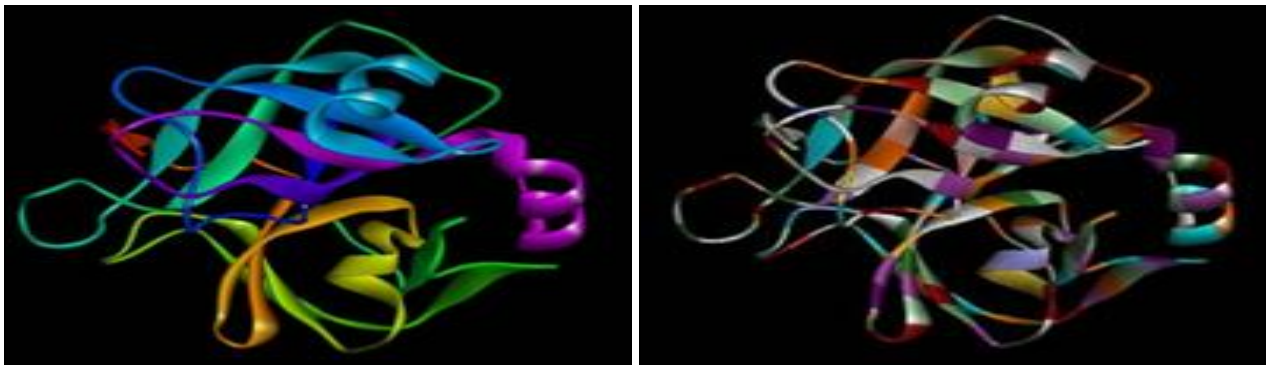


FIG. 2: CPH SERVER

DISCUSSION: This study aims to use bioinformatics methods to predict protein structure based on motifs. To determine the functional region of the protein sequence, we downloaded the crustin protein sequence and ran it through the GYM motif server. One of the most well-studied patterns in proteins is the Helix-Turn-Helix (HTH) motif. Transcription factors are usually proteins containing such patterns [14].

They attach to DNA and interfere with RNA polymerase's action, hence controlling gene expression. The HTH motifs of these proteins are found to be responsible for DNA binding. The GYM motif prediction sequence is shown. Detecting motifs, such as the HTH motif, has become a major problem in biochemistry. Statistically based profile approaches are the most extensively employed, with the Dodd & Egan (DE) method being the most widely used for detecting the HTH pattern. The programme also compares and contrasts these two ways. Cph model server was used to estimate the 3D structure. The 3D structure of the crustin protein sequence was created. We have provided notes on the applications of the CPh model server in Figure 2. CPHmodels-3.0 is a web server that uses single template homology modelling to predict protein 3D structure. The server combines CPHmodels-2.0's scoring capabilities with a revolutionary remote homology-modeling approach. The rapid CPHmodels-2.0 profile-profile scoring function, ideal for near homology modelling, is used to first model a query sequence. The new computationally expensive distant homology modeling approach is used if no acceptable PDB template is found in the initial search. In the CASP8 competition, CPHmodels-3.0 delivered models for 94 percent of the targets (117 out of 128) with 74 percent predicted as high reliability models (87 out of 117). These had an RMSD of 4.6 on average. When overlaid on a three-dimensional structure. With an average RMSD of 9.3, the remaining 26% low-reliability models (30 out of 117) could superimpose to the genuine 3D structure [15]. The CPHmodels-3.0 technique joins the ranks of high-performing 3D-prediction tools thanks to these results. Aside from its accuracy, one of the method's most essential advantages is its quickness. The server's response time for most queries is less than 20 minutes. All of the above findings were discussed.

CONCLUSION: Antimicrobial peptides AMPs (Crustin) are a class of chemicals produced by the host immune system that has a role in demonstrating antimicrobial action against invading pathogens. Crustin has been reported to have an immunological function to prevent disease in the aquaculture business. If an empirical 3D "template" structure with >30 percent sequence identity is known, comparative ("homology") modelling approximates the 3D structure of a target protein for which only the sequence is given. When

the target and template are closely related, homology modelling can create high-quality structural models, which has prompted the development of a structural genomics cooperative committed to producing representative experimental structures for all classes of protein folds. In the crustin protein, we use a motif sequence to predict structure (*Portunus trituberculatus*). A comprehensive analysis of protein could be employed in future studies.

ACKNOWLEDGEMENT: I am profusely thankful to Dr. P. N. Sudha, Principal, Department of Chemistry, DKM College for Women (Autonomous), for the valuable suggestion, guidance and support.

CONFLICTS OF INTEREST: We declare that we have no conflicts of interest.

REFERENCES:

1. Janeway CA and Medzhitov R: Innate immune recognition. Annu Rev Immunol 2002; 20: 197–216.
2. Zhang, Liu Y, Song C, Ning J and Cui ZM: Characterization and functional analysis of a novel mannose-binding lectin from the swimming crab *Portunus trituberculatus*. Fish Shellfish Immunol 2019; 89: 448–57.
3. Wang XW, Vasta GR and Wang JX: The functional relevance of shrimp C-type lectins in host-pathogen interactions. Dev. Comp Immunol 2020; 109: 103708.
4. Kwankaew P, Praparatana R, Runsaeng P and Utarabhand P: An alternative function of C-type lectin comprising low-density lipoprotein receptor domain from *Fenneropenaeus merguiensis* to act as a binding receptor for viral protein and vitellogenin. Fish and Shellfish Immunol 2018; 74: 295–308.
5. Qin Y, Jiang S, Huang J, Zhou F, Yang Q, Jiang S and Yang L: C-type lectin response to bacterial infection and ammonia nitrogen stress in tiger shrimp (*Penaeus monodon*). Fish Shellfish Immunol 2019; 90: 188–198.
6. Liu H, Liu Y and Cui Z: Functional characterization of two clip-domain serine proteases in the swimming crab *Portunus trituberculatus*. Fish Shellfish Immunol 2019; 89: 98–107.
7. Zhang XW, Man X, Huang X, Wang Y, Song QS, Hui KM and Zhang HW: Identification of a C-type lectin possessing both antibacterial and antiviral activities from red swamp crayfish. Fish Shellfish Immunol 2018; 77: 22–30.
8. Shoba K, Manjuladevi M and Mazher Sultana: Biochemical analysis and gene expression profiling on collagenase protein in fiddler crab, World Journal of Pharmacy and Pharmaceutical Sciences 2278 –4357, 6, Issue 3, 747-756
9. Shoba K, Sowmiya S and Mazher Sultana: World Journal of Pharmaceutical and Life Sciences, ISSN 2454-2229, Vol. 3, Issue 1, 427-436. 10.
10. Shoba K and Kalpana K: Protein modeling and drug docking studies on potential protein target (e. coli–dosp) and compound aldehyde (sumatriptan) using bioinformatics tools. Research & Reviews A Journal of Bioinformatics 2018; 5(3): 9–18.
11. Bulet P, Hetru C and Dimarcq JL: Antimicrobial peptides in insects: structure and function. Dev Comp Immunol 1999; 23: 329–344.
12. Velayutham M, Kamanuri SK and Saravanan K: Cation metals specific hemocyanin exhibits differential antibacterial property in mud crab, Scylla serrata. Biologia 2016; 71(2): 176–183.
13. Shoba K and Mazher Sultana: Three - dimensional structure and motif prediction studies on collagenase protein in fiddler crab. International Journal of Novel Trends in Pharmaceutical Sciences Issn: 2277 – 2782, 6(4): 79-83.
14. Shoba K and Lavanya G: Identification of de novo peptide and motif prediction on porphyria protein (HMBS) using *in-silico* tools. Universal Journal of Pharmacy 2018; 8(1): Issn 2320-303.
15. Su Y, Liu Y, Gao F and Cui Z: A novel C-type lectin with a YPD motif from *Portunus trituberculatus* (PtCLec1) mediating pathogen recognition and opsonization. Dev. Comp. Immunol 2020; 106: 103609.

How to cite this article:
Moganapriya S and Shoba K: *In-silico* homology modeling of crustin protein in *Portunus trituberculatus* crab species. Int J Pharm Sci & Res 2022; 13(10): 4239-43. doi: 10.13040/IJPSR.0975-8232.13(10).4239-43.