Molecular Cloning of *clpX* Gene from *Vibrio alginolyticus* HY9901 and Its Bioinformatics Analysis

Xiaoxin WEN[1,2], Yuyan HE[1,2], Jiajie MA[1,2], Weijie ZHANG[1,2], Huanying PANG[1,2*], Na WANG[3]

1. Fisheries College, Guangdong Ocean University, Zhanjiang 524088, China; 2. Guangdong Provincial Key Laboratory of Aquatic Animal Disease Control and Healthy Culture & Key Laboratory of Control for Diseases of Aquatic Economic Animals of Guangdong Higher Education Institutes, Zhanjiang 524088, China; 3. Chinese Academy of Inspection and Quarantine, Beijing 100176, China

Abstract According to the *clpX* gene sequence of *Vibrio alginolyticus* HY9901, a pair of specific primers were designed, and the full length was cloned by PCR and subjected to bioinformatics analysis. The results showed that the *clpX* gene was 1 281 bp in length and encoded 426 amino acids. Its molecular structure formula was $C_{3\,842}H_{6\,405}N_{1\,281}O_{1\,598}S_{260}$, with a theoretical protein molecular weight of approximately 1 044 473.4 kDa and a theoretical pI value of 5.04. The *clpX* gene was predominantly situated within the cytoplasm, exhibiting unstable and hydrophilic protein characteristics. It possessed a signal peptide cleavage site, lacked a transmembrane region, and was not associated with any KEGG metabolic pathway. Additionally, it possessed 2 glycine phosphorylation sites, a CAMP-dependent protein kinase phosphorylation site, a C-terminal amidation modification site, 6 protein kinase C phosphorylation sites, 7 microbody C-terminal target signal sites, and an ATP/GTP site. The *clpX* phylogenetic tree was constructed using the MEGA 5.0 software via the neighbor-joining method. The results demonstrated that the *clpX* of *V. alginolyticus* exhibited up to 100% affinity with the *clpX* of *Vibrio* spp. The single subunit 3D structure model of the ClpX protein was obtained using the SWISS-MODEL program. A structural and functional analysis of the protein revealed the presence of three distinct ClpX structural and functional domains. In the prediction of secondary structure, the proportions of α-helix, random coil, β-sheet and extended strand were 40.38%, 37.09%, 5.40% and 17.14%, respectively. The analysis of the ClpX protein through the STRING database revealed that the proteins interacting with the ClpX protein were Tig, Atpd, Hflb, Msrb-2, Rpod, Clpp, Clpa, Lon-1, Hfq, and ANP63951.1. A computational analysis of the ClpX protein identified a number of post-translational modification sites, including phosphorylation, acetylation, ubiquitination, glycosylation, methylation, S-palmitoylation, and lactylation. The significance of this study is to analyze the function of the *clpX* gene and establish a robust foundation for subsequent investigations into the mechanism of the *clpX* gene in *Vibrio alginolyticus*.

Key words *Vibrio alginolyticus*, *clpX* gene, Bioinformatics analysis

1 Introduction

Vibrio alginolyticus, a member of the genus *Vibrio* and family Vibrionaceae, is a gram-negative anaerobic bacterium with a terminal flagellum and motile characteristics. It is a common pathogen in the process of mariculture. *V. alginolyticus* is one of the most prevalent bacterial diseases[1], with a wide distribution in marine and estuarine environments and other water bodies. It has a significant impact on marine organisms. In the temperature range of 25 - 35 ℃, *V. alginolyticus* is a significant contributor to the development of serious vibriosis in fish, shrimp, shellfish, and other economically valuable aquatic animals[2]. *V. alginolyticus* has been associated with a range of human diseases, including diarrhea, otitis media, enteritis, and others. The pathogenicity of *V. alginolyticus* is contingent upon its interaction with the host. The invasive and proliferative processes of *V. alginolyticus* on the host cause damage to the host cell's tissues and disrupt the host cell's metabolism. A variety of virulence factors have been demonstrated to be associated with the pathogenicity of *V. alginolyticus*, including lipopolysaccharides[3], extracellular products[4], biofilms[5], adhesion factors[6], the ferric uptake system[7], the group sensing system[8], and the type III secretion system[9].

ClpX is an AAA ATPase that is responsible for distinguishing potential substrates. This is achieved by recognizing short motifs at the C-terminus of proteins and delivering the substrate to the ClpP protease for degradation. Among the Clp ATPases, ClpX is the smallest and simplest of the ATPases, consisting of an ATPase hexamer (ClpX) and 2 peptidase heptamers (ClpP)[10]. ClpX is approximately 45 kDa in size and contains a highly conserved zinc-binding structural domain at the N-terminus, an AAA + ATPase structural domain with Walker A and B motifs in the middle, and a tripeptide IGF motif that is necessary for interaction with ClpP. The biochemical functions of ClpX include the binding of substrates, junctions, and ClpP, as well as protein unfolding and peptide translocation. The unfolding and translocation of substrates necessitate the binding and hydrolysis of ATP, which drive conformational changes in the enzymes that drive these mechanical processes. In gram-positive bacteria, the Clp complex plays a cen-

Received: March 15, 2024 Accepted: July 28, 2024

Supported by National Natural Science Foundation of China (32073015); Graduate Education Innovation Program of Guangdong Province (YJYH[2022] 1); Undergraduate Innovation and Entrepreneurship Training Program of Guangdong Ocean University (CXXL2024007); Undergraduate Innovation Team of Guangdong Ocean University (CCTD201802).

Xiaoxin WEN, master candidate, research fields: prevention and control of aquatic animal diseases.

* Corresponding author. Huanying PANG, PhD., associate professor, research fields: prevention and control of aquatic animal diseases.

tral role in regulating the production of virulence factors. In contrast, in gram-negative bacteria, conserved proteases typically contribute to pathogenesis by influencing biological processes within the bacterial envelope.

In recent years, there has been a proliferation of studies on the *clpX* gene in a number of bacterial species, including *Staphylococcus aureus*[11], *Mycobacterium tuberculosis*[12], *Streptococcus mutans*[13], and *Chlamydia trachomatis*[14]. However, there has been a paucity of research on the *clpX* gene in *V. alginolyticus*. In order to further elucidate the mechanism of action of the *clpX* gene in *V. alginolyticus*, this study cloned the *clpX* gene of *V. alginolyticus* and conducted a bioinformatic analysis of its sequence. This lays the foundation for the subsequent study on the mechanism of virulence regulation of its protein in *V. alginolyticus*.

2 Materials and methods

2.1 Materials

2.1.1 Strains and vectors. The virulent strain of *V. alginolyticus* HY9901[15] was preserved by the Guangdong Provincial Key Laboratory of Aquatic Animal Disease Control and Healthy Culture. The *Escherichia coli* DH5α competence was maintained by the same laboratory. The cloning vector pMD18-T was procured from Takara Biotechnology (Dalian) Co., Ltd.

2.1.2 Reagents. The Ex*Taq* DNA polymerase was procured from Takara Biotechnology (Dalian) Co., Ltd. The bacterial genomic DNA extraction kit and DNA gel recovery kit were purchased from Beijing TransGen Biotech Co., Ltd. The remaining reagents were imported or domestically produced and were of an analytically pure quality. The PCR primers were synthesized and sequenced by Invitrogen Shanghai Trading Co., Ltd. The antibiotics were procured from Sigma.

2.2 Methods

2.2.1 Extraction of total DNA from *V. alginolyticus* HY9901. The methodology proposed by Zeng Fuyuan[16] was employed, with genomic DNA extracted in accordance with the instructions provided by the kit and stored at a temperature of −20 ℃.

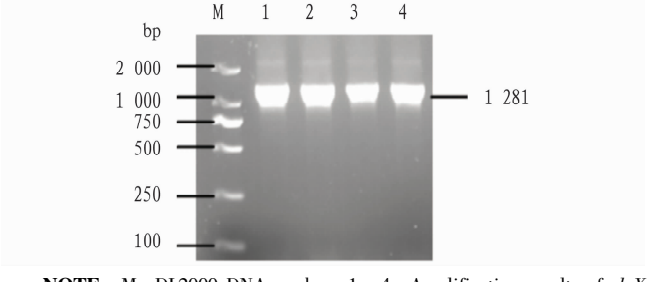
2.2.2 Cloning of *clpX* gene. A pair of primers was designed according to the sequence of the *clpX* gene from *V. alginolyticus*. The upstream primer, designated P1, was ATGACAGATAAAAGCAAAGAAAGT, while the downstream primer, designated P2, was TTACTCTGCTCCAGCTGCCT. A polymerase chain reaction (PCR) was conducted using the total DNA extracted from *V. alginolyticus* HY9901 as the template. The reaction was conducted under the following conditions: pre-denaturation at 98 ℃ for 3 min; denaturation at 98 ℃ for 30 sec, annealing at 54 ℃ for 30 sec, and extension at 72 ℃ for 40 sec, 33 cycles; and extension at 72 ℃ for 5 min. The gel was cut and recovered using a gel-cutting kit.

2.2.3 Sequencing of PCR products. The PCR product was ligated to the pMD18-T vector and subsequently transformed into DH5α receptor cells in accordance with the procedure outlined in the manual, to promote its entry and stable presence in these cells. The colonies were coated on LB plates containing Kna$^+$ resistance in order to select those that had successfully undergone transformation and carried the target plasmid. Subsequently, the colonies were dispatched to Guangzhou Sangon Biotech Co., Ltd. for analysis.

2.2.4 Bioinformatics analysis of *clpX* gene from *V. alginolyticus* HY9901. The methodology employed in this experiment was in accordance with that proposed by Pang Huanying[17]. Sequence homology comparison and similarity analysis were conducted via the NCBI website. Physicochemical properties were analyzed using the ExPASy Proteomics Server. Nucleic acid homology comparison analysis was performed using DNAMAN Version 6.0. Open reading frames were determined using ORF Finder. The signal peptide sequences were predicted by online analysis software SignalP 5.0 Server. The functional site distribution in amino acid sequences was predicted by SoftBerry-Psite. The transmembrane structural domains were predicted by using TMHMM Server 2.0. The lactation modification site was predicted by using the Kla website. The protein structure and functional domains were analyzed using the SMART website. The subcellular localization was predicted using the Gneg-PLoc software. Phylogenetic trees were constructed using the Clastal 2.0 and MEGA 5.0 software. The protein modeling was carried out using the SWISS-MODEL program of the ExPASy server. The metabolic pathways were searched using the KEGG website. The protein-protein network interactions were searched in the STRING database. Finally, the PTM modification sites were predicted using the musite website.

3 Results and analysis

3.1 Gene amplification A specific band of approximately 1 281 bp was successfully obtained by PCR (Fig. 1), which encoded 426 amino acids (NCBI accession No.: PP526748.1).



NOTE M. DL2000 DNA marker; 1 − 4. Amplification results of *clpX* gene.

Fig. 1 Cloning of *clpX* gene

3.2 Physicochemical properties of *clpX* The results indicate that the *clpX* had a total of 13 386 atoms, with a molecular structure formula of $C_{3\,842}H_{6\,405}N_{1\,281}O_{1\,598}S_{260}$, a theoretical protein molecular weight of 104 447.34 kDa, and a theoretical pI value of 5.04. The ClpX protein was composed exclusively of alanine (Ala, 30.3%), glycine (Gly, 24.7%), cysteine (Cys, 20.3%), and threonine (Thr, 24.7%), with no negatively charged residues (Asp + Glu) or positively charged residues (Arg + Lys). The protein had an extinction coefficient of 9 970 mol/cm at 280 nm. The half-life of expression in yeast and *E. coli* was grea-

ter than 20 and 10 h, respectively, while that in mammalian reticulocytes in *in vitro* culture was found to be 4.4 h. ClpX is a protein with a high degree of instability, as indicated by an instability index (II) of 40.50 and a lipid index of 30.29. The mean value of hydrophilicity (GRAVY) was 0.781, indicating that the protein is generally hydrophilic.

This evidence indicates that the HY9901 ClpX protein is a relatively large, unstable hydrophilic protein composed of specific amino acids with specific pI values, extinction coefficients, and half-lives. The protein's composition of only four amino acids suggests the possibility of a unique biological function or structural feature. The absence of charged residues may be indicative of the protein having no significant role in charge regulation or ion binding. These properties may be closely related to the specific intracellular functions of the protein, and further studies are needed to reveal its specific mechanism of action.

3.3 Sequence analysis of *clpX* The N-terminal signal peptide structure of the *clpX* amino acid sequence was predicted using the SignalP 50 Server program, which revealed the presence of a signal peptide cleavage site and the presence of a signal peptide. The probability of the signal peptide being a Tat signal peptide was 47.72%, the probability of it being a Sec signal peptide was 43.91%, and there was also a probability of it being some other signal peptide of 8.15%. The protein was found to lack a transmembrane region, as predicted by the TMHMM Server 2.0 program. Subcellular localization was performed through the GnegmPLoc server (URL: sjtu.edu.cn), which predicted that the *clpX* gene was located in the cytoplasm. The SoftBerry-Psite program predicted the presence of 2 glycine sites (8-11 aa, 205-208 aa), a CAMP-dependent protein kinase site (59-62 aa), 6 protein kinase C functional sites (2-4 aa, 10-12 aa, 67-69 aa, 125-127 aa, 192-194 aa, 242-244 aa), a C-terminal amidation site (228-231 aa), 7 microsomal C-terminal target signaling sites (11-13 aa, 96-97 aa, 135-137 aa, 176-178 aa, 243-245 aa, 308-310 aa, 396-398 aa), and an ATP_GTP site (121-128 aa) in the amino acid sequence (Fig.2).

```
1    ATGACAGATAAAAGCAAAGAAAGTGGCAGCGGTAAATTGCTGTACTGTTCTTTCTGCGGC
1     M  T  D  K  S  K  E  S  G  S  G  K  L  L  Y  C  S  F  C  G
61   AAAAGTCAGCACGAAGTTCGCAAGCTAATCGCAGGTCCGTCAGTGTACATTTGCGACGAA
21    K  S  Q  H  E  V  R     L  I  A  G  P  S  V  Y  I  C  D  E
121  TGTGTCGACTTATGTAACGATATTATTCGCGAAGAAATCAAGGATGTTCTCCCTAAGAAA
41    C  V  D  L  C  N  D  I  I  R  E  E  I  K  D  V  L  P  K  K
181  GAATCTGAAGCGTTACCAACGCCAAAACAGATCCGTGAACACCTTGATGACTATGTGATC
61    E  S  E  A  L  P  T  P  K  Q  I  R  E  H  L  D  D  Y  V  I
241  GGGCAAGATTACGCGAAAAAAGTGCTCGCAGTTGCGGTATATAACCACTACAAGCGTTTA
81    G  Q  D  Y  A  K  K  V  L  A  V  A  V  Y  N  H  Y  K  R  L
301  CGCAATGGTGATACAACGAGCGAAGGTGTGGAGCTTGGTAAGAGTAACATCCTTCTAATT
101   R  N  G  D  T  T  S  E  G  V  E  L  G  K  S  N  I  L  L  I
361  GGTCCTACAGGTAGCGGTAAAACACTACTTGCGGAGACATTAGCTCGATTCTTGGATGTG
121   G  P  T  G  S  G  K  T  L  L  A  E  T  L  A  R  F  L  D  V
421  CCATTTACAATGGCAGACGCGACGACACTAACCGAAGCCGGTTACGTAGGTGAAGACGTT
141   P  F  T  M  A  D  A  T  T  L  T  E  A  G  Y  V  G  E  D  V
481  GAAAACATCATCCAAAAGCTTCTGCAAAAATGTGATTACGATGTAGCTAAAGCGGAACGC
161   E  N  I  I  Q  K  L  L  Q  K  C  D  Y  D  V  A  K  A  E  R
541  GGCATTGTTTACATTGATGAAATTGATAAGATTTCACGCAAAGCTGAAAACCCATCTATT
181   G  I  V  Y  I  D  E  I  D  K  I  S  R  K  A  E  N  P  S  I
601  ACGCGTGACGTATCTGGTGAAGGTGTTCAGCAGGCGCTACTAAAACTTATCGAAGGTACG
201   T  R  D  V  S  G  E  G  V  Q  Q  A  L  L  K  L  I  E  G  T

661  GTTGCATCGGTTCCACCTCAAGGTGGTCGTAAGCATCCTCAGCAAGAATTCCTACAAGTG
221   V  A  S  V  P  P  Q  G  G  R  K  H  P  Q  Q  E  F  L  Q  V
721  GACACATCTAAGATCCTGTTCATCTGTGGTGGTGCTTTTGCTGGCTTAGACAAAGTGATT
241   D  T  S  K  I  L  F  I  C  G  G  A  F  A  G  L  D  K  V  I
781  GAGCAGCGTGTTGCAACAGGTACAGGTATTGGTTTTGGTGCTGAAGTTCGTTCGAAGAAT
261   E  Q  R  V  A  T  G  T  G  I  G  F  G  A  E  V  R  S  K  N
841  GAATCGAAGACAGTAGGTGAGCTATTTACTCAAGTTGAACCTGAAGACTTAGTGAAGTAC
281   E  S  K  T  V  G  E  L  F  T  Q  V  E  P  E  D  L  V  K  Y
901  GGCCTCATTCCAGAATTCATTGGTCGTCTACCTGTAACAACAACACTGACTGAGCTTGAT
301   G  L  I  P  E  F  I  G  R  L  P  V  T  T  T  L  T  E  L  D
961  GAAGAAGCGTTGATTCAGATCCTATGTGAACCGAAGAACGCACTGACCAAGCAATATGCA
321   E  E  A  L  I  Q  I  L  C  E  P  K  N  A  L  T  K  Q  Y  A
1021 GCATTGTTTGAGCTAGAAAACGCCGAACTTGAATTCCGTGAAGATGCGCTGCGCGCTATC
341   A  L  F  E  L  E  N  A  E  L  E  F  R  E  D  A  L  R  A  I
1081 GCGAAAAAAGCAATGGAGCGTAAAACAGGTGCTCGTGGCTTACGCTCGATTCTTGAAGGC
361   A  K  K  A  M  E  R  K  T  G  A  R  G  L  R  S  I  L  E  G
1141 GTGCTACTTGAAACCATGTACGAATTACCATCTGCAACAGACGTTAGTAAAGTCGTGATT
381   V  L  L  E  T  M  Y  E  L  P  S  A  T  D  V  S  K  V  V  I
1201 GATGAGTCCGTCATTAATGGCGAGTCAGAGCCACTGCTTATTTACAGTAACGCAGACAAT
401   D  E  S  V  I  N  G  E  S  E  P  L  L  I  Y  S  N  A  D  N
1261 CAGGCAGCTGGAGCAGAGTAA
421   Q  A  A  G  A  E  *
```

NOTE Blue sections represent C-terminal amidation site; double underline indicates microsomal C-terminal target signaling site; wavy line indicates ATP-GTP site; yellow sections represent protein kinase C site; green sections represent CAMP-dependent protein kinase phosphate site; and single underline indicates glycine site.

Fig.2 Nucleotides of *clpX* gene and their encoded amino acid sequences

This hypothesis posits that ClpX is a cytoplasmic protein that may possess a signal peptide enabling its entry into the cytoplasm via a specific secretion or localization mechanism. It lacks a transmembrane region, yet may function by interacting with other membrane proteins or organelles. ClpX contains multiple functional sites that may be involved in a variety of biological processes, including protein stability, enzyme catalytic activity, and signal transduction. The predicted ATP_GTP site provides further evidence that ClpX functions as a subunit of the ATP-dependent protease complex. These predictions provide a preliminary understanding of the structure and function of ClpX proteins. However, the specific mechanism of action requires further experimental verification.

3.4 Homology and evolutionary analysis of *clpX* A comprehensive analysis conducted using BLAST software revealed a high degree of homology in the *clpX* gene between *V. alginolyticus* and other *Vibrio* species. It is particularly noteworthy that the amino acid sequence homology with the *clpX* gene of *V. harveyi* was 99%, which demonstrated an extremely high degree of similarity. Furthermore, additional results of multiple sequence similarity

comparisons substantiated that the *clpX* gene in *V. harveyi* displayed a considerable degree of evolutionary conservation, which is crucial for elucidating its biological function and the genetic relationships among *Vibrio* spp. (Fig. 3).

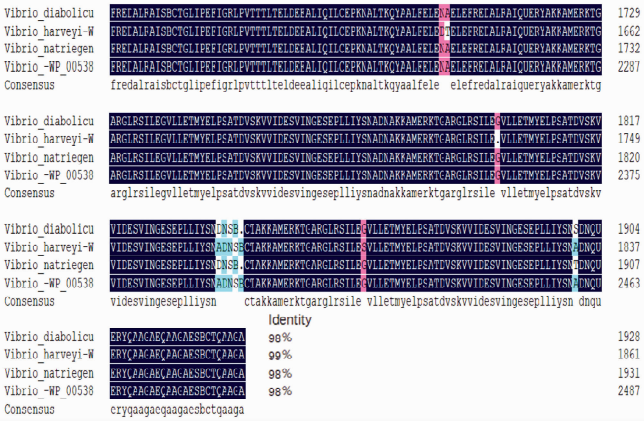


Fig. 3 Homology comparison of amino acid sequences derived from *clpX* gene

The Neighbor-joining (NT) method of MEGA 5.0 was employed to construct a phylogenetic tree using the amino acid sequences of the *clpX* gene in *V. alginolyticus* and those of other *Vibrio* sp. Fig. 4 demonstrated that *V. alginolyticus* and *V. harveyi* were situated within the same subfamily, indicating a high degree of homology between them.

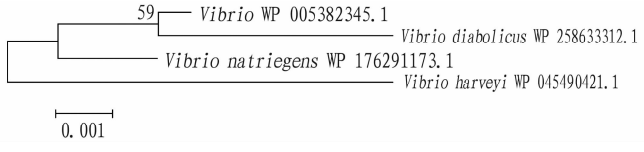


Fig. 4 Phylogenetic tree constructed based on the NJ method

3.5 Functional domains, secondary and tertiary structure prediction of ClpX The SMART program identified the presence of three distinct functional domains within ClpX (Fig. 5): a ZF-C4_ClpX functional domain (13-52 aa), an AAA functional domain (113-266 aa), and a ClpB functional domain (319-414 aa). In the secondary structure prediction, the α-helix was identified with a frequency of 40.38%, while the β-sheet, random coil, and extended strand were observed with the frequency of 5.40%, 37.09%, and 17.14%, respectively (Fig. 6).

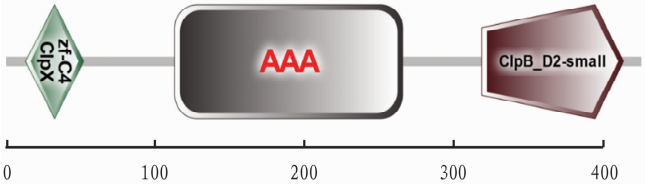
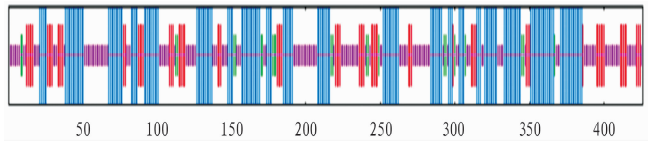


Fig. 5 Functional domains of ClpX

The zf-C4_ClpX functional domain may be involved in protein-protein interactions or substrate recognition. The AAA functional domain suggests that ClpX is an ATP-dependent molecular chaperone or protease, and these enzymes are often involved in a variety of cellular activities, such as the maintenance of protein homeostasis. The ClpB functional domain provides further evidence of the relevance of ClpX to protein degradation and refolding processes. The findings indicate that the ClpX protein is a multifunctional protein that may play a significant role in protein homeostasis, degradation, folding, and the regulation of cellular activities.



NOTE Blue: α-helix; Green: β-sheet; Purple: random coil; Red: extended strand.

Fig. 6 Secondary structure prediction of ClpX

The ClpX amino acid sequence was submitted to the SWISS-MODEL program, which employed an automated search for homologous proteins as templates to obtain a tertiary structure model of the single subunit of ClpX (Fig. 7). The template was 8e91.1.D, and the similarity between the ClpX tertiary structure model and the template was 83.73%, indicating a high degree of structural conservation between the ClpX protein and the template protein (8e91.1.D). This conservation may be indicative of the functional similarity of these proteins, which may exhibit analogous biological functions or participate in analogous biological processes.

NOTE Template: 8e91.1.D; Similarity: 83.73%.

Fig. 7 Tertiary structure prediction of ClpX

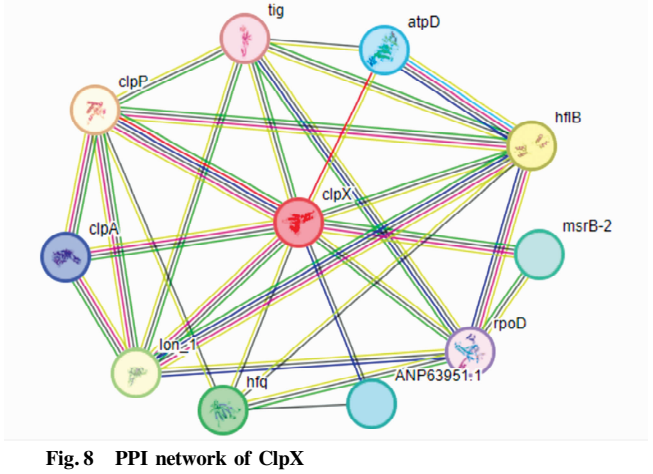


Fig. 8 PPI network of ClpX

3.6 Protein-protein interaction (PPI) network of ClpX Among the PPI networks, it can be observed that the proteins adjacent to the ClpX protein included Tig, AtpD, HflB, MsrB-2, RpoD, ClpP, ClpA, Lon-1, Hfq and ANP63951.1 (Fig. 8).

Tig protein is a peptidyl proline cis-trans isomerase, which belongs to the FKBP-type PPIase family. It catalyzes the cis/trans isomerization of proline and acts as a regulatory switch in the folding, activation, and/or degradation of numerous proteins. Additionally, Tig's protein interactions with ClpX suggest that ClpX may play a role in regulating protein changes. The AtpD protein is typically associated with ATP synthesis, which suggests that the function of ClpX may be related to energy metabolism. HflB proteins play a role in protein quality control and may also be involved in protein degradation and reuse, in collaboration with ClpX. The MsrB-2 protein is a member of the Msr (methionine sulfoxide reductase) family and is responsible for repairing methionine damage caused by oxidative stress. The interactions between ClpX and MsrB-2 suggest that ClpX plays a role in the cellular response to oxidative stress. The RpoD protein is a subunit of RNA polymerase that is associated with DNA transcription. The interactions between ClpX and RpoD may imply a role for ClpX in the regulation of gene expression. Two proteins, ClpP and ClpA, are members of the same Clp protease family as ClpX and are involved in protein degradation processes. Their interactions serve to further confirm the central role of ClpX in protein degradation. The Lon-1 protein is another protease that, in conjunction with ClpX, is implicated in protein degradation and quality control. The Hfq protein is an RNA-binding protein that plays a role in regulating RNA stability and translation. The interactions observed between ClpX and Hfq may suggest that ClpX has a function in RNA metabolism or translation regulation. The precise function of the ANP63951.1 protein remains to be elucidated. However, given its interaction with ClpX, it may represent a hitherto unidentified member of the ClpX functional network, warranting further investigation.

3.7 Prediction analysis of PTM modification sites The amino acid sequence of ClpX was submitted to automatic analysis using the programs Musite Deep (https://www.musite.net/) and FSL-Kla (http://kla.zbiolab.cn/). This analysis revealed the presence of phosphorylation, acetylation, ubiquitination, glycosylation, methylation, S-palmitoylation, and lactylation modification sites.

4 Discussion and conclusions

4.1 Discussion The results of the bioinformatics analysis indicate that the ClpX protein is composed of only four amino acids: alanine, cysteine, glycine, and threonine. This suggests that the protein is unstable, with a simple composition and an unstable structure. While ClpX proteins have been linked to both the T3SS and T6SS systems[18-19], research into protein modification methods involving alanine, cysteine, glycine, and threonine is limited. Consequently, the impact of ClpX on virulence remains unclear. Subcellular localization predictions indicates that ClpX is present in the cytoplasm and absent from other cellular structures, suggesting that ClpX, as a molecular chaperone of Clpp, likely influences the toxic effects of the type III and type VI secretion systems through PPI with Clpp proteins. However, further specific studies are necessary to investigate the histological components of ClpX separately.

In the study of *Bacillus anthracis*, ClpX, in conjunction with the protein homeostasis regulator DnaJ, regulated the maintenance of *B. anthracis* signaling activity through a two-component system. The deletion of ClpX had been observed to increase the susceptibility to innate host defenses and to severely attenuate the virulence of *B. anthracis*, even in fully pathogenic Ames strains[20]. These findings suggest that the role of ClpX is crucial for the pathogenesis of *B. anthracis*. In the study of *Mycobacterium*, it was determined that ClpX was essential for the growth of this organism and that ClpX played a role in the DNA replication process[12]. In studies related to *S. aureus*, it was demonstrated that the ClpX protease was necessary for the inactivation of the CI master repressor and the induction of phage replication[13]. Additionally, a deficiency in ClpX had been shown to result in a reduction in biofilm formation[14]. The virulence effects of the wild-type, deletion, and complemented strains of ClpX were examined in a mouse skin abscess model. The results demonstrated that the virulence of the ClpX deletion strain was significantly reduced.

The PTM modification site prediction results indicated the existence of a multitude of modification sites on ClpX proteins, including phosphorylation, acetylation, ubiquitination, glycosylation, methylation, S-palmitoylation, and lactylation modification sites. It is possible that multiple types of post-translational modifications (PTMs) may be closely related to each other, interdependent and interacting with each other, and working together to regulate the life activities of the organism through antagonistic or synergistic modes of action. The phenomenon of cross-talking refers to the interdependence or influence between multiple types of PTMs[21]. The PTM modification site prediction results for ClpX also revealed the simultaneous existence of different modification sites at the same site. These included lactylation and methylation modification sites at the 4th site K; lactylation, methylation, and ubiquitination modification sites at the 6th site K; phosphorylation and glycosylation modification sites at the 67th site T; lactylation and acetylation modification sites at the 177th site K; lactylation, acetylation and phosphorylation modification sites at the 190th site K; and glycosylation and phosphorylation modification sites at the 205th site S, suggesting the phenomenon of cross-talking between PTM modification sites. The precise manner in which the PTM modification sites interact mutually remains to be the subject of thorough investigation.

ClpX proteins have been observed to act at numerous functional sites, including microsomes, protein kinase C, ATP, and GTP. These observations suggest that ClpX plays a role in a diverse range of biological processes. It was demonstrated that the expression of ClpX was diminished in HCC cancerous tissues[22]. The aberrantly low expression of ClpX is influenced by a multitude of pathways, and the precise mechanism of action remains to be elucidated. In a specific knockout experiment of the *clpX* gene in eukaryotic cells, it was discovered that the deletion of mitochon-

drial ClpX resulted in altered metabolic pathways within the cell, with notable effects on glycolytic processes[23]. The process of glycolysis is a complex biochemical pathway involving numerous enzymes and intermediate metabolites. These components work in concert within the cytoplasmic lysate to facilitate the conversion of glucose to pyruvate, a process that ultimately results in the production of ATP. ClpX plays a crucial role in recognizing and degrading misfolded or damaged proteins within mitochondria. However, further investigation is necessary to elucidate the precise mechanisms and potential implications associated with the absence of ClpX in this process. This will facilitate a more comprehensive understanding of the role of mitochondria in cellular metabolism and the maintenance of normal metabolic activity in cells in the presence of mitochondrial dysfunction. The role of the *clpX* gene in biological process is multifaceted, with a vast array of potential mechanisms of participation. However, the specific mechanism of participation by the *clpX* gene remains unclear, particularly in the context of *V. alginolyticus*, which represents an avenue for further investigation.

4.2 Conclusions The *clpX* gene is 1 281 bp in length and encodes 426 amino acids. Its molecular structure formula is $C_{3\,842}H_{6\,405}N_{1\,281}O_{1\,598}S_{260}$, with a theoretical protein molecular weight of approximately 1 044 473.4 kDa and a theoretical pI value of 5.04. The *clpX* gene is predominantly situated within the cytoplasm, exhibiting unstable and hydrophilic protein characteristics. It possesses a signal peptide cleavage site, lacks a transmembrane region, and is not associated with any KEGG metabolic pathway. The *clpX* gene displays the highest degree of homology with that of *V. harveyi*, exhibiting three distinct structural and functional domains. The proteins interacting with ClpX protein are Tig, Atpd, Hflb, Msrb-2, Rpod, Clpp, Clpa, Lon-1, Hfq, and ANP63951.1. A computational analysis reveals the presence of phosphorylation, acetylation, ubiquitination, glycosylation, methylation, S-palmitoylation, and lactylation modification sites.

The *clpX* gene plays a role in a multitude of metabolic pathways in eukaryotic cells, as well as in prokaryotic cells, where it is involved in regulating a variety of substances. Nevertheless, the precise manner in which the *clpX* gene exerts its influence remains to be elucidated. Bioinformatics analysis may offer a theoretical foundation for further investigation into the role of the *clpX* gene.

References

[1] AGUIRRE-GUZMÁN G, ASCENCIO F, SAULNIER D. Pathogenicity of Vibrio penaeicida for white shrimp *Litopenaeus vannamei*: A cysteine protease-like exotoxin as a virulence factor[J]. Diseases of Aquatic Organisms, 2005, 67(3): 201 – 207.

[2] MENG L, CHE JY, HUANG XX, *et al.* Deletion of *RseB* gene attenuates the pathogenicity of *Vibrio alginolyticus* in *Litopenaeus vannamei*[J]. Journal of Shanghai Ocean University, 2023(4): 681 – 689. (in Chinese).

[3] YUAN SB, ZHU AY. Progress on pathogenicity research on *Vibrio alginolyticus* to aquatic products[J]. Journal of Zhejiang Ocean University (Natural Science Edition), 2012, 31(3): 256 – 264. (in Chinese).

[4] MEI B, LU X, WANG LN, *et al.* Progress on virulence factor and virulence gene research of *Vibrio alginolyticus*[J]. Liaoning Agricultural Sciences, 2015(5): 58 – 60. (in Chinese).

[5] FLEMMING HC, WINGENDER J, SZEWZYK U, *et al.* Biofilms: An emergent form of bacterial life[J]. Nature Reviews Microbiology, 2016, 14(9): 563 – 575.

[6] WANG HY, YUAN Y, ZHOU QL, *et al.* Advance in studies on the main virulence factors of the pathogenic *Vibrios*[J]. Biotechnology Bulletin, 2009(S1): 75 – 78. (in Chinese).

[7] RUTSCHLIN S, BOTTCHER T. Engineering siderophores[M]//SHUKLA AK. Chemical and synthetic biology approaches to understand cellular functions-Part C[J]. Life Sciences, 2020(633): 29 – 47.

[8] ZHANG J, HAO Y, YIN KY, *et al.* VqsA controls exotoxin production by directly binding to the promoter of asp in the pathogen *Vibrio alginolyticus*[J]. Fems Microbiology Letters, 2019, 366(6): fnz056.

[9] LI J. Study on the associated gene of the needle-like injectisome in the type II secretion system from *Vibrio alginolyticus*[D]. Zhanjiang: Guangdong Ocean University, 2016. (in Chinese).

[10] BAKER TA, SAUER RT. ClpXP, an ATP-powered unfolding and protein-degradation machine[J]. Biochimica Et Biophysica Acta, 2012, 1823(1): 15 – 28.

[11] FREES D, GERTH U, INGMER H. Clp chaperones and proteases are central in stress survival, virulence and antibiotic resistance of *Staphylococcus aureus*[J]. International Journal of Medical Microbiology, 2014, 304(2): 142 – 149.

[12] KESTER JC, KANDROR O, AKOPIAN T, *et al.* ClpX is essential and activated by single-strand DNA binding protein in mycobacteria[J]. Journal of Bacteriology, 2021, 203(4): DOI:10.1128/JB.00608 – 20.

[13] GURUNG V, BISWAS I. ClpX/P-dependent degradation of novel substrates in *Streptococcus mutans*[J]. Journal of Bacteriology, 2022, 204(4): e0059421.

[14] WOOD NA, BLOCKER AM, SELEEM MA, *et al.* The ClpX and ClpP2 orthologs of *Chlamydia trachomatis* perform discrete and essential functions in organism growth and development[J]. Microbiology, 2020, 11(5): DOI:10.1128/mBio.02016 – 20.

[15] CAI SH, WU ZH, JIAN J, *et al.* Cloning and expression of gene encoding the thermostable direct hemolysin from *Vibrio alginolyticus* strain HY9901, the causative agent of vibriosis of crimson snapper (*Lutjanus erythopterus*)[J]. Journal of Applied Microbiology, 2007, 103(2): 289 – 296.

[16] ZENG FY. Effects of lysine succinylation modification of *Vibrio alginolyticus* PEPCK on drug resistance[D]. Zhanjiang: Guangdong Ocean University, 2022. (in Chinese).

[17] PANG HY, ZHOU ZJ, DING J, *et al.* Molecular cloning and bioinformatics analysis of T3SS chaperone escort protein VscO from *Vibrio alginolyticus*[J]. Biotechnology Bulletin, 2014(6): 155 – 161. (in Chinese).

[18] LI Y, YAMAZAKI A, ZOU L, *et al.* ClpXP protease regulates the type III secretion system of *Dickeya dadantii* 3937 and is essential for the bacterial virulence[J]. Molecular Plant-Microbe Interactions, 2010, 23(7): 871 – 878.

[19] LIN HH, YU M, SRIRAMOJU MK, *et al.* A high-throughput interbacterial competition screen identifies ClpAP in enhancing recipient susceptibility to type VI secretion system-mediated attack by *Agrobacterium tumefaciens*[J]. Frontiers in Microbiology, 2019(10): 3077.

[20] LAUT CL, LEASURE CS, PI H, *et al.* DnaJ and ClpX are required for HitRS and HssRS two-component system signaling in *Bacillus anthracis*[J]. Infection and Immunity, 2022, 90(1): e00560 – 00521.

[21] VENNE AS, KOLLIPARA L, ZAHEDI RP. The next level of complexity: Crosstalk of posttranslational modifications[J]. Proteomics, 2014, 14(4 – 5): 513 – 524.

[22] WANG DL, SONG J, CONG CL, *et al.* Expression and significance of Clp protease complex subunit ClpX in hepatocellular carcinoma[J]. Chinese Journal of Oncology Prevention and Treatment, 2020(3): 291 – 296. (in Chinese).

[23] SUZUKI K, KUBOTA Y, KANEKO K, *et al.* CLPX regulates mitochondrial fatty acid β-oxidation in liver cells[J]. The Journal of Biological Chemistry, 2023, 299(10): 105210.