Gene Cloning and Bioinformatics Analysis of *phoR* Gene from *Vibrio alginolyticus* HY9901

Xiangyu LIU[1,2△], Peng ZHOU[1,2△], Haiyun FENG[1,2], Weijie ZHANG[1,2], Huanying PANG[1,2*], Na WANG[3], Xiaonan LU[4]

1. Fisheries College, Guangdong Ocean University, Zhanjiang 524025, 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 524025, China; 3. Chinese Academy of Inspection and Quarantine, Beijing, 100176, China; 4. Department of Food Science and Agricultural Chemistry, Faculty of Agricultural and Environmental Sciences, McGill University, Sainte-Anne-de-Bellevue H9X 3V9, Canada

Abstract PhoR is a histidine kinase in a two-component regulatory system that regulates phosphorus metabolic pathways and undertakes the key mission of information transmission in pathogenic bacteria. The full-length *phoR* gene was successfully cloned from the *Vibrio alginolyticus* HY9901 strain. A comprehensive analysis of the cloned gene was conducted using bioinformatics. Sequence analysis revealed that the total length of the *phoR* gene (GenBank accession No.: KJ958404.1) is 1299 bp, with the coding region containing a total of 432 amino acid residues. The phylogenetic tree of *PhoR* revealed that it belongs to the same subclade as *V. diabolicus*. The SMART program was employed for the purpose of functional domain prediction, which revealed that *PhoR* possesses three major functional domains: PAS (amino acids 98-166), His-KA (amino acids 205-272), and HATPase_c (amino acids 317-429).

Key words *Vibrio alginolyticus*, *phoR* gene, Gene cloning, Bioinformatics analysis

1 Introduction

Vibrio alginolyticus is a gram-negative bacterium belonging to the family Vibrionaceae and is characterized by its rod-shaped morphology. It is typically found in environments such as seawater and estuaries, but can also be found in some marine animals and seafood. In recent years, with the ongoing expansion of the aquaculture industry, the issue of aquatic diseases has emerged as a significant concern. Among the aforementioned issues, those caused by *V. alginolyticus* have been a significant challenge for the aquaculture industry. *V. alginolyticus* is known to cause tissue damage in the host organism, which in turn affects the host's normal metabolic processes[1]. *V. alginolyticus* is a highly pathogenic bacterium found in both humans and aquatic animals. Contact with water containing this bacterium may result in various infections, including those of the skin and external ear. It can also be transmitted through uncooked seafood, which may cause digestive organ infections, endangering human health[2]. The pathogenicity of *V. alginolyticus* is contingent upon its interaction with host cells[3]. The currently known virulence factors of *V. alginolyticus* include extracellular products[4], adhesion factor[5], siderophore[6], and lipopolysaccharide[7]. The two-component regulatory system (TCRS) plays an equally important role in the pathogenesis of diseases caused by *V. alginolyticus* and is an important virulence factor[8].

TCRS is a signaling system that exists within bacteria. It maintains the bacteria's own survival by sensing changes in the external environment, regulating physiological functions, and expressing virulence factors. This mechanism allows bacteria to adapt to selective pressures[9]. The PhoBR TCRS serves as a fundamental stimulus-response coupling mechanism that allows organisms to sense and respond to changes in a multitude of environmental conditions. The system is typically composed of a histidine kinase, PhoR, which senses specific environmental stimuli, and a response-regulating protein, PhoB, which mediates cellular responses[10]. It is widely distributed in both gram-negative bacteria, such as *Escherichia coli*, and gram-positive bacteria, such as *Bacillus subtilis* and *Staphylococcus aureus*[11]. PhoR, as a histidine protein kinase, is responsible for recognizing and translating external environmental signals. Additionally, it regulates and phosphorylates the response-regulating protein PhoB[12]. Inorganic phosphate is an essential nutrient for organisms consumed by animals as a protective mechanism against bacterial infection. However, this consumption is perceived by some pathogens as a signal to turn on the expression of virulence genes. Under Pi-restricted conditions, PhoB is activated by PhoR as a kinase, but under Pi-filled conditions, PhoB activation is interrupted by PhoR as a phosphatase[13]. The results of experimental studies indicate that in the plant pathogen *Agrobacterium tumefaciens*, biofilm formation is en-

Received: February 25, 2024 Accepted: May 7, 2024
Supported by Outstanding Graduate Entering Laboratory Project of College of Fisheries, Guangdong Ocean University; National Natural Science Foundation of China (32073015); Undergraduate Innovation Team of Guangdong Ocean University (CCTD201802); Undergraduate Innovation and Entrepreneurship Training Program of Guangdong Ocean University (CXXL2024007).

△These authors contributed equally to this work.
* Corresponding author. Huanying PANG, PhD., associate professor, research fields: aquatic veterinary medicine.

hanced under conditions of low inorganic phosphate concentration by the action of PhoR-PhoB, thus promoting bacterial survival and dissemination[14]. At low Pi concentrations, PhoB exerts its direct or indirect inhibitory effects on the secretion and expression of proteins in the type III and type VI secretion systems through esrC control[13]. Furthermore, PhoR is involved in the resistance of pathogens to acidity, toxin production, and antimicrobial drugs, which is influenced by the concentration of inorganic phosphate[15].

To date, PhoR has been investigated in *E. coli* and *B. subtilis*, with some findings emerging. However, there are fewer corresponding results in *V. alginolyticus*. This study was to clone and bioinformatically analyze the *phoR* gene of *V. alginolyticus* HY9901, in order to provide a theoretical basis for further research on drugs against *V. alginolyticus*.

2 Materials and methods

2.1 Materials

2.1.1 Strains. *V. alginolyticus* virulent strain HY9901 was isolated and preserved from a diseased *Lutjanus sanguineus* specimen collected from the sea area of Zhanjiang Harbor in Guangdong Province, China.

2.1.2 Reagents. The Ex*Taq* DNA polymerase was procured from Takara, while the bacterial genomic DNA extraction kit and DNA gel recovery kit were sourced from Tiangen Biotech (Beijing) 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 Shanghai Sangon Biotechnology Services Co., Ltd.

2.2 Methods

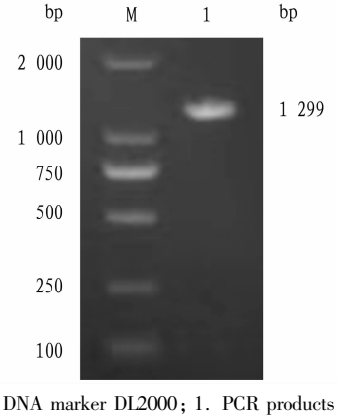
2.2.1 Extraction of total DNA from *V. alginolyticus* HY9901. The *V. alginolyticus* HY9901 strain was coated on TSA plates. Single colonies were selected and inoculated in TSB (5% NaCl) medium, and then incubated at 28 ℃ with oscillation for a period exceeding 12 h. The appropriate amount of bacterial liquid was taken and transferred to an EP centrifuge tube. Centrifugation at 10 000 rpm/min for 1 min was performed to collect the thalli. The instructions for the kit were consulted in order to extract genomic DNA, and the samples were stored at −20 ℃ for future use.

2.2.2 Cloning of *phoR* gene. A pair of primers was designed according to the sequence of *phoR* gene from *V. alginolyticus*. The upstream primer, designated P1, was GTGGTTGAAAGATTAACGTGGA, while the downstream primer, designated P2, was TTATTTCACCACCAAACGTTCA. 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 94 ℃ for 4 min; denaturation at 94 ℃ for 30 sec, annealing at 62 ℃ for 40 sec, and extension at 72 ℃ for 60 sec, 35 cycles; and extension at 72 ℃ for 10 min. Following the examination of the PCR products by 1% agarose gel electrophoresis, the gel was cut and recovered using a gel-cutting kit. The recovered fragments were then cloned into the pMD18-T vector, which was named pMD-*phoR*.

2.2.3 Bioinformatics analysis of *phoR* gene of *V. alginolyticus* HY9901. Sequence homology comparison and similarity analysis were conducted using the National Center for Biotechnology Information (NCBI) (http://blast.ncbi.nlm.nih.gov/Blast.cgi). Amino acid homology comparison analysis was performed using the software program DNAMAN Version 6.0 (Lynnon Biosoft). ORF Finder (http://www.ncbi.nlm.nih.gov/gorf/gorf.html) and ExPASy Proteomics Server (http://ca.expasy.org) were employed for the purpose of deducing amino acid sequences, determining open reading frames (ORFs), calculating molecular weight values (Mw), and making theoretical isoelectric point (pI) predictions. The signal peptide sequences were predicted by online analysis software SignalP 4.0 Server (http://www.cbs.dtu.dk/services/SignalP). The transmembrane structural domains were predicted by TMHMM Server 2.0 (http://www.cbs.dtu.dk/services/TMHMM). The distribution of functional sites in amino acid sequences was predicted by SoftBerry-Psite (http://linux1.softberry.com/berry.phtml?topic=psite&group=programs&subgroup=proloc). The protein structure-function domains were analyzed using InterProScan Sequence Search (http://www.ebi.ac.uk/Tools/InterProScan), while the subcellular localization prediction was performed using PSORT II Prediction (http://psort.hgc.jp/form2.html). A Phylogenetic tree was constructed using the neighbor-joining method with the Clastal 2.0 and MEGA 5.0 software. Modeling was conducted using the SWISS-MODEL program (http://www.swissmodel.expasy.org/) on the ExPASy server and analyzed using the 3D structural analysis software PyMOL Viewer[16].

3 Results and analysis

3.1 Full-length cloning of *phoR* gene The PCR amplification procedure yielded a *phoR* gene-specific band of approximately 1 300 bp in size (Fig. 1). The sequencing results demonstrated that the *phoR* gene contained an open reading frame of 1 299 bp, which encoded 432 amino acids. The gene was subsequently submitted to GenBank, where it was assigned the accession number KJ958404.1.

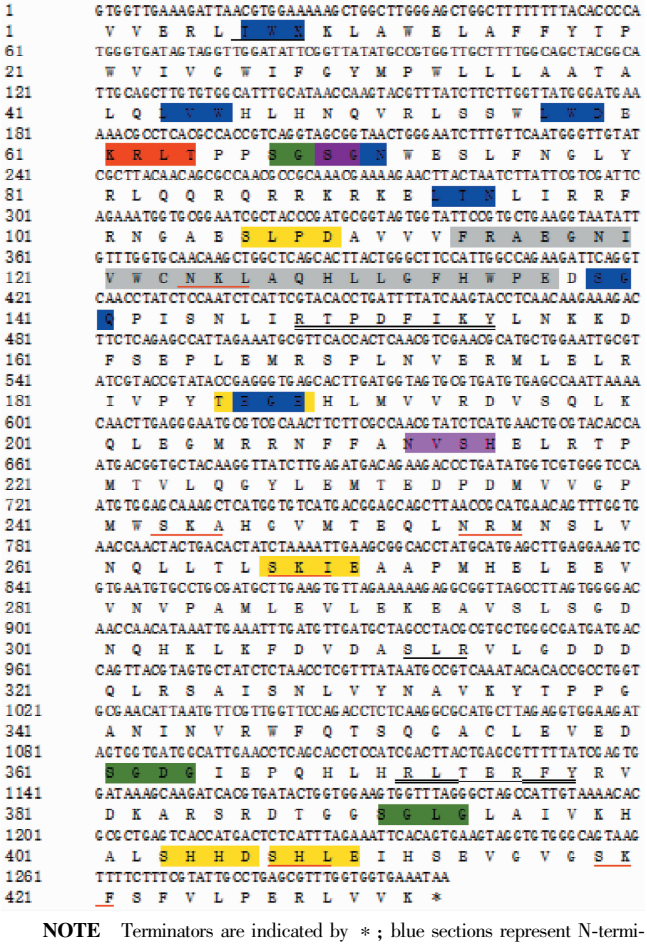


NOTE M. DNA marker DL2000; 1. PCR products of *phoR* gene.

Fig. 1 Amplification of *phoR* gene

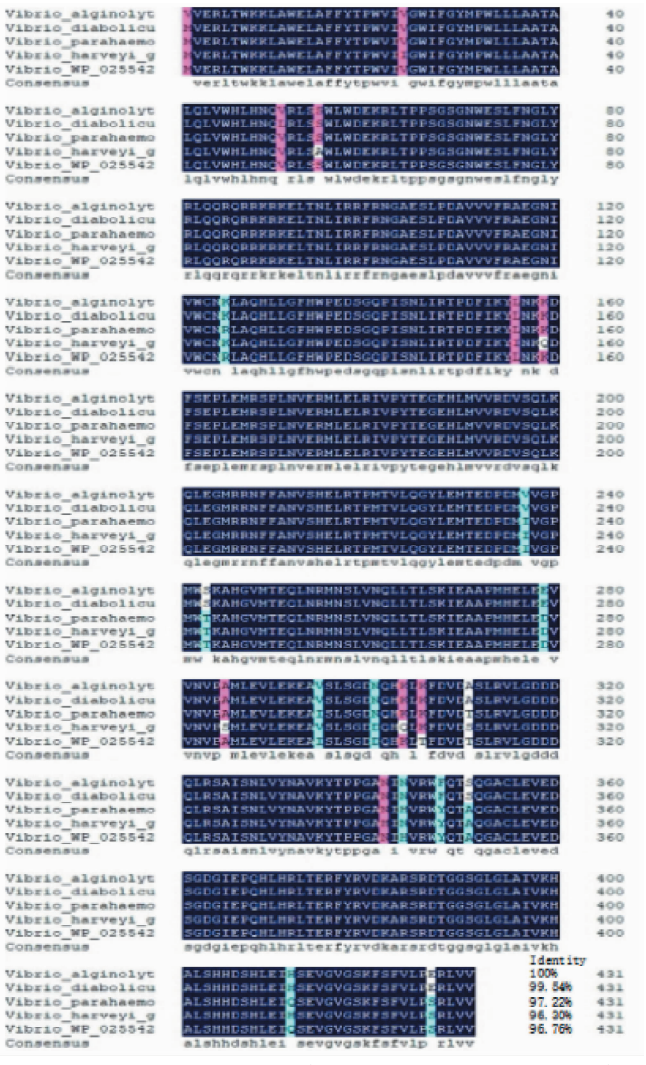
3.2 Physicochemical properties of PhoR The PhoR protein of *V. alginolytici* HY9901 was analyzed using the ExPASy soft-

ware. The results indicate that the protein has a total of 6 996 atoms, with a molecular structure formula of $C_{2232}H_{3495}N_{625}O_{629}S_{15}$, a theoretical molecular weight of 49.630 01 kDa, a theoretical pI value of 6.81, an instability coefficient of 45.08 (unstable), a fat coefficient of 94.70, a total average hydrophilicity of −0.287, and that the protein is hydrophobic in general. The protein in question does not contain pyrrole lysine (Pyl), selenocysteine (Sec), and has a molar extinction coefficient of 85 034 mol/cm at 280 nm. The total number of acidic amino acids (Asp + Glu) is 52, and the total number of basic amino acids (Arg + Lys) is 50, with valine (Val) at the N-terminus. The half-life of expression in yeast and *E. coli* is greater than 20 and 10 h, respectively, while the half-life of expression in mammalian reticulocytes in *in vitro* culture is found to be 100 h.



NOTE Terminators are indicated by ＊; blue sections represent N-terminal cardamoylation sites; black underlined sections represent protein kinase C phosphorylation sites; yellow sections represent casein kinase II phosphorylation sites; red sections represent cAMP- and cGMP-dependent protein kinase phosphorylation sites; red underlined sections represent microbial C-terminal target signaling sites; green sections represent glycosaminoglycan attachment sites; gray portion represents class I Aminoacyl-tRNA synthases 1; pink portion represents N-glycosylation site; double underlined portion represents tyrosine kinase phosphorylation site.

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



NOTE *Vibrio alginolyticus* (KJ 958404.1); *V. diabolicus* (WP 257889693.1); *V. parahaemolyticus* (WP 140257442.1); *V. harveyi* group (WP 038885644.1); *Vibrio* (WP 025542575.1).

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

3.3 Sequence analysis of PhoR The N-terminal signal peptide structure of the PhoR amino acid sequence was predicted using the SignalP 4.0 Server program. This analysis revealed that there was no obvious signal peptide cleavage site and that no signal peptide was present. The protein was found to contain one transmembrane region, as predicted by the TMHMM Server 2.0 program. SoftBerry-Psite program prediction revealed that the amino acid sequence contains one cAMP- and cGMP-dependent protein kinase phosphorylation site (61-64 aa); 3 protein kinase C phosphorylation sites (6-8 aa, 312-314 aa, 374-376 aa); 5 casein kinase II phosphorylation sites (106-109 aa, 185-188 aa, 267-270 aa, 403-406 aa, 407-410 aa); one N-glycosylation site (212-215 aa); 7 N-myristoylation sites (6-8 aa, 43-45 aa, 57-59 aa, 69-71 aa, 93-95 aa, 139-141 aa, 186-188 aa); 3 glycosaminoglycan attachment sites (67-70 aa, 361-364 aa, 391-394 aa); 2 tyrosine kinase phosphorylation sites (148-155 aa, 372-378 aa); 6 micro-

bial C-terminal target signals (124-126 aa, 243-245 aa, 254-256 aa, 267-269 aa, 407-409 aa, 419-421 aa); and one class I Aminoacyl-tRNA synthases 1 (114-137 aa) (Fig. 2). The predicted subcellular localization of the protein indicated that PhoR had the highest likelihood of being located on the plasma membrane (65.2%), with a possibility of being located on the endoplasmic reticulum at 34.8%.

3.4 Homology and evolutionary analysis of *phoR* A BLAST analysis revealed that the *phoR* of *V. alginolyticus* exhibited high homology with the *phoR* of other *Vibrio* species. Among these, it exhibited the highest homology with the amino acid sequence of the *phoR* of *V. diabolicus* (99.54%). Furthermore, multiple sequence similarity comparisons indicated that the *phoR* in *Vibrio* are highly conserved (Fig. 3).

The deduced amino acid sequences of PhoR were employed to construct a phylogenetic tree with other microorganisms using the Neighbor-joining method with MEGA 5.0 software. The results demonstrated that the PhoR proteins of *V. alginolyticus* HY9901 clustered into the same subfamily with *V. diabolicus*, indicating a closer kinship (Fig. 4). This finding corroborates the results of the traditional classification of morphologic and biochemical characteristics.

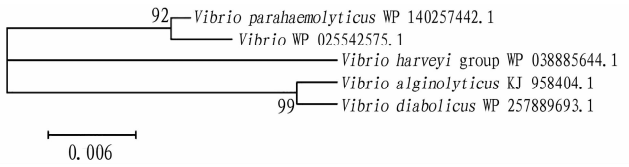
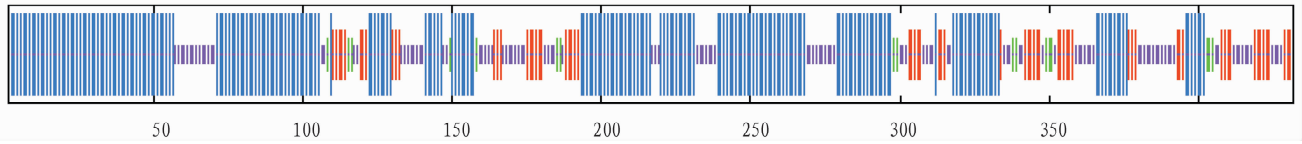


Fig. 4 PhoR amino acid phylogenetic tree constructed based on the NJ method

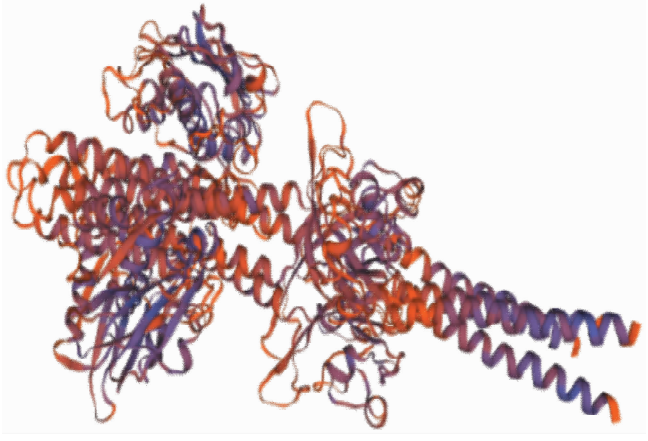
3.5 Functional domains, secondary and tertiary structure prediction of PhoR The SMART program was employed for the purpose of functional domain prediction, which revealed the existence of three principal functional domains within PhoR (Fig. 5). These domains were identified as PAS (amino acids 98-166), HisKA (amino acids 205-272), and HATPase_c (amino acids 317-429).



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

Fig. 5 Secondary structure prediction of PhoR

The PhoR amino acid sequence was submitted to the SWISS-MODEL program, which employed an automated search for homologous proteins as templates to generate a tertiary structure model of the PhoR single subunit (Fig. 6).



NOTE Template: 4i5s.1.A; Similarity: 28.49%.

Fig. 6 Tertiary structure prediction of PhoR

3.6 Protein-protein interaction (PPI) network of PhoR Among the PPI networks, it can be observed that the proteins adjacent to PhoR protein included phoB-1, pstC, ompR, ntrB, mprA-2, ANP63802.1, ANP65289.1, ANP67691.1, ANP68099.1, and ANP66965.1 (Fig. 7). The following proteins were identified in the text mining process: hoB-1, pstC, ompR, ntrB, mprA-2, ANP63802.1, ANP65289.1, ANP67691.1, and ANP68099.1.

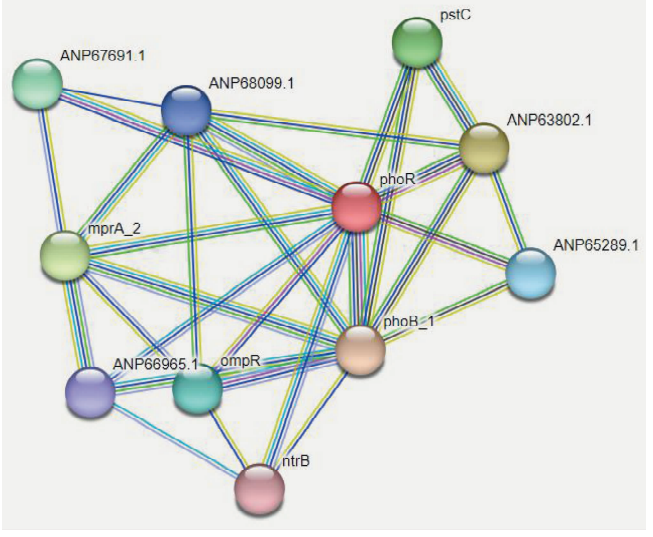


Fig. 7 PPI network of PhoR

4 Discussion and conclusions

4.1 Discussion This study examined the homology of *V. alginolyticus* PhoR with other *Vibrio* PhoR proteins. It was found that the amino acid sequence of the PhoR of *V. diabolicus* exhibited a homology of 99.54%, indicating a high degree of homology. The phylogenetic tree demonstrated that PhoR proteins were grouped into the same subfamily as those of *V. diabolicus*. The multifaceted results indicate that PhoR in *Vibrio* is highly conserved, which suggests that PhoR plays an important role in the process of biological

evolution.

The PhoBR TCRS is a prevalent gene regulatory system in bacteria that is primarily involved in regulating bacterial perception of and response to phosphorus concentrations in the environment. In this system, PhoR acts as a bacterial receptor protein, sensing the inorganic phosphorus concentration in the external environment[17] and regulating phosphorus metabolism pathways within the cell to adapt to environmental changes. These information are then transmitted to the response regulator protein PhoB[18]. The predicted subcellular localization of the protein and the predicted transmembrane structure in this study indicated that PhoR is most likely to be located on the plasma membrane and contains a transmembrane region. This spatial distribution is more conducive to PhoR receiving signals from environmental stimuli outside the cell membrane and transmitting these signals to intracellular PhoB, thereby regulating changes in metabolic pathways.

The results of PhoR functional domain prediction indicate that PhoR possesses a HATPase_c functional domain. Christopher *et al.*[19] demonstrated that the HATPase_c functional domain contains conserved residues that are involved in ATP and metal ion coordination. When the PhoR receptor binding region is subjected to external environmental signals, the HATPase_c functional domain hydrolyzes ATP to ADP. This allows the histidine kinase to autophosphorylate and subsequently phosphorylates PhoB by modulating the phosphoryl group. This, in turn, alters the bacterial phosphorus metabolism pathway and the expression of virulence genes. In a study by Li Nan *et al.*[20], it was found that certain small molecule lead compounds can compete with ATP for the binding site and inhibit the autophosphorylation of histidine protein kinase. Consequently, further investigation into the structure of the binding site and the identification of compounds capable of binding to it may prove beneficial in reducing the expression of virulence genes and contributing to the reduction of the emergence of drug-resistant bacteria.

4.2 Conclusions In this study, the *phoR* gene sequence was amplified from *V. alginolyticus* HY9901, which exhibited a length of 1 299 bp and encoded 432 amino acids. The protein is hydrophobic, with a theoretical pI value of 6.81, and a theoretical molecular weight of 49.630 01 kDa. A homology and evolutionary analysis of PhoR revealed that PhoR is highly conserved in *V. alginolyticus*. The SMART program was employed for the purpose of functional domain prediction, which revealed the existence of three principal functional domains within PhoR. These domains include the PAS (amino acids 98-166), HisKA (amino acids 205-272), and HATPase_c (amino acids 317-429). These domains are of significant importance in the context of bacterial life activities and the expression of virulence genes.

References

[1] RUI HP. From fish to human beings: Pathogenesis investigation of several important *Vibrio* species[D]. Shanghai: East China University of Science and Technology. (in Chinese).

[2] ZHENG HY, YAN L, YANG C, *et al.* Population genomics study of *Vibrio alginolyticus*[J]. Hereditas (Beijing), 2021, 43(4): 350 – 361. (in Chinese).

[3] 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).

[4] JIN S, ZHENG TL, WANG GL, *et al.* Pathogenicity of extracellular products of *Vibrio alginolyticus* to great yellow croaker, *Psedosciaena crocea*[J]. Chinese Journal of Veterinary Science, 2004(5): 439 – 441. (in Chinese).

[5] LIU X. Prokaryotic expression, antigenicity identification and bioinformatics analysis of *Vibrio alginolyticus* accessory colonization factor ACFA[J]. Southwest China Journal of Agricultural Sciences, 2016, 29(7): 1755 – 1760. (in Chinese).

[6] WANG PB. Preliminary study on iron uptake and regulatory mechanism of *Vibrio alginolyticus*[D]. Shanghai: East China University of Science and Technology (in Chinese).

[7] XUAN XZ, XIE ZY, ZHOU YC, *et al.* The polymorphism of lipopolysaccharides isolated from *Vibrio alginolyticus* strains with different virulence [J]. Progress in Fishery Sciences, 2011, 32(1): 104 – 108. (in Chinese).

[8] GELLATLY SL, BAINS M, BREIDENSTEIN EBM, *et al.* Novel roles for TCRSs in cytotoxicity and virulence-related properties in *Pseudomonas aeruginosa*[J]. AIMS Microbioly, 2018, 4(1): 173 – 91.

[9] GROISMAN EA. Feedback control of TCRSs[J]. Annual Review of Microbiology, 2016(70): 103 – 124.

[10] ZHENG D, XUE B, SHAO Y, *et al.* Activation of PhoBR under phosphate-rich conditions reduces the virulence of *Xanthomonas oryzae* pv. *oryzae*[J]. Molecular Plant Pathology, 2018, 19(9): 2066 – 2076.

[11] WINKLER ME, HOCH JA. Essentiality, bypass, and targeting of the YycFG (VicRK) TCRS in gram-positive bacteria[J]. Journal of Bacteriology, 2008, 190(8): 2645 – 2648.

[12] MENG X, AHATOR SD, ZHANG LH. Molecular mechanisms of phosphate stress activation of *Pseudomonas aeruginosa* quorum sensing systems[J]. mSphere, 2020, 5(2).

[13] CHAKRABORTY S, SIVARAMAN J, LEUNG KY, *et al.* Two-component PhoB-PhoR regulatory system and ferric uptake regulator sense phosphate and iron to control virulence genes in type III and VI secretion systems of *Edwardsiella tarda*[J]. Journal of Biological Chemistry, 2011, 286(45): 39417 – 39430.

[14] DANHORN T, HENTZER M, GIVSKOV M, *et al.* Phosphorus limitation enhances biofilm formation of the plant pathogen *Agrobacterium tumefaciens* through the PhoR-PhoB regulatory system[J]. Journal of Bacteriology, 2004, 186(14): 4492 – 4501.

[15] SANTOS-BENEIT F. The Pho regulon: A huge regulatory network in bacteria[J]. Frontiers in Microbiology, 2015(6): 402.

[16] BORDOLI L, SCHWEDE T. Automated protein structure modeling with SWISS-MODEL workspace and the protein model portal[M]. Homology Modeling, 2011: 107 – 136.

[17] SCHRAMKE H, LAERMANN V, TEGETMEYER HE, *et al.* Revisiting regulation of potassium homeostasis in *Escherichia coli*: The connection to phosphate limitation[J]. Microbiology Open, 2017, 6 (3): e00438 – e00438.

[18] CHOI S, CHOI E, CHO YJ, *et al.* The *Salmonella* virulence protein MgtC promotes phosphate uptake inside macrophages[J]. Nature Communications, 2019, 10(1): 3326.

[19] ZSCHIEDRICH CP, KEIDEL V, SZURMANT H. Molecular mechanisms of two-component signal transduction[J]. Journal of Molecular Biology, 2016, 428(19): 3752 – 3775.

[20] LI N, WANG F, NIU S, *et al.* Discovery of novel inhibitors of *Streptococcus pneumoniae* based on the virtual screening with the homology-modeled structure of histidine kinase (VicK)[J]. BMC Microbiology, 2009, 9(1): 129.