PLOS ONE

RESEARCH ARTICLE

Molecular docking and dynamics simulation studies uncover the host-pathogen protein-protein interactions in *Penaeus vannamei* and *Vibrio parahaemolyticus*

Nur Fathiah Rosilan[1], Muhamad Arif Mohamad Jamali[2], Siti Aishah Sufira[3], Khor Waiho[4,5], Hanafiah Fazhan[4,5], Noraznawati Ismail[1], Yeong Yik Sung[1], Zeti-Azura Mohamed-Hussein[6,7], Azzmer Azzar Abdul Hamid[3]*, Nor Afiqah-Aleng[1]*

1 Institute of Climate Adaptation and Marine Biotechnology (ICAMB), Universiti Malaysia Terengganu, Kuala Nerus, Terengganu, Malaysia, **2** Faculty of Science and Technology, Universiti Sains Islam Malaysia, Bandar Baru Nilai, Nilai, Negeri Sembilan, Malaysia, **3** Research Unit for Bioinformatics and Computational Biology (RUBIC), Kuliyyah of Science, International Islamic University Malaysia, Bandar Indera Mahkota, Kuantan, Pahang, Malaysia, **4** Higher Institution Centre of Excellence (HICoE), Institute of Tropical Aquaculture and Fisheries, Universiti Malaysia Terengganu, Kuala Nerus, Terengganu, Malaysia, **5** Centre for Chemical Biology, Universiti Sains Malaysia, Minden, Penang, Malaysia, **6** UKM Medical Molecular Biology Institute, UKM Medical Centre, Jalan Yaacob Latiff, Cheras, Kuala Lumpur, Malaysia, **7** Department of Applied Physics, Faculty of Science and Technology, Universiti Kebangsaan Malaysia, UKM Bangi, Selangor, Malaysia

* afiqahaleng@umt.edu.my (NAA); azzmer@iium.edu.my (AAAH)

OPEN ACCESS

Citation: Rosilan NF, Jamali MAM, Sufira SA, Waiho K, Fazhan H, Ismail N, et al. (2024) Molecular docking and dynamics simulation studies uncover the host-pathogen protein-protein interactions in *Penaeus vannamei* and *Vibrio parahaemolyticus*. PLoS ONE 19(1): e0297759. https://doi.org/10.1371/journal.pone.0297759

Editor: Dharmendra Kumar Meena, CIFRI: Central Inland Fisheries Research Institute, INDIA

Received: August 3, 2023

Accepted: January 11, 2024

Published: January 24, 2024

Data Availability Statement: All relevant data are within the manuscript and its Supporting information files.

Funding: NA-A Fundamental Research Grant Scheme (FRGS) (FRGS/1/2021/STG01/UMT/03/1, Vot No. 59667) Ministry of Higher Education (MoHE) Malaysia https://www.mohe.gov.my/en YES - Data collection and analysis, preparation of manuscript.

Abstract

Shrimp aquaculture contributes significantly to global economic growth, and the whiteleg shrimp, *Penaeus vannamei*, is a leading species in this industry. However, *Vibrio parahaemolyticus* infection poses a major challenge in ensuring the success of *P. vannamei* aquaculture. Despite its significance in this industry, the biological knowledge of its pathogenesis remains unclear. Hence, this study was conducted to identify the interaction sites and binding affinity between several immune-related proteins of *P. vannamei* with *V. parahaemolyticus* proteins associated with virulence factors. Potential interaction sites and the binding affinity between host and pathogen proteins were identified using molecular docking and dynamics (MD) simulation. The *P. vannamei-V. parahaemolyticus* protein-protein interaction of Complex 1 (Ferritin-HrpE/YscL family type III secretion apparatus protein), Complex 2 (Protein kinase domain-containing protein-Chemotaxis CheY protein), and Complex 3 (GPCR-Chemotaxis CheY protein) was found to interact with -4319.76, -5271.39, and -4725.57 of the docked score and the formation of intermolecular bonds at several interacting residues. The docked scores of Complex 1, Complex 2, and Complex 3 were validated using MD simulation analysis, which revealed these complexes greatly contribute to the interactions between *P. vannamei* and *V. parahaemolyticus* proteins, with binding free energies of -22.50 kJ/mol, -30.20 kJ/mol, and -26.27 kJ/mol, respectively. This finding illustrates the capability of computational approaches to search for molecular binding sites between host and pathogen, which could increase the knowledge of *Vibrio*

Competing interests: The authors have declared that no competing interests exist.

spp. infection on shrimps, which then can be used to assist in the development of effective treatment.

Introduction

Shrimp aquaculture is a rapidly growing animal food-producing sector because it is in high demand and contributes significantly to the global economy [1, 2]. Among various cultured shrimp species, the whiteleg shrimp, *Penaeus vannamei*, is the most commonly cultured species worldwide, with a reported production of 6.3 million tonnes in 2021 [3]. However, bacterial infection has led to a high mortality rate, and it has become one of the main constraints to the success of shrimp aquaculture, which often results in a significant global economic loss [4]. *Vibrio parahaemolyticus*, a gram-negative and rod-shaped bacteria [5], is one of the most common marine bacteria that infect *P. vannamei*. It can potentially cause disease outbreaks, resulting in mass mortality in cultured shrimps [6]. One of the prevalent diseases caused by *V. parahaemolyticus* is acute hepatopancreatic necrosis disease (AHPND), causing the infected shrimps to be lethargic and exhibit stunted growth. Ultimately, this infection is able to cause 100% mortalities within 30 to 40 days after stocking [7, 8]. In addition to AHPND, *V. parahaemolyticus* can also cause white faeces disease (WFD) in the infected *P. vannamei* [9]. WFD is characterised by the presence of white faecal strings and golden brown intestines among infected individuals [10]. WFD causes mass mortalities after approximately 50 days of culture, reducing the survival rate by 20% to 30% among infected individuals [11]. *V. parahaemolyticus* is also able to cause glass post-larvae disease (GPD) in *P. vannamei*, with more than 90% of the mortality observed between 24 to 48 hours after the first appearance of abnormal symptoms among the infected shrimps, such as pale hepatopancreas and a colourless digestive tract [12, 13].

The pathogenicity of *V. parahaemolyticus* can be determined with the presence of virulence factors, which are specific molecules produced by the bacteria that assist in host colonisation [14]. One of the common virulence factors in *V. parahaemolyticus* is the type III secretion system (T3SS) [15]. T3SS is a needle-like bacterial structure known as an injectisome that permits the gram-negative bacteria to inject and secrete effector proteins into the cytosol of *P. vannamei* [16]. T3SS consists of T3SS1 and T3SS2, located in chromosome 1 and chromosome 2, respectively [17], and both of them mediate toxin secretion, leading to cytotoxicity and enterotoxicity in *P. vannamei* during infection [18]. Besides T3SS, flagella is another important virulence factor in *V. parahaemolyticus* that provides motility and confers various interactions such as invasion, adhesion, and biofilm formation between bacteria and the hosts [19, 20]. These virulence factors may attack immune-related proteins to suppress the host immune systems, which then potentially leads to disease development [21]. Therefore, studying specific interactions between *V. parahaemolyticus* proteins involved in virulence and *P. vannamei* immune-related proteins may enhance the knowledge of bacteria infection in shrimps.

Molecular docking, specifically protein-protein docking, followed by molecular dynamics (MD) simulation, are valuable computational approaches that can be used to study protein-protein interaction (PPI). Protein-protein docking predicts possible binding sites between two protein structures [22], while MD simulations are used to validate the prediction of protein-protein docking and its binding affinity [23]. Protein-protein docking prediction and MD simulations are used to reveal potential mechanisms of pathogen entry into shrimp and identify the interacting residues involved in infection [24–26]. Ji et al. [26] used this approach to identify the binding of white spot syndrome virus (WSSV) envelope protein VP28 to the secondary beta-sheet of Rab7 residues of *P. chinensis* at LEU73 to ASP86. They also predicted GLU81,

PHE77, and ASP76 of *Pc*Rab7 as possible interaction sites for the *Pc*Rab7-VP28 complex, suggesting their potential roles in WSSV infection of *P. chinensis* [26]. Similarly, the possible interaction sites for WSSV and Rab7 protein of *P. monodon* were predicted at ARG69-SER74, VAL75-ILE143, LEU73-ILE143, ARG79-ASN144, and ALA198-ALA182, and these interaction sites contributed to the formation of *Pm*Rab7-VP28 complex [24]. Chitin-binding protein (CBP) binds to envelop protein VP24 of WSSV *via* the interactions between GLY1-TYR5, THR5-HIS4, ASP88-ASN2, and CYS86-LEU7, which suggested the potential WSSV entry mechanism in *P. monodon* [25]. These findings illustrate the ability of protein-protein docking and MD simulation to determine specific points of interactions between host and pathogen proteins and identify binding modes and the dynamics of protein complexes.

Rosilan et al. [27] have conducted a host-pathogen PPI (HP-PPI) network analysis between *P. vannamei* and *V. parahaemolyticus* using interolog- and domain-based methods and several pairs of interactions that might be involved in *V. parahaemolyticus* infection in *P. vannamei* were identified. Those interactions are identified between ferritin-HrpE/YscL family type III secretion apparatus protein, protein kinase domain-containing protein-chemotaxis CheY protein, and G protein-coupled receptor kinase (GPCR)-chemotaxis CheY protein [27]. This study aims to elucidate the above-mentioned interactions by determining the binding sites and evaluating their affinity using structure-based approaches such as protein-protein docking and MD simulation. The binding sites on the host (*P. vannamei*) and pathogen (*V. parahaemolyticus*) proteins and their binding affinities are highlighted, and the information will provide an in-depth understanding of the molecular interactions between host and pathogen, which may be useful in future studies to improve shrimp aquaculture production, such as developing strategies to impede or prevent pathogen entry into the shrimp immune system.

Computational methods

Retrieval of three-dimensional (3D) models

The protein sequence of *P. vannamei* (ferritin–A0A423SK59; protein kinase domain-containing protein–A0A423SQ07; G protein-coupled receptor kinase (GPCR)–A0A423TB58) and *V. parahaemolyticus* (HrpE/YscL family type III secretion apparatus protein–A0A6H0JL36; Chemotaxis CheY protein–Q79YX1) were retrieved from UniProt database (https://www.uniprot.org/) [28]. The three-dimensional (3D) structure model of respected proteins were retrieved from the AlphaFold protein structure database (https://alphafold.ebi.ac.uk/) [29, 30]. AlphaFold is an artificial intelligence (AI) system developed by DeepMind, has been ranked as a top structure prediction method in Critical Assessment of Structure Prediction 14 (CASP14) [30] due to its ability to accurately predict 3D model of a protein from its amino acids. The information on each protein structure, such as alpha (α)-helices, beta (β)-sheets, domain, and two-dimensional (2D) topology diagram, was retrieved from PDBsum (http://www.ebi.ac.uk/thornton-srv/databases/pdbsum/) [31]. PDBsum is a web server that provides 3D structure information such as protein domains, interactions, secondary structures, PROCHECK analysis, enzyme reaction, and domain architecture [31]. The quality of the 3D protein structure of *P. vannamei* and *V. parahaemolyticus* was validated to predict the reliability of the retrieval 3D protein structure using ERRAT (https://saves.mbi.ucla.edu/) [32], PROCHECK (https://saves.mbi.ucla.edu/) [33, 34], MolProbity (http://molprobity.biochem.duke.edu/index.php) [35], protein quality predictor (ProQ) (https://proq.bioinfo.se/cgi-bin/ProQ/ProQ.cgi) [36], and protein structure analysis (ProSA) (https://prosa.services.came.sbg.ac.at/prosa.php) [37, 38].

Protein-protein docking of *Penaeus vannamei—Vibrio parahaemolyticus* complex

In protein-protein docking, possible binding sites between the protein complexes, i.e. ferritin-HrpE/YscL family type III secretion apparatus protein (Complex 1), protein kinase domain-containing protein-chemotaxis CheY protein (Complex 2), and GPCR-chemotaxis CheY protein (Complex 3) were determined using HawkDock server (http://cadd.zju.edu.cn/hawkdock/) [39–41]. HawkDock uses ATTRACT docking algorithm to analyse PPI and identify key docking residues involved in binding complexes. The HawkRank scoring function determines the docked score and MM/GBSA for free energy decomposition analysis based on weighted energy terms, such as van der Waals potentials, electrostatic potentials, and desolvation potentials [40]. HawkDock generates top ten docked proteins, and the best-docked model is selected based on the lowest docked score [42]. The best-docked model was visualised, and the hydrogen bond (H-bond) was identified within 5 angstroms (Å) of *P. vannamei* and *V. parahaemolyticus* protein using PyMOL (https://pymol.org/2/) [43]. The best-docked model was then submitted to the PDBsum server to obtain the total and type of other interactions, i.e., salt bridges and non-bonded contacts between protein residues.

Molecular dynamics (MD) simulation of protein complexes

MD simulation was conducted on the best-docked models using GROMACS software (https://www.gromacs.org/) [44, 45] to evaluate the protein complexes' binding stability *via* Optimised Potential for Liquid Simulations (OPLS) force fields for each protein complex [46]. MD simulation run was completed after the completion of three stages, i.e., (i) system neutralisation, (ii) system energy minimisation, and (iii) system equilibration. Each protein complex or system was solvated using the spc216 water model in a cubic box of 1.0 nanometer (nm). Sodium ion (NA^+) was added to neutralise the system for energy minimisation. The steepest descent minimisation algorithm was applied to set up the system energy minimisation with the maximum force below 1000 kJ/mol of 50,000 maximum steps. The systems equilibration was performed on 50 picoseconds (ps) using the Particles Mesh Ewald (PME) electrostatics approach to compute long-range electrostatics, contributing to reliable energy estimates [47]. The 100 nanoseconds (ns) MD simulations were conducted in triplicate by collecting data every 100 ps after the systems were equilibrated at a constant temperature of 300 Kelvin (K). The trajectory file of root-mean-square deviation (RMSD), the radius of gyration (R_g), H-bond, and distance was analysed by computing the average from triplicate MD simulation runs to measure the binding affinity between protein complexes. Pre- and post-MD structures were superimposed to observe conformational changes, structural stability, and PPI. Several tools were utilised to analyse the MD simulation results, such as visual molecular dynamics (VMD) (https://www.ks.uiuc.edu/Research/vmd/) [48] for displaying and animating MD simulation trajectories using 3D graphics, Xmgrace (https://plasma-gate.weizmann.ac.il/Grace/) [49] for plotting the data of RMSD, R_g, H-bond, and distance and PyMOL to visualise the extracted 3D structures of complex after MD simulation.

Principal component analysis (PCA)

The collective motion of each protein complex from the MD simulation trajectories was analysed using principal component analysis (PCA) over 100 ns in triplicate MD simulation using GROMACS [44]. PCA was conducted by creating the covariance matrix of backbones atoms using the gmx covar tool. The first two principal components, PC1 (projection on eigenvector 1) and PC2 (projection of eigenvector 2), were analysed using the gmx anaeig tool to observe

the global motion of Complex 1, Complex 2, and Complex 3. The results of PCA (PC1 and PC2) were visualised using the scatter plot for each replicate.

Cluster analysis

The most representative protein conformation throughout the 100 ns simulation was determined by further analysing the MD simulation trajectories of each complex through clustering analysis using the Gromos clustering algorithm [50] in GROMACS. The trajectory file (.xtc) of each replicate was combined for each complex using the gmx trjcat module before clustering. The gmx cluster module was used to cluster the MD simulation trajectories based on 0.2 nm RMSD cut-off values. The .pdb file of top cluster coordinates (first coordinates, middle coordinates, last coordinates) for each complex was extracted from MD simulation trajectories. The interacting residues involved in *P*. *vannamei*-*V*. *parahaemolyticus* PPI of each coordinate was retrieved by PDBsum to compare the interactions and visualised using PyMOL.

Molecular mechanics poisson-boltzmann surface area (MM/PBSA) calculation

Molecular Mechanics Poisson-Boltzmann surface area (MM/PBSA) approach was used to calculate the binding free energy of several energy components, i.e., (i) van der Waals force (VDWAALS), (ii) electrostatic energy (EEL), (iii) the electrostatic contribution to the solvation free energy (EGB), and (iv) non-polar contribution to the solvation free energy (ESURF) between *P*. *vannamei* and *V*. *parahaemolyticus* proteins. Kollman et al. [51] developed MM-PBSA as the end-point for free energy methods as it is more efficient and accurate than scoring functions and computationally inexpensive [52]. The binding free energy analysis was calculated using the MM-PBSA.py tool in Amber (https://ambermd.org/) [53]. The overall workflow is illustrated in Fig 1.

Results

Predicted structure of *Penaeus vannamei* and *Vibrio parahaemolyticus* proteins

The 3D protein structures and 2D topology diagrams of ferritin, protein kinase domain-containing protein, and GPCR protein from *P*. *vannamei* (Fig 2) and HrpE/YscL family type III secretion apparatus protein, chemotaxis CheY protein from *V*. *parahaemolyticus* (Fig 3), which starts with N-terminal and end with C-terminal were retrieved from AlphaFold and PDBsum. For *P*. *vannamei* proteins, ferritin was found to be a single-domain structure with six α-helices (Fig 2a), protein kinase domain-containing protein is a single-domain structure with 19 α-helices and two β-sheets that contain seven β-strands (Fig 2b), and GPCR was determined to be a multi-domain structure with 24 α-helices and four β-sheets that comprise 18 β-strands (Fig 2c). Meanwhile, both *V*. *parahaemolyticus* proteins, chemotaxis CheY protein and HrpE/YscL family type III secretion apparatus, were identified as single-domain structures with five α-helices and one β-sheet containing five β-strands (Fig 3a), four α-helices and one β-sheet containing four β-strands, respectively (Fig 3b). The 3D protein structures of *P*. *vannamei* and *V*. *parahaemolyticus* retrieved from AlphaFold demonstrated that these proteins are reliable and reasonable as good protein structures, as they are located within a range of good protein structures according to several 3D structure validation tools, including ERRAT, PROCHECK, MolProbity, ProQ, and ProSA web server (S1 File).

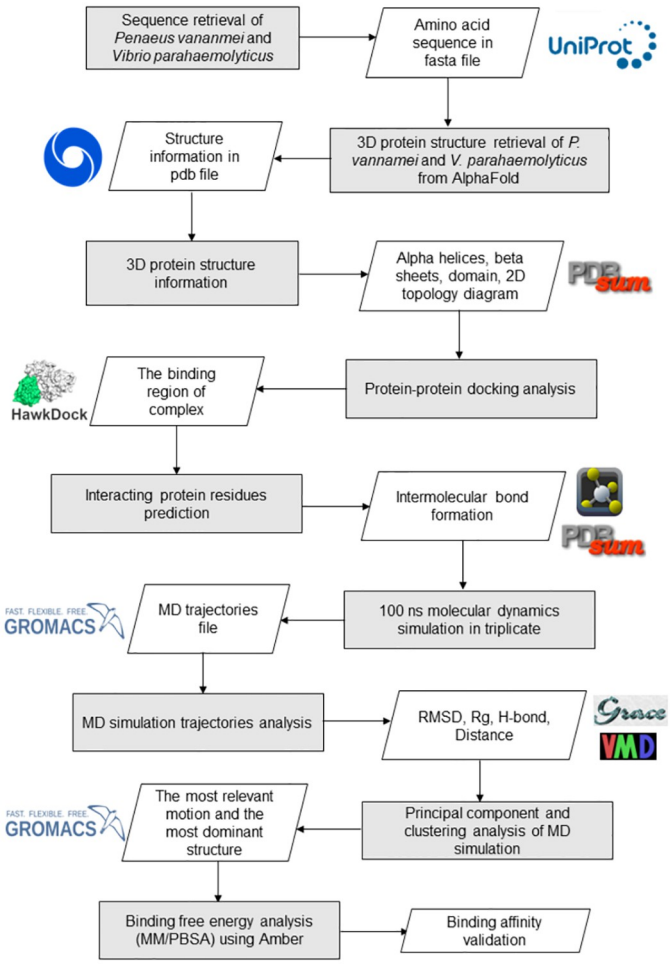

Fig 1. Workflow of the study. The rectangle box represents the method. The parallelogram indicates the output.

https://doi.org/10.1371/journal.pone.0297759.g001

Protein complexes of *Penaeus vannamei-Vibrio parahaemolyticus*

The best-docked pose of protein complex for Complex 1 (Ferritin-HrpE/YscL family type III secretion apparatus protein) (S1 Fig), Complex 2 (Protein kinase domain-containing protein-Chemotaxis CheY protein) (S2 Fig), and Complex 3 (GPCR-Chemotaxis CheY protein) (S3 Fig) were predicted by HawkDock server with a docking score of -4319.76, -5271.39, and -4725.57, respectively (S1 Table).

H-bond, salt bridge (Table 1), and non-bonded contact interaction (S2 Table) with a distance within 5Å between *P. vannamei* and *V. parahaemolyticus* proteins residues were obtained using PyMOL tools and PDBsum. In Complex 1, the H-bond was formed between VAL154-SER42 and ARG153-VAL27 (S1a Fig), and 109 non-bonded contacts were formed in 24 interacting residues (S1b Fig). In Complex 2, H-bond was formed between THR20-HIS79, SER23-ASP51, GLY25-MET1, GLY25-ASN2, and PRO94-LYS78 (S2a Fig). Meanwhile, the salt bridge formation was identified between ASP100-ARG72, and 88 non-bonded contacts were formed between 24 interacting residues (S2b Fig). In complex 3, H-bond was formed between ASN35-SER14, ARG38-ASP12, ARG48-GLU34, and GLN368-ASP74 (S3a Fig), salt bridge bonds were identified between ARG38-ASP36, ARG41-ASP36, ARG48-GLU34, and 139 non-bonded contacts were formed in 29 interacting residues (S3b Fig).

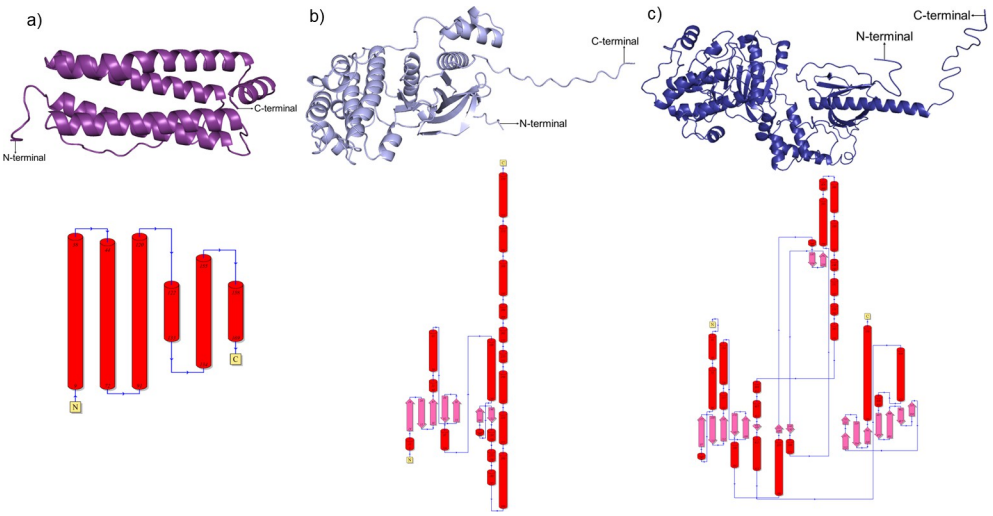

Fig 2. 3D models and 2D topology diagrams of *Penaeus vannamei* proteins. (a) Ferritin (A0A423SK59), (b) Protein kinase domain-containing protein (A0A423SQ07), (c) G protein-coupled receptor (A0A423TB58). The red cylinder in the topology diagram represents α-helix, and the pink arrow indicates the β-strand to form β-sheet. A small blue arrow illustrates the direction of the protein chain from N- to C-terminal.

https://doi.org/10.1371/journal.pone.0297759.g002

Molecular dynamics (MD) simulation of protein complexes

The binding affinity of each protein complex of *P. vannamei* and *V. parahaemolyticus* was determined using triplicate MD simulation over 100 ns (Figs 4–6) by analysing the average RMSD, H-bond, R_g, and distance. During the 100 ns simulation, RMSD values of Complex 1, Complex 2, and Complex 3 were 0.530 nm, 0.632 nm, and 1.088 nm, respectively (Fig 7a). The rigidity of each complex was analysed based on the R_g with a value of 2.302 nm, 2.381 nm, and 3.076 nm, accordingly (Fig 7b). Nine H-bonds were identified in Complex 1 and Complex 2,

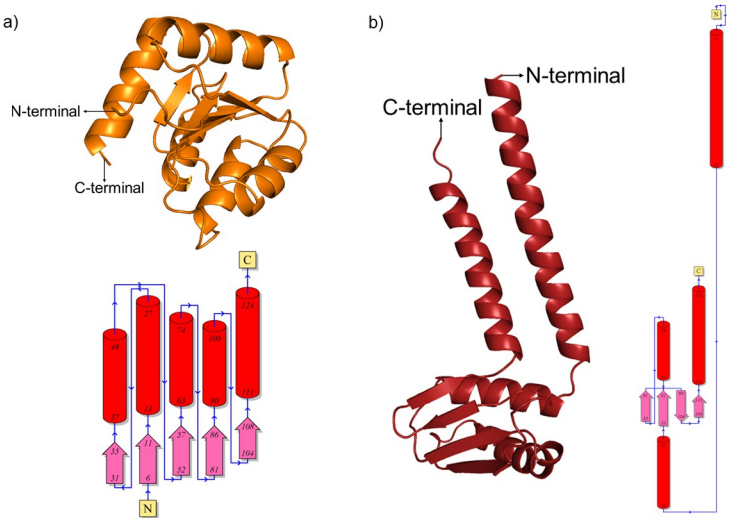

Fig 3. 3D models and 2D topology diagrams of *Vibrio parahaemolyticus* proteins. (a) Chemotaxis CheY protein (Q79YX1), (b) HrpE/YscL family type III secretion apparatus protein (A0A6H0JL36). The red cylinder in the topology diagram represents α-helix, and the pink arrow indicates the β-strand to form β-sheet. A small blue arrow illustrates the direction of the protein chain from N- to C-terminal.

https://doi.org/10.1371/journal.pone.0297759.g003

Table 1. Distance of a hydrogen bond and a salt bridge between the interacting residues.

Protein complex	*Penaeus vannamei* residue	*Vibrio parahaemolyticus* residue	Interaction bond	Distance (Å)
Complex 1	VAL154	SER42	Hydrogen	2.60
	ARG153	VAL27	Hydrogen	2.60
Complex 2	THR20	HIS79	Hydrogen	2.99
	SER23	ASP51	Hydrogen	2.83
	GLY25	MET1	Hydrogen	2.74
	GLY25	ASN2	Hydrogen	2.74
	PRO94	LYS78	Hydrogen	2.95
	ASP100	ARG72	Salt bridge	3.81
Complex 3	ASN35	SER14	Hydrogen	2.89
	ARG38	ASP12	Hydrogen	3.23
	ARG48	GLU34	Hydrogen	3.14
	GLN368	ASP74	Hydrogen	3.14
	ARG38	ASP36	Salt bridge	3.64
	ARG41	ASP36	Salt bridge	3.64
	ARG48	GLU34	Salt bridge	3.14

https://doi.org/10.1371/journal.pone.0297759.t001

and eight H-bonds were identified in Complex 3 (Fig 7c) during 100 ns simulation, where the significance of interaction refers to high H-bond formation. After 100 ns simulation, the distance between *P. vannamei* and *V. parahaemolyticus* proteins in Complex 1 and Complex 3 was 0.150 nm, while in Complex 2 the distance between both proteins was 0.153 nm (Fig 7d). After 100 ns MD simulation, each complex undergoes conformational changes in each replicate, which is represented in Fig 8.

Principal component analysis (PCA) of molecular dynamics (MD) simulation trajectories

The scatter plot represented in Fig 9 visualises the projection of the MD simulation trajectory onto the first two principal components, PC1 and PC2, which showed a significantly large

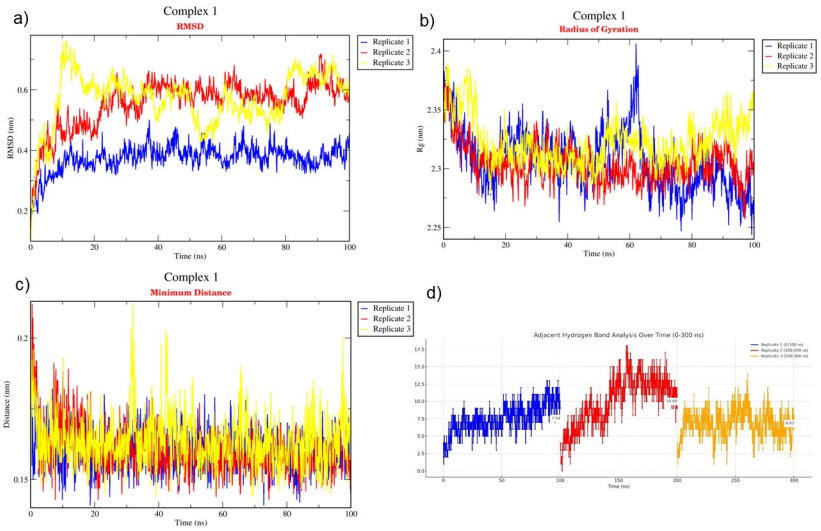

Fig 4. Triplicate of 100 ns MD simulation of Complex 1. (a) RMSD, (b) R_g, (c) Distance, (d) H-bond. The blue plot indicates replicate 1, the red plot represents replicate 2, and the yellow plot illustrates replicate 3.

https://doi.org/10.1371/journal.pone.0297759.g004

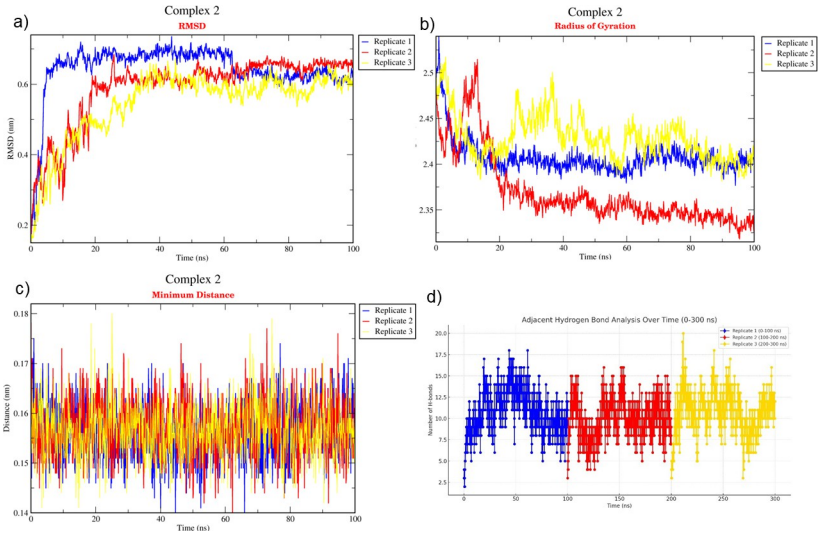

Fig 5. Triplicate of 100 ns MD simulation of Complex 2. (a) RMSD, (b) R_g, (c) Distance, (d) H-bond. The blue plot indicates replicate 1, the red plot represents replicate 2, and the yellow plot illustrates replicate 3.

https://doi.org/10.1371/journal.pone.0297759.g005

motion of Complex 3 compared with Complex 1 and Complex 2. Referring to Fig 9, each data point in the plot represents a distinct conformation of the protein complex at a specific time point in the simulation, mapped in terms of its deviation along the major axes of conformational variability. The PC1 axis (horizontal) captures the most significant mode of structural variation, encompassing large-scale conformational changes, while the PC2 axis (vertical) represents the second most dominant mode, often associated with more subtle or local fluctuations. The distribution and density of the points across this 2D space provide a visual representation of the conformational diversity explored by the protein, highlighting regions of

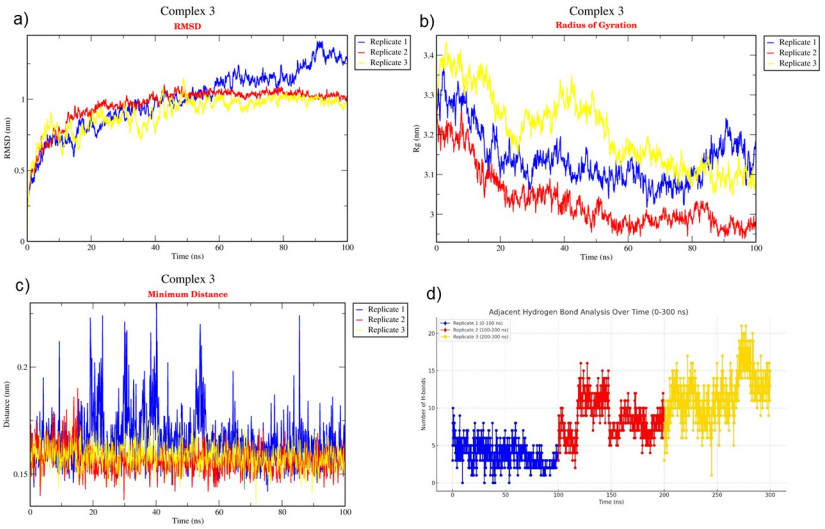

Fig 6. Triplicate of 100 ns MD simulation of Complex 3. (a) RMSD, (b) R_g, (c) Distance, (d) H-bond. The blue plot indicates replicate 1, the red plot represents replicate 2, and the yellow plot illustrates replicate 3.

https://doi.org/10.1371/journal.pone.0297759.g006

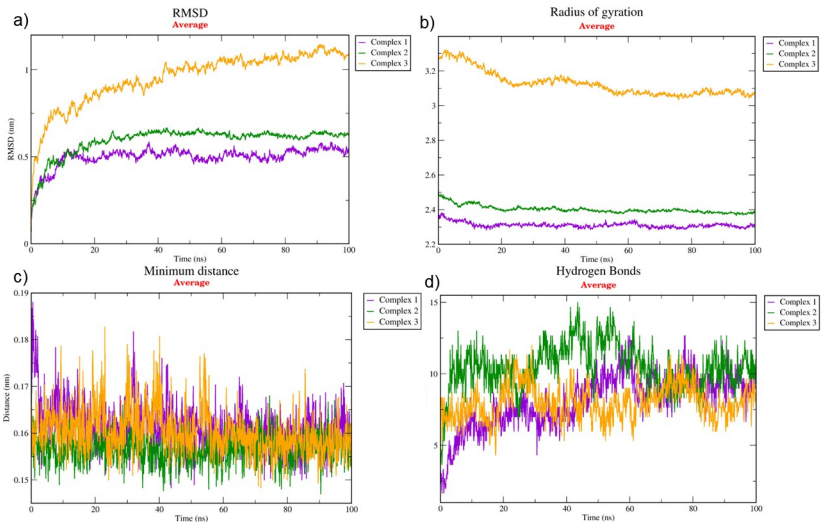

Fig 7. Average of triplicate 100 ns MD simulation. (a) RMSD, (b) R_g, (c) H-bond, and (d) distance. The purple plot indicates Complex 1 (Ferritin-HrpE/YscL family type III secretion apparatus protein), the red plot represents Complex 2 (Protein kinase domain-containing protein-Chemotaxis CheY protein), and the orange plot illustrates Complex 3 (GPCR-Chemotaxis CheY protein).

https://doi.org/10.1371/journal.pone.0297759.g007

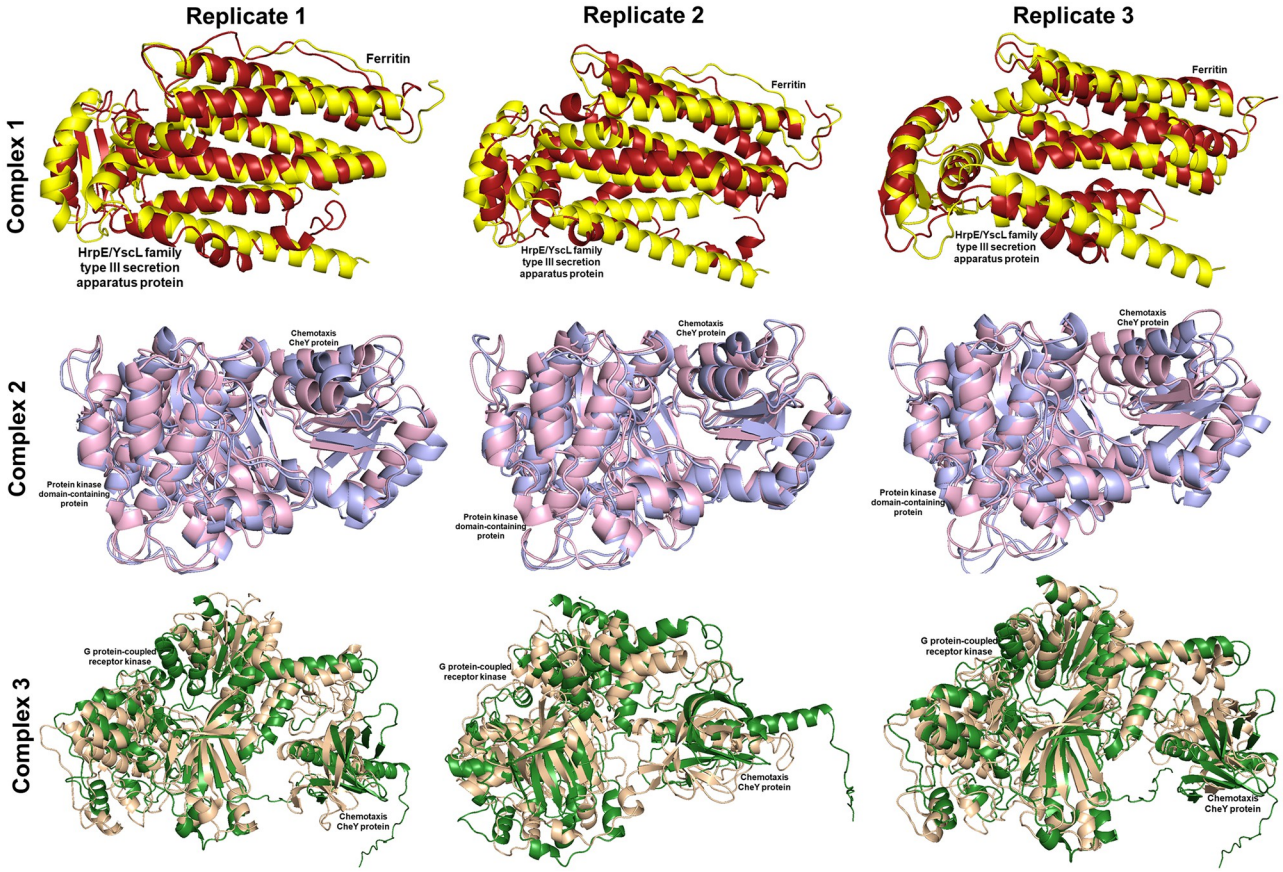

Fig 8. Superimposition of pre-MD and post-MD. The pre-MD for Complex 1, Complex 2, and Complex 3 are represented with yellow, pink, and green, respectively. Meanwhile, the post-MD for Complex 1, Complex 2, and Complex 3, are illustrated with red, purple, and brown, respectively.

https://doi.org/10.1371/journal.pone.0297759.g008

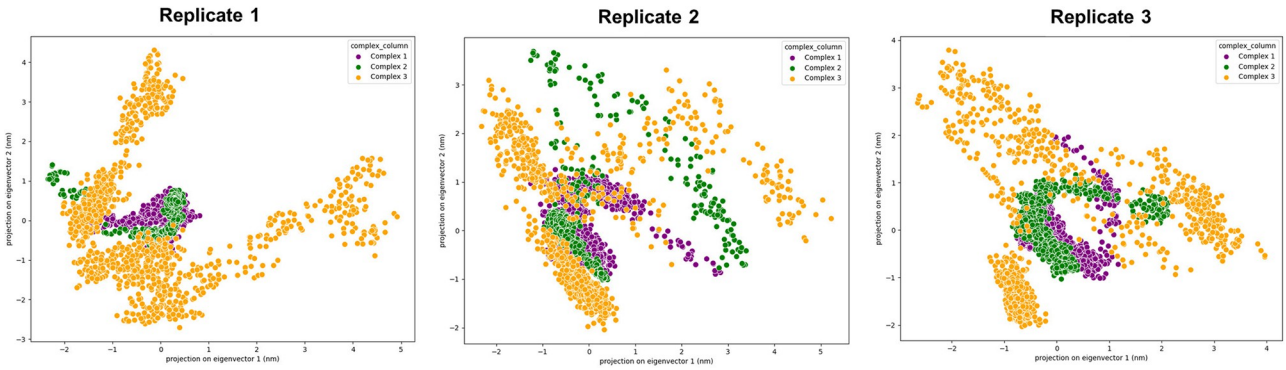

Fig 9. Scatter plot of PCA over 100 ns MD simulation. The purple plot represents Complex 1, the green plot indicates Complex 2, and the yellow plot shows Complex 3.

https://doi.org/10.1371/journal.pone.0297759.g009

conformational stability (dense clusters) and transition (sparse or bridge-like connections between clusters).

Clustering analysis of molecular dynamics (MD simulation trajectories)

The clustering analysis of Complex 1 formed 11 clusters with an average RMSD of 0.251 nm, Complex 2 formed 8 clusters with a 0.229 nm average RMSD, and 73 clusters with a 0.560 nm average RMSD formed in Complex 3. From the clusters, the most dominant structure of Complex 1, Complex 2, and Complex 3, located in time frames of 62.60 ns, 71.20 ns, and 94.80 ns, respectively, throughout the 100 ns simulation, which these dominant structure were located in replicate 1 in each complex. The comparative interaction of different cluster coordinates, Complex 1 (0 ns, 62.60 ns, 100 ns), Complex 2 (0 ns, 71.20 ns, 100 ns), and Complex 3 (0 ns, 94.80 ns, 100 ns), showed the changes of protein conformational and the total of *P. vannamei*-*V. parahaemolyticus* PPI (Fig 10).

Molecular mechanics poisson-boltzmann surface area (MM/PBSA) calculation

MM/PBSA calculated the binding free energy in VDWAALS, EEL, EGB, and ESURF (Table 2) after each complex (Complex 1, Complex 2, Complex 3) was stabilised *via* simulation. The average total binding energy of Complex 1, Complex 2, and Complex 3 is -22.50 kJ/mol, -30.20 kJ/mol, and -26.27 kJ/mol, respectively.

Discussion

In-depth knowledge of the interactions between host and pathogen proteins is essential to uncover the mechanisms of bacterial infection that may contribute to the advancement of shrimp farming production [54]. The information on the PPI between proteins in *P. vannamei* and *V. parahaemolyticus* is able to deepen the understanding of how *V. parahaemolyticus* interacts, evades, and invades the host's immune system. This knowledge may be essential in the development of treatment and prevention strategies to fight pathogenic infection in shrimp farming, as this problem remains a major challenge to the success of shrimp aquaculture production [55]. In this study, the interactions between *P. vannamei* and *V. parahaemolyticus* proteins, i.e., ferritin-HrpE/YscL family type III secretion apparatus protein, protein kinase domain-containing protein-chemotaxis CheY protein, GPCR-chemotaxis CheY protein were

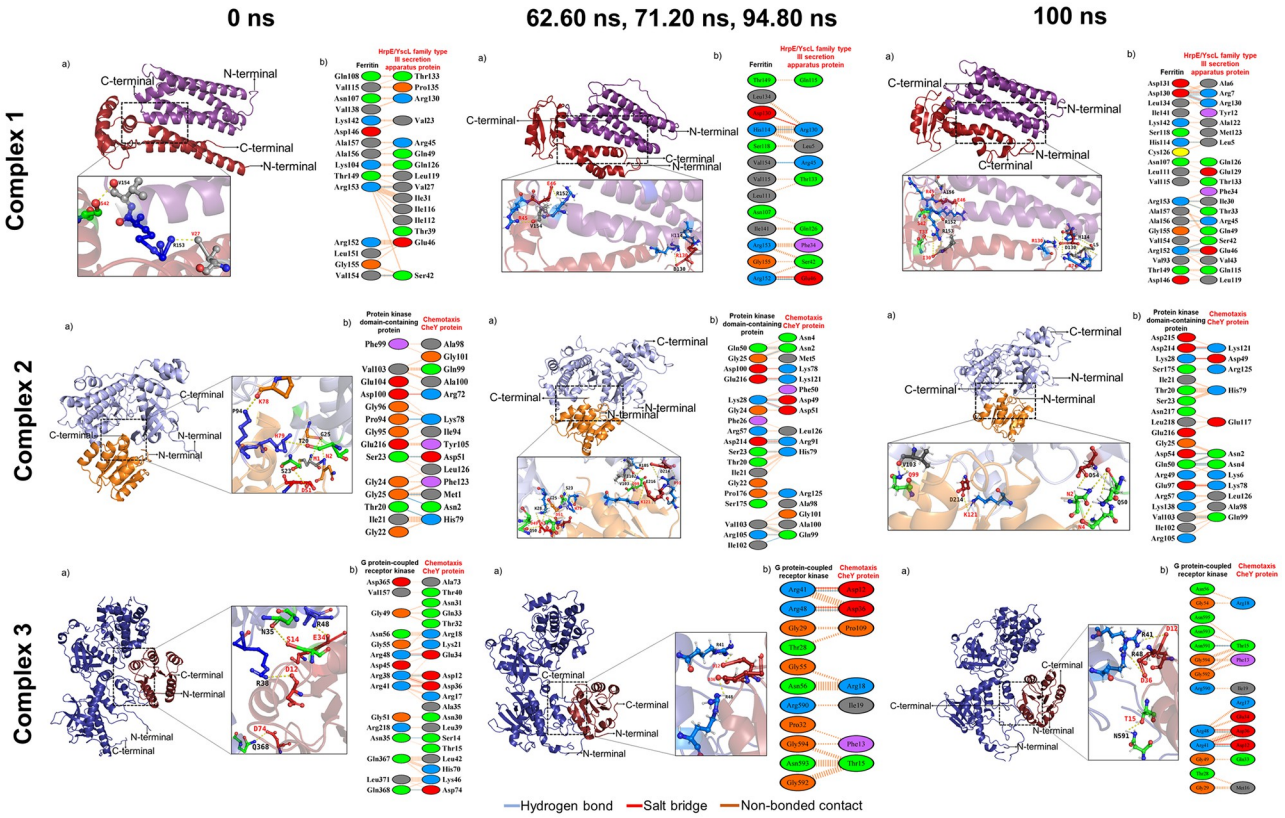

Fig 10. Comparative PPI of the top cluster coordinates from MD simulation trajectories.

https://doi.org/10.1371/journal.pone.0297759.g010

elucidated using a structure-based approach by identifying the binding region that involved several amino acid residues, which suggest these interacting residues might taking part in infection.

Ferritin, a major storage protein for iron, is crucial in energy metabolism, nucleic acid synthesis, and cell proliferation in all organisms [56]. This protein can also protect cells from the detrimental effects of free iron, where the availability of free iron promotes cell and tissue damage *via* free radical formation [57]. During an infection, both host and pathogen compete for available nutrients, such as iron, which is stored within the ferritin protein. Thus, in response to infection, the host protects its iron metabolism by limiting its availability to the bacteria to prevent bacteria replication [58]. However, in several bacteria species, including *V*. *parahaemolyticus*, *Yersinia pestis*, and *Salmonella enterica serovar* Typhimurium, ferritin is used as an iron source during infection and damages the host's membrane by bacterial secretion system, T3SS [59].

Table 2. Average MM/PBSA free energy of *Penaeus vannamei*-*Vibrio parahaemolyticus* complex of triplicate MD simulation.

Complex	VDWAALS (kJ/mol)	EEL (kJ/mol)	EGB (kJ/mol)	ESURF (kJ/mol)	GGAS (kJ/mol)	GSOLV (kJ/mol)	Average (kJ/mol)
Complex 1	-65.47	115.71	-64.32	-8.41	50.24	-72.74	-22.50
Complex 2	-89.30	-395.90	467.52	-12.52	-485.21	455.00	-30.20
Complex 3	-48.81	-239.48	271.10	-9.07	-288.30	262.02	-26.27

https://doi.org/10.1371/journal.pone.0297759.t002

Meanwhile, protein kinases are vital in crustacean cellular processes, such as cell division, proliferation, apoptosis, and differentiation, especially during immune response [60]. Protein kinases are also involved in the regulation of flagella length during the initiation of infection in the host [61, 62]. In bacterial infection, kinase proteins, known as mitogen-activated protein kinases, suppress the inflammatory reactions in orange-spotter grouper, *Epinephelus coioides*, which were infected with the flagellar basal-body rod protein (*flgC*) gene associated with chemotaxis and motility regulation [63]. In addition, CheY is a gene of chemotaxis that acts as a response regulator, which controls the motility of bacterial flagellar through the binding process to the cytoplasmic switch complex of the flagellar motor [64].

GPCR is vital in crustacean immune processes and aids in host defence, as indicated by its up-regulation after being challenged by gram-positive bacteria [65]. A GPCR family member, CXCR2, was also found to bind with the interleukin-8 (IL8) receptor during *V*. *parahaemolyticus* infection in large yellow croaker, *Larimichthys crocea*, in which the interaction initiated both chemotaxis and phagocytosis processes [66]. The involvement of these proteins during infection suggests their interactions might mediate the cross-talks between *P*. *vannamei* and *V*. *parahaemolyticus*. Hence, the interactions between these proteins were structurally observed to determine their specific binding sites and affinity.

Protein-protein docking is essential in the identification of possible binding sites of protein pairs [67]. Using the validated retrieved 3D structures from AlphaFold, the interactions between three protein pairs in *P*. *vannamei* and *V*. *parahaemolyticus*, namely Complex 1 (Ferritin-HrpE/YscL family type III secretion apparatus protein), Complex 2 (Protein kinase domain-containing protein-Chemotaxis CheY protein), and Complex 3 (GPCR-Chemotaxis CheY protein), were docked to predict their binding site and the best-docked of complexes were ranks according to docking score. Docked complexes were ranked based on the top 10 docked scores, which is important to validate the binding mode and to predict the binding affinity [68], and the docked pose with the lowest score in the top 10 ranks indicates the best-docked complexes [42]. The specific binding sites identified in the *P*. *vannamei*-*V*. *parahaemolyticus* protein complexes with a distance of interacting residues within 5Å are defined as closely interacting [69]. The *P*. *vannamei* and *V*. *parahaemolyticus* complexes that consist of Complex 1, Complex 2, and Complex 3 indicate that the H-bond is formed in all complexes between specific interacting residues. H-bond is pivotal in controlling molecular interaction and is significant in protein-protein binding, thereby determining the specificity between *P*. *vannamei* and *V*. *parahaemolyticus* [70, 71]. The identified H-bond in the host-pathogen interaction involving specific residues indicates strong binding due to the higher binding affinity, suggesting the binding is important for the infection of pathogenic agents into the host cells. This mechanism is illustrated in the respiratory syncytial virus (RSV) and the severe acute respiratory syndrome coronavirus (SARS-CoV) infection in humans [72, 73], and the WSSV infection in crustaceans [25, 74].

Other types of intermolecular bonds in PPI, such as salt bridges and non-bonded contacts, are also important, and the information can be obtained in PDBSum. These bonds are important to maintain the structure's stability and rigidity [71, 75–77]. Salt bridges are also highly involved in protein folding, as highlighted in host-pathogen studies investigating the innate immune response and adaptation of an antarctic notothenioid fish, *Trematomus bernacchii*, during bacterial infection in low-temperature environments [78]. Non-bonded contacts contribute to the stabilisation of insulin-like growth factor binding proteins (IGFBPs) and insulin-like growth factor (IGF) binding, suggesting the specific interaction interface that mediates the understanding of insulin endocrinology in the eastern spiny lobster, *Sagmariasus verreauxi* [79]. These examples show the importance of the salt bridges and non-bonded contact interactions in understanding how proteins interact with functional residues that are involved in the

stabilisation of the protein complex and binding to the specific residues in protein complexes [71].

The results from this study identified significant interacting residues in specific binding sites and bond formation that occur between selected proteins of *P. vannamei* and *V. parahaemolyticus*, which can be experimentally validated to measure their strength and involvement in executing their functions. However, the interaction of H-bond is significantly relevant in drug discovery and development in future studies, which affects the biological activity, pharmacokinetics, and physicochemical properties of drugs [80] that are able to inhibit the pathogen infection compared to the salt bridge and non-bonded interactions as their roles in drug design have remained unclear [76]. In current study, the interacting residues that involved H-bond formation in each complex, Complex 1 (VAL154-SER42, ARG153-VAL27), Complex 2 (THR20-HIS79, SER23-ASP51, GLY25-MET1, GLY25-ASN2, PRO94-LYS78), and Complex 3 (ASN35-SER14, ARG38-ASP12, ARG48-GLU34, GLN368-ASP74) might be a crucial region for inhibition of pathogenic infection. In addition, the MD simulation can further validate the binding affinity of *P. vannamei*-*V. parahaemolyticus* complexes that reveal the strength of their interaction by several parameters, including the H-bond.

The triplicate MD simulation with 100 ns trajectories generates RMSD, R_g, H-bond, and distance that are used to measure the stability and behaviour of protein complexes. The RMSD values of Complex 1, Complex 2, and Complex 3 were 0.530 nm, 0.632 nm, and 1.088 nm, respectively. The RMSD graph (Fig 7a) shows a steady and constant plot without huge spikes, suggesting that the protein complex remained stable throughout the 100 ns simulation. Low RMSD values indicate that the protein complex is highly stable [81]. However, based on the results, Complex 3 has a significantly higher RSMD value than Complex 1 and Complex 2 due to the larger size of the protein complex, which could influence the higher RMSD [82]. In terms of molecular dynamics, a larger protein complex frequently exhibits more dynamic behaviour by involving conformational changes, protein movements, and interactions. The R_g is an important parameter that indicates the compactness and proper folding of the protein [83, 84]. In the present study, the R_g values of each protein complex were 2.250 nm, 2.398 nm, and 3.161 nm, which depicted the lower R_g values, proving the high stability of the complex and revealing the compactness is consistent during 100 ns simulation [84]. H-bond between selected proteins in *P. vannamei* and *V. parahaemolyticus* proteins are calculated to evaluate the binding affinity and confirm the stable interaction [85]. Nine H-bonds in Complex 1 and 2 and eight H-bonds in Complex 3 suggest strong binding stability [23]. The distances between selected proteins in *P. vannamei* and *V. parahaemolyticus* proteins are calculated to confirm their binding affinities. We calculated the distances between two proteins in each complex, i.e., 0.148 nm for Complex 1 and Complex 2 and 0.151 nm for Complex 3. A distance below 3.5Å reveals a strong binding affinity between proteins [86]. Based on the RMSD, R_g, H-bond, and distance values, Complex 1, Complex 2, and Complex 3 are highly stable and exhibit strong binding with each other.

PCA is a widely used approach to determine the dominant molecular motion in the MD trajectory [87] and offers a compact yet insightful view into the dynamic behaviour of Complex 1, Complex 2, and Complex 3 during the simulation period performed in this study. Notably, several distinct clusters can be observed, indicative of preferred conformational states or energy wells where the protein spends a significant portion of its simulation time [88]. The central cluster, characterised by a high density of data points, likely represents the most energetically favourable and most frequently adopted conformation under the simulated conditions. Its centrality and density suggest a conformational state with considerable stability, potentially correlating with the protein's functional resting or inactive state. The extent of spread along the PC1 axis underscores the protein's ability to undergo large-scale

conformational rearrangements, which could be critical for its biological function, such as ligand binding, allosteric regulation, or PPI [89]. In contrast, the spread along the PC2 axis, while covering a smaller range, indicates additional degrees of conformational freedom [89]. These might correspond to more localised movements, such as side-chain rearrangements, loop flexibilities, or domain swiveling, complementing the broader structural transitions captured by PC1. Furthermore, the presence of points forming less dense regions or paths between the main clusters is particularly intriguing. These points may represent transitional conformations that the protein adopts as it shifts from one stable state to another. Such transitions are often crucial in understanding the mechanism of action of the protein, as they can reveal the pathways through which conformational changes propagate [90]. Identifying these transitional states can provide key insights into the dynamic aspects of the protein's function, offering potential targets for therapeutic intervention or modulation.

In addition, the MD simulation generates millions of data, making it difficult to determine each complex's most representative protein conformations. The clustering analysis is one of the advanced analyses used to identify the most dominant protein conformation throughout the simulation in the specific time frames, which is pivotal in elucidating the behaviour of protein complexes [91]. The comparative interaction across trajectories demonstrated changes in the interaction between *P*. *vannamei* and *V*. *parahaemolyticus* proteins as a result of protein complex conformational changes throughout the simulation, which could aid in a better understanding of the interactions. The MM-PBSA analysis is an end-point for docking pose evaluation, structure stability, and binding affinity determination by calculating the binding free energy in VDWAALS, EEL, EGB, and ESURF between the selected proteins in *P*. *vannamei* and *V*. *parahaemolyticus* [52]. In this study, VDWAALS, EEL, and ESURF contribute negatively and favour protein binding in Complex 2 and Complex 3, while the contribution of EGB was very low. In complex 1, the protein binding was negatively contributed by VDWAALS, EGB, and ESURF, while there are comparatively low EEL contributions. The results indicate the high binding affinity was -30.20 kJ/mol in Complex 2, followed by Complex 3 and Complex 1 with the average binding energy of -26.27 kJ/mol and -22.50 kJ/mol, respectively. It suggests the high binding stability of the interaction between *P*. *vannamei* and *V*. *parahaemolyticus* due to the contribution of VDWAALS, EEL, EGB, and ESURF.

Conclusion

Protein-protein docking and MD simulation might provide the structural insight into host-pathogen interactions between *P*. *vannamei* and *V*. *parahaemolyticus*. The protein-protein docking predicts the structure of protein complexes and identifies the several interacting residues, specifically, the binding sites formed by H-bond during bacterial infection consist of VAL154-SER42, ARG153-VAL27 in Complex 1, THR20-HIS79, SER23-ASP51, GLY25--MET1, GLY25-ASN2, PRO94-LYS78 in Complex 2, and ASN35-SER14, ARG38-ASP12, ARG48-GLU34, GLN368-ASP74 in Complex 3 that might be a target region for drug development in the future study. From this study, protein binding in each complex is stable and strong, based on their RMSD, R_g, H-bond, and distance value during 100 ns MD simulation and MM-PBSA analysis used as an end-point in validating binding affinity revealed that the interaction of Complex 1, Complex 2, and Complex 3 was highly stable and strongly bind. Hence, this suggests that significant binding causes infection in *P*. *vannamei*. In-depth knowledge of the molecular interactions between the host and pathogen has improved our understanding of the mechanisms of bacterial infection, and it can be used in the biomarker

identification in shrimp diseases and to develop strategies for effective treatment by targeting specific binding sites involved in the host-pathogen interactions.

Supporting information

S1 Fig. Complex 1 (ferritin-HrpE/YscL family type III secretion apparatus protein). The purple chain indicates the ferritin protein from *P. vannamei*, and the red chain indicates the HrpE/YscL family type III secretion apparatus protein from *V. parahaemolyticus*.
(TIF)

S2 Fig. Complex 2 (protein kinase domain-containing protein-chemotaxis CheY protein). The blue chain indicates the protein kinase domain-containing protein from *P. vannamei*, and the maroon chain indicates the chemotaxis CheY protein from *V. parahaemolyticus*.
(TIF)

S3 Fig. Complex 3 (GPCR-chemotaxis CheY protein). The purple chain indicates the GPCR protein from *P. vannamei*, and the orange chain indicates the chemotaxis CheY protein from *V. parahaemolyticus*.
(TIF)

S1 File. Protein structure validation using ERRAT, PROCHECK, MolProbity, ProQ, and ProSA.
(DOCX)

S1 Table. Protein-protein docking score using the HawkDock server.
(DOCX)

S2 Table. Distance of non-bonded contact between *Penaeus vannamei* and *V. parahaemolyticus* residues.
(DOCX)

Acknowledgments

The authors thank the Bioinformatics program of the Institute of Climate Adaptation and Marine Biotechnology (ICAMB), Universiti Malaysia Terengganu (UMT), for providing computational facilities.

Author Contributions

Conceptualization: Nor Afiqah-Aleng.

Data curation: Nur Fathiah Rosilan, Muhamad Arif Mohamad Jamali, Azzmer Azzar Abdul Hamid, Nor Afiqah-Aleng.

Formal analysis: Nur Fathiah Rosilan, Muhamad Arif Mohamad Jamali, Siti Aishah Sufira, Azzmer Azzar Abdul Hamid.

Funding acquisition: Nor Afiqah-Aleng.

Supervision: Nor Afiqah-Aleng.

Writing – original draft: Nur Fathiah Rosilan.

Writing – review & editing: Muhamad Arif Mohamad Jamali, Khor Waiho, Hanafiah Fazhan, Noraznawati Ismail, Yeong Yik Sung, Zeti-Azura Mohamed-Hussein, Azzmer Azzar Abdul Hamid, Nor Afiqah-Aleng.

References

1. Kumar V, Roy S, Meena DK, Sarkar UK. Application of Probiotics in Shrimp Aquaculture: Importance, Mechanisms of Action, and Methods of Administration. Rev Fish Sci Aquac. 2016; 24(4):342–68. https://doi.org/10.1080/23308249.2016.1193841
2. Prathiviraj R, Rajeev R, Fernandes H, Rathna K, Lipton AN, Selvin J, et al. A gelatinized lipopeptide diet effectively modulates immune response, disease resistance and gut microbiome in *Penaeus vannamei* challenged with *Vibrio parahaemolyticus*. Fish Shellfish Immunol. 2021; 112:92–107. https://doi.org/10.1016/j.fsi.2021.02.018 PMID: 33675990
3. FAO. Fisheries & Aquaculture—Global aquaculture production. 2023 [cited 2023 Jul 27]. Database: FishStatJ [Internet]. https://wnternaw.fao.org/fishery/statistics-query/en/aquaculture.
4. Yu YB, Choi JH, Kang JC, Kim HJ, Kim JH. Shrimp bacterial and parasitic disease listed in the OIE: A review. Microb Pathog. 2022; 166:105545. https://doi.org/10.1016/j.micpath.2022.105545 PMID: 35452787
5. Letchumanan V, Chan KG, Lee LH. *Vibrio parahaemolyticus*: A review on the pathogenesis, prevalence, and advance molecular identification techniques. Front Microbiol. 2014; 5:1–13. https://doi.org/10.3389/fmicb.2014.00705 PMID: 25566219
6. Maralit BA, Jaree P, Boonchuen P, Tassanakajon A, Somboonwiwat K. Differentially expressed genes in hemocytes of Litopenaeus vannamei challenged with *Vibrio parahaemolyticus* AHPND (VP_{AHPND}) and VP_{AHPND} toxin. Fish Shellfish Immunol. 2018; 81:284–96. https://doi.org/10.1016/j.fsi.2018.06.054 PMID: 29966688
7. Kumar R, Ng TH, Wang HC. Acute hepatopancreatic necrosis disease in penaeid shrimp. Rev Aquac. 2020; 12(3):1867–80. https://doi.org/10.1111/raq.12414
8. Peña-Navarro N, Castro-Vásquez R, Vargas-Leitón B, Dolz G. Molecular detection of acute hepatopancreatic necrosis disease (AHPND) in *Penaeus vannamei* shrimps in Costa Rica. Aquaculture. 2020; 523:735190. https://doi.org/10.1016/j.aquaculture.2020.735190
9. Wang H, Wan X, Xie G, Dong X, Wang X, Huang J. Insights into the histopathology and microbiome of Pacific white shrimp, *Penaeus vannamei*, suffering from white feces syndrome. Aquaculture. 2020; 527:735447. https://doi.org/10.1016/j.aquaculture.2020.735447
10. Sriurairatana S, Boonyawiwat V, Gangnonngiw W, Laosutthipong C, Hiranchan J, Flegel TW. White Feces Syndrome of Shrimp Arises from Transformation, Sloughing and Aggregation of Hepatopancreatic Microvilli into Vermiform Bodies Superficially Resembling Gregarines. PLoS One. 2014; 9(6):2–9. https://doi.org/10.1371/journal.pone.0099170 PMID: 24911022
11. Alfiansah YR, Peters S, Harder J, Hassenrück C, Gärdes A. Structure and co-occurrence patterns of bacterial communities associated with white faeces disease outbreaks in Pacific white-leg shrimp *Penaeus vannamei* aquaculture. Sci Rep. 2020; 10(1):1–13. https://doi.org/10.1038/s41598-020-68891-6 PMID: 32686764
12. Yang F, Xu L, Huang W, Li F. Highly lethal *Vibrio parahaemolyticus* strains cause acute mortality in *Penaeus vannamei* post-larvae. Aquaculture. 2022; 548(P1):737605. https://doi.org/10.1016/j.aquaculture.2021.737605
13. Zou Y, Xie G, Jia T, Xu T, Wang C, Wan X, et al. Determination of the Infectious Agent of Translucent Post-Larva Disease (TPD) in *Penaeus vannamei*. Pathogens. 2020; 9(9):1–17. https://doi.org/10.3390/pathogens9090741 PMID: 32927617
14. Wang R, Zhong Y, Gu X, Yuan J, Saeed AF, Wang S. The pathogenesis, detection, and prevention of *Vibrio parahaemolyticus*. Front Microbiol. 2015; 6:1–13. https://doi.org/10.3389/fmicb.2015.00144 PMID: 25798132
15. Castellanos A, Restrepo L, Bajaña L, Betancourt I, Bayot B, Reyes A. Genomic and Evolutionary Features of Nine AHPND Positive *Vibrio parahaemolyticus* Strains Isolated from South American Shrimp Farms. Microbiol Spectr. 2023;e04851–22. https://doi.org/10.1128/spectrum.04851-22 PMID: 37272817
16. Ma R, Wang Y, Huang L, Ying N, Ling H, Fang W. Complete genome sequence of highly pathogenic *Vibrio parahaemolyticus* isolated from mariculture *Penaeus vannamei* reveals virulence factor genes. Aquac Res. 2021; 52(4):1401–13. https://doi.org/10.1111/are.14995
17. Li L, Meng H, Gu D, Li Y, Jia M. Molecular mechanisms of *Vibrio parahaemolyticus* pathogenesis. Microbiol Res. 2019; 222:43–51. https://doi.org/10.1016/j.micres.2019.03.003 PMID: 30928029
18. Prithvisagar KS, Krishna Kumar B, Kodama T, Rai P, Iida T, Karunasagar I, et al. Whole genome analysis unveils genetic diversity and potential virulence determinants in *Vibrio parahaemolyticus* associated with disease outbreak among cultured *Litopenaeus vannamei* (Pacific white shrimp) in India. Virulence. 2021; 12(1):1936–49. https://doi.org/10.1080/21505594.2021.1947448 PMID: 34415829

19. Khan F, Tabassum N, Anand R, Kim YM. Motility of Vibrio spp.: regulation and controlling strategies. Appl Microbiol Biotechnol. 2020; 104(19):8187–208. https://doi.org/10.1007/s00253-020-10794-7 PMID: 32816086

20. Zhong X, Lu Z, Wang F, Yao N, Shi M, Yang M. Characterization of GefA, a GGEEF Domain-Containing Protein That Modulates *Vibrio parahaemolyticus* Motility, Biofilm Formation, and Virulence. Appl Environ Microbiol. 2022; 1. https://doi.org/10.1128/aem.02239-21 PMID: 35108083

21. Denzer L, Schroten H, Schwerk C. From Gene to Protein-How Bacterial Virulence Factors Manipulate Host Gene Expression During Infection. Int J Mol Sci. 2020; 21(10):1–37. https://doi.org/10.3390/ijms21103730 PMID: 32466312

22. Kangueane P, Nilofer C. Protein-Protein Docking: Methods and Tools. In: Protein-Protein and Domain-Domain Interactions. Singapore: Springer Singapore; 2018. p. 161–8. https://doi.org/10.1007/978-981-10-7347-2_14

23. Mirzadeh A, Kobakhidze G, Vuillemot R, Jonic S, Rouiller I. In silico prediction, characterization, docking studies and molecular dynamics simulation of human p97 in complex with p37 cofactor. BMC Mol Cell Biol. 2022; 23(1):1–12. https://doi.org/10.1186/s12860-022-00437-2 PMID: 36088301

24. Verma AK, Gupta S, Verma S, Mishra A, Nagpure NS, Singh SP, et al. Interaction between shrimp and white spot syndrome virus through PmRab7-VP28 complex: an insight using simulation and docking studies. J Mol Model. 2013; 19(3):1285–94. https://doi.org/10.1007/s00894-012-1672-0 PMID: 23179770

25. Yadav K, Verma AK, Pathak AK, Awasthi A. *In Silico* Prediction of Molecular Interaction Within PmCBP-VP24 Complex to Understand Initial Instigation of WSSV into Shrimps. Genet Aquat Org. 2021; 5(2):87–93. https://doi.org/10.4194/2459-1831-v5_2_05

26. Ji X, Zheng Y, Sheng J. Prediction of Interaction Sites of PcRab7-VP28. Comput Mol Biosci. 2021; 11 (02):50–6. https://doi.org/10.4236/cmb.2021.112003

27. Rosilan NF, Waiho K, Fazhan H, Sung YY, Mohd Nor SA, Nor Muhammad NA, et al. Protein-protein interaction network analysis on the whiteleg shrimp *Penaeus vannamei* and *Vibrio parahaemolyticus* host-pathogen relationship reveals possible proteins and pathways involved during infection. Aquac Reports. 2023; 30:101583. https://doi.org/10.1016/j.aqrep.2023.101583

28. Consortium TU. UniProt: the Universal Protein Knowledgebase in 2023. Nucleic Acids Res. 2023; 51:523–31. https://doi.org/10.1093/nar/gkac1052 PMID: 36408920

29. Varadi M, Anyango S, Deshpande M, Nair S, Natassia C, Yordanova G, et al. AlphaFold Protein Structure Database: massively expanding the structural coverage of protein-sequence space with high-accuracy models. Nucleic Acids Res. 2022; 50(D1):D439–44. https://doi.org/10.1093/nar/gkab1061 PMID: 34791371

30. Jumper J, Evans R, Pritzel A, Green T, Figurnov M, Ronneberger O, et al. Highly accurate protein structure prediction with AlphaFold. Nature. 2021; 596(7873):583–9. https://doi.org/10.1038/s41586-021-03819-2 PMID: 34265844

31. Laskowski RA, Jabłońska J, Pravda L, Vařeková RS, Thornton JM. PDBsum: Structural summaries of PDB entries. Protein Sci. 2018; 27(1):129–34. https://doi.org/10.1002/pro.3289 PMID: 28875543

32. Colovos C, Yeates TO. Verification of protein structures: patterns of nonbonded atomic interactions. Protein Sci. 1993; 2(9):1511–9. https://doi.org/10.1002/pro.5560020916 PMID: 8401235

33. Laskowski RA, MacArthur MW, Moss DS, Thornton JM. *PROCHECK*: a program to check the stereochemical quality of protein structures. J Appl Crystallogr. 1993; 26(2):283–91. https://doi.org/10.1107/S0021889892009944

34. Laskowski RA, Rullmann JAC, MacArthur MW, Kaptein R, Thornton JM. AQUA and PROCHECK-NMR: programs for checking the quality of protein structures solved by NMR. J Biomol NMR. 1996; 8 (4):477–86. https://doi.org/10.1007/BF00228148 PMID: 9008363

35. Williams CJ, Headd JJ, Moriarty NW, Prisant MG, Videau LL, Deis LN, et al. MolProbity: More and better reference data for improved all-atom structure validation. Protein Sci. 2018 Jan 1; 27(1):293–315. https://doi.org/10.1002/pro.3330 PMID: 29067766

36. Wallner B, Elofsson A. Can correct protein models be identified?. Protein Sci. 2003; 12(5):1073–86. https://doi.org/10.1110/ps.0236803 PMID: 12717029

37. Wiederstein M, Sippl MJ. ProSA-web: interactive web service for the recognition of errors in three-dimensional structures of proteins. Nucleic Acids Res. 2007; 35(suppl_2):W407–10. https://doi.org/10.1093/nar/gkm290 PMID: 17517781

38. Sippl MJ. Recognition of errors in three-dimensional structures of proteins. Proteins Struct Funct Bioinforma. 1993; 17(4):355–62. https://doi.org/10.1002/prot.340170404 PMID: 8108378

39. Zacharias M. Protein-protein docking with a reduced protein model accounting for side-chain flexibility. Protein Sci. 2003; 12(6):1271–82. https://doi.org/10.1110/ps.0239303 PMID: 12761398

40. Feng T, Chen F, Kang Y, Sun H, Liu H, Li D, et al. HawkRank: a new scoring function for protein-protein docking based on weighted energy terms. J Cheminform. 2017; 9(1):1–15. https://doi.org/10.1186/s13321-017-0254-7 PMID: 29282565

41. Hou T, Qiao X, Zhang W, Xu X. Empirical Aqueous Solvation Models Based on Accessible Surface Areas with Implicit Electrostatics. J Phys Chem B. 2002; 106(43):11295–304. https://doi.org/10.1021/jp025595u

42. Weng G, Wang E, Wang Z, Liu H, Zhu F, Li D, et al. HawkDock: a web server to predict and analyze the protein-protein complex based on computational docking and MM/GBSA. Nucleic Acids Res. 2019; 47 (W1):W322–30. https://doi.org/10.1093/nar/gkz397 PMID: 31106357

43. Schrödinger L. The PyMOL Molecular Graphics System. 2015.

44. Abraham M, Alekseenko A, Bergh C, Blau C, Briand E, Doijade M, et al. GROMACS 2023 Manual. 2023 Feb 6 [cited 2023 Apr 16]. https://zenodo.org/record/7588711.

45. Abraham MJ, Murtola T, Schulz R, Páll S, Smith JC, Hess B, et al. Gromacs: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. SoftwareX. 2015; 1–2:19–25. https://doi.org/10.1016/j.softx.2015.06.001

46. Jorgensen WL, Tirado-Rives J. The OPLS [optimized potentials for liquid simulations] potential functions for proteins, energy minimizations for crystals of cyclic peptides and crambin. J Am Chem Soc. 1988; 110(6):1657–66. https://doi.org/10.1021/ja00214a001 PMID: 27557051

47. Darden T, York D, Pedersen L. Particle mesh Ewald: An N·log(N) method for Ewald sums in large systems. J Chem Phys. 1993; 98(12):10089–92. https://doi.org/10.1063/1.464397

48. Humphrey W, Dalke A, Schulten K. VMD: visual molecular dynamics. J Mol Graph. 1996; 14:33–8. https://doi.org/10.1016/0263-7855(96)00018-5 PMID: 8744570

49. Turner PJ. XMGRACE, Version 5.1. 19. Vol. 2, Center for Coastal and Land-Margin Research, Oregon Graduate Institute of Science and Technology, Beaverton, OR. 2005. https://plasma-gate.weizmann.ac.il/Grace/.

50. Daura X, Gademann K, Jaun B, Seebach D, van Gunsteren WF, Mark AE. Peptide Folding: When Simulation Meets Experiment. Angew Chemie Int Ed. 1999; 38(1–2):236–40. https://doi.org/10.1002/(SICI)1521-3773(19990115)38:1/2%3C236::AID-ANIE236%3E3.0.CO

51. Kollman PA, Massova I, Reyes C, Kuhn B, Huo S, Chong L, et al. Calculating Structures and Free Energies of Complex Molecules: Combining Molecular Mechanics and Continuum Models. Acc Chem Res. 2000; 33(12):889–97. https://doi.org/10.1021/ar000033j PMID: 11123888

52. Wang E, Sun H, Wang J, Wang Z, Liu H, Zhang JZH, et al. End-Point Binding Free Energy Calculation with MM/PBSA and MM/GBSA: Strategies and Applications in Drug Design. Chem Rev. 2019; 119 (16):9478–508. https://doi.org/10.1021/acs.chemrev.9b00055 PMID: 31244000

53. Case DA, H.M. A, Belfon K, Ben-Shalom IY, J.T. B, S.R. B, et al. AMBER 2023. 2023. https://ambermd.org/index.php.

54. Gnanagobal H, Santander J. Host-Pathogen Interactions of Marine Gram-Positive Bacteria. Biology (Basel). 2022; 11(9). https://doi.org/10.3390/biology11091316 PMID: 36138795

55. Roy S, Bossier P, Norouzitallab P, Vanrompay D. Trained immunity and perspectives for shrimp aquaculture. Rev Aquac. 2020; 12(4):2351–70. https://doi.org/10.1111/raq.12438

56. Knovich MA, Storey JA, Coffman LG, Torti S V. Ferritin for the Clinician. Blood Rev. 2010; 23(3):95–104. https://doi.org/10.1016/j.blre.2008.08.001 PMID: 18835072

57. Moreira AC, Mesquita G, Gomes MS. Ferritin: An Inflammatory Player Keeping Iron at the Core of Pathogen-Host Interactions. Microorganisms. 2020; 8(4):1–20. https://doi.org/10.3390/microorganisms8040589 PMID: 32325688

58. Nairz M, Weiss G. Iron in infection and immunity. Mol Aspects Med. 2020; 75:100864. https://doi.org/10.1016/j.mam.2020.100864 PMID: 32461004

59. Gehrer CM, Mitterstiller AM, Grubwieser P, Meyron-Holtz EG, Weiss G, Nairz M. Advances in Ferritin Physiology and Possible Implications in Bacterial Infection. Int J Mol Sci. 2023; 24(5). https://doi.org/10.3390/ijms24054659 PMID: 36902088

60. Huang Y, Ren Q. A newly identified Hippo homologue from the oriental river prawn *Macrobrachium nipponense* is involved in the antimicrobial immune response. Vet Res. 2021;1–15. https://doi.org/10.1186/s13567-021-00945-7 PMID: 34078461

61. Erdmann M, Scholz A, Melzer IM, Schmetz C, Wiese M. Interacting Protein Kinases Involved in the Regulation of Flagellar Length. Mol Biol Cell. 2006 Apr; 17(4):2035–45. https://doi.org/10.1091/mbc.e05-10-0976 PMID: 16467378

62. Baker N, Catta-Preta CMC, Neish R, Sadlova J, Powell B, Alves-Ferreira EVC, et al. Systematic functional analysis of *Leishmania* protein kinases identifies regulators of differentiation or survival. Nat Commun. 2021; 12(1):1244. https://doi.org/10.1038/s41467-021-21360-8 PMID: 33623024

63. Yang D, Zhao L, Li Q, Huang L, Qin Y, Wang P, et al. *flgC* gene is involved in the virulence regulation of *Pseudomonas plecoglossicida* and affects the immune response of *Epinephelus coioides*. Fish Shellfish Immunol. 2023;132. https://doi.org/10.1016/j.fsi.2022.108512 PMID: 36587883

64. Rossmann FM, Hug I, Sangermani M, Jenal U, Beeby M. In situ structure of the *Caulobacter crescentus* flagellar motor and visualization of binding of a CheY-homolog. Mol Microbiol. 2020; 114(3):443–53. https://doi.org/10.1111/mmi.14525 PMID: 32449846

65. Zhu D, Yang L, Huang J, Zhou F, Yang Q, Jiang S, et al. The comprehensive expression analysis of the G protein-coupled receptor from *Penaeus monodon* indicating it participates in innate immunity and anti-ammonia nitrogen stress. Fish Shellfish Immunol. 2018; 75:17–26. https://doi.org/10.1016/j.fsi.2018.01.019 PMID: 29410275

66. Wang T, Liang J, Xiang X, Yuan J, Chen X, Xiang X, et al. Functional identification and expressional responses of large yellow croaker (Larimichthys crocea) interleukin-8 and its receptor. Fish Shellfish Immunol. 2019; 87:470–7. https://doi.org/10.1016/j.fsi.2019.01.035 PMID: 30708055

67. Guo F, Li SC, Wang L, Zhu D. Protein-protein binding site identification by enumerating the configurations. BMC Bioinformatics. 2012; 13(1). https://doi.org/10.1186/1471-2105-13-158 PMID: 22768846

68. Li J, Fu A, Zhang L. An Overview of Scoring Functions Used for Protein–Ligand Interactions in Molecular Docking. Interdiscip Sci Comput Life Sci. 2019; 11(2):320–8. https://doi.org/10.1007/s12539-019-00327-w PMID: 30877639

69. Rossi A, Marti-Renom MA, Sali A. Localization of binding sites in protein structures by optimization of a composite scoring function. Protein Sci. 2006; 15(10):2366–80. https://doi.org/10.1110/ps.062247506 PMID: 16963645

70. Chen D, Oezguen N, Urvil P, Ferguson C, Dann SM, Savidge TC. Regulation of protein-ligand binding affinity by hydrogen bond pairing. Sci Adv. 2016; 2(3). https://doi.org/10.1126/sciadv.1501240 PMID: 27051863

71. Tam JZ, Palumbo T, Miwa JM, Chen BY. Analysis of Protein-Protein Interactions for Intermolecular Bond Prediction. Molecules. 2022;1–14. https://doi.org/10.3390/molecules27196178 PMID: 36234723

72. Hamza A, Samad A, Imam MA, Faizan MI, Ahmed A, Almajhdi FN, et al. Structural Characterization of Ectodomain G Protein of Respiratory Syncytial Virus and Its Interaction with Heparan Sulfate: Multi-Spectroscopic and In Silico Studies Elucidating Host-Pathogen Interactions. Molecules. 2021; 26(23). https://doi.org/10.3390/molecules26237398 PMID: 34885979

73. Mitra D, Pal AK, Das Mohapatra PK. Intra-protein interactions of SARS-CoV-2 and SARS: a bioinformatic analysis for plausible explanation regarding stability, divergency, and severity. Syst Microbiol Biomanufacturing. 2022; 2(4):653–64. https://doi.org/10.1007/s43393-022-00091-x

74. Hasan MM, Hoque MN, Hasan MR, Asaduzzaman M, Juliana FM. Rab7 Investigation Insights into the Existence of White Spot Syndrome Virus in Crustaceans: An *In Silico* Approach. Adv Virol. 2022;2022. https://doi.org/10.1155/2022/3887441 PMID: 36313590

75. Hubbard RE, Kamran Haider M. Hydrogen Bonds in Proteins: Role and Strength. eLS. 2010. https://doi.org/10.1002/9780470015902.a0003011.pub2

76. Spassov DS, Atanasova M, Doytchinova I. A role of salt bridges in mediating drug potency: A lesson from the N-myristoyltransferase inhibitors. Front Mol Biosci. 2023; 9:1–15. https://doi.org/10.3389/fmolb.2022.1066029 PMID: 36703920

77. Al-Hamdani YS, Tkatchenko A. Understanding non-covalent interactions in larger molecular complexes from first principles. J Chem Phys. 2019; 150(1):1–22. https://doi.org/10.1063/1.5075487 PMID: 30621423

78. Dara M, Giulianini PG, Manfrin C, Parisi MG, Parrinello D, La Corte C, et al. F-type lectin from serum of the Antarctic teleost fish *Trematomus bernacchii* (Boulenger, 1902): Purification, structural characterization, and bacterial agglutinating activity. Comp Biochem Physiol Part B Biochem Mol Biol. 2021; 256:110633. https://doi.org/10.1016/j.cbpb.2021.110633 PMID: 34126205

79. Chandler JC, Gandhi NS, Mancera RL, Smith G, Elizur A, Ventura T. Understanding Insulin Endocrinology in Decapod Crustacea: Molecular Modelling Characterization of an Insulin-Binding Protein and Insulin-Like Peptides in the Eastern Spiny Lobster, *Sagmariasus verreauxi*. Int J Mol Sci. 2017; 18 (9):1–19. https://doi.org/10.3390/ijms18091832 PMID: 28832524

80. Ghiandoni GM, Caldeweyher E. Fast calculation of hydrogen-bond strengths and free energy of hydration of small molecules. Sci Rep. 2023; 13(1):4143. https://doi.org/10.1038/s41598-023-30089-x PMID: 36914670

81. Patoliya J, Thaker K, Rabadiya K, Patel D, Jain NK, Joshi R. Uncovering the Interaction Interface Between Harpin (Hpa1) and Rice Aquaporin (OsPIP1;3) Through Protein–Protein Docking: An In Silico Approach. Mol Biotechnol. 2023;(0123456789). https://doi.org/10.1007/s12033-023-00690-6 PMID: 36807270

82. Sargsyan K, Grauffel C, Lim C. How Molecular Size Impacts RMSD Applications in Molecular Dynamics Simulations. J Chem Theory Comput. 2017 Apr 11; 13(4):1518–24. https://doi.org/10.1021/acs.jctc.7b00028 PMID: 28267328

83. Duan L, Guo X, Cong Y, Feng G, Li Y, Zhang JZH. Accelerated Molecular Dynamics Simulation for Helical Proteins Folding in Explicit Water. Front Chem. 2019; 7(August):1–18. https://doi.org/10.3389/fchem.2019.00540 PMID: 31448259

84. Alazmi M, Alshammari N, Alanazi NA, Sulieman AME. In silico characterization, docking, and simulations to understand host–pathogen interactions in an effort to enhance crop production in date palms. J Mol Model. 2021; 27(11). https://doi.org/10.1007/s00894-021-04957-0 PMID: 34731299

85. Sudharsana S, Rajashekar Reddy CB, Dinesh S, Rajasekhara Reddy S, Mohanapriya A, Itami T, et al. Molecular docking and simulation studies of 3-(1-chloropiperidin-4-yl)-6-fluoro benzisoxazole 2 against VP26 and VP28 proteins of white spot syndrome virus. J Fish Dis. 2016; 39(10):1231–8. https://doi.org/10.1111/jfd.12454 PMID: 26850228

86. Oselusi SO, Fadaka AO, Wyckoff GJ, Egieyeh SA. Computational Target-Based Screening of Anti-MRSA Natural Products Reveals Potential Multitarget Mechanisms of Action through Peptidoglycan Synthesis Proteins. ACS Omega. 2022; 7(42):37896–906. https://doi.org/10.1021/acsomega.2c05061 PMID: 36312373

87. Kitao A. Principal Component Analysis and Related Methods for Investigating the Dynamics of Biological Macromolecules. Vol. 5, J. 2022. p. 298–317. https://doi.org/10.3390/j5020021

88. Lee YT, Glazer EC, Wilson RF, Stout CD, Goodin DB. Three Clusters of Conformational States in P450cam Reveal a Multistep Pathway for Closing of the Substrate Access Channel,. Biochemistry. 2011; 50(5):693–703. https://doi.org/10.1021/bi101726d PMID: 21171581

89. David CC, Jacobs DJ. Principal Component Analysis: A Method for Determining the Essential Dynamics of Proteins. In: Livesay DR, editor. Protein Dynamics: Methods and Protocols. Totowa, NJ: Humana Press; 2014. p. 193–226.

90. Orellana L, Yoluk O, Carrillo O, Orozco M, Lindahl E. Prediction and validation of protein intermediate states from structurally rich ensembles and coarse-grained simulations. Nat Commun. 2016; 7(1):12575. https://doi.org/10.1038/ncomms12575 PMID: 27578633

91. Damjanovic J, Murphy JM, Lin YS. CATBOSS: Cluster Analysis of Trajectories Based on Segment Splitting. J Chem Inf Model. 2021 Oct 25; 61(10):5066–81. https://doi.org/10.1021/acs.jcim.1c00598 PMID: 34608796