

RESEARCH ARTICLE

# Expression, purification and characterization of the dimeric protruding domain of *Macrobrachium rosenbergii* nodavirus capsid protein expressed in *Escherichia coli*

Li Chuin Chong<sup>1</sup>, Hagilaa Ganesan<sup>2</sup>, Chean Yeah Yong<sup>2</sup>, Wen Siang Tan<sup>2,3</sup>, Kok Lian Ho<sup>4\*</sup>

**1** Department of Biomedical Sciences, Faculty of Medicine and Health Sciences, Universiti Putra Malaysia, UPM Serdang, Selangor, Malaysia, **2** Department of Microbiology, Faculty of Biotechnology and Biomolecular Sciences, Universiti Putra Malaysia, UPM Serdang, Selangor, Malaysia, **3** Institute of Bioscience, Universiti Putra Malaysia, UPM Serdang, Selangor, Malaysia, **4** Department of Pathology, Faculty of Medicine and Health Sciences, Universiti Putra Malaysia, UPM Serdang, Selangor, Malaysia

\* [klho@upm.edu.my](mailto:klho@upm.edu.my)



**OPEN ACCESS**

**Citation:** Chong LC, Ganesan H, Yong CY, Tan WS, Ho KL (2019) Expression, purification and characterization of the dimeric protruding domain of *Macrobrachium rosenbergii* nodavirus capsid protein expressed in *Escherichia coli*. PLoS ONE 14 (2): e0211740. <https://doi.org/10.1371/journal.pone.0211740>

**Editor:** Paulo Lee Ho, Instituto Butantan, BRAZIL

**Received:** October 10, 2018

**Accepted:** January 18, 2019

**Published:** February 1, 2019

**Copyright:** © 2019 Chong et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Data Availability Statement:** All relevant data are within the manuscript and its Supporting Information file.

**Funding:** This work was supported by the Ministry of Education of Malaysia through the Fundamental Research Grant Scheme (Grant number: 01-01-18-1988FR) and Universiti Putra Malaysia (UPM) Putra Grant (Grant number: GP-IPS/2018/9602500). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

## Abstract

*Macrobrachium rosenbergii* nodavirus (*MtNV*) is the causative agent of white tail disease (WTD) which seriously impedes the production of the giant freshwater prawn and has a major economic impact. *MtNV* contains two segmented RNA molecules, which encode the RNA dependent RNA polymerase (RdRp) and the capsid protein (*MtNV*-CP) containing 371 amino acid residues. *MtNV*-CP comprises of the Shell (S) and the Protruding (P) domains, ranging from amino acid residues 1–252 and 253–371, respectively. The P-domain assembles into dimeric protruding spikes, and it is believed to be involved in host cell attachment and internalization. In this study, the recombinant P-domain of *MtNV*-CP was successfully cloned and expressed in *Escherichia coli*, purified with an immobilized metal affinity chromatography (IMAC) and size exclusion chromatography (SEC) up to ~90% purity. Characterization of the purified recombinant P-domain with SEC revealed that it formed dimers, and dynamic light scattering (DLS) analysis demonstrated that the hydrodynamic diameter of the dimers was ~6 nm. Circular dichroism (CD) analysis showed that the P-domain contained 67.9% of beta-sheets, but without alpha-helical structures. This is in good agreement with the cryo-electron microscopic analysis of *MtNV* which demonstrated that the P-domain contains only beta-stranded structures. Our findings of this study provide essential information for the production of the P-domain of *MtNV*-CP that will aid future studies particularly studies that will shed light on anti-viral drug discovery and provide an understanding of virus-host interactions and the viral pathogenicity.

## Introduction

The giant freshwater prawn or scientifically known as *Macrobrachium rosenbergii* (de Man) is one of the most commercially important farmed aquatic invertebrates in subtropical countries

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

including Malaysia, Indonesia and Thailand [1, 2]. Aquaculture farming has significantly boosted the global aquaculture industry which recorded the highest output value of USD 2.13 billion in 2008. However, the prawn production has declined tremendously since 2012 due to the white tail disease (WTD) that mainly affects prawn hatcheries, and this disease causes great economic loss [1]. The causative agent of WTD is *Macrobrachium rosenbergii* nodavirus (*MrNV*) which can lead to 100% mortality in larvae, post-larvae, and early juvenile stages of the fresh water prawn [3, 4]. As the name of the disease implies, WTD originates from its milky whitish appearance at the tail of an infected prawn. Other symptoms such as lethargy, anorexia, abdominal muscle opaqueness, as well as a weakened ability to feed and swim [5–7].

*MrNV* is a non-enveloped virus that belongs to the family *Nodaviridae*. This virus contains two bipartite single-stranded RNA (ssRNA) molecules, known as RNA 1 (3.2 kb) and RNA 2 (1.2 kb), which encode for the RNA dependent RNA polymerase (RdRp) and the viral capsid protein (*MrNV*-CP), respectively [8, 9]. Currently, there are two established genera in the family of *Nodaviridae*, namely *Alpha-nodavirus* and *Beta-nodavirus*. *Alpha-nodavirus*, one of the causal mortality agents among insect populations, includes black beetle virus (BBV), Boolarra virus (BoV), Flock House virus (FHV), Nodamura virus (NoV), and Pariacoto virus (PaV), while *Beta-nodavirus*, has a major impact on marine fish species, includes barfin flounder nervous necrosis virus (BFNNV), redspotted grouper nervous necrosis virus (RGNNV), striped jack nervous necrosis virus (SJNNV) and tiger puffer nervous necrosis virus (TPNNV) [10]. Interestingly, *MrNV*, *Penaeus vannamei* nodavirus (*PvNV*) and covert mortality nodavirus (CMNV) are not clustered in neither *Alpha*- nor *Beta-nodavirus* due to some of their unique characteristics which distinguish them from these two genera [11]. The distinct criterion that differentiates crustacean-infected nodaviruses from the *Alpha*- and *Beta-nodavirus* is the amino acid sequence identity of their capsid proteins. Amino acid sequence alignment revealed that the capsid protein of *MrNV* shares more than 80% sequence similarity with that of *PvNV*. However, neither *MrNV* nor *PvNV* capsid protein shows similarity higher than 20% with that of *Alpha*- or *Beta-nodavirus*. Therefore, a new genus known as *Gamma-nodavirus* was proposed [11, 12].

*MrNV*-CP is a polypeptide comprises of 371 amino acid residues which can be divided into two major domains: the Shell (S) and the Protruding (P) domains, ranging from amino acid residues 1–252 and 253–371, respectively. The full-length and several truncated derivatives of *MrNV*-CP expressed in *E. coli* and insect cells self-assembled into virus like particles (VLPs) [13–15]. The purified recombinant *MrNV*-CP that assembled into VLPs was used for structural and functional studies. The arginine (Arg) and lysine (Lys)-rich N-terminal region of *MrNV*-CP is believed to interact with the negatively charged RNA molecules during the viral morphogenesis [16]; the middle portion of *MrNV*-CP composed of hydrophobic and polar amino acids, whereas the C-terminal region containing the P-domain, has been shown to be involved in host cell attachment and internalization [17]. To understand these mechanisms, a high-resolution atomic structure of the P-domain of *MrNV*-CP is required. Our previous cryo-electron microscopic analysis revealed that the P-domain is exposed on the surface of VLPs as dimerized protruding blade-like spikes which is distinct from the trimeric spikes found in *Alpha*- and *Beta-nodaviruses* [12]. Recently, an atomic-resolution model of the *MrNV*-CP was built based on cryo-EM data with an improved resolution at 3.3 Å [18]. However, the local resolution of the dimeric spikes formed by the P-domain of *MrNV*-CP remained poorer than 4 Å.

In this article, we describe the production of the recombinant P-domain of *MrNV*-CP in *E. coli*. The recombinant protein was then purified to ~90% pure with a two-step chromatography and characterized with size exclusion chromatography (SEC), dynamic light scattering (DLS) and circular dichroism (CD). SEC, DLS and CD analyses demonstrated that the purified *MrNV*-CP P-domain formed dimer with a hydrodynamic diameter ~6 nm and contained

~68%  $\beta$ -stranded structures. The capability of the P-domain to form a dimer indicates that it mimics the dimeric state of its endogenous form exposed on the viral capsid.

## Materials and methods

### Construction of recombinant plasmid encoding the P-domain of *MrNV*-CP

The recombinant plasmid, pTrcHis2-TARNA2 [13] encoding the full-length *MrNV*-CP was extracted using the alkaline lysis method. The nucleotide sequence encoding the P-domain of *MrNV*-CP (amino acids 253–371) was amplified by PCR using the forward primer (5' AAT CCT ACA CCA GCC ATG GGC TCA CAG TTA ACT 3'; *Nco*I cutting site is underlined) and the reverse primer (5' ACG TAA GCT TCG AAT TCG CCC TTA TTA TTG CCG ACG ATA G 3'; *Eco*RI cutting site is underlined). The PCR product was digested with *Nco*I and *Eco*RI (Thermo Scientific, Waltham, USA) and ligated to pTriHis2-TOPO which had been digested with the same restriction enzymes. The ligated plasmid was introduced into *E. coli* TOP10 competent cells for protein expression. The nucleotide sequence of the recombinant plasmid was confirmed by DNA sequencing.

### Expression and purification of the recombinant P-domain of *MrNV*-CP

Expression of the *MrNV*-CP P-domain in *E. coli* was performed according to Goh et al. [13] with some modifications. *E. coli* TOP 10 cells harboring the recombinant plasmid was inoculated in Luria-Bertani (LB) broth (500 ml) supplemented with (50  $\mu$ g/ml) ampicillin and incubated overnight at 37°C with shaking at 200 rpm. The expression of recombinant protein was induced by adding isopropyl  $\beta$ -D-1-thiogalactopyranoside (IPTG, 1 mM) at 30°C for 5 hours. Cells were harvested by centrifugation at 16,200  $\times$ g, at 4°C. The pelleted cells were resuspended in lysis buffer (25 mM HEPES, 500 mM NaCl, pH7.4) and lysed by adding  $MgCl_2$  (4 mM), Triton-100 [0.1% (v/v)], freshly prepared lysozyme (0.2 mg/ml), DNase 1 (0.02 mg/ml) and phenylmethylsulfonyl fluoride (PMSF, 2 mM) for 2 hours at room temperature. The cell suspension was then sonicated at 20 kHz for 25 seconds for 10 cycles with 15 seconds intervals between pulses using an ultrasonicator (Qsonica, USA). The crude lysate was recovered by centrifugation at 4°C at 16,200  $\times$ g for 20 minutes.

The poly-prep chromatography column (Bio-Rad, USA) was packed by adding the proficiency IMAC  $Ni^{2+}$ -charged resin (2 ml) and equilibrated with Tris-HCl buffer (20 mM Tris-HCl, 100 mM NaCl, pH 7.6). The crude lysate was loaded onto the column. Weak binders were removed by Tris-HCl buffer containing imidazole (10 mM) and bound proteins were eluted using Tris-HCl buffer containing Imidazole (100 mM). Eluted fractions were collected for protein analysis using SDS-PAGE [12% (w/v)] and western blot (see below). The positive fractions containing the P-domain of *MrNV*-CP were pooled, concentrated to ~1.5 ml using a centrifugal protein concentrator with molecular mass cut-off 3 kDa (Merck Millipore, USA) prior to gel filtration chromatography. The concentrated sample was loaded onto a Tris-HCl buffer-equilibrated HiPrep 16/60 Sephacryl S-200 HR column (GE Healthcare, Sweden) mounted on an AKTA Purifier (GE Healthcare, Sweden). Eluents of the column were collected (2 ml per fraction) using a Frac-950 fraction collector at a constant flow rate of 0.5 ml/min. Eluents were analyzed using SDS-PAGE and western blotting. The positive fractions containing the P-domain of *MrNV*-CP were pooled and concentrated.

### Western blot analysis

The separated proteins on an SDS-polyacrylamide gel were transferred onto a nitrocellulose membrane using a semi dry blotter (Trans-Blot SD, BioRad, USA). The nitrocellulose

membrane was then blocked with skimmed milk [10% (w/v)] for 1 hour at room temperature. After three times of washing with TBS-T [50 mM Tris-HCl, pH 7.6; 150 mM NaCl; 0.1% (v/v) Tween 20], the nitrocellulose membrane was incubated with the anti-His monoclonal antibody (Merck, Germany) for 1 hour at room temperature. Afterwards, the nitrocellulose membrane was washed with TBS-T and incubated with the alkaline phosphatase conjugated goat anti-mouse antibody (Merck, Germany; 1:5000 dilutions) for 1 hour at room temperature. The immunoblotted bands were developed by adding 5-bromo-4-chloro-3-indolyl phosphate/nitro blue tetrazolium (BCIP/NBT) in alkaline phosphatase buffer [100 mM Tris-HCl; pH 9.5, 100 mM NaCl, 5 mM MgCl<sub>2</sub>]. Color development was terminated by incubating the nitrocellulose membrane in distilled water and the membrane was allowed to dry at room temperature.

### Estimation of protein molecular mass with size exclusion chromatography

To determine the molecular weight and the multimeric state of the purified P-domain, the gel filtration chromatograms of the P-domain and the protein standard markers [blue dextran (2000 kDa), Ferritin (F; 440 kDa), Aldolase (Ald; 158 kDa), Conalbumin (C; 75 kDa), Ovalbumin (O; 44 kDa), Carbonic Anhydrase (CA; 29 kDa), Ribonuclease A (R; 13.7 kDa), Aprotinin (Apr; 6.5 kDa); GE Healthcare, USA] were overlaid. The molecular mass of the P-domain of MrNV-CP was estimated by comparing the elution profile of the P-domain with the elution profile of the standard markers.

### Quantification of the P-domain of MrNV-CP

The quantity and quality of the P-domain of MrNV at each step of the purification were determined according to Yoon et al. [19]. Briefly, after each step of the protein purification, the P-domain of MrNV-CP was quantified with the Bradford assay [20]. The amount of the P-domain of MrNV-CP from SDS-PAGE gels was measured using the Image J software [21] by calculating protein band intensities. The purity, purification factor and protein recovery were calculated with Eqs 1, 2 and 3, respectively.

$$\text{Purity (\%)} = \frac{\text{Area of graph of the P-domain}}{\text{Area of graph of total protein}} \times 100\% \quad (1)$$

$$\text{Purification factor} = \frac{\text{Purity of the P-domain}}{\text{Purity of the P-domain in crude lysate}} \quad (2)$$

$$\text{Recovery (\%)} = \frac{\text{Amount of the P-domain in first step purification (SEC)}}{\text{Amount of the P-domain in second step purification (IMAC)}} \times 100\% \quad (3)$$

### Dynamic light scattering (DLS)

The purified protein sample (0.3 mg/ml) was filtered using a 0.2 μm-syringe filter membrane and then loaded into a folded capillary Zeta cell and placed into a Zetasizer Nano ZS (Malvern Instruments) equipped with He-Ne laser (633 nm) to measure the diameter of the P-domain of MrNV-CP. All Zetasizer data were then analyzed using the Zetasizer Software 7.12.

### Estimation of secondary structure

The percentage of the secondary structure of the P-domain of MrNV-CP was estimated using circular dichroism (CD). The buffer of the purified protein sample was exchanged with HEPES buffer (4 mM HEPES, pH 7.6). The protein sample (5 μM) was loaded into a cuvette



**Fig 1. Recombinant expression vector, nucleotide and amino acid sequences of the protruding (P) domain of *Macrobrachium rosenbergii* nodavirus capsid protein (*MrNV-CP*).** (a) Recombinant expression vector containing the nucleotide sequence encoding for the P-domain of *MrNV-CP*. (b) Nucleotide coding sequence (Nucl seq) for the P-domain of *MrNV-CP* and the deduced amino acid sequence (AA seq) of the P-domain of *MrNV-CP*. Numbers above the amino acid sequence refer to amino acid positions of the P-domain *MrNV-CP*. Amino acid sequence of the P domain is colored in red.

<https://doi.org/10.1371/journal.pone.0211740.g001>

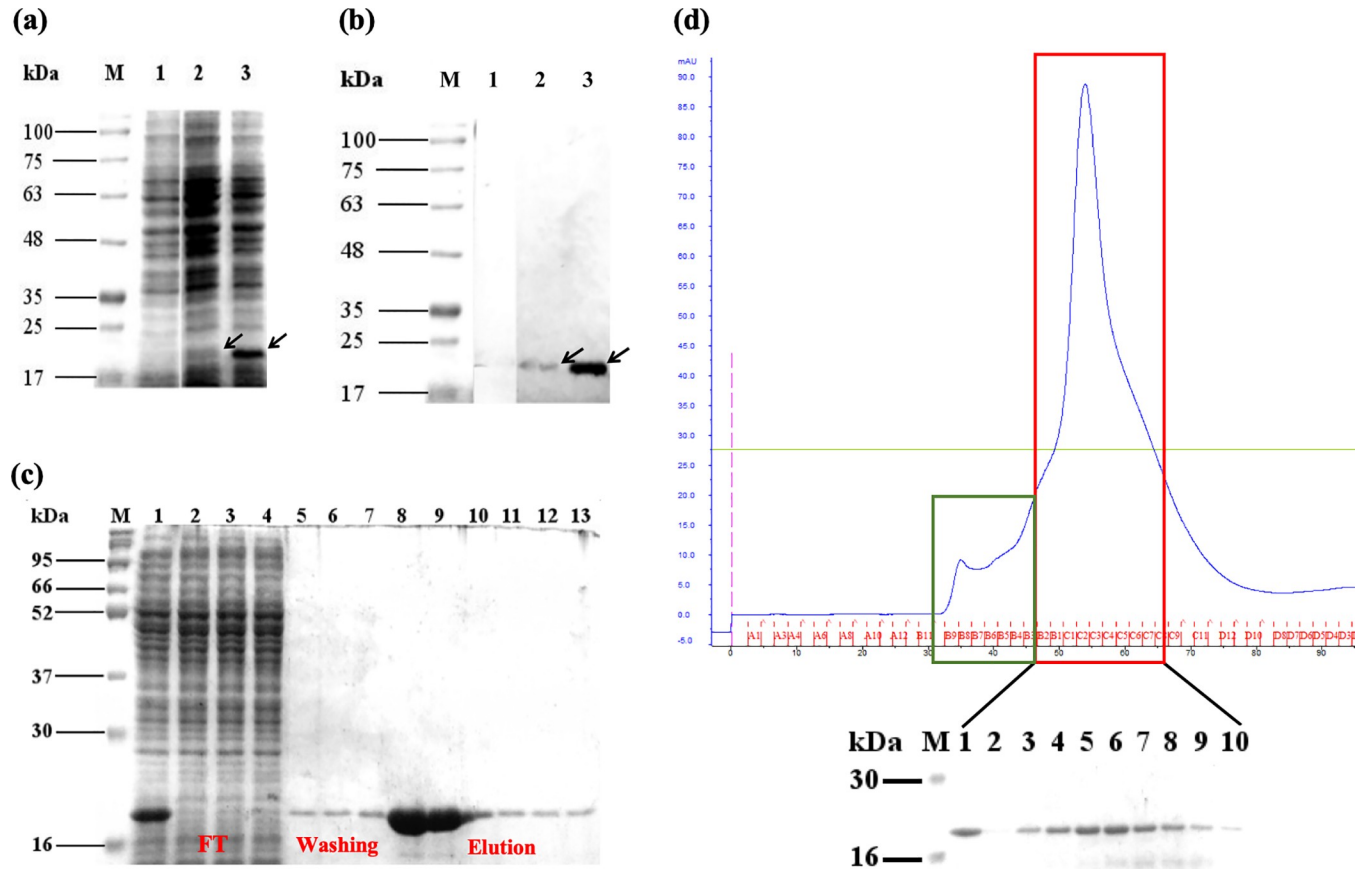
with 0.1 cm path length and placed in the measurement chamber of a JASCO J-815 CD spectropolarimeter (Jasco International, Japan) after being filtered through a 0.2 μm-syringe filter membrane. The sample was scanned from wavelengths 190 nm to 260 nm, with a speed of 10 nm/min for three times. The data were collected and analysed using the Secondary Structure Estimation Software which was installed in the Spectra Manager version 2. The calculations of the secondary structure elements were carried out based on Yang et al [22] and Reed and Reed [23].

## Results and discussion

### Cloning, expression and purification of the P-domain of *MrNV-CP*

The coding sequence of the P-domain of *MrNV-CP* was ligated to the expression vector pTrcHis2-TOPO at *NcoI* and *EcoRI* endonuclease restriction sites (Fig 1A). The recombinant plasmid was introduced into *E. coli* Top10 cells for protein expression. The recombinant protein comprises 150 amino acid residues including amino acids 253–371 of the P-domain of *MrNV-CP*, a *myc* epitope and a 6xHis-tag at its C-terminus (Fig 1B). Protein expression was optimized with several parameters including IPTG induction time and induction temperature. A time course study was performed, and an overnight IPTG induction of ~16 hours resulted in the highest protein expression level. However, by considering factors such as the protein stability and time efficiency, induction time of 5 hours was chosen as the most appropriate induction period, which is in good agreement with that reported by Goh et al. [13] for the expression of the full-length *MrNV-CP*. To obtain an optimum induction temperature, three different induction temperatures (25°C, 30°C and 37°C) were compared. The expression of *MrNV-CP* P-domain was at its highest expression level at 37°C, followed by 30°C, and the lowest at 25°C. In the expression of full-length *MrNV-CP*, Goh et al. [13] induced the protein expression at 25°C in order to improve the protein solubility. This temperature mimics the climate of the natural environment where *MrNV* propagates in the giant freshwater prawn. As a result, we chose 30°C as the most appropriate induction temperature in order to maintain the protein solubility, and at the same time, not compromising the protein expression level.





**Fig 2. SDS-PAGE and western blot analyses of the recombinant protruding (P) domain of *Macrobrachium rosenbergii* nodavirus capsid protein (*MrNV-CP*) expressed in *Escherichia coli*.** (a) SDS-PAGE, and (b) western blot analysis of the P-domain of *MrNV-CP* after isopropyl β-D-1-thiogalactopyranoside (IPTG) induction at 30°C. Lane M: protein markers in kDa; lane 1: cell lysate before IPTG induction; lane 2: cell lysate without IPTG induction; lane 3: cell lysate after being induced with IPTG for 2 hours. (b) the blot was probed with anti-His monoclonal antibody. Arrowheads indicate additional protein band ~18 kDa in both IPTG induced and uninduced cells. (c) immobilized metal ion affinity chromatography (IMAC) purification. Lane M: protein markers in kDa; lane 1: crude lysate; lanes 2 and 3: flow through; lanes 4–7: fractionated protein samples washed out with washing buffer; lanes 8–13: fractionated protein samples eluted out with elution buffer. (d) Gel filtration chromatogram of the P-domain of *MrNV-CP* fractionated by HiPrep 16/60 Sephacryl S-200 HR on an FPLC system. The first peak (green box) comprises of elution volume between 33 ml and 45 ml while the second peak (red box) comprises of elution volume between 45 ml and 63 ml. SDS-PAGE analysis of the eluted fractions of the second peak is shown under the chromatogram. Lane M: protein markers in kDa; lane 1: injected sample; lanes 2–10: eluted fractions of the second peak.

<https://doi.org/10.1371/journal.pone.0211740.g002>

The recombinant P-domain of *MrNV-CP* inclusive of a *myc*-tag and a cleavable 6xHis-tag was expressed as a soluble protein ~18 kDa at 30°C for 2 hours (Fig 2A). Interestingly, the bacterial culture without IPTG induction also produced the recombinant P-domain of *MrNV-CP* as detected in the western blot analysis despite the quantity of this leaky expression was relatively lower than that of the IPTG induced culture. To facilitate protein purification with Ni<sup>2+</sup> ion chelated nitriloacetic acid (Ni-NTA) immobilized metal ion affinity chromatography (IMAC), a 6xHis-tag was fused to the C-terminal region of the recombinant P-domain of *MrNV-CP*. As shown in Fig 2C, weakly bound proteins such as host proteins were mostly washed out using a low concentration of imidazole (10 mM) while the 6xHis-tag *MrNV-CP* P-domain was eluted later with 100 mM imidazole, producing a partially purified protein with protein purity of ~77% (Table 1).

To achieve a better protein purity for downstream analyses such as protein characterizations and structural analysis, another step of protein purification is normally required. In this study, a size exclusion chromatography was performed using HiPrep 16/60 Sephacryl S-200

**Table 1. Purification table of the protruding (P) domain of *Macrobrachium rosenbergii* nodavirus capsid protein (*MrNV*-CP) expressed in *Escherichia coli*.**

	Amount of <i>MrNV</i> -CP P-domain (mg) <sup>a</sup>	Purity of <i>MrNV</i> -CP P-domain (%) <sup>b</sup>	Purification factor <sup>c</sup>	Recovery (%) <sup>d</sup>
Crude lysate	35.6	14.8	-	-
IMAC <sup>e</sup>	12.4	77.0	5.2	34.8
SEC <sup>f</sup>	5.6	88.8	6.0	15.7

<sup>a</sup> The primary measurement of protein concentration using the Bradford assay by multiplying with the total volume.

<sup>b</sup> The percentage of the amount of *MrNV*-CP P-domain against the amount of total protein.

<sup>c</sup> The ratio of purity of *MrNV*-CP P-domain in eluted fraction and purity of protein in the crude lysate.

<sup>d</sup> The amount of *MrNV*-CP P-domain recovered in relation to the initial amount of protein.

<sup>e</sup> IMAC; Immobilized metal ion affinity chromatography

<sup>f</sup> SEC; Size exclusion chromatography

<https://doi.org/10.1371/journal.pone.0211740.t001>

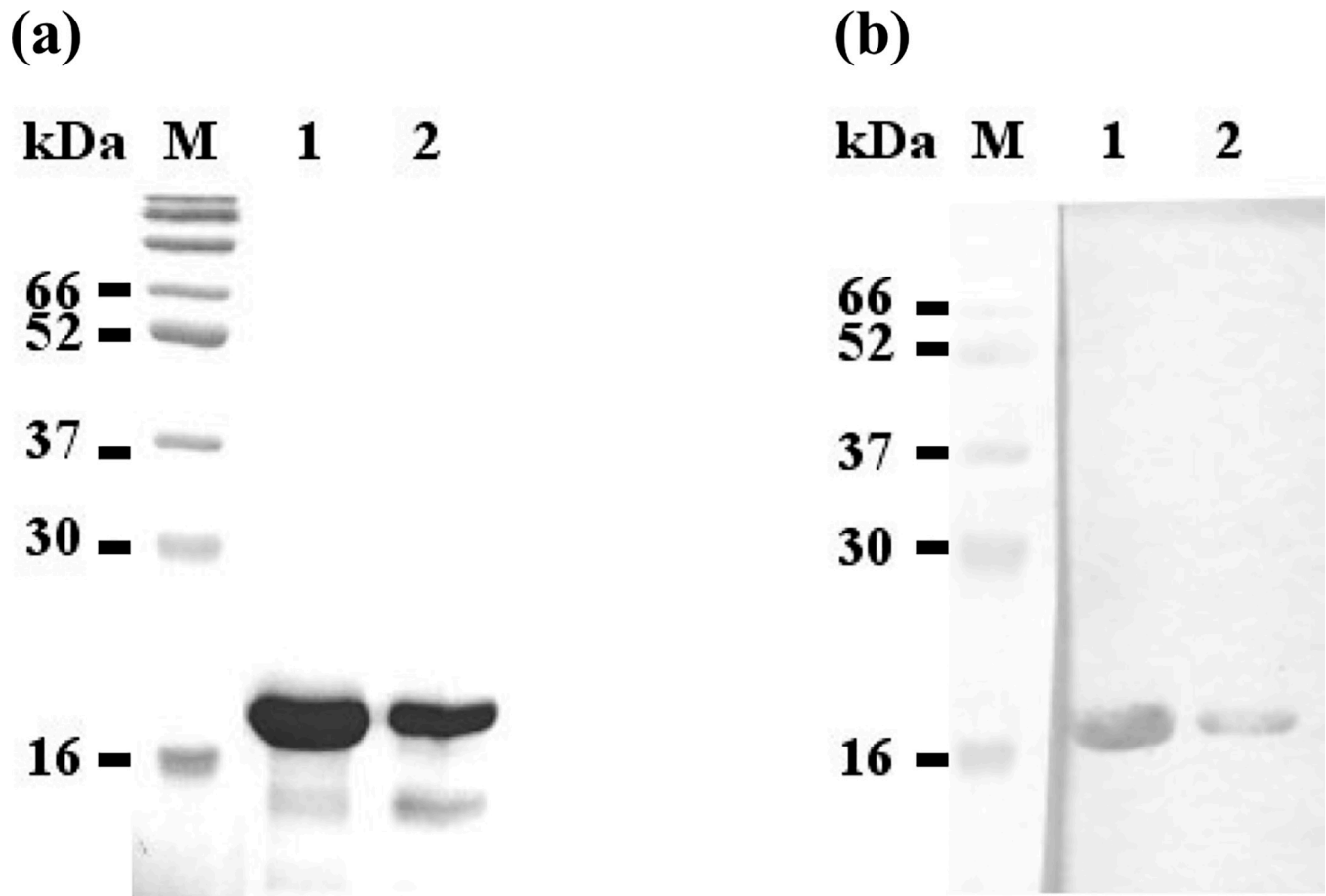
HR on an FPLC system. The fractionation range of this column is between 5 kDa to 250 kDa. Therefore, this column is suitable to fractionate the P-domain of *MrNV*-CP with molecular weight ~34 kDa, assuming the *MrNV*-CP P-domain forms dimers in solution. The size exclusion chromatography fractionated the protein sample containing the P-domain of *MrNV*-CP into two peaks between 33 ml and 65 ml of elution volume (as indicated by green and red colored boxes in Fig 2D). The first peak (green box) corresponds to the void volume ( $V_0$ ) which did not give a protein band when analyzed with SDS-PAGE whereas the second peak (red box) of the profile gave a protein band with a molecular weight ~18 kDa, which is in good agreement with the molecular mass of the monomer of *MrNV*-CP P-domain (Fig 2D). Following gel filtration, the purity of the recombinant protein increased to ~89%. Nevertheless, approximately 45% of the protein was lost in the purification process. The purification factor improved from 5.2 in IMAC to 6.0 after the second step of purification. The percentage of recovery, however, reduced from ~35% to ~16% after the gel filtration. During the IMAC purification, host proteins which bound to Ni-NTA column were removed when the protein mixtures were fractionated by a gel filtration column. In addition, certain amount of proteins might be degraded during the process of gel filtration. As a result, it is not surprised to recover only ~16% of the recombinant protein. This is also in accord with Yoon et al. [19] who reported that the recovery of the highly expressed hepatitis B virus capsid was ~20% after the protein was purified with gel filtration chromatography.

### Western blot analysis

After gel filtration, concentrated protein sample was analyzed by western blotting. Fig 3 shows that the P-domain of *MrNV*-CP has a molecular mass of ~18 kDa as detected by the anti-His monoclonal antibody. Lower molecular mass protein bands detected by Coomassie brilliant blue (CBB) staining on the polyacrylamide gel (Fig 3A) were not detected by the anti-His antibody, indicating that these smaller protein bands could be the *MrNV*-CP P-domain fragments without the C-terminus harboring the 6xHis-tag, although the protein was initially purified with an IMAC column. Protein degradation might occur when the protein was concentrated with a centrifugal protein concentrator and other purification processes before the size exclusion chromatography.

### Purified P-domain of *MrNV*-CP exists homogeneously as a dimeric form

The molecular mass and multimeric state of the P-domain of *MrNV*-CP were estimated by overlaying the gel filtration chromatograms of the P-domain and the standard markers. The overlaid chromatograms (Fig 4) shows that the peak containing the P-domain is located



**Fig 3. SDS-PAGE and western blot analyses of the purified protruding (P) domain of *Macrobrachium rosenbergii* nodavirus capsid protein (*MrNV*-CP).** (a) SDS-PAGE, and (b) western blot analysis. Lane M: protein marker in kDa; lane 1: concentrated protein sample after IMAC purification; lane 2: concentrated protein sample after gel filtration.

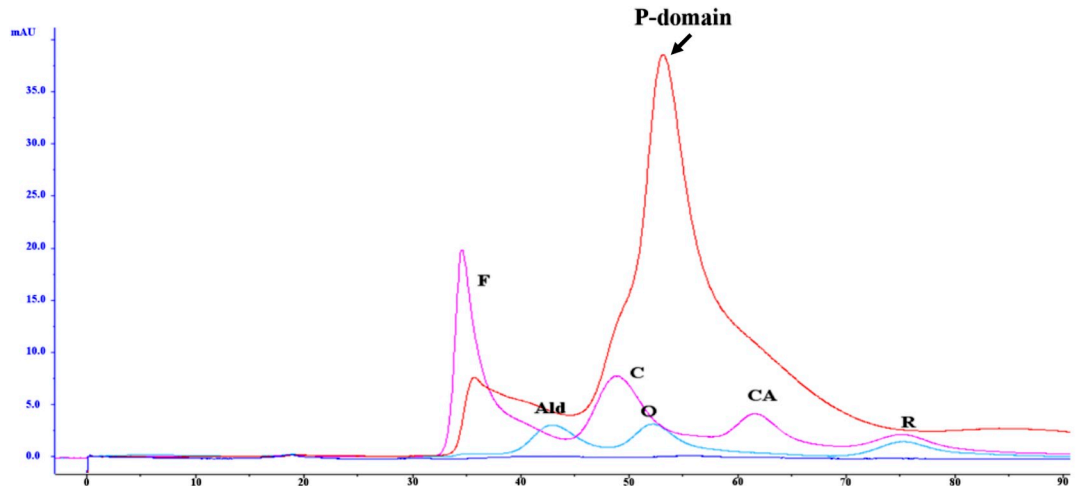
<https://doi.org/10.1371/journal.pone.0211740.g003>

between the peaks for carbonic anhydrase (CA, 29 kDa) and ovalbumin (O, 44 kDa), indicating that the P-domain may exist in the form of dimer with a molecular mass of ~36 kDa. This is in good agreement with the cryo-EM analysis reported by Ho et al. [12] who reported that the P-domain of *MrNV*-CP indeed self-assembled to form dimers. The purified P-domain was then characterized by dynamic light scattering (DLS). As shown in Fig 5A, the purified P-domain of *MrNV*-CP has a hydrodynamic diameter ~6 nm. The measured diameter of the dimeric particle corresponded well with the cryo-EM structure showing that the blade-shaped spike formed by the P-domain of *MrNV*-CP is around 4.7 nm in both width and height. It is believed that in the absence of the S-domain, the dimeric conformation of the P-domain alone is possibly distinct from that observed in cryo-EM analysis. In addition, the diameter of a particle determined by DLS is normally bigger than that measured by EM due to the fact that ions in solution associate with the particle and contribute to the size determined by DLS [24].

### P-domain of *MrNV*-CP is mainly composed of beta-stranded structures

Circular dichroism (CD) analysis was carried out to estimate the content of the secondary structures of the recombinant P-domain. Fig 5B shows the CD spectra scanned from near-UV wavelength 240 nm to 200 nm. The CD spectra showed that the protein contains a high



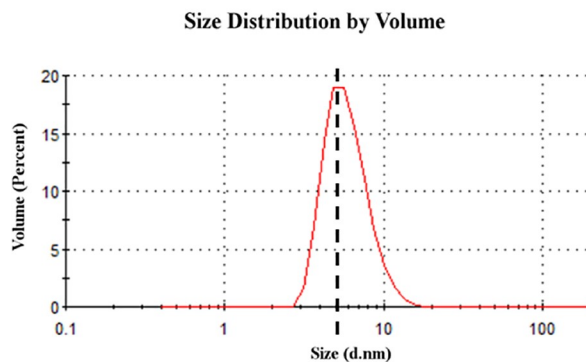


**Fig 4. Determination of the multimeric state of the protruding (P) domain of *Macrobrachium rosenbergii* nodavirus capsid protein (*MrNV*-CP) using gel filtration.** The protein fractionation profile of native protein markers and the P-domain of *MrNV*-CP fractionated by HiPrep 16/60 Sephacryl S-200 HR. Peaks of each protein marker are indicated [Abbreviation: F (Ferritin;  $M_r$  440 kDa), Ald (Aldolase;  $M_r$  158 kDa), C (Conalbumin;  $M_r$  75 kDa), O (Ovalbumin;  $M_r$  44 kDa), CA (Carbonic Anhydrase;  $M_r$  29 kDa), R (Ribonuclease A;  $M_r$  13.7 kDa)]. Red line indicates the chromatogram of the P-domain of *MrNV*-CP. Purple and blue lines indicate protein markers' chromatograms.  $M_r$  indicates relative molecular mass.

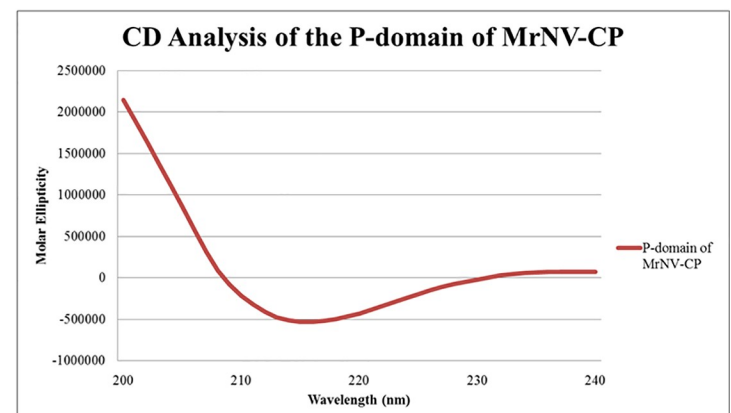
<https://doi.org/10.1371/journal.pone.0211740.g004>

percentage of beta-stranded structures, which are indicated by the negative bands around 218 nm—a characteristic spectrum for beta-stranded structures [25]. Other unique features for alpha-helical and floppy structures are not observed in the spectra. Based on Reed and Reed [23], the P-domain contained ~67.9% beta-stranded structures and 32.1% turns and unstructured loops. *In silico* analysis performed using Phyre 2 identified the crystal structure of cucumber necrosis virus (CNV) (PDB code: 4LLF [26]) as the closest homologous template to *MrNV*-CP, the same structure was used as a template to model a 3D structure of *MrNV* capsid [17, 18]. Amino acid sequence comparison between the P-domain of *MrNV* and CNV (amino acids 267–380) showed 19% identity and 27.5% similarity, respectively (Fig 6). The P-domain

(a)

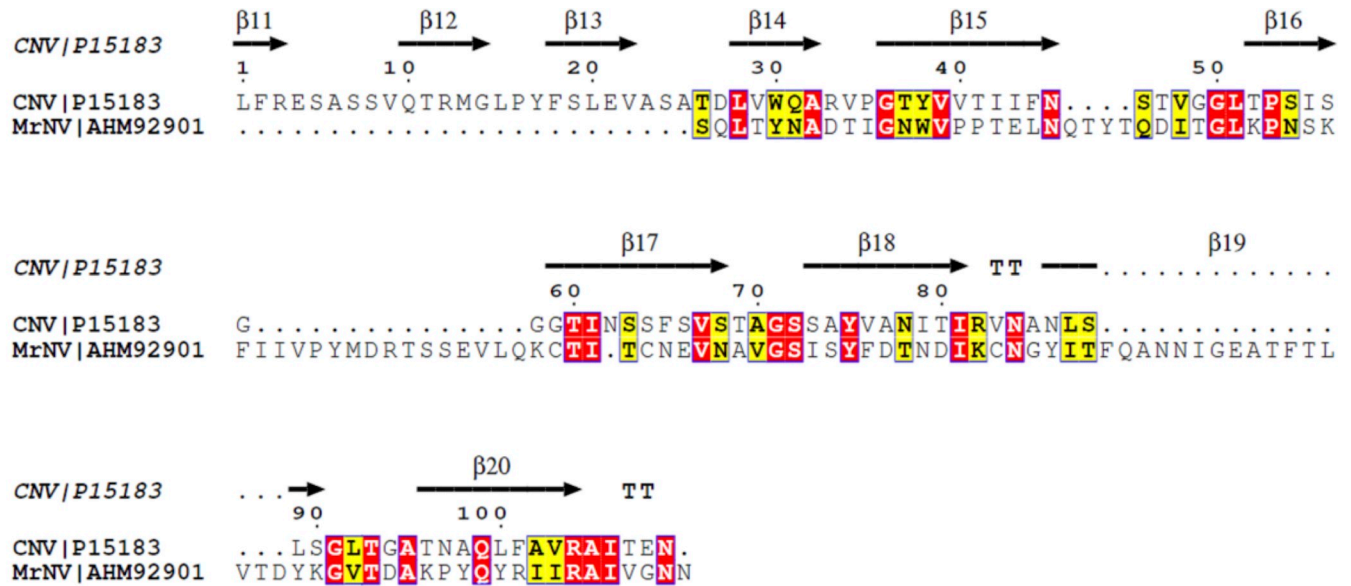


(b)



**Fig 5. Biophysical characterizations of the protruding (P) domain of *Macrobrachium rosenbergii* nodavirus capsid protein (*MrNV*-CP).** (a) Dynamic light scattering (DLS) analysis of the P-domain of *MrNV*-CP. The dotted line indicates the diameter of the P-domain. (b) Circular dichroism (CD) spectra of the P-domain of *MrNV*-CP scanned from near-UV wavelengths 240 nm to 200 nm. The CD spectra display the unique characteristic of beta-stranded structures, indicated by negative bands at wavelength 218 nm.

<https://doi.org/10.1371/journal.pone.0211740.g005>



**Fig 6. Pairwise sequence alignment of the protruding (P) domain of *Macrobrachium rosenbergii* nodavirus capsid protein (*MrNV*-CP) and cucumber necrosis virus (CNV).** The homology model used was cucumber necrosis virus (CNV). Beta-strands, turns and gaps are indicated by arrows, TT letters and dotted lines, respectively. Identical and similar amino acid sequences are highlighted in red and yellow, respectively.

<https://doi.org/10.1371/journal.pone.0211740.g006>

of the crystal structure of CNV contains 57% of beta-stranded structure and also has no alpha-helical structure. Structural similarity for protein pairs with amino acid sequence identity between 20–35% is often referred to as in the ‘twilight zone’, as the likelihood that protein pairs with sequence identity below 25% to have similar structures is less than 10% [27, 28]. In the present study, structural modeling using CNV as a template successfully generated a model with a confidence score of 98% (S1 Fig). The model covers 90% of the amino acid sequence of the P-domain of *MrNV*-CP. The result suggests that the P-domain of CNV could represent a homologous template for future structural studies of the P-domain of *MrNV*-CP.

## Conclusion

The P-domain of *MrNV*-CP was expressed as a soluble protein in *E. coli*. The recombinant P-domain formed dimers with a hydrodynamic diameter ~6 nm which resemble the full-length capsid protein. Our current data provide crucial biophysical parameters for downstream analyses of the P-domain of *MrNV*-CP such as structural analysis using X-ray crystallography or nuclear magnetic resonance. Availability of a high-resolution 3D structure of the P-domain is essential to elucidate the mechanisms involved in host cell recognition and penetration.

## Supporting information

**S1 Fig. Putative structure of the protruding (P) domain of *Macrobrachium rosenbergii* nodavirus capsid protein (*MrNV*-CP).** The model was generated by program Phyre 2 using the P-domain of cucumber necrosis virus (CNV; PDB code 4LLF) as a template. The confidence score of this model is 98% over 90% of the amino acid sequence of the P-domain of *MrNV*-CP.

(TIFF)

## Acknowledgments

We thank Prof. Dr. Malcolm Walkinshaw for critical reading of the manuscript.

## Author Contributions

**Conceptualization:** Kok Lian Ho.

**Data curation:** Li Chuin Chong, Hagilaa Ganesan.

**Formal analysis:** Li Chuin Chong, Chean Yeah Yong, Kok Lian Ho.

**Funding acquisition:** Wen Siang Tan, Kok Lian Ho.

**Methodology:** Hagilaa Ganesan, Chean Yeah Yong.

**Project administration:** Kok Lian Ho.

**Supervision:** Chean Yeah Yong, Wen Siang Tan, Kok Lian Ho.

**Writing – original draft:** Li Chuin Chong.

**Writing – review & editing:** Wen Siang Tan, Kok Lian Ho.

## References

1. Banu R, Christianus A. Giant Freshwater Prawn *Macrobrachium rosenbergii* Farming: A Review on its Current Status and Prospective in Malaysia. *Journal of Aquaculture Research & Development*. 2016; 7(4):1–5. <https://doi.org/10.4172/2155-9546.1000423>
2. New MB. Freshwater prawn culture: a review. *Aquaculture*. 1990; 88(2):99–143. [https://doi.org/10.1016/0044-8486\(90\)90288-X](https://doi.org/10.1016/0044-8486(90)90288-X).
3. Sahul Hameed AS, Bonami JR. White Tail Disease of Freshwater Prawn, *Macrobrachium rosenbergii*. *Indian J Virol*. 2012; 23(2):134–40. <https://doi.org/10.1007/s13337-012-0087-y> PMID: 23997437; PubMed Central PMCID: PMC3550746.
4. Sahul Hameed AS, Yoganandhan K, Sri Widada J, Bonami JR. Experimental transmission and tissue tropism of *Macrobrachium rosenbergii* nodavirus (MrNV) and its associated extra small virus (XSV). *Dis Aquat Organ*. 2004; 62(3):191–6. <https://doi.org/10.3354/dao062191> PMID: 15672874.
5. Arcier J-M, Francois H, Donald VL, Rita MR, Jocelyne M, Jean-Robert B. A viral disease associated with mortalities in hatchery-reared postlarvae of the giant freshwater prawn *Macrobrachium rosenbergii*. *Diseases of Aquatic Organisms*. 1999; 38(3):177–81.
6. Murwantoko M, Bimantara A, Roosmanto R, Kawaichi M. *Macrobrachium rosenbergii* nodavirus infection in a giant freshwater prawn hatchery in Indonesia. *Springerplus*. 2016; 5(1):1729. <https://doi.org/10.1186/s40064-016-3127-z> PMID: 27777864; PubMed Central PMCID: PMC5053962.
7. Qian D, Shi Z, Zhang S, Cao Z, Liu W, Li L, et al. Extra small virus-like particles (XSV) and nodavirus associated with whitish muscle disease in the giant freshwater prawn, *Macrobrachium rosenbergii*. *J Fish Dis*. 2003; 26(9):521–7. <https://doi.org/10.1046/j.1365-2761.2003.00486.x> PMID: 14575370.
8. Bonami JR, Shi Z, Qian D, Sri Widada J. White tail disease of the giant freshwater prawn, *Macrobrachium rosenbergii*: separation of the associated virions and characterization of MrNV as a new type of nodavirus. *J Fish Dis*. 2005; 28(1):23–31. <https://doi.org/10.1111/j.1365-2761.2004.00595.x> PMID: 15660790.
9. Hayakijkosol O, Owens L. Non-permissive C6/36 cell culture for the Australian isolate of *Macrobrachium rosenbergii* nodavirus. *J Fish Dis*. 2013; 36(4):401–9. <https://doi.org/10.1111/j.1365-2761.2012.01414.x> PMID: 23134578.
10. Yong CY, Yeap SK, Omar AR, Tan WS. Advances in the study of nodavirus. *PeerJ*. 2017; 5:e3841. Epub 2017/10/04. <https://doi.org/10.7717/peerj.3841> PMID: 28970971; PubMed Central PMCID: PMC5622607.
11. NaveenKumar S, Shekar M, Karunasagar I, Karunasagar I. Genetic analysis of RNA1 and RNA2 of *Macrobrachium rosenbergii* nodavirus (MrNV) isolated from India. *Virus Res*. 2013; 173(2):377–85. <https://doi.org/10.1016/j.virusres.2013.01.003> PMID: 23318596.
12. Ho KL, Kueh CL, Beh PL, Tan WS, Bhella D. Cryo-Electron Microscopy Structure of the *Macrobrachium rosenbergii* Nodavirus Capsid at 7 Angstroms Resolution. *Sci Rep*. 2017; 7(1):2083. <https://doi.org/10.1038/s41598-017-02292-0> PMID: 28522842; PubMed Central PMCID: PMC5437026.

13. Goh ZH, Tan SG, Bhassu S, Tan WS. Virus-like particles of *Macrobrachium rosenbergii* nodavirus produced in bacteria. *J Virol Methods*. 2011; 175(1):74–9. <https://doi.org/10.1016/j.jviromet.2011.04.021> PMID: 21536072.
14. Kueh CL, Yong CY, Masoomi Dezfooli S, Bhassu S, Tan SG, Tan WS. Virus-like particle of *Macrobrachium rosenbergii* nodavirus produced in *Spodoptera frugiperda* (Sf9) cells is distinctive from that produced in *Escherichia coli*. *Biotechnol Prog*. 2017; 33(2):549–57. Epub 2016/11/20. <https://doi.org/10.1002/btpr.2409> PMID: 27860432.
15. Jariyapong P, Chotwiwatthanakun C, Somrit M, Jitrapakdee S, Xing L, Cheng HR, et al. Encapsulation and delivery of plasmid DNA by virus-like nanoparticles engineered from *Macrobrachium rosenbergii* nodavirus. *Virus Res*. 2014; 179:140–6. <https://doi.org/10.1016/j.virusres.2013.10.021> PMID: 24184445.
16. Goh ZH, Mohd NA, Tan SG, Bhassu S, Tan WS. RNA-binding region of *Macrobrachium rosenbergii* nodavirus capsid protein. *J Gen Virol*. 2014; 95(Pt 9):1919–28. <https://doi.org/10.1099/vir.0.064014-0> PMID: 24878641.
17. Somrit M, Watthammawut A, Chotwiwatthanakun C, Ounjai P, Suntimanawong W, Weerachatanukul W. C-terminal domain on the outer surface of the *Macrobrachium rosenbergii* nodavirus capsid is required for Sf9 cell binding and internalization. *Virus Res*. 2017; 227:41–8. Epub 2016/11/05. <https://doi.org/10.1016/j.virusres.2016.09.017> PMID: 27693291.
18. Ho KL, Gabrielsen M, Beh PL, Kueh CL, Thong QX, Streetley J, et al. Structure of the *Macrobrachium rosenbergii* nodavirus: A new genus within the Nodaviridae? *PLoS Biol*. 2018; 16(10):e3000038. Epub 2018/10/23. <https://doi.org/10.1371/journal.pbio.3000038> PMID: 30346944; PubMed Central PMCID: PMC6211762.
19. Yoon KY, Tan WS, Tey BT, Lee KW, Ho KL. Native agarose gel electrophoresis and electroelution: A fast and cost-effective method to separate the small and large hepatitis B capsids. *Electrophoresis*. 2013; 34(2):244–53. Epub 2012/11/20. <https://doi.org/10.1002/elps.201200257> PMID: 23161478.
20. Bradford MM. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal Biochem*. 1976; 72:248–54. PMID: 942051.
21. Rueden CT, Schindelin J, Hiner MC, DeZonia BE, Walter AE, Arena ET, et al. ImageJ2: ImageJ for the next generation of scientific image data. *BMC Bioinformatics*. 2017; 18(1):529. Epub 2017/12/01. <https://doi.org/10.1186/s12859-017-1934-z> PMID: 29187165; PubMed Central PMCID: PMC5708080.
22. Yang JT, Wu CS, Martinez HM. Calculation of protein conformation from circular dichroism. *Methods Enzymol*. 1986; 130:208–69. Epub 1986/01/01. PMID: 3773734.
23. Reed J, Reed TA. A set of constructed type spectra for the practical estimation of peptide secondary structure from circular dichroism. *Anal Biochem*. 1997; 254(1):36–40. Epub 1998/01/31. <https://doi.org/10.1006/abio.1997.2355> PMID: 9398343.
24. Eaton P, Quaresma P, Soares C, Neves C, de Almeida MP, Pereira E, et al. A direct comparison of experimental methods to measure dimensions of synthetic nanoparticles. *Ultramicroscopy*. 2017; 182:179–90. Epub 2017/07/12. <https://doi.org/10.1016/j.ultramic.2017.07.001> PMID: 28692935.
25. Greenfield NJ. Using circular dichroism spectra to estimate protein secondary structure. *Nat Protoc*. 2006; 1(6):2876–90. Epub 2007/04/05. <https://doi.org/10.1038/nprot.2006.202> PMID: 17406547; PubMed Central PMCID: PMC2728378.
26. Li M, Kakani K, Katpally U, Johnson S, Rochon D, Smith TJ. Atomic structure of Cucumber necrosis virus and the role of the capsid in vector transmission. *J Virol*. 2013; 87(22):12166–75. Epub 2013/09/06. <https://doi.org/10.1128/JVI.01965-13> PMID: 24006433; PubMed Central PMCID: PMC3807921.
27. Kinjo AR, Nishikawa K. Eigenvalue analysis of amino acid substitution matrices reveals a sharp transition of the mode of sequence conservation in proteins. *Bioinformatics*. 2004; 20(16):2504–8. Epub 2004/05/08. <https://doi.org/10.1093/bioinformatics/bth297> PMID: 15130930.
28. Rost B. Twilight zone of protein sequence alignments. *Protein Eng*. 1999; 12(2):85–94. Epub 1999/04/09. PMID: 10195279.