PHILIPP AGRIC SCIENTIST
Vol. 104 No. 4, 371-387
December 2021

ISSN 0031-7454

Serine/Threonine Protein Phosphatase 1-α (STPP1-α) from Black Tiger Shrimp, *Penaeus monodon*, an Immune-Related Gene

Jacqueline Marjorie R. Pereda[1,2,*], Neil Andrew D. Bascos[3], and Mudjekeewis D. Santos[1]

[1]Genetic Fingerprinting Laboratory, National Fisheries Research and Development Institute, 101 Mother Ignacia Ave., South Triangle, Quezon City, 1103, Philippines
[2]Far Eastern University, Nicanor Reyes Street, Sampaloc, Manila, 1015, Philippines
[3]National Institute of Molecular Biology and Biotechnology, University of the Philippines Diliman, Quezon City, 1101, Philippines

[*]Author for correspondence; E-mail: jacquerpereda@gmail.com

Received: June 14, 2021/ Revised: October 02, 2021/ Accepted: November 17, 2021

Reversible protein phosphorylation is a significant regulatory mechanism in many cellular functions, such as the dephosphorylation of Serine/Threonine protein residues catalyzed by protein phosphatase. In this study, the full-length STPP1-*α* gene from *Penaeus monodon* was cloned, characterized, and analyzed for its constitutive expression in WSSV-negative *P. monodon* organs. The gene was originally an isotig isolated from the gills of *P. monodon* that survived WSSV infection. *PmSTPP1-α* gene (GenBank: KX385833) has a total of 2,171 bp, with a 990 open reading frame (ORF) that encodes 329 amino acids (aa), sharing a 93% sequence identity with human Serine/Threonine PP1-α catalytic subunit. The protein has a single conserved catalytic domain and shares almost all the conserved sites and functional residues of the human protein phosphatase, which particularly might have putative functions to viral protein synthesis. Using clustering analysis, *PmSTPP1-α* was verified to be the –*α* isoform while the reported *L. vannamei* STPP1 is the β form. *In silico* homology modelling predicts similar structures for *PmSTPP1-α* and STPP1 from *H. sapiens*. Conserved functional domains for metal binding, target protein interaction, and toxin binding sites were identical in sequence and predicted structure. The observed variations in amino acid sequence were outside these conserved domains but should be further studied to determine their potential effects on function. Current molecular docking predictions for *PmSTPP1-α* against three proteins from known *P. monodon* pathogens suggest specific functional interactions with the target protein binding domain and the molecular toxin interaction sites. *PmSTPP1-α* was found to be ubiquitously and highly expressed in organs of WSSV-negative *P. monodon*, further investigations on the interactions of this protein will help validate its predicted involvement in the *P. monodon* immune response.

Keywords: molecular cloning, immunity, disease, white spot syndrome virus, Crustaceans

Abbreviations: WSSV – white spot syndrome virus, *PmSTPP1-α* – *Penaeus monodon* Serine/Threonine Protein Phosphatase 1-α, PP1 – Protein Phosphatase 1, STPP1-α – Serine/Threonine Protein Phosphatase 1-α

INTRODUCTION

In the cell, reversible phosphorylation of proteins is an essential process that is triggered by a stimulus from its surface causing various changes in the function and activities of the intracellular proteins. Regulated changes in the state of protein phosphorylation and dephosphorylation are caused by either the protein kinases or phosphatases (Cieśla et al. 2011). Serine/Threonine Protein Phosphatase 1 (PP1) has been found to be highly involved in a variety of cellular processes such as protein synthesis, apoptosis, meiosis and cell division, cytoskeletal reorganization, regulation of membrane receptors and channels, and metabolism (Bollen et al. 2010). Collectively, PP1 isoforms have been shown to act on a broad substrate, although there are some such as PP1 holoenzyme, catalyzes a specific substrate to affect a definite biological response (Shi 2009). Though, the three subunits of Serine/Threonine Protein Phosphatase 2A (PP2A) have been fully elucidated in black tiger shrimp (Zhao et al. 2016), further studies on shrimp's PP1 mechanism are still lacking.

White spot syndrome virus (WSSV) is one of the most dreaded viruses in the shrimp industry as it causes massive mortality and major damage to many types of shrimp farming operations (Dieu et al. 2004). The limited

knowledge regarding the disease's pathogenicity and the shrimp's immunity warrants further understanding in order to find means of combating the disease, and help the shrimp industry. While studies have been conducted to understand WSSV-shrimp pathogen interaction, and identifying and analyzing genes that are related to WSSV (Li and Xiang 2013; Song and Li 2014), these studies have likewise produced unknown, partial gene fragments that are potentially involved in WSSV infection (Maralit et al. 2014). It is important that these unknown, partial fragments be fully elucidated, as these fragments, when completed, could lead to a better understanding of the shrimp immunity in general. The generation of fully annotated novel genes, if confirmed to occur in shrimp tissues, could then be used later for possible functional *in vivo* experiments.

In the study of Maralit et al. (2014), an unknown partial gene fragment Isotig00463 (i00463) was singled out from a suppression subtractive hybridization (SSH)-next-generation sequencing (NGS) transcriptome database of gene isotigs isolated from the gills of black tiger shrimp, *P. monodon* that survived WSSV infection. Here, this i00463 was fully-cloned and characterized to be the shrimp ortholog of Serine/Threonine Protein Phosphatase 1-α, named here as *PmSTPP1-α*. We provide its complete nucleotide sequence, protein structure, and its constitutive organ expression in WSSV-negative *P. monodon*. This is the first cloning report of the α isoform of STPP1 in shrimp following the report of the β isoform in Pacific white shrimp *Litopenaeus vannamei*.

MATERIALS AND METHODS

Shrimp Samples A live adult *P. monodon* used for the cloning part of this study, was collected in a local market in Pasay City, Philippines. On the other hand, three (3) live adult *P. monodon*, procured from a shrimp pond located in barangay Sta. Monica, Hagonoy, Bulacan, Philippines, were used for the organ expression analysis in gills, hepatopancreas, intestine, muscle, lymphoid organ, heart and hemolymph. All of the *P. monodon* samples tested negative for WSSV by PCR using the primers of Flegel et al. (2006) (136198F: 5′ GTACGGCAATACTGGAGGAGGT 3′ and 136429R: 5′ GGAGATGTGTAAGATGGACAAG 3′). These tissues and organs were dissected out and preserved in screw cap tubes containing RNALater. The collected tissue samples were shipped to the Genetic Fingerprinting Laboratory and stored in a -80˚C Ultralow Refrigerator prior to the experiment.

Cloning of Full Length *PmSTPP1-α* cDNA Gene

An unknown partial fragment Isotig00463 (i00463) was obtained from transcriptome database of Maralit et al. (2014) accessible at NCBI Sequence Read Archive with accession number SRR57708030 (http://www.ncbi.nlm.nih.gov/sra). RNA from gills of *P. monodon* was extracted using QIAzol® Lysis Reagent (QIAGEN) following the manufacturer's protocol. Total RNA of the samples was purified using RNeasy Mini Cleanup Kit (QIAGEN) following the manufacturer's protocol while its concentration (µg per µl) was determined using Implen™ nanophotometer P-Class.

The total RNA template used in the first strand cDNA synthesis was 0.103 µg. A total of 10 µl reaction mix was composed of the following: 0.5X first strand buffer, 2 mM DTT, 0.1 mM dNTP mix, SMARTerOligo IIA (for 5′ cDNA synthesis), 5U RNAse inhibitor, 10 U SMARTScribe Reverse Transcriptase, and 1 mM of respective primers for 5′ and 3′ cDNA synthesis. The reaction mix was placed in a thermal cycler and programmed to 42°C for 90 min and heated to 70°C for 10 min. After which, 20 µl of tricine-EDTA buffer was added to the samples for dilution.

Rapid Amplification of cDNA Ends (RACE) PCR was employed to complete the full length of i00463. Gene specific primers were designed (Table 1) and primer annealing sites are shown in Fig. S1. Reactions were done utilizing a total of 25 µl reaction mix using Advantage 2 Kit (Clontech) with composition as follows: 1X PCR buffer, 1 mM dNTP mix, 1X polymerase mix, 1X universal primer, and the generated gene specific primers. The reaction mix was amplified in a thermal cycle with conditions as follow: initial denaturation at 94°C for 3 min, 38 cycles of 94°C for 30 sec, 70°C for 30 sec and 72° C for 3 min, then final extension at 72°C for 10 min. The quality of the PCR amplicons were assessed and documented by visualizing on a 1% gel stained with ethidium bromide in SynGene G:BOX. The resulting amplicons were sent to 1st Base (Malaysia) for sequencing.

In Silico Analysis

In silico analysis was performed following Santos et al. (2006 and 2007) with modifications. The sequences, including

Table 1. Primer sequences used to amplify and sequence the full-length serine/threonine protein phosphatase 1–α gene from *P. monodon* using RACE PCR.

PRIMER NAME	SEQUENCE (5' → 3')
PmSTPP1-α 5RACE-1	CTCTGGACTTGAGGCATAGTCCACG
PmSTPP1-α 3RACE-1	GGTGACATCCACGGACAGTACTACG
PmSTPP1-α 3RACE-2	CCAGACCAGGGCTTATTGTGCGATC
PmSTPP1-α 3RACE-3	GCAAGTCTGAGGTTGGCCTTTGTTC
PmSTPP1-α 3RACE-4	GTCCAGTTGCCGTAGTAGCAGGAGC

the nucleotide and translated amino acids, average molecular weight, isoelectric point, and extinction coefficient were examined and obtained using Geneious 6.1.8 (Biomatters). The identities of nucleotide and amino acid sequence were analyzed through regions of similarity between other biological sequences using BLASTn and BLASTp. The functional and conserved domain of the amino acid sequence was determined using the Conserved Domain Architecture Retrieval Tool (CDART) (http://www.ncbi.nlm.nih.gov/Structure/lexington/lexington.cgi). Complete multiple alignments were obtained using ClustalW with default parameters. The alignment (.aln) file format produced from the former tool was utilized for the clustering analysis in MEGA 6 (http://www.megasoftware.net/) using the Neighbor-joining (NJ) tree based on Poisson model with 500 bootstrap replications and complete deletion of sites (Tamura et al. 2013). The sequences of the members of MPP superfamily (GENBANK: cl13995) were retrieved from NCBI.

Structural Analysis

BLAST Analysis

Structures related to the *PmSTPP1-α* sequence were determined by submitting the polypeptide sequence for BLAST analysis at the UniProt website (uniprot.org 2021). Related sequences with available 3D-structures were extracted and compared with the target sequence to determine percentage identity, and sites of conservation and variation. Structures of the related sequences were acquired from the Protein Data Bank (www.rcsb.org; Berman et al. 2000) for further analysis. Structures of the *PmSTPP1-α* sequence were predicted using the Magic Fit function of the DeepView Molecular Viewer (Guex et al. 1997). This function fits the submitted sequence unto a reference structure (e.g. *H. sapiens* STPP1-α; PDBID 3e7a (Kelker et al. 2009)). Comparison of the fit and reference sequences was done to observe predicted locations of conservation and variance between the two proteins.

Homology Modelling

Predictions on the potential 3D-structure of the target protein sequence (i.e. *PmSTPP1-α*) were made using the ITASSER Homology Modelling server (https://zhanggroup.org/I-TASSER/). This system makes predictions on the potential structures and functions of target proteins based on comparisons with related/homologous proteins with curated structures in the protein data bank. Model structures for the target protein are returned, as well as identities of the top threading templates that served as references for the structure prediction. Predictions for enzyme class and ligand binding sites are also provided in the results. The returned models for the submitted sequence were acquired and subsequently analyzed.

In Silico Protein Docking Predictions

The predicted involvement of *PmSTPP1-a* with *P. monodon* immune response was assessed based on predicted associations with protein targets from known Shrimp pathogens. Specifically, protein docking experiments were done with three proteins (PirA, PirB, and VP24). Both PirA and PirB are toxins from *V. parahaemolyticus* (Lee et al. 2015). While V224 is a major envelope protein of White Spot Syndrome Virus (Sun et al. 2016). Molecular structures for these target proteins were acquired from the Protein Data Bank (www.rcsb.org: Berman, et al. 2000); PirA and PirB: PDBIDs: 3X0T and 3X0U; Wang et al. 2014; VP24: PDBID 5HLJ; Sun et al. 2016).

Docking experiments were conducted using the ClusPro Protein Docking server (https://cluspro.org/; Kozakov et al. 2017). This service generates predictions of docking interactions between submitted "receptor" and "ligand" protein partners. The predictions are then assessed for relevance based on their prevalence for the multiple replicate predictions. Models are generated based on these prevalent dock clusters and returned to the user. Models of the top 10 docked structures based on specific criteria (balanced, electrostatic-favored; hydrophobic-favored; VDW + electrostatic-favored) are then used to assess the interactions between the submitted protein pairs. As interactions between proteins are often based on combined interaction types (i.e electrostatic, hydrophobic, and VDW interactions) the top ranked models based on "balanced" factors were used as bases for comparing the different partner proteins. However, additional information on the potential bases that promote specific docking conformations were derived on observations of similar structures in models that were ranked highly when considering the other factors.

To assess the relevance of the predicted docking models for *PmSTPP1-a* and the target proteins, the placement of the docked structures was compared relative to documented functional domains for STPP1-α (e.g. toxin-binding-sites; target-protein binding site; etc.).

Constitutive Expression Analysis

Expression in tissues and organs was determined using the Reverse Transcription–Polymerase Chain Reaction (RT–PCR) method following Santos et al. (2006 and 2007)

with some modifications. Total RNA from various organs and tissues of three (3) WSSV-negative *P. monodon* samples were obtained using QIAzol® Lysis Reagent (QIAGEN) following the manufacturer's protocol. The reverse transcription (RT) reaction mix of each sample was composed of the following: 1X Buffer RT, 0.5 mM dNTPs, 1 μM Oligo-dT primer, 10 units RNase inhibitor, 4 units Omniscript Reverse Transcriptase (QIAGEN), with 0.254 μg of total RNA template and amplified using Gene Specific Primers (GSP) forward (5'-CATCCGAGTCTGTGGTGTAG-3') and reverse (5'-GATCGCACAATAAGCCCTGG-3'). The Elongation Factor 1- α (EF1-α) was amplified using the following primers: EF1-α forward primer (5'-ATGGTTGTCAACTTTGCCCC-3') and EF1-α reverse primer (5'-TTGACCTCCTTGATCACACC-3') (Dang et al. 2010). The cycling parameters used for the amplification were: initial denaturation at 95°C for 5 min, 30 cycles of denaturation at 95°C for 30 sec, annealing at 60°C for 30 sec, 72°C for 30 sec, and final extension at 72°C for 10 min. The negative control which contained no template was also included. PCR amplicons were assessed and visualized on a 1% gel stained with ethidium bromide and documented with SynGene G:BOX.

RESULTS AND DISCUSSION

PmSTPP1-α cDNA and Protein

BLAST and *in silico* analysis showed that the unknown partial gene fragment as Serine/Threonine Protein Phosphatase 1 Catalytic Subunit – α isoform, named here as *PmSTPP1-a* (GenBank: KX385833) (Fig. 1). *PmSTPP1-a* cDNA sequence has a total of 2,171 base pairs (bp) with an open reading frame (ORF) of 990 bp encoding for a 329 aa protein with a molecular weight of 37.6 kDA, an isoelectric point of 6.75, and an extinction coefficient of 36, 620. The upstream- and downstream untranslated region (UTR/DTR) consisted of 139 base pairs and 1,042 base pairs, respectively. The promoter sequence, such as the downstream promoter element (AGACA), was in the location of +27 to +31 nucleotides downstream from the initial codon (ATG). The Kozak's sequence (ATCATGG) is also present in the generated full-length gene sequence. Whereas, two (2) of the poly-adenylation signals (AATAAA) have occurred near the C-terminus non-coding region of the gene, which ended with poly-A tail.

PmSTPP1-a has 93% identity with Serine/Threonine Protein Phosphatase 1 Catalytic Subunit – α of humans (*Homo sapiens*) (Accession no.: NP_002699.1) and 92% identity with that of African clawed frog (*Xenopus laevis*) (Accession no.: NP_001080222.1) (Fig. S2). A recent study identified a novel protein phosphatase (PPs) derived from the Pacific white shrimp (*Litopenaeus vannamei*) cell (Lu and Kwang 2004). It was reported to only consist of 199 aa, contains almost all the functional catalytic domains of human protein phosphatase except the C-terminal non catalytic sequence. Protein Phosphatase 1 gene was also found in other invertebrates such as in parasitic protozoan (*Trypanosoma brucei brucei*) (Accession no.: AAZ10947.1), and fruit fly (*Drosophila melanogaster*) (Accession no.: NP_001262919.1); however, these two sequences have low identity match with the *PmSTPP1-α*. Interestingly, BLAST search showed that *PmSTPP1-a* is most closely identical to a hypothetical protein from water flea (*Daphnia pulex*) (Accession no.: EFX67868.1) with 98% identity. Moreover, *PmSTPP1-a* has 99.77% identity with 100% query cover matched with a recent NCBI sequence record of a predicted sequence of *Penaeus monodon* Serine/Threonine-Protein Phosphatase alpha-2 isoform (Accession no.: XM_037948540.1), which was derived from a genomic sequence of *Penaeus monodon* isolate on its chromosome 41 with a whole genome shotgun sequence (Accession no.: NC_051426.1) (The NCBI sequences of the latter two (2) mentioned accessions are not included in the clustering tree).

NJ analysis showed *PmSTPP1-α* clustered with STPP1 clade (Fig. 2). Moreover, a 100% bootstrap value separated the 2 subclades, the STPP1-α/-γ (where the *PmSTPP1-α* clustered) and the STPP1-β (where the previously reported *L. vannamei* PPs grouped). Within the STPP1-α/-γ subclade, *PmSTPP1-α* clustered with STPP1-α, hence we concluded that *PmSTPP1-α* was indeed of the –α isoform while the *L. vannamei* is the STPP1- β termed here as *LvSTPP1- β* (Lu and Kwang 2004).

CDART shown *PmSTPP1-α* to be under superfamily Metallophosphatase (MPP) (GenBank: ofcl13995). The said gene possesses a single catalytic domain homologous with other STPP subfamily that starts with Leucine (L7) and ended with Alanine (A299) (Fig. 1 and Fig. S2). And, based on the result of the tertiary structure (3D) prediction (Phyre^2 software), *PmSTPP1-α* was predicted as not conserved from Lysine (K301) up until Lysine (K319). This structure is very similar to that of mammalian PP1s. In mammals, all PP1 catalytic domains (PP1Cs) are affected by alternative splicing, which causes its non-conserved N and C termini to be cleaved. The N-terminus of PP1Cα and PP1Cγ are nearly similar but the C-terminus of all PP1C isoforms is divergent (Korrodi-Gregório et al. 2014). On the C-terminal end of PP1Cα structures, about 25 - 30 aa shown a simple coil (Terrak et al. 2004). This structure is thought to form a complex with some regulatory subunits to mediate isoform specificity (Ma et al. 2015); and also functions as an epitope for

PmSTPP1-α

```
          ACATGGGGGGCGCCATCTTGAATAGTCTATGGGGTTGGGCTGTTGTGAGGTTGTGGCAGTGTATCGGCT   69
         CTTATCTGCCTCTCATTGAACCATCCGAGTCTGTGGTGTAGTGCAAAGTGATCAGGATAACCTTCTCATC  139
ATGGCGGAGACGGATAAGCTTAATATAGACAGCATCATCGCCAGATTATTAGAAGTTCGAGGGTCTCGCCCAGGA  214
 M  A  E  T  D  K  L  N  I  D  S  I  I  A  R  L  L  E  V  R  G  S  R  P  G   25
AAGAATGTGCAGTTAACCGAGAATGAAATTCGTGGACTCTGCCTCAAGTCCAGAGAAATCTTCCTCTCACAACCA  289
 K  N  V  Q  L  T  E  N  E  I  R  G  L  C  L  K  S  R  E  I  F  L  S  Q  P   50
ATCCTTCTGGAATTGGAAGCT CT CTCAAAATTTGTGGTGACATCCACGGACAGTACTACGATTTGCTTCGATCC  364
 I  L  L  E  L  E  A  P  L  K  I  C  G  D  I  H  G  Q  Y  Y  D  L  L  R  L   75
TTTGAATACGGCGGCTTCCCCCCAGAGTCCAACTACTTGTTTTTAGGAGATTACGTCGATCGTGGTAAACAGTCC  439
 F  E  Y  G  G  F  P  P  E  S  N  Y  L  F  L  G  D  Y  V  D  R  G  K  Q  S  100
TTGGAAACCATATGTCTCCTCTTGGCCTACAAGATCAAATATCCAGAAAATTTCTTCCTTCTCAGAGGCAACCAT  514
 L  E  T  I  C  L  L  L  A  Y  K  I  K  Y  P  E  N  F  F  L  L  R  G  N  H  125
GAGTGTGCATCCATCAACAGAATCTATGGTTTCTACGACGAATGCAAGCGGCGGTATAACATCAAACTGTGGAAG  589
 E  C  A  S  I  N  R  I  Y  G  F  Y  D  E  C  K  R  R  Y  N  I  K  L  W  K  150
ACTTTCACAGACTGTTTCAATTGCTTACCTGTCGCGGCCATTGTGGATGAAAAGATCTTCTGCTGCCACGGAGGC  664
 T  F  T  D  C  F  N  C  L  P  V  A  A  I  V  D  E  K  I  F  C  C  H  G  G  175
TTGAGCCCCGACCTCCAGAGCATGGAACAGATCCGTCGCATTATGAGGCCTACCGACGTGCCAGACCAGGGCTTA  739
 L  S  P  D  L  Q  S  M  E  Q  I  R  R  I  M  R  P  T  D  V  P  D  Q  G  L  200
TTGTGCGATCTTTTGTGGTCAGATCCAGACAAGGACACAATGGGCTGGGGAGAAAATGATAGAGGTGTATCCTTC  814
 L  C  D  L  L  W  S  D  P  D  K  D  T  M  G  W  G  E  N  D  R  G  V  S  F  225
ACTTTTGGGGCAGAAGTAGTTGCAAAGTTCCTCCACAAGCATGACTTTGACCTCATCTGTCGAGCTCATCAGGTT  889
 T  F  G  A  E  V  V  A  K  F  L  H  K  H  D  F  D  L  I  C  R  A  H  Q  V  250
GTTGAAGATGGTTACGAGTTCTTTGCCAAGCGACAACTAGTAACATTATTCTCAGCACCCAACTACTGTGGGGAA  964
 V  E  D  G  Y  E  F  F  A  K  R  Q  L  V  T  L  F  S  A  P  N  Y  C  G  E  275
TTTGACAATGCAGGTGCAATGATGTCTGTAGATGAGACACTTATGTGTTCCTTCCAGATACTTAAGCCAGCAGAT  1039
 F  D  N  A  G  A  M  M  S  V  D  E  T  L  M  C  S  F  Q  I  L  K  P  A  D  300
AAAAAGAAGTTCCCCTATGGAGGATTGAACACAGGGCGACCCGTAACGCCGCCGCGAGGAGCAGCCAATCAGAAA  1114
 K  K  K  F  P  Y  G  G  L  N  T  G  R  P  V  T  P  P  R  G  A  A  N  Q  K  325
AACAAGAAGAAATAAAACTAGTGTTGGTGTCACTGTCTTGGTGTTGGCAAGTCTGAGGTTGGCCTTTGTTCATCCT  1191
 N  K  K  K  *                                                                329
TGCCTTAAGTTTTGTGGATTGCTATCAGGGGATGTGACTTCAAAAAGAAATAAATATTGAATTTAACTGGGTTATTTA  1269
TAATTAACATATTTTAATCCTGGATGAACAAGGGTAATTAGACATTGCAACCCAGGACATAAATCTGTTAATGAAATA  1347
ATTTTATGAATTTCAGTTGCTAGATTATCCATGATAGTTATATTTATTGGTACATTGGCCCAGGTTCCCTCAGTAAAG  1425
TTATTTTTCTATTTTGAACCTTCATTTTTGTTTTATTCATTTATAATTATTAATTACTGTATTTTTTAGGGCAAAGTC  1503
AAACTTTTGTTTTATCCGAAATAGGTTTTATCAGCAGTATTAATGCTAAAAGGTTTGTCTATTGGTTCCAAATTTTCT  1581
AGAAAAGAATTGGCTTGCAACTGTGATTATCGGAAAATTTGTTCAACTTCAGTTTATATCTTGGAAGATAAGTTTGAC  1659
TTGTATTAATATATATATATTTCAGTGTGTAAATAGCCATGGTATTTAACCGCATATAAGCATTTACCATTGTTCTTG  1737
TATAAACTTTAGGGCCCAAAGAAGACTCGAGTGAAGCAGCGAGGGGATGATAGTTAGAGAAGCTGCTGGTTAGGTTCA  1815
TGTACAAATGTCCAGTTGCCGTAGTAGCAGGAGCATCTGGCTTTTGCTATCGCTTGTATAGTCAGAACAAGAGCGCAC  1893
TGGTGGTGGTATGTCGAGAGCCCAACCCCTACCCAAAACCTTTCGAGAGAGAAAGAAATGACTCTCTTTCCCCACACC  1971
CGAGGACTTGCATCATACATCCACCAGGTAAACCAGGTATTGGCTAATATACCCATGCTAGATTTGTGGAGTTTGGGC  2049
ATGGTGTGAAAAGTCCCTTAATGTGCAAAAGTAAAATACATGTACCAAAGCAGATGTAAATTGTGGTTGTTAATTCAA  2127
TAAAAGCGTTCATCTCGGAAAAAAAAAAAAAAAAAAAAAAAAAAA                                  2171
```

Fig. 1. Full-length cDNA and *PmSTPP1-α* amino acid sequence showing single conserved domain (highlighted in gray) – CDART result in multi-domain was PP2Ac which is a homologue of PP1 catalytic subunit. The Kozak's sequence and down promoter element are colored red. The two poly-adenylation sites (AATAAA) and poly-A tail were underlined.

developing isoform-specific antibodies (Korrodi-Gregório et al. 2014). The PP1C C-terminus has been documented to interact with myosin phosphatase (MYPT1) to promote dephosphorylation in the smooth muscle tissue. This interaction has been shown to promote structural changes in the PP1C C-terminus, from a simple coil to a more ordered state (Terrak et al. 2014).

PmSTPP1-α possesses two (2) aspartic acids (D64 and D92), three (3) histidines (H66, H173, and H248), and one (1) asparagine (N124) that are conserved in human PPPs. These conserved amino acids are thought to catalyze reactions when two metal ions bind and activate a water molecule, inducing a nucleophilic attack on the phosphorus atom. The two metal ions that contributed to the catalysis for *PmSTPP1- α* were assumed to be either Zn^{2+}, Cu^{2+}, Ni^{2+}, Fe^{2+}, and Mn^{2+} (Dancheck et al. 2011; Kolupaeva 2019), while the bimetallic Iron (Fe^{2+}) and Zinc (Zn^{2+}) were specific for PPP family (Heroes et al. 2015; Verbinnen et al. 2017). Structures of the PP1 complexes showed that the α-isoform was determined to contain two Mn^{2+} ions that interacts with spinophilin, which is a neuronal regulator protein bound (Ragusa et al. 2010). The bimetallic center is surrounded by the said metal ligands and is known to be the active site of the catalytic sub-unit.

Within the catalytic sub-unit of *PmSTPP1-α*, a total of three (3) characteristic conserved sequence motifs and three (3) conserved sequence residues were present; GDIHG, GDYVDRG, GNHE, HGG, RG, and H (G, glycine; D, aspartic acid; I, isoleucine; H, histidine; V, valine; R, arginine; N, asparagine; E, glutamic acid; Y, tyrosine). *PmSTPP1-α* showed a conserved RVxF motif suggesting that it functions similarly with the RVxF motif of PP1 in humans and the shrimp could possess that same PP1 regulatory proteins that would be interesting to study further. The said motifs and sequences were inclusive in PPP family (Brautigan 2013) and functions in metal coordination and phosphate binding (Shi 2009). Since the said motifs and sequences were conserved in STPP1 as shown in the sequence alignment (Fig. S2), it is likely that *PmSTPP1-α* also functions the same way.

PP1 regulatory proteins act as either substrate-independent activity regulator, targeting subunits, substrate specifier or substrate that interacts to the catalytic subunit (Heroes et al. 2013). They mostly interact with PP1C via RVxF binding motif (Korrodi-Gregório et al. 2014). The RVxF motif is composed of the consensus sequence K/R][K/R][V/I][x][F/W], denoted in standard single letter amino acid code: Lysine (K) or Arginine (R), Valine (V) or Isoleucine (I), Phenylalanine (F) or Tryptophan (W); where x is any residue other than Phenylalanine (F), Isoleucine (I), Methionine (M), Tyrosine (T), Aspartic Acid (D), or Proline (P). The RVxF

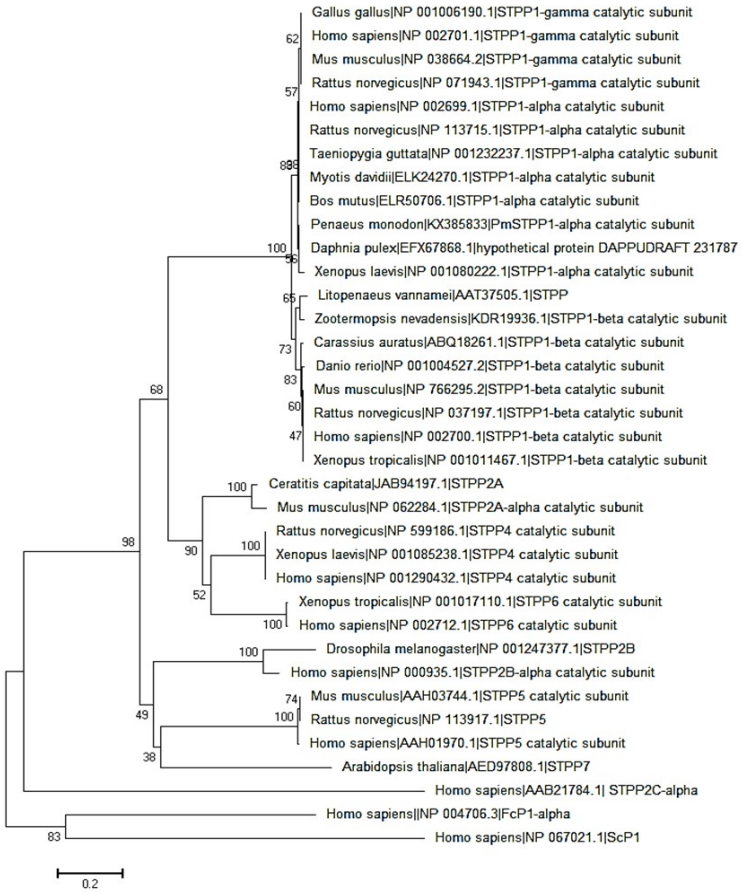


Fig. 2. Neighbor-Joining Tree of *PmSTPP1-α* protein sequence using Poisson Model. *PmSTPP1-α* was denoted with yellow-filled box while *LvSTPP1-β* was marked with broken-line unfilled box. Voucher sequences of Serine/Threonine Protein Phosphatase (PPP, PPM and FCP/SCP families) amino acid of the species indicated above were included in the analysis.

interaction functions in the protein binding regulation and doesn't influence the catalytic activity of the active site of PP1C. The glycogen-targeting subunit – muscle (GM) and inhibitor-2 (I-2) regulatory protein binds to the PP1C with the RVxF binding pockets (Peti et al. 2013) as specified in the sequence alignment (Fig 4).

PmSTPP1-α Structural Analysis

BLAST Analysis

The UniProt BLAST search results identified 2 entries in the database with available 3D structures. These entries were for STPP1-α proteins from *H. sapiens* and *O. cuniculus*. The sequences of these related proteins were determined and aligned pairwise with *PmSTPP1-α* using Emboss Needle (https://www.ebi.ac.uk/Tools/psa/emboss_needle/). Initial alignment of the related sequence from the *H. sapiens* STPP1a structures (PDBID 3e7a Chain A; 293 aa) and the *PmSTPP1-α* sequence (329aa) revealed 86.9% identity. The variance mainly occurred in the N and C terminal sections which did not align for the two sequences. Pairwise alignment between length matched sequences (293aa) of *PmSTPP1-α* and STPP1-α from *H. sapiens* was found to have 97.6% identity.

Model Fitting

The polypeptide sequence of *PmSTPP1-α* was fit unto the related sequence of *H. sapiens* STPP1α, HSSTPP1a, (PDBID 3e7a, Kelker et al. 2009) using the Magic Fit function of the DeepView Molecular Viewer (Guex et al. 1997). The generated structures showed good alignment of the conserved regions of PP1α structure as listed in Table 2. Locations of the variant residues between the two structures are shown in Figure 3 B-D. The sites of variation are observed to be outside the conserved areas in STPP1-α, proteins. A comparison of the line models for the *H. sapiens* STPP1-α residues at the sites of variation in the proposed *P. monodon* STPP1-α shows similar orientations for most of the residue sidechains (Figure 3 B-D). The most variance is observed for the position of Met208 and Phe235 in the fit *PmSTPP1-α* structure (Fig 3 C).

Homology Model Analysis

Structural models were also predicted using the ITASSER Homology Modelling Server (https://zhanggroup.org/I-TASSER/). The polypeptides sequence of *PmSTPP1-α* was submitted for modelling and several models were generated. Structure model predictions are made by threading the submitted sequence through related/homologous reference structures. The top threading reference for *PmSTPP1-α* was a structure for STPP1-β from *H. sapiens* (PDBID 1s70; Terrak et al. 2004). This top model was given a confidence score (C-score) of 0.16. C-scores range from -5 to 2, with scores nearing 2 representing higher confidence for the predicted structure. As noted in the sequence alignment, most of the variance between *PmSTPP1-α* and its references occur in the unmatched N and C-terminal sections. In fact, a resubmission of a trimmed *PmSTPP1-α* sequence without these sections returns only one predicted model with a near perfect C-score of 1.82.

To compare the structures generated by model fitting and by homology modeling, the generated models were analyzed based on a common reference structure, *H.*

Table 2. Conserved functional domains in STPP1. Identical residues were found for the models of *HsSTPP1-α* and *PmSTPP1-α* for these conserved functional domains. The color scheme used for the functional domains is like that used for the structures shown in Figures 3 and 4.

Conserved Functional Domain	Residues in *HsSTPP1-α*	Residues in *PmSTPP1-α*
Conserved active site/ Metal coordination site	D64; H66;D92; N124; H173; H248 *Identical Residues	D64; H66;D92; N124; H173; H248 *Identical Residues
Target protein binding domain	D71; R74 *Identical Residues	D71; R74 *Identical Residues
RVxF binding pockey	K168; I169; F170; D242; L243; F257; R261; L289; M290; C291; F293 *Identical Residues	K168; I169; F170; D242; L243; F257; R261; L289; M290; C291; F293 *Identical Residues
Molecular toxin interaction sites	R96; E126; S129; I130; Y134; W206; D220; R221; G222; V223; Y272 *Identical Residues	R96; E126; S129; I130; Y134; W206; D220; R221; G222; V223; Y272 *Identical Residues
Protruding Loop	N271; D277 *Identical Residues	N271; D277 *Identical Residues

Table 3. Neighboring residues for the original and variant residues between *HsSTPP1-α* and *PmSTPP1-α*. Most of the observed residue contacts are similar. But potential residue interactions are observed to be altered due to some variations. These include changes in potential residue contacts as seen with the change from Leu241 to Phe241. The modelling of the C-terminal domain also predicts interactions with other residues in the *PmSTPP1-α* protein. Investigations on the relevance of these contacts for affecting the protein function will be included in future studies.

***HsSTPP1-α* (3e/7a/ *PmSTPP1-α* (model)**	***H. sapiens* STPP1 Residue Contacts**	***P. monodon* STPP1 Residue Contacts**
Leu9/ Ile9	7, 8, 9, 10, 11, 12, 13, 14, 38, 109, 112, 113 *Identical contacts	7, 8, 9, 10, 11, 12, 13, 14, 38, 109, 112, 113 *Identical contacts
Gly14/ Ala14	9-18 *Identical Contacts	9-18 *Identical Contacts
Glm20/Arg20	17-22, 73, 77, 81	16, 17-22, 70, 73, 77, 81 *Additional Contacts: Leu16; Tyr70
Ile161/ Val161	51, 121, 122, 159-163, 171-173, 186, 201, 204, 205 *Identical contacts	51, 121, 122, 159-163, 171-173, 186, 201, 204, 205 *Identical contacts but the following residues were predicted to form sheets in the model. Sheets: 51, 121, 122;160-162
Val213/ Thr213	210-215; **216**; 217; 226-229 *Additional contact: Trp216	210-215; 217; 226-229
Glm214/ Met214	212-216; 228; **229**; 230; **231** *Additional contact: Ala229, Val231	212-216; 228; 230
Leu241/ Phe241	164, 167, 168, **169**; 170; 183; 235; 236; 239-243 *Additional contact: Ile169	164, 167-168; 170; 183; 235; 236; **237**; 239-243 *Additional contact: His237
C-terminal domain (Asp300– Lys329)	N/A	66, 68, 71, 96-98, 125, 133, 134, 138, 142, 250, 267, 270-277, 298, 299

sapiens STPP1-α, (PDBID 3e7a). Upon fitting the top ITASSER predicted model for the full-length *PmSTPP1-α* sequence unto 3e7a, similar structures were observed between Leu7-Ala299 (Fig 4A). Identical residues were observed for both proteins in the predicted functional domains (i.e. metal coordination sites, target protein binding site, RVxF domain, Toxin binding domains, and the protruding loop; Table 2; Fig. 4 B). The variant residue identities in *PmSTPP1-α* compared to *HsSTPP1-α*, occurred outside these conserved functional domains (Fig 4 C).

While the variant residues in *PmSTPP1-α* were found outside the conserved functional domains, it was still important to note potential changes in residue interaction that may be brought about by the changes. Table 3 lists the residues within a 5 distance from the original and variant residues in *HsSTPP1-α* and *PmSTPP1-α*, respectively. This proximity allows the formation of contacts between residues. Most of the variance was conservative (e.g. L9I; I161V, etc.) but several changes in residue contacts were observed (Table 3). One change, from Leu241 to Phe241 results in the loss of an interaction with Ile169, and an additional contact with His 237. While both the Leu241-Ile169 and Phe241-His237 involve non-polar interactions, the potential creation of Pi bonds through the aromatic rings in the latter pair may provide a stronger bond. It is interesting to note that the Leu241 to Phe241 variation has not been observed in the other related STPP1 forms (Table 2).

It must also be noted that the modeled structure included predictions for the conformation of the flanking N and C-terminal sections. The N-terminal sequences (Met1-Lys6); and the C-terminal section from Asp300-Lys329 did not have corresponding structures in the *H. sapiens* STPP1-α structure (Fig 4 D). The C-terminal

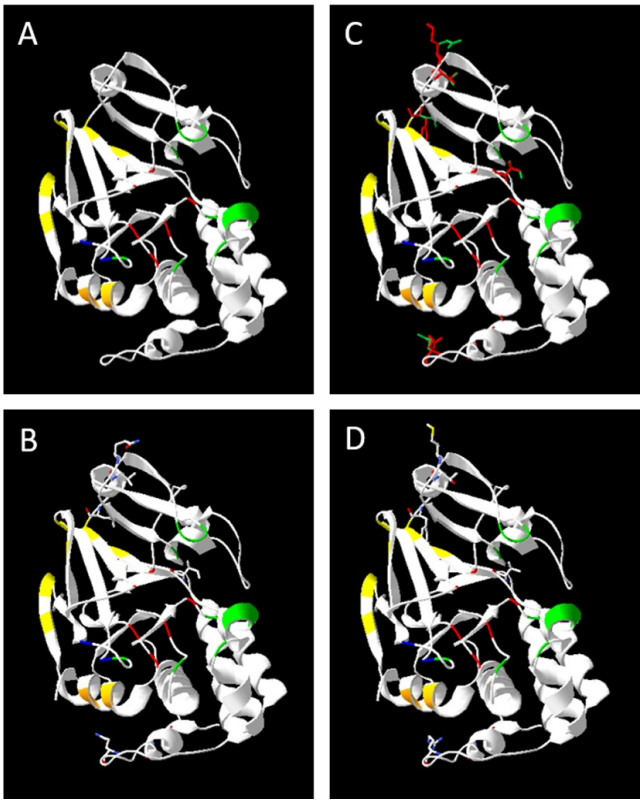


Fig. 3. Variant locations between STPP1-α in *H. sapiens* and *P. monodon* in Fit Model structures. The residues observed to vary between the two source species do not occur in the conserved functional domains for STPP1-*α*. (A) Conserved functional domains in STPP1-*α*, colored as follows in the ribbon structure: Metal coordination site (Red); Target protein binding site (Orange); RVxF domain (Yellow) and Toxin Binding Sites (Green). (B) Residues in *H. sapiens* STPP1-*α* at the variant locations rendered as line models and colored in CPK format. (C) Residues in the predicted *P. monodon* STPP1α at the variant locations rendered as line models and colored in CPK format. (D) Variant residue line models in STPP1-α from *H. sapiens* (green) and *P. monodon* (red).

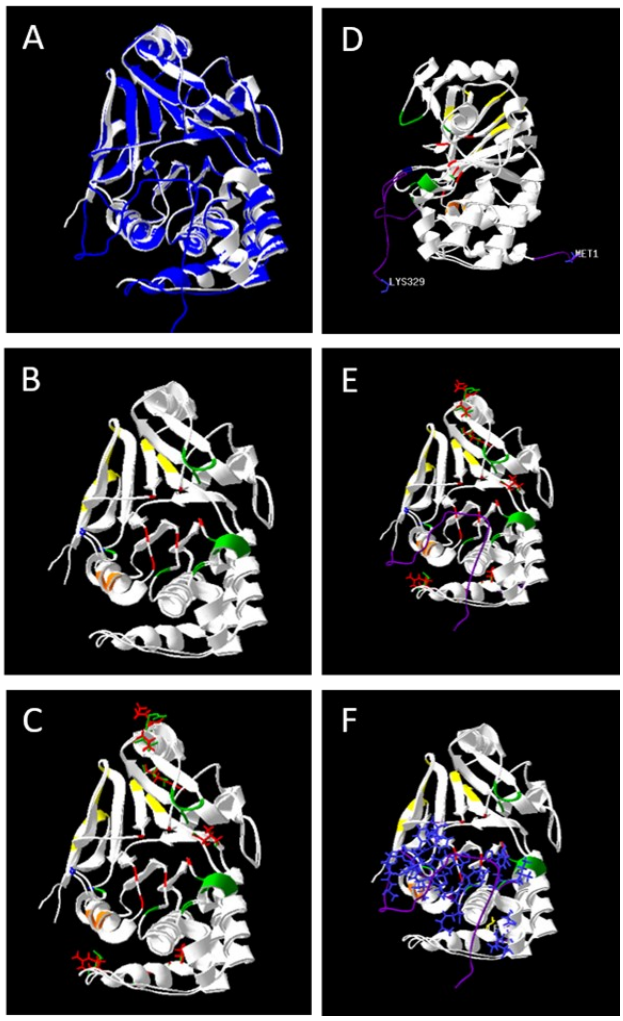


Fig. 4. Homology Modeled Structure of *PmSTPP1-α.* (A) Homology modelling predicts a similar structure for *PmSTPP1-α* (blue) as that of *HsSTPP1-α* (white). (B) Identical residues and locations were observed for the functional domains (i.e. Metal coordination site (Red); Target protein binding site (Orange); RVxF domain (Yellow); Toxin binding sites (Green); Protruding Loop (Blue)). (C) The variant residues for *PmSTPP1-α* occur outside these conserved functional domains. Original residues in the *HsSTPP1-α* are shown in green; and variant residues in *PmSTPP1-α* are shown in red. (D) The generated homology models included N and C-terminal domains (violet) which were not observed for the reference *HsSTPP1-α* structure. (E, F) The modelled C-terminal domain is predicted to interact with other regions of the protein, including the Metal Coordination Site, the Target Protein Binding Domain, and the Toxin binding sites.

section is in fact predicted to interact with other sections of the protein, including areas that are associated with toxin binding, the metal coordination site, and the expected target protein binding domain (Fig 4 E, F). How these predicted interactions translate to effects on the expected function of *PmSTPP1-α* would be the subject of further studies.

Functional Domain Predictions

In addition to generating structure models for the submitted sequences, ITASSER also provides predictions on potential function based on the presence of characteristic domains. The submitted sequence for *PmSTPP1-α* was predicted to have a conserved metal coordination site involving D64 and H248. These corresponds with the predicted metal coordination sites predicted based on the polypeptide sequence (Table 2). The sequence was also associated with structures of enzymes with class designations, EC 3.1.3.16, which corresponds to Protein Phosphatases, and S/T specific Protein Phosphatases. Both results suggest that the function of the predicted *PmSTPP1-α* protein matches the expected S/T phosphatase function of its related proteins.

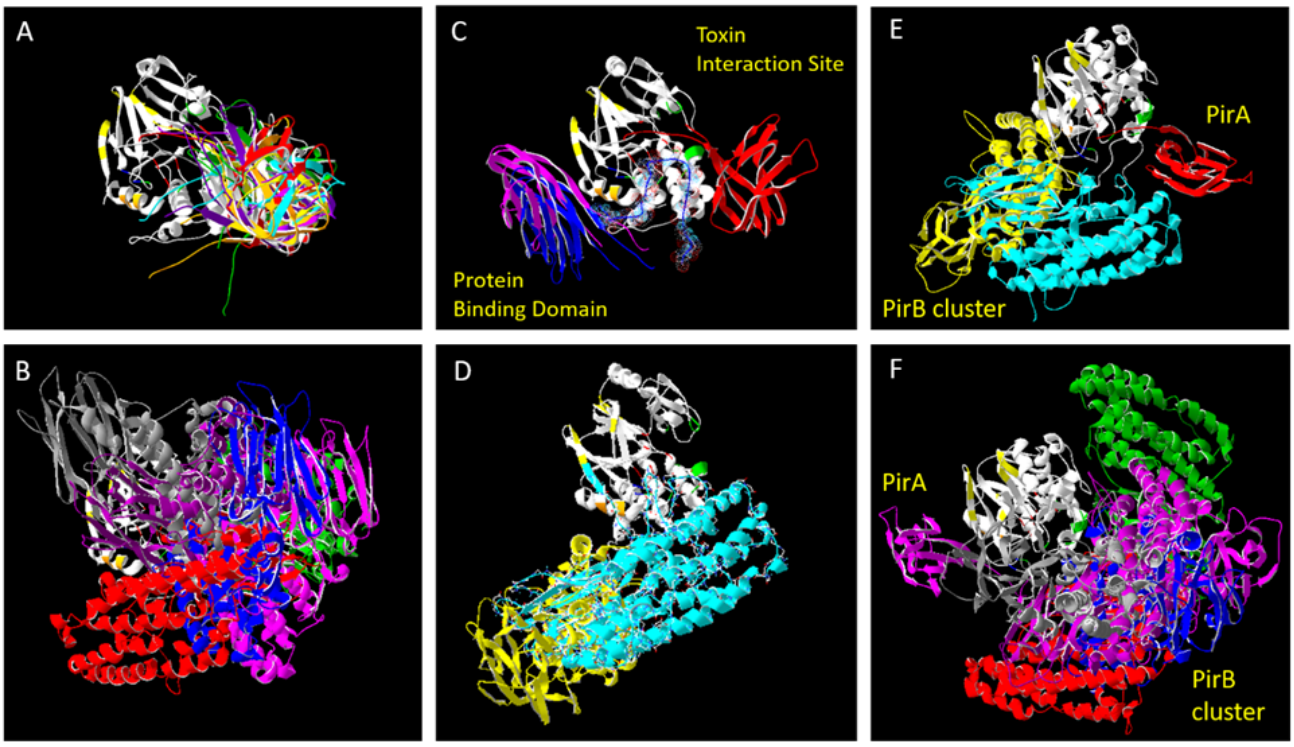


Fig. 5. Predicted Dock positions for PirA and PirB. The functional domains for *PmSTPP1-α* are colored similarly as in Figure 3 (i.e. Green: Molecular Interaction Sites; Orange: Target Protein Binding domain; Yellow: PVxF region, etc.). (A) Eight of ten docks based on balanced factors were predicted to bind the molecular toxin binding region for PirA. (B) Six of ten docks based on balanced factors were predicted to bind the molecular toxin binding region for PirB. (C) Two docks (colored blue and magenta) were predicted to bind with the Target Protein Binding domain for PirA. (D) Two docks (colored cyan and yellow) were predicted to bind with the Target Protein Binding domain for PirB. (E) The predicted position for PirA binding to the molecular toxin interaction site allows PirB binding with the target binding region without stearic clashes. (F) The converse is true for some docks where PirA is bound to the protein binding region and PirB is bound to the molecular toxin interaction site. (A-D) Structures are shown in side-view perspective; (E,F) Structures are shown from the top-view perspective. (C) The position of the modelled C-terminal domain is rendered as a blue ribbon with visible VDW spheres. This is seen to be placed between the docking positions for the molecular toxin interaction sites, and the target protein binding domain. The molecular structures used in this figure were rendered using the DeepView molecular viewer (Guex et al. 1997).

In Silico Protein Docking Predictions

The predicted involvement of *PmSTPP1-α* with *P. monodon* immune response was assessed based on predicted associations with protein targets from known Shrimp pathogens. Specifically, protein docking experiments were done with three proteins (PirA, PirB, and VP24). Both PirA and PirB are toxins from *V. parahaemolyticus* (Lee et al. 2015), while VP24 is a major envelope protein of White Spot Syndrome Virus (Sun et al. 2016).

Predicted Dock Models: PirA

Eight out of the top 10 models generated based on "balanced" factor contribution placed the PirA structure in close association with the expected molecular toxin interaction sites of *PmSTPP1-α* (Table 2; Fig. 5 A). The remaining two models in the top 10 showed predictions for association with the target protein binding domain of *PmSTPP1-α* (Table 2). These observed dock positions occur at opposite flanking sides of the *PmSTPP1-α* structure (Fig. 5 C).

The position of the toxin interaction site docked models was observed to be similar to the top-docked structures in two other groups, those ranked based on electrostatics and hydrophobics. These results suggest the relevant involvement of these types of interactions in the predicted association of *PmSTPP1-α* and PirA.

Predicted Dock Models: PirB

Six out of the top 10 models generated based on "balanced" factor contribution placed the PirA structure in close association with the expected molecular toxin interaction sites of *PmSTPP1-α* (Table 2; Fig. 5 B). Two of the remaining models in the top 10 showed predictions for association with the target protein binding domain of *PmSTPP1-α* (Table 2; Fig. 5 D). The remaining two models were predicted to bind different positions in *PmSTPP1-α*.

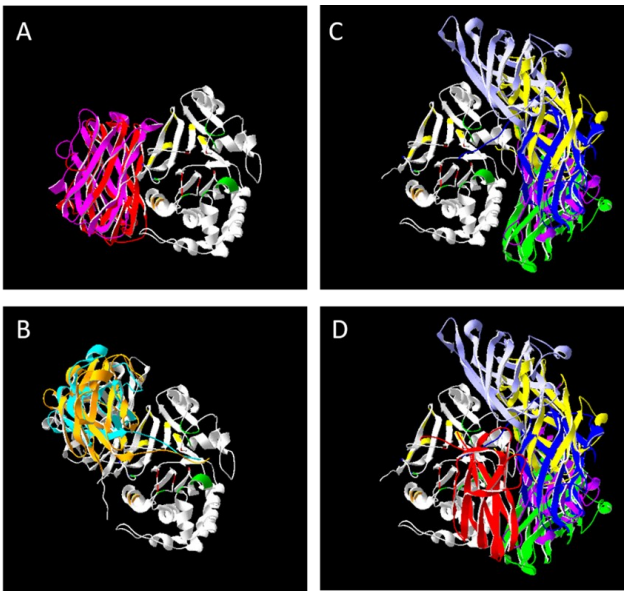


Fig. 6. Predicted Dock positions for VP24. The functional domains for *PmSTPP1-α* are colored similarly as in Figure 3 (i.e. Green: Molecular Interaction Sites; Orange: Target Protein Binding domain; Yellow: PVxF region, etc.). (A) Two of ten docks based on balanced factors (red and magenta) were predicted to bind the Target Protein Binding domain of *PmSTPP1-α*. The red ribbon represents the position of the top ranked dock. (B) Two of ten docks based on balanced factors were predicted to bind the RVxF domain. (C) Six of ten docks based on balanced factors were predicted to bind near the toxin binding site of *PmSTPP1-α*. (D) It must be noted, however, that the binding positions for these six docks were not observed to coincide with the predicted toxin interaction site bound by PirA (foreground; red ribbon). The molecular structures used in this figure were rendered using the DeepView molecular viewer (Guex et al. 1997).

The position of the target protein binding domain docked models was observed to be similar to the top-docked structure for the models ranked based on hydrophobics. While not matching the top dock, the position predicted for the 6 models bound to the toxin interaction site was similar as the 2nd ranked model based on hydrophobics. These results suggest the relevant involvement of hydrophobic interactions in the predicted association of *PmSTPP1-α* and PirB through either the toxin interaction sites or the target protein binding surface.

Predicted Dock Models: VP24

Two out of the top 10 models generated based on "balanced" factor contribution placed the VP24 structure in close association with the expected target protein binding domain of *PmSTPP1-α* (Table 2; Fig. 6 A). Two of the remaining models were also near this site but were shifted higher towards the PVxF domain (Table 2; Fig. 6 B). The remaining six models were predicted to bind different positions in *PmSTPP1-α* near the toxin interaction site, but at distinctly different positions as that predicted for PirA binding. (Fig. 6 C, D)

The position of the docked models near the PVxF domain was observed to be similar to the top-docked structure for the models ranked based on electrostatics. The position predicted for the 6 models near the toxin-binding site was similar as the top ranked model based on hydrophobics. These results suggest the relevant involvement of both electrostatic and hydrophobic interactions in guiding the predicted associations of *PmSTPP1-α* and VP24 through contacts near the toxin interaction sites or the target protein binding surface.

Predicted Dock Models: PirAB vs VP24

The predicted involvement of *PmSTPP1-α* in the immune response for *P. monodon* was investigated based on its potential interactions with two types of viral pathogens. These were a membrane protein from WSSV (VP24) and components of a toxin protein (PirAB) from *V. parahaemolyticus*. The observed interactions with the two target types had similarities and differences. Both target types were observed to have predicted interactions with the documented target protein binding surface of *PmSTPP1-α*. Associations were also observed for VP24 and both PirAB components (PirA and PirB) near the toxin interaction sites of *PmSTPP1-α*. However, the predicted docking regions near this site for VP24 were distinctly separate from those favored by PirA (Fig. 6 D). This suggests specificity for binding target types bound by the toxin interaction sites. Our current results show selective binding for toxin protein components over other targets (VP24 membrane protein) in the predicted toxin-binding site for *PmSTPP1-α*. The functional relevance of this selective function must be further analyzed to understand the role of *PmSTPP1-α* in the immune response.

Predicted Dock Models: PirA and PirB

It is interesting to note that some of the predicted docks for PirA binding to the *PmSTPP1-α* toxin binding site allows the binding of PirB to the target protein binding region (Fig. 5 E, F). As both PirA and PirB are components of the PirAB toxin from *V. parahaemolyticus* these observed binding combinations predict the possibility for *PmSTPP1-α* to associate with the PirAB toxin through multiple surfaces, both as single-surface interactions (i.e. toxin binding site only; target protein binding site only); combined action on two different surfaces; or as a linker protein binding two PirAB toxin molecules. The presence of the modelled C-terminal domain between the toxin interaction site and the protein

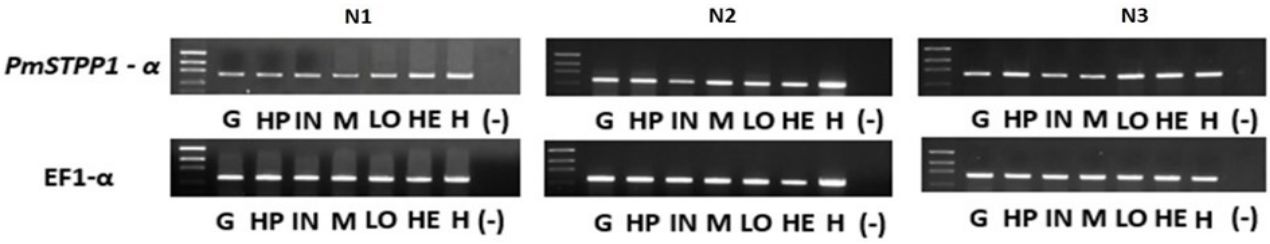


Fig. 7. Basal expression at organ level of *PmSTPP1-α* gene relative with the EF1-α, as the internal control. The seven (7) organs from three (3) WSSV-negative *P. monodon* were analyzed (1st lane: low mass DNA ladder; 2nd to 8th lanes were as follows: *G*, gills; *HP*, hepatopancreas; *IN*, intestine; *M*, muscle; *LO*, lymphoid organ; *HE*, heart; *H*, hemolymph; -, negative control).

binding domain (Fig. 5 C) may influence the relative positioning of the PirA and PirB components. The significance of these different interaction possibilities on the potential role of *PmSTPP1-α* for the immune response will be the subject of future studies.

PmSTPP1-α Gene Constitutive Expression

PmSTPP1-α gene was expressed ubiquitously in tissues and organs of WSSV-negative *P. monodon* (Fig. 7), suggesting a probable involvement of the gene in the overall physiology of the species. In general, STPP1-α is persistently distributed in tissues, and enriched in the brain and heart (cardiomyocytes) (Aoyama et al. 2011; Peters et al. 2009; Vafiadaki et al. 2013). Specifically, in humans, PP1 was found to be highly involved in a variety of cellular processes such as protein synthesis, apoptosis, meiosis and cell division, cytoskeletal reorganization, regulation of membrane receptors and channels, and metabolism (Bollen et al. 2010). Korrodi-Gregorio et al. (2014) described the distribution of PP1-α in tissues such as the brain (ubiquitous but enriched in striatum and hippocampus) and testis (cytoplasm of Leydig and peritubular cells, spermatogonia, and preleptotene spermatocytes); in polarized cells such as spermatozoa and neurons (dendritic spines, perikaryal cytoplasm and nucleus); in cellular processes such as interphase (cytoplasm especially in centrosome, and nucleus particularly the nuclear matrix and nucleoplasm) and mitosis (also in centrosome which is throughout the mitosis stage, another is in kinetochores which is during the metaphase, and the midbody which is in telophase). Such varied PP1-α gene function could also be found in *P. monodon*, since *PmSTPP1-α* is constitutively expressed in gills, hepatopancreas, intestine, muscle, lymphoid organ, heart and hemolymph.

CONCLUSION

To date, this study is the first cloning report of alpha-subunit in STPP1 present in *P. monodon* encoding the complete C-terminal end which is found to be very similar to human PP1Cα (Kim et al. 2015). In silico predictions of *PmSTPP1-α* structure were in observed to share a similar form as human STPP1a, especially in the conserved functional domains. Some residue variations were still observed outside these areas, and their significance for *PmSTPP1-α* function must be further studied. *PmSTPP1-α* was found to be ubiquitously and highly expressed in organs of WSSV-negative *P. monodon* and appears to be very critical for the maintenance of its general physiology. With this, it may also specifically be involved in shrimp-virus interaction and should be studied further on its probable role in shrimp immunity. Current molecular docking predictions for *PmSTPP1-α* against three proteins from known *P. monodon* pathogens suggest specific functional interactions with the target protein binding domain and the molecular toxin interaction sites. Further investigations on the interactions of this protein will help validate its predicted involvement in the *P. monodon* immune response.

REFERENCES CITED

AOYAMA H, IKEDA Y, MIYAZAKI Y. et al. 2011. Isoform-specific roles of protein phosphatase 1 catalytic subunits in sarcoplasmic reticulum-mediated Ca (21) cycling. Cardiovascular Research, 89:79–88. DOI:10.1093/cvr/cvq252

BERMAN H, WESTBROOK J, FENG Z, GILLILAND G, BHAT T, WEISSIG H, SHINDYALOV I, BOURNE P. 2000. The protein data bank. Nucleic Acids Research, 28:235-242.

BOLLEN M, PETI W, RAGUSA MJ, BEULLENS M. 2010. The extended PP1 toolkit: designed to create specificity. Trends in Biochemical Science, 35:450–458. DOI:10.1016/j.tibs.2010.03.002

BRAUTIGAN D. 2013. Protein Ser/Thr phosphatases—the ugly ducklings of cell signaling. The Federation of European Biochemical Societies Journal, 280:324-345. DOI: 10.1111/j.1742-4658.2012.08609.x

CIEŚLA J, FRACZYK T, RODE W. 2011. Phosphorylation of basic amino acid residues in proteins: important but easily missed. Acta Biochimica Polonica, 58:137-148. PMID: 21623415.

DANCHECK B, RAGUSA M, ALLAIRE M, NAIRN A, PAGE R, PETI W. 2011. Molecular investigations of the structure and function of the protein phosphatase 1-spinophilin-inhibitor 2 heterotrimeric complex. Biochemistry, 50:1238–1246. DOI:10.1021/bi101774g

DANG LT, KOYAMA T, SHITARA A, KONDO H., AOKI, T., HIRONO, I., 2010. Involvement of WSSV shrimp homologs in WSSV infectivity in kuruma shrimp: *Marsupenaeus japonicus*. Antiviral Research, 88:217-226. DOI: 10.1016/j.antiviral.2010.08.017

DIEU B, MARKS H, SIEBENGA J, GOLDBACK R, ZUIDEMA D, DUONG T, VLAK J. 2004. Molecular epidemiology of white spot syndrome virus within Vietnam. Journal of General Virology, 85:3607-3618. PMID: 15557233.

FLEGEL T. 2006. Detection of major penaeid shrimp viruses in Asia, a historical perspective with emphasis on Thailand. Aquaculture, 258:1-33. DOI:10.1016/j.aquaculture.2006.05.013

GEUX N, PEITSCH M. 1997. Swiss-model and the Swiss-Pdb viewer: An environment for comparative protein modeling. Electrophoresis, 18:2714 - 2723.

HEROES E, LESAGE B, GORNEMANN J, BEULENNS M, MEERVELT L, BOLLEN M. 2013. The PP1 binding code: a molecular-lego strategy that governs specificity. The Federation of European Biochemical Societies Journal, 280:584-595. DOI:10.1111/j.1742-4658.2012.08547.x

HEROES E, RIP J, BEULLENS M, VAN MEERVELT L, DE GENDT S, BOLLEN, M. 2015. Metals in the active site of native protein phosphatase-1. Journal of Inorganic Biochememistry, 149:1–5. doi:10.1016/j.jinorgbio.2015.03.012

KELKER M, PAGE R, PETI W. 2009. Crystal structures of protein phosphatase-1 bound to Nodularin-R and Tautomycin: a novel scaffold for structure-based drug design of serine/threonine phosphatase inhibitors. Journal of Molecular Biology, Volume 385, Issue 1. http://doi.org/10.2210/pdb3E7A/pdb

KOLUPAEVA V. 2019. Serine-threonine protein phosphatases: Lost in translation. Biochimica et Biophysica Acta (BBA) – Molecular Cell Research, 1866(1):83-89. DOI: 10.1016/j.bbamcr.2018.08.006

KORRODI-GREGORIO L, ESTEVES SLC, FARDILHA M. 2014. Protein phosphatase 1 catalytic isoforms: specificity toward interacting proteins. Translational Research, 164(5):366-391. doi:10.1016/j.trsl.2014.07.001

KOZAKOV D, HALL D, XIA B, PORTER K, PADHORNY D, YUEH C, BEGLOV D, VAJDA S. 2017. The ClusPro web server for protein-protein docking. Nature Protocols, 12(2):255-278.

LEE C, CHEN I, YANG Y, KO T, HUANG Y, HUANG J, HUANG M, et al. 2015. The opportunistic marine pathogen Vibrio parahaemolyticus becomes virulent by acquiring a plasmid that expresses a deadly toxin. Proceedings of the National Academy of Sciences, 112 (34):10798-10803; DOI: 10.1073/pnas.1503129112

LI F, XIANG J. 2013. Signaling pathways regulating immune responses in shrimp. Fish & Shellfish Immunology, 34:973-980. DOI:10.1016/j.fsi.2012.08.023

LU L, KWANG J. 2004. Identification of a novel shrimp protein phosphatase and its association with latency - related ORF427 of white spot syndrome virus. FEBS Letters, 577:141–146. DOI:10.1016/j.febslet.2004.08.087

MA G, ZHOU R, SONG H, ZHU H, ZHOU Z, ZENG Y. 2015. Molecular mechanisms of serine/threonine protein phosphatase 1 (PP1ca-PP1r7) in spermatogenesis of *Toxocara canis*. Acta Tropical, 149:148-154. doi:10.1016/j.actatropica. 2015.05.026

MARALIT B, VENTOLERO M, MANINGAS M, AMAR E, SANTOS M. 2014. Substracted transcriptome profile of black tiger shrimp (*Penaeus monodon*) that survived WSSV challenge. Dataset Papers in Science, 2014:1–11. DOI: 10.1155/2014/807806

NEEDLEMAN S, WUNSCH C. 1970. A general method applicable to the search for similarities in the amino acid sequence of two proteins. Journal of Molecular Biology, 48(3):443-53. DOI: 10.1016/0022-2836(70) 90057-4. https://www.ebi.ac.uk/Tools/psa/emboss_needle/

PETERS M, BLETSCH M, CATAPANO R, ZHANG X, TULLY T, BOURTCHOULADZE R. 2009. RNA interference in hippocampus demonstrates opposing roles for CREB and PP1 alpha in contextual and temporal long term memory. Genes, Brain and Behavior, 8:320-329. DOI:10.1111/j.1601-183X.2009.00474.x

PETI W, NAIRN A, PAGE R. 2013. Structural basis for protein phosphatase 1 regulation and specificity. The Federation of European Biochemical Societies Journal, 280:596-611. DOI: 10.1111/j.1742-4658.2012.08509.x

RAGUSA MJ, DANCHECK B, CRITTON DA, NAIRN AC, PAGE R, PETI W. 2010. Spinophilin directs protein phosphatase 1 specificity by blocking substrate binding sites. Nature Structural and Molecular Biology, 17:459–464. DOI: 10.1038/nsmb.1786

ROY A, KUCUKURAL A, ZHANG Y. 2010. I-TASSER: a unified platform for automated protein structure and function prediction. Nature Protocols, 5:725-738. https://zhanggroup.org/I-TASSER/

SANTOS M, YASUIKE M, HIRONO I, AOKI T. 2006. The granulocyte colony-stimulating factors (CSF3s) of fish and chicken. Immunogenetics, 58(5-6):422-432. DOI: 10.1007/s00251-006-0106-5

SANTOS M, YASUIKE M, HIRONO I, AOKI T. 2007. A novel type-1 cytokine receptor from fish involved in the Janus kinase/Signal transducers and activators of transcription (Jak/STAT) signal pathway. Molecular Immunology, 44:3355-3363. DOI: 10.1016/j.molimm.2007.02.018

SHI Y. 2009. Serine/Threonine Phosphatases: mechanism through structure. Cell, 139. DOI:10.1016/j.cell.2009.10.006

SONG Y, LI C. 2014. Shrimp immune system – special focus on Penaeidin. Journal of Marine Science and Technology, 22(1):1-8. DOI: 10.6119/JMST-013-0813-1

SUN L, SU Y, ZHAO Y. et al. 2016. Crystal structure of major envelope protein VP24 from white spot syndrome virus. Scientific Reports 6, 32309. https://doi.org/10.1038/srep32309

TERRAK M, KERFF F, LANGSETMO K, TAO T, DOMINGUEZ R. 2004. Structural basis of protein phosphatase 1 regulation. Nature, 429:780-784. DOI:10.1038/nature02582

THE UNIPROT CONSORTIUM. 2021. UniProt: the universal protein knowledgebase in 2021. Nucleic Acids Research, 49:D1.

TAMURA K, STECHER G, PETERSON D, FILIPSKI A, KUMAR S. 2013. MEGA6: Molecular evolutionary genetics analysis version 6.0. Molecular Biology and Evolution, 30:2725-2729. DOI:10.1093/molbev/mst197

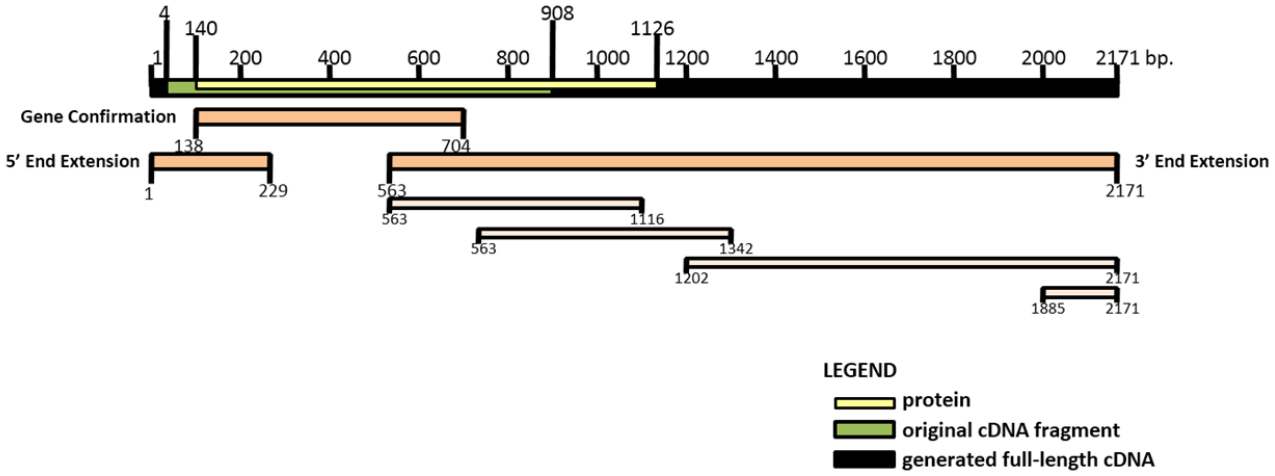
VAFIADAKI E, ARVANITIS DA, SANOUDOU D, KRANIAS EG. 2013. Identification of a protein phosphatase-1/phospholamban complex that is regulated by cAMP-dependent phosphorylation. PLoS ONE, 8(11).e80867. DOI:10.1371/journal.pone.0080867

VERBINNEN I, FERREIRA M, BOLLEN M. 2017. Biogenesis and activity regulation of protein phosphatase 1. Biochemical Society Transactions, 45:89-99. DOI: 10.1042/BST-2016-0154C_COR

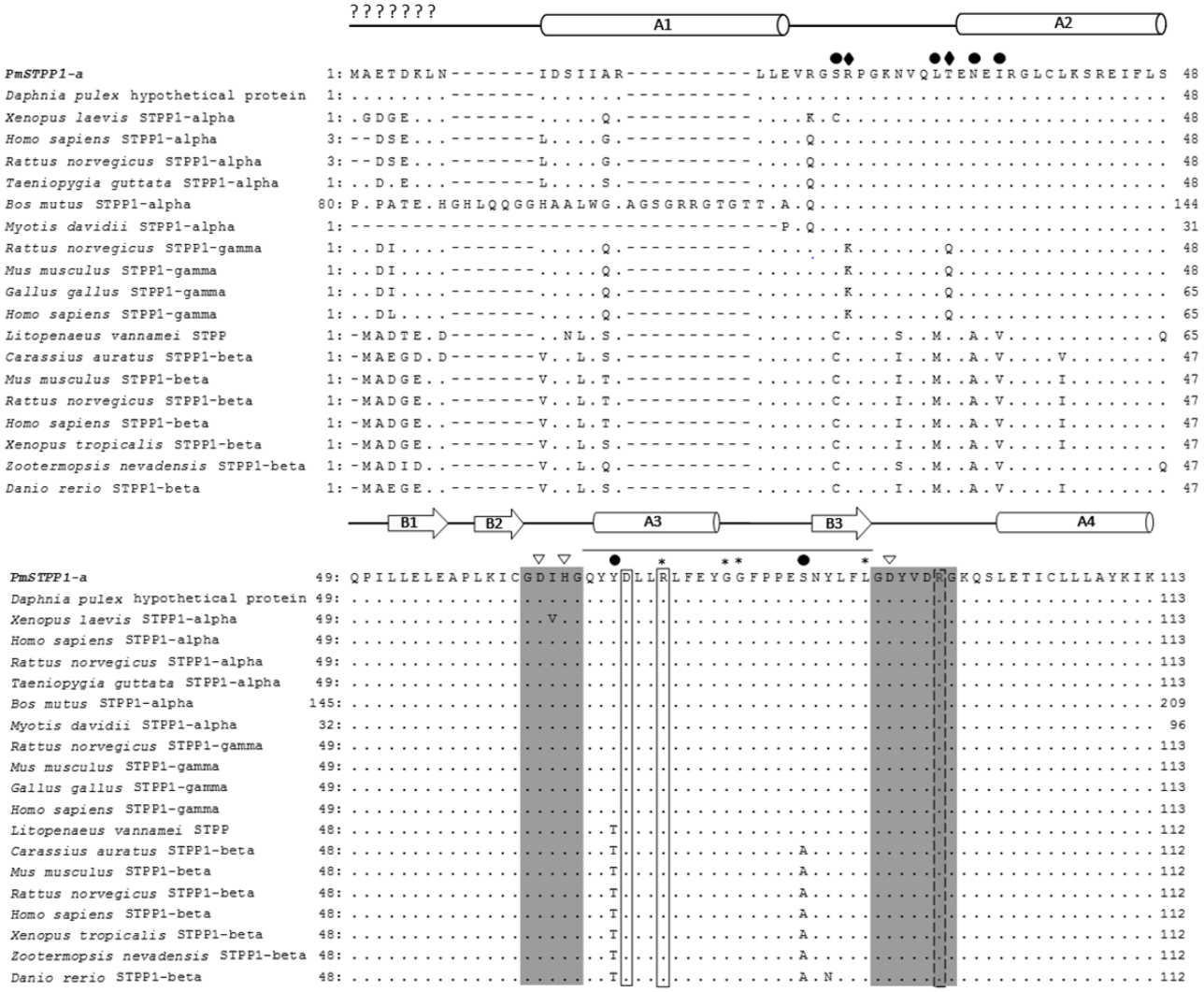
WANG H, KO T, WANG A, LO C. 2014. 3X0T:Crystal structure of PirA. Deposited into the PDB 10/22/2014. DOI: 10.2210/pdb3X0T/pdb

ZHAO C, WANG Y, FU M, YANG K, QIU L. 2016. Molecular cloning, expression and functional analysis of three subunits of protein phosphatase 2A (PP2A) from black tiger shrimps (*Penaeus monodon*). Comparative Biochemistry and Physiology, Part B, 204:77-89. DOI:10.1016/j.cbpb.2016.11.011

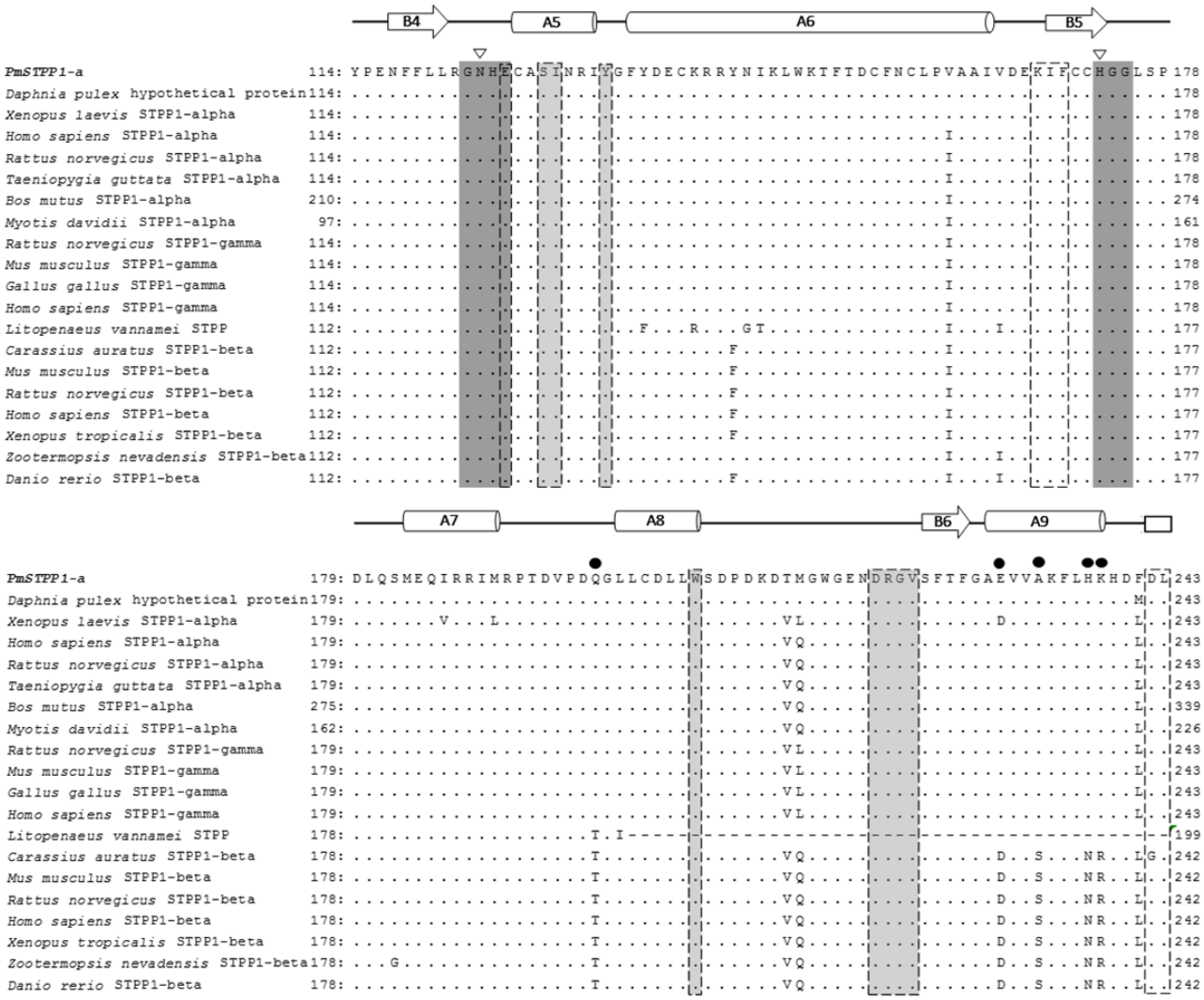
Supplemental Figure 1. Illustration of the derived full-length *PmSTPP1-α* gene sequence. The position of the original cDNA fragment and the gene product were indicated in the generated full-length gene.



Supplemental Figure 2. Complete Multiple Alignment of *PmSTPP1-α* with reference PP1 amino acid sequences. Secondary structure elements were represented as cylinders for the *α*-helices and as arrows for the β-strands as indicated above the alignment. The unordered N and C-termini were marked with question mark (?). Amino acids similar with *PmSTPP1-α* sequence were represented by dots (…); gaps on the aa positions were designated by dashes (---); conserved aa sequence motifs and other conserved sequences (unique for PPP family) were denoted with filled box; conserved active site or metal-coordinating residues were marked with inverted unfilled triangle; residues responsible in binding targeting protein or substrate that bind in PP1C (a regulatory protein bound, spinophilin, also binds to these residues) were denoted with unfilled box, residues that form the RVxF binding pocket were denoted with broken-line unfilled box, residues critical for associating molecular toxins with STPP1-α [within 5 A° of the bound toxin in each of the molecular toxin structures determined as described in the paper of Peti et al. (2013)] were denoted with broken-line filled box, aa residues that are distinct feature to STPP1 together with PP2A and PP2B subfamilies were marked with asterisk (*); position of the twenty three (23) aa residues in between of the first and second conserved motifs (which is the same position and number of residues with the human STPP1-α) were marked with a line on top; STPP1-α with similar aa sequences STPP1-γ and different with STPP1-β were marked with filled circle; STPP1-α with similar aa sequences STPP1-β and different with STPP1-γ were marked with filled diamond; STPP1-α with unique aa sequences from STPP1- γ and STPP1- β were marked with unfilled diamond; aa residues that separated the acidic and C-terminal groove [noted as a protruding loop which is from N271 – D277 in the paper of Terrak et al. (2004)] were marked with unfilled circle.



Supplemental Figure 2. Continuation.



Supplemental Figure 2. Continuation.