

***In silico* molecular interaction studies of shrimp
antiviral protein, PmAV with WSSV RING
finger domain**

A Major Project dissertation submitted

In partial fulfilment of the requirement for the degree of

Master of Technology

In

Bioinformatics

Submitted by

Garima Soni (2K11/Bio/05)

Delhi Technological University, Delhi, India

Under the supervision of

Dr. Navneeta Bharadvaja



Department of Biotechnology
Delhi Technological University
(Formerly Delhi College of Engineering)
ShahbadDaulatpur, Main Bawana Road,
Delhi-110042, INDIA



CERTIFICATE

This is to certify that the M. Tech. dissertation entitled “***In silico* molecular interaction studies of shrimp antiviral protein, PmAV with WSSV RING finger domain**”, submitted by **GARIMA SONI (2K11/BIO/05)** in partial fulfilment of the requirement for the award of the degree of Master of Engineering, Delhi Technological University (Formerly Delhi College of Engineering, University of Delhi), is an authentic record of the candidate’s own work carried out by her under my guidance.

The information and data enclosed in this dissertation is original and has not been submitted elsewhere for honouring of any other degree.

Date:

Dr Navneeta Bharadvaja

(Project Mentor)

Department of Bio-Technology

Delhi Technological University

(Formerly Delhi College of Engineering, University of Delhi)

DECLARATION

I **Ms. Garima Soni (2K11/BIO/05)**, student of M. Tech. (Bioinformatics), Delhi Technological University have completed the project titled “***In silico* molecular interaction studies of shrimp antiviral protein, PmAV with WSSV RING finger domain**” for the award of Degree of M. Tech. (Bioinformatics), for academic session 2011-13. The information given in this project is true to the best of my knowledge.

Garima Soni

2K11/BIO/05

ACKNOWLEDGEMENT

It gives me immense pleasure to thank all those people who have been instrumental in the completion of my project.

I express my deep sense of thankfulness to Dr Navneeta Bharadvaja whose enthusiastic zeal boosted me for the successful completion of my work.

In the first place I sincerely thank Dr. Anil Rai, Head, Centre for Agricultural Bioinformatics, ICAR-IASRI, for allowing me to do this research work.

I would like to express my heartfelt gratitude to my supervisor Dr. MNV Prasad Gajula (Scientist-Bioinformatics, DST Ramanujan Fellow) for allowing me to work under his guidance and providing me with all the facilities during my project work.

Finally, yet importantly, I would like to express my heartfelt thanks to my beloved parents, for their blessings and my friends, for their help and wishes for the successful completion of this project.

Garima Soni

2K11/BIO/05

CONTENTS

TOPIC	PAGE NO
<i>LIST OF FIGURES</i>	1
<i>LIST OF TABLES</i>	3
<i>LIST OF ABBREVIATIONS</i>	4
1.ABSTRACT	6
2.INTRODUCTION	7
3.REVIEW OF LITERATURE	9
3.1 White spot syndrome Baculovirus.	9
3.1.1 White spot syndrome virus in Shrimps.	9
3.1.2 Genetic analysis of WSSV.	11
3.1.3 WSSV Pandemic.	13
3.2 Shrimp diseases	15
3.2.1 Defense system of Insects.	15
3.2.2 Current diagnostic methods.	16
3.3 Economic importance	18
3.4 Ring Domain of WSSV	20
3.5 PmAV protein.	20
3.5.1 C-type lectins	20
3.5.2 Carbohydrate-Mediated Mechanisms	24
3.6 Techniques used	25
3.6.1 Protein 3D structure prediction	25
3.6.2 Protein-Protein Interactions	26

4.	METHODOLOGY	30
4.1	Molecular modeling of the proteins	30
4.2	Minimization and validation.	30
4.3	Molecular dynamics simulations of PmAV and RING DOMAIN	30
4.4	Protein-Protein interaction study by Docking	31
4.5	Molecular dynamics simulation of PmAV-RING DOMAIN complex.	32
5.	RESULTS	33
6.	DISCUSSION	43
7.	CONCLUSION AND FUTURE PERSPECTIVE	45
8.	REFERENCES	46
9.	APPENDIX	51

LIST OF FIGURES

Figures	Page No.
Figure 1: Schematic representation of the morphology of the WSSV viral particle.	9
Figure 2 : A model of the apoptotic interactions between a shrimp host cell and WSSV.	10
Figure 3 : The target tissues of WSSV in shrimp..	11
Figure 4: Cladogram based on genetic distances between aa sequences from DNA polymerase of wssv and other large DNA viruses.	12
Figure 5: Schematics map of ds DNA-WSSV strain CN genome showing the genomic organization.s	13
Figure 6: The white spot virus pandemic (Years of first occurrence by location).	14
Figure 7: Asian pandemic of WSSV.	14
Figure 8: WSSV copy number/mg DNA in shrimp with acute phase WSSV infection.	15
Figure 9: Schematic representation of the in vivo metabolic changes in shrimp hemocytes at the viral genome replication.	16
Figure 10: Source of information and available methods for diagnosis.	17
Figure 11: Penaeus monodon Hatchery Practices in india.s	23
Figure 12: Flowchart of I-TASSER method for protein structure prediction.s	26
Figure 13: Example of protein-protein interaction by docking analysis.	27
Figure 14: Creating water environment for molecular dynamics	28

simulation.

Figure 15: Flowchart for setting up an MD run.	29
Figure 16: Modeled structure of (a) PMAV (b) RING domain of wsv199.	34
Figure17. The Ramachandran plots for the PmAV and RING domain of wsv199,	35
Figure 18: RMSD graph of PmAV.	36
Figure 19 : RMSD graph of Ring Domain.	36
Figure 20: Radius of gyration graph of PmAV.	37
Figure 21: Radius of gyration graph of RING DOMAIN.	37
Fig 22:..Docking interaction between PmAV (green) and RING domain of wsv199 (red).	38
Figure 23: RMSD graph of PmAV-ring domain complex.	39
Figure 24: Radius of gyration graph of docked comple.	40
Figure 25: Vander Waals energy graph of complex.	40
Figure 26: Electrostatic energy graph of complex.	41
Figure 27: Interaction between two chains.	42

LIST OF TABLES

Tables	Page No.
Table 1: Methods available for diagnosis of shrimp diseases.	17
Table 2. Overview of structural and functional relationships of subfamilies of C-type lectins involved in pathogen sensing.	21
Table 3: Shrimp Farming Potential, Usage and Production in India	23
Table 4: List of the selected templates.	34
Table 5: FireDock Results.	38

LIST OF ABBREVIATIONS

WSD	White spot disease
WSSV	White spot syndrome virus
DD	Differential display
CRD	Carbohydrate recognition domain
PRRs	Pattern recognition receptors
WSBV	White spot syndrome baculovirus
SEMBV	Systemic ectodermal and mesodermal baculovirus
RFPs	RING Finger Proteins
CTLD	C-type lectin-like domain
CLRs	C-type lectin-like receptors
DCs	Dendritic cells
NCBI	National Center for Biotechnology Information
PDB	Protein Data Bank
BLAST	Basic Local Alignment Search Tool
C-score	Confidence score
RMSD	Root-mean square deviation

Rg	Radius of gyration
pI	Isoelectric point
GRAVY	Grand Average hydropathicity
II	The instability index
MD	Molecular Dynamics

***In silico* molecular interaction studies of shrimp antiviral protein, PmAV with WSSV RING finger domain**

Garima Soni

Delhi Technological University, Delhi, India

1. ABSTRACT

Background: White spot disease (WSD) is an overwhelming syndrome of shrimp *Penaeus monodon* in which receptor protein of *P. monodon* interacts with viral envelope protein, and causes commencement of the disease. PmAV is the first and only anti-viral protein found in shrimp. It was isolated from a virus unaffected shrimp, *P. monodon* and identified to be up-regulated at the time of viral infection.

Methods: In the present study, we applied molecular modeling and molecular docking to determine the interaction pattern of amino acid residues between PmAV protein and WSSV RING finger domain.

Results and conclusion: Our result showed that PmAV interacts with WSSV RING finger domain and prevents their activity thus inhibiting the establishment of viral infection. Further studies like molecular dynamics simulation of the complex might be applied to open new possibilities for preventing WSD. The quantifiable calculations offer an opportunity for experimental analysis in future as well as provide with an upfront evidence to understand cellular mechanisms underlying the syndrome.

Keywords: Docking, Molecular dynamics Simulation, Shrimp, White spot disease.

2. INTRODUCTION

Black tiger shrimp, *Penaeus monodon* is one of the most widely cultured shrimp species in the world. However, the major setback to shrimp farming industry worldwide is white spot disease (WSD) caused by white spot syndrome virus (WSSV). Several efforts have been made in preventing and controlling the disease in recent years (Lightner *et al.*, 1998; Lo *et al.*, 1996; Van Hulten *et al.*, 2001).

However, development of an effective therapy against this deadly disease has been a great challenge due to lack of specific immune system in shrimp (Yang *et al.*, 2001). The shrimp defense system is believed to rely largely on innate immunity, which is little understood. In this context, it is important to understand the host-virus interaction at molecular level to develop effective therapies. Several authors have reported the effect of WSSV on host gene expression profiles to determine their role in pathogenesis (Wang *et al.*, 2005; He *et al.*, 2006).

PmAV is the first and the only antiviral protein identified in shrimp. It was cloned from a virus-resistant shrimp, *P. monodon* by differential display (DD) and found to be up-regulated in response to viral infection (Luo *et al.*, 2003). *PmAV* is a Calcium-dependent (C-type) lectin containing a single characteristic carbohydrate recognition domain (CRD). C-type lectins contain a prototypic lectin fold, consisting of two antiparallel β -strands and two α -helices. C-type lectins are regarded as primary candidates for pattern recognition receptors (PRRs) in innate immunity and play an important role in the clearance of pathogens (Drickamer *et al.*, 1993; Hoffmann *et al.*, 1999; Yu *et al.*, 2004; Cerenius *et al.*, 2004;).

The PmAV protein was found to be located mainly in the cytoplasm and showed no interaction with intact WSSV particles (Luo *et al.*, 2003). It implies that the antiviral mechanism of PmAV protein is not by inhibiting the attachment of virus to target host cell. Instead it may counteract the virus molecular strategies that establish the infection in a cell. It is implicit that many viruses including WSSV have evolved the ability to utilize the host protease machinery to direct cellular protein degradation for their survival and

replication (Wang *et al.*, 2005;). Four proteins of WSSV, namely WSSV199, WSSV222, WSSV249 and WSSV403 contain RING-H2 domain, which is involved in specific ubiquitination events by acting as the E3 ubiquitin protein ligase (Wang *et al.*, 2005;).

Therefore, in this study we attempted *in silico* studies to explore the possible interaction between shrimp antiviral protein, *PmAV* and RING finger domain proteins of WSSV. Here we present 3D modeling, molecular docking and MD simulations of *PmAV* with RING finger domain of WSSV.

3. REVIEW OF LITERATURE

3.1 White spot syndrome Baculovirus

White spot syndrome baculovirus (WSBV) has been found across different shrimp species and in different Asian countries. The principal clinical sign of the disease associated with this virus is the presence of white spots in the exoskeleton and epidermis of the diseased shrimp. The causative virus itself consists of an envelope, rod-shaped nucleocapsid. It is extremely virulent, has a wide host range and targets various tissues.

The rapid onset and lethality of this disease are remarkable. Based on the histopathology of the affected animals and the characteristics of the virus, WSBV is either identical or closely related to Thailand's systemic ectodermal and mesodermal baculovirus (SEMBV). WSBV infection and its associated mortality are emerging as one of the most significant problems for the global shrimp industry (Lo *et al.*, 1996).

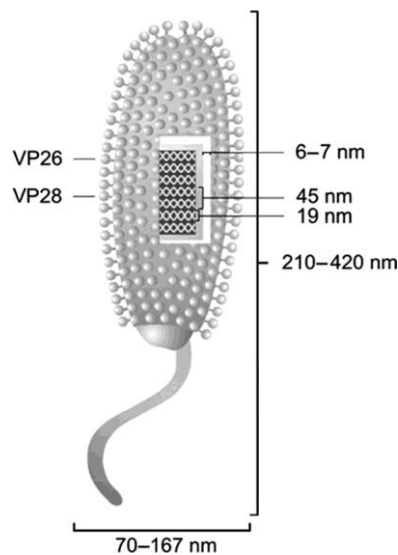


Figure 1: Schematic representation of the morphology of the WSSV viral particle.

3.1.1 White spot syndrome virus in Shrimps

Shrimp is one of the most important species in aquaculture. During the last decade, the worldwide shrimp culture was greatly puzzled by diseases caused by viruses particularly by white spot syndrome virus (WSSV) and suffered significant economic losses. Due to the extreme virulence of WSSV and a wide host range covering almost all crustaceans, it is difficult to prevent and inhibit the spread of the virus (Luo *et al.*, 2003). People also noticed

that although most of the WSSV infected shrimps died, a few of them still survived. Therefore it is interesting to find out the immune factors responsible for the shrimp resistance against WSSV (Gross *et al.*, 2001).

Like other invertebrates, shrimp lacks specific immunity, and its disease resistance relies on its innate defense system including a series of humoral and cellular immune factors, so it is regarded as an appropriate species for studying the innate immunity. Nowadays it becomes clear that the innate immune system plays an important role in host defense reaction. In recent years, some exciting progresses at the molecular level were made on shrimp innate immunity, such as proPO activating system and antimicrobial peptides. However, they all aimed at bacteria, fungi or parasites rather than viruses. (Liu *et al.*, 2007)

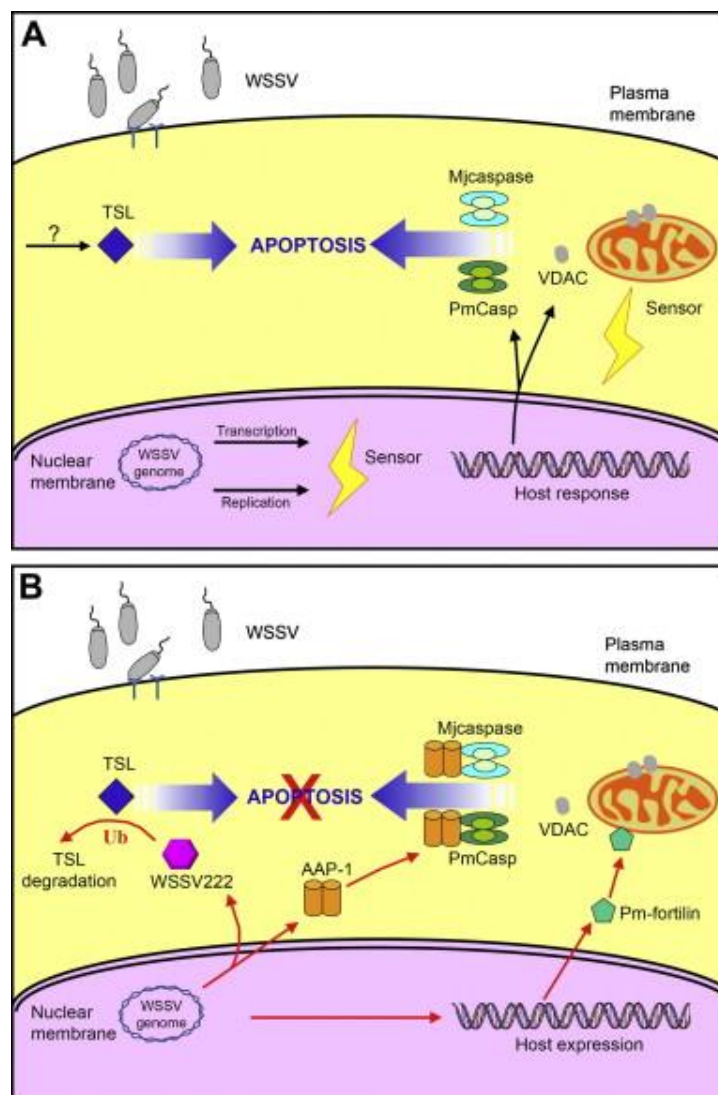


Figure 2 : A model of the apoptotic interactions between a shrimp host cell and WSSV. (Erridge *et al.*,2008)

(A) When WSSV infection occurs, the cellular sensors detect the presence of the virus and trigger the expression of PmCasp (the effector caspase), MjCaspase (the initiator caspase) and voltage-dependent anion channel (VDAC). Together with the mitochondrial stresses and the proapoptotic protein TSL, this causes the apoptosis program to be initiated. (B) Meanwhile, WSSV produces the two anti-apoptosis proteins, AAP-1 and WSV222. AAP-1 binds to and inhibits PmCasp and MjCaspase, and WSV222 attaches ubiquitin (Ub) to the TSL protein, which is then degraded. WSSV also induces the expression of shrimp Pm-fortilin, which acts to inhibit mitochondria-triggered apoptosis. Together, these three proteins function to block the occurrence of apoptosis, so that WSSV can complete its replication cycle. (Koizumi *et al.*,1999)

The target tissues of WSSV infected shrimp that are mainly of ectoderm and mesoderm, the initial infection starts in the stomach, gills, cuticular epidermis and the connective tissue of the hepatopancreas. Then, the lymphoid organ, antennal gland, muscle tissue, hematopoietic tissue, heart, hindgut also become infected. The nervous system and the compound eyes are only infected at the very late stages. The stomach, gills, cuticular epidermis, lymphoid organ, hematopoietic tissue and antennal gland are all heavily infected with WSSV at late stages of infection and become necrotic. (Rahman *et al.*,2007)

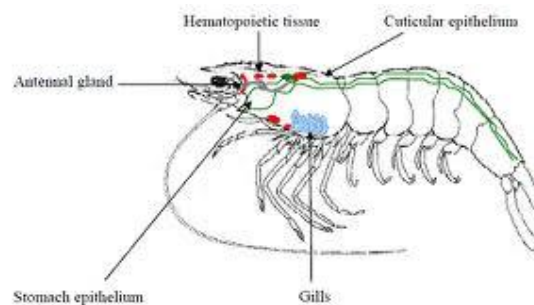


Figure 3 : The target tissues of WSSV in shrimp.

3.1.2 Genetic analysis of WSSV

White spot syndrome virus (WSSV) is a major disease agent of penaeid shrimp in Southeast Asia, the Indian continent, and in South and Central America. The disease is caused by an ovoid-to-bacilliform virus with a rod-shaped nucleocapsid and a tail-like appendix at one end of the virion. The virus contains a double stranded DNA with an estimated size of 290 kbp. Genetic analysis indicates that WSSV is a representative of a new virus group provisionally

named **whispovirus**. WSSV virions are ovoid-to-bacilliform in shape with a tail-like appendage at one end. They circulate ubiquitously in the haemolymph of infected shrimp. The virions contain a rod-shaped nucleocapsid, typically measuring 65–70 nm in diameter and 300–350 nm in length. The nucleocapsids, which contain a DNA-protein core bounded by a distinctive capsid layer giving it a crosshatched appearance, are wrapped singly into an envelope to shape the virion.

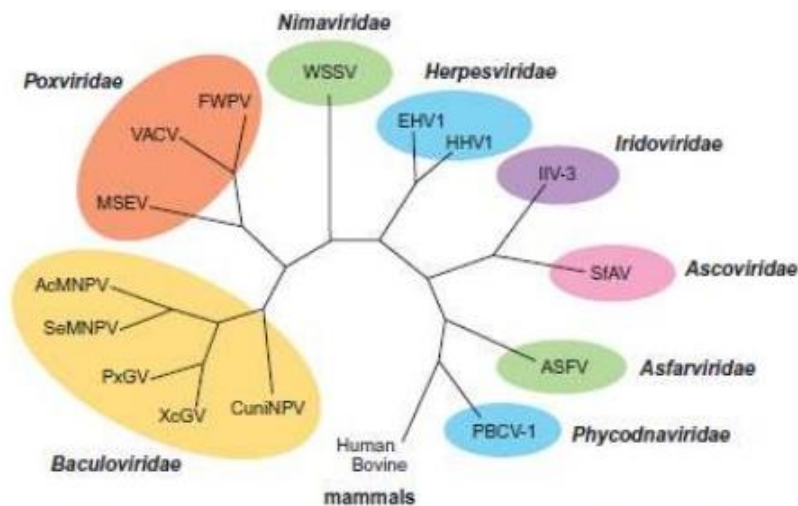


Figure 4: Cladogram based on genetic distances between aa sequences from DNA polymerase of wssv and other large DNA viruses.

The virus particle consists of at least five major proteins with estimated sizes of 28 kDa (RING DOMAIN), 26 kDa (VP26), 24 kDa (VP24), 19 kDa (VP19), and 15 kDa (VP15). RING DOMAIN and VP19 are associated with the virion envelope and VP26, VP24, and VP15 with the nucleocapsid. Amino acid analysis of RING DOMAIN, VP26 and VP24 indicated that these proteins have about 40% amino acid identity and that their genes may have evolved from a common ancestral gene. The role of the envelope and its proteins in the establishment of the systemic infection process has not been determined. Neutralization experiments have often been performed to study the role of virion proteins or their domains in the infection process. Neutralizing antibodies bind to envelope spikes on the virion and prevent attachment of the virus to the cell surface, cell entry, or virus uncoating.

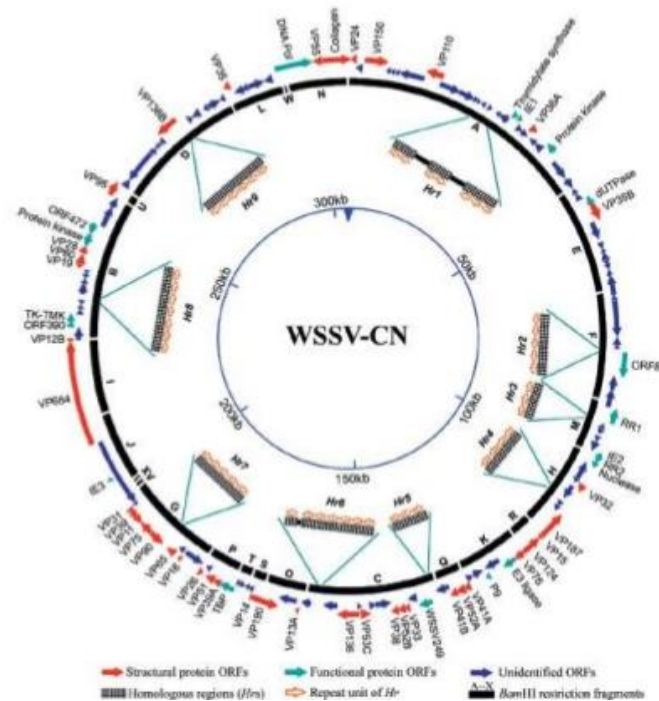


Figure 5: Schematics map of ds DNA-WSSV strain CN genome showing the genomic organization.

WSSV has a broad host range, infecting several crustacean species, like shrimp, crab, and crayfish. Little is known about WSSV infection and morphogenesis *in vivo*. Upon infection *per os*, infected cells are observed first in the stomach, gill, and cuticular epidermis of the shrimp. The infection subsequently spreads systemically in the shrimp to other tissues of mesodermal and ectodermal origin. Research on virus replication and virion morphogenesis shows that DNA replication and *de novo* envelope formation take place in the nucleus. The mechanism of virus entry into the shrimp and of the spread of the virus in the crustacean body is not known (Van Hulst *et al.*, 2001).

3.1.3 WSSV Pandemic

White spot syndrome virus (WSSV) is a pathogen of major economic importance in cultured penaeid shrimp. The virus is not only present in shrimp but also occurs in other freshwater and marine crustaceans, including crabs and crayfish. In cultured shrimp, WSSV infection can reach a cumulative mortality of up to 100% within 3–10 days and can cause large economic losses to the shrimp-culture industry. The virus was first discovered in Taiwan, from where it quickly spread to other shrimp-farming areas in Southeast Asia. WSSV initially appeared to be limited to Asia until it was found in Texas and South Carolina in November 1995.



Figure 6: the white spot virus pandemic (Years of first occurrence by location).



Figure 7: Asian pandemic of WSSV.

In early 1999, WSSV was also reported from Central and South America, and it has now also been detected in Europe and Australia. Intensive shrimp cultivation, inadequate sanitation, and worldwide trade has aggravated the disease incidence in crustaceans and enhanced disease dissemination. As such, WSSV has become an epizootic disease and is not only a major threat to shrimp culture but also to marine ecology (Van Hulten *et al.*, 2001).

3.2 Shrimp diseases

3.2.1 Defense system of Insects

Insects have a defense system against microorganisms similar to the innate immune system of vertebrates, involving both humoral and cellular responses. The humoral immune response of insects includes the induced synthesis of a group of antimicrobial peptides and proteins, lectins, and cell adhesion molecules (Yu *et al.*, 1999).

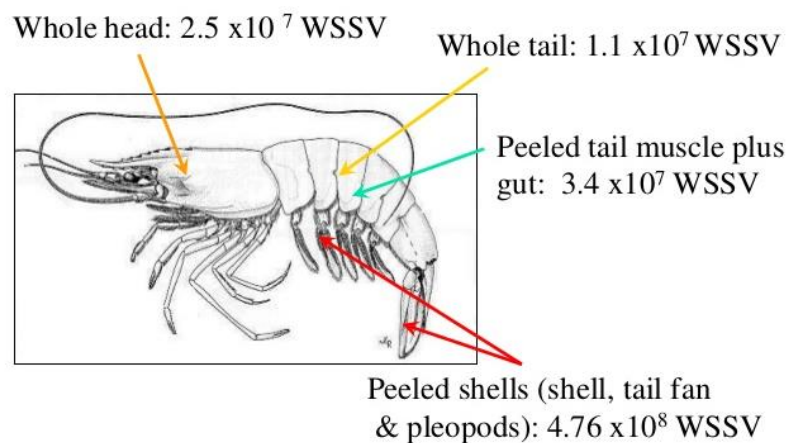


Figure 8: WSSV copy number/mg DNA in shrimp with acute phase WSSV infection.

Innate immunity not only in vertebrates but also in invertebrates is now attracting a significant amount of attention. Unlike vertebrates, invertebrates are believed to lack adaptive immunity and to rely completely on their innate immune system including a set of humoral and cellular immune reactions (Luo *et al.*, 2006). Recognition of non-self-materials in the innate immune system is mediated by a group of proteins named pattern recognition proteins (PRPs), which recognize and bind to different molecules on the surface of invading microorganisms. PRPs include lectins, lipopolysaccharide (LPS)-binding proteins, peptidoglycan-binding proteins, and β -1,3-glucanbinding proteins, and so on. Lectins are proteins that have the ability to bind to specific carbohydrates and they are present in almost all living organisms. Lectins have been known as playing a central role in nonself recognition and clearance of invaders in invertebrate immunity. Some lectins from invertebrates were reported to be involved in various biological responses, for instance promotion of phagocytosis, antibacterial activity, activation of the proPO system and ommatidial formation. From shrimps, a few lectins have been purified and characterized, but no lectin genes have been identified to date (Luo *et al.*, 2006).

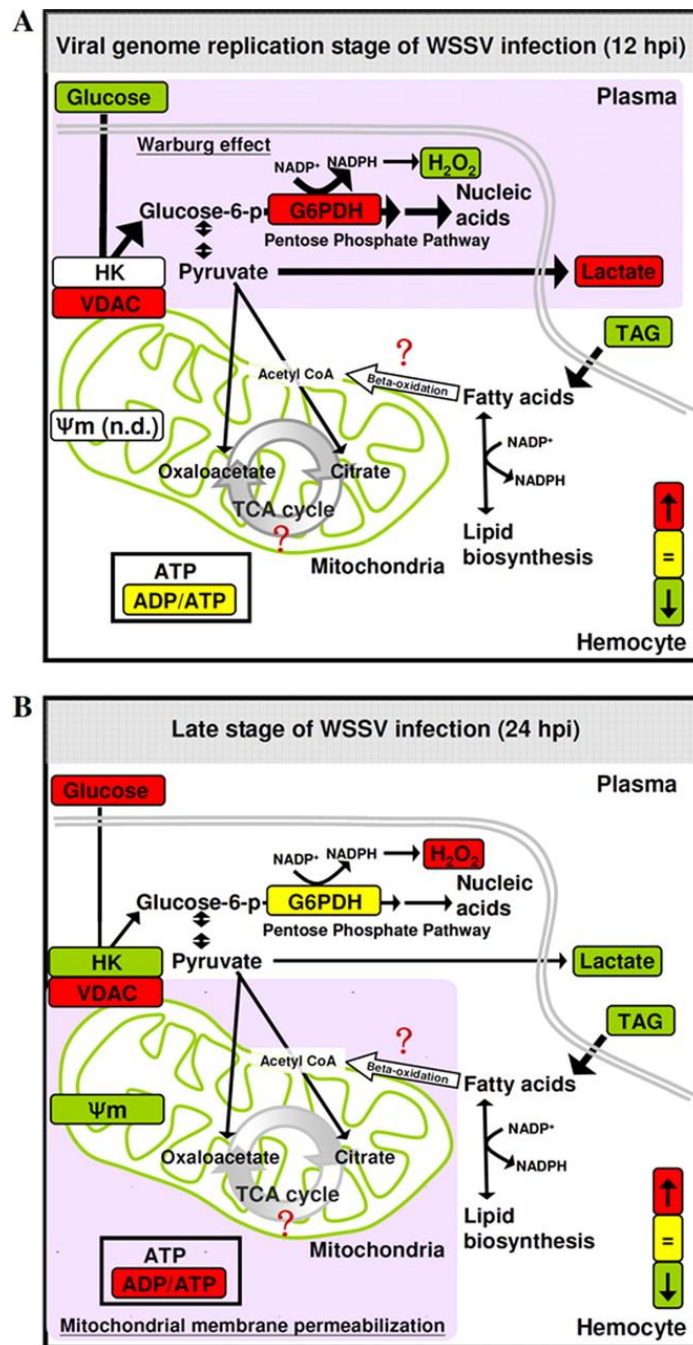


Figure 9: Schematic representation of the in vivo metabolic changes in shrimp hemocytes at the viral genome replication stage (12 h) (A) and the late stage (24 h) (B) of the first WSSV replication cycle.

3.2.2 Current diagnostic methods

The most important diseases of cultured penaeid shrimp have had viral or bacterial etiologies, but a few important diseases have fungal and protozoan agents as their cause. Diagnostic methods for these pathogens include the traditional methods of morphological pathology direct light microscopy, histopathology, electron microscopy, enhancement and bioassay

methods, traditional microbiology, and the application of serological methods. While tissue culture is considered to be a standard tool in medical and veterinary diagnostic labs, it has never been developed as a useable, routine diagnostic tool for shrimp pathogens. The need for rapid, sensitive diagnostic methods led to the application of modern biotechnology to penaeid shrimp disease.

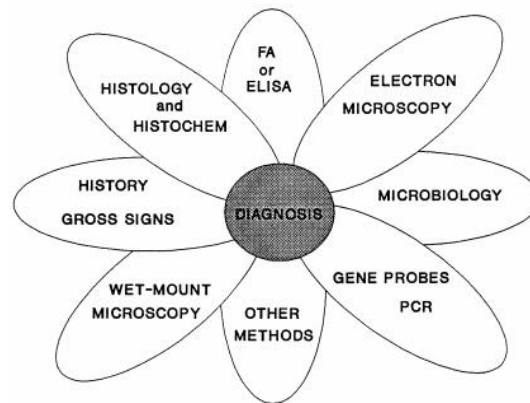


Figure 10: Source of information and available methods for diagnosis.

The industry now has modern diagnostic genomic probes with nonradioactive labels for viral pathogens like IHHNV, HPV, TSV, WSSV, MBV, and BP. Additional genomic probes for viruses, for bacterial pathogens like NHP and certain *Vibrio* spp., and Microsporidia have also been developed. Highly sensitive detection methods for some pathogens that employ DNA amplification methods based on the polymerase chain reaction PCR now exist, and more PCR methods are being developed for additional agents. These advanced molecular methods promise to provide badly needed diagnostic and research tools to an industry reeling from catastrophic epizootics and which must become poised to go on with the next phase of its development as an industry that must be better able to understand and manage disease.

Table 1: Methods available for diagnosis of shrimp diseases.

Method	Tests and data obtained
History	History of disease at facility or in region, facility design, source of seed stock (e.g. wild or domestic specific pathogen-free, SPF, or resistant, SPR), type of feed used, environmental conditions, etc.
Gross, clinical signs	Lesions visible, behavior, abnormal growth, feeding or food conversion efficiency, etc.
Direct microscopy	Bright-field, phase contrast or dark-field microscopic examination of stained or unstained tissue smears, whole-mounts, wet-mounts, etc. of diseased or abnormal specimens.
Histopathology	Routine histological or histochemical (with special stains) analysis of tissue sections.
Electron microscopy	Ultrastructural examination of tissue sections, negatively stained virus preparations, or sample surfaces.
Culture and biochemical identification	Routine culture and isolation of bacterial isolates on artificial media and identification using biochemical reactions on unique substrates.
Enhancement	Rearing samples of the appropriate life stages of shrimp under controlled, stressful conditions to 'enhance' expression of latent or low grade infections.
Bioassay	Exposure of susceptible, indicator shrimp to presumed carriers of a pathogenic agent.
Serological methods	Use of specific antibodies as diagnostic reagents in immunoblot, agglutination, IFAT, ELISA, or other tests.
Hematology and clinical chemistry	Determination of hemocyte differential count, hemolymph clotting time, glucose, lactic acid, fatty acids, certain enzymes, etc.
Toxicology/analysis	Detection of toxicants by analysis and verification of toxicity by bioassay.
DNA Probes	Detection of unique portions of a pathogen's nucleic acid using a labeled DNA probe.
PCR	Amplification of unique sections of a pathogen's genome to readily detectable concentrations using specific primer pairs.
Tissue culture	In vitro culture of shrimp pathogens in non-shrimp tissue culture systems or in primary cell cultures derived from shrimp.

3.3 Economic importance

White spot disease of shrimp caused by white spot syndrome virus (WSSV) is responsible for a major proportion of the diseases plaguing commercial shrimp ponds, and has resulted in high mortality and economic losses. It is well known that shrimp, other crustaceans and invertebrates in general lack a truly adaptive immune response system and appear to rely on a variety of innate immune response systems to rapidly and efficiently recognize and destroy "non-self" materials. It has been shown that shrimp that recover from viral infection contain a humoral neutralizing factor that can reduce the virus. This led to the possibility of using an agent or immunostimulant to stimulate this humoral factor and enhance the shrimp defense mechanisms (Söderhäll *et al.*, 1998).

Such applications have been reported against WSSV infection, for example, oral administration of lipopolysaccharide to *Penaeus japonicus*, glucan to *Penaeus monodon*, and fucoidan to *P. monodon*. Shrimp hemocyanin was also proved to be a useful antiviral agent against both DNA and RNA viruses. Novel antimicrobial peptides have also been discovered

that have potential uses to reduce mortality in aquaculture. WSSV envelope and structural proteins have been studied intensively as possible immunotherapeutic agents to prevent WSSV infections in *Penaeus* spp. Recently, a number of genes that were involved in the shrimp immune system have been discovered and characterized. (Tonganunt *et al.*,2008)

Among the shrimp viral pathogens, white spot syndrome virus (WSSV) is highly pathogenic and responsible for huge economic losses in the shrimp culture industry worldwide. WSSV is an enveloped DNA virus that causes 100% mortality in cultured shrimp within 3–4days. Due to the extreme virulence of WSSV and a wide host range covering almost all crustaceans, it is difficult to prevent and inhibit the spread of the virus. Due to the current intensity of aquaculture practices and the broad host range of WSSV, novel control strategies including “vaccination” against the virus would be highly desirable (Schneidman-Duhovny *et al.*, 2003).

However, invertebrates lack a true adaptive immune response system and rely on various innate immune responses. Nevertheless, some studies have reported the protection of shrimp using either inactivated pathogens or recombinant proteins against this viral pathogen suggesting some aspects of specific immunity appear to be present in some cases. In the case of inactivated pathogen and protein vaccines, the protection is for a maximum period of 3 weeks which might be due to the lack of immunologic memory in shrimp. The protection against WSSV in shrimp may require continuous exposure of immunogenic protein in shrimp tissues. DNA constructs may help for persistent expression of immunogenic proteins in shrimp tissue and be more suitable for long-term protection instead of protein vaccines in the case of shrimp (Rajeshkumar *et al.*,2009).

Researchers also reported that intramuscularly DNA vaccinated shrimp developed resistance against this virus for about 7 weeks. However, most of the vaccination studies have relied on intramuscular injection into shrimp. This route of delivery is not practically feasible in shrimp farming. Hence, work needs to be carried out to develop simple and cost effective delivery systems to deliver DNA constructs orally for mass immunostimulation in shrimp farms (Rajeshkumar *et al.*,2009).

In aquaculture, several studies have reported the feasibility of gene transfer through oral route using chitosan nanoparticles in fish. Royet al. reported that oral administration of DNA nanoparticles prepared by complexing plasmid DNA with chitosan resulted in transduced gene expression in the intestinal epithelium *in vivo*. Chitosan nanoparticles have also been

used for delivery of bacterially expressed double stranded (ds) RNA (Rajeshkumar *et al.*,2009).

3.4 Ring Domain of WSSV

So far, approximately 58 structural proteins, including 9 nucleocapsid proteins and 32 envelope proteins, have been identified in the WSSV virions, and they molecular mechanism, the dynamic behaviour of proteins and its interactions in shrimp infected with WSSV disease is still unknown. However four proteins of WSSV, namely WSSV199, WSSV222, WSSV249 and WSSV403 are known to contain RING-H2 domains. Many RING finger domains simultaneously bind ubiquitination enzymes and their substrates and hence function as ligases. Therefore, it is known that viral proteins containing RING finger domain might play a key role in the modification of host's ubiquitination pathway. Many findings have shown that viral infection can up-regulate expression of ubiquitin, suggesting that the ubiquitin system may play a key role in the course of viral infection. Increasing number of RFPs (RING Finger Proteins) has been identified in viruses.

3.5 PmAV protein

Previous studies showed the existence of nonprotein antiviral substance in crustaceans, but so far little is known about the possible innate antiviral factor generated by the interaction between host cell and virus, and neither antiviral gene nor antiviral protein has been characterized from crustaceans. It has been proved that *Penaeus monodon* is the first virus-resistant shrimps having antiviral gene PmAV. The PmAV protein might play an important role in the defense mechanism against viruses.

The PmAVcDNA encodes a 170 amino acid peptide with a C-type lectin-like domain (CTLCD) and was found to be up-regulated in virus resistant shrimps. Furthermore, recombinant PmAV protein can inhibit virus-induced cytopathic effects in cultured fish cell. The PmAV protein might therefore play an important role in the antiviral defense of shrimp. These findings also suggest that PmAV gene is inducible by infection, in contrast to the constitutively expressed shrimp penaeidin genes (Luo *et al.*,2007).

3.5.1 C-type lectins

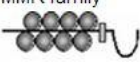



Classical C-type lectins contain so-called carbohydrate recognition domains (CRDs) that bind carbohydrate structures in a calcium (Ca²⁺)-dependent manner. Ca²⁺ ions are directly

involved in both ligand binding as well as in maintaining the structural integrity of the CRD that is necessary for the lectin activity. C-type lectins contain a prototypic lectin fold, consisting of two anti-parallel β -strands and two α -helices. The C-type CRDs form a subfamily of the larger group of protein domains called C-type lectin-like domains (CTLDs). Some CTLDs bind protein or lipid moieties instead of carbohydrates; often these are Ca^{2+} -independent. However, there have been examples described of Ca^{2+} -independent carbohydrate binding. CTLDs containing receptors are also indicated as C-type lectin-like receptors (CLRs) (Cambi *et al.*, 2005).

C-type lectins are either produced as transmembrane proteins or secreted as soluble proteins (Table 2). Examples of soluble C-type lectins include members of the collectins family, such as the lung surfactant proteins A (SP-A) and SP-D, which are secreted at the luminal surface of pulmonary epithelial cells, and the mannose-binding protein (MBP), a collectin present in the plasma. Transmembrane C-type lectins can be divided into two groups, depending on the orientation of their amino (N)-terminus. These are type I and type II C-type lectins depending on their N-terminus pointing outwards or inwards into the cytoplasm of the cell respectively. (Weis *et al.*, 1998)




Examples of transmembrane C-type lectins are the selectins, the mannose receptor (MMR) family and the dendritic cell-specific ICAM-3 grabbing non-integrin (DC-SIGN).

Table 2. Overview of structural and functional relationships of subfamilies of C-type lectins involved in pathogen sensing. (Cambi *et al.*, 2005)

Group and molecular structure ^(a)	C-type lectin	Pathogen	Ligand specificity
	MMR	HIV, <i>Pneumocystis</i> , <i>Mycobacterium tuberculosis</i> , <i>C. albicans</i>	Mannose, fucose, sLe ^{x(b)}
	DEC-205	Unknown	Unknown
	Endo-180	Unknown	Collagen, mannose, fucose, GlcNAc
	MBL	HIV, IAV, <i>Staphylococcus aureus</i> , <i>Streptococcus pneumoniae</i> , <i>C. albicans</i> , <i>Aspergillus fumigatus</i>	GlcNAc, ManNAc, fucose, glucose
	SP-A	IAV, RSV, HSV-1, <i>S. aureus</i> , <i>S. pneumoniae</i> , <i>A. fumigatus</i>	ManNAc, fucose, glucose, GlcNAc
	SP-D	IAV, RSV, <i>M. tuberculosis</i> , <i>Pseudomonas aeruginosa</i> , <i>A. fumigatus</i> , <i>C. albicans</i>	Maltose, mannose, glucose, lactose, galactose, GlcNAc
	DC-SIGN	HIV, HCV, CMV, Dengue, <i>H. pylori</i> , <i>M. tuberculosis</i> , <i>S. mansoni</i> , <i>C. albicans</i> , <i>A. fumigatus</i> , <i>Leishmania</i>	Mannan, Le ^x , Le ^a , Le ^y , Le ^b , sLe ^a , ManLAM
	L-SIGN	HIV, HCV, <i>S. mansoni</i>	Mannan, Le ^a , Le ^y , Le ^b
	DCIR	Unknown	Unknown
	Langerin	Unknown	Mannose, GlcNAc, fucose, 6SLe ^x
	DCAL-1	Unknown	Unknown
	BDCA-2	Unknown	Unknown
	β-GR (Dectin-1)	<i>Pneumocystis</i> , <i>C. albicans</i>	β-Glucan
	CLEC-1	Unknown	Unknown
	CLEC-2	Unknown	Unknown

(a) . Based on nomenclature <http://ctld.glycob.ox.ac.uk>

(b) . This interaction occurs through the cystein-rich domain, not the CRD, of the MMR.

 C-type lectin domain;  fibronectin type II repeat;  collagen-like triple helix.

IAV, influenza A virus; RSV, respiratory syncytial virus; HSV-1, herpes simplex virus type 1; β-GR, β-glucan receptor; GlcNAc, N-acetyl-D-glucosamine; ManNAc, N-acetyl-D-mannosamine; ManLAM, mannosyl-lipoarabinomannan; Le^a, LewisA; Le^b, LewisB; Le^y, LewisY; 6SLe^x, 6-sulpho sialyl-Lewis^x.

In the immune system, C-type lectins and CLR have been shown to act both as adhesion and as pathogen recognition receptors. While cell–cell contact is a primary function of selectins, other C-type lectins, like collectins, are specialized in recognition of pathogens (Table 2). Interestingly, DC-SIGN is a cell adhesion receptor as well as a pathogen recognition receptor. As adhesion receptor, DC-SIGN mediates the contact between dendritic cells (DCs) and T lymphocytes, by binding to ICAM-3, and mediates rolling of DCs on endothelium, by interacting with ICAM-2. As pathogen uptake receptor, DC-SIGN recognizes a variety of microorganisms, including viruses, bacteria, fungi and several parasites. Recently, the type I CLR MMR, mainly known as pathogen recognition receptor, was discovered to mediate adhesion between human lymphatic endothelium and lymphocytes through L-selectin (Cambi *et al.*, 2005; Weis *et al.*, 1998).

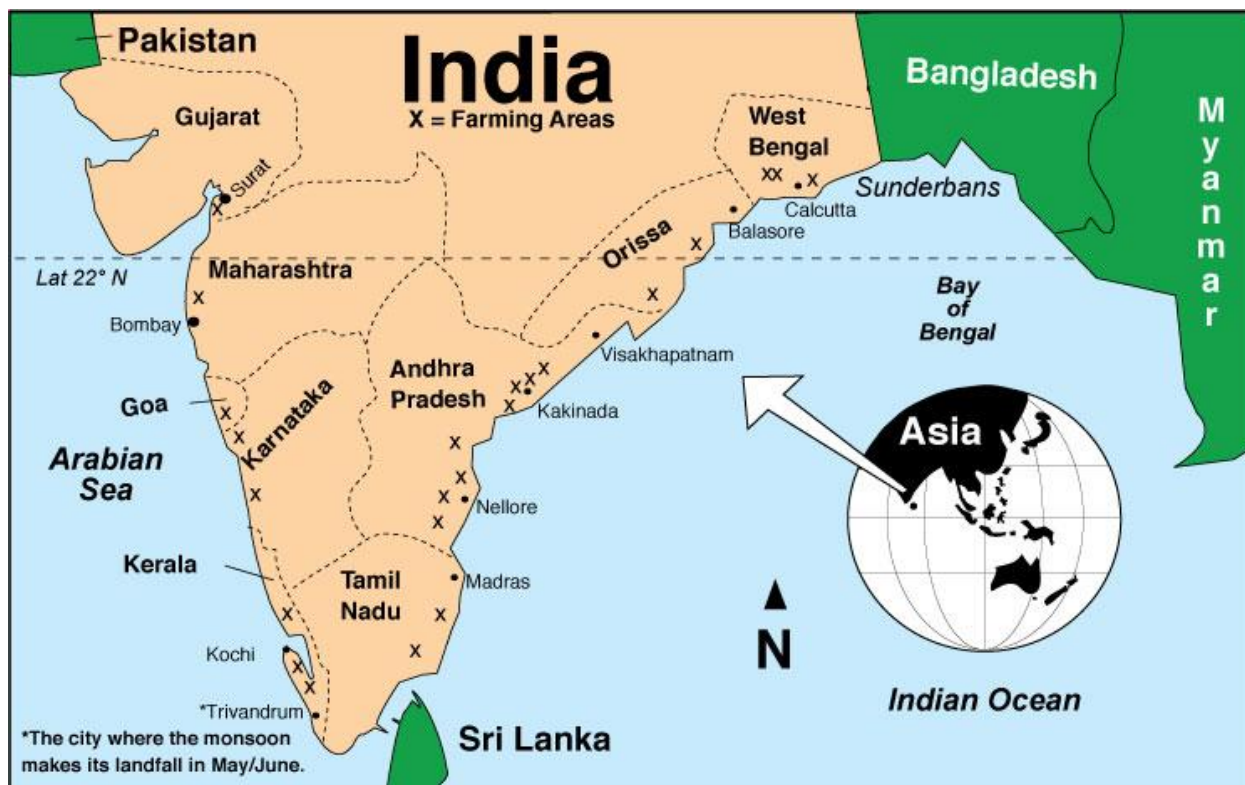


Figure 11: *Penaeus monodon* Hatchery Practices in India.

Although there are 30 shrimp hatcheries in Andhra Pradesh equipped with PCR laboratories, only a few screen the broodstock for whitespot. Random checking is performed by selecting one or two individuals from a batch of 50–60 broodstock. Wild-caught broodstock is the only source of shrimp seed. Studies indicate that about a quarter of wild-caught spawners are infected with the whitespot virus. Moreover, quality appears to vary with the season, and there are some indications that average size has become smaller and quality poorer.

Table 3: Shrimp Farming Potential, Usage and Production in India

State	Potential Area for Shrimp Farming
Andhra Pradesh	150,000
West Bengal	405,000
Orissa	31,000
Kerala	65,000
Tamil Nadu and Pondicherry	56,800
Karnataka	8,000
Gujarat	376,000
Goa	18,500
Maharashtra	80,000
Totals	1,190,900

3.5.2 Carbohydrate-Mediated Mechanisms

Complex carbohydrates are widely distributed in animal tissues. It has long been appreciated that there is extensive glycosylation of proteins in the serum and extracellular matrix, and of proteins and lipids at the surface of cells. More recently, it has become clear that proteins within the cell, including many found within the nuclear pore complex and in the cytosolic compartment, are also conjugated to sugars (Drickamer 1988; Sheriff *et al.*, 1994). The potential of these complex carbohydrate structures for encoding information has been repeatedly noted. Alternative isomeric linkages between sugars and the formation of branched structures can generate an enormous number of structures from even a small number of saccharide units. As increasingly refined analytical methods become available, more and more diversity in the complex carbohydrate structures associated with animal cells is being documented. What we know of the biosynthetic machinery which generates the carbohydrate structures suggests that the glycosyltransferases responsible for the assembly process have highly specific substrate requirements and that each enzyme generates only a very limited number of bonds. Thus, the diversity in carbohydrate structure presumably arises from a multiplicity of synthetic enzymes. Taken together, these facts suggest that information can be and probably is encoded in carbohydrates (Sheriff *et al.*, 1994).

Experimental evidence directly implicates complex carbohydrates in recognition processes including adhesion between cells, adhesion of cells to the extracellular matrix, and specific recognition of cells (such as egg and sperm) by one another (all). In addition, carbohydrates are recognized as differentiation markers and as antigenic determinants. Thus, both theoretical and experimental considerations suggest the need for biological molecules, presumably proteins, which can decode the information found in complex carbohydrate structures.

Animal lectins, more or less by definition, have the potential to do this. Lectins are generally considered to be nonenzymatic (and nonimmune) proteins which selectively bind to specific carbohydrate structures. Plant lectins, which have the ability to distinguish subtle differences in carbohydrate structures found in animals, provide a paradigm for endogenous animal proteins which make similar distinctions. An increasing body of evidence demonstrates the existence of a large number of animal lectins. Recently obtained structural information makes it possible to organize the known lectins into several categories. (Drickamer 1988a; 1999b)The C-type (Ca²⁺-dependent) animal lectins are structurally related to the

asialoglycoprotein receptor, while the S-type (thiol-dependent) animal lectins form a distinct group. Most of the animal carbohydrate-binding proteins which do not fall into these two groups are proteins best known in other contexts. These include some of the proteins which mediate the interactions of cells with the extracellular matrix (fibronectin and laminin), serum immunoglobulins which have specificity for carbohydrates, the mannose-phosphate receptor, which directs transport of proteins to lysosomes, viral hemagglutinins, and serum amyloid protein, a member of the pentraxin family (Sheriff *et al.*, 1994).

Carbohydrate-recognition domains of C-type (Ca²⁺-dependent) animal lectins serve as prototypes for an important family of protein modules. Only some domains in this family bind Ca²⁺ or sugars. A comparison of recent structures of C type lectin-like domains reveals diversity in the modular fold, particularly in the region associated with Ca²⁺ and sugar binding. Some of this diversity reflects the changes that occur during normal physiological functioning of the domains. C-type lectin-like domains associate with each other through several different surfaces to form dimers and trimers, from which ligand-binding sites project in a variety of different orientations (Feizi 2000; Sheriff *et al.*, 1994).

3.6 Techniques used

3.6.1 Protein 3D structure prediction

Protein structures have proven to be a crucial piece of information for biomedical research. Of the millions of currently sequenced proteins only a small fraction is experimentally solved for structure and the only feasible way to bridge the gap between sequence and structure data is computational modeling. In the late 1950s Anfinsen and coworkers suggested that an amino acid sequence carries all the information needed to guide protein folding into a specific spatial shape of a protein. In turn, protein structure can provide important insights into understanding the biological role of protein molecules in cellular processes. Thus, in principle it should be possible to form a hypothesis concerning the function of a protein just from its amino acid sequence, using the structure of a protein as a stepping stone toward reaching this goal.

Modern experimental methods for determining protein structure through X-ray crystallography or NMR spectroscopy can solve only a small fraction of proteins sequenced by the large-scale genome sequencing projects, because of technology limitations and time constraints. Currently, there are more than 6,800,000 protein sequences accumulated in the no

redundant protein sequence database (NR; accessible through the National Center for Biotechnology Information: <ftp://ftp.ncbi.nlm.nih.gov/blast/db/>) and fewer than 50,000 protein structures in the Protein Data Bank (PDB; <http://www.rcsb.org/pdb/>). With these numbers at hand, it seems that the only way to bridge the ever-growing gap between protein sequence and structure is computational structure modeling (Kryshtafovych *et al.*, 2009).

Here we have used I-Tasser server for the prediction of 3D coordinates of proteins. I-TASSER server is an on-line platform for protein structure and function predictions. 3D models are built based on multiple-threading alignments by LOMETS and iterative template fragment assembly simulations; function insights are derived by matching the 3D models with BioLiP protein function database. Based on the statistical significance of the threading alignments and the structure convergence of the Monte Carlo simulations, a new confidence score (C-score) is introduced and benchmarked for the I-TASSER server, which demonstrates a strong correlation with the real quality of the final models. The strong correlation data allows us to make quantitative estimates of the accuracy of the I-TASSER predictions. Using a 2-order polynomial equation fit from 300 training proteins, it can predict the TM-score and RMSD of the final models with an average error of 0.08 and 2.0 Å respectively in a large scale benchmark test (Zhang *et al.*, 2008).

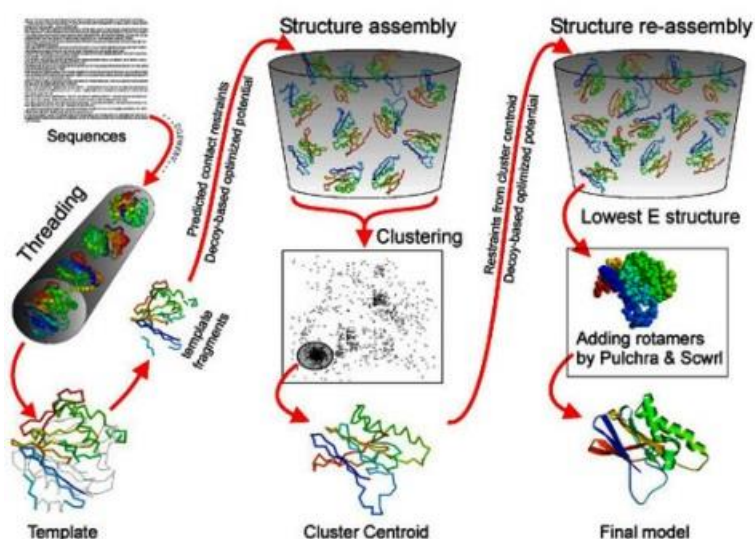


Figure 12: Flowchart of I-TASSER method for protein structure prediction.

3.6.2 Protein-Protein Interactions

The field of protein–protein interactions has rapidly progressed in the past 10 years. Proteins seldom act alone and most cellular functions are regulated through intricate protein–protein

interaction networks. Large efforts have been made to unveil these interactions in a high-throughput manner, with the first interactome maps for several model organisms. We can identify binary interactions—that involve only two proteins by multiple methods such as array technology, cross-linking study, cytoplasmic complementation assay, NMR, two hybrid or X-ray crystallography)—and multi-component interactions. It is to be noted that if the experimental protein–protein structures are known, there are computer-based tools available to design proteins to bind faster and tighter to their protein-complex partner by electrostatic optimization between the two proteins.

Protein–protein docking aims to predict the three dimensional structure of a complex from the knowledge of the structure of the individual proteins in aqueous solution. This motion generally involves the displacement or the internal rearrangement of loops or domains and can also be characterized by the simultaneous movement of several flexible parts (up to three in known cases).

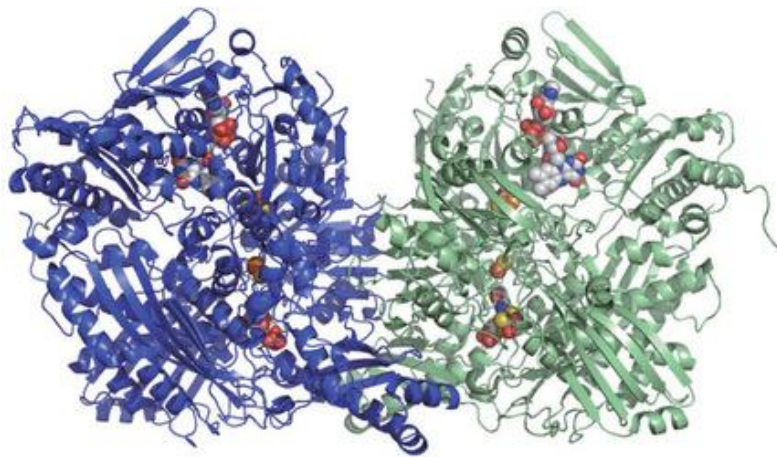


Figure 13: Example of protein-protein interaction by docking analysis.

3.6.3 Molecular Dynamics Simulation

MD simulations are the most common and widely used computer simulations techniques to study the conformational dynamics of biological macromolecules such as proteins. Newton's classical equations of motion are solved iteratively to simulate the motion of a system of particles as a function of time. Due to the remarkable resolution in space (single atom), time (femtosecond), and energy, MD represents a powerful complement to experimental techniques, providing mechanistic insight into experimentally observed processes.

Consequently, MD simulations can be treated as a virtual experimental method that provides an interface between the theory and a real laboratory experiment. In a typical MD simulation, a starting configuration is generated from an experimentally determined structure, and put into an environment that best mimics its natural environment. Obviously, the quality of the obtained dynamic model depends on the quality of the starting model. Therefore, MD can be used to address specific questions about the properties of a model system, often more straightforwardly than the experiments on real systems.

MD primarily concerns with the energy of a given molecular system based on the structure. Thus, the initial step in designing a molecular modeling study is to describe the problem as one involving a structure-energy relationship. Theoretically, there are two different possible ways to study the energy of such molecular systems, namely (i) a molecular mechanics model that describes the energy of a molecule in terms of a classical force field that is used to determine the distortions from equilibrium values such as bond lengths, angles, and also non-bonded interactions; (ii) a quantum mechanics based model that describes the energy of a molecule in terms of interactions among nuclei and electrons as given by the Schrödinger equation. In principle, quantum mechanics would be the best approach to compute the dynamics and functional properties of molecular systems.

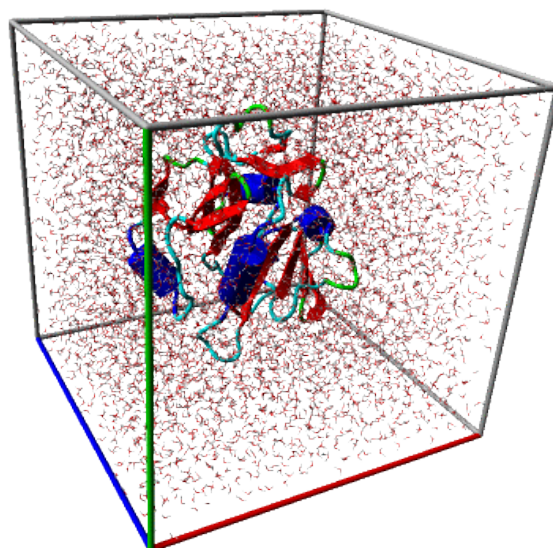


Figure 14: Creating water environment for molecular dynamics simulation.

The flowchart below describes the general procedure for setting up an MD run. Please refer to Appendix IV for the step by step procedure for the MDsimulations within the GROMACS simulation suite.

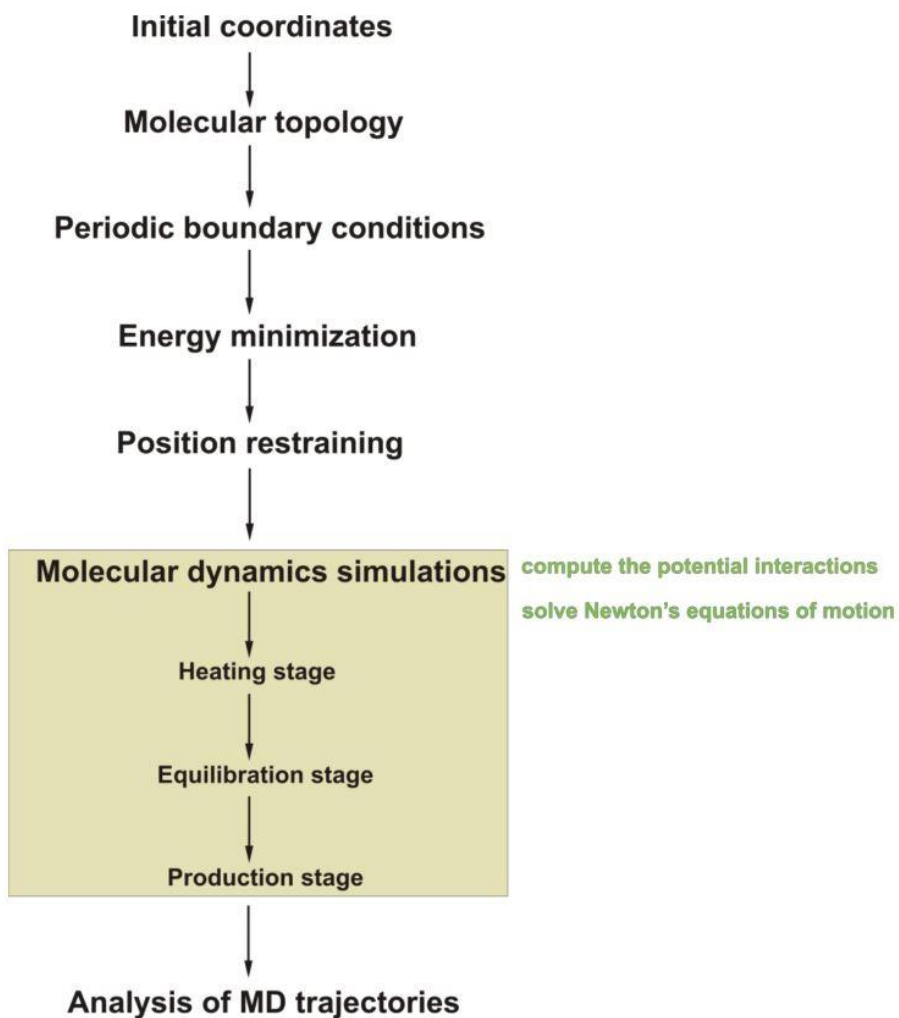


Figure 15 : flowchart for setting up an MD run.

4. METHODOLOGY

4.1 Molecular modeling of the proteins

The protein sequences of *Penaeus monodon* PmAV (Accession No.AAQ75589.1) and RING domain of wsv199 (white spot syndrome virus) (NP_477666.1) were taken from National Centre for Biotechnology Information (NCBI) [www.ncbi.nlm.nih.gov]. Each of the FASTA sequences was subjected to Protein-BLAST [www.blast.ncbi.nlm.nih.gov] against PDB [www.rcsb.org] database to select the suitable templates. The best template was selected on the basis of proportion of similarity & identity, E-value, bit scores and query coverage region. The three dimensional coordinates of PmAV was created using Modeller 9v10 (Eswar *et al.*, 2008).5-5 models for both of the proteins were generated through modeler and consequently the good 3D model was obtained. Another modeling method using I-TASSER server (Zhang *et al.*, 2008) was applied, in order to compare the results obtained by Modeller 9v10. I-TASSER is based on model building using multiple-template fragment alignments.

4.2 Minimization and validation

Energy minimization of all the models was performed using steepest descent method by SwissPDB Viewer (Guex *et al.*,1997). All the minimized models were applied for validation. Stereo- chemistry of all the models was evaluated by PROCHEK and WHAT-IF softwares (Laskowski *et al.*, 1993; Vriend *et al.*,1990). The model showing best Ramachandran plot was chosen for further analysis.

4.3 Molecular dynamics simulations of PmAV and RING DOMAIN

Molecular dynamics simulations of PmAV and RING DOMAIN were performed using the GROMACS 4.5.3 (Berendsen *et al.*, 1995) (<http://www.gromacs.org/>) under OPLS 2005 atoms force field on Pentium (R) dual-core processor machine. The 3D structure of PmAV was immersed in a cubic box of 0.85 nm and periodic boundary conditions were applied using edit conf tool followed by addition of 21,668 SPC water molecules.

System was made electrically neutral by adding 2Na⁺ using the 'genion' tool. The system was first minimized for energy in 1000 steps by steepest descent method to remove excessive strain. The minimized system was then subjected to MD in two steps. Initially NVT ensemble

(constant number of particles, volume, and temperature) was performed for 500 ps, followed by NPT ensemble (constant number of particles, pressure, and temperature) for 500 ps. The well-equilibrated system was then subjected to molecular dynamics simulations for 9 ns. The total number of atoms in the system was 72,691. Temperature was kept constant at 300 K with Andersen thermostat, pressure coupling of 1 bar with Berendsen algorithm, and the system was further allowed to undergo production runs. LINCS algorithm (Hess *et al.*, 1997) was used to constrain the lengths of all bonds while the water molecules were restrained using the SETTLE algorithm (Miyamoto *et al.*, 1992).

The trajectory files were analyzed by using `g_rms` and `g_gyr` utilities of GROMACS to obtain the root-mean square deviation (RMSD) and radius of gyration (R_g) values. RMSD values for $C\alpha$ atoms from the initial structure were considered as a necessary condition to determine the convergence of the proteins toward equilibrium and calculated by:

$$RMSD(t_1, t_2) = \left[\frac{1}{M} \sum_{i=1}^N m_i \|r_i(t_1) - r_i(t_2)\|^2 \right]^{\frac{1}{2}}$$

Where $\sum_{i=1}^N m_i$ and $r_i(t)$ is the position of atom i at time t . The shape of the protein molecule at all instants of simulation is indicated through the hydrodynamic radius obtained using the radius of gyration calculated by:

$$R_g = \left(\frac{\sum_i \|r_i\|^2 m_i}{\sum_i m_i} \right)^{\frac{1}{2}}$$

m_i is the mass of atom i and r_i is the position of atom i with respect to the center of mass of the molecule. The same approach was used to simulate RING DOMAIN under similar conditions (300 K temperature and 1 bar pressure) for 9 ns. 8 Na^+ and 19,256 SPC water molecules were used during simulation of RING DOMAIN and a total of 60,323 atoms were present in the system.

4.4 Protein-Protein interaction study by Docking

Shape Complementarity Principle based docking was performed using PatchDOCK and FireDock servers (19, 20, 21). The output files were subjected to FiberDock server for flexible induced-fit backbone and side chain refinement of the protein complexes (22, 23). The molecular interaction plots between PmAV and RING domain of wsv199 were generated

using Dimplot in LIGPLOT software (24, 25). The 3D structure and detailed interaction of all protein complexes were visualized using Accelrys Viewerlite software [accelrys.com/products/discovery-studio/visualization-download.php].

4.5 Molecular dynamics simulation of PmAV-RING DOMAIN complex

Molecular dynamics simulations of PmAV-RING DOMAIN complex was performed using the GROMACS 4.5.3 undergromos43a1 force field (Berendsen *et al.*, 1995). The whole simulation experiment was done for 13 ns by using 26,553 SPC water molecules and 10 Na⁺ ions. The same methodology was adopted for simulation of this complex as done in the MD of individual proteins PmAV and RING DOMAIN. The trajectory files were analyzed by using *g_rms* and *g_gyrate* utilities of GROMACS to obtain the RMSD and Rg values while van der Waals and short range electrostatic energies were computed in order to obtain information about the stability of the complex (Verma *et al.*, 2012). Snapshots of the docked complex were generated by PyMol 0.99 and Discovery Studio Visualizer 2.5 software in order to visualize the pictorial depiction of the interaction between both proteins.

5. RESULTS

Knowledge of the shrimp defense at the molecular level has focused mainly on antimicrobial factors (Feizi 2000) while little is known about the possible antiviral factors. So far, only one antiviral protein, PmAV has been identified from shrimp (Luo *et al.*,2003). However, the molecular mechanism of this protein in controlling WSSV infection is still unknown. WSSV RING finger domain proteins, WSSV199, WSSV222, WSSV249 and WSSV403 functions as E3 ubiquitin ligases and mediates the degradation of important host immune proteins for disease manifestation (Yang *et al.*, 2001). Therefore, an attempt was made in this study to determine molecular interaction between PmAV and RING finger domain of WSSV proteins using *in silico* approach.

Physico-chemical characterization

All C-type lectins contain a carbohydrate recognition domain (CRD), which mediates sugar binding with Ca²⁺ (Feizi 2000). The results of the primary structure analysis as predicted using ExPASy's ProtParam server showed that 170 amino acids long PmAV has an estimated molecular weight of 19317.5 daltons. The total number of negatively charged residues (Asp + Glu) was 22 while the total number of positively charged residues (Arg + Lys) was 11. The calculated isoelectric point (pI) was 4.8. As the computed pI value of protein was less than 7 (pI < 7), this indicates that this protein can be considered as acidic in character. This information will be useful for its purification by isoelectric focusing point in a lab facility. Extinction coefficient of PmAV protein at 280 nm was found to be 49765 M⁻¹ cm⁻¹ with respect to the concentration of Cys residues. This information can help in the quantitative study of protein–protein and protein–ligand interactions in solution. GRAVY index of PmAV protein was found to be -0.245. The Grand Average hydropathicity (GRAVY) value for a peptide or protein is calculated as the sum of hydropathy values of all the amino acids, divided by the number of residues in the sequence. This low range of value indicates the possibility of better interaction with water. The instability index (II) was computed to be 39.74 indicating that the protein is stable.

Comparative modeling of the proteins

The modeling of the 3D structure of PMAV and RING finger domain (Fig. 16) was performed by modeller 9v10. In order to generate the 3 dimensional coordinates of protein, BLASTp was performed.

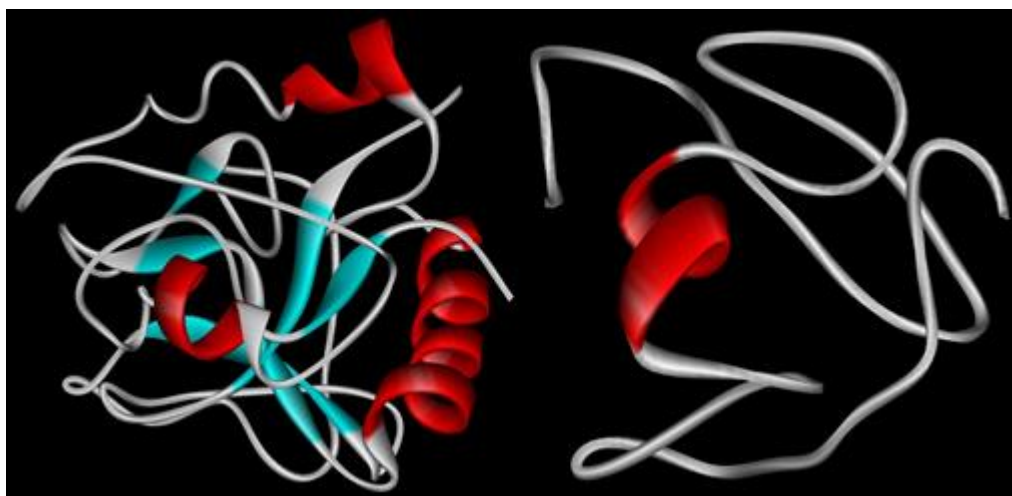


Figure 16. Modeled structure of (a) PMAV (b) RING domain of wsv199. Visualization by AccelrysViewerLite. All the validating parameters were assuring the quality and reliability of models.

Table 4: List of the selected templates.

Name of the protein	AA	Template_chain	Name of the template
PMAV	170	1TN3_A	The c-type lectin carbohydrate recognition domain of human tetranectin
		1ZU4_A	Crystal structure of FtsY from Mycoplasma mycoides- space group P21212.
RING domain of wsv199	133	1V87_A	Solution Structure of the Ring-H2 Finger Domain of Mouse Deltex Protein 2.
		2CT2_A	Solution Structure of the RING domain of the Tripartite motif protein 32

Table 4 shows the templates chosen for the modeling process. Multiple template modeling was performed using modeler 9v10. In order to compare the modeling results threading alignment based server I-TASSER was used. It uses 10 most appropriate templates to generate the 3D coordinates. Amongst all the models, best structure was selected on the basis of DOPE score as well as Ramachandran plot. Ramachandran plot was created using PROCHECK in order to confirm the reliability and excellence of the generated models, by evaluating the backbone conformation, angles and bond lengths (Fig. 17).

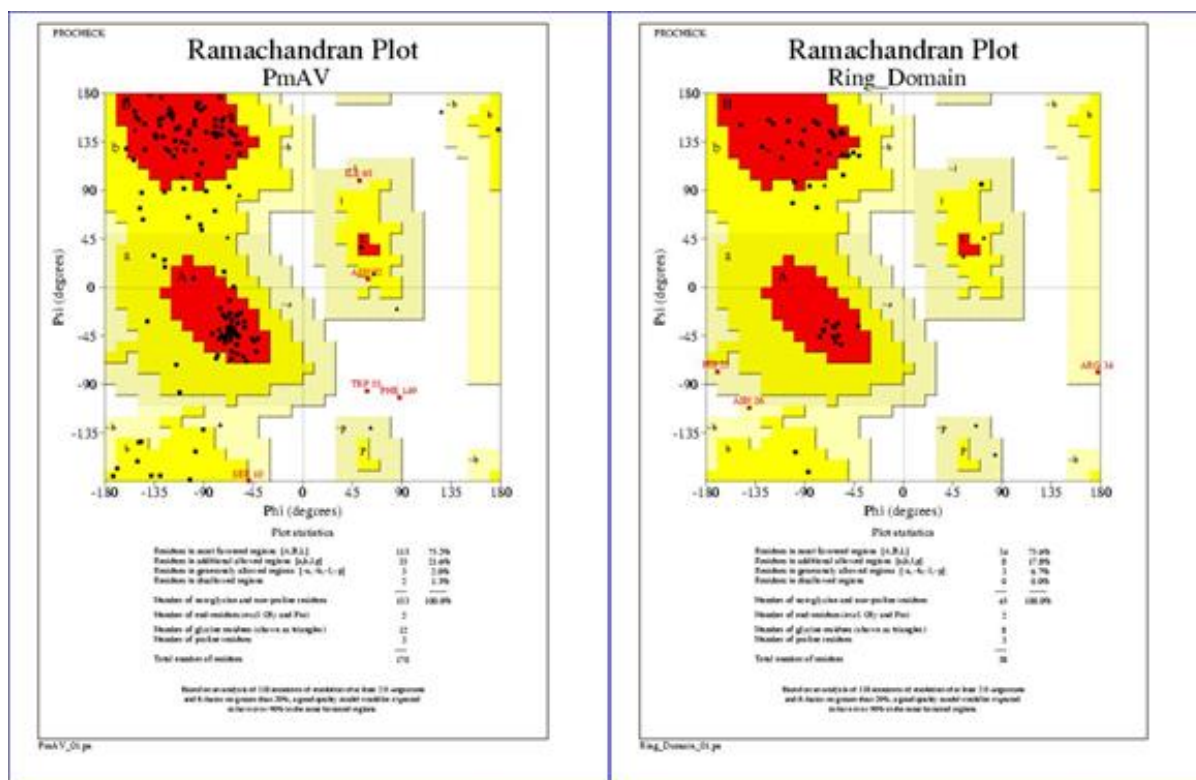


Figure 17.a and b shows the Ramachandran plots for the PmAV and RING domain of wsv199, respectively. It can be seen that the models generated fit very well for a good model. Overall residue energies are all well within the limits. Analysis of each model with ProSA web interface brings out -3.63 and -4.85 Z-Score for PmAV and RING domain, respectively. The z-score indicates overall model quality and measures the deviation of the total energy of the structure with respect to an energy distribution derived from random conformations (Wiederstein and Sippl, 2007). The Z-score value of -3.363 and -4.85 obtained fits well within the array of high quality models.

The RMSD trajectories showed that PmAV and RING DOMAIN became stable at 7.7 ns and 6.0 ns respectively (Fig. 18).

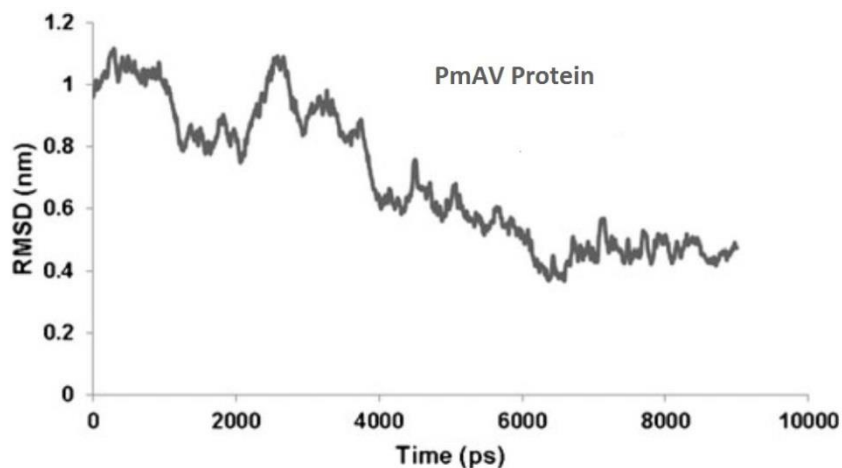


Figure 18: RMSD graph of PmAV.

RMSD values for PmAV increased from 0.99 nm to 1.044 nm at 2.4 ns followed by atomic fluctuation within protein that leveled off around 7.7 ns (0.51 nm) and then demonstrated a stable trajectory between 0.49 nm and 0.51 nm. The RMSD value for RING DOMAIN starts from 0.001 nm and increases to 0.19 nm around 1.3 ns followed by sharp increase of 0.26 nm at 4.8 ns and decrease to 0.13 nm at 5.6 ns and finally becomes stable after 6.0 ns between 0.17 nm–0.18 nm (Fig. 19).

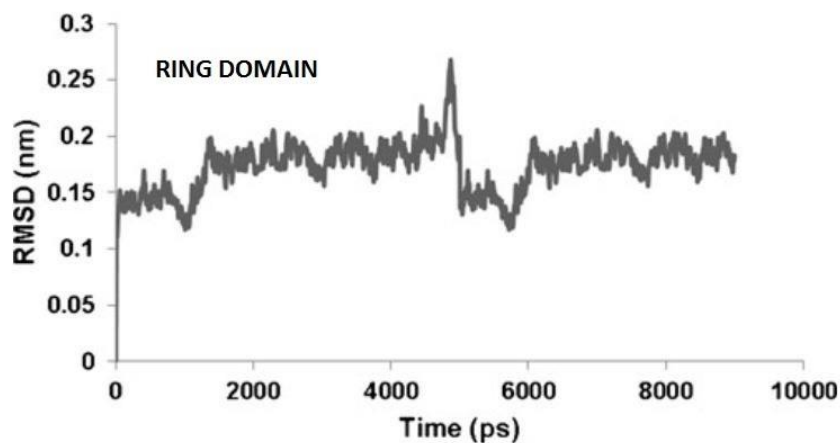


Figure 19 : RMSD graph of Ring Domain.

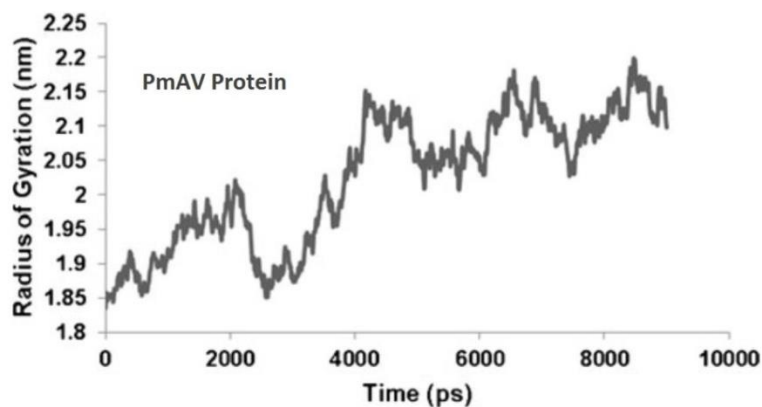


Figure 20: Radius of gyration graph of PmAV.

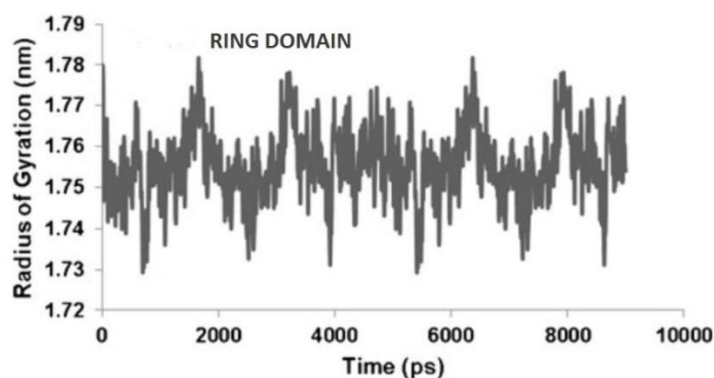


Figure 21: Radius of gyration graph of RING DOMAIN.

RING DOMAIN maintained its compactness during the entire simulation while PmAV faced minor changes, but overall compactness was maintained throughout simulation. Initially the Rg trajectory for PmAV was more compact showing dimension of 1.8 nm which steadily increased to 1.95 nm at 1.8 ns and elevated to 2.01 nm at 1.9 ns and finally attained stability after 6.3 ns between 2.0 nm and 2.2 nm (Fig. 20). Rg profile generated for RING DOMAIN is quite stable between 1.73 nm and 1.78 nm (Fig. 21). At the end of simulation, we obtained the stable conformations of the proteins that served as optimized input for the docking algorithm.

Protein-Protein interaction study by Docking

The molecular mechanism, the dynamic behavior of proteins and its interactions in this shrimp disease is still unknown. In the present study, we applied molecular modelling

and docking to study the nature and interaction between PmAV and wssv199 proteins (Fig. 22). Docking study was performed using PatchDock online server (cluster RMSD 4.0). 20 different solutions with their score, area and transformation file were obtained. The output file is applied to another online server ie. FireDock. This server was used to refinement the docking interaction. FireDock gave the 10 best results on the basis of minimum Global, Attractive and Repulsive Van-der wall, Atomic contact and Hydrogen bond energies as shown Table 5. In order to analyze the backbone induced-fit refinement of docking interactions, FiberDock server was used. Rigid body protein-protein solutions, in the form of protein files and transformations file, were taken as an input and 1 refined solution file was obtained as a result.

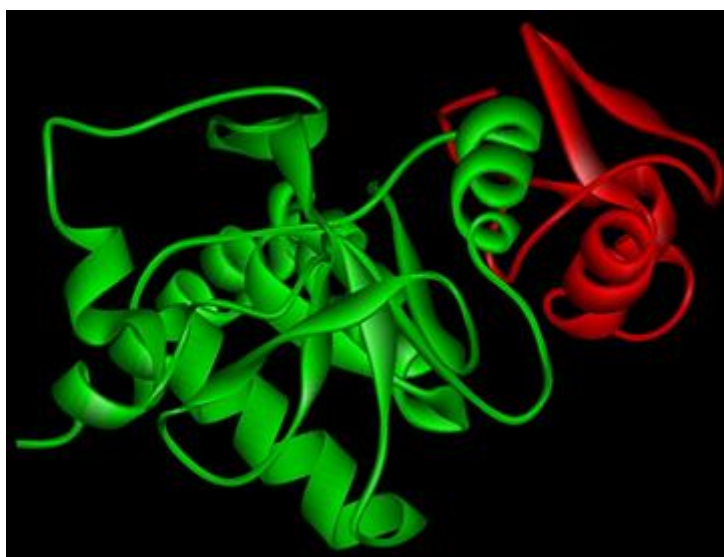


Figure 22: Docking interaction between PmAV (green) and RING domain of wssv199 (red).

This result file was used to generate molecular interaction plots certifying the protein-protein interaction between protein complexes, using DIMPLOT tool. Fig. 4 illustrates the interaction between the two proteins. Residues with red circle belong to PmAV protein and residues with pink circle belong to RING domain.

Table 5: FireDock Results.

Rank	Sol. Number	glob	aVdW	rVdW	ACE	HB
		↓				
1	10	0.00	0.00	0.00	0.00	0.00
2	3	0.00	0.00	0.00	0.00	0.00
3	4	0.00	0.00	0.00	0.00	0.00
4	5	0.00	0.00	0.00	0.00	0.00

5	7	0.00	0.00	0.00	0.00	0.00
6	8	0.00	0.00	0.00	0.00	0.00
7	9	0.00	0.00	0.00	0.00	0.00
8	6	3.41	-9.83	3.15	3.88	0.00
9	1	3.90	-9.26	15.03	1.51	-2.77
10	2	2799.19	-22.64	3558.90	-8.39	-3.16

glob - Global Energy, the binding energy of the solution

aVdW, rVdW - softened attractive and repulsive van der Waals energy

ACE - atomic contact energy (ACE)

HB - hydrogen and disulfide bonds

Transformation - ligand transformation after refinement

Molecular dynamics of the complex

The PmAV and RING DOMAIN complex with least binding energy was further used for carrying out MD. RMSD was calculated for every single backbone atoms, Rg, electrostatic energy, van der Waals energy of PmAV–RING DOMAIN complex were calculated in the form of MD trajectories. RMSD profiles constantly stayed not as much of 0.6 nm for the whole simulation. The RMSD value for the PmAV–RING DOMAIN complex was larger from 0.059 nm to 0.41 nm at 4.1 ns, further constantly augmented to reach 0.51 nm value at 10 ns and eventually achieve 0.52 nm nearby 11 ns illustrating a stable RMSD profile throughout the simulation (Fig. 23).

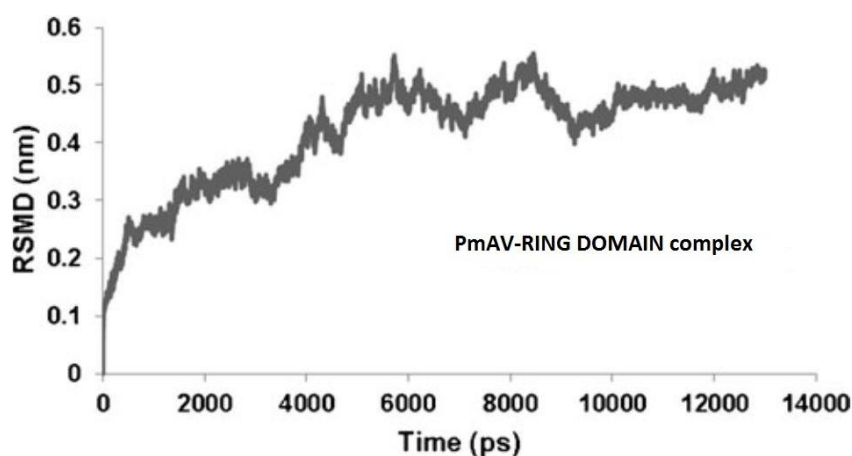


Figure 23: RMSD graph of PmAV-ring domain complex.

The continual trajectory illustrated a stable complex formation which consecutively showed strong bonding amongst both the proteins. Rg of PmAV-RING DOMAIN complex was examined to define its compactness. Rg value of primary complex configuration is 2.25 nm

after that decrement in value to 2.14 nm around 7 ns. Afterward the Rg further reduced to reach a constant value of 2.13 nm around 10 ns (Fig. 24).

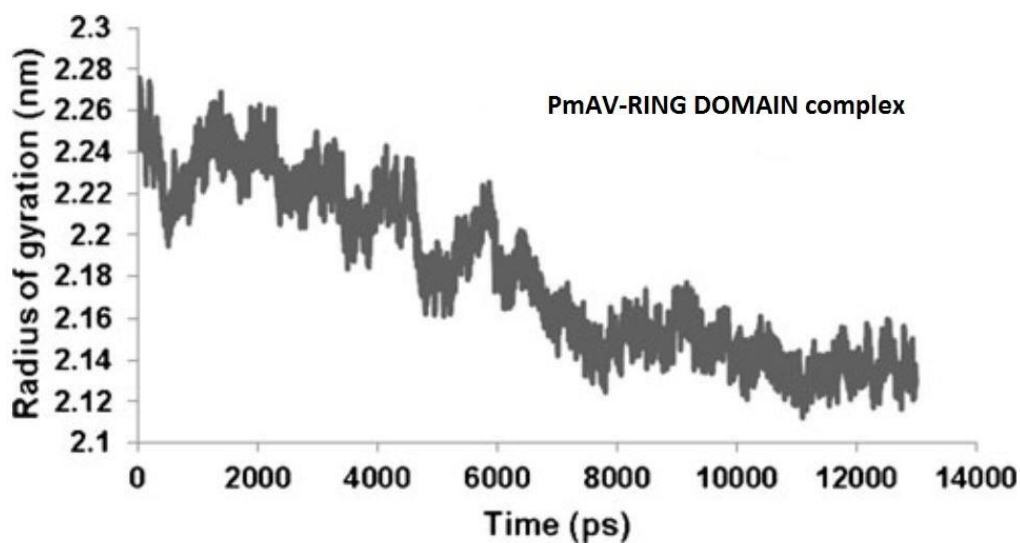


Figure 24: Radius of gyration graph of docked comple.

The trajectory displays that preliminary binding between both the proteins is loose then with progress complex turn into more compact in turn signifying the close-fitting bonding between the PmAV and Ring Domain. To get a deeper understanding into the constancy as well as non-bonded energies of complex van der Waals energy was determined that ranged from -459.2 to -577.49 kJmol^{-1} (Fig. 25).

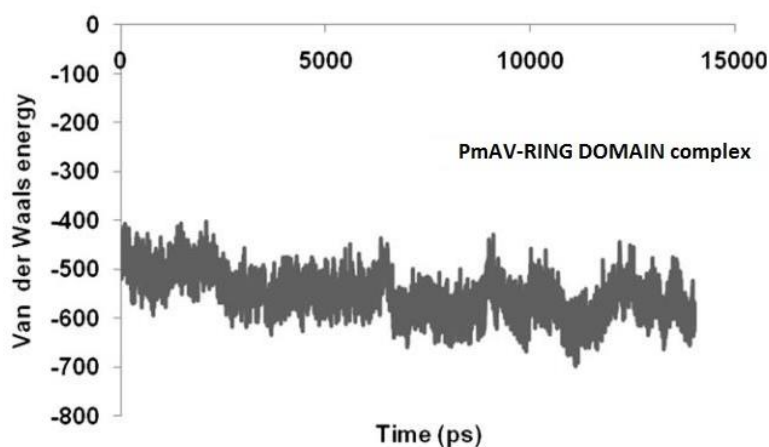


Figure 25: VanderWaals energy graph of complex.

The trajectory intended for van der Waals energy is more or less constant during simulation. Short-range electrostatics of the complex was delivered using electrostatic energy that extended from -341.79129 to -648.61853 kJmol^{-1} (Fig. 26).

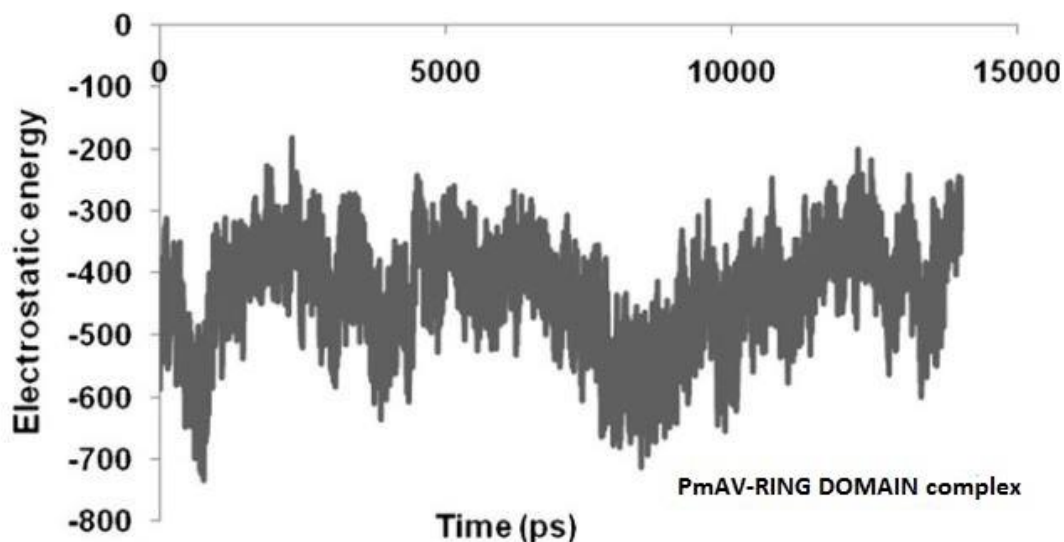
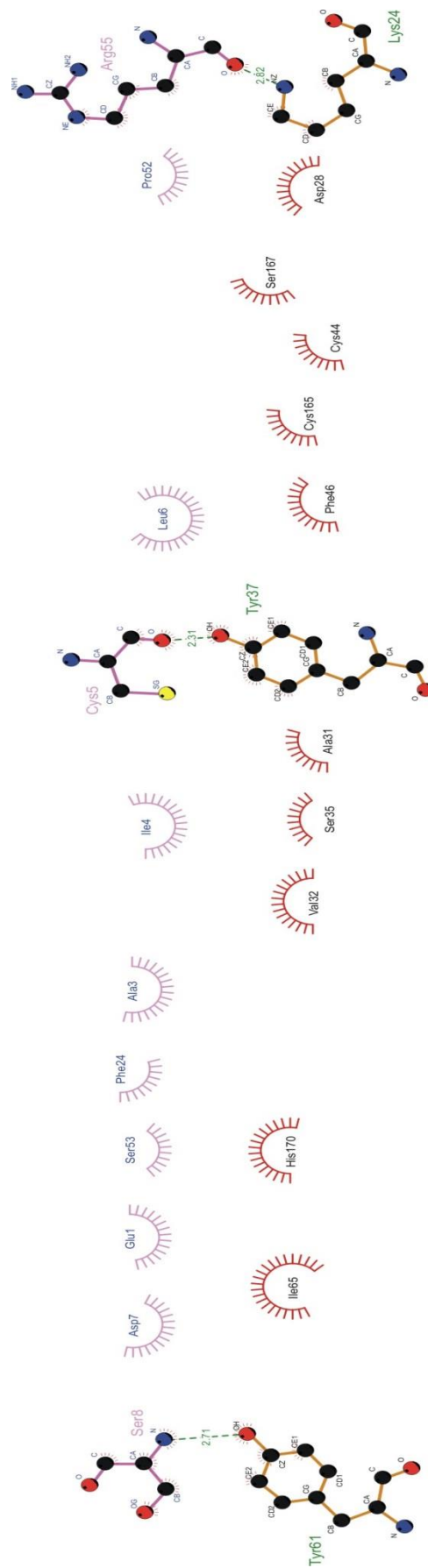


Figure 26: Electrostatic energy graph of complex.

Before the simulation of the complex 3H-bonds were drawn among both the proteins, but by the start of simulation the total of H-bonds amplified to 9 and π - π interactions could also be outlined sandwiched inside the complex. With rise in time the total number of H-bonds increased to 17 at 4 ns. In addition, fluctuation in tendencies of H-bond and π - π formation was detected up to 7 ns, 19 H-bonds and 3 π - π connections were perceived. Later 9 ns of simulation, a slight decrease in arrangement of H-bond and π - π bond formation were perceived. Subsequently, the number of H-bonds reached between 9 and 11 and the π - π connections were mark out. The outcomes altogether represented that, the amount of H-bonds and π - π connections were nearly comparable at start and end of simulation, but this complex was found more constant at the end of simulation as recommended by stable RMSD and Rg values.

In contrast, the RMSD and Rg graph in the commencement of simulation propose that primarily the complex was not stable. When H-bonds, π - π interaction, RMSD and Rg trajectory were considered about 7 ns, all four constraints signified constancy of the complex. The higher value of H-bonds as well as π - π interaction also suggested that both complement proteins interacted closely to each other at this instant of simulation.



complexT

Figure 27: Interaction between two chains.

6. DISCUSSION

The present study offers an appropriate 3D model of PmAV protein that provides a first participation for PmAV-RING DOMAIN complex formation. When this 3D model is aligned to the 3D coordinates of its template illustrated finest structural superimposition as well as minimal RMSD proposes a strong homology between both of proteins. The 3D model of PmAV was again confirmed for stereo-chemical constraints and overall structure geometry by different energy validation servers and found to fit well into the acceptance criteria of these servers, consequently implying that the 3D structure generated is excellent mutually in terms of structure as well as energy factors.

MD of PmAV and RING DOMAIN were prepared to know individual behavior of PmAV and RING DOMAIN at dynamic conditions. RMSD values for PmAV and RING DOMAIN implied that while both proteins experience fluctuations at the start but at final stage of simulation the trajectories turn into stable. Rg value for PmAV raises with time and stabilizes after some period denoting that with increase in time the dimension of PmAV amplifies but the protein retains its shape after that specific time. Rg of RING DOMAIN stays almost constant all through the simulation thus, signifying that RING DOMAIN continues of equal size throughout the simulation. The outcome from Procheck explains the residues in the allowed region of the Ramachandran plot and Errat illustrates the information of non-bonded contacts between different atom types. VERIFY3D hypothesizes the compatibility of the 3D atomic structure with its own amino acid sequence. The energy validations of the proteins ensure that both are energetically favorable and hence serve as an optimized input to the docking experiment.

Docking and MD tools have driven out an effective way for calculating the binding sites participated in protein–protein docking. The effectiveness of docking algorithm has been authenticated by taking unprocessed data from already published work and further running the same docking method. The concordant outcomes have been attained in both cases suggesting the capability of PatchDock as an excellent docking server. The docking algorithm uses the surface complementarities of amino acid residues nearby in the binding pocket based on how the scoring functions rank the a variety of docking outputs.

The most excellent docked conformation is evaluated on the basis of most favorable geometric score, interface area size and desolvation energy produced by the complex. The *insilico* docking approaches based on surface geometry complementarities and amino acid pairwise affinities between PmAV and RING DOMAIN predict formation of hydrogen bond arrangements between amino acid residues lying in the edge of both proteins.

7. CONCLUSION AND FUTURE PERSPECTIVE

The *in silico* study suggested that PmAV might interact with WSSV RING finger domain containing proteins and inhibits their activity thus preventing the establishment of viral infection. Additional studies similar to molecular dynamics simulation (26, 27) of the protein-protein complex are found to be useful for revealing new aspects for prevention of WSD. This can further lead to reduced modification of host's ubiquitination pathway resulting in decreased diseased condition. However, detailed *in vitro* and *in vivo* studies will be continued to elucidate the specific anti-viral activity of PmAV.

8. REFERENCES

Andrusier, N;Nussinov, R;Wolfson, HJ; (2007).FireDock: Fast Interaction Refinement in Molecular Docking. *Proteins: Structure, Function, and Bioinformatics*, Volume 69, 139–159.

Berendsen, HJC; VanderSpoel, D; VanDrunen, R; (1995). GROMACS—A message passing parallel molecular dynamics implementation. *Computer Physics Communications*.91, 43–56.

Bombyxmori is a novel member of the C-type lectin superfamily with two different tandem carbohydrate-recognition domains.*FEBS Letters*.443(2),139-43.

Cerenius, L; Soderhall, K;(2004). The prophenoloxidase-activating system in invertebrates.*Immunological Reviews*. 198, 116–26.

Clett, Eriddge; Elliot Bennett, Guerrero; Ian R, Poxton ;(2002).Structure and function of lipopolysaccharides.*Microbes and Infection*. 4(8), 837-851.

Cambi, A; Koopman, M;Figdor, CG ;(2005).How C-type lectins detect pathogens.*Cellular microbiology*.7 (4), 481-8.

Dina, SD;Inbar,Y;Nussinov,R; Wolfson, HJ; (2005). PatchDock and SymmDock: servers for rigid and symmetric docking. *Nucleic Acids Research*.33, W363-W367.

Drickamer, K;(1993). Ca²⁺ -dependent carbohydrate-recognition domains in animal proteins. *Current Opinion in Structural Biology*.3, 393–400.

Drickamer, K ;(1988).Two distinct classes of carbohydrate-recognition domains in animal lectins.*Journal of Biological Chemistry*.263 (20), 9557-60.

Drickamer, K ;(1999).C-type lectin-like domains.*Current Opinion in Structural Biology*.9 (5), 585-90.

Eswar, N;Eramian, D; Webb, B;Shen, MY;Sali, A; (2008).Protein structure modeling with MODELLER.*Methods in Molecular Biology*. 426, 145-159.

Feizi, T;(2000).Carbohydrate-mediated recognition systems in innate immunity. *Immunological Reviews*.173, 79–88.

Gajula, P;Borovykh, IV;Beier,C;Shkuropatova, T;Gast, P; Steinhoff, HJ; (2007). Spin-labeled photosynthetic reaction centers from Rhodobactersphaeroides studied by electron paramagnetic resonance spectroscopy and molecular dynamics simulations. *Applied Magnetic Resonance*.31 (1), 167-178:8.

Guex, N;Peitsch, MC; (1997). SWISS-MODEL and the Swiss-PdbViewer: An environment for comparative protein modeling. *Electrophoresis* 18(15), 2714-23.

Gross, PS; Bartlett TC; Browdy, CL; Chapman, RW; Warr, GW ;(2001).Immune gene discovery by expressed sequence tag analysis of hemocytes and hepatopancreas in the Pacific White Shrimp, *Litopenaeusvannamei*, and the Atlantic White Shrimp, *L. setiferus*.*Developmental & Comparative Immunology*.25 (7), 565-77.

He, F;Fenner, B; Godwin J; A, K; Kwang, J;(2006). White spot syndrome virus open reading frame 222 encodes a viral E3 ligase and mediates degradation of a host tumor suppressor via ubiquitination, *J. Virology*.80, 3884–3892.

Hess, B;Bekker, H;Berendsen, HJC;Fraaije, JGEM; (1997). LINCIS: A linear constraint solver for molecular simulations. *Journal of Computational Chemistry*.16:273–284

Hoffmann, JA; Kafatos, FC; Janeway, CA; Ezekowitz, RA; (1999).Phylogenetic perspectives in innate immunity.*Science*. 284, 1313–1318.

Koizumi, N; Imamura, M;Kadotani, T;Yaoi, K;Iwahana, H; Sato, R;(1999).The lipopolysaccharide-binding protein participating in hemocyte nodule formation in the silkworm

Lightner, DV; Redman, RM;(1998).Shrimp diseases and current diagnostic methods.*Aquaculture*.164, 201–220.

Lo, CF; Ho, CH; Peng, SE; Chen, CH; Hsu, HC; Chiu, YL; Chang, CF; Liu, KF; Su, MS; Wang, CH; Kou, GH; (1996). White spot syndrome baculovirus detected in cultured and captured shrimp, crabs and other arthropods. *Diseases of Aquatic Organisms*.27, 215–225.

Luo, T; Zhang, X; Shao, Z;Xu, X;(2003).*PmAV*, a novel gene involved in virus resistance of shrimp *Penaeusmonodon*. *FEBS Letters*. 551, 53–57.

Laskowski, RA; MacArthur, MW; Moss, DS; Thornton, JM; (1993). PROCHECK: a program to check the stereo chemical quality of protein structures. *Journal of Applied Crystallography*. 26, 283-291

Laskowski, RA; Swindells, MR; (2011) LigPlot+: multiple ligand-protein interaction diagrams for drug discovery. *Journal of Chemical Information and Modeling*, 51 (10), 2778–2786.

Liu, YC; Li, FH; Dong, B; Wang, B; Luan, W; Zhang, XJ , Zhang, LS; Xiang, JH;(2007). Molecular cloning, characterization and expression analysis of a putative C-type lectin (Fclectin) gene in Chinese shrimp *Fenneropenaeuschinensis*. *Molecular Immunology*.44 (4), 598-607.

Luo, T; Yang, H; Li, F; Zhang, X;Xu,X;(2006).Purification, characterization and cDNA cloning of a novel lipopolysaccharide-binding lectin from the shrimp *Penaeus monodon*.*Developmental & Comparative Immunology*.30 (7), 607-17.

Luo, T; Li, F; Lei, K;Xu, X ;(2007).Genomic organization, promoter characterization and expression profiles of an antiviral gene PmAV from the shrimp *Penaeus monodon*.*Molecular Immunology*.44 (7), 1516-23.

Mashiach, E; Dina, SD; Andrusier, N; Nussinov, R; Wolfson HJ; (2008). FireDock: A web server for fast interaction refinement in molecular docking.*Nucleic Acids Research*.36, W229-W232.

Mashiach, E;Nussinov, R;Wolfson HJ; (2009) FiberDock: Flexible induced-fit backbone refinement in molecular docking. *Proteins: Structure, Function, and Bioinformatics*. Volume 78, 1503–1519.

Mashiach, E; Nussinov, R; Wolfson HJ; (2010) FiberDock: A web server for flexible induced-fit backbone refinement in molecular docking. *Nucleic Acids Research*.38, W457-W461.

Miyamoto, S;Kollman, PA; (1992). SETTLE: An analytical version of the SHAKE and RATTLE algorithms for rigid water models. *Journal of Computational Chemistry*.13:952–962

PrasadGajula,MNV; Vogel, KP; Rai, A; Dietrich, F; Steinhoff, HJ; (2013).How far *in-silico* computing meets real experiments. A study on the structure and dynamics of spin labeled vinculin tail protein by molecular dynamics simulations and EPR spectroscopy. BMC Genomics. 14, (Suppl 2):S4.

Rahman, MM;(2007). Difference in Virulence between White Spot Syndrome Virus (WSSV) Isolates and Testing of Some Control Strategies in WSSV Infected Shrimp, Doctoral Thesis, Department of Virology, Parasitology and Immunology, Faculty of Veterinary Medicine, Ghent University, pp. 12.

Rajesh Kumar, S;Venkatesan, C;Sarathi, M;, Sarathbabu, V; Thomas, J;AnverBasha, K;, SahulHameed, AS;(2009).Oral delivery of DNA construct using chitosan nanoparticles to protect the shrimp from white spot syndrome virus (WSSV).Fish and Shellfish Immunology.26(3),429-37.

Schneidman-Duhovny ,D;Inbar, Y;Polak, V;Shatsky, M; Halperin, I; Benyamini, H;Barzilai, A;Dror, O;Haspel, N;Nussinov, R;Wolfson, HJ;(2003).Taking geometry to its edge: fast unbound rigid (and hinge-bent) docking.Proteins.52 (1), 107-12.

Söderhäll, K; Cerenius, L ;(1998).Role of the prophenoloxidase-activating system in invertebrate immunity.Current Opinion in Immunology.10 (1), 23-8.

Sheriff, S; Chang, CY; Ezekowitz, RA; (1994).Human mannose-binding protein carbohydrate recognition domain trimerizes through a triple alpha-helical coiled-coil. Nature Structure Biology.1 (11),789-94.

Tonganunt , M;Nupan, B; Saengsakda, M;Suklour, S; Wanna, W;Senapin, S; Chotigeat, W; Phongdara, A;(2008).The role of Pm-fortilin in protecting shrimp from white spot syndrome virus (WSSV) infection.Fish and Shellfish Immunology.25 (5), 633-7.

Van Hulten, MC; Witteveldt, J; Snippe, M; Vlak, JM; (2001). White spot syndrome virus envelope protein RING DOMAIN is involved in the systemic infection of shrimp. Virology. 285, 228–233.

Verma, S; Singh, A; Mishra, A; (2012). Dual inhibition of chaperoning process by taxifolin: molecular dynamics simulation study. Journal of Molecular Graphics & Modelling.37:27–38

Verma, S; Singh A; Mishra, A; (2012). The effect of fulvic acid on pre- and postaggregation state of A β (17–42): molecular dynamics simulation studies. *Biochimica et Biophysica Acta*.1834 (1), 24-33.

Vriend, G; (1990). WHAT IF: A molecular modeling and drug design program. *Journal of Molecular Graphics*.8, 52–56.

Wallace,AC; Laskowski, RA; Thornton, JM; (1995). LIGPLOT: A program to generate schematic diagrams of protein-ligand interactions. *Protein Engineering*.8 (2), 127-134.

Wang, Z; Chua, HK; Gusti, AA; He, F; Fenner, B; Manopo, I; Wang, H; Kwang J; (2005).RING-H2 protein WSSV249 from white spot syndrome virus sequesters a shrimp ubiquitin-conjugating enzyme, PvUbc, for viral pathogenesis. *Journal of Virology*.79 (14), 8764-72.

Weis, WI; Taylor, ME;Drickamer, K;(1998).The C-type lectin superfamily in the immune system.*Immunological Reviews*.163, 19–34.

Yang, F; He, J; Lin, X; Li, Q; Pan, D; Zhang, X; Xu X;(2001). Complete genome sequence of the shrimp white spot bacilliform virus. *Journal of Virology*.75, 11811–11820.

Yang, Zhang ;(2008) I-TASSER server for protein 3D structure prediction.*BMC Bioinformatics* 9, 40.

Yu, XQ; Kanost, MR; (2004).Immulectin-2, a pattern recognition receptor that stimulates hemocyte encapsulation and melanization in the tobacco hornworm, *Manduca sexta*.*Developmental & Comparative Immunology*. 28, 891–900

Yu, XQ;Gan, H; Kanost,MR;(1999).Immulectin, an inducible C-type lectin from an insect, *Manduca sexta*, stimulates activation of plasma prophenoloxidase.*Insect Biochemistry and Molecular Biology*.29(7),585-97

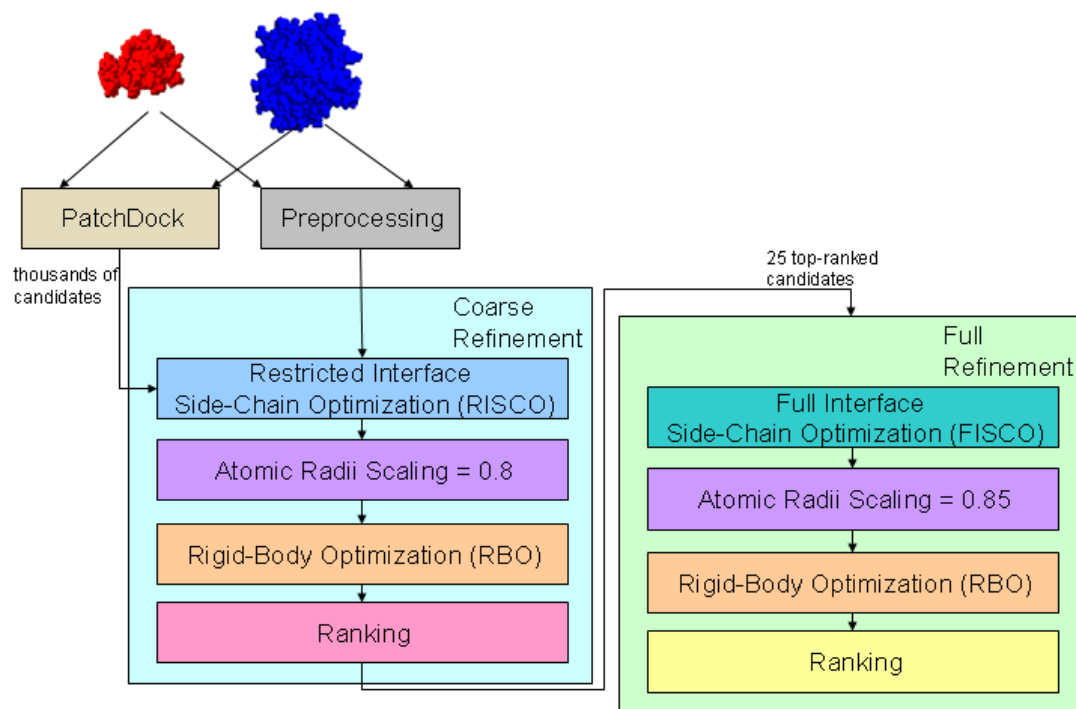
9. APPENDIX

Fire-Dock : FireDock is an efficient method for the refinement and rescoring of rigid-body docking solutions. The refinement process consists of two main steps:

- (1) rearrangement of the interface side-chains and
- (2) adjustment of the relative orientation of the molecules.

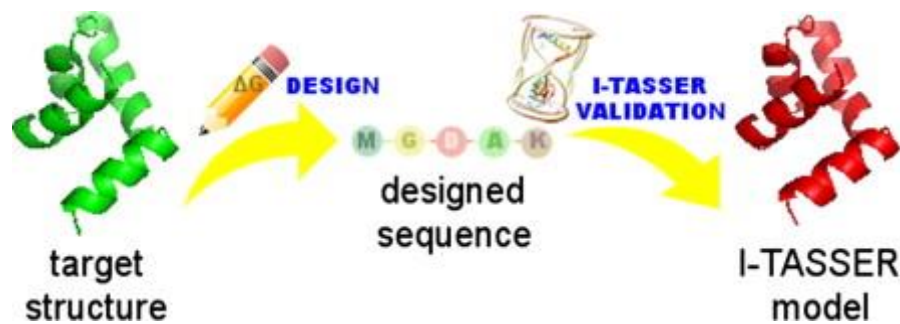
This method accounts for the observation that most interface residues that are important in recognition and binding do not change their conformation significantly upon complexation. Allowing full side-chain flexibility, a common procedure in refinement methods, often causes excessive conformational changes. These changes may distort preformed structural signatures, which have been shown to be important for binding recognition. This server restricts side-chain movements, and thus manages to reduce the false-positive rate noticeably.

In the later stages (orientation adjustments and scoring), it smooths the atomic radii. This allows for the minor backbone and side-chain movements and increases the sensitivity of the algorithm. FireDock succeeds in ranking a near-native structure within the top 15 predictions for 83% of the 30 enzyme-inhibitor test cases, and for 78% of the 18 semiunbound antibody-antigen complexes. This refinement procedure significantly improves the ranking of the rigid-body PatchDock algorithm for these cases. The FireDock program is fully automated. The method is available at <http://bioinfo3d.cs.tau.ac.il/FireDock/>.



The figure below shows an example of FireDock usage. Transformations of docking candidates are generated by PatchDock and are given as an input to FireDock. First a coarse refinement is performed, using a restricted interface side-chain optimization with atomic radii scaling of 0.8, in order to allow a certain amount of steric clashes. The refined candidates are scored and ranked according to the energy function and are returned as an output. Then, FireDock is run again on the best 25 solutions for a final refinement. In this second run, a full interface side-chain optimization is performed with atomic radii scaling of 0.85, in order to reduce the amount of clashes.

I-Tasser: The I-TASSER server has been developed to generate automated full-length 3D protein structural predictions where the benchmarked scoring system helps users to obtain quantitative assessments of the I-TASSER models. The output of the I-TASSER server for each query includes up to five full-length models, the confidence score, the estimated TM-score and RMSD, and the standard deviation of the estimations. The I-TASSER server is freely available to the academic community at <http://zhang.bioinformatics.ku.edu/I-TASSER>.



GROMACS: The software suite GROMACS (Groningen MACHine for Chemical Simulation) was developed at the University of Groningen, The Netherlands, in the early 1990s. The software, written in ANSI C, originates from a parallel hardware project, and is well suited for parallelization on processor clusters. By careful optimization of neighbor searching and of inner loop performance, GROMACS is a very fast program for molecular dynamics simulation. It does not have a force field of its own, but is compatible with GROMOS, OPLS, AMBER, and ENCAD force fields.

In addition, it can handle polarizable shell models and flexible constraints. The program is versatile, as force routines can be added by the user, tabulated functions can be specified, and analyses can be easily customized. Nonequilibrium dynamics and free energy determinations are incorporated. Interfaces with popular quantum-chemical packages (MOPAC, GAMES-UK, GAUSSIAN) are provided to perform mixed MM/QM simulations. The package includes about 100 utility and analysis programs. GROMACS is in the public domain and distributed (with source code and documentation) under the GNU General Public License. It is maintained by a group of developers from the Universities of Groningen, Uppsala, and Stockholm, and the Max Planck Institute for Polymer Research in Mainz. Its Web site is <http://www.gromacs.org>.