Molecular Cloning and Characterization of Lectin and Lipopolysaccharide and β-1,3-Glucan Binding Protein Genes from Banana Shrimp, *Fenneropenaeus merguiensis*

Onnicha Rattanaporn

A Thesis Submitted in Partial Fulfillment of the Requirements

for the Degree of Doctor of Philosophy in Biochemistry

Prince of Songkla University

2009

Thesis Title	Molecular Cloning and Characterization of Lectin and Lipopolysaccharide and β-1,3-Glucan Binding Protein Genes from Banana Shrimp, *Fenneropenaeus merguiensis*
Author	Miss Onnicha Rattanaporn
Major Program	Biochemistry

Major Advisor

……

(Assoc. Prof. Dr. Prapaporn Utarabhand)

Co-advisor

……

(Assoc. Prof. Dr. Porntip Prapunpoj)

Examining Committee:

……..Chairperson

(Assoc. Prof. Dr. Rapepun Wititsuwannakul)

…….

(Assoc. Prof. Dr. Prapaporn Utarabhand)

…….

(Assoc. Prof. Dr. Porntip Prapunpoj)

…….

(Prof. Dr. Vichai Boonsaeng)

The Graduate School, Prince of Songkla University, has approved this thesis as partial fulfillment of the requirements for the Degree of Doctor of Philosophy in Biochemistry

……

(Assoc. Prof. Dr. Krerkchai Thongnoo)

Dean of Graduate School

ชื่อวิทยานิพนธ์	การโคลนและศึกษาคุณสมบัติของยีนเลคตินและโปรตีนที่จับจำเพาะกับลิโพโพลีแซคคาไรด์และเบตา-1,3-กลูแคนจากกุ้งแชบ๊วย
ผู้เขียน	นางสาวอรณิชา รัตนาภรณ์
สาขาวิชา	ชีวเคมี
ปีการศึกษา	2551

บทคัดย่อ

เลคตินจากฮีโมลิมฟ์เป็นโมเลกุลที่มีศักยภาพซึ่งมีความเกี่ยวข้องกับการจดจำของระบบภูมิคุ้มกันและกระบวนการกลืนกินจุลินทรีย์ (microorganism phagocytosis) ผ่านการเกิดออปโซไนซ์ (opsonization) ในสัตว์ไม่มีกระดูกสันหลัง (invertebrate) รวมทั้งในครัสเตเชียน (crustacean) การเกิดออปโซไนซ์เป็นขั้นตอนแรกที่ส่งเสริมการยึด การย่อย และตามด้วยการกำจัดจุลินทรีย์ที่เกิดขึ้นเมื่อเลคตินจับกับแบคทีเรียและผิวของเซลล์เม็ดเลือด (hemocyte) การศึกษาครั้งนี้ได้ทำให้เลคตินบริสุทธิ์จากจากฮีโมลิมฟ์ของกุ้งแชบ๊วยโดยการแยกผ่านคอลัมน์ Fetuin-agarose 2 ครั้ง พบว่ามีค่าความจำเพาะในการเกาะกลุ่มเม็ดเลือดแดงเพิ่มขึ้นเป็น 2,890 เท่าของซีรัมเริ่มต้น เลคตินบริสุทธิ์มีมวลโมเลกุล 316.2 กิโลดัลตัน เมื่อวิเคราะห์ด้วยโพลีอะคริลาไมด์เจลอิเล็กโทรโฟรีซิสแบบมีเอสดีเอส (SDS-PAGE) พบว่าเลคตินบริสุทธิ์ปรากฏ 2 หน่วยย่อยที่มีขนาด 32.3 และ 30.9 กิโลดัลตัน จากวิธี Edman degradation สามารถหาลำดับของกรดอะมิโน 10 หน่วย ทางด้านปลายอะมิโนของหน่วยย่อย 32.3 และ 30.9 กิโลดัลตัน พบว่ามีลำดับกรดอะมิโนเหมือนกันเป็น LAGAGTVLAG เมื่อหาลำดับกรดอะมิโนภายในสายเปปไทด์โดยวิธี matrix-assisted laser desorption ionization-mass spectrometry-time of flight analysis (MALDI-TOF) พบว่ามีความคล้ายคลึงกับส่วนอนุรักษ์ (conserved motif) ที่พบใน fibrinogen-related domains (FReDs) ของ โปรตีน ficolin

นอกจากนั้นการศึกษาครั้งนี้ได้ทำการโคลนและศึกษาคุณลักษณะของยีนที่มีความเกี่ยวข้องกับระบบภูมิคุ้มในกุ้งแชบ๊วย 2 ยีน คือยีนของเลคติน (lectin, FmL) และยีนของโปรตีนที่จับจำเพาะกับลิโพโพลีแซคคาไรด์และเบตา-1,3-กลูแคน (lipopolysaccharide and β-1,3-glucan binding protein, LGBP, FmLGBP) จาก ห้องสมุด cDNA (cDNA library) ที่เตรียมจากตับ (hepatopancreas) และเซลล์เม็ดเลือด (hemocyte) ของกุ้งแชบ๊วย ตามลำดับ โดยวิธี PCR (polymerase chain reaction) และวิธี RACE (rapid amplification reaction of cDNA ends) ผลการศึกษาสามารถโคลนยีน FmL และ FmLGBP เส้นเต็มได้

ความยาวของลำดับนิวคลีโอไทด์ทั้งหมด (full-length) ของ cDNA ของยีนเลคตินมีขนาด 1,118 คู่เบส ประกอบด้วยส่วนปลาย 5’ ที่ไม่แปลรหัส 42 คู่เบส ส่วนปลาย 3’ ที่

ไม่แปลรหัส 74 คู่เบส และมี open reading frame (ORF) ที่มีขนาด 1,002 คู่เบส ซึ่งถอดรหัสได้เป็นสายโพลีเปปไทด์ที่มีกรดอะมิโน 333 หน่วย มีมวลโมเลกุล 36.3 กิโลดัลตัน และมีค่า pI เป็น 4.14 ภายในโมเลกุลของ FmL พบโดเมนที่จดจำคาร์โบไฮเดรต (carbohydrate recognition domain, CRD) 2 โดเมน โดยพบที่กรดอะมิโนตำแหน่งที่ 36-169 และ 197-330 ทั้ง 2 โดเมนมีความเหมือนกัน 35% (มีความคล้ายกัน 47%) พบส่วนอนุรักษ์ QPD (QPD motif) อยู่ภายใน CRD แรก ขณะที่พบส่วนอนุรักษ์ EPN (EPN motif) อยู่ภายใน CRD ที่สอง เมื่อเปรียบเทียบลำดับกรดอะมิโนของเลคตินที่แยกได้จากกุ้งแชบ๊วยกับเลคตินจากกุ้ง *Penaeus monodon, Litopenaeus vanamei* และ *Penaeus semisulcatus* พบว่ามีความเหมือนกัน 82%, 68% และ 65% ตามลำดับ จากการศึกษาการแสดงออกของยีนเลคติน โดยวิธี RT-PCR (Reverse Transcription-PCR) พบการแสดงออกของยีนเลคตินในตับเท่านั้นและไม่พบการแสดงออกของยีนนี้ในเนื้อเยื่ออื่น ๆ ของกุ้งแชบ๊วย โปรตีน FmL สามารถแสดงออกได้ในแบคทีเรีย *Escherichia coli* สายพันธุ์ BL21(DE3)star โดยโปรตีนลูกผสม FmL ที่ผลิตได้ ถูกทำให้บริสุทธิ์ภายใต้ภาวะแปลงสภาพด้วยคอลัมน์ nickel-based affinity และมีมวลโมเลกุล 55 กิโลดัลตัน

ความยาวของลำดับนิวคลีโอไทด์ทั้งหมดของ cDNA ของยีน LGBP จากกุ้งแชบ๊วย (FmLGBP) มีขนาด 1,280 คู่เบส ประกอบด้วยส่วนปลาย 5’ ที่ไม่แปลรหัส 30 คู่เบส ส่วนปลาย 3’ ที่ไม่แปลรหัส 152 คู่เบส และมี ORF ขนาด 1,101 คู่เบส ซึ่งถอดรหัสได้เป็นสายโพลีเปปไทด์ที่มีกรดอะมิโน 366 หน่วย ที่ประกอบด้วยส่วน signal peptide ที่มีกรดอะมิโน 17 หน่วย มีมวลโมเลกุล 41.6 กิโลดัลตัน และมีค่า pI เท่ากับ 4.43 ภายในโมเลกุลของ FmLGBP พบส่วนอนุรักษ์ของ integrin-binding motif 2 บริเวณ ซึ่งมีลำดับกรดอะมิโนเป็น RGD (Arg-Gly-Asp) และบริเวณจดจำสำหรับโพลีแซคคาไรด์แบบ β-1,3-linkage นอกจากนั้น ยังพบบริเวณที่เติมน้ำตาลแบบ *N*-link (*N*-linked glycosylation site) ซึ่งมีกรดอะมิโนเป็น NRS และ NLS บ่งชี้ว่า LGBP เป็นไกลโคโปรตีน เมื่อเปรียบเทียบลำดับกรดอะมิโนของ LGBP ที่แยกได้จากกุ้งแชบ๊วย กับ LGBP จากกุ้ง *Fenneropenaeus chinensis,* BGBP จาก *P. monodon* และ LGBP จาก *Litopenaeus styrilostris* พบว่ามีความเหมือนกัน 96%, 93% และ 91% ตามลำดับ จากการศึกษาการแสดงออกของยีน LGBP โดยใช้วิธี RT-PCR พบการแสดงออกของยีน LGBP ทั้งในตับและเซลล์เม็ดเลือด โดยไม่พบการแสดงออกของยีนนี้ในเนื้อเยื่ออื่นๆ ของกุ้งแชบ๊วย

Thesis Title	Molecular Cloning and Characterization of Lectin and Lipopolysaccharide and β-1,3-Glucan Binding Protein Genes from Banana Shrimp, *Fenneropenaeus merguiensis*
Author	Miss Onnicha Rattanaporn
Major Program	Biochemistry
Academic Year	2008

Abstract

Lectins from the hemolymph of invertebrates, including crustaceans, have been regarded as potential molecules involved in immune recognition and microorganism phagocytosis through opsonization. Opsonization that occurs when a lectin binds bacteria to blood cell (hemocyte) surfaces is considered to be the first step that promotes adherence, ingestion and subsequent elimination of microbes. In the present study, a natural lectin from the hemolymph of banana shrimp *Fenneropenaeus merguiensis* was purified to homogeniety by repetitive affinity chromatography on Fetuin-agarose column with an increasing in hemagglutinating activity up to 2,890-fold of the initial serum. The molecular mass of purified lectin is 316.2 kDa and it showed two protein bands (32.3 and 30.9 kDa) in SDS-PAGE. Edman degradation indicated that the NH_2-terminal amino acid sequence of both subunits is identical as LAGAGTVLAG. The internal amino acid sequence of purified lectin was deduced from peptides obtained after tryptic treatment and by matrix-assisted laser desorption ionization-mass spectrometry-time of flight analysis (MALDI-TOF). The predicted amino acid sequence of the lectin is similar to conserved motifs present in fibrinogen-related domains (FReDs) of ficolin protein.

Moreover, two immune genes of *F. merguiensis*, were isolated and characterized. Hereby, this thesis reports the cloning and characterization of lectin and lipopolysaccharide binding protein and β-1, 3-glucan binding protein (LGBP) cDNAs from the hepatopacreas and haemocyte cDNA library of *F. merguiensis*, respectively. The cloning strategy was based on polymerase chain reaction (PCR) and rapid amplification of cDNA ends (RACE) methods. As a result, two full-length cDNAs of

F. merguiensis lectin (FmL) and FmLGBP genes were obtained.

The full length of *FmL* cDNA was 1118 bp in length with a 5' untranslated region of 42 bp and a 3' untranslated region of 74 bp. *FmL* cDNA contained an open reading frame (ORF) of 1,002 bp encoding a polypeptide of 333 amino acids with a calculated molecular mass of 36.3 kDa and pI value of 4.14. FmL consists of two predicted carbohydrate-recognition domains (CRDs) located at amino acid residues 36-169 and 197-330. The two CRDs shared 35% identity (or 47% similarity) with each other. The first CRD contains a QPD motif with specificity for galectose binding, while the second CRD contains an EPN motif for mannose binding. A sequence comparison showed that the deduced amino acids of *F. merguiensis* lectin has an overall similarity of 82%, 68% and 65% to those of tiger shrimp *Penaeus monodon*, white shrimp *Litopenaeus vannamei*, and green tiger shrimp *Penaeus semisulcatus*, respectively. The RT-PCR showed that FmL gene was expressed only in the hepatopancreas, with no detection in other tissues of the banana shrimp. FmL protein was successfully expressed in the *Escherichia coli* strain BL21(DE3)star expression system. The recombinant FmL protein purified under denaturing conditions using the nickel-based affinity chromatography showed a major band corresponding to the molecular mass of 55 kDa.

The full length of *FmLGBP* cDNA was 1,280 bp in length with a 5' untranslated region of 30 bp and a 3' untranslated region of 152 bp. The ORF of FmLGBP is 1,101 bp and encodes a 366 amino acid polypeptide with a signal peptide of 17 amino acid residues. Its calculated molecular mass is 41.6 kDa with an estimated pI of 4.43. Two putative integrin-binding motifs (cell adhesion site), RGD (Arg-Gly-Asp) and a potential recognition motif for β-(1,3) linkage of polysaccharides were observed in FmLGBP. Moreover, two putative *N*-linked glycosylation sites, NRS and NLS, for *N*-linked carbohydrate chains are present in FmLGBP, suggesting that LGBP of *F. merguiensis* is a glycoprotein. A sequence comparison showed that the deduced amino acids of FmLGBP has an overall similarity of 96%, 93% and 91% to *Fenneropenaeus chinensis* LGBP, tiger shrimp *P. monodon* BGBP, and *Litopenaeus stylirostris* LGBP, respectively. The RT-PCR showed that LGBP gene was expressed in both hemocytes and hepatopancreas, with no detection in other tissues.

Acknowledgments

I am very excited to finally write this page to express my gratitude to the people who contributed in many different ways to the accomplishment of this thesis, and without them this work would not have been possible. First and foremost, I would like to express my deepest gratitude to my advisor, Assoc. Prof. Dr. Prapaporn Utarabhand for giving me great opportunity to study in Ph.D. Program, correction and valuable suggestions in writing this dissertation. Her guidance, suggestions, constructive criticism and encouragement throughout my work in Department of Biochemistry, Faculty of Science, Prince of Songkla University have been invaluable. I am gratefully indebted to Assoc. Prof. Dr. Porntip Prapunpoj, my co-advisor, for her advice, guidance, and kindness through the course of this study. I really feel grateful to the examining committee, Assoc. Prof. Dr. Rapepun Wititsuwannakul and Prof. Dr. Vichai Boonsaeng.

I would like to express my appreciation to my other co-advisor, Prof. Dr. Jerry L. Hedrick and his lab members at the Deparment of Animal Science, University of California, Davis, USA, for their kind advice and suggestion through my laboratory work in USA. I would like to specially thank Dr. Gustavo A. Barisone for providing his expert guidance for the molecular cloning part of my thesis research work. I also would like to thank Prof. Dr. M.R. Jisnuson Svasti and his co-workers at Laboratory of Biochemistry, Biomedical Research Section, Chulabhorn Research Institute (CRI) for the use of equipment and their advice for N-terminal amino acid sequencing of lectin. I would like to express my great thanks to Dr. Kallaya Sritunyalucksana at Center of Excellence for Shrimp Molecular Biology and Biotechnology (CENTEX) for her valuable advice on the cDNA library construction.

I gratefully thank to the members of ST. 414 and all staff members of Department of Biochemistry, Faculty of Science, PSU, for all helps and best friendship. I would like to specially thank Dr. Wanida Rittidach for being a very good sister, nearby to help me during the difficult times and for her continued faith in me. I would like to thank P'noy for always taking care of me and help me everything while

I have been in USA. Many thanks for all of my loving friends; Toom, Tangmo, E, Nim, Nampeung for being great to me.

I am grateful for the generous financial supports from Thailand Research Fund (TRF) through The Royal Golden Jubilee Ph.D. Program, the Biochemistry Excellence Strengthening Program supported by Department of Biochemistry, Faculty of Science, and Graduate School, Prince of Songkla University through the Graduate Research Fund.

Last but certainly not the least, I would like to thank my father for being nearby and protecting me in spirit from heaven. My heartfelt and profound gratitude thank go to my family; my mother, my aunt and my sister for their endless love, understanding, encouragements and faith in me that helped me greatly particular in moments of doubt. You have all been there for me every step of the way, and I'm truly grateful. I couldn't have done it without your support. Love you always.

Onnicha Rattanaporn

Contents

List of Tables

List of Figures

List of Figures (continued)

List of Figures (continued)

List of Figures (continued)

List of Abbreviations and Symbols

A	=	absorbance
AMP	=	antimicrobial peptide
BGBP	=	β-1,3-glucan binding protein
bp	=	base pair
BSA	=	bovine serum albumin
CAPS	=	3-(cyclohexylamino)-1-propanesulfonic acid
°C	=	degree Celsius
cDNA	=	complementary deoxyribonucleic acid
CHAPS	=	3-[(3-cholamidopropyl)dimethylammonio]-1-propanesulfonate
CM-cellulose	=	carboxymethyl-cellulose
CRD	=	carbohydrate recognition domain
Da	=	Dalton
DEPC	=	diethyl pyrocarbonate
dNTP	=	deoxyribonucleotide triphosphate
DMSO	=	dimethylsulfoxide
dsRNA	=	double stranded ribonucleic acid
DTT	=	dithiothreitol
EDTA	=	ethylamine diamine tetraacetic acid
g	=	gram
x g	=	gravitation acceleration
GalNAc	=	N-acetyl galactosamine
GlcNAc	=	N-acetyl glucosamine
Glc	=	Glucose
GNBP	=	Gram negative bacteria-binding protein
h	=	hour
HA	=	hemagglutinating activity

List of Abbreviations and Symbols (Continued)

Ig	=	immunoglobulin
kb	=	kilobase
kDa	=	kilodalton
LB	=	Luria Bertani
LBP	=	Lipopolysaccharide binding protein
LGBP	=	Lipopolysaccharide binding protein and β-1,3-glucan binding protein
LPS	=	Lipopolysaccharide
M	=	molar
M_r	=	apparent molecular mass
mA	=	milliampere
MALDI-TOF	=	matrix-assisted laser desorption ionization time-of-flight
Man	=	mannose
Man-6-P	=	mannose-6-phosphate
ManNAc	=	N-acetyl mannosamine
MBP	=	mannose-binding protein
mg	=	milligram
min	=	minute
ml	=	milliliter
mm	=	millimeter
mM	=	millimolar
mRNA	=	messenger ribonucleic acid
nm	=	nanometer
NeuNAc	=	N-acetyl neuraminic acid
ORF	=	open reading frame
O.D.	=	optical density

List of Abbreviations and Symbols (Continued)

PAGE	=	polyacrylamide gel electrophoresis
PCR	=	polymerase chain reaction
pH	=	-log hydrogen ion concentration
pI	=	isoelectric pH
PMSF	=	phenylmethylsulfonylfluoride
PPAE	=	proPO activating enzyme
proPO	=	prophenoloxidase activating
PRP	=	pattern recognition protein
PVDF	=	polyvinyl difluoride
RACE	=	rapid amplification of cDNA ends
RNA	=	ribonucleic acid
RT-PCR	=	reverse transcription-polymerase chain reaction
rRNA	=	ribosomal ribonucleic acid
s	=	second
SDS-PAGE	=	sodium dodecyl sulfate-polyacrylamide gel electrophoresis
TBS	=	tris buffer saline
TEMED	=	N,N,N',N'-tetramethylethylenediamine
Tris	=	tris(hydroxymethyl)aminomethane
U	=	unit
µl	=	microliter
µg	=	microgram
V	=	voltage
WSSV	=	white spot syndrome virus
X-gal	=	5-bromo-4-chloro-3-indolyl-beta-D-galactopyraniside
%	=	percentage

Chapter 1

Introduction

Introduction

Shrimp, once regarded as a luxury food by most people, has become more affordable and available as demand for it steadily rises in the major consuming markets of Japan, the United States, and various European countries. Shrimp culture, especially penaeid shrimp, has shown to be the fastest growing section in aquaculture industry. The aquaculture of penaeid shrimp has rapidly grown to a major industry, which on a worldwide basis, provides not only economic income and a high quality food product, but also employment to hundreds of thousands of skilled and unskilled workers. The shrimp producing countries are mainly in Asia and South America. Thailand is the world's leading shrimp farming nation exporting the most shrimp in terms of volume and value. It is the top supplier of farmed shrimp to the US and Japan. However, the rapid expanding shrimp industry has faced problems since 1988. Diseases have emerged as a major constraint to the sustainable growth of shrimp aquaculture industry. Many diseases are linked to environmental deterioration and stress associated with culture intensification. The increase of disease problems that have devastated and continue to threaten production of several species throughout the world has emphasized the need to develop tools for the rapid recognition and control of pathogens. Disease prevention is more important than treatment. Studies towards a better understanding of defense mechanisms in shrimp constitute one approach to overcome disease problem.

In view of the economic importance of shrimp aquaculture and the number of pathological problems in industry is now suffering. Therefore, studying and developing tools for the rapid recognition and control of pathogens are needed.

Correct diagnosis, including knowledge of the life cycle and ecology of the pathogen, is obviously a critical step in any control program. Epidemiological surveys of pathogens are still marginally performed, partly due to a lack of suitable diagnostic methods. However, technologies for quick recognition of pathogens in shrimp culture are developing rapidly and diagnostic probes, which can be used in screening of captured broodstock and their postlarvaes prior to their stocking, are now available for many of the severe shrimp pathogens (Lightner, 1996).

Control of diseases is an essential importance and can be reached in different way, however, until now chemotherapy and management practices are the only methods available to decrease the infection stress in shrimp farming. Therefore, research on quantitative assays to control the defense system of penaeid shrimp and accordingly the health status has a high priority. Disease prevention is more important than treatment. Studying about defense molecules in shrimp is one approach to overcome disease problem. Studies related to invertebrate immunity are infrequent mainly due to the idea that they do not require an efficient immune system because of their short life period and high reproductive rate. However, in recent years, the defense response of penaeid shrimp has surprisingly formed subject of minor interest. The economical importance of shrimps has prompted the necessity to study their immune components.

Many species of marine shrimps are commercially cultured worldwide. In this study, we selected *Fenneropenaeus merguiensis* to be a model for study; *F. merguiensis* shrimps tolerate in low water quality, can be grown at high densities, and they are readily available in the wild. In addition, *F. merguiensis* brood stock is cheapter than *Penaeus monodon,* the most important economically cultural penaeids in Thailand. The banana prawn, *F. merguiensis* has a relatively high market value, and considerable commercial importance throughout the Indo-West Pacific. However, outbreaks of diseases caused by parasites, bacteria and viruses have caused severe economic losses to the aquaculture industry, and in some cases the mortality can reach as high as 100% (Hoang, 2001). It is well known that some gram-negative *Vibrio* spp. including *Vibrio harveyii*, *Vibrio alginolyticus* and *Vibrio parahemolyticus* are highly pathogenic to penaeid shrimps which can cause significant economic loss in cultured shrimp (Goarant *et al.*, 1999). Despite the economical importance of the shrimp and

severe economic losses caused by diseases, little research has been carried out on the banana prawn immune factors. To our knowledge, only a few immune factors have been so far purified and characterized from banana prawn, and those immune factors were reported to act important roles in pattern recognition and immune defense response against microorganism infection (Rittidach *et al.,* 2007; Mai and Hu, 2008).

The isolation, purification, characterization of purified lectin to generate specific anti-lectin antibodies, used for determining amount of lectin in hemolymph at various stages of ovarian development including that of bacterial infectous of *P. merguiensis* by enzyme linked immunosorbent assay (ELISA) were published in the Ph.D. thesis of Wanida Rittidach (2006). In this study, the present work is concerned with extending the knowledge about lectin in this shrimp by characterizing the purified lectin in more detail such as determining the N-terminal and internal amino acid sequence. Moreover, two immune genes of the banana shrimp, *F. merguiensis*, were isolated and characterized. Here, this thesis reports the cloning of lectin and lipopolysaccharide binding protein and β-1,3-glucan binding protein (LGBP) cDNA from the hepatopacreas and haemocyte cDNA library of the banana shrimp *F. merguiensis*, respectively. This work also determined the sites of both gene synthesis and several bioinformatics tools were used to perform analyses such as sequences comparison, and phylogenetic tree analysis. Expression of recombinant *F. merguiensis* lectin (FmL) protein in *Escherichia coli* was also studied. The expressed FmL protein was purified. These results contribute to an improved understanding of the shrimp immune response genes, a necessary prerequisite for development of rational strategies to improve health management in aquaculture. The results are summarized and discussed below.

Review of literatures

1. Banana shrimp, *Fenneropenaeus merguiensis*

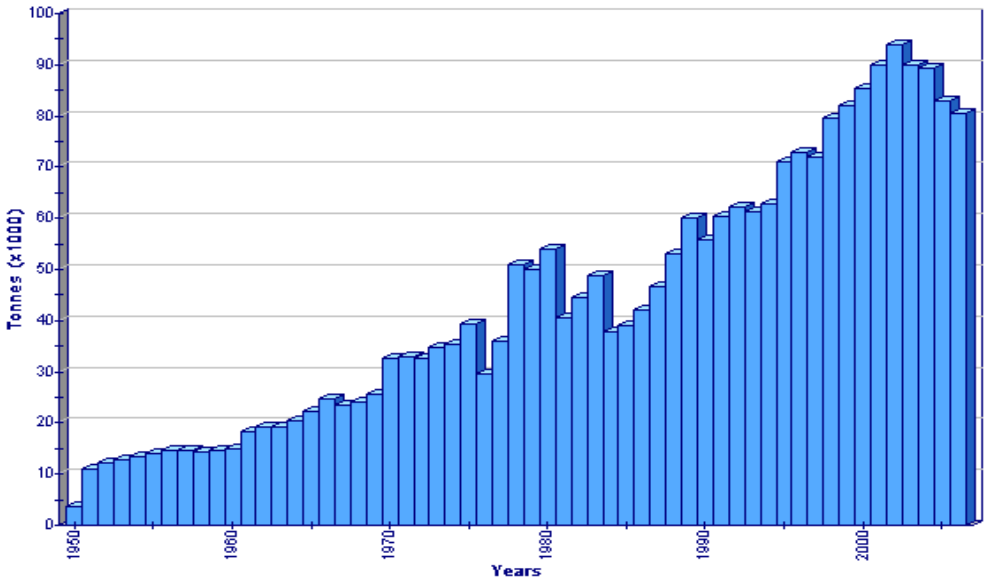
1.1 Economic importance and geographic distribution

Fenneropenaeus merguiensis or banana shrimp is one of the important species for fisheries and aquaculture in the Indo-West Pacific. The species ranges from the Persian Gulf and Pakistan through the Malay Archipelago and South China Sea to Australia, where it is found from Western Australia all the way around the north coast to northern New South Wales as shown in Fig. 1 (Holthuis, 1980). The total banana shrimp capture around the world reported to FAO for 2006 was 80,380 tons with aquaculture production of 96,833 tons (Fig. 2A and 2B) (Food and Agriculture Organization of the United Nations (FAO), 2008, available: http://www.fao.org/fishery/species/2583/en, accessed 12/12/08). The production from aquaculture of this species keeps increasing and it has become an important species for aquaculture since 2001. Therefore, the banana shrimp is a good candidate for shrimp farming. The countries with the largest total catches were Indonesia (65,230 tons) and Thailand (9,200 tons) as of 1999. An article in the December 2001 issue of World Aquaculture Society (WAS) (2006, available: http://www.was.org, accessed 12/12/08) reviewed banana shrimp and its prospects as a farmed species. This shrimp has several advantages: easy larval rearing, good survival in extensive and semi-intensive ponds, toleration of a wide range of salinities and temperatures, minimal size variation, natural maturation and spawning in captivity, it can grow at high densities and is readily available in the wild. It means that wild-caught breeders are cheaper than *P. monodon*. However, banana shrimp also has some disadvantages such as a slow growth rate, limited information in biology and culture, rapid death at harvest and lack of species-specific commercial feeds.

Fig. 1 Distribution of *F. merguiensis* in the world.

Species distribution of *F. merguiensis*, banana shrimp, is shown as red areas. (Fisheries and Aquaculture Department, Food and Agriculture Organization [FAO], 2007).

A. Capture



B. Aquaculture

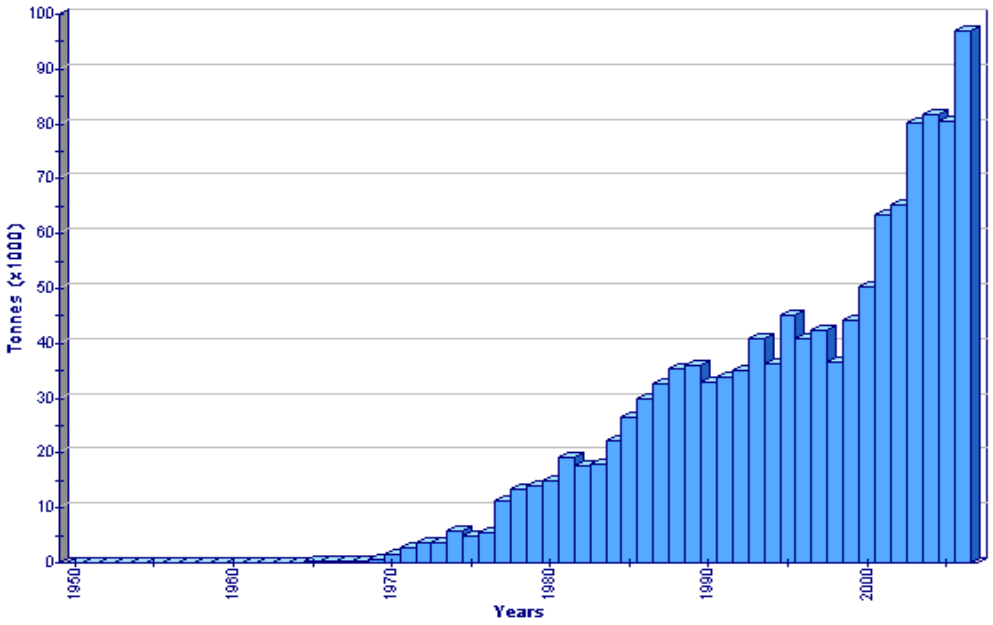


Fig. 2 Global wild capture and aquaculture production of *F. merguiensis.*

(FAO, 2008)

1.2 Taxonomy of *F. merguiensis*

F. merguiensis (Fig. 3) is a member of the family Penaeidae. Penaeid shrimps belong to the largest phylum in the animal kingdom, Arthropoda. Arthropods have a rigid cuticle that covers whole body, made largely of chitin and proteins, forming an exoskeleton that may or may not be further stiffened with calcium carbonate. The phylum takes its name from its distinctive jointed appendages (Hickman *et al.*, 2006). The subphylum Crustacea is predominantly aquatic species, that belongs to 10 classes. Within the class Malacostraca, shrimp, together with crayfish, lobsters and crabs, belong to the order Decapoda as shown below. Many species in the genus *Penaeus* have been recently reassigned, based on morphological differences, to new genera in the family Penaeidae: *Farfantepenaeus*, *Fenneropenaeus*, *Litopenaeus* and *Marsupenaeus* (Perez-Farfante and Kensley, 1997). The genus name of *P. merguiensis* was then changed from *Penaeus* to *Fenneropenaeus* in 1997 by Perez-Farfante and Kensley (1997).

Phylum-Arthropoda
Subphylum Crustacea
Class Malacostraca
Order Decapoda
Suborder Dendrobranchiata
Superfamily Penaeoidea
Family Penaeidae
Genus *Fenneropenaeus*
Species *merguiensis*

Commercial names:

Australia: Banana prawn, White prawn
Japan: Tenjikuebi, Bananaebi
Malasia: Udang kaki merah, Udang pasir
Iran: Banana shrimp, Brow tiger shrimp
Thailand: Kung chaebauy

F. merguiensis is a member of the genus *Fenneropenaeus* which is distinguished by rostrum with high blade and teeth above and below. Gastro-Orbital ridge was absent or very feebly defined. Body was often yellow or translucent with no banding on body or antennae and speckled with reddish brown dots. Penaeid shrimps are different from other decapods because they hatch out in a nauplius stage and females deposit their eggs instead of carrying them until they hatch.

F. merguiensis was raised on extensive farms throughout Southeast Asia. *F. merguiensis* is also a "white" shrimp that has attracted attention because it tolerates to low water quality better than *P. monodon.* It can grow at high densities and it is readily available in the wild. Native to the Indian Ocean from Oman to Western Australia, to Southeast Asia from the Philippines to Indonesia, and to Eastern Australia, *F. merguiensis* is heavily fished throughout its range, especially in Australia (http://www.shrimpnews.com/FarmedSpecies.html, accessed 13/12/08).

Fig. 3 Banana shrimp, *Fenneropenaeus merguiensis*.

1.3 Morphology

The exterior of penaeid shrimp is distinguished by a cephalothorax with a characteristic hard rostrum, and by a segmented abdomen (Fig. 4). Most organs, such as gills, digestive system and heart, are located in the cephalothorax, while the muscles concentrate in the abdomen. Appendages of the cephalothorax vary in appearance and function. In the head region, antennules and antennae perform sensory functions. The mandibles and the two pairs of maxillae (not visible in Fig. 4) form the jaw-like structures that are involved in food uptake (Solis, 1988). In the thorax region, the maxillipeds are the first three pairs of appendages, modified for food handling, and the remaining five pairs are the walking legs (pereiopods). Five pairs of swimming legs (pleopods) are found on the abdomen (Bell and Lightner, 1988; Baily-Brock and Moss, 1992).

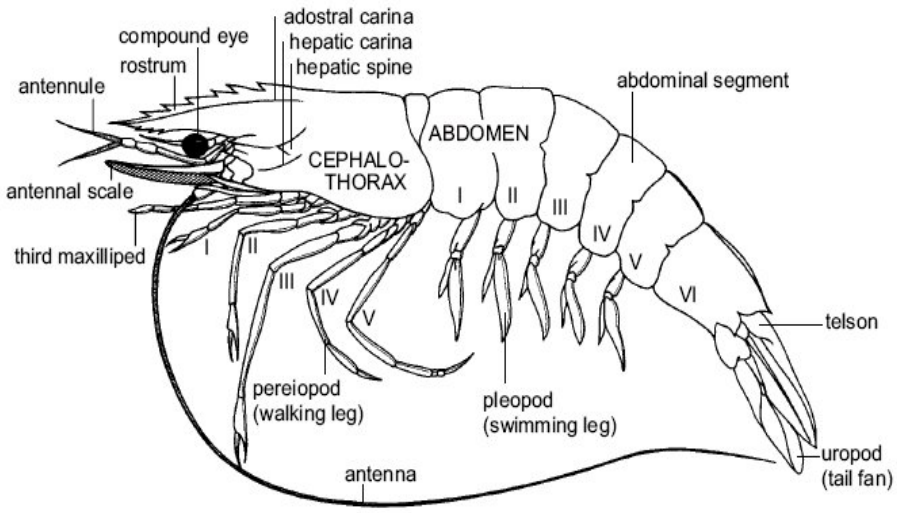


Fig. 4 Lateral view of the external morphology of penaeid shrimp.
(Primavera, 1988)

The internal morphology of penaeid shrimp is outlined in Fig. 5. Penaeids and other arthropods have an open circulatory system and, therefore, the blood and the blood cells are called haemolymph and haemocytes, respectively. Crustaceans have a muscular heart that is dorsally located in the cephalothorax. The valved haemolymph vessels leave the heart and branch several times before the haemolymph arrives at the sinuses that are scattered throughout the body, where exchange of substances takes place. After passing the gills, the haemolymph returns in the heart by means of three wide non-valved openings (Bauchau, 1981). A large part of the cephalothorax in penaeid shrimp is occupied by the hepatopancreas. This digestive gland consists of diverticula of the intestine. Spaces between these hepatopancreatic tubules are haemolymph sinuses. The main functions of the hepatopancreas are the absorption of nutrients, storage of lipids and production of digestive enzymes (Johnson, 1980). One of the haemolymph vessels that leaves the heart ends in the lymphoid organ, where the haemolymph is filtered. This organ is located ventro-anteriorly to the hepatopancreas. The haemocytes are produced in haematopoietic tissue. This organ is dispersed in the cephalothorax, but mainly present around the stomach and in the onset of the maxillipeds. Lymphoid organ and haematopoietic tissue are not shown in Fig. 5.

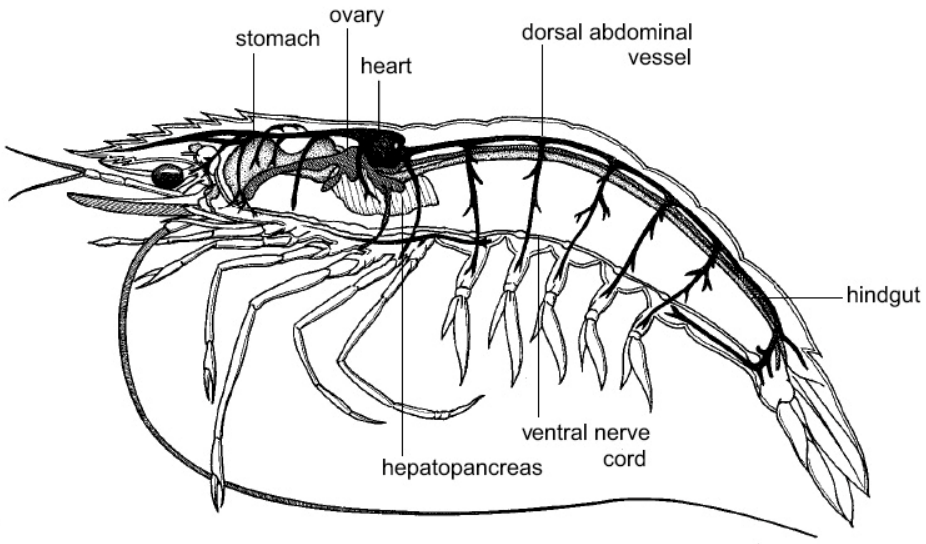


Fig. 5 Lateral view of the internal anatomy of a female penaeid shrimp.
(Primavera, 1988)

1.4 Distribution and life cycle of shrimp

The banana shrimps are widely distributed and importantly Indo-West Pacific species live in shallow water between 10 and 45 meters on muddy bottoms. Juveniles are estuarine, adults mostly marine. The penaeid life cycle includes several distinct stages found in a variety of habitats (Fig. 6). Juveniles prefer brackish shore areas and mangrove estuaries as their natural environment. Most of the adults migrate to deeper offshore areas that have higher salinity, where mating and reproduction take place. Females produce between 50,000-1,000,000 eggs per spawning (Rosenberry, 1997). Within 16 h after fertilization, the eggs hatch into the first larval stage, the nauplis (6 stages in 2 days). The nauplii feed on their yolk reserves within their bodies for a few days and develop into the protozoeae (3 stages in 5 days). The protozoeae feed on algae and metamorphose into myses (3 stages in 4-5 days). The mysis, which have many of the characteristics of adult shrimp, feed on algae and zooplankton such as rotifers (*Artemia* nauplii). They then develop into megalopas (6-35 days). The megalopa and early juveniles are call postlarvae. Transition from juvenile to subadult takes 135-255 days and subsequently completion of sexual maturity occurs within 10 months (Motoh, 1984).

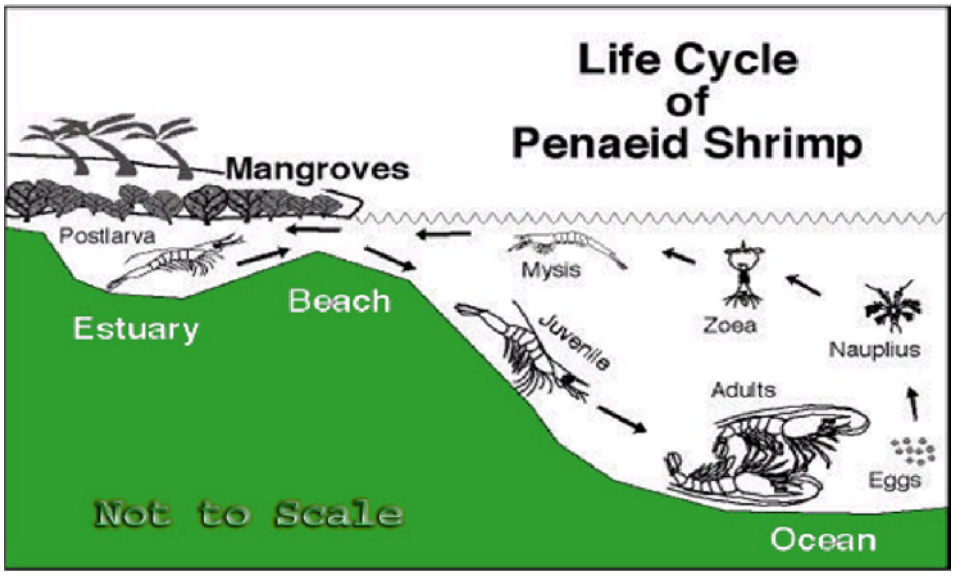


Fig. 6 Schematic of a typical penaeid shrimp life cycle.

(Bailey-Brock and Moss, 1992)

2. The invertebrate immune system

Invertebate animals do not have antibodies and therefore rely on an innate immune system. However, they have to be able to recognize foreign materials and respond to them so that appropriate measures are initiated to combat and destroy microorganisms (Loker *et al*., 2004). To initially recognize and fight infections of microorganisms, both vertebrates and invertebrates have an effective innate immune system. This system is comprised of both cellular and humoral components, and these components interplay to recognize and counter antigens. In addition to the innate system, vertebrates have a powerful acquired immune system composed of lymphocytes and different classes of immunoglobulins (Ig). These Ig molecules can adapt to almost any antigen. Once an Ig is produced in response to an antigen, immunological memory develops. If challenged again with the same antigen, the immunological response is much more rapid.

The innate immune system is the first line of inducible host defense against bacterial, fungal, and viral pathogens (Hoebe *et al.*, 2004). This defense system is essential for the survival and perpetuation of all multicellular organisms (Hoffmann *et al.*, 1999; Salzet 2001). Invertebrates, which do not possess immunoglobulins, have developed unique modalities to detect and respond to microbial surface antigens like lipopolysaccharides (LPS), lipoteichoic acids, lipoproteins, peptidoglycan (PGN) and (1→3) β-D-glucans (Begum *et al.*, 2000). Because both invertebrates and vertebrates respond to these substances, it is likely that a system recognizing these epitopes emerged at an early stage in the evolution of animals (Medzhitov and Janeway, 2000; Aderem and Ulevitch, 2000). Moreover, it is well known that various microbial cell wall components elicit a variety of responses that depend on species and cell type (Hoffmann *et al.*, 1999; Cooper *et al.*, 2002). Table 1 summarizes the major biological host defense systems of invertebrates; such systems are also found in mammals (Aderem and Ulevitch, 2000). In invertebrates, toll-like receptor-mediated antimicrobial peptide production (Lemaitre *et al.*, 1996; Imuler and Hoffmann, 2000; Krutzik *et al.*, 2001; Underhill and Orinsky, 2002), hemolymph coagulation (Iwanaga *et al.*, 1978), melanin formation (Sugumaran, 2002), and lectin-mediated complement activation (Fujita, 2002) are prominent immune responses. In addition to these enzyme cascades, a variety of agglutinin-

lectins and reactive oxygen producing and phagocytic systems cooperate with immune reactions to kill invading pathogens (Bogdan *et al.*, 2000). Fig.7 shows the principal defense systems associated with phagocytosis. Invaders detected by these systems are ultimately engulfed by phagocytes, such as macrophage-like, neutrophil-like and dendritic cells, and are then internalized as phagosomes and finally killed (Greenberg and Grinstein, 2002).

Table 1 Major host defense systems in invertebrate animals.

1. Hemolymph coagulation system
2. Pro-phenoloxidase (pro-PO) activating system
3. Lectin-complement system
4. Agglutinin-lectin system
5. Antibacterial, antifungal, and antiviral systems mediated by Toll-like receptors and peptidoglycan binding protein (PGBP)
6. Reactive oxygen-producing system
7. Phagocytic system

(adopted from Iwanaga and Lee, 2005).

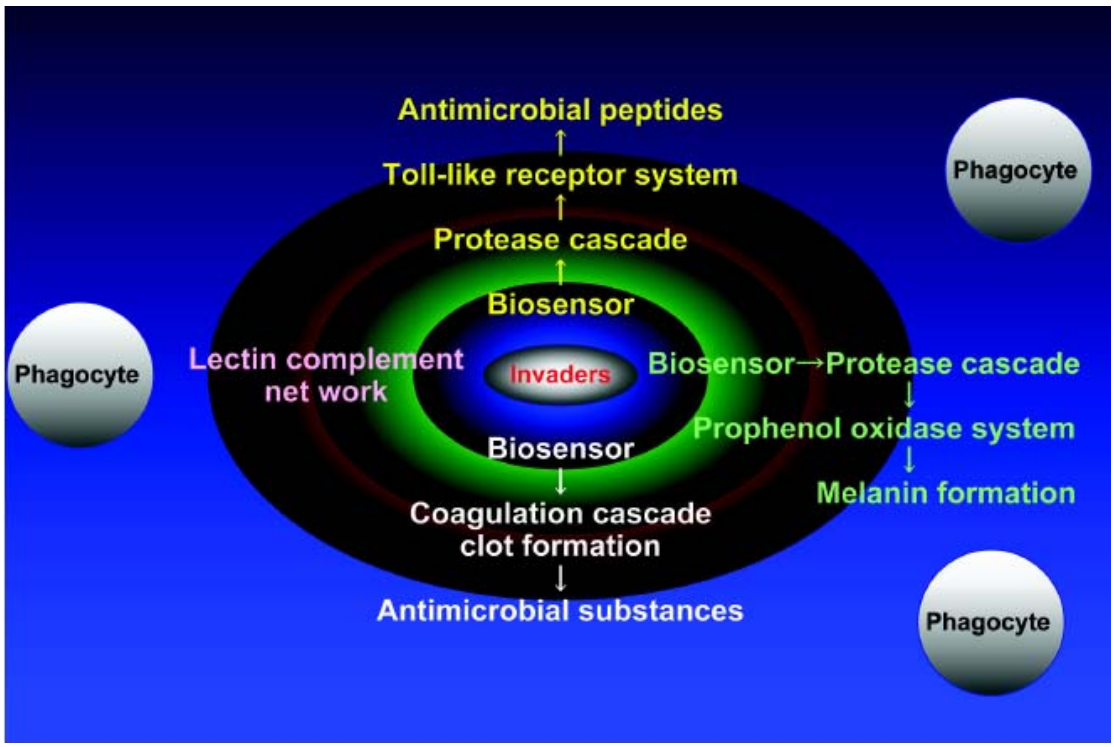


Fig. 7 The principal host defense systems associated with phagocytosis in invertebrates.

The major innate immune systems include; hemolymph coagulation, melanization mediated by phenoloxidase, the expression of antimicrobial peptides mediated by Toll-like receptors and the immuno deficiency (IMD) pathway, and the lectin/complement pathway mediated by bacterial cell wall components. Invaders detected by these systems are ultimately engulfed by phagocytic cells, such as macrophage-like, neutrophil-like, or dendrotic cells, and then internalized, processed, and killed (adopted from Iwanaga and Lee, 2005).

3. The crustacean immune system

Crustacean immune system is non-specific immune system so-called the innate immunity (Söderhäll and Cerenius, 1998). The innate immune system is an ancient defense mechanism and can be found in all multicellular organisms. This system is based on cellular and humoral components of the circulatory system. Although invertebrates cannot produce antibodies and hence have no immune memory, their innate immune mechanisms are sufficient to protect and preserve themselves from intruding microorganisms. The hard cuticle which covers all external surfaces of crustaceans is the first line of defense to provide an effective physical and chemical barrier against the attachment and penetration of pathogens. The digestive tract, a main route of invasion, is partly lined with chitinous membranes and its hostile environment of acids and enzymes is able to inactivate and digest many viruses and bacteria. In addition to their rigid and wax-covered cuticle, which serves as a mechanical barrier, shelled animals can also produce rapidly and effective innate immune responses during infection. In most cases the cuticular defences are sufficient to protect against even quite virulent pathogens, which often only produce disease when the integument has been physically damaged. However, the innate immunity responses rapidly if microorganisms can invade the animals. Once pathogens gain entry into the hemocoel of the host, they encounter a complex system of innate defence mechanisms involving cellular and humoral responses. Humoral defences include the production of antimicrobial peptides (AMPs), reactive intermediates of oxygen or nitrogen, and the complex enzymatic cascade that regulate clotting or melanization of hemolymph. In contrast, a cellular immune response that involves different types of hemocytes, which participate in pathogen clearance by phagocytosing microorganisms, or trapping them in hemocyte aggregates or nodules, or by encapsulation of larger microorganisms and cytotoxic reactions are also triggered. Nodule and capsule formation are often observed to become melanized, through the action of phenoloxidase (Lavine and Strand, 2002; Cerenius and Söderhäll, 2004; Iwanaga and Lee, 2005). In fact, there is an overlap between humoral and cellular defence, since many humoral factors affect hemocyte function and hemocytes are an important source of many humoral molecules (Elrod-Erickson

et al., 2000; Lavine and Strand, 2002). An illustration of the innate immune reactions is shown in Fig. 8.

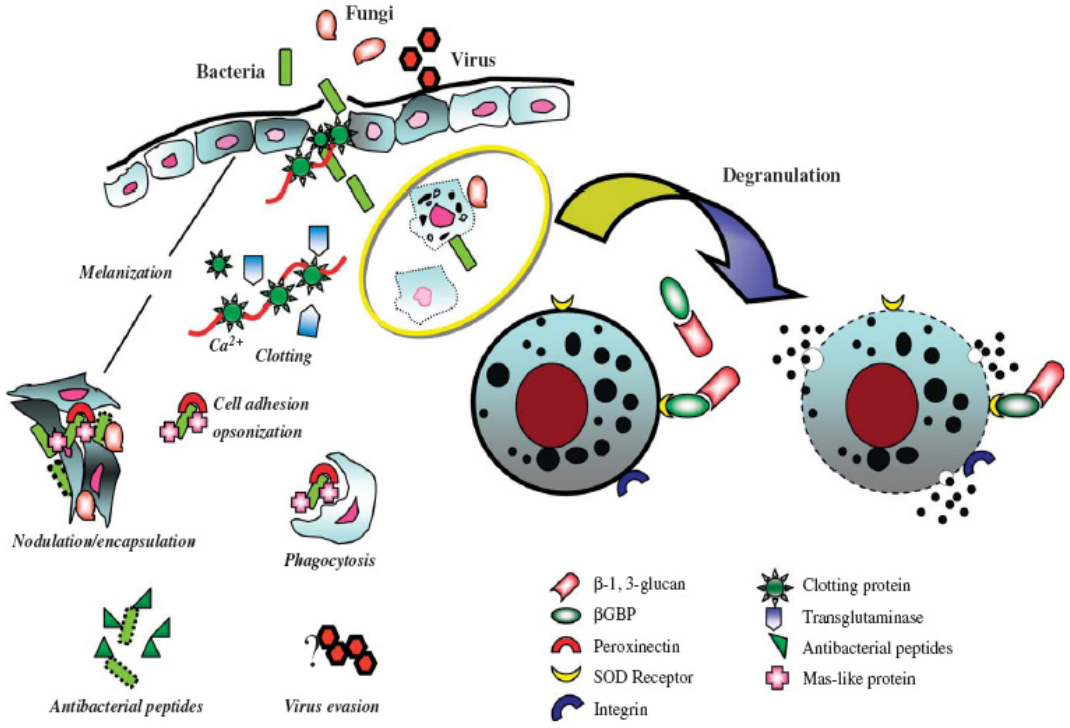


Fig. 8 Overview of present-day knowledge of the most important defense factors in decapods crustaceans.

(adopted from Jiravanichpaisal *et al*., 2006)

3.1 Cellular defenses

Crustaceans are not different from other animals in that the host defense is largely based on activities of the blood cells or haemocytes, which perform functions such as phagocytosis (Foukas *et al*., 1998), encapsulation (Kobayashi *et al*., 1990; Asgari *et al*., 1998; Cho *et al*., 1999) and nodule formation (Koizumi *et al*., 1999). They play extremely important roles not only by direct sequestration and killing of infectious agents but also by synthesis and exocytosis of a battery of bioactive molecules. Three types of circulating hemocytes are morphologically recognized in crustaceans (Johansson *et al*., 2000). The smallest and least numerous hyaline cells (HCs) are considered as phagocytes (Söderhäll *et al*., 1986) and also involve in coagulation by releasing transglutaminase (Omori *et al* 1989; Söderhäll and Cerenius, 1992) that can change coagulogen into dissolved coagulin. They also produce reactive oxygen intermediates (ROIs) after phagocytosis (Persson *et al*., 1987). The semigranular cells (SGCs) that contain small eosinophilic granules possess the prophenoloxidase activating system (proPO system). The SGCs, which are responsible for encapsulation and a limited function in phagocytosis, are specialized in particle (Persson *et al*., 1987; Kobayashi *et al*., 1990). The largest granular cells (GCs) contain a large amount of eosinophilic secretory granules and are the major storage cell for the proPO system, but are not associated with phagocytosis function.

3.2 Humoral immunity

The humoral response elements include proPO activation system (PAS), precipitins, agglutinins or lectins, opsonins and antimicrobial peptides (Söderhäll and Cerenius, 1992). The shrimp's full immune reactivity is always achieved through co-operation and interaction between hemocytes of different types and their porducts. In non-challenge or uninfected condition, most of these defense factors exist in hemocytes as inactive forms. They are activated and released by exocytosis when non-self materials present in the hemolymph.

Innate immune activation is based on recognition of pathogen molecules not present in the host. An essential early step in an immune response is recognition of an invading pathogen as dangerous non-self (Janeway and Medzhitov, 2002; Sansonetti, 2006). Once such recognition has taken place, it may trigger a protective response involving blood cells or soluble plasma proteins. Innate immune

systems in both mammals and arthropods utilize proteins known as pattern recognition proteins (PRPs), which perform a surveillance function by binding to molecules (molecular patterns) common to the groups of microorganisms, but absent from animals (Janeway, 1989 and Hoffmann *et al*., 1999). A number of soluble molecules binding to and lysing microbes have been found in invertebrates. For instance, lectin-like proteins can bind to carbohydrates present on microbial cell walls followed by activation of immune response to kill/remove the invading microbes. These recognition molecules for non-self materials have been named as pattern recognition proteins (Medzhitov and Janeway 1997). Pathogen-associated molecules not found in other multicellular organisms, eg. Lipopolysaccharide (LPS)/peptidoglycan (PGN) from bacterial cell walls, β-1,3-glucan from fungal cell walls, and double stranded RNA (dsRNA) of viruses, can be recognized by these recognition proteins based on the particular structures of the invading microbes. Binding of these pattern recognition proteins in turn leads to activation of other immune responses. In several cases, pattern recognition proteins have been either discovered or studied by investigating plasma proteins that bind to bacteria *in vitro*. Current studies have emphasized the important roles of non-self recognition molecules both in the vertebrate and the invertebrate immune systems. Recognition of microbial carbohydrates that are distinct from those of host cells is the first step to trigger the innate immunity. Many recognition molecules, such as LPS or/and β-1, 3-glucan binding proteins (LBP, BGBP or LGBP), peptidoglycan recognition proteins (PGRPs), and lectins have been identified in various invertebrates and different biological functions have been suggested for these molecules following their binding to the microbes (Lee and Söderhäll 2002).

This section gives a brief description of some of most important groups of innate immune receptors from invertebrates including a more thoroughly description of the lectin and the lipopolysaccharide and β-1,3-glucan binding protein.

3.2.1 Lectin

Lectin or agglutinin is a sugar-binding protein of non-immune origin that agglutinates cells or precipitates glycoconjugates (Goldstein *et al.*, 1980). A lectin molecule contains at least two sugar-binding sites; sugar-binding proteins with a single site will not agglutinate or precipitate structures that contain sugar residues, so

are not classified as lectins. The specificity of lectin is usually defined by the monosaccharides or oligosaccharides that are best at inhibiting the agglutination or precipitation the lectin causes. Lectins have been found in plants, viruses, microorganisms and animals, but despite their ubiquity, their function in nature is unclear. Although lectins share the common property of binding to defined sugar structures, their roles in various organisms are not likely to be the same.

Lectins may interact with carbohydrates through hydrogen bonds, metal coordination, Van der Waals, and hydrophobic interactions (Weis and Drickamer, 1996; Elgavish and Shaanan, 1997). Basically, hydroxyl groups on sugar molecules can serve as both a donor and an acceptor to cooperate in hydrogen bonds (Elgavish and Shaanan, 1997). Hemagglutination activity is a major property of lectins and has been widely used for their detection and characterization (Lis and Sharon, 1998). However, several lectins possess only one binding site for specific sugars, therefore they cannot agglutinate red blood cells (Sharon and Lis, 1989).

Literally thousands of articles as well as various reviews on lectins and their application in affinity chromatography have been published during the past forty years examining hundreds of different aspects and uses of lectins (Sharon and Lis, 1972; Lotan and Nicolson, 1979). But only a few of the uses and properties of each lectin are described. Concurrently, it was shown that lectins function as recognition molecules in cell–molecule and cell–cell interactions in a variety of biological systems (Sharon and Lis, 2004).

3.2.1.1 Structures, properties and classification of lectin

Most lectins studied to date are multimeric, consisting of non-covalently associated subunits. A lectin may contain two or more of the same subunit, such as concanavalin A (Con A) (Agrawal and Goldstein, 1972), or different subunits, such as *Phaseolus vulgaris* agglutinin (Takahashi *et al*., 1980). It is this multimeric structure which gives lectins their ability to agglutinate cells or form precipitates with glycoconjugates in a manner similar to antigen-antibody interactions.

One major property of lectins is their specific saccharide- binding sites. Some lectins are composed of subunits with different binding sites. The specificity of the binding sites of the lectins suggestes that there are endogenous saccharide receptors in the tissues from which they are derived or on other cells or

glycoconjugates with which the lectin is specialized to interact. Because of the specificity that each lectin has toward a particular carbohydrate structure, even oligosaccharides with identical sugar compositions can be distinguished or separated. Some lectins will bind only to structures with mannose or glucose residues, while others may recognize only galactose residues. Some lectins require that the particular sugar be in a terminal non-reducing position in the oligosaccharide, while others can bind to sugars within the oligosaccharide chain. The affinity between a lectin and its receptor may vary a great deal due to small changes in the carbohydrate structure of the receptor.

Another property of some lectins is an ability to induce mitosis in cells which are normally not dividing. This property has been exploited extensively in an attempt to understand the process of lymphocyte blastogenesis and the biochemical and structural alterations associated with mitogenesis. It is not clear why some lectins are mitogenic since the structures to which mitogenic lectins bind are not necessarily the same, and not all lectins with similar binding specificities are mitogenic. It is likely that binding to the cell surface alone is not sufficient to cause mitosis but that other interactions on the cell surface are equally important.

In general most lectins are considered glycoproteins. However, non-glycoprotein lectins are believed to be synthesized as glycosylated precursors. All glycoprotein lectins contain a peptide sequence: asparagine-X-threonine/serine, which is characteristic of glycosylation sites. These sequences are different in the non-glycoprotein lectins. Also, peptide sequences, which in one glycoprotein lectin contain the glycosidic side-chains, are not necessarily conserved in another glycoprotein lectin. This may suggest that the biological activity of the lectins may not be determined by carbohydrate part of their structure (Barondes, 1981).

Animal lectins are a heterogeneous class of molecules, which exhibit a high structural diversity. They have been initially classified in two different groups: C- and S-type lectins ("C" stands for Ca-requiring and "S" stands for SH-requiring) based on amino acid sequence similarities, particularly in the carbohydrate recognition domain, along with overall domain organisation and physico-chemical properties, such as divalent cation dependence and free thiol requirement (Drickamer, 1988).

C-type lectins are calcium-dependent animal lectins that are carbohydrate-binding proteins of animal origin. Carbohydrate-binding activity of C-type lectins is based on the function of the carbohydrate recognition domain whose structure is highly conserved among this family (Drickamer, 1988). Calcium is not only directly involved in the carbohydrate binding itself at the binding site (Weis *et al*., 1992) but also contributes to the structural maintenance of the lectin domain that is essential for the lectin activity (Kimura, 1995).

Galectins are defined as lectins having both galactose-binding ability and amino acid sequences which characterize galectins. Since these lectins usually require a thiol reducing reagent for the maintenance of their activity, Drickamer (1988) designated them as "S-type" lectins.

C-type lectins and galectin that recognize complex structures at the cell surface found in invertebrate organisms as well as vertebrates but the functions of these proteins have evolved differently in different animal lineages (Dodd and Drickamer, 2001).

Lectins are often complex, multi domain proteins, but sugar-binding activity can usually be ascribed to a single protein module within the lectin polypeptide. Such a module is designated a carbohydrate recognition domain (CRD). Lectins can be divided by sequence comparison of the CRD. Some of the best characterized of these CRD groups are summarized in Table 2.

Table 2 Summary of lectin categories.

Lectin group	Structure of CRD	Typical ligands	Examples of functions
Calnexin	Unknown	Glc_1Man_9	Protein sorting in the endoplasmic reticulum
L-type lectins	β-sandwich	Various	Protein sorting in the endoplasmic reticulum
P-type lectins	Unique β-rich structure	Man-6-P	Protein sorting post Golgi
M-type lectins		Man_8	Endoplasmic reticulum-associated degradation of glycoproteins
C-type lectins	Unique mixed α/β structure	Various	Cell adhesion (Selectins), Glycoprotein clearance, Innate immunity (Collectins)
Galectins	β-sandwich	β-Galactosides	Glycan crosslinking in the extracellular matrix
I-type lectins	Immunoglobulin superfamily	Sialic acid	Cell adhesion (Siglecs)
R-type lectins	β-trefoil	Various	Enzyme targeting, Glycoprotein hormone turnover

Glc, glucose; Man, mannose; Man-6-P, mannose-6-phosephate.

The functions listed are primarily those that have been identified in vertebrates.

(Drickamer and Taylor, 1998)

The lectins that contain CRDs listed in the Table 2 fall broadly in two categories. Lectins that contain CRDs in the first three of the structural groups are located mostly intracellularly, in luminal compartments. They function in the trafficking, sorting and targetting of glycoproteins in the secretory and other pathways. CRDs in the remaining structural groups are found in lectins that function largely outside the cell and are either secreted or localized to the plasma membrane (Dodd and Drickamer, 2001).

Relatively simple invertebrate organism may serve as useful model for some of the functions of sugar-binding the protein in mammals. The early intracellular sorting events involving calnexin and L-type lectins as well as the role of R-type CRDs in glycosyltransferases are likely to be quite similar, whereas later sorting events involving the mannose-6-phosephate receptors will probably be different. At the cell surface, the role of some of the galectins may be similar in all animals so that genetic and developmental analysis of the model invertebrates is likely illuminated studies of the vertebrate proteins as well. In contrast, the greater diverstion of vertebrates and invertebrates proteins containing C-type lectin like domains (CTLDs) suggests that these proteins probably participate in more specialized functions of glycan that are unique to different groups of animals (Dodd and Drickamer, 2001).

3.2.1.2 Functional aspects of lectins

In subsequent years numerous lectins have been isolated from plants as well as from microorganisms and animals, and during the past two decades the structures of hundreds of them have been established. Concurrently, it was shown that lectins function as recognition molecules in cell–molecule and cell–cell interactions in a variety of biological systems. In a broader sense, the foregoing discussion implies that lectins possess the ability to act as recognition molecules inside cells, on cell surfaces, and in physiological fluids (Fig. 9 and Table 3).

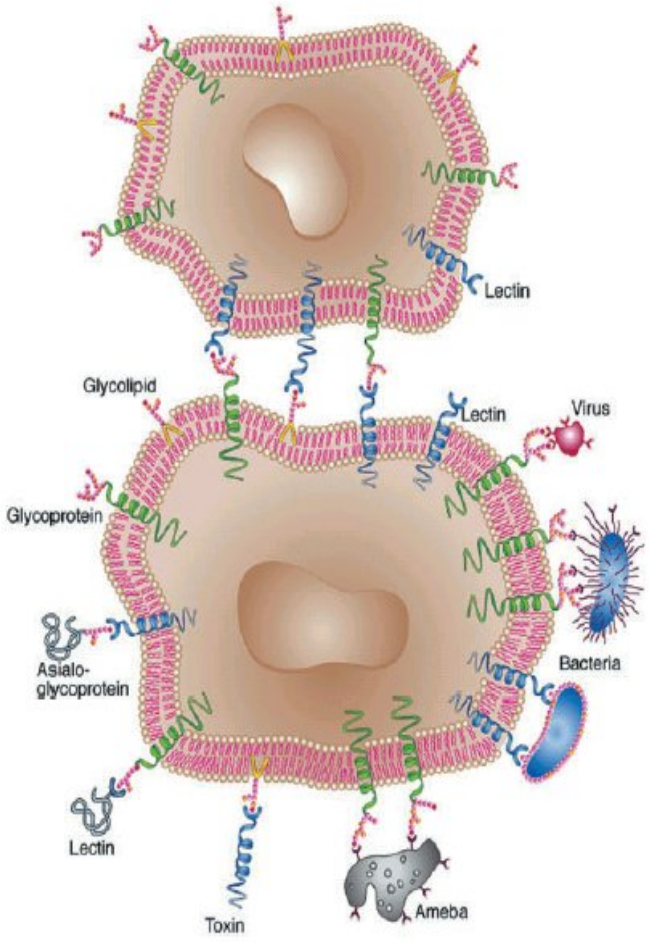


Fig. 9 Cell surface lectin-carbohydrate interactions.

Lectins serve as means of attachment of different kinds of cell as well as viruses to other cells via the surface carbohydrates of the latter. In some cases, cell surface lectins bind particular glycoproteins (e.g., asialoglycoproteins), whereas in other cases the carbohydrates of cell surface glycoproteins or glycolipids serve as sites of attachment for biologically active molecules that themselves are lectins (e.g., carbohydrate-specific bacterial and plant toxins, or galectins) (Sharon and Lis, 2004).

Table 3 Functions of lectins.

Lectin sources	Roles
Microorganisms	
Amoeba	Infection
Bacteria	Infection
Influenza virus	Infection
Plants	
Various	Defense
Legumes	Symbiosis with nitrogen-fixing bacteria
Animals	
Calnexin, calreticulin, ERGIC-53	Control of glycoprotein biosynthesis
Collectins	Innate immunity
Dectin-1	Innate immunity
Galectins	Regulation of cell growth and apoptosis; regulation of the cell cycle; modulation of cell–cell and cell–substratum interactions
Macrophage mannose receptor	Innate immunity; clearance of sulfated glycoprotein hormones
Mannose-6-phosephate receptors	Targeting of lysosomal enzymes
L-selectin	Lymphocyte homing
E- and P-selectins	Leukocyte trafficking to sites of inflammation
Siglecs	Cell-cell interactions in the immune and neural system
Sperm adhesin	Sperm-egg interaction

(Sharon and Lis, 2004)

From the functional point of view, lectins can functionally be distinguished by whether they recognize endogenous or exogenous ligands. The former appear to play an important role in fertilization and development, and their function often involves cell-to-cell or cell-to-matrix interaction. The latter probably evolved for self/non-self discrimination and they may be soluble or surface bound (Arason, 1996).

Selectins and the asialoglycoprotein receptor can recognize endogenous ligands. Selectins belong to the C-type lectin family and a similarity in the domain organization allows them to be recognized as a family of cell-cell adhesion molecules. They interact with carbohydrate ligands on leukocytes and endothelial cells. These operate in the vascular and hematologic systems. Selectins are essential for leukocyte recruitment into inflamed tissue (Graves *et al*., 1994) and the function of lymphocyte selectin plays a role in the homing of lymphocytes to the peripheral lymph nodes (Arason, 1996).

The high affinity of selectins for their ligands is probably achieved by forming dimers or oligomers in the cell membrane. This is a common theme in mammalian lectins and is known to result in increased affinity for multivalent ligands. A point quite well illustrated by the asialoglycoprotein receptor of hepatocytes (Lodish, 1991). This receptor is thought to function in the removal of effete serum glycoproteins with oligosaccharides terminating in galactose.

The macrophage mannose receptor and the collectins can recognize endogenous ligands. The macrophage mannose receptor binds to terminal mannose, fucose or *N*-acetyl glucosamine (GlcNAc). This ligand is common in the glycocalyx of viruses, bacteria, fungi and parasites. But it is not normally found at the termini of mammalian glycoproteins other than intracellular ones (Kornfeld and Kornfeld, 1985). The mannose receptor of mammalian macrophages and hepatic endothelial cells (Stahl, 1992) mediates phagocytosis of pathogens as well as receptor-mediated pinocytosis of potentially harmful high-mannose glycoproteins, released from cells in response to pathological events.

Collectin (collagen-like lectin) is a subgroup of C-type (i.e. Ca^{2+}-dependent) animal lectins characterized by the presence of collagen-like sequences (Gly-Xaa-Yaa triplet). They play important roles in innate immunity without

involvement of antibodies (Lu, 1997). The collectins are known to act as opsonin in various circumstances. They recognize and bind to non-host carbohydrates structures present on the surface of a range of microorganisms including bacteria, yeasts, fungi, parasitic protozoa and viruses on microorganisms and particles and participate in the processing or elimination of such material by interaction with phagocytic cell receptors (Lu, 1997).

The collectin serum protein called mannan-binding protein, MBP can activate the complement system through the classical pathway. This MBP-mediated complement activation, called the lectin pathway, provides an additional mechanism of microorganism recognition by the complement system in the absence of specific antibodies (Ikeda *et al*., 1987) (Fig. 10).

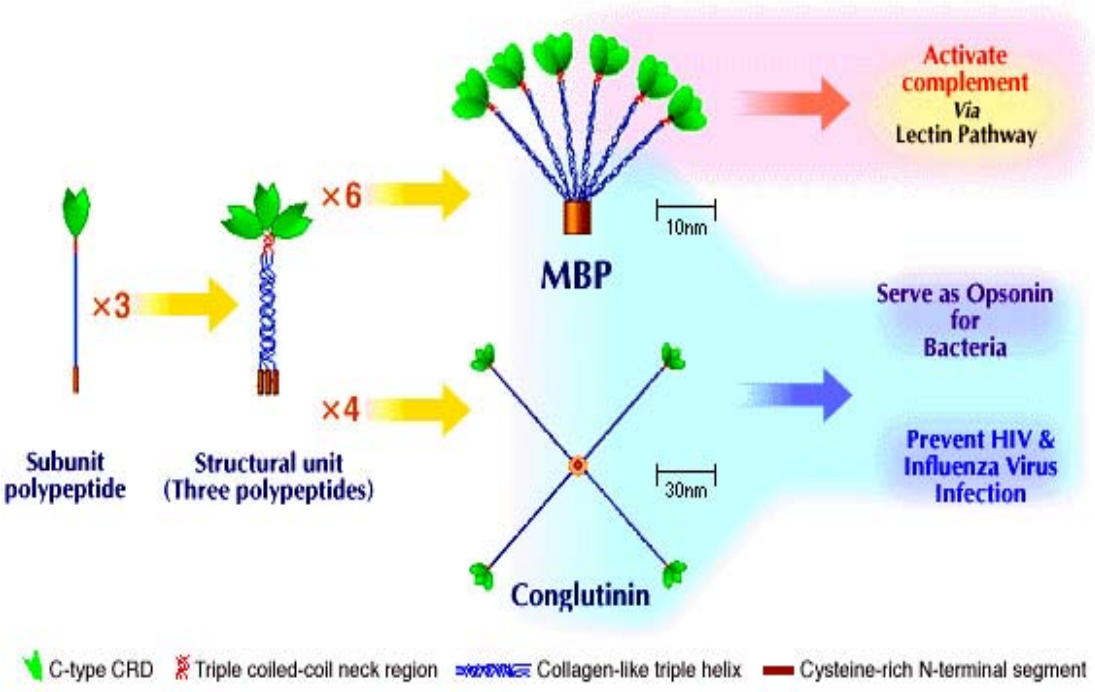


Fig. 10 Molecular structures and biological activities of collectins (MBP and conglutinin).

MBP, mannose-binding protein**;** CRD, carbohydrate recognition domain (Sharon and Lis, 1995; Holmskov *et al*., 1994)

The first substantial evidence of multiple lectins in crustaceans was reported more than two decades ago in the lobster *Homarus americanus* by Hall and Rowlands (1974). The authors purified two different plasma agglutinins and characterized the sugar specificity of both purified lobster lectins. The authors emphasized the fact that multiple lectins with different sugar specificity could constitute an important evidence of agglutinin heterogeneity and thus, support the hypothesis of the role of agglutinins as non-self-recognition molecules in crustacean defense.

Numerous reviews of agglutinin involvement in arthropod defense and recognition mechanisms have been reported (Amirante, 1986; Olafsen, 1986; 1988; Vasta, 1992; Millar and Ratcliffe, 1994; Vargas-Albores, 1995; Vasta *et al*., 1996; Natori *et al*., 1999). The two recent examples in arthropods appear to support the view that invertebrate lectins may exhibit the required binding diversity to efficiently discriminate non-self-particles. The first, concerns the cockroach *Blaberus discoidalis* which contains multiple plasma lectins, each with different carbohydrate-binding specificities (Chen *et al*., 1993; Wilson *et al*., 1999) and consequently, able to potentially recognize different invading pathogens. It was demonstrated that each of these purified molecules was capable to induce a specific and enhanced phagocytic response towards different microorganisms, such as yeast *Saccharomyces cerevisiae* and bacteria *E. coli* and *Bacillus cereus*. This response was related to the carbohydrate exposed on the microorganism surface and to the sugar specificity of each lectin (Wilson *et al*., 1999).

Another interesting example of multiple lectins with different sugar specificities and playing distinct functional roles in non-self-recognition was also provided by horseshoe crabs. In the Japanese horseshoe crab *Tachypleus tridentatus*, several lectins were purified from hemolymph. They promoted the agglutination of certain strain of Gram-positive *Staphylococcus* and recognised several kinds of lipopolysaccharides, LPS (Okino *et al*., 1995; Kawabata and Iwanaga, 1999).

Besides lectins, the immune system of arthropods includes also other defense mechanisms involving microorganism-carbohydrate recognition. A complex proteolytic cascade known as prophenoloxidase (proPO) system is triggered by components of microbial cell walls. (Söderhäll and Cerenius, 1992; Söderhäll *et al*.,

1996; Ashida and Brey, 1997). Recently an interesting connection between lectins and the proPO activating system of arthropods was reported in insects. In the cockroach *B. discoidalis*, Wilson *et al*. (1999) obtained evidences that the level of phagocytosis of the distinct microorganisms mentioned above could be increased by the several lectins, in a proPO-dependent and/or proPO-independent mechanism.

The observations of Vazquez *et al*. (1993; 1996; 1997) on the lectins of the hemolymph of the freshwater prawn *Macrobrachium rosenbergii* are of particular interest. The authors demonstrated that the granulocytes of *M. rosenbergii*, in spite of expressing a surface receptor, which seemed to correspond to the humoral purified lectin, had the ability to recognize foreign cells in an apparently non-mediated sugar recognition basis.

It is important to note that apart from the potential involvement of arthropod lectins in non-self-recognition and opsonization, recent findings are emerging on the role of these molecules as immune effectors in microorganism neutralization.

3.2.1.3. Synthesis and induction of lectins

There are evidences that arthropod lectins might be synthesised by hemocytes. The agglutinating activity in the lobster *H. americanus* was strongly associated to hemocyte extracts and the authors suggested that they should be the main agglutinin source (Hall and Rowlands, 1974). In the horseshoe crab *T. tridentatus*, also referred to above, four of the five purified lectins came from hemocytes and were released from the granules upon LPS stimulation.

In the majority of arthropods, attempts to stimulate their production have met only limited success. The potential inducibility of lectins, especially in species of economic interest, such as shrimps, could be of particular relevance since if these molecules are really involved in immune defense reactions, the increase in their concentration could virtually confer a better protection to the host against invading pathogens. In the black tiger shrimp *P. monodon*, Ratanapo and Chulavatnatol (1992) reported an elevation of the lectin monodin level in most of the shrimps suffering from bacterial *Vibrio vulnificus* infection. However, this finding could not be clearly associated to a possible inducible mechanism, since this increased lectin concentration was not observed in all infected shrimps. On the other hand, in the same shrimp

species, Sritunyalucksana *et al*. (1999) failed to induce an increase of lectin concentration by using components of microorganism cell wall, such as LPS, β-glucans, peptidoglycan and also commercial stimulants.

3.2.1. 4. Lectins in marine invertebrates

Although a number of lectins of various molecular weights have been found in marine invertebrates, very limited information concerning their structures has thus far been obtained. One of the most probable roles of marine invertebrate lectins is to act as humoral factors in the defense mechanism, as do immunoglobulins in vertebrates. This is suggested from some observations such as the activation of phagocytes by the binding of lectin to foreign cells (opsonin activity) or the enhancement of lectin production in body fluids after injection of foreign substances. On the other hand, direct hemolytic activity has recently been found for a sialic acid-specific lectin from horseshoe crab *Tachypleus tridentatus* (Beisel *et al*., 1999) and a galactose-specific lectin from the sea cucumber *Cucumaria echinata* (Hatakeyama *et al*., 1995). After binding to the specific carbohydrate chains on the erythrocyte surface, these lectins damage the cell membrane, leading to cell lysis (Fig. 11).

These lectins may play an important role against bacterial infections or natural enemies. A lectin with biological activities such as mitogenic and chemotactic activities was found in the venom of the pedicellariae (spines) of the sea urchin *Toxopneustes pileolus*, suggesting its involvement in toxic action (Nishimura *et al*., 2001). Other functions of marine invertebrate lectins have also been suggested for some C-type lectins: calcium carbonate crystallization (the acorn barnacle, *Megabalanus rosa*), morphogenesis (the tunicate, *Polyandrocarpa misakiensis*) (Matsumoto *et al.,* 2001), and vitelline coat lysis (the blue mussel, *Mytilus edulis* sperm) (Takagi *et al.,* 1994). These examples suggest that there may be various lectins with different physiological roles in marine invertebrates.

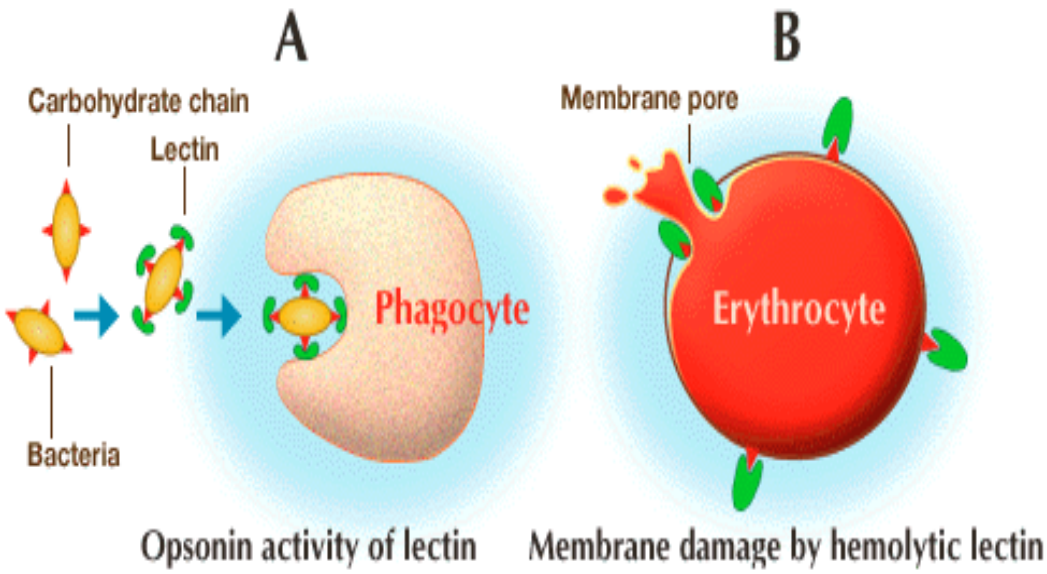


Fig. 11 Binding of lectins to foreign cells.

Binding of lectins with opsonin activity to foreign substances such as bacteria promote phagocytosis (A). After binding to the carbohydrate chains on erythrocyte surface, hemolytic lectins induce membrane damage (e.g., by forming pores), leading to hemolysis (B) (Hatakeyama *et al*., 1995).

3.2.1.5 Lectins as defense molecules in crustaceans

In invertebrates, the mechanism of non-self recognition is still largely unknown. Scientists focus their attention on two systems which could produce recognition molecules: the proPO system and lectins of hemolymph. It is well documented that invertebrate hemolymph contains internal defense factors against potential pathogens, and humoral factors such as lectins may contribute to such defense.

Lectins in invertebrates have been detected in the blood components, namely cell-free hemolymph and hemocytes (Yeaton, 1981; Ratclife *et al*., 1985; Renwrantz, 1986; Richards and Renwrantz, 1991; Smith and Chisholm, 1992), in tissues (Mullainadhan and Renwrantz, 1986; Suzuki and Mori, 1991) and in mucus (Fountain and Campbell, 1984). The actual functions of invertebrate lectin are not yet well understood. However, they have been implicated in several diverse physiological processes, including host immune responses, such as non-self recognition, phagocytosis, encapsulation, and hemocoelic clearance of foreign cells (Renwrantz, 1986; Olafsen, 1988; Vasta, 1991; Mercy and Ravindranath, 1994).

Lectins from the hemolymph of invertebrates, including crustaceans, have been regarded as potential molecules involved in immune recognition and microorganism phagocytosis through opsonization. Opsonization that occurs when a lectin binds bacteria to blood cell (hemocyte) surfaces is considered to be the first step that promotes adherence, ingestion and subsequent elimination of microbes (Martin *et al*., 1993). Concerning carbohydrate specificity, invertebrate lectins can vary from specific molecules to others that exhibit a broader spectrum of recognition. The specificity of a lectin is related to the carbohydrate for which it shows the highest affinity. But most lectins that are considered specific for one monosaccharide may also bind to other carbohydrates which are structurally related. The agglutination of many different types of cells/glycoconjugates may actually reflect the ubiquity of the ligand. A lectin from the horseshoe crab, *Limulus polyphemus* that binds sialic acid can illustrate this statement and be considered as a lectin of a broad spectrum of recognition, since the fact that it agglutinates a variety of cells can be associated to the wide distribution of this monosaccharide among cell surfaces (Cohen *et al*., 1983; Vasta, 1992).

Among decapod crustaceans, the species whose lectins have been examined are the freshwater prawn, *M. rosenbergii* (Huang *et al*., 1981), crayfish, *Pacifastacus leniusculus* (Kopacek *et al*., 1993), *Penaeus japonicus* (Muramoto *et al*., 1995), *Penaeus californiensis* (Vargas-Albores *et al*., 1993) and freshwater crab, *Parathelphusa hydrodromus* (Nalini *et al*., 1994). The specificity of lectins towards carbohydrates is mainly related to *N*-acetylated carbohydrates, such as *N*-acetyl neuraminic acid, *N*-acetyl glucosamine and *N*-acetyl-D-galactosamine as can be depicted by the data partially summarised in Table 4.

Table 4 Lectins characterized in different species of penaeids.

Species	Divalent cation	MW (kDa)	Subunits (kDa)	Carbohydrate/glycoprotein/ LPS inhibition	Biological activity	References
Penaeus japonicus	CD	330	33	GlcNAc; GalNAc NeuAc; Ribose BSM; PSM; Fetuin	Opsonic activity-serum	Kondo *et al*., 1992; 1998
	CI	n.d.	n.d.	GlcNAc; GalNAc NeuAc; BSM PSM; Fetuin	n.d.	
Penaeus indicus	CI	n.d.	n.d.	GlcNAc; GalNAc ManNAc; NeuAc BSM; Fetuin; LPS	n.d.	Maheswari *et.al*., 1997
Penaeus californiensis	CD	175	41	GlcNAc; GalNAc NeuAc; BSM Fetuin; LPS	Agglutination of bacteria (*V. fischeri*; *parahaemolyticus*; *vulnificus*); Opsonic activity	Vargas-Albores *et al*., 1993; Vargas-Albores, *et al*., 1995
Penaeus stylirostris	CD	30		n.d.	n.d.	Vargas-Albores *et al*., 1993
Penaeus monodon	CD	420	27	GlcNAc; GalNAc ManNAc; NeuAc BSM; Fetuin	Agglutination of bacteria (*V. vulnificus* *V. parahaemolyticus*; *A. hydrophila*; *P. shigelloides*; *P. aeruginosa*; *E. coli*; *B. subtilis*)	Ratanapo and Chulavatnatol l990; 1992; Sritunyalucksana *et al*., 1999
Penaeus paulensis	CI	153	31	NeuAc; GalNAc GlcNAc; Fetuin; LPS	Opsonic activity (serum) (*V. harveyi*; *B. cereus*)	Marques and Barracco, 2000
Penaeus schmitti	CPD	153	31, 34	NeuAc; GalNAc GlcNAc; Fetuin; LPS	n.d.	Marques and Barracco, 2000
Penaeus longirostris	CD	440 210	27 36	NeuAc; Gal; GalNAc; BSM; Fetuin	Agglutination of bacteria (*P. aeruginosa*; *E. coli*)	Fragkiadakis and Stratakis, 1995

Abbreviations: CD: cation dependent; CI: cation independent; CPD: cation partially dependent; BSM: bovine mucin; OSM: ovine mucin; ESM: equine mucine; PSM: porcine mucine; GlcNAc: *N*-acetylglucosamine; GalNAc: *N*-acetylgalactosamine; ManNAc: *N*-acetylmannosamine; NeuAc: *N*-acetylneuraminic acid; LPS – lipopolysaccharide (the type/source is not specified); n.d.: not determined.

(Marques and Barracco, 2000)

3.2.1.6 Purification and characterization of lectins from invertebrates

A number of reports regarding the purification and characterization of crustacean lectins are now available (Marques and Barracco, 2000). However, compared to other arthropod groups, such as insects and horseshoe crabs, the current knowledge on lectin involvement in crustacean non-self-recognition is still much less well established. In contrast, studies on the proPO-activating system and related molecules (Söderhäll and Cerenius, 1992; Söderhäll *et al*., 1996) and, more recently, on clotting proteins (Kopacek *et al*., 1993; Komatsu and Ando, 1998; Hall *et al*., 1999) in crustaceans have well progressed. Most of reported lectins belong to the C-type lectins. The occurrence of C-type lectins has been reported in crustaceans for two species of the acorn barnacles, *M. rosa* and *Balanus rostratus* (Muramoto and Kamiya, 1989; Toda *et al*., 1998). It is relevant to point out that among C-type lectins, little or no homology may be present in domains other than that involved in CRD (Vasta *et al*., 1996).

Most of the literature dealing with crustacean lectins concerns the purification and characterisation of a unique lectin from the hemolymph and very rarely is its biological activity determined. This is quite intriguing, in view of the great economic interest of crustaceans, especially penaeid shrimps. Some examples of crustacean lectins, mainly from shrimps, functioning as potential non-self-recognition factors were shown in Table 4. In the penaeid *P. monodon*, Ratanapo and Chulavatnatol (1992) reported the agglutination of the pathogenic bacteria *V. vulnificus* by a purified lectin called monodin. The ability of the purified lectin from *P. californiensis* was investigated (Vargas-Albores *et al*. 1993) to react with different marine species of *Vibrio*. They demonstrated that the agglutinin of this penaeid was able to react to at least three different *Vibrio* species, *V. vulnificus*, *Vibrio fischeri and V. para-haemolyticus*. This reaction was specific and the agglutination of *V. parahaemolyticus* could be inhibited by N-acetyl galactosamine and LPS. The inhibition by LPS suggested that this natural ligand of the penaeid lectin could be one effective sign that triggered the shrimp immune system (Vargas-Albores *et al*., 1993). In the shrimp *Parapenaeus longirostris*, Fragkiadakis and Stratakis (1995) also

referred that purified lectins from the hemolymph that recognised *N*-acyl aminosugars strongly agglutinated formalin-fixed bacteria, *Pseudomonas aeruginosa* and *E. coli*.

Bacterial agglutination has been demonstrated with lectins from a wide variety of invertebrates such as the sea hare *Aplysia* sp. (Zipris *et al*., 1986), solitary ascidian *Halocynthia roretzi* (Azumi *et al*., 1991), black tiger shrimp *P. monodon* (Ratanapo and Chulavatnatol, 1992), Pacific oyster *C. gigas* (Hardy *et al*., 1977; Olafsen *et al*., 1992), freshwater prawn *M. rosenbergii* (Vazquez *et al*., 1996), edible crab *Scylla serrata* (Chattopadhyay *et al*., 1996) and horse mussel *Modiolus modiolus* (Tunkijjanukij and Olafsen, 1998).

The report of Ravindranath *et al*. (1985) has shown the specificity of the marine crab *Cancer antennarius* hemolymph lectin to 9-*O*-4-*O*-acetyl sialic acid, whereas the hemolymph lectins of *M. rosenbergii* and the marine crab *Liocarcinus depurator* have been shown to specifically recognize 9-*O*-acetyl sialic acid (Vazquez *et al*., 1993; Fragkiadakis and Stratakis, 1997). The agglutination of the bacterium *B. cereus* by *M. rosenbergii* hemolymph lectin can be related to the recognition of these *O*-acetylated sugars on the bacteria cell surface (Vazquez *et al*., 1996). On the other hand, a sialic acid-binding lectin with specificity for *N*-glycolyl neuraminic acid was purified from the hemolymph of the marine crab *S. serrata* (Mercy and Ravindranath, 1993). The unique binding specificity of this lectin distinguishes it from other known.

3.2.2 Lipopolysaccharide and β-1,3-glucan-binding protein (LGBP)

Lipopolysaccharide and β-1,3-glucan-binding protein (LGBP) play a crucial role in the innate immune response of invertebrates as a pattern recognition protein. LGBP or gram-negative binding protein (GNBP) family was originally discovered in the silkworm *Bombyx mori* (Ochiai and Ashida, 1988) and then from *Drosophila melanogaster*, and functions as a recognition protein for LPS and β-1,3-glucan (Kim *et al*., 2000). In addition, coelomic cytolytic factor-1 (CCF-1) from annelid *Eisenia fetida* which binds LPS and β-1,3-glucan is LGBP. Surface patterns of potential pathogens, such as bacteria and fungi, include β-1,3-glucan, peptidoglycan (PGN), LPS, and lipoteichoic acid (LTA), which can be recognized by pattern recognition proteins (Kanost *et al*, 2004; Yu *et al*, 2002). LPS is the principal component of the cell wall of gram negative bacteria whereas a common fungal call

wall molecule is β-1,3-glucan, which is a polymer of D-glucose β(1, 3) linkage with β(1, 6) linked side chains of varying length and distribution. Binding of these pattern recognition proteins in turn leads to activation of other immune responses. Over expression of DGNBP-1 in a *Drosophila* cell line enhanced antimicrobial gene expression induced by LPS- and β-1,3-glucans (Kim *et al*., 2000). In the mosquito *Anopheles gambiae*, GNBP messenger RNA (mRNA) level was elevated both locally in the midgut and systemically after infection of the mosquito with *Plasmodium berghei,* a parasite that enters the mosquito through the midgut (Richman *et al*., 1997). These binding proteins from insects appear to be functionally similar by having affinity to the gram negative bacterial cell wall and are inducible during injury or infection.

In crustaceans, BGBP was firstly cloned and identified for its properties involved in the activation of the proPO system (Cerenius *et al*., 1994). Crayfish BGBP is present in plasma and recognizes β-1, 3-glucan to promote the activation of the proPO system and also to act as an opsonin to increase the phagocytic activity. The BGBP and glucan complex may bind to crayfish granular hemocytes surface via its RGD (Arg-Gly-Asp) motif, suggesting binding to an integrin-like protein and then spreading and degranulation of crayfish hemocytes is induced (Lee and Söderhäll, 2001). An integrin β-subunit has been characterized as a candidate receptor for the BGBP and glucan complex (Cerenius *et al*., 1994; Duvic and Söderhäll 1990; Duvic and Söderhäll 1993; Thörnqvist *et al*., 1994). BGBP also interacts with an extracellular superoxide dismutase (SOD) involved in binding with peroxinectin (a cell adhesive and opsonic peroxidase) to the cell surface. Binding of peroxinectin to the cell surface may signal into the cell via an integrin for further cellular responses such as cell adhesion and phagocytosis (Johansson *et al*., 1999). Other LGBPs and BGBPs were also shown to mediate the activation of the proPO-system (Ma and Kanost, 2000) as well as induction of AMPs (Kim *et al*., 2000). These molecules do not exhibit any glucanase activity although they have similar primary structure to bacterial glucanases (Beschin *et al*., 1998; Cerenius *et al*., 1994; Ochiai and Ashida, 2000).

The first LGBP of shrimp was purified and cloned from *P. leniusculus*, which was proved to possess activities of binding to LPS and β-1,3-glucan (Lee *et al*.,

2000). It is present in hemocytes with a calculated molecular mass of 39.5 kDa. The function of *P. leniusculus* LGBP is bind to LPS or β-1,3-glucans such as laminarin and curdlan, but not to peptidoglycan, and to initiate activation the proPO activating system in crayfish as well as other insect carbohydrate binding protein (Lee *et al*., 2000). The primary structure and cloning results showed that LGBP has significant homology with several Gram-negative bacteria binding proteins and bacterial glucanases (Lee *et al*., 2000; Kim *et al*., 2000; Ochiai and Ashida, 2000; Ma and Kanost, 2000 and Beschin *et al*., 1998), but they lack glucanase activity. LGBP cDNA clones were also obtained from a hepatopancreas cDNA library of blue shrimp *Litopenaeus stylirostris* (Roux *et al*., 2002), a hemocyte cDNA library of *P. monodon* (Sritunyalucksana *et al*., 2002), from the hemocyte and hepatopancreas of *Litopenaeus vannamei* (Cheng *et al*., 2005) and from the hemocytes of fleshy shrimp *Fenneropenaeus chinensis* (Du *et al*., 2007) and the estimated molecular mass of these protein are 44, 39.5, 39.92 and 46 kDa, respectively. Analysis of the shrimp LGBP deduced amino acid sequence identified conserved features of this gene Family including a potential recognition motif for β-1,3-linkage of polysaccharides and putative RGD cell adhesion sites. The LGBP of *P. monodon* can bind curdlan and zymosan, which are β-1,3-glucan polymer, but not to LPS, so it was named β-1, 3-glucan binding protein (BGBP). Its expression is found constitutive in hemocytes by northern hybridization (Sritunyalucksana *et al*., 2002). Quantitative real-time RT-PCR analysis showed that LGBP gene expression in hepatopancreas of *L. stylirostris* was upregulated as the white spot syndrome virus (WSSV) infection progressed suggesting that shrimp LGBP is an inducible acute-phase protein that may play a critical role in shrimp-WSSV interaction (Roux *et al*. 2002) and the LGBP transcript in hemocyte of *L. vannamei* increased in 3 and 6 h after *V. alginolyticus* injection (Cheng *et al*., 2005). Similarly, the LGBP expression in hemocytes of the fleshy shrimp *F. chinensis* was significantly upregulated 6 h after being injected with mixture of *V. alginolyticus* and yeast, but returned to the original value after 12–24 h. Reported isolations and identifications of pattern recognition proteins including BGBP, LBP, and LGBP and their clones are shown in Table 5.

Table 5 Characteristics of pattern recognition proteins in crustaceans.

Species	Year	Name	Status	Size	References
Astacus astacus	1993	BGBP	Purified	95–105 kDa	Duvic *et al*., 1993
Carcinus maenas	1994	BGBP	Purified	110 kDa	Thörnqvist *et al.,* 1994
Farfantepenaeus californiensis	1993	LBP	Purified	175 kDa	Vargas-Albores *et al.,* 1993
Farfantepenaeus californiensis	1996	BGBP	Purified	100 kDa	Vargas-Albores *et al.,* 1996
Farfantepenaeus californiensis	1998	BGBP	Purified	112 kDa	Yepiz *et al.,* 1998
Pacifastacus leniusculus	1990	BGBP	Purified	100 kDa	Duvic and Söderhäll, 1990
Pacifastacus leniusculus	2000	LGBP	Purified	40 kDa	Lee *et al.,* 2000
Procambarus clarkia	1993	BGBP	Purified	100 kDa	Duvic *et al.,* 1993
Litopenaeus stylirostris	1997	BGBP	Purified	100 kDa	Vargas-Albores *et al.,* 1998
Litopenaeus vannamei	1997	BGBP	Purified	100 kDa	Vargas-Albores *et al.,* 1998
Litopenaeus vannamei	1998	BGBP	Purified	112 kDa	Yepiz *et al.,* 1998
Penaeus monodon	2002	BGBP	Purified	31 kDa	Sritunyalucksana *et al*., 2002
Pacifastacus leniusculus	1994	BGBP	Cloned	4,679 bp	Cerenius *et al.,* 1994
Pacifastacus leniusculus	2000	LGBP	Cloned	1,650 bp	Lee *et al.,* 2000
Penaeus monodon	2002	BGBP	Cloned	1,314 bp	Sritunyalucksana et *al*., 2002
Litopenaeus stylirostris	2002	LGBP	Cloned	1352 bp	Roux *et al*., 2002
Litopenaeus vannamei	2004	BGBP–HDL	Cloned	6,379 bp	Romeo-Figueroa *et al*., 2004
Litopenaeus vannamei	2005	LGBP	Cloned	1,272 bp	Cheng *et al*., 2005
Fenneropenaeus chinensis	2007	LGBP	Cloned	1,253 bp	Du *et al*., 2007
Fenneropenaeus merguiensis	2008	LGBP	Cloned	1,280 bp	**Present study**

The characteristics of pattern recognition proteins: LBP (lipopolysaccharide binding protein), BGBP (β-1, 3-glucan binding protein), and LGBP (lipopolysaccharide- and β-1, 3-glucan binding protein), in crustaceans that have been purified and cloned so far.

3.2.3 The prophenoloxidase activating system

Invertebrates, including crustaceans, do not produce antibodies and rely on innate immune system to recognize and respond to pathogens or environmental antigens (Söderhäll and Cerenius, 1998; Sritunyalucksana and Söderhäll, 2000; Johansson *et al*., 2000; Cerenius and Söderhäll, 2004). One of their major defense mechanisms is the melainzation of pathogens and damaged tissues. A melanization reaction can be triggered by the prophenoloxidase cascade (Cerenius and Söderhäll, 2004). Moreover, melanization prevents or retards the growth of intruders because highly reactive and toxic quinone intermediates are produced when melanin is formed (Söderhäll and Cerenius, 1998; Ashida and Brey, 1998; Sugumaran, 1996). An enzyme involved in this important process is phenoloxidase (PO). PO is synthesized as a zymogen, prophenoloxidase (proPO), which can be activated by specific proteolysis. A cascade of serine proteinase and other factors has been recognized to control the activation of proPO into PO and is known as proPO activation system (PAS) (Söderhäll and Cerenius, 1998; Sritunyalucksana and Söderhäll, 2000; Cerenius and Söderhäll, 2004). PO is a bifunctional copper containing enzyme, also known as tyrosinase, which catalyses two successive reaction; the first is the hydroxylation of a monophenol to an *o*-diphenol, and the second is the oxidation of *o*-diphenol to *o*-quinone (Sugumaran, 2002). The production of toxic quinone intermediates and *o*-quinones by phenoloxidase is an initial step in the biochemical cascade of melanin biosynthesis, and is also important in cuticular scleotization, wound healing, and in the encapsulation of foreign emterials for host defense (Cerenius and Söderhäll, 2002). PO converts tyrosine to 3,4-dihydroxy-L-phenylalanine (DOPA), from DOPA to DOPA-quinone, and several subsequent intermediate steps that lead to the synthesis of melanin, a brown pigment (Fig. 12). The conversion of inactive proPO to active PO is mediated by a serine protease named the proPO activating enzyme (PPAE). PPAE has been characterized and cloned from four different animals; the beetle *Holotrichia diomphalia* (Lee *et al*., 1998a; b), silkworm *B. mori* (Satoh *et al*., 1999), tobacco hornworm *Manduca sexta* (Jiang *et al*., 1998), and crayfish *P. leniusculus* (Wang *et al*., 2001). Several proteinase inhibitors that prevent over activation of PPAE (Aspan *et al*., 1990; Liang *et al*., 1997; Park *et al*., 2000) and phenoloxidase inhibitors, belonged to serpins

(serine proteinase inhibitors), such as pacifastin that inhibit directly the activity of phenoloxidase (Daquing *et al*., 1999; Sugumaran and Nellaiappan., 2000) have been identified in crayfish and several other arthropod species (Fig. 12).

Carbohydrates on microbial cell wall such as β-1,3-glucan (Vargas-Albores and Yepiz-Plascencia, 2000) and LPS (Lee *et al*., 2000), through pattern-recognition molecules, can trigger the PAS by initiating the serine proteinase cascade. Several proteins such as peroxinectin, transglutaminase and clotting protein are components of or associated with the PAS. They have been elucidated to be involved cell adhesion, degranulation, encapsulation enhancement, phagocytosis and cytotoxic reactions (Sritunyalucksana and Söderhäll, 2000; Cerenius and Söderhäll, 2004). There is evidence that the PAS and its associated factors provide an extensive defense against foreign materials in invertebrates.

Since the first invertebrate proPO gene was cloned from the freshwater crayfish *P. leniusculus* in 1995 (Aspan *et al*., 1995) such enzymes have been cloned from about 20 different arthropod species, including the tiger shrimp (Sritunyalucksana *et al*., 1999), the Pacific white shrimp *L. vannamei* (Lai *et al*., 2005) and the giant freshwater prawn *M. rosenbergii* (Liu *et al*., 2006). There also are unpublished proPO sequences of *Marsupenaeus japonicus*, green tiger prawn *Penaeus semisulcatus* and American lobster *H. americanus* that are deposited in the GenBank as of May 2007.

The crustacean proPO contains highly conserved regions that correspond to the two potential copper-binding domains of arthropod hemocyanin, a hexameric oxygen carrier protein in lower arthropods. PO is one of these proteins that are included in the hemocyanin superfamily and is suspected to have evolved from the same ancestral binuclear copper protein as arthropod hemocyanin (Burmester and Scheller, 1996). Although the proPO possesses two conserved copper binding sites, it is still unknown whether it is capable of, or is involved in oxygen transport. Hemocyanin, however, was reported to have phenoloxidase activity after the proteolytic cleavage of its N-terminal part (Nagai *et al*., 2001; Pless *et al*., 2003).

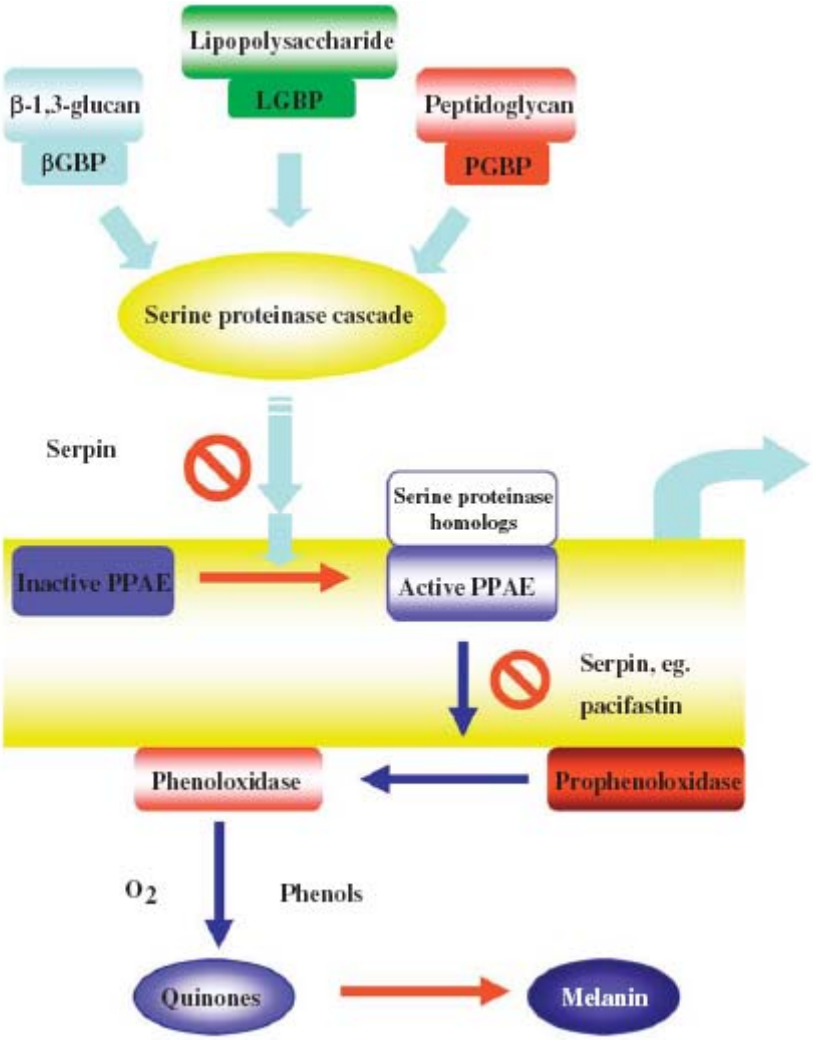


Fig. 12 Overview of prophenoloxidase activating system. ProPO, prophenoloxidase activating: PPAE, proPO-activating enzyme. (adopted from Jiravanich-paisal *et al*., 2006).

3.2.4 The clotting system

The clotting mechanism is an essential defense response of crustaceans that entraps foreign material and prevents loss of hemolymph (Martin *et al*., 1991). Only two different coagulation mechanisms have been characterized in molecular detail in invertebrates so far; the hemocyte-derived clotting cascade in horseshoe crab Tachypleus tridentatus (Kawabata *et al*., 1996) and the transglutaminase dependent clotting reaction in crayfish *P. leniusculus* (Kopacek *et al*., 1993; Hall *et al*., 1999).

Transglutaminases (TGase) are important for blood coagulation and post-translation remodeling of proteins. The TGase-dependent clotting reaction (Fig. 13) in crayfish is induced when TGase is released from hemocytes of tissues, becomes activated by Ca^{2+} in plasma and starts to crosslink the clotting protein molecules found in the plasma, to form large aggregates (Kopacek *et al*., 1993; Yeh *et al*., 1998). TGases are Ca dependent enzymes capable of forming covalent bonds between the side chains of free lysine and glutamine residues on certain proteins. TGase was first characterized and cloned from *Limulus* hemocytes (Tokunaga *et al*., 1993a; b). Its role during clotting is unclear. Crustacean TGases were first cloned from crayfish *P. leniusculus* (Wang *et al*., 2001) and then from *P. monodon* (Huang *et al*., 2004). The crayfish hemocytes, both SGCs and GCs, as well as the muscle tissue, show TGase activity, whereas the hepatopancreas and plasma have no TGase activity (Wang *et al*., 2001). In contrast, *P. monodon* showed TGase activity in all organs tested (Huang *et al*., 2004). The widespread and low-level expression of TGase mRNA is ubiquitous. However, high levels of TGase expression were detected in hematopoietic tissue of *P. monodon*. Crayfish TGase is expressed exclusively in the hemocytes and expressions in muscle and the hepatopancreas are very low (Wang *et al*., 2001).

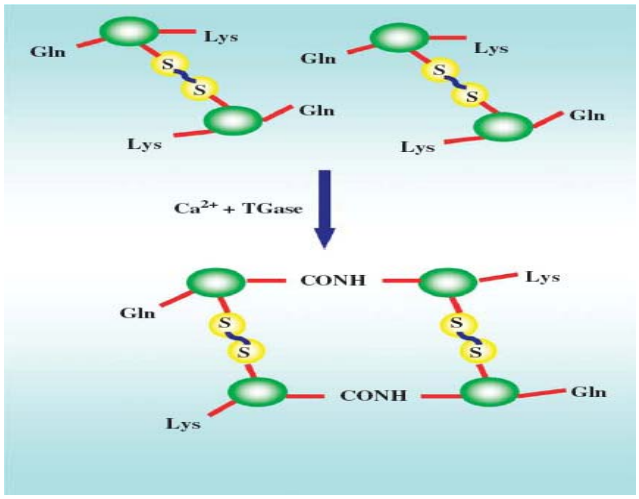


Fig. 13 Transglutaminase-mediated clotting reaction in crayfish and shrimp.

Green oval: clotting protein which is a dimeric protein consisting of two identical 210-kDa subunits held together by disulfide bond (●-s-s-●) are cross-linking by TGase in the presence of Ca^{2+} filled hexagon (adopted from Jiravanichpaisal *et al*., 2006).

3.2.5 Antimicrobial peptides (AMPs)

Antimicrobial peptides appear to be ubiquitous and are multipotent components of the innate immune defense arsenal used by both prokaryotic and eukaryotic organisms (Bulet *et al*., 1999). As the incidence of microbial resistance to existing antibiotics increase, AMPs are a promising resource for therapeutic and pharmaceutical alternatives. Some AMPs have been clinically tested against multidrug-resistant strains of bacteria (Giacometti *et al*., 2000) and some AMPs may also be useful in treatment of some infectious viruses (Murakami *et al*., 1991) and fungi (Destoumieux *et al*., 1999). Of the more than 400 AMPs reported, over 50% have been identified in invertebrates and more than 170 AMPs in have been identified in insects. In contrast, few AMPs were characterized in crustaceans.

Antimicrobial activity of AMPs has been demonstrated in the hemocytes, plasma and in the hepatopancreas of several crustaceans (Evans *et al*., 1968; Adams, 1991; Chisholm and Smith, 1995; Noga *et al*., 1996a; b). Even though the evidence of their antimicrobial activity has existed for many years, characterization of the first AMP in crustaceans did not occur until 1996 when Schnapp *et al*. (1996) isolated and partially sequenced a proline-rich AMP from the shore crab *Carcinus maenas*. This 6.5 kDa peptides exhibited activity against both gram-negative and gram-positive bacteria. Another 3.7 kDa AMP named callinectin was also partially characterized from blue crab *Callinectes sapidus* (Khoo *et al*., 1999). In penaeid shrimps, two kinds of AMPs have been fully characterized, namely the penaeidins and crustin from hemocytes.

3.2.5.1 Penaeidins

Penaeidins were initially characterized from *L. vannamei* by biochemical approaches and molecular cloning. The three peptides PEN2-1, PEN2-2, and PEN-3 were isolated in their active and mature form (5.48-6.62 kDa) from the hemocytes of the animals collected from intensive shrimp farms (Destoumieux *et al*., 1997). According to database and sequence alignments, additionally to PEN-2 and PEN-3, a third distinct subgroup, penaeidin-4 (PEN-4), has been identified (Cuthbertson *et al*., 2002). By biochemical approach, the PEN-3 subgroup was shown to be abundantly produced (Destoumieux *et al*., 1997), representing more than 90% of

all the penaeidin mRNA sequences detected in both *L. vannamei* and *L.setiferus* (Cuthbertson *et al*., 2002).

Many reports now showed that penaeidins appear to be a family of AMPs ubiquitous among penaeid shrimp. Indeed, recent studies based on a genomic approach have revealed the presence of penaeidins in different penaeid shrimp species (Gross *et al*., 2001; Rojtinnakorn *et al*., 2002; Supungul *et al*., 2002). To date, penaeidin sequences have been described in *L. vannamei*, Atlantic white shrimp *L. setiferus*, *L. stylirostris*, *P. semisulcatus*, *M. japonicus*, *P. monodon* and *F. chinensis* (Bachere *et al*., 2004). All these peptides share high sequence similarities and contain a highly conserved signal peptide, a proline-rich amino terminal domain and a carboxy-terminal domain containing 6 cysteines, which form 3 intramolecular disulfide bridges. The C-terminal provides the penaeidins with their broad-spectrum activity against filamentous fungi (Destoumieux *et al*., 1999).

Penaeidins have a broad spectrum of anti-fungal activities against filamentous fungi. Their antibacterial activity is predominantly directed against gram-positive bacteria via a strain-specific inhibition mechanism and through multiple modes of action (Destoumieux *et al.,* 1999). They also display weak antibacterial activity in vitro against gram-negative strains including Vibrionaceae species, a family of bacteria commonly associated with shrimp mortalities. They also presented chitin-binding ability through their C-terminal domain (Destoumieux *et al*., 2000). This ability might be related to their anti-fungal activity or involue in chitin assembly and wound healing processes. A model for shrimp immune response to microbial challenge via penaeidin was depicted in Fig. 14 (Bachere *et al*., 2004). Under normal conditions, penaeidin-producing hemocytes are located either in blood vessels or widespread in the shrimp tissues (Fig. 14A). The response to microbial infection can be divided in two phases. Phase I, or local reaction (Fig. 14B), is characterized by the migration of the hemocytes towards the site of infection and by a massive release of penaeidins. Free penaeidins and hemocyanin derived peptides are seen in circulation. Phase II, or systemic reaction (Fig. 14C), is characterized by a massive accumulation of penaeidin producing hemocytes, as a result of the activation of hematopoiesis. Penaeidins are seen bound to cuticular surfaces.

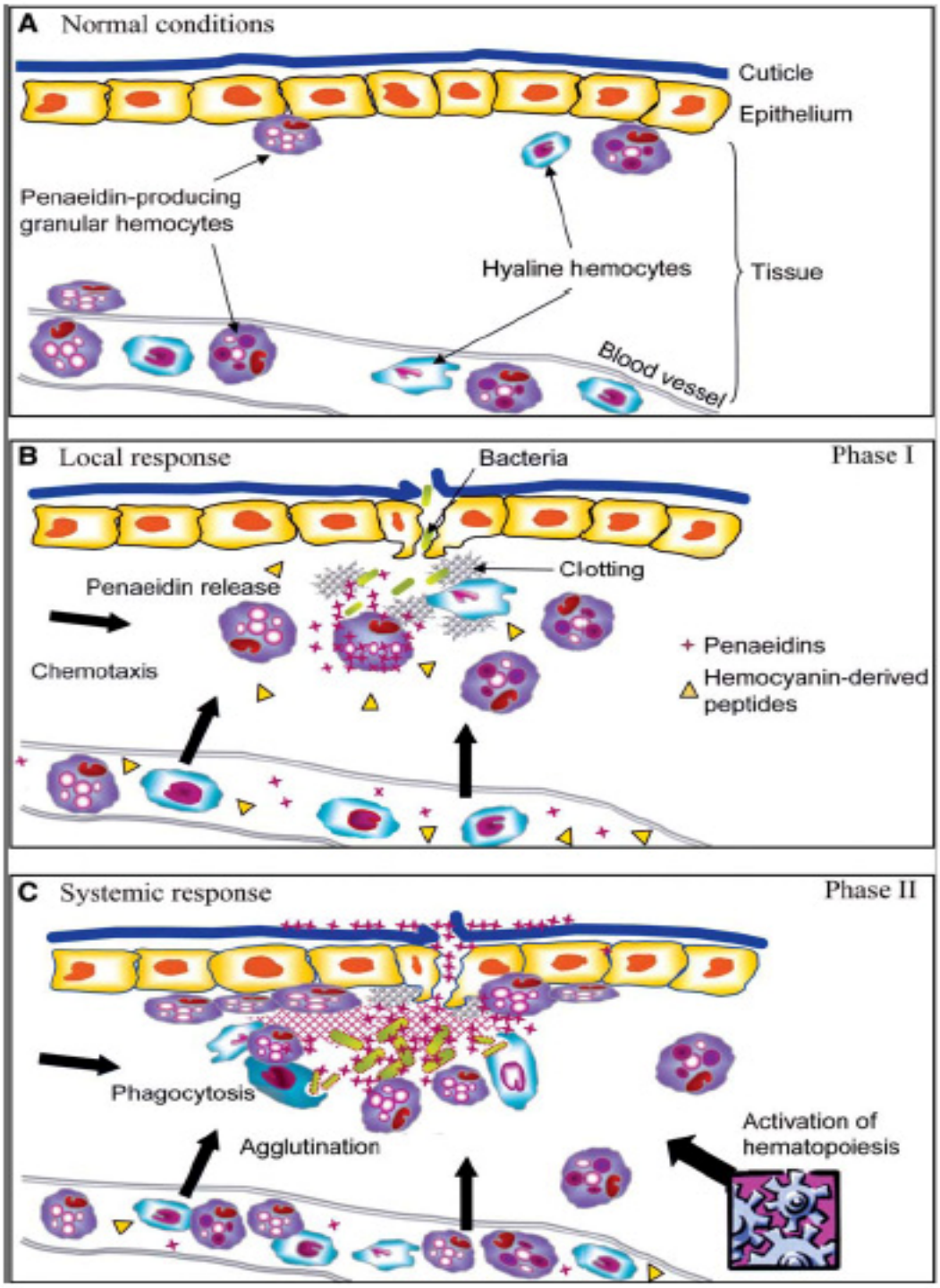


Fig. 14 Proposed model of the antimicrobial response in penaeid shrimp, involving the antimicrobial peptides.

(adopted from Bachère *et al*., 2004)

3.2.5.2 Crustins

Another common group of AMPs of decapod crustaceans is the crustin family. Granular hemocytes of *C. maenas* were the first source of the crustin family of putative AMPs (Relf *et al*., 1999). The 11.5 kDa crustin peptide is cysteine-rich, hydrophobic and exhibits specific activity towards gram-positive marine bacteria (Relf *et al*., 1999). Crustins have also been reported in the *L. vannamei*, *L. setiferus* (Bartlett *et al*., 2002), spiny lobster *Panulirus argus* (Stoss *et al*., 2003), *M. japonicus* (Rattanachai *et al*., 2004) and *P. monodon* (Supungul *et al*., 2004). Recent genomic studies have revealed the presence of crustins in other crustacean species including crayfish *P. leniusculus* (Jiravanichpaisal et al., 2007), *F. chinensis* (Zhang *et al.,* 2007), European lobster *Homarus gammarus*, (Hauton *et al.,* 2007) and American lobster *H. americanus* (Christie *et al.,* 2007). Two different isoforms (i.e. I and P) of crustin were found in the Pacific white shrimp. They are differentiated based on two nucleotide differences that result in the change in amino acid sequence (Ile/Pro) at position 108 (Vargas-Albores *et al.,* 2004). The steady state levels of crustin I mRNA, were down regulated by *V. alginolyticus* inoculation but not of crustin P. The expressions of crustins in crayfish and lobster also responded to microbial infection (Jiravanichpaisal *et al.,* 2007; Hauton *et al*., 2007). The recombinant *F. chinensis* crustins could inhibit *in vitro* the growth of gram-positive bacteria and the four-disulfide core domain of crustin may play an important role in its biological function (Zhang *et al.,* 2007).

3.2.5.3 Lysozyme

Lysozyme (muamidase, EC.3.2.1.17) is widely distributed among eukaryotes and prokaryotes. It catalyzes the hydrolysis of bacterial cell walls and acts as a nonspecific innate immunity molecule against the invasion of bacterial pathogens (Jolles and Jolles, 1984). Lysozymes are classified as six types: chicken-type lysozyme (c-type), which includes stomach lysozyme and calcium-binding lysozyme, goose-type lysozyme (g-type), plant lysozyme, bacteria lysozyme, T4 phage lysozyme (phage-type), and invertebrate lysozyme (i-type) (Jolles, 1996). The i-type lysozyme was first described on the starfish *Asterias rubens*, by Jolles and Jolles (1975). Recently, it has been extensively studied as a novel member of the lysozyme family (Ito *et al*., 1999; Nilsen and Myrnes, 2001; Bachali *et al.,* 2002).

Although the role of lysozyme in insect immunity has been extensively studied and reviewed (Hultmark, 1996), its role in crustaceans has only been reported in the freshwater crayfish (Fenouil and Roch, 1991) and brine shrimp *Artemia franciscana* (Stabili *et al.,* 1999). It was not found in the plasma of *L. vannamei* (Alabi *et al*., 2000). Recently, the C-type lysozyme cDNA was cloned and characterized from the *L. vannamei* and *M. japonicus* (Sotelo-Mondo *et al*., 2003; Hikima *et al*., 2003). A complete cDNA sequence for *P. monodon* C-type lysozyme is also available in GenBank (AF539466). The presence of lysozyme was confirmed directly from shrimp hemocytes and mRNA was detected in the hemocytes by RT-PCR. The enzymatic activity of shrimp lysozyme was detected in hemocyte lysate, but not in plasma, indicating that lysozyme is expressed, translated and stored inside the cells. Recombinant *M. japonicus* lysozyme displayed lytic activities against several *Vibiro* species. Sequences of I-type lysozyme are available for the *L. vannamei* (BF023863, BF024192) and the Atlantic white shrimp *L. setiferus* (BF024309) in EST databases in the National Center for Biotechnology Information (NCBI). However, no cDNA for I-type lysozyme has been successfully cloned till now.

Objectives

The aims of this thesis were to study genes in the immune system especially those of lectin and LGBP in *F. merguiensis*, as followed.

1. To purify lectin from the hemolypmh of *F. merguiensis*.
2. To characterize purified lectin.
3. To clone and characterize full-length *F. merguiensis* lectin (*FmL*) and LGBP (*FmLGBP*) cDNAs.
4. To analyze primary structures of deduced amino acid sequences of lectin and LGBP of *F. merguiensis*.
5. To determine the expression sites of lectin and LGBP genes in *F. merguiensis* by RT-PCR.
6. To construct phylogenetic tree and study comparatively the sequences of lectin and LGBP of *F. merguiensis* with those of other shrimps.
7. To express the recombinant FmL proteins in *E. coli*.

Chapter 2

Materials and methods

Materials

1. Animals

1.1 Banana shrimp *Fenneropenaeus merguiensis*

Live mature female and male *F. merguiensis* were collected from Trang Coastal Aquaculture Station and the Sichon district in Nakhon Si Thammarat province. They were 10-14 cm long and weighed between 30-40 g.

1.2 Rabbits

To draw the blood for hemagglutination test, albino rabbits weighing 1.5 kg was used. The rabbit was fed with normal food at the Southern Laboratory Animal Facility, Prince of Songkla University (PSU).

2. Chemicals

2.1 Analytical grade

A chemicals used were of analytical grade and purchased from the following companies.

Chemical	Company
Acetic acid	Merck
Acrylamide	Fluka
Agarose	Sigma
Ammonium persulphate	Merck
Ammonium sulphate	Fluka
Beta-Mercaptoethanol	Fluka
Bisacrylamide (N,N' methylenediacrylamide)	Fluka

Chemical	**Company**
Bovine serum albumin	Sigma
Bromophenol blue	Merck
Calcium chloride	Ajax
Carboxymethyl-Cellulose	Sigma
Chloroform: isoamyl alcohol (49:1)	Fluka
Citric acid	Ajax
Coomassie brilliant blue G-250	Bio-Rad
Coomassie brilliant blue R-250	Sigma
Coomassie plus protein assay reagent kit	Pierce
3-(Cyclohexylamino)-1-propane-sulfonic acid	Sigma
Diethylpyrocarbonate	Sigma
Ethylenediaminetetraacetic acid	Fluka
DL-dithiothreitol	USB Corporation
Ethanol	Scharlau
Ethidium bromide	Promega
Fetuin-agarose	Sigma
Formaldehyde	Sigma
Formamide	Sigma
Glycerol	Sigma
Glycine	Fluka
Glucose	Ajax
Guanidine thiocyanate	Sigma
Heparin	Difco Laboratories
Hydrochloric acid	Merck
Imidazole	Sigma
Isopropanol	Lab-scan
Low molecular weight standard	Amersham Pharmacia Biotech
Lysozyme	Sigma
Magnesium chloride (hexahydrate)	Sigma
Methanol	Merck
N-Acetyl glucosamine	Sigma
N,N,N',N'-Tetramethylethylenediamine	Fluka

Chemical	Company
3-(N-morpholino)-propanesulfonic acid	Sigma
Phenylmethylsulphonylfluoride	Sigma
Phosphoric acid	BDH Laboratory
Polyethylene glycol	Sigma
Potassium bromide	Merck
Potassium dihydrogen phosphate	Merck
N-Lauroylsarcosine sodium salt	Sigma
Silver staining kit	Bio-Rad
Sodium acetate (trihydrate)	Fisher scientific
Sodium chloride	Sigma
Sodium citrate (dihydrate)	Ajax
Sodium dodecyl sulphate	Riedel de-Haen
Sodium hydroxide	Lab-scan
Sodium metabisulphite	Mallinckrodt
Standard bovine serum albumin	Pierce
Tris (hydroxymethyl) aminomethane	Sigma
Triton X-100	Merck
Urea	Amresco

2.2 Molecular biology grade

All chemicals used below were molecular biology grade and purchased from the following companies.

Chemical	Company
Agar	BD Bioscience
Agarose	Promega
Ampicillin	Sigma
100 Base pair DNA ladder	Biolabs
1 kb DNA ladder	Biolabs
5-Bromo-4-chloro-3-indolyl-beta-D-galactopyranoside	USB Corporation
Isopropylthiogalactoside	USB Corporation
Luria Bertani broth	USB Corporation
Tetracycline	Sigma

2.3 Plasmid vectors

Plasmid vectors used in this study are listed as below and the physical maps of the vectors pDrive, pCR 4-TOPO, pGEM T-easy and pET32a(+) are shown in Fig. 15, Fig. 16, Fig. 17 and Fig. 18, respectively.

Plasmids	Description of use	Antibiotic	Company
pCR 4-TOPO	Cloning PCR products	Ampicillin	Invitrogen
pDrive	Cloning PCR products	Ampicillin	Qiagen
pGEM T-Easy	Cloning PCR products	Ampicillin	Promega
pET32a(+)	Expression vector	Ampicillin	Novagen

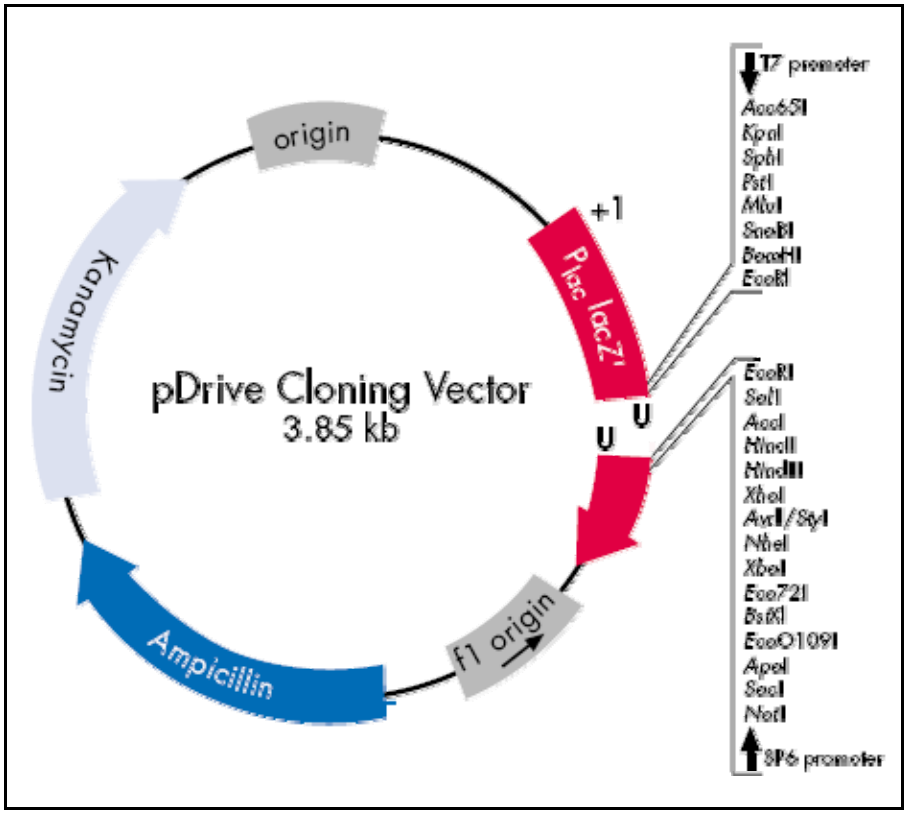


Fig. 15 The physical map of pDrive cloning vector.

The linearized vector has U overhangs and the restriction endonuclease recognition sites of the cloning site are listed.

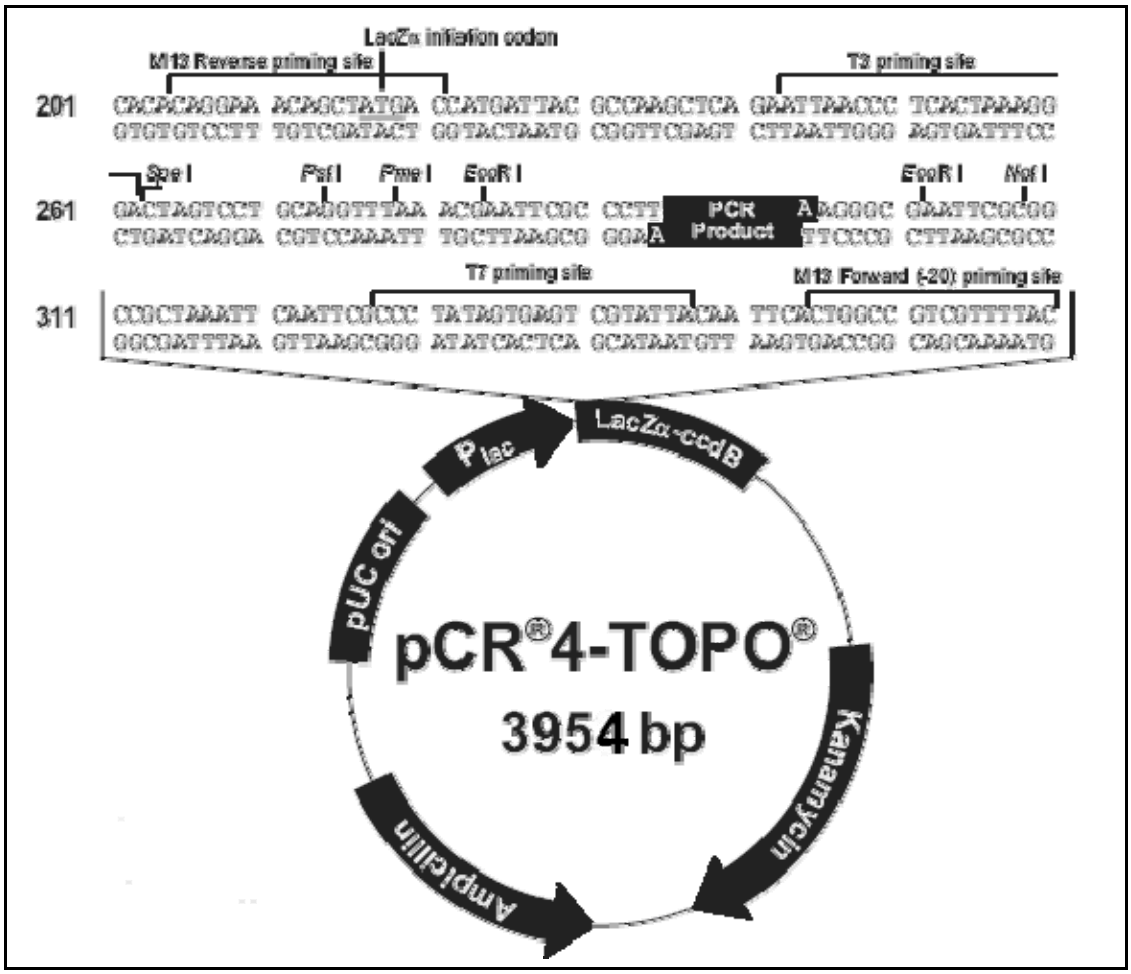


Fig. 16 The physical map of pCR 4-TOPO cloning vector.

The linearized vector has T overhangs and the restriction endonuclease recognition sites of the cloning site are listed.

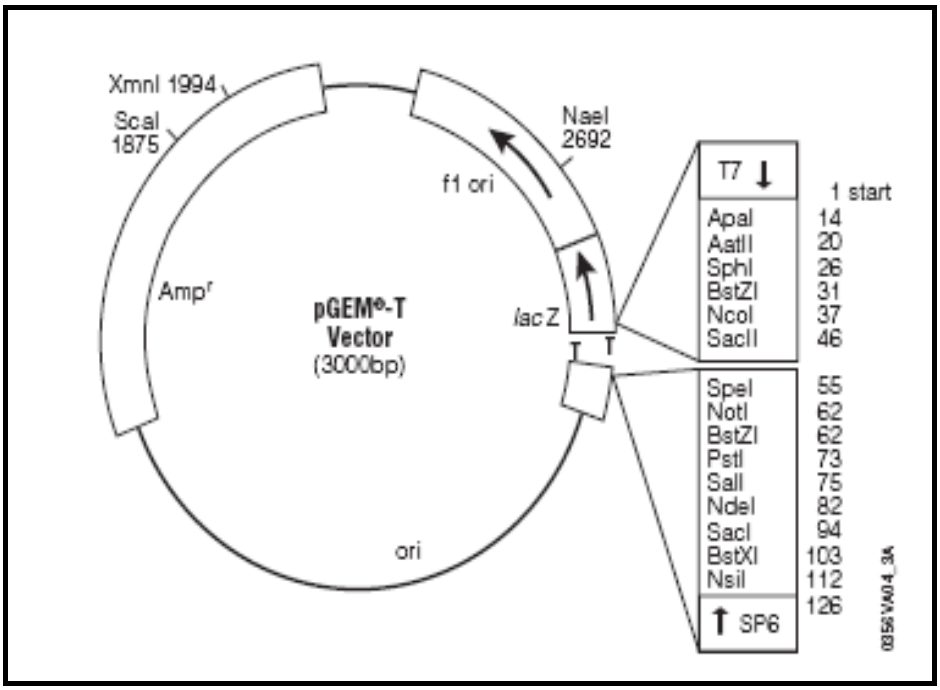
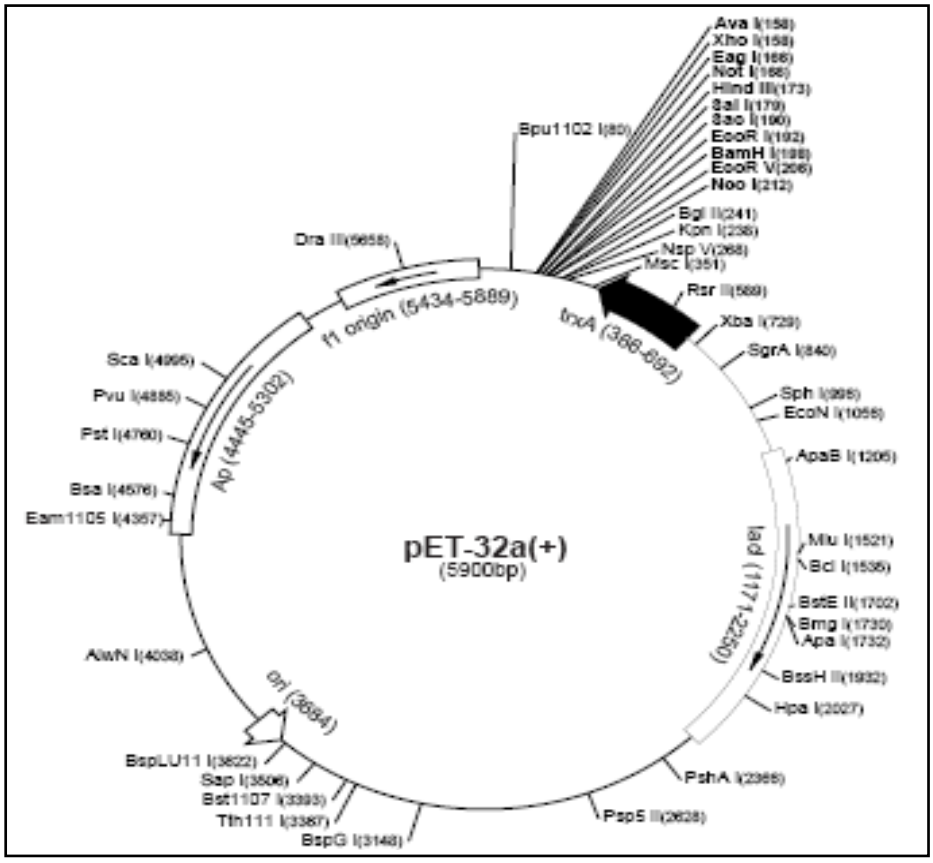


Fig. 17 The physical map of pGEM T-easy cloning vector.

The linearized vector has T overhangs and the restriction endonuclease recognition sites of the cloning site are listed.



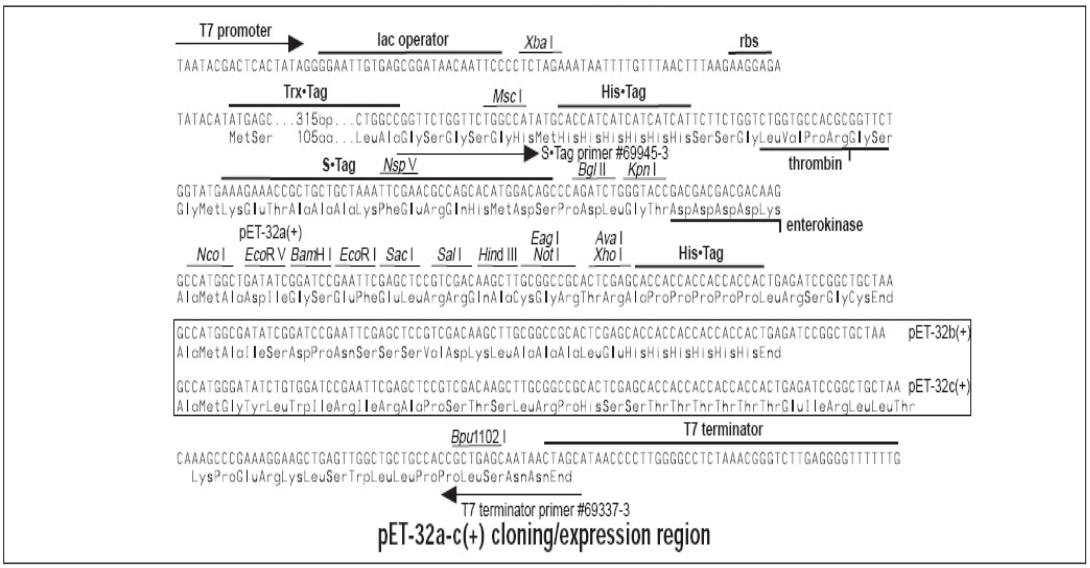


Fig. 18 The physical map of pET32a(+) expression vector.

The pET32a(+) vector is designed for cloning and high-level expression of peptide sequences fused with the 109 amini acids of thioredoxin (Trx•Tag). Cloning sites are available for producing fusion proteins also containing cleavable His6•Tag and S•Tag sequence for detection and purification.

2.4 Bacterial strains

Five strains of *Escherichia coli* were used and purchased from Invitrogen, Qiagen and Stratagene as follows.

Invitrogen : *E. coli* strain TOP10 [F^- *mcr* A Δ(*mrr-hsd*RMS-*mcr*BC) Φ80*lac*ZΔM15 Δ*lac*X74 *rec*A1 *ara*D139 Δ(*ara-leu*)7697 *gal*U *gal*K *rps*L (Str^R) *end*A1 *nup*G],

Invitrogen : *E. coli* strain BL21 (DE3) star F^- *ompT hsdS*$_B$ (rB^-mB^-) *gal dcm rne*131 (DE3)

Invitrogen : *E. coli* strain DH5α F^- Φ80*lac*ZΔM15 Δ(*lac*ZYA *arg*F) U169 *deo*R *rec*A1 *end*A1 *hsd*R17(rk^-, mk^+) *pho*A *sup*E44 *thi*-1 *gyr*A96 *rel*A1 λ^-

Qiagen : *E. coli* strain Qiagen EZ [F'::Tn*10*(Tc^r) *pro*A^+B^+ *lacI*^{q}ZΔM15] *recA1 end A1 hsdR 17*($r_{K12}^-m_{K12}^+$) *lac glnV44thi-1 gyrA96 relA1*.

Stratagene : *E. coli* strain XL1-Blue MRF' Δ(*mcrA*)*183* Δ(*mcrCB-hsdSMR-mrr*)*173 endA1 supE44 thi-1 recA1 gyrA96 relA1 lac* [F′ *proAB lacI*qZΔ*M15* Tn*10* (Tetr)]

2.5 Primers

The oligonucleotide primers were custom-ordered from either Qiagen and Life Technologies, Invitrogen life technologies or Bio Service Unit (BSU), National Science and Technology Development Agency (NSTDA), Thailand.

2.6 Enzymes

Restriction enzymes and other modification enzymes were purchased from New England Biolabs, Invitrogen, Roche Applied Science, Clontech and Promega.

3. Instruments

Instrument	Model	Company
Autoclave	ES-315	Tomy
Automated DNA sequencer	ABI PRISM 3100 or 377	Applied Biosystems
Bacteria incubator	1510E	Shel Lab
Balance (4 digits)	AB204-S	Mettler
Balance (3 digits)	GT410	Ohaus
Balance (2 digits)	PG5002-S	Mettler
Centrifuge	Avanti J 30I; J2-21	Beckman
Centrifuge	5415C; 5415R; 5804R	Eppendorf
Gel document	BioDoc-It^{TM}System	UVP
Glass-teflon homogenizer	-	Thomas
Gradient makers	SG 15	Hoefer
Heating box	AccuBlock	Labnet
Imaging densitometer	GS-700	Bio-Rad
Microcentrifuge	SD220	Clover
Micropipettes	-	Gilson, Eppendorf
Oven	-	Binder
Orbital shaker	-	GallenKamp
pH meter	713	Metrohm
Protein/peptide sequencer	ABI473A	Applied Biosystems
Proteomic analyzer	4700	Applied Biosystems
Power supply	AE-8150; AE-8450	Atto
Power supply	1000/500	Bio-Rad
PCR	Master Cycler	Eppendorf
Shaking incubator	SI-23MC	Bioer technology
Slab gel electrophoresis	AE-6450	Atto
Slab gel electrophoresis	AE-6530	Atto
Submarine electrophoresis	Mupid-ex	Advance
UV-VIS Spectrophotometer	160A	Shimadzu
Vortex-mixer	G-560E	Scientific Industries
Water bath	EcoTempTW20	Julabo

Methods

1. Preparation of hemolymph samples

Hemolymph was collected from pericardial sinus with a 26G gauge needle attached to a syringe and transferred immediately to tubes. Serum was obtained by allowing the hemolymph to clot at 4°C overnight. Serum was cleared by centrifugation at 16,000xg for 30 min at 4°C and the supernatant was immediately used or stored as aliquots at -20°C for further use.

2. Preparation of hemocytes

Shrimp hemolymph was collected from the ventral sinus located at the third abdominal segment mixing with an equal volume of Modified Aiservier Solution (MAS) [27 mM Na-citrate, 336 mM NaCl, 115 mM glucose, 9 mM ethylenediaminetetraacetic acid (EDTA), pH 7.0] used as anticoagulant, and centrifuged at 800xg for 10 min. Hemocyte pellets were suddenly preserved in liquid nitrogen for RNA extraction.

3. Hemaggutination assays

Rabbit erythrocytes were obtained from a healthy animal. Prior to use, the erythrocytes were washed several times by centrifugation at 800xg for 10 min at 4°C with TBS (50 mM Tris-HCl, pH 7.5-0.15 M NaCl). A 2% suspension of the erythrocytes was prepared in the same buffer. The hemagglutinating activity (HA) of lectin was assayed in a microtiter U plate (NUNC, Denmark) according to a two-fold serial dilution procedure. To each well, 25 µl of the serum diluted two-fold in TBS was mixed with an equal volume of 2% rabbit erythrocyte suspension. Hemagglutination was observed 1 h after incubation at room temperature. The activity was defined as the titer value of maximum dilution with positive agglutination of 2% rabbit erythrocytes. Positive hemagglutination was obtained when the erythrocytes did not sediment to the bottom of the well forming a red button. Specific hemagglutinating activity was defined as HA unit/mg protein.

4. Repetitive purification of lectin from serum by Fetuin-agarose column

Lectin was purified from *F. merguiensis* as reported in Rittidach *et al.* (2007) with some modifications. In brief, five milliliters of clarified serum were loaded onto a column previously equilibrated with TB-NaCa (50 mM Tris-HCl, pH 7.5-0.3 M NaCl-0.1 M $CaCl_2$). After being retained in the column at 4°C for 12 h to allow for complete binding, unbound material was eluted from the column with the equilibrating buffer at a 15ml/h flow rate. Fractions of 1 ml were collected and the optical density was measured at 280 nm ($O.D._{280}$) until absorbance of the eluent at 280 nm was below 0.01. The lectin was desorbed by 100 mM GlcNAc (N-acetyl glucosamine) in TB-NaCa. The eluted fractions were dialyzed overnight against TBS at 4°C. The samples were assayed for their HA by the 2-fold dilution procedure in the presence of a 2% solution of rabbit erythrocytes. The active fractions with high HA were pooled, concentrated and dialyzed overnight against 50 mM Tris-HCl, pH 7.5 at 4°C. The active samples from the first Fetuin-agarose column with high HA were subjected to the second Fetuin-agarose column. The steps which used in the first column were utilized to purify lectin. The purity of lectin was analyzed by non-denaturing PAGE and kept at -20°C for further analysis.

5. Determination of protein

5.1 Bradford method (Bradford, 1976)

The Bradford assay is a rapid and accurate method commonly used to determine total protein concentrations of samples. One milliliter of Coomassie plus protein assay reagent kit or Bradford reagent (0.01% Coomassie brilliant blue G-250-4.7% ethanol- 8.5% phosphoric acid) was added to each sample tube and mixed well by inversion or gentle vortex mixing for 2-3 min. The $O.D._{595}$ of the samples was measured during 2-60 min after mixing. Protein standards in a range of 0 to 9 μg BSA were prepared from a stock of 0.5 mg/ml bovine serum albumin (BSA) and used to determine protein concentration in the test samples taking into consideration any dilution factor.

5.2 Bicinchoninic acid (BCA) method (Smith *et al.*, 1985)

Twenty-five microliters of each standard or samples was added into the appropriate microtiter plate wells. The diluents were mixed with 200 μl of BCA assay

reagent kit for 30 second. The plate was covered and incubated at 37°C for 30 min. The absorbance was measured at 562 nm by an ELISA plate reader after the plate was cooled to room temperature. The standard curve was prepared from 0-250 µg/ml of BSA.

6. Polyacrylamide gel electrophoresis (PAGE)

6.1 Non-denaturing PAGE

A slab gel (10 x 10 x 0.1 cm) composed of stacking gel (3 cm) and separating gel (6 cm) was used. Non-denaturing PAGE with 4-10% gradient gel was performed according to the method of Davis (1964). Compositions of the gels are shown as below.

Composition	**Stacking gel** 3% (5 ml)	**Separating gel**	
		4% (3 ml)	10% (3 ml)
30% Acrylamide-0.8% bisacrylamide	0.50 ml	0.40 ml	1.00 ml
0.5 M Tris-HCl, pH 6.8	0.63 ml	-	-
1.5 M Tris-HCl, pH 8.8	-	1.50 ml	1.50 ml
10% Ammonium persulphate	50 µl	30 µl	30 µl
TEMED	5 µl	3 µl	3 µl
Distilled water	3.82 ml	1.07 ml	0.47 ml

The purity of lectin during purification steps was carried out by non-denaturing PAGE using the Davis buffer system (Davis, 1964). Samples were mixed with a sample buffer (0.2 M Tris HCl, pH 6.8, 8 mM EDTA, 40% glycerol and 0.04% bromophenol blue) in a ratio of 3:1. Gel was run with a constant current of 15 mA for 2 h at room temperature. The molecular mass of purified lectin was estimated by measuring its relative mobility in non-denaturing PAGE as compared to those of high molecular weight standards (Amersham Pharmacia Biotech): thyroglobulin (669 kDa), ferritin (440 kDa), catalase (232 kDa), lactate dehydrogenase (140 kDa) and BSA (67 kDa).

6.2 Sodium dodecyl sulphate-PAGE (SDS-PAGE)

SDS-PAGE was prepared and run following the method of Laemmli (1970) using a slab gel (10 x 10 x 0.1 cm), composed of stacking gel (3 cm) and separating gel (6 cm) in the presence or absence of 1% β-mercaptoethanol. Compositions of the gels are shown as below.

Composition	**Stacking gel** 3% (5 ml)	**Separating gel** 12% (6 ml)
30% Acrylamide-0.8% bisacrylamide	0.5 ml	2.4 ml
0.5 M Tris-HCl, pH 6.8	1.25 ml	-
1.5 M Tris-HCl, pH 8.8	-	1.50 ml
0.2 M EDTA	50 µl	60 µl
10% SDS	50 µl	60 µl
10% Ammonium persulphate	50 µl	60 µl
TEMED	5 µl	6 µl
Distilled water	3.10 ml	1.92 ml

Samples were mixed with a sample buffer (0.2 M Tris-HCl, pH 6.8, 4% SDS, 40% glycerol in the presence or absence of 4% β-mercaptoethanol and 0.04% bromophenol blue as the tracking dye) in a ratio of 3:1. They were then boiled for 30 min. Electrophoresis was carried out in the electrode buffer (25 mM Tris-0.192 M glycine, 0.1% SDS, pH 8.3). The gel was run at a constant voltage of 250 V for 2 h. The molecular masses of subunits of purified lectin were estimated by measuring their relative mobility in SDS-PAGE as compared to those of low molecular weight standards (Amersham Pharmacia Biotech): phosphorylase b (94 kDa), BSA (67 kDa), ovalbumin (43 kDa), carbonic anhydrase (30 kDa), trypsin inhibitor (20 kDa) and α-lactalbumin (14 kDa).

6.3 Protein staining

6.3.1 Coomassie brilliant blue R-250 staining

Following electrophoresis, protein bands were stained with Coomassie blue (0.1% Coomassie brilliant blue R-250-50% methanol-7.5% acetic acid) for 2 h. Excess dye was removed and then gel was fixed for 30 min in a fixing solution (50% methanol-7.5% acetic acid). The gel was destained with several changes of destaining solution (5%methanol-7.5% acetic acid) until the background was transparent.

6.3.2 Silver staining

After electrophoresis, proteins were fixed in the gel with 40% methanol-10% acetic acid for 15 min and washed twice with 10% ethanol-5% acetic acid for 15 min with gentle shaking. The gel was transferred to an oxidizing solution, shaken gently for 3 min then washed several times with deionized water until its background became clear. The gel was transferred into a silver staining kit for 20 min and washed with deionized water for 15 sec. It was then immersed in a developing solution while shaking gently and the solution was changed several times until protein bands appeared in the gel. This reaction was then stopped by 3% acetic acid.

7. N-Terminal amino acid sequence analysis

Fifty micrograms of purified lectin were subjected to SDS-PAGE and transferred to a polyvinyl difluoride (PVDF) membrane in 10 mM 3-(cyclohexylamino)-1-propane-sulfonic acid (CAPS) containing 1% methanol at a constant current of 500 mA for 1.5 h. After Coomassie blue staining and extensive rinsing in distilled water, the 30.9 kDa and 32.3 kDa bands were cut out separately and subjected to N-terminal sequencing by Edman Degradation (Edman and Begg, 1967) in a protein/peptide sequencer.

8. Peptide analysis by MALDI-TOF-TOF MS

One hundred micrograms of purified lectin were subjected to SDS-PAGE. The protein bands (30.9 and 32.3 kDa) were excised and subjected to in-gel tryptic digestion. Mass spectra of these two samples were obtained on a matrix assisted laser desorption and ionization / time-of-flight mass spectrometer (MALDI-TOF), TOF/TOF 4700 Proteomics analyzer (Applied Biosystems). MALDI spectra

were calibrated using a peptide mixture provided by the manufacturer (Applied Biosystems). Protein identifications were realized by the method of peptide mass fingerprinting (PMF) and database searches.

9. Extraction of total RNA

To prepare cDNA library, the guanidinium thiocyanate method (method section 9.1) was used to isolate total RNA from hepatopancreas and hemocyte, whereas total RNAs, from muscle, heart, stomach and lymphoid organ, were extracted by using the RNeasy Protect Mini Kit (method section 9.2)

9.1 Total RNA extraction using the guanidinium thiocyanate method

Total RNA from the hepatopancreas or hemocytes was extracted using the acid guanidinium thiocyanate phenol-chloroform extraction method as described by Chomczynski and Sacchi (1987). In brief, samples were homogenized in 500 µl of 4 M guanidinium thiocyanate, 0.025 M sodium citrate (pH 7.0), 0.5% sarcosyl and 0.1 M β-mercaptoethanol using syringe and needle. Homogenate was added subsequently with the following reagents: 50 µl of 2 M sodium acetate (pH 4.0), 500 µl of water-saturated phenol and 100 µl of chloroform-isoamyl alcohol mixture (49:1). The suspension was mixed by inversion after each addition, then shaken vigorously for 10 sec and cooled on ice for 15 min. The homogenate was then centrifuged at 10,000xg for 20 min at 4°C. The aqueous phase was transferred to a new tube. To precipitate RNA, 1 ml of isopropanol was added and mixture was kept at -80°C for at least 1 h. RNA was pelleted by centrifugation at 10,000xg for 20 min. The RNA pellet was then washed once with 70% ethanol. Ultimate RNA pellet was collected, air dried and dissolved in RNase-free water. RNA was then stored at -80°C until used.

9.2 Extraction of total RNA using RNeasy Protect Mini Kit

Total RNA was extracted from muscle, heart, stomach, hepatopancreas, ovary, intestine and lymphoid organ by using RNeasy Protect Mini Kit (Qiagen). In brief, each tissue was disrupted and homogenized in buffer RLT containing guanidine isothiocyanate supplemented with 1% (v/v) β-mercaptoethanol using syringe and needle. Homogenate was then centrifuged at 8,000xg for 3 min. The supernatant was transferred to a new tube, and 70% ethanol was then added to precipitate DNA. The mixture was applied to RNeasy spin column and centrifuged at 8,000xg for 15 sec.

Contamination of DNA and proteins were removed by washing the column with a solution containing guanidine salt and ethanol. RNA was eluted in the final step with RNase-free water. Total RNA concentration was determined by spectrophotometer at $O.D._{260}$ (1 $O.D._{260}$ RNA = 40 µg/ml) and purity of isolated RNA was determined using an $O.D._{260}$/ $O.D._{280}$ ratio of 1.8-2.1. RNA was finally stored at -80°C until used.

9.3 Analysis of RNA by formaldehyde agarose gel electrophoresis

The quality of RNA was analyzed by formaldehyde agarose gel electrophoresis after RNAs were measured for their concentrations by spectrophotometer at wavelength 260 and 280. To prepare 1% gel for RNA analysis, 0.75 g of agarose was mixed with 47 ml of diethylpyrocarbonate (DEPC)-treated water and 15 ml of 5x running buffer (100 mM 3-(N-morpholino)-propanesulfate (MOPS), 40 mM sodium acetate, 5 mM EDTA, pH 7). Gel mixture was heated until the agarose was dissolved and then cooled to 65-70°C. After cooling, 12 ml of 37% (12.3 M) formaldehyde was added and mixed well by swirling in a fume hood. Gel electrophoresis chamber was set up. Agarose solution was poured into a gel tray in a fume hood to get a 3-5 mm thickness and stood for at least 30 min. For sample preparation, about 5 µg of each RNA sample was mixed well with 4 µl of 5x running buffer, 3.5 µl of 37% formaldehyde, 10 µl of formamide to a final volume 20 µl. Each sample was incubated at 65°C for 15 min and immediately chilled to denature RNA. The samples were mixed with 6x loading dye [0.4% orange G, 0.03% bromophenol blue, 0.03% xylene cyanol FF, 15% Ficoll 400, 10 mM Tris-HCl (pH 7.5) and 50 mM EDTA (pH 8.0)] in a ratio of 5:1 and then 10 mg/ml ethidium bromide (EtBr) solution was added to a final concentration of 0.2 µg/ml before loading into gel slots. Then, electrophoresis was performed at 60 volts for 3 h in 1x running buffer. Afterwards, the gel was destained with DEPC-treated water for 30 min with gentle agitation. RNA pattern was visualized under an ultraviolet (UV) lamp.

9.4 Digestion of contaminated DNA with deoxynueleotidase (DNase)

Although the company claims that silica-membrane technology of the RNeasy kit efficiently removed most of contaminated DNA, additional DNA removal is necessary for certain RNA application such as RT-PCR and real-time PCR which are sensitive to very small amounts of DNA. In order to get rid of residual DNA, the digestion with DNase I was performed prior to RT-PCR or reverse transcription using

the basic protocol recommended by the company with some modifications. Reaction mixture (20 µl) was composed of total RNA 2 µg, DNase I buffer (40 mM Tris-HCl, pH 8.4, 4 mM $MgCl_2$, 100 mM KCl), RNaseOut (RNase inhibitor, 40 units) and 2 units of DNase I. Digestion of DNA was allowed to occur at room temperature for 15 min before 1 µl of 25 mM EDTA solution was added and the mixture was incubated at 65°C for 10 min to inactivate the DNase I. The reaction mixture was kept at -20°C or immediately used for cDNA synthesis.

10. Isolation of mRNA from total RNA

After extraction of RNA from the hepatopancreas and hemocytes, mRNA was isolated from total RNA using PolyATtract®mRNA isolation system (Promega) (Fig. 19) as follows: total RNA (about 0.1-1.0 mg) was adjusted to 500 µl with RNase-free water and then placed in 65°C heating block for 10 min. A 3 µl aliquot of Biotinylated-Oligo(dT) probe and 13 µl of 20xSSC were added. The tube was gently shaken and stood at room temperature until cooled down. Streptavidin-paramagnetic particles (SA-PMPs) were resuspended by gently flicking bottom of the tube until completely dispersed. Then it was captured by placing in the magnetic stand until SA-PMPs were collected at the inner wall of the tube. The supernatant was carefully removed. SA-PMPs were repeatedly washed three times with 0.5xSSC 0.3 ml each. Washed SA-PMPs were re-suspended in 0.1 ml of 0.5xSSC. The entire content of the annealing reaction was added to the tube containing the washed SA-PMPs. The tube was incubated at room temperature for 10 min. The SA-PMPs were captured by using the magnetic stand and the supernatant was carefully removed. The particles were washed four times with 0.1xSSC (0.3 ml per wash) by gently flicking the bottom of the tube until all the particles were re-suspended. After a final washing, the supernatant was removed as much as possible without disturbing the SA-PMP particles. To elute mRNA, the final SA-PMP pellet was re-suspended in 0.1 ml of the RNase-free water, the particles were gently re-suspended by flicking the tube. The SA-PMPs were magnetically captured and the eluted mRNA was transferred to a sterile RNase-free tube. The elution step was repeated by re-suspension the SA-PMPs in 0.15 ml of RNase-free water. The capture step was repeated. The eluate was pooled with a previously eluted RNA (0.25 ml total volume).

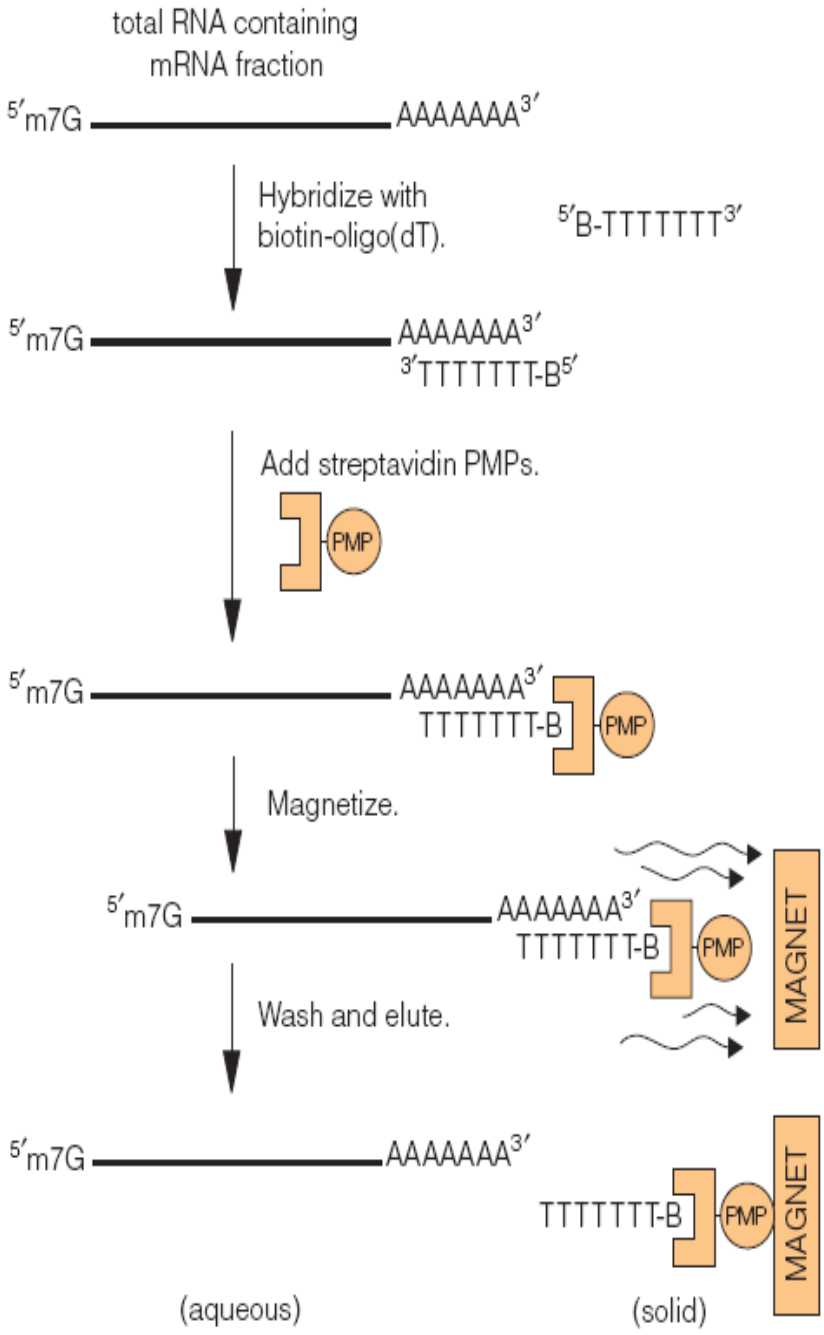


Fig. 19 Schematic diagram of PolyATtract® mRNA isolation procedure.

11. cDNA library construction

Complementary DNA library was prepared from purified mRNA and subjected to a directional cloning into UniZAP®XR Vector (Stratagene, USA) using ZAP-cDNA synthesis kit (Stratagene, USA) with some modification (Fig. 20). The cDNA ligated into UniZAP®XR Vector was packaged using the Gigapack® III Gold Cloning Kit (Stratagene, USA) packaging system.

11.1 First-strand cDNA synthesis

Poly A^+ RNA (5 µg) was primed in the first strand cDNA synthesis with an oligo (dT) linker-primer that contains an *Xho* I restriction site, then transcribed using StrataScript™ reverse transcriptase and 5-methyl dCTP. First-strand cDNA synthesis reaction was performed in a 50 µl mixture containing 5 µl of 10x first-stand buffer [0.5M Tris-HCl (pH 8.3), 0.75 M KCl, 0.1 M dithiothreitol (DDT) and 30 mM $MgCl_2$], 3 µl of first-stand methyl nucleotide mixture (10 mM dATP, dGTP, and dTTP plus 5 mM 5-methyl dCTP), 2 µl of linker-primer (5'-GAG AGAGAGAGAGAGAGAACTAGTCTCGAGTTTTTTTTTTTTTTTT-3') (1.4 µg/µl), 1 µl of RNase Block Ribonuclease Inhibitor (40 U/µl) and 5 µg of mRNA. The reaction was mixed and left for 10 min at room temperature, to allow the primer to anneal to the template. Then 1.5 µl of Strata Script RT (50 U/µl) was added to the reaction mixture and incubated at 42°C for 1h. Afterwards, the first-strand synthesis reaction mixture was placed on ice.

11.2 Second-strand cDNA synthesis

An addition into the ice-cold first-stand synthesis reaction mixture 45 µl, was done as followed; 20 µl of 10x second-strand buffer, 6 µl of second-strand dNTP mixture (10 mM dATP, dGTP and dTTP plus 26 mM dCTP), 116 µl of sterile distilled water, 2 µl of RNase H (1.5U/µl) and 11 µl of DNA polymerase I (9 U/µl). The mixture was gently mixed and incubated at 16°C for 2.5 h. The second-strand synthesis reaction mixture was then placed immediately on ice.

11.3 Blunting of cDNA termini

Blunting dNTP mix 23 µl (2.5 mM dATP, dGTP, dTTP and dCTP) and 2 µl of Cloned *Pfu* DNA polymerase (2.5 U/µl) were added to the second-stand synthesis reaction and incubated at 72°C for 30 min. Afterwards, cDNA was purified to remove primers, nucleotides, polymerase and salts, as well as concentrated using

1:1(v/v) phenol-chloroform which 200 µl of phenol-chloroform was added and mixed by vortex. The tube was spun at maximum speed for 2 min at room temperature. The upper aqueous layer, containing cDNA, was transferred to a new tube. An equal volume of chloroform was added and vortexed. The tube was spun at maximum speed for 2 min at room temperature and the upper aqueous layer was transferred to a new tube. The cDNA was precipitated by adding 20 µl of 3 M sodium acetate (pH 5.2) and 400 µl of 100% (v/v) ethanol to the saved aqueous layer. The reaction mixture was mixed by vortex and kept to precipitate overnight at -20°C. The microcentrifuge tube was spun at maximum speed at 4°C for 60 min. The pellet was gently washed by adding 500 µl of 70% (v/v) ethanol. The tube was re-spun at maximum speed for 2 min at room temperature. Ethanol was aspirated while the pellet was air dried and re-suspended in 9 µl of *Eco*R I adapters (5'-OH-AATTCGGCACGAGG-3'). To allow

3'-GCCGTGCTCCp-5'

the cDNA to re-suspend, the reaction was incubated at 4°C for 30 min. One µl of this second-strand synthesis reaction was transferred to a separate tube as a control reaction.

11.4 Ligation of *Eco*R I adapters

The following components: 1 µl of 10x ligase buffer (500 mM Tris-HCl, pH 7.5, 70 mM $MgCl_2$, 10 mM DTT), 1 µl of 10 mM rATP and 1 µl of T4 DNA ligase (4 U/µl) were added to the tube containing the blunted cDNA and *Eco*R I adapters. The reaction was mixed and incubated at 4°C for 2 days. The ligation reaction was heat-inactivated by placing the tubes in a 70°C water bath for 30 min.

11.5 Phosphorylation at *Eco*R I ends

After heat inactivation, the reaction mixture was spun in a microcentrifuge for 2 sec. The reaction was then cooled at room temperature for 5 min, and the adapter ends were phosphorylated by addition of the following components: 1 µl of 10x ligase buffer, 2 µl of 10 mM rATP, 6 µl of sterile water, and 1 µl of T4 polynucleotide kinase (10 U/µl). The reaction was incubated at 37°C for 30 min. The kinase was heat inactivated at 70°C for 30 min, and then left for 5 min to equilibrate to room temperature.

11.6 Digestion with *Xho* I

The reaction (from method section 11.5) was added with 28 µl of *Xho* I buffer supplement and 3 µl of *Xho* I (40 U/µl) and incubated at 37°C for 1.5 h. After that, 5 µl of 10x STE buffer [200 mM Tris-HCl (pH 7.5), 1 M NaCl, 100 mM EDTA] and 125 µl of 100% ethanol were added. The cDNA was precipitated overnight at -20°C. Following precipitation, the reaction was spun in a microcentrifuge at maximum speed for 60 min at 4°C. The supernatant was discarded. The pellet was completely dried, and resuspend in 14 µl of 1x STE buffer.

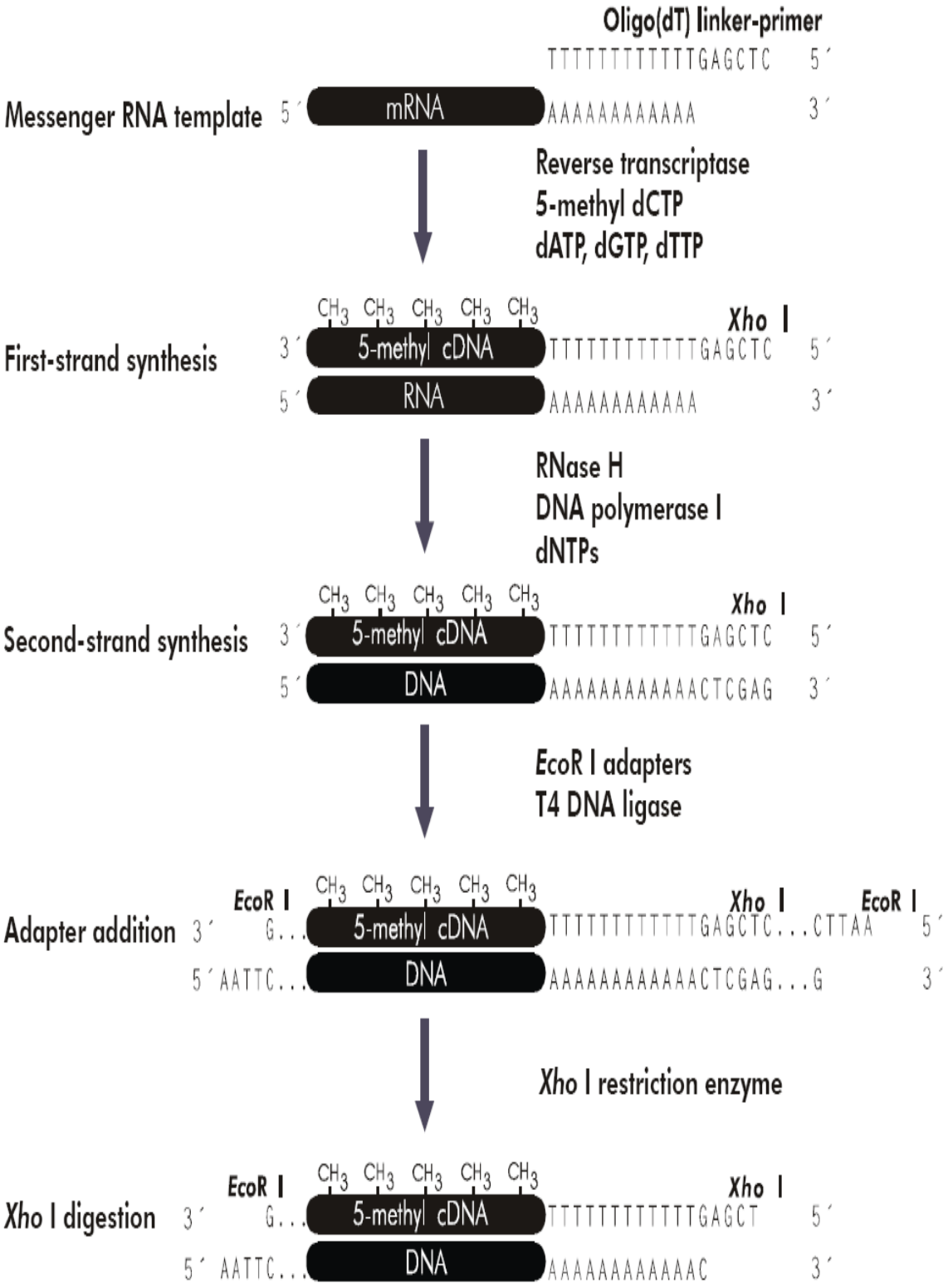


Fig. 20 The cDNA synthesis flow chart.

11.7 Size fractionation

To remove residual oligonucleotides, restriction enzymes, and cDNAs < 400 bp, which might compete or interfere with an insertion or ligation of the cDNA to the vector, SizeSep™ 400 Spin Column (Sepharose CL-4B) was used to selected cDNA >400 bp. The column was pre-equilibrated with 1x STE buffer, and then centrifuged at 400xg for 2 min. The sample was applied slowly to the center at the top of the column, cDNA was collected in a 1.5 ml microcentrifuge tube after centrifugation at 400xg for 2 min. The size fractionated cDNA was precipitated and ligated to Uni®-ZAP XR Vector.

11.8 Ligation of cDNA into Uni®-ZAP XR vector

Ligation reaction was performed in 5 µl mixture containing 2.5 µl resuspened cDNA (100 ng), 0.5 µl of ligase buffer, 0.5 µl of 10 mM rATP (pH 7.5), 1 µl of the predigested Uni®-ZAP XR vector (1 µg/µl) and 0.5 µl of T4 DNA ligase (4U/µl). The reaction was incubated at 4°C for 2 days.

11.9 Packaging

After ligation into Uni-ZAP® XR vector, the resulting DNA was packaged *in vitro* using Gigapack III Gold packaging extract. The packaging extract was removed from -80°C and quickly thawed in hand, then 1-4 µl containing 0.1-1 µg of the ligated DNA was immediately added to the packaging solution. The reaction was incubated at room temperature for 2 h before being added 500 µl of SM buffer [50 mM Tris-HCl (pH 7.5), 100 mM NaCl, 8 mM $MgSO_4$, 0.002% (w/v) gelatin] and 20 µl of chloroform. The reaction mixture was mixed gently and then the tube was briefly centrifuged to remove sediment. The supernatant which contained lambda phage library was transferred to a new tube and kept at 4°C and ready for titering.

11.10 Plating and titering

It is important to determine a phage titer of the library being screened in order to plate the correct number of plaque forming units (pfu) per plate. A culture of XL1-Blue MRF' host strain, in Luria-Bertani (LB) medium, supplemented with 10 mM $MgSO_4$ and 0.2% maltose, was grown to an $O.D._{600}$ of 1.0, then the bacteria was pelleted by centrifugation at 1,000 x g for 10 min. The cells were gently resuspended to half of an original volume with sterile 10 mM $MgSO_4$, and then diluted to a final $O.D._{600}$ of 0.5. LB top agar (0.7% agarose/NZY medium) was melted and cooled to

48°C in water bath. One µl of lambda phage packaging material was added to 300 µl of the diluted host cells. Also, 1µl of 1:10 serial dilution of packaging material in SM buffer was added to 300 µl of host cells. The phages and bacteria were incubated 15 min at 37°C to allow the phages to attach to host cells. After that 3 ml of LB top agar (0.7% agarose/NZY medium which pre-melted and cooled to ~48°C in water bath) were added and immediately poured onto prewarmed NZY agar plates. These plates were allowed to set at room temperature for 10 min. They were then inverted and incubated at 37°C for 6-8 h. Subsequently, plaque formation could be observed as a clear zone of bacterial cells on NZY plates. So that, a titer of primary library was determined by counting the number of plaques and calculated by this equation.

$$\left[\frac{\text{Number of plaques (pfu)} \times \text{dilution factor}}{\text{Volume plated (µl)}}\right] \times 1000\ \mu\text{l / ml}$$

The phage library was amplified by combine aliquots of the library suspension containing ~$5x10^4$ pfu of bacteriophage with 600 µl of XL1-Blue MRF' cells (bacteria cells were prepared in the same manner as library titering) at $O.D._{600}$ of 0.5 in 15 ml tubes, incubated at 37°C for 15 min. Thereafter the reaction tube were plated on NZY plates by mixing with 6.5 ml of NZY top agar that melted and cooled to ~48°C, then were spreaded onto a 150-mm NZY agar plate. These plates were allowed to set for 10 min at room temperature, inverted and incubated at 37°C for 6-8 h. After incubation, the plaques were observed. They were recovered by being added 8-10 ml SM buffer onto NZY plates and incubated at 4°C overnight. The phage particles that diffused to SM buffer were collected and pooled together in a sterile 50 ml tube. The plates were rinsed with an additional 2 ml of SM buffer. The bacteriophage suspension was added with chloroform to a 5% (v/v) final concentration, mixed well and incubated at room temperature for 15 min. The cell debris were removed by centrifugation at 500xg for 10 min. The supernatant was recovered and transferred to a new tube. The amplified cDNA library was added with dimethyl sulfoide (DMSO) to a 7% (v/v) final concentration and aliquoted (1 ml) into 1.5 ml microcentrifuge tubes and stored at -80°C.

12. Molecular cloning of c-type lectin and a lipopolysaccharide and β-1,3-glucan binding protein

F. merguiensis Lectin (*FmL*) and *F. merguiensis* lipopolysaccharide and β-1,3-glucan binding protein (*FmLGBP*) genes were isolated by amplification method using PCR. Complementary DNA fragments at 5'- and 3'- ends were amplified by rapid amplification of cDNA ends (RACE) method.

12.1 Designing primer for gene fragment amplification

For *F.merguiensis* lectin gene (*FmL*), primers were designed from the Blastp and Blastn (NCBI) results using C-type lectin from black tiger shrimp as a query sequence search. The results of Blastp and Blastn analysis showed amino acid and nucheotide sequences form various C-type lectin sources. Primers were designed by multiple alignments of amino acid and nucleotide sequences of C-type lectins form various sources. Conserved regions of amino acid sequences were chosen for primer designing then compared with nucleotide sequences. The conservation between amino acid and nucleotide sequences was used to design degenerate primers.

For *F.merguiensis* lipopolysaccharide and β-1,3-glucan binding protein gene (*FmLGBP*), primers were designed based on the highly conserved nucleotide sequences of LGBP or BGBP of *P. monodon* (AAM21213), *P. leniusculus* (CAB65353), *L. stylirostris* (AAM73871) and *Marsupenaeus japonicus* (BAD36807) in the GenBank database. The primers for both gene are shown in Table 6. All of these degenerate primers were commercially synthesized by Invitrogen.

Table 6 Nucleotide sequences of primers used for amplification of partial *FmL* and *FmLGBP* cDNA fragments from *F. merguiensis*.

Primer names	Primer sequence
FmL-F1	5'-TGGACGGACSMGACGARGMAAGAGA-3'
FmL-R1	5'-CTGTGGTGTAGACCCATGTCCCTTCA-3'
FmLGBP-F1	5'-TCCGGCGGAGGAAACTGGGAATTCCA-3'
FmLGBP-R1	5'-GATCAGGTAGAACTTCTGGTCGGAGG-3'

12.2 First-strand cDNA synthesis by reverse transcription

First-strand cDNA was synthezied using SuperScript™ III Reverse Transcriptase (Invitrogen, USA) according to the manufacturer's instructions. Five micrograms of total RNA (from method section 9.2) were mixed with 1 µl of oligo $(dT)_{12-18}$, 10 mM dNTP mix and DEPC-treated water to a total volume of 13 µl. Reaction was incubated at 65°C for 5 min and chilled on ice for at least 1 min to remove secondary structure of RNA. The mixture was added with 4 µl of 5xFirst-Strand Buffer [250 mM Tris-HCl (pH 8.3 at room temperature), 375 mM KCl, 15 mM $MgC1_2$], 1 µl of 0.1 M DTT, 1 µl of RNaseOUT™ Recombinant RNase Inhibitor (40 units/µl), 1 µl of SuperScript™ III RT (200 units/µl) to give a total volume of 20 µl. It was incubated further at 50°C for 1 h followed by heating at 70°C for 15 min to inactivate the enzyme reverse transcriptase. First-strand cDNA was used as a template for subsequent PCR reactions.

12.3 Amplification of DNA fragments by polymerase chain reaction (PCR)

PCR was carried out with *Taq* DNA polymerase according to the instructions of the manufacturer (Promega, USA). PCR was performed using the cDNA library prepared from hepatopancreas and hemocyte as a template (from method section 11) in a final volume of 50 µl containing 1xPCR buffer (10 mM Tris-HCl, pH 9.2, 50 mM KCl), 3 mM $MgCl_2$, 0.2 mM dNTP mix, 0.2 µM forward primer (FmL F1 for *FmL* or FmLGBP F1 for *FmLGBP*), 0.2 µM reverse primer (FmL-R1 for *FmL* or FmLGBP-R1 for *FmLGBP*), 2 µl of preheated (5 min at 95°C) phage cDNA library template, 5 units of *Taq* DNA polymerase and sterile water to make the volume to 50 µl. The amplification was started with an initial denaturation step at 94°C for 2 min. This was followed by 30 cycles of denaturation at 94°C for 30 sec, annealing at 60°C for *FmL* or 61°C for *FmLGBP* for 45 sec and extension at 72°C for 1 min. The final extension was carried out for 1 cycle of 72°C for 5 min. The PCR products were analyzed on 1.2% agarose gel electrophoresis.

12.4 Nested amplification

Often a single PCR of 25-35 cycles will not generate enough specific products to be detected by ethidium bromide staining. A dilution (1:10) of the primary PCR product (above) was reamplified using the nested primer. The products of the second

PCR were analysed on agarose gel, then cloned into pDrive vector (as described in method section 12.7.1) and sequenced (method section 12.10).

12.5 DNA analysis by agarose gel electrophoresis

Agarose gel electrophoresis was used to visualize the interesting DNA fragments. The concentration of agarose gel depended on the size of DNA to be analyzed. Agarose was melted in TAE buffer (40 mM Tris-acetate buffer, pH 8, 1 mM EDTA), then 10 mg/ml ethidium bromide solution was added to a final concentration of 0.2 µg/ml. The solution was poured into a casting plate with an appropriate comb. After the agarose gel was completely solidified, the comb was removed. The gel was immersed in an electrophoresis chamber containing TAE buffer. Five volumes of the DNA samples were mixed with one volume of 6 x loading dye [0.4% orange G, 0.03% bromophenol blue 0.03% xylene cyanol FF, 15% Ficoll 400, 10 mM Tris-HCl (pH 7.5) and 50 mM EDTA (pH 8)] and then loaded into wells in the gel. Electrophoresis was carried out with constant voltage of 50-100 V. The running time depended on the size of the DNA to be separated. After electrophoresis, DNA patterns were visualized under an ultraviolet (UV) lamp.

12.6 Purification of DNA fragment from agarose using QIAquick Gel Extraction kit (QIAGEN)

For isolation of DNA from agarose gel, the DNA band was cut out from the agarose gel with a clean and sharp scalpel and then transferred to a 1.5 ml microcentrifuge tube. Three volumes of buffer QG (guanidine thiocyanate, pH 7.5) were added to 1 volume of the gel (e.g. 100 mg of gel was mixed with 300 µl of buffer QG). The mixture was incubated at 50°C for 10 min and mixed by vortex every 3 min during incubation. If the gel slice contained <500 bp or > 4 kb DNA fragment, 1 gel volume of isopropanol was added to the mixture before being transferred to a QIAquick column that inserted in a 2 ml collection tube. The column was centrifuged at 8,000xg for 1 min a room temperature. The flow-through solution was discarded and the column was washed with 0.75 ml of buffer PE (2 mM Tris-HCl, pH 7.5-20 mM NaCl in 80% ethanol), followed by centrifugation at 8,000xg for 1 min. The flow-through solution was discarded and the column was centrifuged for an additional 1 min. In order to collect the eluted DNA, the column was transferred and placed on a new 1.5 ml microcentrifuge tube, then 30 µl of buffer EB (10 mM Tris-HCl, pH 8.5)

was applied to the column. Afterwards, the column was left at room temperature for 1 min before centrifugation at 8,000 x g for 1 min. The eluted DNA was stored at 4°C until used.

12.7 Cloning of PCR products into T-overhang vectors

PCR products (as described in method section 12.6) were subcloned through the TA-ligation (TA: thymidine-adenine), if PCR was performed with a *Taq* polymerase that leaves a deoxyadenosine at the 3´end of the amplified fragment in the elongation phase of the PCR cycle. The PCR fragments were TA-ligated to the following vectors. These vectors contain an open insertion site inside the multiple cloning sites (MCS) that has a 3´overhanged thymidine where the PCR fragment ligates. The ligation reaction was done as described in method section 12.7.1-12.7.3.

12.7.1 Ligation of purified PCR products into pDrive vector

The ligation reaction was performed by mixing the following components: 4 µl of eluted DNA, 1 µl of pDrive cloning vector (Qiagen, Germany) and 5 µl of 2x Ligation master mix. The reaction was briefly mixed and incubated at 4°C for 30 min. Then the ligation mixture was used to transform *E.coli* competent cells.

12.7.2 Ligation of purified PCR products into pCR4-TOPO vector

Ligation of the purified cDNA into pCR4-TOPO vector (Invitrogen, USA) was carried out in a total volume of 6 µl containing 1 µl of pCR4-TOPO vector (10 ng/µl), 4 µl of purified cDNA and 1 µl of salt solution (200 mM NaCl and 10 mM $MgCl_2$). The reaction was incubated at room temperature for 5 min, and then transformed into *E.coli* competent cells.

12.7.3 Ligation of purified PCR products into pGEM-T Easy vector

Ligation of the purified cDNA into pGEM-T Easy vector (Promega, USA) was carried out in a total volume of 10 µl containing 1 µl of pGEM-T Easy vector (50 ng/µl), 3 µl of purified cDNA, 1 µl of T4 DNA ligase, 5 µl of 2xRapid ligation buffer [60 mM Tris-HCl (pH 7.8), 20 mM $MgCl_2$, 20 mM DTT, 2 mM ATP and 10% polyethylene glycol, MW 8000, and T4 DNA ligase]. The reaction was incubated at 16°C for 1 h or 4°C overnight, and then transformed into *E. coli* competent cells.

12.8 Plasmid DNA transformation into *E. coli* by $CaCl_2$ method

12.8.1 Preparation of *E. coli* competent cells

By $CaCl_2$ method (Sambrook *et al*., 1989), the competence of *E.* coli cells to accept free DNA is artificially increased. This is achieved by modifying the cell wall with $CaCl_2$. *E. coli* competent cells were prepared either from strain EZ (Qiagen, Germany), strain DH5α (Invitrogen, USA), TOP10 (Invitrogen, USA), or strain BL21(DE3) (Novagen, USA) using $CaCl_2$ method. *E. coli* cells were taken from a frozen stock at -80°C and incubated overnight at 37°C on an antibiotic-free LB agar plate. An isolated single colony of *E.* coli was inoculated into 5 ml of LB medium containing 100 µg/ml amplicilin and grown overnight at 37°C with vigorous shaking at 250 rpm. The culture was diluted (1: 100 dilutions) and incubated at 37°C at 250 rpm for 2-3 h. When the optical density ($O.D._{600}$) of the bacteria reached 0.4 unit, the culture was placed on ice for 10 min and then the cells were collected by centrifugation at 3,000 x g at 4°C for 10 min. The cells were gently resuspended in 10 ml of iced-cold 0.1 M $CaCl_2$, incubated further of ice for 30 min, and then centrifuged at 3,000 x g at 4°C for 10 min. After that the cells were gently resuspended in 2.5 ml of ice-cold 0.1 M $CaCl_2$ containing 15% (v/v) glycerol, aliquoted (100 µl) into 1.5 ml microcentrifuge tubes and stored at -80°C.

12.8.2 Transformation of plasmid into *E. coli* competent cells

One hundred microliters of competent cells were thawed on ice (10 min), mixed with 2 µl of ligation reaction mixture and placed on ice for 30 min with occasional mixing. The transformation reaction mixture was then incubated at 42°C for 90 sec for a heat-shock and placed on ice for 2 more min. The resulting cells were mixed with 250 µl of ambient SOC medium and incubated at 37°C for 1 h with shaking. The transformed culture was spreaded on LB agar plates containing 100 µg/ml ampicillin, 50 µM isopropylthiogalactoside (IPTG) and 80 µg/ml 5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside (X-gal), and incubated at 37°C for 16 h. Among the recombinant clones identified by blue-white color screening, white colonies generally contained inserted DNA while the blue colonies had no insertion (self-ligated plasmid).

12.8.3 Analysis of possitive clones from transformation

The transformants were tested for the presence of desired DNA by colony PCR. For the colony PCR screening of transformants, five to ten colonies were picked up and resuspened individually in 5 µl of sterile water. The cell suspension (5 µl) was heated at 95°C for 5 min to lyze cells and inactivate nucleases. PCR cocktail was composed of 2.5 µl of 10xPCR buffer, 2.5 µl of 25 mM $MgCl_2$, 0.5 µl of 10 mM dNTPs, 100 ng of each primer, 0.1 µl of *Taq* DNA polymerase (0.5 unit), and sterile distilled water to adjust a total volumn of the PCR mixture to 20 µl. The PCR cocktail was added to each PCR tube. The target DNA was amplified by using the colony PCR temperature program of 2 minutes at 95°C, followed by 30 cycles of 30 seconds at 95°C, 45 seconds at 55°C and 1 min at 72°C, then final 5 min at 72°C. Ten microliters of PCR products were analyzed on 1.2% agarose gel electrophoresis and positive clones were shown by the presence of DNA bands of appropriate size. The colonies with positive clones were cultured overnight in LB medium and plasmid DNA was purified using the QIAprep Spin Miniprep kit (as described in method section 12.9.1) and subsequently sequenced (as described in method section 12.10) to confirm the presence and orientation of the target insert.

12.9 Plasmid DNA extraction

12.9.1 Plasmid DNA extraction using the QIAprep Spin Miniprep kit

E. coli plasmids were extracted using the QIAprep Spin Miniprep kit (Qiagen). An overnight culture at 37°C of the *E. coli* was prepared in the presence of antibiotic (100 µg/ml culture of ampicillin). For harvesting, the cells were centrifuged at 3,000 rpm for 5 min. The sediment was resuspended in 250 µl of buffer P1 (50 mM Tris-HCl, pH 8.0-10 mM EDTA) containing 100 µg/ml RNase A. After addition of 250 µl of buffer P2 (NaOH-SDS), the tubes were gently inverted 4-6 times. To stop lysis reaction, 350 µl of neutralized buffer N3 (acidic K-AC) were added to each tube after which they were inverted 4-6 times. In the meanwhile, each spin column was placed in a 2 ml collection tube. The supernatant was carefully applied onto the QIAprep column followed by a 45 sec centrifugation step. The flow-through fraction was discarded and 750 µl of washing buffer PE (20 mM NaCl and 2 mM Tris-HCl, pH 7.5 in 80% ethanol) were added into the column followed by a 45 sec

centrifugation. After discarding the flow-through fraction, a 60 sec centrifugation step completed the washing procedure. A clean 1.5 ml microcentrifuge tube was placed under the column and the plasmid was eluted with buffer EB (10 mM Tris-HCl, pH 8.5). The concentration of DNA was determined by spectrophotometer at $O.D._{260}$ (1 $O.D._{260}$ DNA = 50 µg/ml) and the purity of isolated DNA was judged by $O.D._{260}$ /$O.D._{280}$ (pure DNA solution has an $O.D._{260}/O.D._{260}$ ratio of 1.9-2.0). The plasmid DNA was stored at -20°C until used.

12.9.2 Plasmid DNA extraction by alkaline lysis with SDS

E. coli plasmids were extracted by alkaline lysis with SDS according to the method of Maniatis *et al*. (1989). An overnight bacterial culture was centrifuged at 12,000xg for 1 min and the cell pellet was completely resuspended in 100 µl of ice-cold solution I (50 mM glucose, 10 mM EDTA, 25 mM Tris-HCl, pH 8.0) with vigorous mixing by vortex. Cells were lysed by an addition of 200 µl of a freshly prepared solution II (0.2 N NaOH and 1% SDS). The mixtures were mixed by inverting the tube rapidly five times, and incubated at room temperature for 5 min. Samples were neutralized by addition of 150 µl of ice-cold solution III (potassium acetate solution, pH 4.8), mixed by inversion and incubated on ice for 10 min. The cell lysate was precipitated by centrifugation at 15,300 x g for 5 min and the supernatant was transferred into a new tube. Isopropanol (540 µl, 0.6 volume) was added to the supernatant, and the mixture was incubated at room temperature for 2 min. The nucleic acid was pelleted by centrifugation at 12,000xg for 5 min and the supernatant was gently removed by decanting. The DNA pellet was washed twice with 1 ml of 70% ethanol and the tube was inverted several times. To precipitate DNA, the mixture was centrifuged at 17,950 x g for 5 min and the supernatant was removed by gentle aspiration. The DNA pellet was briefly air-dried and dissolved in 50 µl of sterile water. Then the DNA solution was treated by RNase A (BioLab) which 1 µl of 10 mg/ml RNase A was added and incubated at 37°C for 1 h..

12.9.3 Determination of recombinant clones by enzymatic digestion

DNA restriction was used to verify a recombinant clone, the restriction endonuclease *Eco*R I binds and cleaves double stranded DNA at the specific sequence 5'-G↓AATT C-3' with the other strand of the cleavage site being 3'-C TTAA↑G-5'. The digestion reaction was composed of 1µg of palsmid DNA, 5 unit of *Eco*R I,

reaction buffer (50 mM Tris-HCl, pH 8.0, 10 mM $MgCl_2$, 100 mM NaCl) and sterile water in a total volume of 10 µl, and was incubated at 37°C for 1 h. After the digestions, the digested products were analyzed by agarose gel electrophoresis.

12.10 Automated DNA sequencing

All inserted plasmid fragments were sequenced on ABI Prism 3100 Genetic DNA Sequencer (Applied Biosystems, Perkin-Elmer) by the Bioservice Unit (BSU), National Science and Technology Development Agency (NSTDA) or they were sequenced on ABI Prism 377 Automated DNA Sequencer (Applied Biosystems) by Scientific Equipment Center, Prince of Songkla University (PSU) based on the dideoxynucleotide chain themination method of Sanger *et al.* (1977). The plasmid sequencing primers (T7, SP6 and T3), internal primers and a Dye Terminator Cycle Sequencing Kit (Applied Biosystems) were used to sequence both strands of the insert in the plasmid.

12.11 Amplification of 5' and 3'ends of cDNAs

Rapid amplification cDNA ends (RACE) is a procedure for amplifying cDNA from a mRNA template between a defined internal site and unknown sequences at either 3' or 5'end of the mRNA. Rapid amplification of the 5' and 3' cDNA ends (RACE) was used to obtain a full-length cDNA (Fig. 21). This experiment was carried out using the GeneRacer^{TM} Kit (Invitrogen, USA).

12.11.1 Construction of first-strand cDNA

12.11.1.1 Dephosphorylation of uncapped RNA

Total RNA (5 µg, from method section 9.2) was treated with calf intestinal phosphatase (CIP) to remove the 5' terminal phosphates of any RNA (i.e. mRNA, truncated mRNA and non-mRNA). This eliminates truncated mRNA and non-mRNA from subsequent ligation with GeneRacer^{TM} RNA Oligo. Dephosphorylation reaction was carried out in a 10 µl reaction. The reaction contained 1 µl of 10x CIP buffer (0.5 M Tris-HCl, pH 8.5, 1 mM EDTA), 1 µl of RNaseOut (40 units/µl), 1 µl CIP [10 units/µl in 25 mM Tris-HCl (pH 7.6), 1 mM $MgCl_2$, 0.1 mM $ZnCl_2$, and 50% (w/v) glycerol], and DEPC-treated water to get a total volume of 10 µl. The reaction was mixed by pipetting, briefly centrifuged and then incubated at 50 °C for 1 h. After that the reaction tube was removed, briefly centrifuged and placed on ice.

12.11.1.2 RNA precipitation

To precipitate the dephosphorylated RNA, 90 µl of DEPC-treated water and 100 µl of phenol : chloroform : isoamyl alcohol (25:24:1) were added to the dephosphorylation reaction tube. Afterwards, the solution was mixed by vigorously vortex for 30 sec and centrifuged at 16000xg for 5 min at room temperature. The upper aqueous phase (100 µl) was transferred into a new 1.5 ml tube, mixed with 2 µl of 10 mg/ml mussel glycogen, 10 µl of 3 M sodium acetate, pH 5.2 and 220 µl of 96% ethanol, and then freezed on dry ice for 10 min. The RNA was pelleted by centrifugation at 16000 x g for 20 min at 4°C, washed with 500 µl of 70% ethanol by inverting the tube several times and subjected to centrifugation at 16000 x g for 2 min at 4°C. The RNA pellet was air-dried for 2 min at room temperature and then resuspended in 7 µl of DEPC-treated water.

12.11.1.3 Decapping of RNA molecules

The cap structure of the dephosphorylated RNA was removed by tobacco acid pyrophosphatase (TAP) in order to remove the 5' cap structure from intact full-length mRNA. This treatment leaves 5' phosphate required for ligation GeneRacer™ RNA Oligo. Decapping reaction was performed in 10 µl of reaction volume containing 7 µl of dephosphorylated RNA, 1 µl of 10x TAP buffer (0.5 M sodium acetate, pH 6.0, 10 mM EDTA, 1% β-mercaptoethanol and 0.1% Triton X-100), 1 µl of RNaseOut ™ (40 units/µl in 20 mM Tris-HCl, pH 8, 50 mM KCl, 0.5 mM EDTA, 8 mM DTT and 50% (v/v) glycerol), and 1 µl of TAP (0.5 units/µl in 10 mM Tris-HCl, pH 7.5, 0.1 M NaCl, 0.1 mM EDTA, 1 mM DTT, 0.01% Triton X-100, and 50% (w/v) glycerol). The reaction was mixed by pipeting and then incubated at 37°C for 1 h that allowed the enzyme to become activated. The decapped RNA reaction was subjected to RNA precipitation following the method section 12.11.1.2. The RNA pellet was resuspended in 7 µl of DEPC-treated water.

12.11.1.4 Ligation of RNA oligo decapped mRNA

In this step, the GeneRacer™ RNA oligo was ligated to 5' end of decapped mRNA. This oligo contains the priming site for GeneRacer™ 5' Primer and GeneRacer™ 5'Nested Primer that allowed the amplification of 5' end of cDNAs with gene specific primer (GSP). The ligation reaction was performed in a 10 µl reaction volume. Seven µl of dephosphorylated and decapped RNA was added to 0.25

ng of lyophilized GeneRacer™ RNA Oligo, pipette up and down to mix and resuspend RNA Oligo and then briefly centrifuged to collect the fluid. The reaction was incubated for 5 min at 65°C to destroy RNA secondary structure, then placed on ice for 2 min and subsequently centrifuged briefly. Total volume of the reaction should decrease by 1 µl due to the incubation. Afterwards, the relax RNA was added with 1 µl of 10xT4 RNA Ligase Buffer (330 mM Tris-acetate, pH 7.8 at 25°C, 660 mM potassium acetate, 100 mM magnesium acetate, 5 mM DTT), 1 µl of 10 mM ATP, 1 µl of RNaseOut ™, and 1 µl of T4 RNA ligase (5 units/µl in 50 mM Tris-HCl, pH 7.5, 0.1 M NaCl, 0.1 mM EDTA, 1 mM DTT, 0.1% Triton® X-100, 50% (w/v) glycerol), and then incubated at 37°C for 1 h. After that the solution was collected by briefly centrifugation before being placed on ice. The dephosphorylated and decapped RNA was subjected to RNA precipitation (method section 12.11.1.2). The RNA pellet was resuspended in 10 µl of DEPC-treated water.

12.11.1.5 Reverse transcription of mRNA

In this experiment, to obtain the 3' and 5' cDNA ends, the ligated mRNA was reverse transcribed by SuperScript™ III RT (Invitrogen, USA) and GeneRacer RNA oligo dT (5'-GCTGTCAACGATACGCTACGTAACGGCATGAC AGTG$(T)_{24}$-3' 60 bases) as a primer. The reverse transcription (RT) was carried out in a 20 µl reaction mixture containing 10 µl of the ligated mRNA, 1 µl of dNTP mix (10 mM each in 1 mM Tris-HCl, pH 7.5), 1 µl of GeneRacer™ Oligo dT primer and 1 µl of DEPC-treated water. The reaction was incubated at 65°C for 5 min and chilled on ice for 2 min to remove the secondary structure of RNA. Then the mixture was mixed with 4 µl of 5x RT buffer (250 mM Tris-HCl, pH 8.3, 375 mM KCl, and 15 mM $MgCl_2$), 1 µl of 0.1 M DTT, 1 µl of SuperScript™ III RT (200 units/µl in 20 mM Tris-HCl, pH 7.5, 100 mM NaCl, 0.1 mM EDTA, 1 mM DTT, 0.01% (v/v) Nonidet P-40, and 50% (w/v) glycerol) and 1 µl of RNase™Out (40 units/µl) to a total volume of 20 µl. The mixture was incubated at 50°C for 1 h, followed by heating at 70°C for 15 min to inactivate the enzyme SuperScript™ III RT. Then 1 µl of RNase H (2 units/µl in 20 mM Tris-HCl (pH 7.5), 100 mM KCl, 10 mM $MgCl_2$, 0.1 mM EDTA, 0.1 mM DTT, 50 µg/ml BSA and 50% glycerol) was added to the mixture and incubated further at 37°C for 20 min to digest the RNA templates. The first-strand cDNA was used as a template for subsequent PCR reactions.

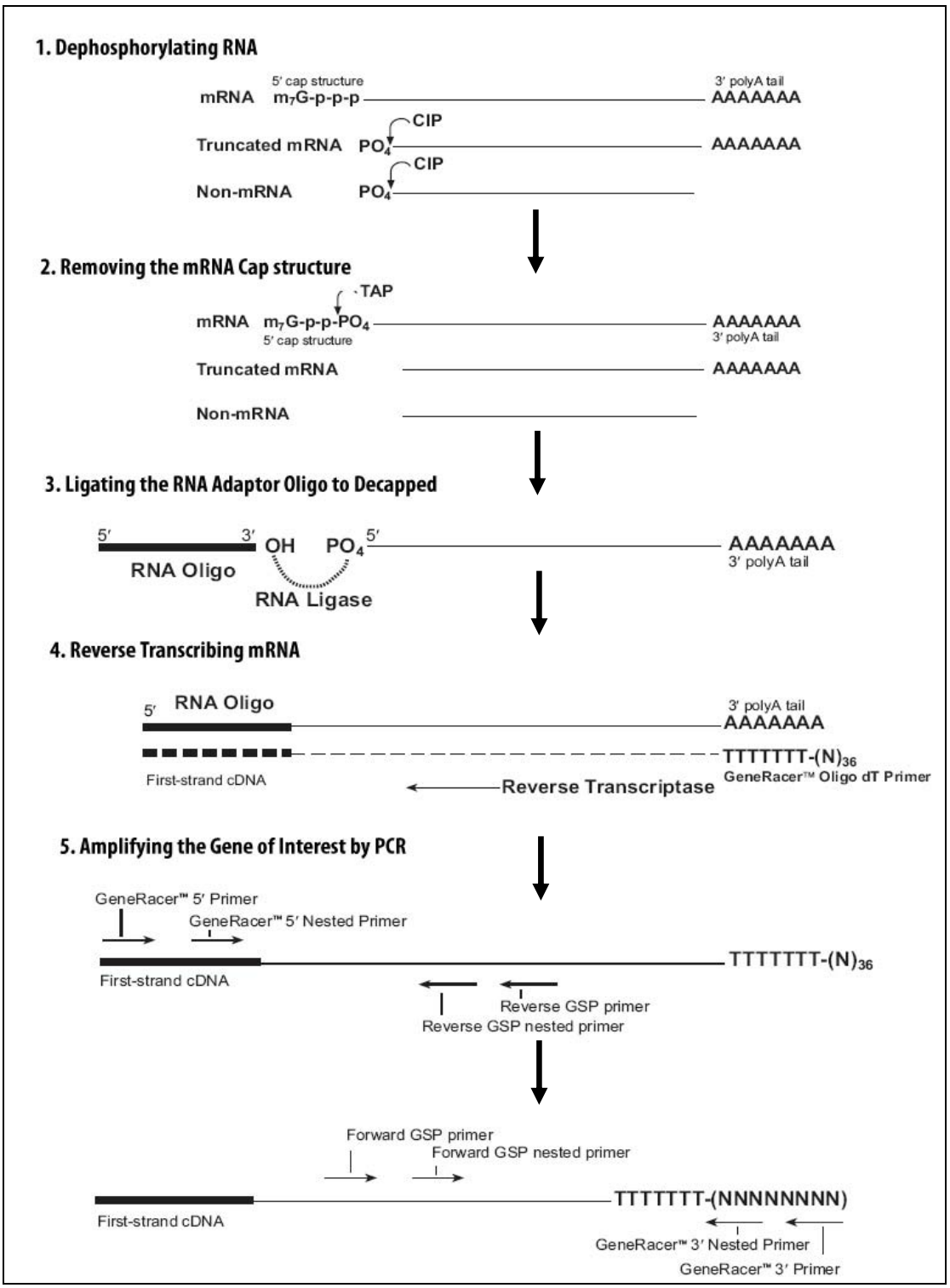


Fig. 21 Diagrammatic representation of first-strand cDNA synthesis for amplification of 5' and 3' ends of cDNA.

CIP, calf intestinal phosphatase; TAP, tobacco acid pyrophosphatase.

12.11.2 Amplification of 5' and 3' ends.

Touchdown PCR was used to amplify 5' and 3' ends of the interesting cDNAs. The GeneRacer™ kit contains four primers (Table 7) designed to amplify 5' and 3' ends of full-length cDNA. GeneRacer™ 5' primer and GeneRacer ™ 5' nested primer for 5' RACE reaction are homologous to sequence in the GeneRacer™ RNA Oligo, whereas GeneRacer ™ 3' primer and GeneRacer ™ 3' nested primer for 3' RACE reaction are homologous to the complementary strand of GeneRacer™ Oligo dT primer. All of primer sequences that used to generate 5' and 3' ends are shown in the Table 7. A PCR reaction was carried out to amplify either both 5' and 3' ends of *FmL* or *FmLGBP* cDNA (Table 8). One microliter of the RT reaction (from method section 12.11.1.5) was used as the template in each PCR reaction. Cycling parameters of touchdown PCR were identical in the amplification of both 5' RACE and 3' RACE and show in Table 9.

Table 7 Primers used for 5' and 3' RACEs of *FmL* and *FmLGBP* cDNAs.

Primer name	Primer sequence
GeneRacer 5' primer	5'-CGACTGGAGCACGAGGACACTGA-3'
GeneRacer 5' nested primer	5'- GGACACTGACATGGACTGAAGGAGTA -3'
GSP 5' FmL primer	5'-CTGTGGTGTAGACCCATGTCCCTTCA-3'
GSP 5' FmLGBP primer	5'-GATCAGGTAGAACTTCTGGTCGGAGG-3'
GeneRacer 3' primer	5'-GCTGTCAACGATACGCTACGTAACG-3'
GeneRacer 3' nested primer	5'-CGCTACGTAACGGCATGACAGTG-3'
GSP 3' FmL primer	5'-TGGACGGACGAGACGAGGAAACAGA-3'
GSP 3' FmLGBP primer	5'- TCCGGCGGAGGAAACTGGGAATTCCA-3'

Table 8 Reagents for primary PCR reaction.

Reagent	5' RACE (µl)	3' RACE (µl)
GeneRacer™ 5' primer (10 µM)	3.0	-
Reverse Gene specific primer (10 µM)	1.0	
GeneRacer™ 3' primer (10 µM)	-	3.0
Forward Gene specific primer (10 µM)	-	1.0
RT template	1.0	1.0
10X High fidelity PCR buffer	5.0	5.0
50 mM $MgSO_4$	2.0	2.0
10 mM dNTP mix	1.0	1.0
Platinum® *Taq* DNA polymerase high fidelity (5U/µl)	0.5	0.5
Sterile water	36.5	36.5
Final volume	50.0	50.0

Table 9 Thermal cycling sequence of a primary PCR.

Temperature	Time	Cycles
94°C	2 min	1
94°C	30 sec	5
72°C	2 min	
94°C	30 sec	5
70°C	2 min	
94°C	30 sec	25
68°C	30 sec	
72°C	1.5 min	
72°C	10 min	1

Nested PCR was used to increase the specificity and sensitivity of the RACE products of 5'- and 3'-ends of *FmL* and *FmLGBP*. The PCR products from the original amplification reaction from the first round PCR were used as the templates in the nested PCR reaction. The PCR reaction mixture composed of reagents shown in Table 10. The cycling program is shown in Table 9. 5'- and 3'- RACE fragments were purified from agarose gel using QIAquick Gel Extraction kit (Qiagen, as described in method section 12.6), cloned into pCR4-TOPO vector (Invitrogen, as described in method section 12.7.2) and transformed into *E coli* TOP10 (as described in method section 12.8.1-12.8.3) and sequenced (as described in method section 12.10) to obtain the 5'- and 3'-end fragments. The identity of the clones was evaluated by matching the sequences to the nucleotide/protein sequences available in GenBank (http://www.ncbi.nlm.nih.gov/, accessed 09/04/07).

Table 10 Reagents for nested PCR reaction.

Reagent	**5' RACE (μl)**	**3' RACE (μl)**
GeneRacer™ 5' Nested Primer (10 μM)	1.0	-
Reverse Gene specific primer (10 μM)	1.0	
GeneRacer™ 3' Nested Primer (10 μM)	-	1.0
Forward Gene specific primer (10 μM)	-	1.0
Initial PCR products (1:10)	1.0	1.0
10x High fidelity PCR buffer	5.0	5.0
50 mM $MgSO_4$	2.0	2.0
10 mM dNTP mix	1.0	1.0
Platinum® *Taq* DNA polymerase high fidelity (5U/μl)	0.5	0.5
Sterile water	38.5	38.5
Final volume	50.0	50.0

12.12 Amplification of a full-length cDNA

Based on the sequence information obtained from the 5' and 3' RACE reaction, gene-specific primers of *FmL* or *FmLGBP* were designed. Each gene was amplified using proofreading *Taq* DNA polymerase. PCR reaction was performed in 50 µl reaction containing 1 µl of first-strand cDNA, 5 µl of 10x Advantage 2 PCR buffer (400 mM Tricine-KOH, pH 8.7 at 25°C), 1 µl of 10 mM of dNTPs, 1 µl of 10 µM forward primer (FmL 1042 forward primer for *FmL* or FmLGBP 1101 forward primer for *FmLGBP*) (Table 11), 1 µl of 10 µM reverse primer (FmL1042 reverse primer for *FmL* or FmLGBP 1101 reverse primer for *FmLGBP*), 1 µl of 50x Advantage 2 polymerase mix and sterile distilled water to give a final volume of 50 µl. The PCR reaction was mixed by vortex (without introducing bubbles). The amplification was started with an initial denaturation step at 95°C for 2 min. This was followed by 30 cycles of denaturation at 95°C for 30 s, annealing at 58ºC (for *FmL*) or 62ºC (for *FmLGBP*) for 30 sec and extension at 72°C for 2 min. The final extension was carried out for 1 cycle of 72°C for 5 min. DNA fragments amplified by PCR were subjected to electrophoresis, purified and subcloned into pDrive vector (as described in method section 12.7.1), The sequences of DNA fragments were determined by ABI Prism 3100 Genetic DNA Sequencer (Applied Biosystems, Perkin-Elmer) by the Bio Service Unit, NSTDA.

Table 11 Nucleotide sequences of gene-specific primers for amplification of the full-length genes.

Primer name	Primer sequence
Primers for *FmL*	
FmL 1042 forward primer	5'- CAGTGCTTTACGTGTTGATGG-3'
FmL 1042 reverse primer	5'- GCTCATCCGCTGTAGACACA-3'
Primers for *FmLGBP*	
FmLGBP 1116 forward primer	5'-ATGAAGGGCTTCATAGCGTCGGTCG -3'
FmLGBP 1116 reverse primer	5'-CTACTGCTCGGCGCTCTCCATCTTC-3'

12.13 Tissue specific expression of *FmL* and *FmLGBP* mRNA

Total RNA was extracted from hemocyte, heart, lymphoid organ, hapatopancreas, stomach, intestine, and muscle (from method section 9.2) and pre-treated with DNAse I for each 2 μg of total RNA as previously described in method sections 9.3. First-strand cDNA synthesis was then performed by using the oligo $(dT)_{12-18}$ primer (from method section 12.2) The primer set for the amplification of the target genes were same to those for cloning of internal fragments. PCR was carried out with *Taq* DNA polymerase according to the instructions of the manufacturer (Promega, USA). PCR was performed using the first-strand cDNA reaction as a template (as described in method section 12.2) in a final volume of 50 μl containing 1x PCR buffer (10 mM Tris-HCl, pH 9.2, 50 mM KCl), 3 mM $MgCl_2$, 0.2 mM dNTP mix, 0.2 μM forward primer (GSP 3' FmL primer for *FmL* or GSP 3' FmLGBP primer for *FmLGBP*, Table 7), 0.2 μM reverse primer (GSP 5' FmL primer for *FmL* or GSP 5' FmLGBP primer for *FmLGBP*), 5% first-strand cDNA, 5 units of *Taq* DNA polymerase and sterile water to make the volume to 50 μl. The amplification was started with an initial denaturation step at 94°C for 2 min. This was followed by 30 cycles of denaturation at 94°C for 30 sec, annealing at 60°C for *FmL* or 61°C for *FmLGBP* for 45 sec and extension at 72°C for 1 min. The final extension was carried out for 1 cycle of 72°C for 5 min. The DNA was analysed on 1.2% agarose gel electrophoresis.

13. Recombinant protein expression in *E. coli*

Expression vector system for *FmL* was constructed using pET-32(a+) vector (Novagen), which was designed to express the gene product as a thioredoxin- and His-tagged fusion protein, and suitable for production of soluble protein in *E. coli* cytoplasm and rapid purification.

13.1 Construction of *Eco*R I and *Xho* I cDNA fragment to insert in expression vectors

The full-length cDNA encoding FmL protein was cloned into pET32a (+) vector and expressed in *E. coli*. The restriction enzymes that should not cut inside the *FmL* but cut at the expression vector cloning site were selected. The specific primers for *FmL* (Table 12) was designed to amplify the cDNA encoding a mature

protein and introduced the restriction enzyme sites complementary to the cloning site of the expression vector to the flank cDNA, allowing inframe translation of a protein. The *FmL* was reamplified using the plasmid containg full-length *FmL* (from method section 12.12) as template with FmL Ex forward and FmL Ex reverse primer, which contain *Eco*R I and *Xho* I sites at the 5' and 3' ends, respectively. PCR reaction was set up in a final volume of 50 µl containing 5 µl of 10x Advantage^{TM} 2 PCR buffer (400 mM Tricine-KOH (pH 8.7 at 25°C), 150 mM KOAc, 35 mM $Mg(OAc)_2$, 37.5 µg/ml BSA, 0.05% Tween 20, 0.05% Nonidet-P40), 1 µl of dNTPs (10 mM each), 1 µl of 10 µM FmL Ex forward primer, 1 µl of 10 µM FmL Ex reverse primer, 1 µl of 1:100 plasmid containing full-length *FmL* (from method section 12.12), and 1 µl of 50x Advantage^{TM} 2 polymerase mix. The sterile water was added to give a final volume of 50 µl. The PCR product was cloned into pDrive vector (as described in method section 12.7.1) and transformed into *E.* coli DH5α (as described in method 12.8.1-12.8.3). The resultant recombinant plasmids were purified and digested with *Eco*R I and *Xho* I. The digestion reaction was performed in a 20 µl volume containing of 0.2 µl of 100 µg/ml BSA, 2 µl of NE Buffer 2 (10 mM Tris-HCl, 50 mM NaCl, 10 mM $MgCl_2$, 1 mM dithiothreitol, pH 7.9 at 25°C), 4 µl of plasmid, 2 µl of *Eco*R I, 2 µl of *Xho* I and 4.9 µl of sterile water to a final volume of 20 µl. The digestion reaction was incubated at 37°C for 2 h. After that the insert fragment was purified using QIAquick Gel Extraction kit (as described in method section 12.6).

Table 12 Primers used for amplification of *FmL* cDNA.

Primer name	Primer sequence
FmL Ex forward	5'-GCGAATTCATGTTGCTCTATCTACTTC-3'
FmL Ex reverse	5'-AACTCGAGTCATCCGCTGTAGACACA-3'

13.2 Construction of expression vector

13.2.1 Preparation of pET32(a+) vector (*Eco*R I and *Xho* I cut)

pET32(a+) vector (2 µl) was mixed with 0.1 µl of 100 µg/ml BSA, 1 µl of NE Buffer 2 (10 mM Tris-HCl, 50 mM NaCl, 10 mM $MgCl_2$, 1 mM dithiothreitol, pH 7.9 at 25°C), 1 µl of *Eco*R I, 1 µl of *Xho* I and 4.9 µl of sterile water to a final volume of 10 µl. The digestion reaction was incubated at 37°C for 2 h. Afterwards, the linear plasmid was purified using QIAquick Gel Extraction kit (QIAGEN).

13.2.2 Construction of recombinant plasmid

The gel-purified insert for *FmL* was cloned into the *Eco*R I and *Xho* I sites of pET32a(+) as described in method section 13.2.1. The ligation mixture was transformed into *E.* coli DH5α and spreaded on LB agar plates containing 100 µg/ml ampicillin as described in method section 12.8.2. Transformation of *E.* coli DH5α with the ligation reaction mix of recombinant plasmid helped to multiply the numper of recombinant plasmids, and these plasmids can be kept permanently (1-4 years) in these host cells.

13.2.3 Identification and isolation of recombinant expression plasmid

The recombinant pET plasmids were identified by colony PCR method (as described in method section 12.8.3). The PCR product was analyzed on 1.2% agarose gel electrophoresis to identify the clones that contained the recombinant plasmids. Then the recombinant colonies were grown in LB broth containing 100 µg/ml ampicillin. The plasmids were isolated using the QIAPrep Spin Miniprep kit (as described in method section 12.9.1) and sequenced (as described in method section 12.10).

13.2.4 Transformation of recombinant plasmid into expression host cells

The recombinant pET32a(+)-*FmL* was transformed into *E.coli* stain BL(DE3) star (Invitrogen). The competent cells of these host strains were prepared as described in method section 12.8.1. Plasdmids (10-100 ng) were used to transform into a 100 µl aliqout of the frozen competent cells, according to the standard method as described in method section 12.8.2 and spreaded on LB agar plate containing 100

µg/ml ampicillin. The recombinant clones were identified by colony PCR method as described in method section 12.8.3

13.3 Protein expression in *E.*coli

For protein expression, the selected clones were grown in LB broth containing 100 µg/ml ampicillin at 37°C with shaking at 180 rpm overnight. A volume of 2 ml of these cultures was used for inoculation into a fresh 200 ml LB broth containing 100 µg/ml ampicillin at 37°C at 180 rpm. When the turbidity $O.D._{600}$ of the culture reached to 0.5-0.7, 1 ml of sample from the culture was taken. It was subsequently centrifuged and the cell pellet was resuspended in 150 µl of 1x SDS-PAGE sample buffer, and then used as the control at 0 h. For the residual culture, IPTG was then added to a final concentration of 0.5 mM to induce the expression of FmL recombinant protein. The culture was further grown for three additional hours. The induced culture was transferred into 50 ml centrifuge tubes, chilled on ice for 10 min and then harvested by centrifugation at 4,000 x g for 10 min at 4°C. The cell pellet was kept at -80°C for 30 min or until used for protein extraction.

To evaluate the FmL expression levels by SDS-PAGE analysis, samples were taken every hour upto four hours after IPTG induction. Whole cell pellets from 1 ml culture were resuspended in 100 µl of 1x SDS-PAGE sample buffer (45 mM Tris-HCl, pH 6.8, 10% glycerol, 1% SDS, 0.01% bromophenol blue, 50 mM DTT). Samples were boiled for 5 min and centrifuged briefly. Five microliters of each sample and standard protein markers were analyzed by a 12% SDS-PAGE as described in method section 6.2. The gel was stained with 0.1% Coomassie brilliant blue R-250 as described in method section 6.3.1.

13.4 Purification of recombinant protein from induced cells

Bacterial culture pellets were thawed, resuspended in 5 ml (per gram of pellet bacteria) freshly prepared extraction buffer (20 mM This-HCl, pH 8.0, 0.2 µg/ml lysozyme, 1% Triton-X100, 1 mM phenylmethylsulfonylfluoride (PMSF), 40 µg/ml DNase I), and incubated for 30 min at room temperature. The cell suspensions were ruptured by 3 cycles of freezing at -80°C and thawing at 37°C following by centrifugation at 10,000 x g for 15 min. The protein lysate was gently mixed with 50% Ni-NTA slurry and shaked at 4°C for 1 h. The mixture was slowly poured into a column (1.0 cm in diameter and 8.0 cm in length) and allowed to flow down by

gravity. The protein bound with Ni-NTA was washed with the washing buffer (20 mM This-HCl, pH 8.0, 1% Triton-X100, 1 mM PMSF, 20 mM imidazole) for 5-10 column volumes. The recombinant protein was eluted by the elution buffer that composed of 20 mM Tris-HCl, pH 8.0, 1% Triton-X100, 1 mM PMSF, 250 mM imidazole. The protein was collected in 1.5 ml microfuge tubes (0.5 ml per each sample). The samples were kept at 4°C until used. In order to determine the purity of the samples, the eluate was subjected to 12% SDS-PAGE and stained with 0.1% Coomassie brilliant blue R-250.

Chapter 3

Results

1. Purification of lectin by repetitive Fetuin-agarose column

Based on hemagglutinating inhibition results demonstrated high affinity of *F. merguiensis* lectin (FmL) for N-acetylated sugars, such as ManNAc, GlcNAc, GalNAc, and especially NeuNAc (Rittidach *et al*., 2007). In this study, lectin from hemolymph of *F. merguiensis* was therefore purified by repetitive affinity chromatography on Fetuin-agarose column. The results of chromatography of *F. merguiensis* hemolymph on the first Fetuin-agarose column are illustrated in Fig. 22. The protein elution profile demonstrated two distinct peaks. The first peak F1, unretained fractions which contained most of serum proteins, was washed extensively with TB-NaCa. They showed no agglutinating activity in the presence of rabbit erythrocytes. This initial purification step was able to exclude a large portion of non-target proteins. The second smaller peak (F2) obtained after elution by 0.1 M GlcNAc, had hemagglutinating activity (HA) that increased by 1,727-fold when compared to that of the initial serum (Table 13). The absorbance profile shows a separation of the HA from a large proportion of contaminating proteins. The fractions with HA were pooled and concentrated by aquasorb to reduce volume. After each chromatographic step, protein contents and HA were measured and the specific HA was calculated. The eluted peak F2 contained 0.337 mg protein. The specific HA of the initial serum was 158.35 units/mg. After separation by the first Fetuin-agarose column (step1), the recovery of HA was 90%, accompanied by a 1,727-fold increase in the specific HA (273,472 units/mg). Concentrated peak F2 fractions were composed of two protein bands with M_r of ~300 and ~200 kDa (which is probably hemocyanin) as determined by PAGE under non-denaturing condition (Fig. 24, lane 3).

The concentrated protein, active fractions (peak F2) with high HA from the first Fetuin-agarose column, was further purified on second Fetuin-agarose column. The chromatographic profile was shown in Fig. 23. Again two peaks of protein appeared (Fig. 23), with the HA present only in peak S2. The fractions of the S2 peak were pooled and concentrated by aquasorb to reduce volume. The HA of the peak S2 fractions increased to 2,890-fold comparing to that of the initial serum (Table 13). This purification step resulted in an increase in the specific HA of *F. merguiensis* lectin from 273,472 to 347,654 units/mg. By means of two purification steps, the *F. merguiensis* lectin was purified with a 27.5% recovery of HA (Table 13) as determined by agglutination of rabbit red blood cells. After the second purification step, a single band was observed in a silver stained non-denaturing PAGE (Fig. 24, lane 4) that exhibited mobility corresponding to ~300 kDa, indicating that the lectin was pure. A summary of the purification results is given in Table 13. The overall recovery of *F. merguiensis* lectin from the hemolymph was 27.5% and it was purified up to 2,890-fold. The lectin from above purification step was kept as aliquots at -20°C for further determination of amino acid sequence.

Table 13 Summary of the purification of lectin from *F. merguiensis* hemolymph.

Fraction	**Protein (mg)**	**HA[a]**	**Specific activity[b]**	**Purification fold**	**Recovey (%)**
Hemolymph	646.65	102,400	158.35	1	100
First fetuin-agarose eluate peak F2	0.337	92,160	273,472	1,727	90
Second fetuin-agarose eluate peak S2*	0.081	28,160	347,654	2,890	27.5

[a] Hemagglutinating activity units tested in the presence of a 2% rabbit erythrocyte suspension.

[b] HA/mg protein.

***** Further purification of Fetuin-agarose eluate peak F2**.**

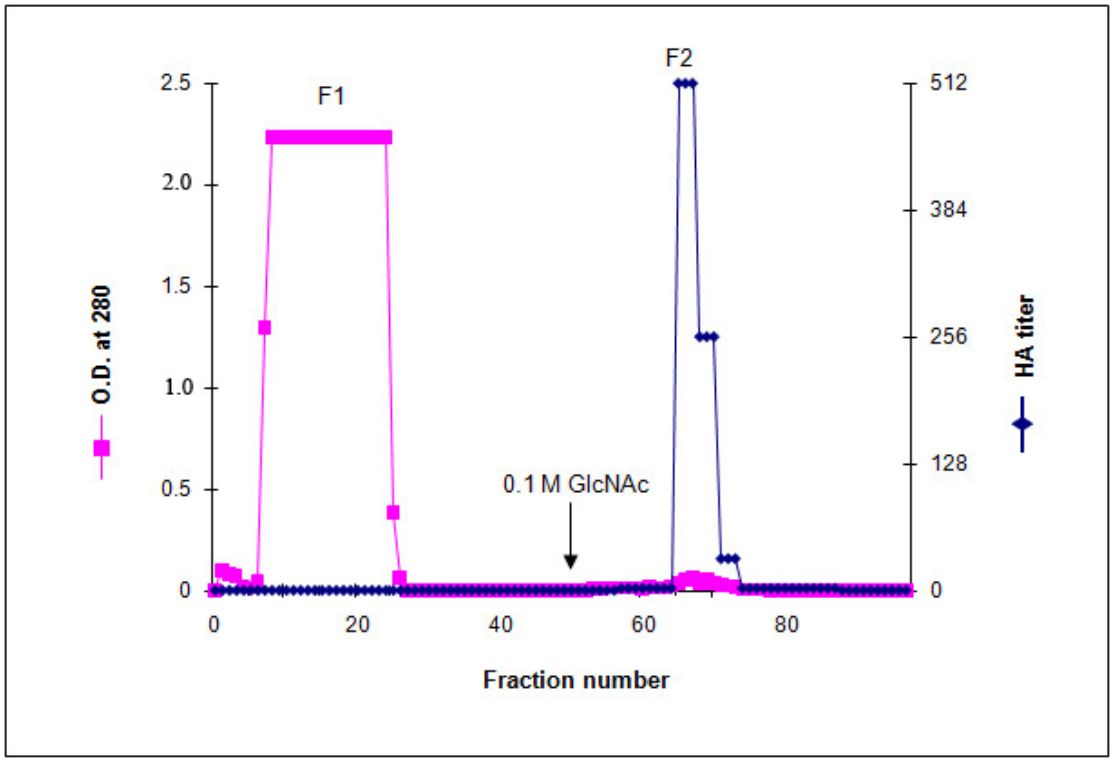


Fig. 22 Purification of lectin from the hemolymph by affinity chromatography on first Fetuin-agarose column.

Serum (646.65 mg protein) was applied onto the column (1.2 x 17 cm) equilibrating with TB-NaCa. After a period of 12 h for binding, unbound material was washed out with TB-NaCa, bound lectin was then eluted with the same buffer containing 0.1 M GlcNAc with a flow rate of 15 ml/h. Each fraction (1 ml) was measured for $O.D._{280}$ and HA in the presence of 2% rabbit erythrocytes after dialysis against TBS.

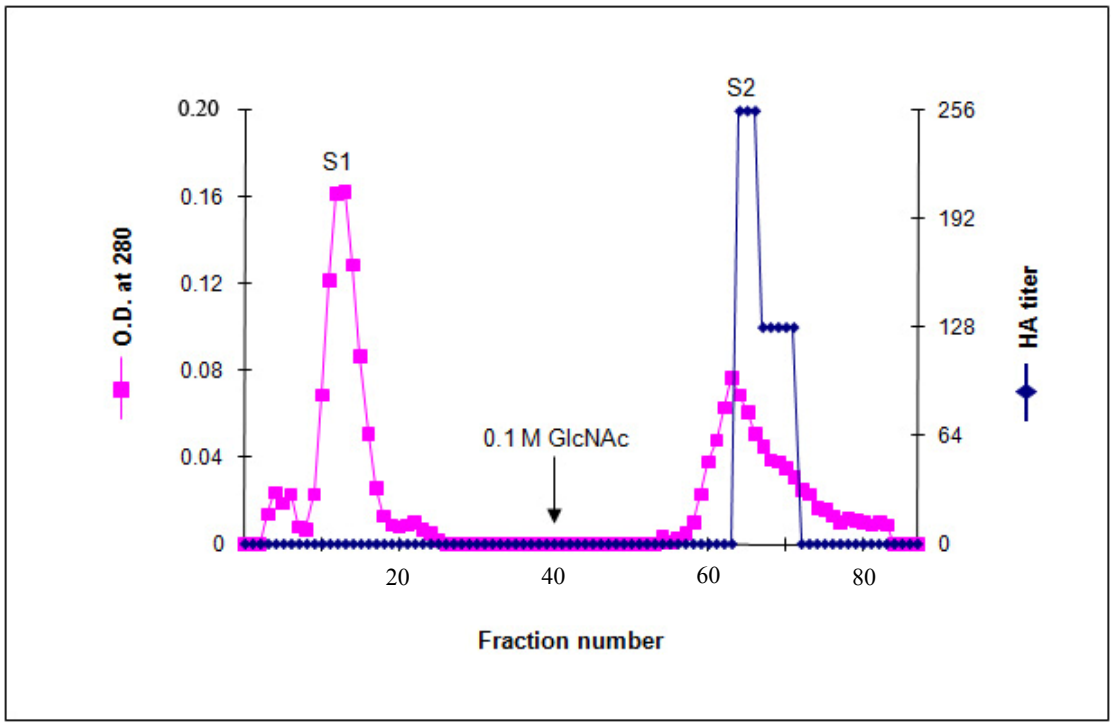


Fig. 23 Purification of lectin from the fractions peak F2 by affinity chromatography on second Fetuin-agarose column.

The dialyzed fractions with high HA (peak F2) from the first Fetuin-agarose column were loaded onto second Fetuin-agarose column, pre-equilibrated and washed by TB-NaCa with a flow rate of 15 ml/h. The $O.D._{280}$ and HA of 1 ml fractions were determined. Fractions 64-72 (peak S2), containing purified lectin, were pooled and concentrated.

2. Characterization of purified lectin

2.1 Determination of lectin purity and protein patterns on non-denaturing PAGE and SDS-PAGE

The composition of samples retained from each purification step was analyzed for purity and protein patterns by non-denaturing PAGE, as shown in Fig. 24. The electrophoretic patterns of fractions from purification process were run in 4-10% non-denaturing PAGE. The final product obtained from above purification procedure consisted only of *F. merguiensis* lectin as represented by a single band on lane 4 of the polyacylamide gel, indicating the purity of the lectin from second Fetuin-agarose. This purification results was satisfactory for the purpose of determining an N-terminal amino acid sequence.

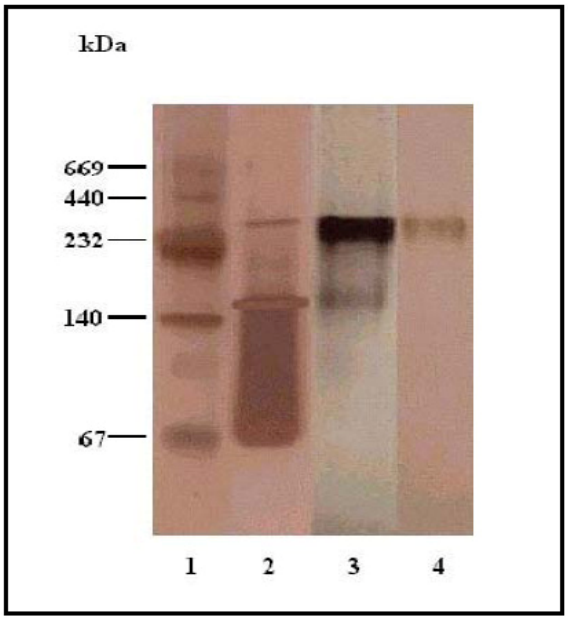


Fig. 24 Non-denaturing 4-10% gradient PAGE of lectin at each purification step.

The gel was stained with a silver staining kit. Lane 1, molecular weight markers; lane 2, hemolymph; lane 3, peak F2 from first Fetuin-agarose column; lane 4, peak S2 from second Fetuin-agarose column.

Subunits of purified lectin were characterized by 12% SDS-PAGE as shown in Fig. 25. When purified lectin was subjected to SDS-PAGE, it resolved into 2 major bands that appear to be arranged in doublets of 30.9 and 32.3 kDa either under reducing or non-reducing conditions (data not shown) as shown in Fig. 25.

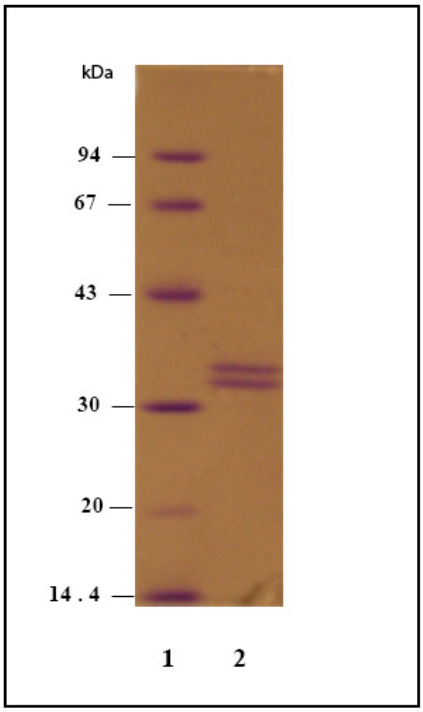


Fig. 25 SDS-PAGE of purified lectin on 12% gel.

The purified lectin was electrophoresed on 12% gel and protein bands were visualized by staining with Coomassie brilliant blue R-250. Lane 1: low molecular weight markers [phophorylase b (94 kDa), BSA (67 kDa), ovalbumin (43 kDa), carbonic anhydrase (30 kDa), trypsin inhibitor (20.1 kDa) and α-lactalbumin (14.4 kDa)]; lane 2, purified lectin.

2.2 Determination of N-terminal amino acid sequences of purified lectin

Twenty micrograms of purified lectin were separated in 6-18 % SDS-PAGE and electrophoretically transferred to a polyvinylidene difluoride (PVDF) membrane. After electroblotting, protein bands on the PVDF membrane were briefly stained with Coomassie brilliant blue R-250, destained and thoroughly washed to remove any glycine bound to the membrane until protein bands appeared (Fig. 26). The 30.9 kDa and 32.3 kDa bands were excised and sequenced using Edman degradation. The protein samples in the blot were sent to Bio Service unit (NSTDA). The N-terminal amino acid sequences of these two bands were determined up to 10 residues for both subunits, and the analysis for sequence homology to other known proteins was carried out by 'gapped basic local alignment search tool search' against the National Center for Biotechnology Information protein database (NCBI database). A homology search in the protein sequence database using the BLAST program did not reveal any high identity of the N-terminal amino acid sequences of *F. merguiensis* lectin to those of the other published lectins. There were no significant matches obtained from the database. The N-terminal sequence obtained was LAGAGTVLAG for both 30.9 and 32.3 KDa bands. The results indicate that the N-terminal sequences of both subunits of *F. merguiensis* lectin were identical.

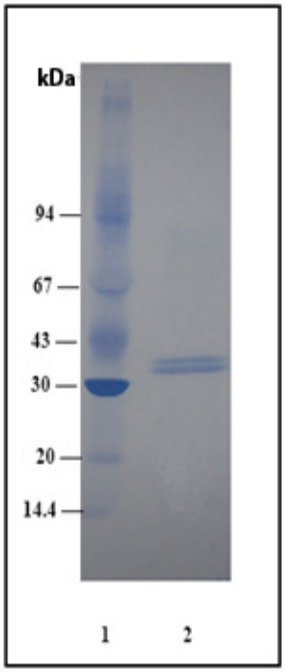


Fig. 26 SDS-PAGE of purified lectin after electroblotting to PVDF membrane.

The purified lectin was electrophoresed in 6-18% SDS-PAGE and electrophoretically transferred to a polyvinylidene difluoride (PVDF) membrane. The protein bands were visualized by staining with Coomassie blue. Lane 1; protein molecular weight markers [phophorylase b (94 kDa), BSA (67 kDa), ovalbumin (43 kDa), carbonic anhydrase (30 kDa), trypsin inhibitor (20.1 kDa) and α-lactalbumin (14.4 kDa)]; lane 2, purified lectin.

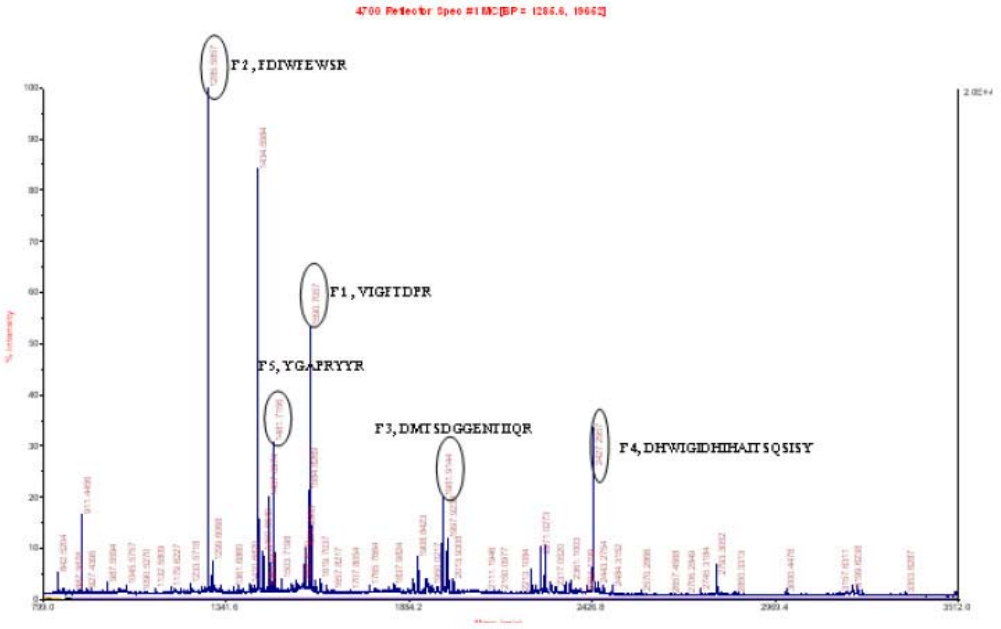
2.3 Peptide analysis by MALDI-TOF-TOF MS of purified lectin

The enzymatic digestion of purified lectin with trypsin and subsequent partially sequencing of the resulting peptides by mass spectrometric analysis was carried out by Dr. Wang Xianhui at the Department of Biological Sciences, National University of Singapore in Singapore.

In order to verify the identity of both protein bands (30.9 and 32.3 kDa), MALDI-TOF/MS was used coupling with peptide mapping fingerprinting (PMF) database search for protein identification. The 30.9 and 32.3 kDa bands from SDS-PAGE gel were separately digested by trypsin. Resulting peptides were then analyzed by MALDI-MS (matrix-assisted laser desorption and ionization-mass spectrometry). Fig. 27A and 27B represent the MALDI-MS data of the digests of 32.3 and 30.9 kDa bands, respectively. The PMFs of these two subunits showed similar tryptic patterns. After trypsin treatment, the MS profile of the peptides from both subunits showed 5 major peaks ranging from 1,285 to 2,428 kDa and these prominent peaks were selected for comparison with protein database. To further characterize the internal sequence of both subunits, the signal at 1590.70, 1285.59, 1981.91, 2427.29, 1481.72 Da were selected for amino acid sequence analysis. All 5 tryptic digested peptide sequences were identical for both subunits of purified lectin. The sequence information from this analysis suggested variation in amino acid sequences as shown in Table 14. The generated peptides were searched against the protein database as described in the experimental section.

The resulting tryptic digested peptide sequence information served for database searching using the NCBInr and MSDB databases. The Blast search revealed that the internal sequence of the peptide F3 displays a significant homology to conserved motif present in fibrinogen-related lectins, especially tachylectins 5A and 5B from the horseshoe crab (Gokudan *et al*., 1999). This search consistently detected significant homology between the peptide fragment F4 with ficolin.

(A) 32.3 kDa



(B) 30.9 kDa

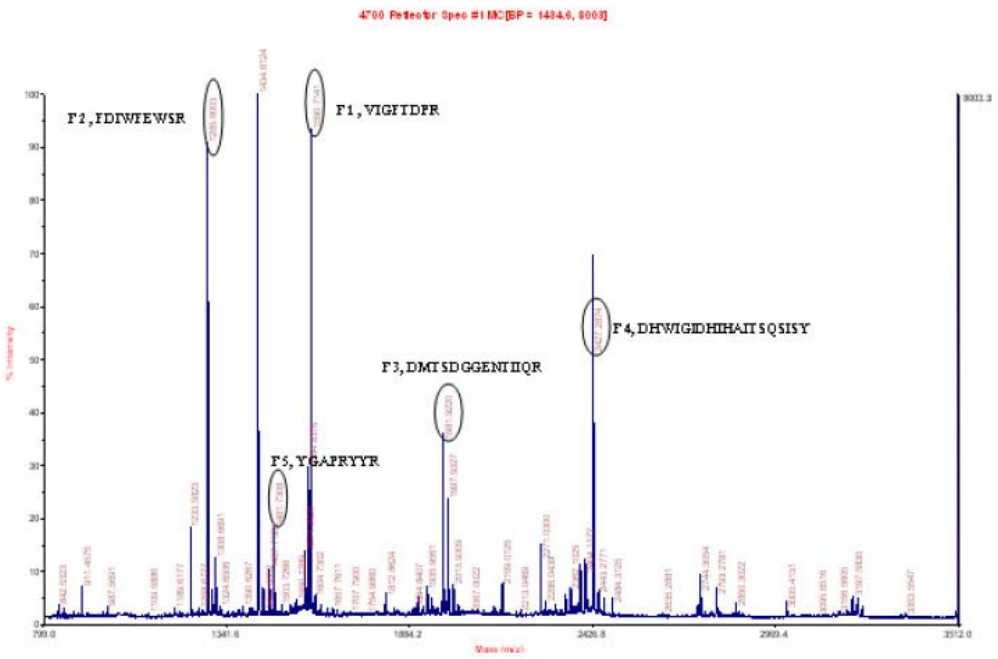


Fig. 27 MALDI-TOF-TOF MS profiles of tryptic digests of purified lectin subunits.

Table 14 Predicted amino acid sequences of tryptic digested peptides of purified lectin subunits.

Peptide number	Mass [m/z]	Peptide sequence
F1	1590.70	…..VIGFTDPR
F2	1285.59	FDIWFEWSR
F3	1981.91	DMTSDGGENTIIQR
F4	2427.29	….DHWIGIDHIHAITSQSISY…..
F5	1481.72	…..YGAPRYYR

3. Molecular approaches

3.1 Determination of RNA quality

Total RNAs were extracted from the hepatopancreas and hemocytes of *F. merguiensis* using the acid guanidinium thiocyanate phenol-chloroform extraction method and then treated with DNase I to remove the contaminated DNA. The $O.D._{260}/O.D._{280}$ ratios of the total RNA samples were in range of 1.7-2.1 which were within the excepted ratio for an acceptable quality of total RNA. The total RNA stability was determined by running in 1 % formaldehyde agarose gel electrophoresis. Bands of 18s and 28s rRNA were clearly observed on formaldehyde gel (Fig. 28). From the gel patterns and the absorbance ratios, the extracted total RNAs were good enough for cDNA library preparation. These total RNA were used for mRNA purification, and then for cDNA library construction.

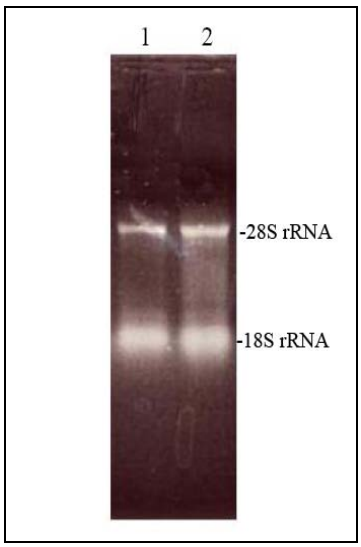


Fig. 28 Agarose electrophoretic patterns of total RNAs extracted from *F. Merguiensis*.

Each 5 μg of total RNA from the hemocytes and hepatopancreas was electrophoresed on 1% formaldehyde agarose gel and subsequently stained by ethidium bromide. Lane 1, total RNA from hemocytes; lane 2, total RNA from hepatopancreas.

3.2 cDNA library construction

The poly (A^+) mRNA was isolated from total RNA which extracted from the hepatopancreas or hemocytes. This mRNA was converted into double-stranded cDNA. Following ligation of an adapter and the release of an *Xho* I restriction site, it was ligated into the *Eco*R I /*Xho* I cut phosphatase-treated ZAP express vector. The resulting DNA was packaged *in vitro* using Gigapack III gold extracts and introduced into *E. coli* XL1-blue- MRF` strain. The titers of the primary library were $1.89x10^6$ and $4.5x10^5$ pfu/ml for hepatopancreas and hemocytes, respectively. These cDNA libraries were used as templates for amplifying the internal fragment of *FmL* and *FmLGBP* cDNA.

4. Molecular cloning and sequence analysis of cDNA encoding *F. Merguiensis* lectin (FmL).

4.1 FmL primer design

F. merguiensis C-type lectin cDNA was cloned by firstly obtained a fragment of the gene through PCR amplification using degenerate primers, designed to anneal to sequences corresponding to conserved regions of the proteins. Hepatopancreas cDNA library of *F. merguiensis* was used as a template. Comparative analysis of nucleotide sequences of C-type lectin gene from several shrimp species showed that stretches of nucleotide sequences were highly conserved among these shrimp lectins. A pair of degenerate primers (FmL-F1 and FmL-R1) (Table 6) were designed corresponding to the highly conserved nucleotide sequences of C-type lectins of tiger shrimp *P. monodon* (DQ871244), white shrimp *L. vannamei* (DQ871245), and green tiger shrimp *P. semisulcatus* (DQ871243)) in the GenBank database.

4.2 Cloning strategy of FmL gene

Cloning strategy and structure of the shrimp *FmL* are demonstrated in Fig. 29. The first-strand cDNA of 5' and 3' RACE is synthesized using a modified oligo (dT) primers and serve as the template for RACE PCR reactions. After the *FmL* cDNA fragment was obtained, 5' and 3' RACE approach was used to obtain clones for both ends. A full-length *FmL* cDNA was reconstructed from the overlapping sequences. The sequence information of both 5' and 3' cDNA ends was used to re-

designed specific PCR primers (FmL-F 1042 and FmL-R 1042) for amplification of a full-length *FmL* cDNA. The full length cDNA amplification of *FmL* was summarized in Fig. 29.

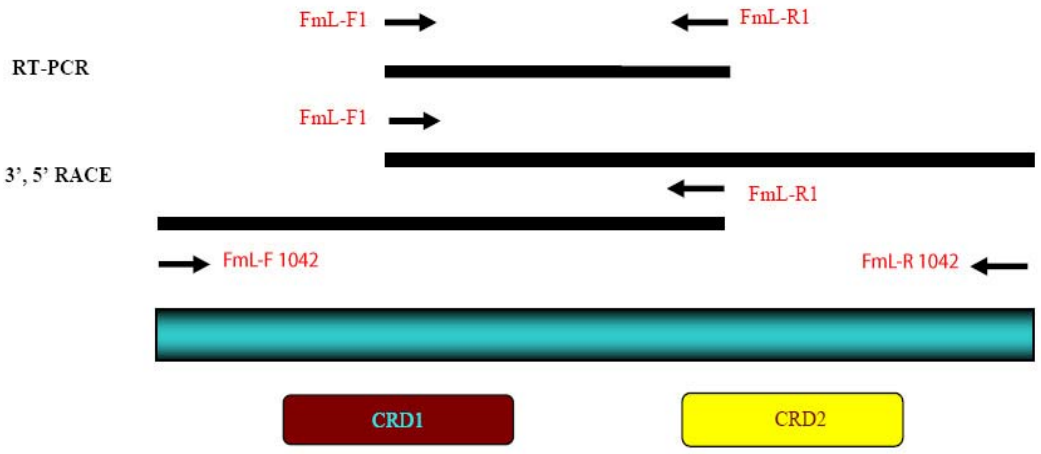


Fig. 29 Cloning strategy and structure of *F. merguiensis* C-type lectin cDNA (*FmL* cDNA).

A first cDNA fragment of *FmL* was cloned from the hepatopancreas cDNA library using primers FmL-F1 and FmL-R1. 5' and 3' RACE approach was used to obtain clones for both ends. The full-length *FmL* cDNA was constructed from the overlapping sequences. The *FmL* consisted of a short signal peptide (SP) and two CRD (carbohydrate recognition domain) domains. RT-PCR was performed using primers from 5' (FmL-F 1042) and 3' untranslated regions (FmL-R 1042) to confirm the full-length cDNA.

4.3 Amplification of *FmL* cDNA fragment

For the core fragment, the amplification of *FmL* cDNA fragment was performed with PCR method using the hepatopancreas cDNA library as a template. The FmL-F1 was used as forward primer and FmL-R1 was used as reverse primer (Table 6). PCR amplification was performed by pre-denaturation at 94°C for 5 minutes, 25 cycles of denaturation at 94°C for 30 seconds, annealing at 61°C for 30 seconds, and extension at 72°C for 1 minute, and the final extension at 72°C for 5 minutes. After amplification, the PCR product was separated in 1.2% agarose gel electrophoresis. A prominent PCR product of *FmL* cDNA fragment with the approximated size of 0.5 kb was observed as shown in Fig. 30. The PCR-amplified product was gel-purified using a QIAquick gel extraction kit (Qiagen). The purified cDNA fragment was ligated into pDrive cloning vector and transformed to *E. coli* cells. After selection the white colonies, the resulting plasmid was extracted and re-checked for size of PCR inserted fragment before nucleotide sequencing. The sequence analysis revealed that *FmL* cDNA fragment contained a length of 531 bp and its sequence was shown in Fig. 31. By using the Blastx analysis, its nucleotide sequence was translated to 177 amino acids and compared with various proteins in the NCBI database. Initial sequence analysis indicated the presence of one open reading frame (ORF) encoding a polypeptide with a high degree of similarity to C-type lectin of many organisms. This fragment contained an ORF for deduced amino acid sequence, which confirmed a similarity to the C-type lectin family. Deduced amino acid sequence comparison revealed a high level of similarity to other shrimp lectins (data not shown). After that this amino acid sequence was subjected to the BlastP program to determine domain(s) containing within *FmL* cDNA. CRDs were found in *F. merguiensis* cDNA fragment and shown in Fig. 32. The partial cDNA of *F. merguiensis* lectin was mostly similar to *P. monodon* C-type lectin with 78% identity. The result suggesting that the obtained cDNA fragment was a part of a gene for shrimp C-type lectin.

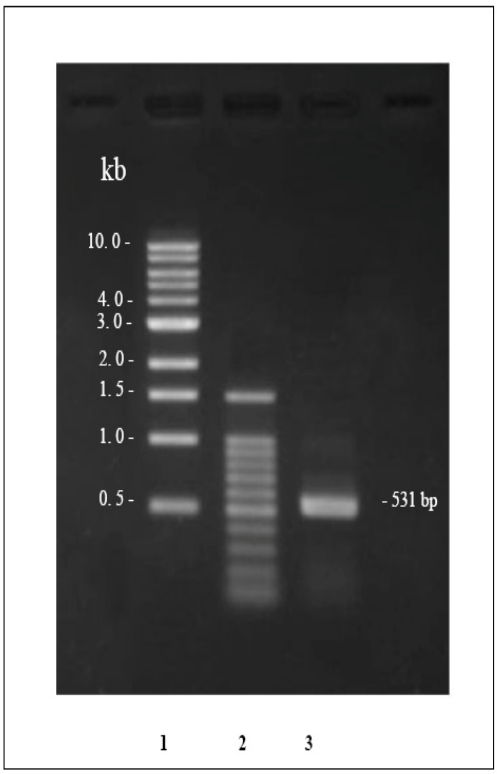


Fig. 30 Amplification of cDNA fragment encoding FmL.

PCR was performed by using a pair of FmL-F1 and FmL-R1 primers and the hepatopancreas cDNA library as template. The PCR product was analyzed in 1.2% agarose gel electrophoresis. Lane 1, 1 kb DNA ladder marker; lane 2, 100 bp DNA ladder markers; lane 3, PCR product amplified of cDNA encoding FmL.

```
      D  G  R  D  E  E  T  E  G  V  W  V  T  A  T  G  E  A  V  P·
1     TGGACGGACGAGACGAGGAAACAGAGGGCGTTTGGGTGACTGCGACAGGAGAAGCCGTGC
      ·P  L  G  T  P  F  W  A  A  F  D  Y  H  Q  Q  P  D  N  S  L  G·
61    CCTTAGGAACCCCCTTCTGGGCCGCCTTTGATTATCACCAGCAACCGGATAACTCCCTCG
      ·G  N  E  H  C  L  S  I  P  S  D  W  F  L  Y  M  N  D  V  P  C·
121   GGAACGAACACTGTCTGTCCATTCCATCGGATTGGTTTCTCTATATGAATGACGTCCCTT
      ·C  S  G  T  K  H  F  I  C  E  A  P  V  Q  R  Q  E  E  A  A  G·
181   GTTCTGGCACAAAACACTTCATATGTGAAGCCCCCGTCCAAAGGCAGGAGGAGGCCGCAG
      ·G  A  A  L  A  P  V  G  P  S  A  V  G  G  R  V  E  C  L  S  P·
241   GCGCAGCGCTGGCTCCCGTCGGCCCGTCGGCGGTGGGCGGCCGCGTCGAGTGCCTCAGCC
      ·P  F  V  E  V  G  D  L  C  L  L  F  I  T  W  A  E  E  T  W  D·
301   CCTTCGTGGAGGTGGGAGACCTGTGCCTGTTGTTCATCACGTGGGCGGAGGAGACCTGGG
      ·D  V  A  Q  Q  L  C  A  G  V  S  S  E  L  L  A  I  T  D  V  E·
361   ACGTGGCTCAGCAGTTGTGTGCGGGAGTCTCCAGCGAGCTCTTGGCCATCACCGATGTTG
      ·E  V  L  R  T  V  Y  L  F  I  Q  S  N  G  L  A  G  N  T  F  W·
421   AAGTGTTGCGAACTGTGTACCTCTTCATACAGTCAAATGGTTTAGCTGGCAACACCTTCT
      ·W  L  G  G  S  D  L  E  N  E  G  T  W  V  Y  T  T
481   GGCTTGGAGGATCCGATCTGGAGAATGAAGGGACATGGGTCTACACCACAG
```

Fig. 31 Nucleotide and deduced amino acid sequence of partial cDNA of *FmL*.

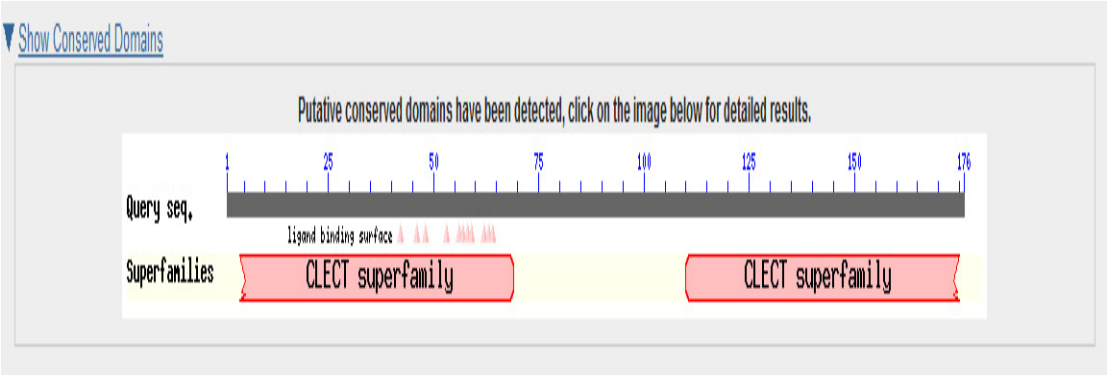


Fig. 32 Conserved domains of partial *FmL* cDNA.

The BLAST results showed two putative conserved domains of partial *FmL* sequences.

4.4 Amplification of 5’ and 3’ cDNA ends of *FmL*

In order to determine the remaining DNA sequence of FmL gene, the 5’ and 3’ ends of the cDNA were completed by means of 3’ and 5’ RACE strategies. The 3’ and 5’ ends of *FmL* cDNA was produced using the RACE system Gene Racer^{TM} Kit (Invitrogen, USA). The first-strand cDNA was synthesized from total RNA isolated from the hepatopancreas. The 3’ end of *FmL* cDNA was amplified with primer GSP 3’ FmL and GeneRacer 3’ primers in the first round of PCR and primers GSP 3’ FmL primer and GeneRacer 3’ nested in the second round of PCR. The 3’ RACE PCRs yielded PCR product of 778 bp in length (Fig. 33). An amplified cDNA product of 3’ end was purified from agarose gel, subcloned into pCR4-TOPO vector and then transformed to *E. coli* cells. After plasmid extraction, the purified plasmid were sequenced and analyzed by the Vector NTI software (version 9.0). The nucleotide and deduce amino acid sequence of the 3’ end fragment showed identical matching with overlapping sequence of the core fragment and showed a high degree of similarity to those of 3’ ends of other shrimp C-type lectins. The 3’UTR displayed a poly(A) tail at 3’ end, confirming that we had obtained the sequences in their integrity at this end. In the similar manner, the 5’ end of *FmL* cDNA was amplified with GSP 5’ FmL and GeneRacer 5’ primers in the first round of PCR and GSP 5’ FmL and GeneRacer 5’ nested primers in the second round of PCR. The 5’ RACE PCRs yielded PCR product of 871 bp in length (Fig. 34). An amplified cDNA product of 5’ end was purified from agarose gel, subcloned into pCR4-TOPO vector and then transformed to *E.coli* cells. After plasmid extraction, the purified plasmid were sequenced and analyzed by Vector NTI software (version 9.0). The nucleotide and deduce amino acid sequence of the 5’ end fragment showed identical matching in overlapping sequence to the core fragment and showed a high degree of similarity to those of 5’ ends of other shrimp C-type lectins. The identity of the clones was confirmed by matching the sequences to the nucleotide/protein sequences available in GenBank (http://www.ncbi.nlm.nih.gov/, accessed 9/04/07). After sequence analysis, PCR products obtained from 5’ RACE amplification showed a high sequence homology to other shrimp C-type lectins.

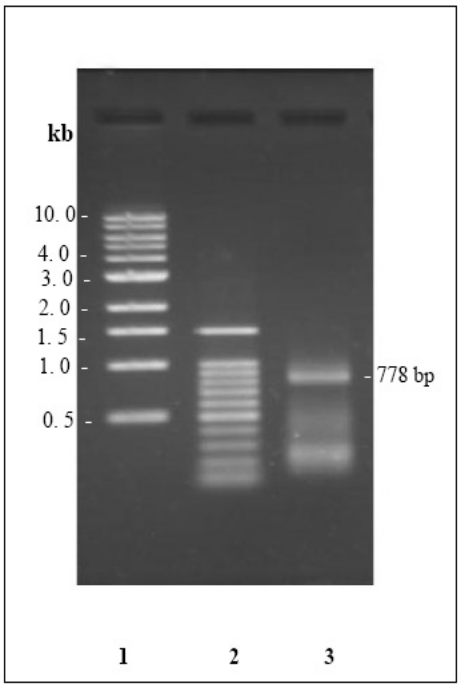


Fig. 33 *FmL* cDNA electrophoretic pattern of 3' RACE fragment.

The PCR product from 3' RACE was analyzed in 1.2% agarose gel electrophoresis. Lane 1, 1 kb DNA ladder markers; lane 2, 100 bp DNA ladder markers; lane 3, DNA fragments of 3' RACE.

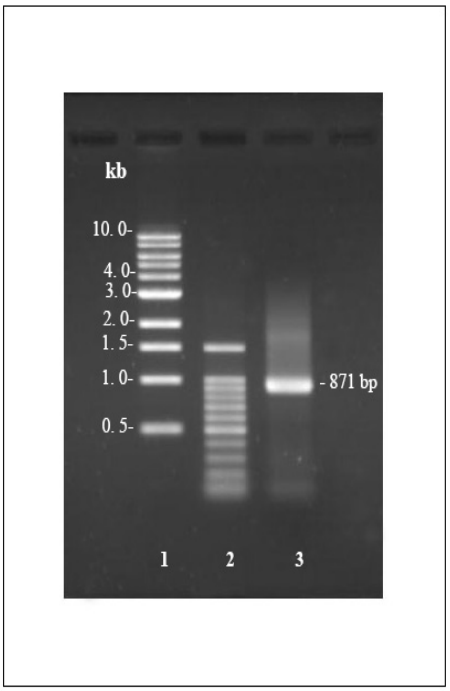


Fig. 34 *FmL* cDNA electrophoretic pattern of 5' RACE fragment.

The PCR product from 5' RACE was analyzed in 1.2% agarose gel electrophoresis. Lane 1, 1 kb DNA ladder markers; lane 2, 100 bp DNA ladder markers; lane 3, DNA fragments of 5' RACE.

4.5 Amplification of open reading frame (ORF) of *FmL*

According to the nucleotide sequence of the 5' and 3' ends, the specific primers were designed (Table 11). A pair of primers was used to amplify the ORF of *FmL*. DNA fragment covering ORF was obtained by PCR with 5' and 3' flanking primers. After amplification, the resulting nucleic acid product was run on a 1.2% agarose gel. From agarose gel electrophoresis result (Fig. 35), the experimental product size appeared related to the expected size. The PCR product that amplified by these specific primers was purified and ligated into pCR4-TOPO vector. The recombinant plasmid was isolated from the white colonies. After that positive clones were determined for the inserted *FmL* appearance with restriction enzymatic digestion. The plasmids were purified, then sequenced and analyzed by Vector NTI software (version 9.0).

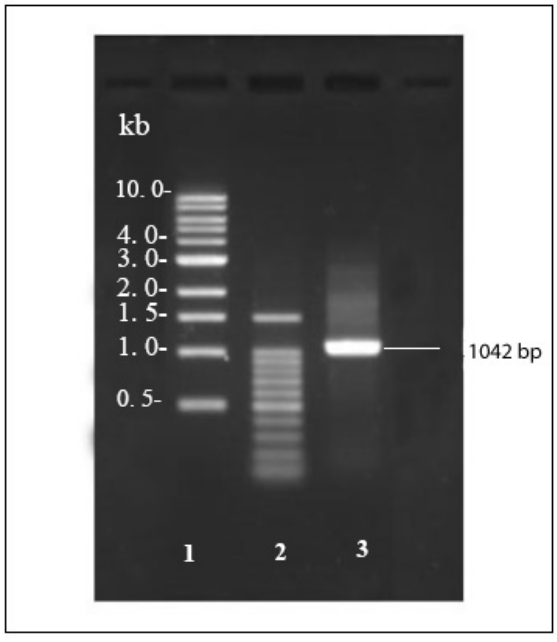


Fig. 35 *FmL* cDNA electrophoretic pattern of open reading frame.

The PCR product was analyzed in 1.2 % agarose gel electrophoresis. Lane 1, 1 kb DNA ladder markers; lane 2, 100 bp DNA ladder markers; lane 3, PCR product of a large fragment of *FmL* open reading frame.

4.6 Sequence analysis of *FmL* cDNA

Analysis of the nucleotide sequences of both 5' and 3' RACE products of *FmL* revealed that these two cDNA fragments encompassed part of 5' untranslated region (5' UTR) and the 3' UTR of *FmL*, as well as a portion of coding region. The full-length cDNA of *FmL* was covered by joining these three fragments. The full-length nucleotide and deduced amino acid sequences are shown in Fig. 36. The full-length cDNA was 1118 bp in length with a 5' UTR of 42 bp and a 3' UTR of 74 bp with a poly(A) tail. The full-length cDNA sequence of *FmL* revealed the presence of one ORF consisting of 1,002 bp. The start and stop codons were located at positions 43 and 1044, respectively. The ORF of *FmL* cDNA encodes for a mature protein of 333 amino acid residues with the calculated molecular mass of 36,263 Da and a pI of 4.14 (Compute pI/MW program available: http://www.expasy.org/cgibin /pi_tool, accessed 09/04/07). It is almost likely a secreted protein with a predicted signal peptide by the Signal P 3.0 server program (Bendtsen *et al.,* 2004). Using the signal peptide prediction software (Signal P: http:// www.cbs.dtu.dk/services/SignalP/, accessed 9/04/07), it was predicted that amino acid residues 1-20 represent the signal peptide region of *FmL*. The prediction result shown the signal peptide was spanning 20 residues from residue 1 to 20 and the cleavage site between Gly^{20} and Arg^{21} of *FmL*. The characterization of the ORF amino acid sequence was accomplished by its comparison to amino acid sequences of other C-type lectins. In *FmL*, three tandem repeats can be detected. The first repeat contains an amino acid sequence of Glu-Gly-(Val/Thr)-Trp-Val located at amino acid residues 108-112 and 269-273. The second repeat consists of Gly-Glu-(Ala/Pro)-Val-Pro-(Leu/Met)-Gly-Thr-Pro-Phe-Trp sequence located at amino acid residues 116-126 and 277-287 (Fig. 36). The third repeat consists of Leu-Ala-Ile-Thr (or Gly) amino acid residues located at 234-237 and 308-331. In addition, SMART prediction revealed that *FmL* also consists of two predicted carbohydrate-recognition domains (CRDs) in the deduced amino acid sequence. The first CRD is located at amino acid residues 44-169 and it contains a putative galectose binding motif (QPD). The second CRD is located at amino acid residues 204-330 and it contains a putative mannose binding motif (EPN) (Fig.36). The two CRDs shared 36.2% identity (or 48% similarity) to each other. The mature protein included one putative *N*-linked glycosylation site (Asn-Xaa-Ser/Thr) at Asn^{25}

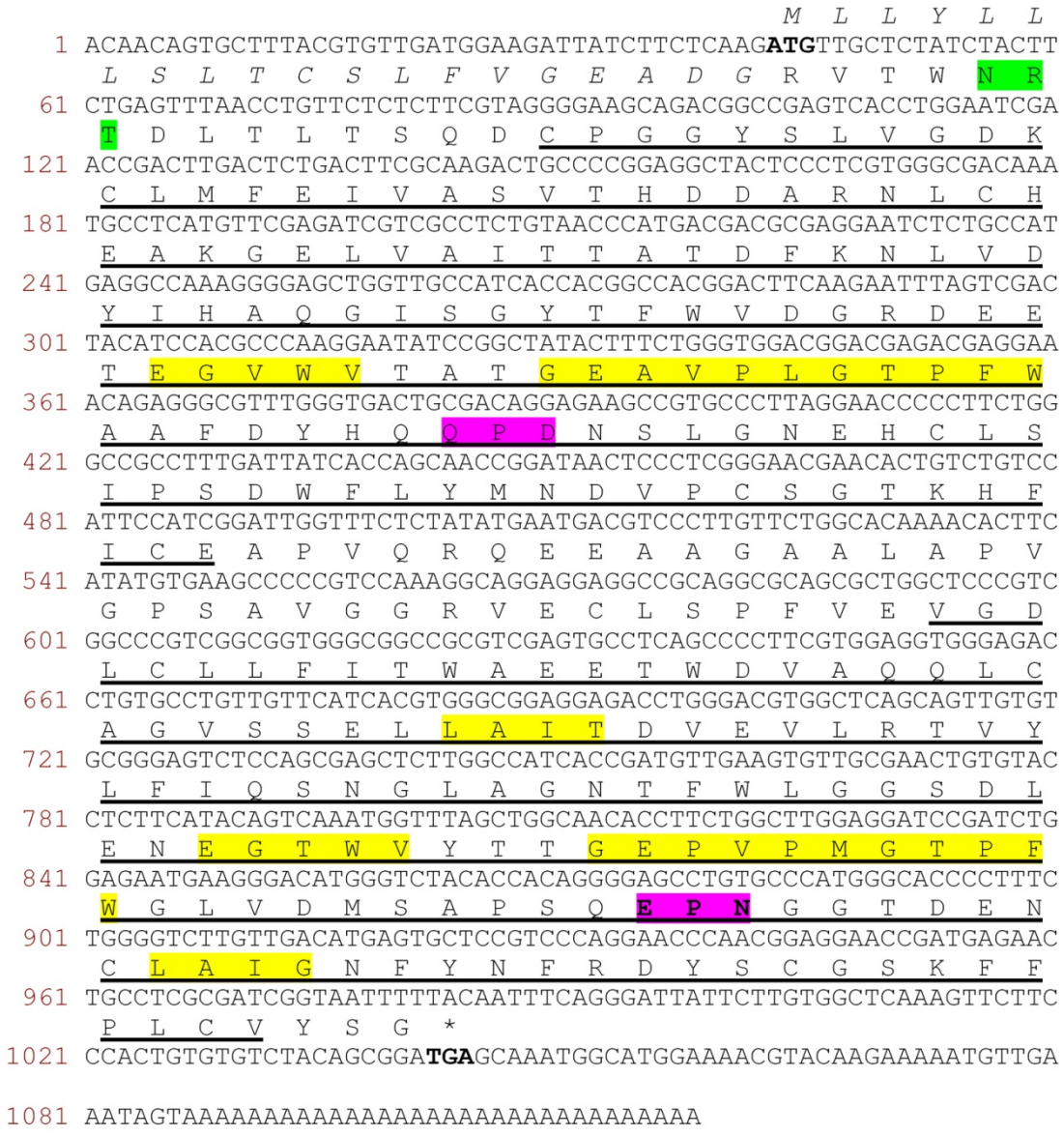


Fig. 36 Nucleotide sequence of cDNA encoding C-type lectin from *F. merguiensis* and its deduced amino acid sequence.

The predicted ORF is from 43 to 1041 bp (the start and stop codons are shown in bold letters). Two CRDs are underlined. The conserved amino acid residues specific to the binding by galactose (QPD) and mannose (EPN) are pink highlighted. The yellow highlighted amino acid sequences represent three tandem repeats of *FmL*. Potential *N*-glycosylation sites are green highlighted. The stop codon is marked with asterisk.

and was suggested to be a glycoprotein. One potential *O*-glycosylation site was found at Thr^{284}.

Using different protein analysis and computer programs contained in the Expasy tools (http://www.expasy.ch/tools, accessed 9/04/07), two putative conserved domains were found in FmL protein : CLECT domains which are homologous to CRDs of the C-type lectins (Fig. 37). They are found in member of C-type lectins. The first CRD is located at amino acid residues 44-169 and it contains a putative galectose binding motif (QPD). The second CRD is located at amino acid residues 204-330 and it contains a putative mannose binding motif (EPN) (Fig.36). The two CRDs shared 36.2% identity (or 48% similarity) to each other.

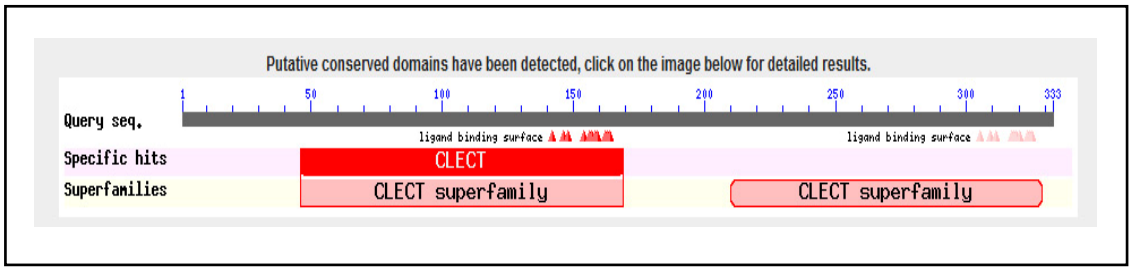


Fig. 37 Putative protein domains present in FmL protein.

The FmL protein contains two putative domains: one CRD1 domain and one CRD2.

4.7 C-type lectin sequence comparisons and phylogenetic analysis

In order to determine the degree of homology, the nucleotide sequence of *F. merguiensis* lectin was compared with all available sequences in the NCBI database using BLAST analysis. The pairwise alignment of *F. merguiensis* lectin sequence to those of lectins from *P. monodon* (PmL), *P. semisulcatus* (PsL), *L. vannamei* (LvL) showed a high identity of 88, 83, and 81%, respectively.

Homology comparison of the *FmL* deduced amino acid sequence with database protein sequences was performed by computer in the NCBI using the BLAST network service. The deduce amino acid sequence of *FmL* revealed highly homologous to those of other shrimp lectins. The comparison of deduced amino acid sequence of *F. merguiensis* lectin with those of other crustacean lectins, showed a relatively high degree of overall homology. Amino acid sequences of 4 shrimp C-type lectins that consist of dual CRDs (*FmL, PmL, LvL, PsL*) were aligned in Fig. 38. The alignment of amino acid sequences also revealed many conserved regions. Percentages of identity and similarity were also assessed between different modular domains (Fig. 39). The amino acid sequence of *FmL* has highest homology (82% identity) with C-type lectin from tiger shrimp *P. monodon* (*PmL*). Similarity with C-type lectin from green tiger shrimp *P. semisulcatus* (*PsL*) and white shrimp *L. vannamei* (*LvL*) were found to be higher than 60% identity. The homology of *FmL* was comparatively lower to the *Fc-hsL* of *F. chinesis* lectin (34% identity, 47% similarity). *FmL* also has low amino acid similarity to the C-type lectin-like domain-containing protein *PtLP* of *Portunus trituberculatus* (31% identity, 50% similarity), and lectin of clam *Chlamys farreri* (26% identity, 39% similarity).

Fig. 38 shows multiple alignments of the deduced amino acid sequences of *FmL* with those from various shrimps. The amino acid sequences of the shrimp lectins that consist of dual CRDs (*PmL, LvL, PsL*) were aligned with *FmL*. In each CRD, the locations of cysteine residues that are important for the formation of disulfide bridges are conserved. In lectins FmL, PmL, PsL and LvL, a QPD motif is located in the first CRD and an EPN motif is located in the second CRD (Fig. 38). Moreover, the CRDs of these lectins share the similarity in range of 64-83% identity (Fig. 39) suggesting that they may also share similar carbohydrate (polysaccharide) binding functions.

A phylogenetic analysis was also performed by using the multiple alignment amino acid sequences from CRDs of lectin from crustacean, rainbow trout, and eel. A pairwise deletion option for gaps was used and groupings in the Neighbor-Joining analyses were supported by bootstrap values of one thousand times. Homologue sequences used in the alignment were obtained from NCBI Genbank and aligned using Clustal W. The Neighbour-Joining tree was produced using the phylogenetic program, MEGA (version 4.1) (Kumar *et al.,* 2004) and minimum evolutionary distance was calculated using pairwise alignments. FmL, PmL, PsL and LvL lectins share a close evolutionary relationship in which each of their CRD1s cluster together separately from their CRD2s and cluster with the single CRDs from Fc-hsL and PmAV. The two CRDs from Fclec and the single CRD from PmLec cluster together separately from the other shrimp (Fig. 40).

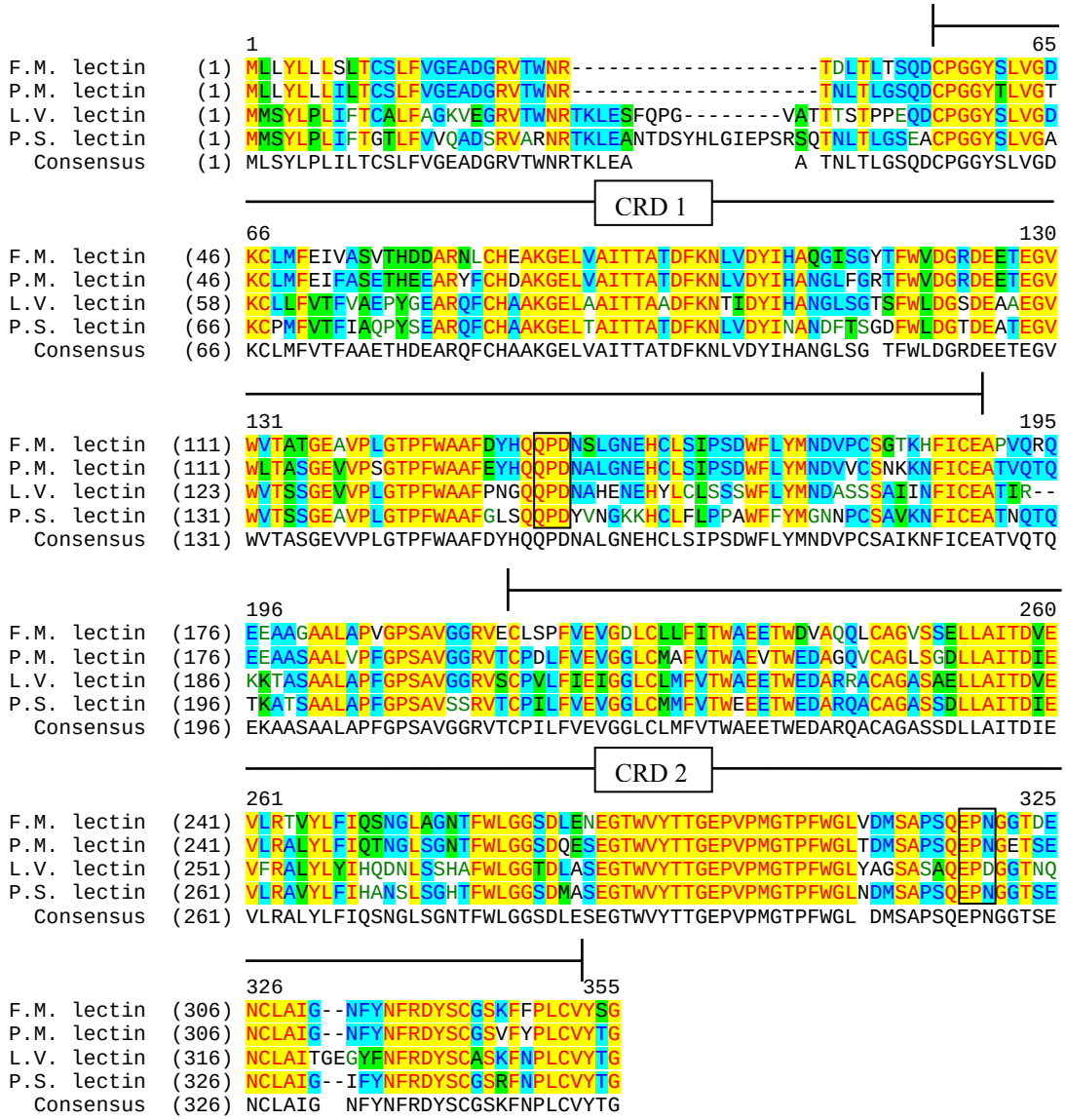


Fig. 38 Multiple alignments of *F. merguiensis* lectin with other shrimp lectins.

F.M., *F. merguiensis*; P.M., *P. monodon*; L.V., *L. vannamei*; P.S., *P. semisulcatus*. Numbers on the left indicate the amino acid position of different sequences. Identical amino acids are shaded in yellow. Other conserved, but not consensus amino acids, are shaded in blue and green.

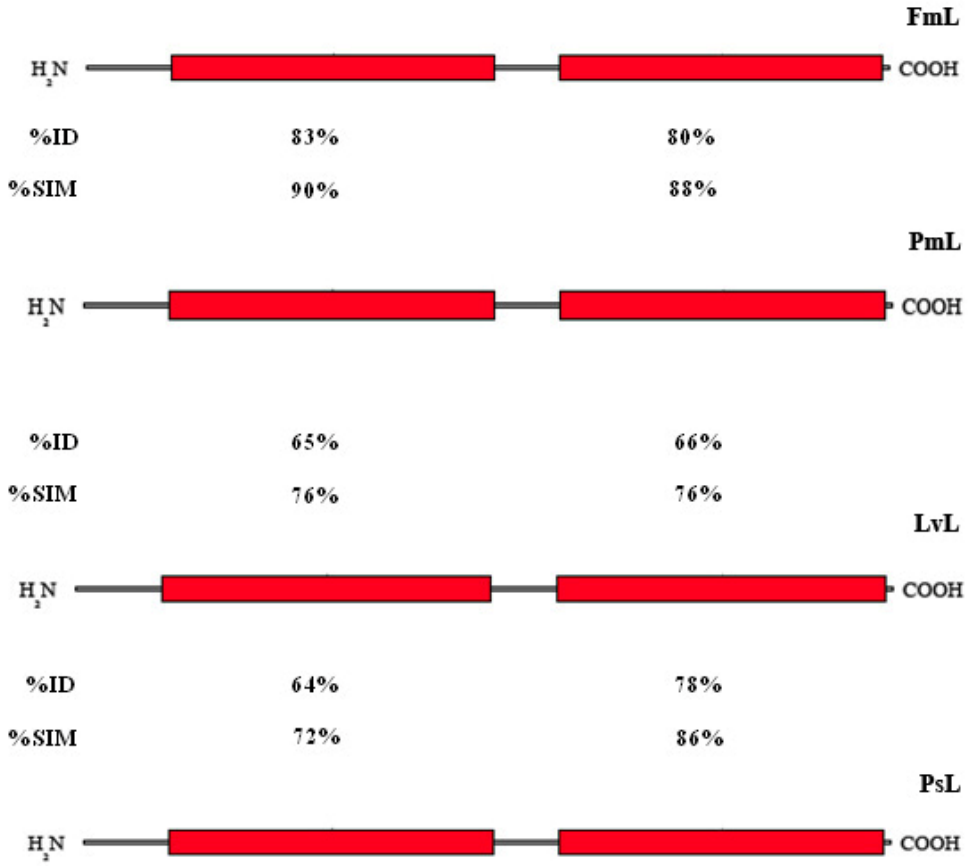


Fig. 39 The arrangement of modular domains for shrimp lectins and the schematic alignment of FmL modular domains with those of other shrimps.

Pairwise alignment was performed with EMBOSS Pairwise Alignment Algorithms of EBI. %ID, percent identity; %SIM, percent similarity; FmL, *F. merguiensis* lectin; PmL, *P. monodon* lectin; LvL, *L. vannamei* lectin; PsL, *P. semisulcatus* lectin

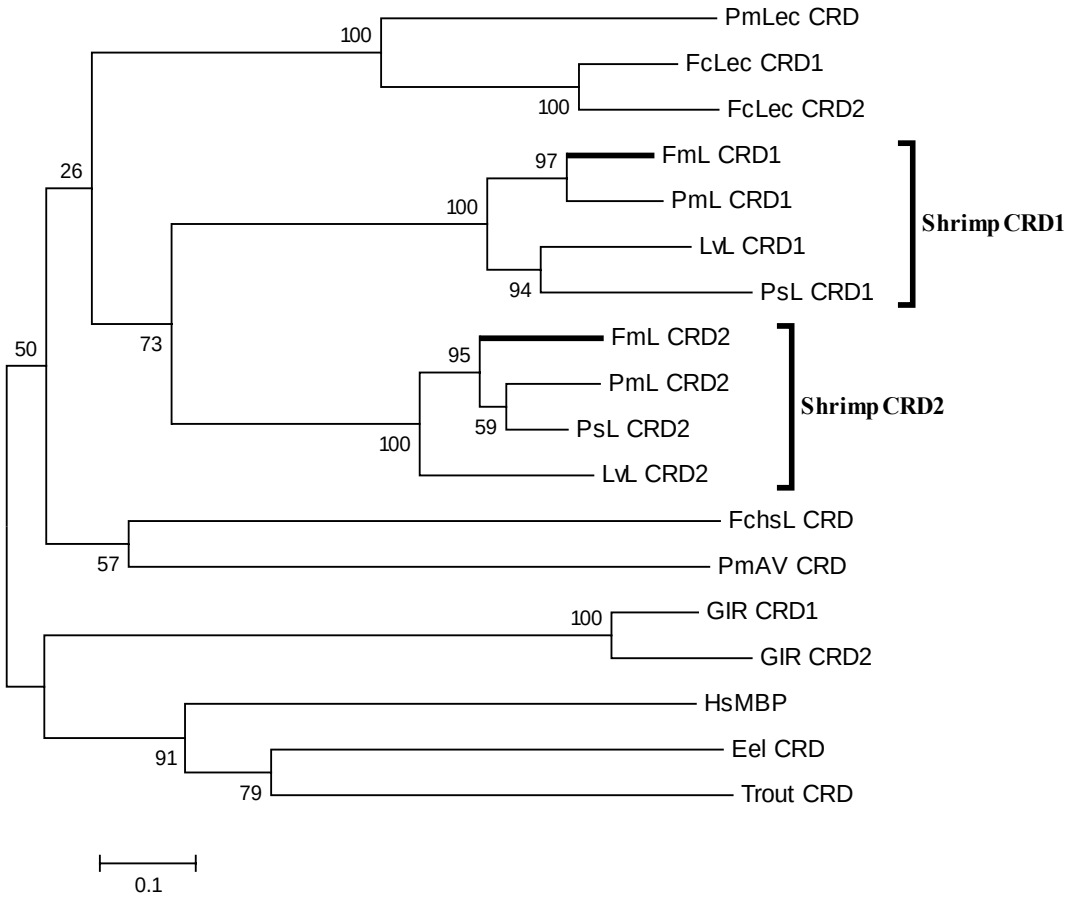


Fig. 40 Phylogenetic analysis based on the amino acid sequences of carbohydrate recognition domains (CRDs) of FmL with other C-type lectin family members.

Neighbour-Joining tree produced by MEGA 4.1. One thousand bootstraps were performed for checking reproducibility of the result. The CRD sequences were obtained from lectin-like molecules of shrimp *P. monodon* (PmAV GenBank # AAQ75589, PmLec GenBank # AAZ29608, PmL GenBank # ABI97373), *F. chinensis* (FchsL GenBank #ABA54612, FcLec GenBank # AAX63905), *P. semisulcatus* (PsL GenBank # ABI97372). *L. vannamei* (LvL GenBank # ABI97374, worm *Girardia* (GIR GenBank # AAL29933), rainbow trout *Oncorhynchus mykiss* (Trout GenBank # AAM21196), Japanese eel *Anguilla japonica* (Eel GenBank # BAC54021) and Human *Homo sapiens*(GenBank # AAG13815).

5. Molecular study of lipopolysaccharide and β-1, 3-glucan binding protein (LGBP) of *F. merguiensis*

5.1 FmLGBP primer design

Comparative analysis of amino acid sequences of LGBP gene from several shrimp species showed that stretches of amino acids are highly conserved among LGBPs (data not shown). This multiple sequence alignment of known shrimp LGBP sequences led to the identification two stretches of conserved amino acids SGGGNWEFQ and PFDQKFYLI. A pair of degenerate primers (FmLGBP-F1 and FmLGBP-R2) was designed corresponding to these two conserved stretches of amino acids (Table 6) and were used in PCR experiments to amplify a segment of *FmLGBP* cDNA.

5.2 Cloning strategy of FmLGBP gene

Cloning strategy and structure of the shrimp *FmLGBP* are demonstrated in Fig. 41. The first-strand cDNA of 5' and 3' RACE is synthesized using a modified oligo (dT) primers and serve as templates for RACE PCR reactions. After the *FmLGBP* cDNA fragment was obtained, 5' and 3' RACE approach was used to get clones for both ends. *FmLGBP* cDNA was reconstructed from the overlapping sequences. The sequence information of both 5' and 3' cDNA ends was used to re-designed specific PCR primers (FmL-F3 and FmL-R3) for amplification of a full-length *FmLGBP*. The full-length cDNA amplification of *FmLGBP* was summarized in Fig. 41.

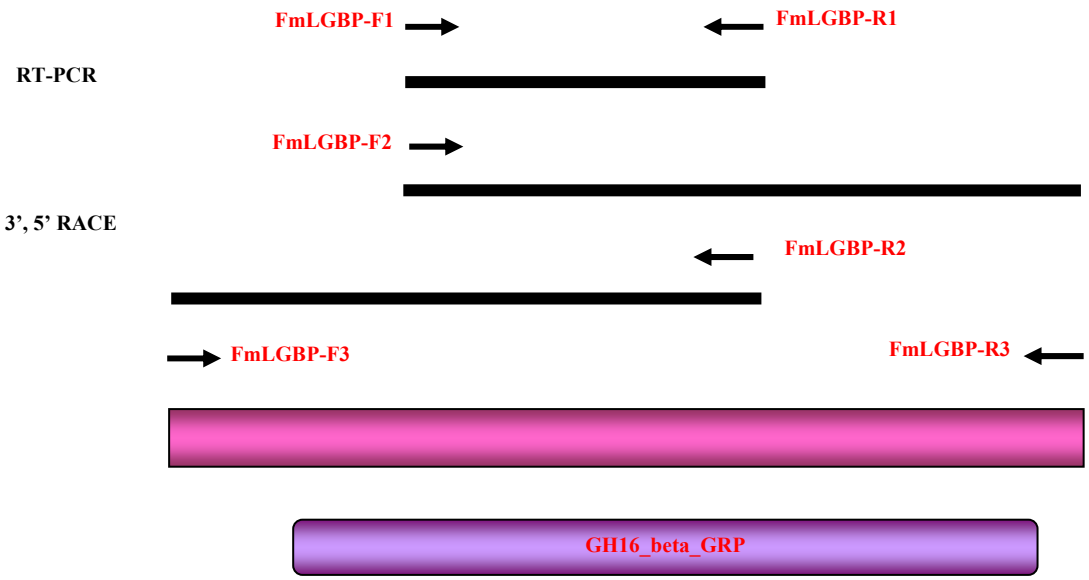


Fig. 41 Cloning strategy and structure of *FmLGBP*.

A partial cDNA of FmLGBP was cloned from the hemocytes using primers FmLGBP-F1 and FmLGBP-R1. 5' and 3' RACE approach is used to obtain clones for both ends and the full-length *FmLGBP* cDNA is reconstructed from the overlapping sequence. *FmLGBP* consists of a short signal peptide (SP) and GH16_beta_GRP domian. RT-PCR was performed using primers from 5' (FmLGBP-F3) and 3' untranslated regions (FmLGBP-R3) to confirm the reconstructed cDNA.

5.3 Amplification of *FmLGBP* cDNA fragment

For the core cDNA fragment, the amplification of *FmLGBP* cDNA was cloned by first obtaining a fragment of the gene through PCR amplification using degenerate primers. We created a hemocyte cDNA library of *F. merguiensis* for using as templates. After amplification, the resulting PCR product was electrophoresed on 1.2% agarose gel. An amplified cDNA product of 729 bp was obtained (Fig. 42). The purified fragment was ligated to pDrive cloning vector and transformed to *E.coli* cells. After selection the white colonies, the resulting plasmids were extracted and re-checked for size of PCR inserted fragment before nucleotide sequencing. The sequence analysis revealed that *FmLGBP* cDNA core fragment contained a length of 729 bp (Fig. 42) and its sequence was shown in Fig. 43.

Initial analysis of nucleotide sequence of the core *FmLGBP* fragment using the BLAST program indicated that the segment of *FmLGBP* cDNA shared high sequence homology to LGBP gene from various shrimps (>89% identity). The highest sequence homology was found to *F. chinensis* LGBP (95 % identity). Sequence analysis indicated the presence of one ORF encoding a polypeptide with a high degree of similarity to LGBP of many shrimp species and almost matching to *F. chinensis* LGBP with 95 % identity.

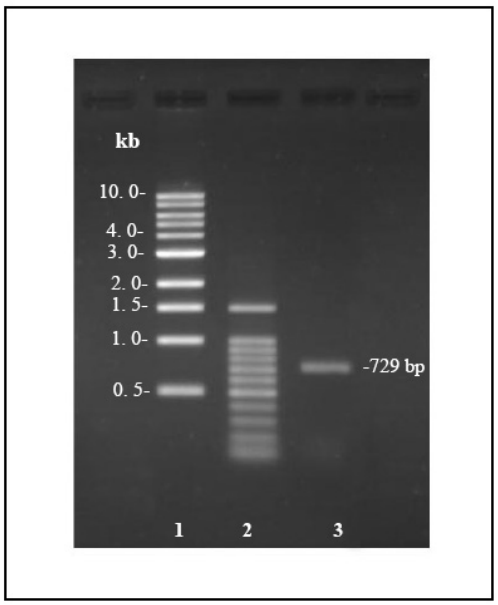


Fig. 42 Amplification of cDNA fragment encoding FmLGBP.

The PCR product was analyzed in 1.2% agarose gel electrophoresis. Lane 1, 1 kb DNA ladder marker; lane 2, 100 bp DNA ladder markers; lane 3, PCR product amplified of cDNA encoding FmLGBP.

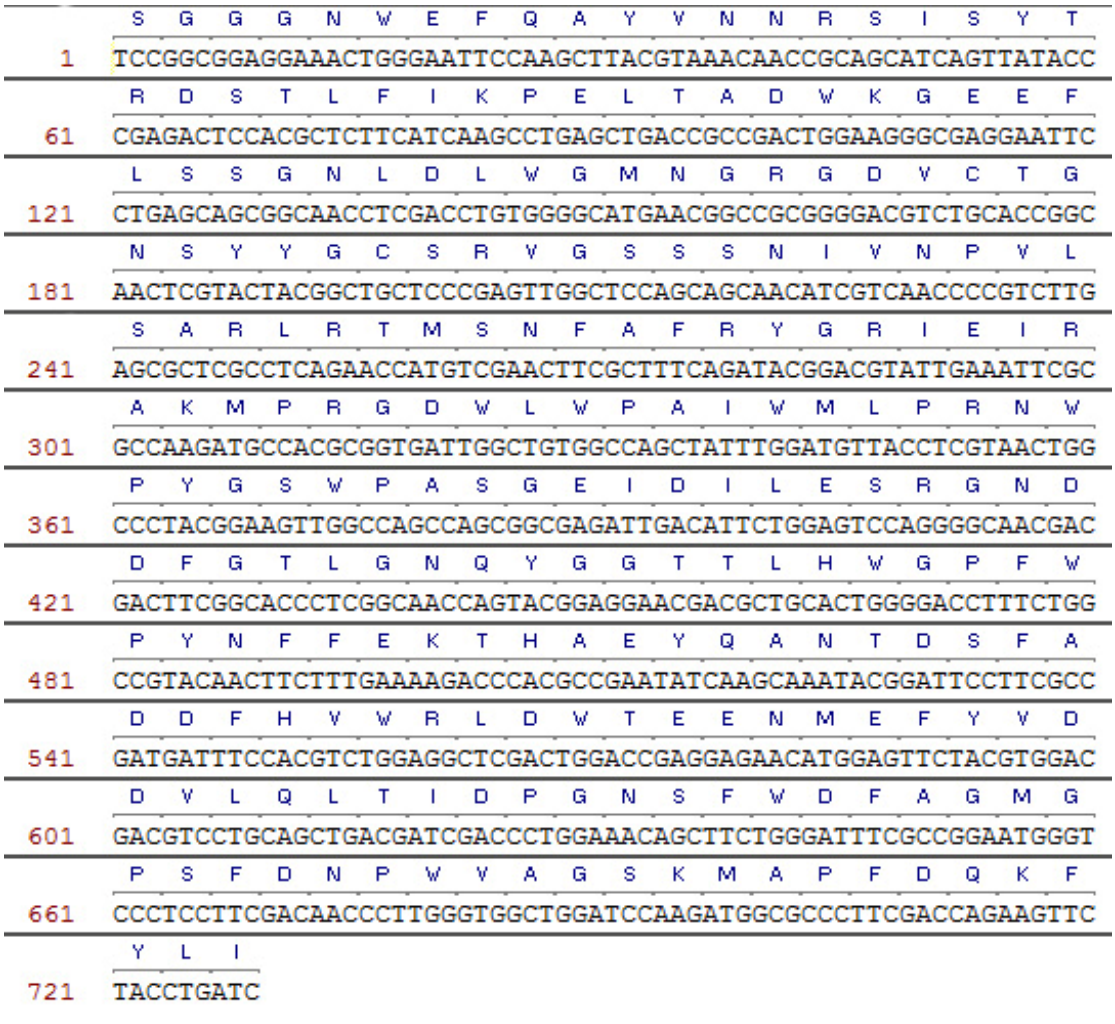


Fig. 43 Nucleotide and deduced amino acid sequences of core *FmLGBP* cDNA.

5.4 Amplification of 5’ and 3’ cDNA ends of *FmLGBP*

In order to determine the remaining DNA sequence of *FmLGBP*, the 5’ and 3’ ends of the cDNA were completed by means of 3’ and 5’ RACE strategies. The 3’ and 5’ end of *FmLGBP* cDNA fragment was produced using the RACE system Gene Racer^{TM} Kit (Invitrogen, USA). The 3’ end of *FmLGBP* cDNA was amplified with FmLGBP-F2 and GeneRacer 3’ primers. An amplified cDNA product of 3’ end (1,094 bp, Fig. 44) was purified from agarose gel, subcloned into pCR4-TOPO vector and then transformed to *E.coli* cells. After plasmid extraction, the purified plasmids were sequenced and analyzed by the Vector NTI software (version 9.0). The nucleotide and deduce amino acid sequences of the 3’ end fragment showed identical matching with overlapping sequence of the core fragment and showed a high degree of similarity to these of 3’ ends of other shrimp C-type lectins. The 3’UTR displayed a poly (A) tail at 3’ end, confirming that we had obtained the sequences in their integrity at this end. In the similar manner, The 5’ end of *FmLGBP* cDNA was amplified with FmLGBP-R2 and GeneRacer 5’ primers in the first round of PCR and FmLGBP-R2 and GeneRacer 5’ primers nested in the second round of PCR. The 5’ RACE PCRs yielded PCR product of 915 bp in length (Fig. 45). An amplified cDNA product of 5’ end was purified from agarose gel, subcloned into pCR4-TOPO vector and then transformed to *E.coli* cells. After plasmid extraction, the purified plasmids were sequenced and analyzed by Vector NTI software (version 9.0). The nucleotide and deduce amino acid sequence of the 5’ end fragment showed identical matching in overlapping sequence to the core fragment and showed a high degree of similarity to those of 5’ ends of other shrimp LGBPs. The identity of the clones was confirmed by matching the sequences to the nucleotide/protein sequences available in GenBank (http://www.ncbi.nlm.nih.gov/, accessed 9/04/07). After sequence analysis, PCR products obtained from 5’ RACE amplification showed a high sequence homology to other shrimp LGBPs.

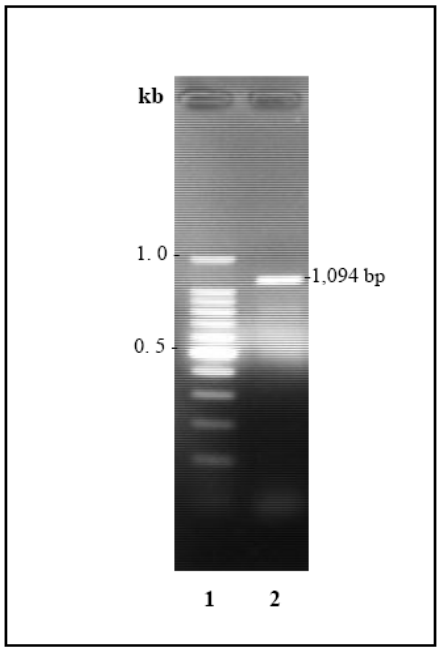


Fig. 44 Electrophoretic pattern of 3' RACE cDNA of *FmLGBP*.

The PCR product was analyzed in 1.5 % agarose gel electrophoresis. Lane 1, 100 bp DNA ladder markers; lane 2, DNA fragment of 3' RACE.

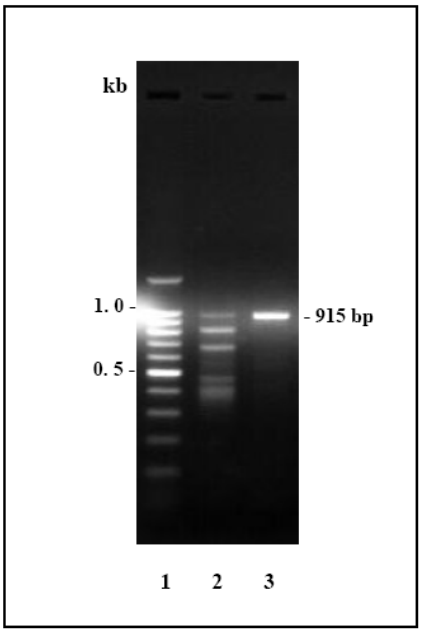


Fig. 45 Electrophoretic pattern of 5' RACE cDNA of *FmLGBP*.

The PCR product was analyzed in 1.5 % agarose gel electrophoresis. Lane 1, 100 bp DNA ladder markers; lane 2, DNA fragments of 5' RACE by using 5’ RACE primer as forward primer; lane 3, DNA fragment of 5' RACE by using 5’RACE nested primer as forward primer.

5.5 Amplification of open reading frame (ORF) of *FmLGBP*

According to the nucleotide sequence of the 5' and 3' ends, the specific primers were designed (Table 7). A pair of primers was used to amplify the *FmLGBP* DNA fragment covering ORF was obtained by PCR with 5' and 3' flanking primers. After amplification, the resulting nucleic acid product was run on a 1.2% agarose gel. From agarose gel electrophoretic result (Fig. 46), the experimental product size appeared related to the expected size. The PCR product that amplified by these specific primers was purified and ligated into pCR4-TOPO vector. The recombinant plasmid was isolated from the white colonies. After that positive clones were determined for the inserted *FmLGBP* appearance with restriction enzymatic digestion. The plasmids were extracted, purified plasmid were then sequenced and analyzed by Vector NTI software (version 9.0).

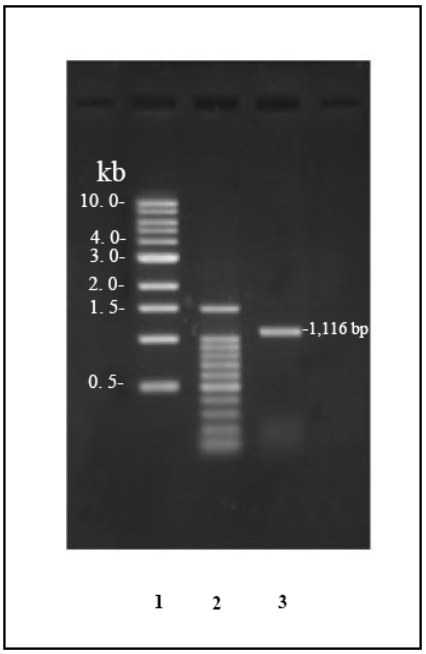


Fig. 46 Electrophoretic pattern of *FmLGBP* open reading frame.

The PCR product was analyzed in 1.2 % agarose gelelectrophoresis. Lane 1, 1 kb DNA ladder markers; lane 2, 100 bp DNA ladder markers; lane 3, PCR product of *FmLGBP* open reading frame

5.6 Sequence analysis of *FmLGBP* cDNA

The whole cDNA of the LGBP was covered by joining these three fragments using Contig Express, part of the Vector NTI software package (version 9). The merged sequence of the cDNA fragment and deduced amino acid sequence are shown in Fig. 47. The 1,280 bp cDNA contains a 1,098 bp ORF with a 5' untranslated region of 30 nucleotides. The termination codon is followed by a long 3' untranslated region of 152 nucleotides that includes a polyadenylation signal (AATAAA) at position 12 nucleotides upstream from poly (A^+) tail. The nucleotide sequence and the deduced amino acid sequence are shown in Fig. 47. The ORF encodes a 366 amino acid protein with a 17 amino acid signal peptide in the N-terminal region. The cleavage site of the 17-residue signal peptide sequence was predicted by the Signal P version 3.0 (Bendtsen *et al.,* 2004) to residue between Ala^{17} and Ala^{18}. The mature LGBP of *F. merguiensis*, therefore, consisted of 349 amino acid residues with a calculated molecular mass of 41.60 kDa and a predicted isoelectic point of 4.43. Two putative glycosylation sites (Asn-Xaa-Ser/Thr) for *N*-linked carbohydrate chains are present in the mature protein sequence at Asn^{49} and Asn^{301}, suggesting that it is a glycoprotein. Two putative cell adhesion sites (integrin binding site; RGD) are located in the N-terminus at Arg^{89} to Asp^{91} and Arg^{140} to Asp^{142} (Fig. 47). A kinase C phosphorylation site (KCPS) at positions Ser^{116} to Arg^{118} and 4 putative amino acid residues at positions Trp^{160}, Glu^{165} , Ile^{166}, and Asp^{167} which are homologous to bacterial glucanase motif (GM) were observed in the *FmLGBP* sequence. The motif scan analysis showed that the amino acid region from Gly^{79} to Phe^{291} belongs to the glycoside hydrolase family 16 (http://expasy.org/tools/#translate, accessed 9/04/07). Additionally, the recognition motif for β-1,3-linkage of the polysaccharide was observed at positions Phe^{235} to Asp^{253}.

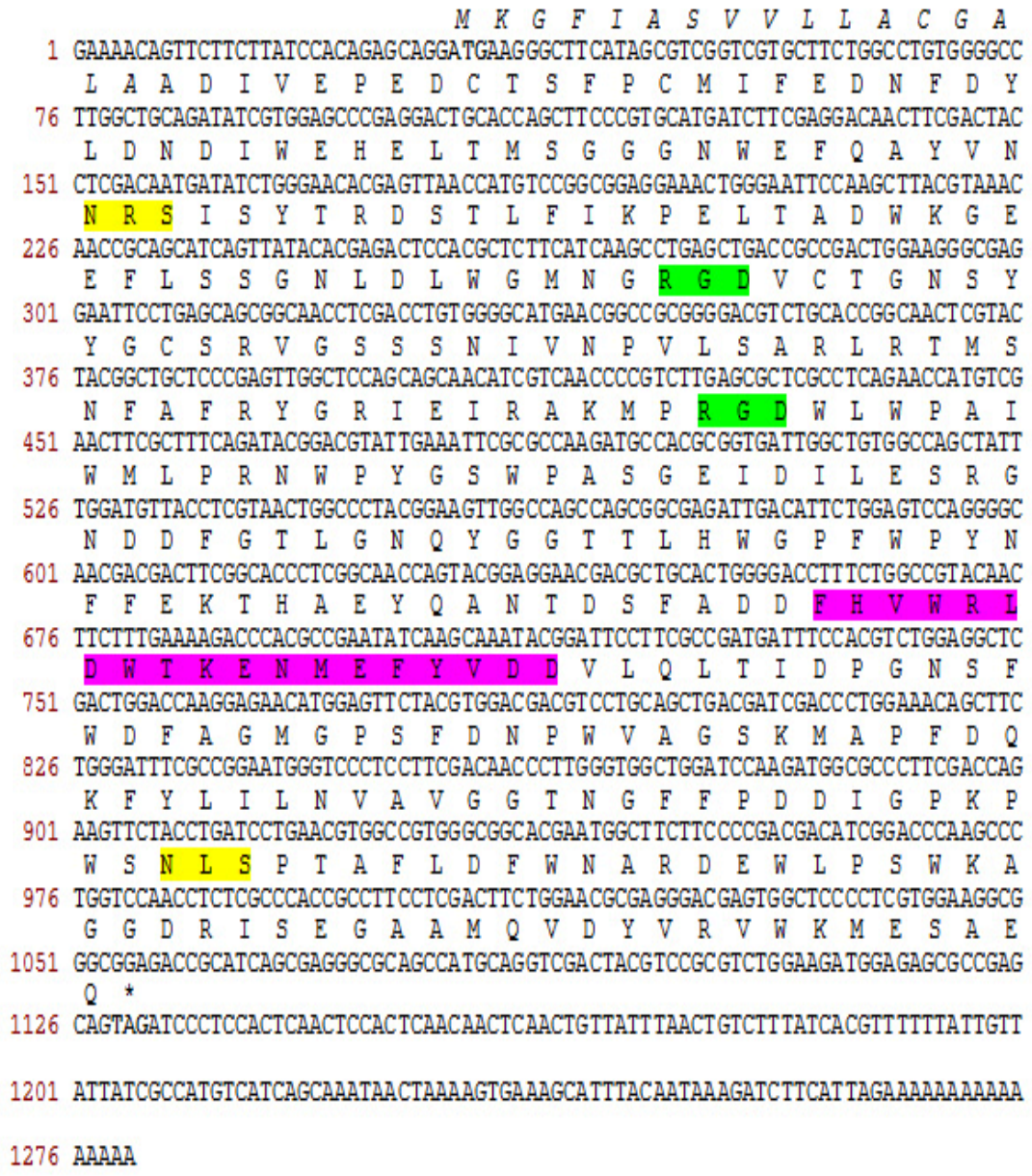


Fig. 47 Nucleotide sequence of cDNA encoding LGBP from *F. merguiensis* and its deduced amino acid sequence.

The sequence of the signal peptide is shown in italics at the N-terminus. Potential *N*-glycosylation sites are yellow highlighted. The pink highlighted amino acid sequence represents a potential recognition motif for β-1,3-linkage of polysaccharides. Two RGD (Arg-Gly-Asp) putative cell adhesive sites (integrin-binding motifs) are green highlighted. The stop codon is marked with asterisk.

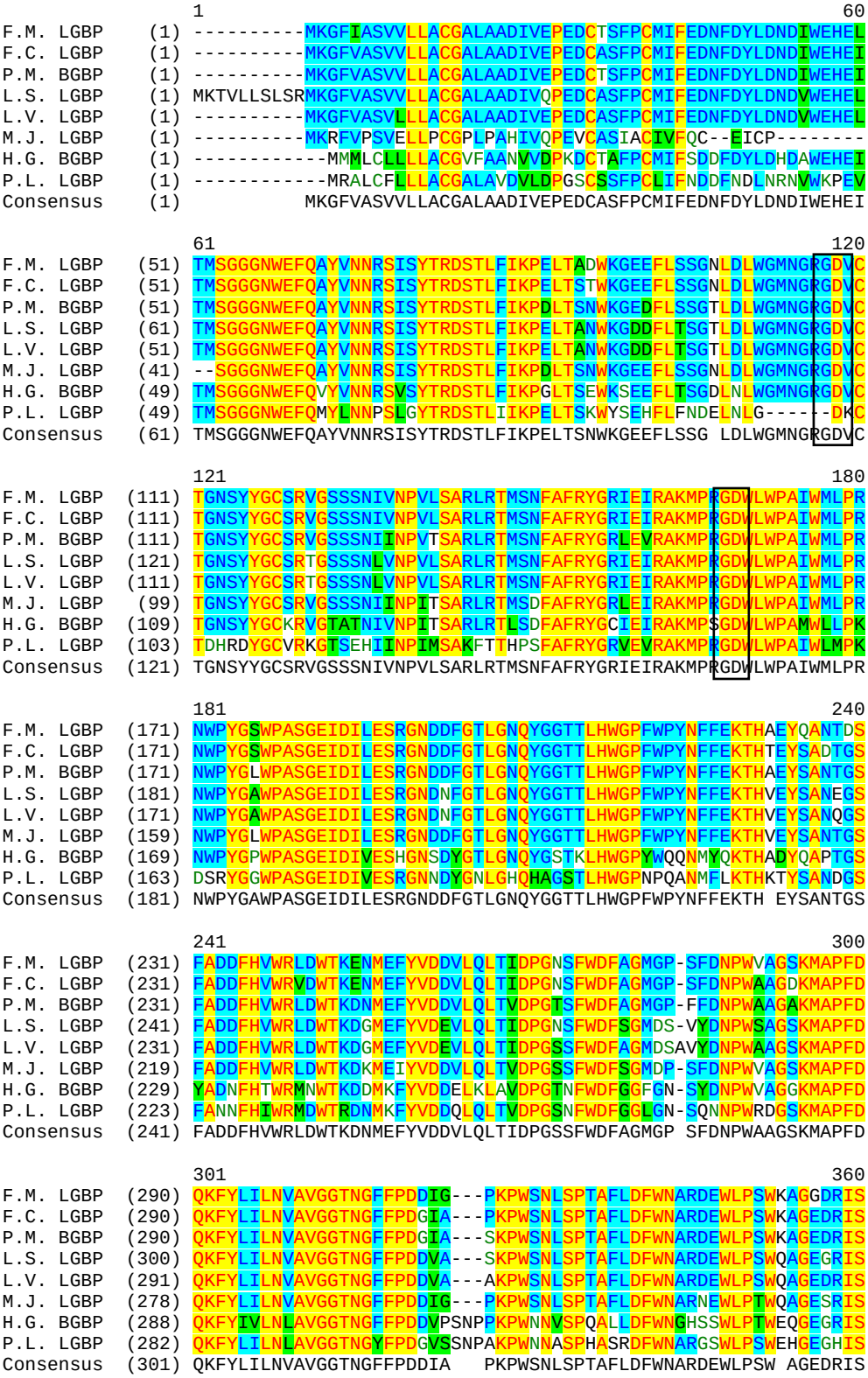
5.7 LGBP sequence comparisons and phylogenetic analysis

In order to determine the degree of homology, the nucleotide sequence of *F. merguiensis* LGBP was compared with all available sequences in the NCBI database using BLAST analysis. The pairwise alignment of *F. merguiensis* LGBP sequence to those of LGBP or BGBP (β-glucan binding protein) from *F. chinensis*, *P. monodon*, *L. vannamei*, *L. stylirostris*, *M. japonicas* show a high identity of 94, 93, 91, 91, and 90%, respectively.

Multiple alignment of *F. merguiensis* LGBP with LGBP of different crustaceans showed conservation of motifs, including two putative integrin binding motifs (RGD) which are located in the N-terminus (residues 89-91 and 140-142) [Ruoslahti, 1996]. A comparison of the deduced amino acid sequence of *F. merguiensis* LGBP with those of other crustacean LGBPs or BGBPs, showed a relatively high degree of overall homology. The amino acid sequence of *F. merguiensis* LGBP has highest homology (95% identity) with LGBP from *F. chinensis*. Similarity to BGBP from *P. monodon* and LGBP from *L. vannamei* were found to be higher than 90%. The homology of *F. merguiensis* LGBP to *M. japonicas* BGBP, *H. gammarus* BGBP and *P. leniusculus* LGBP was comparatively lower but was significant (84%, 73% and 63% identity, respectively). The alignment of the deduced amino acid sequences of these LGBPs revealed a similarity among these 8 species along the entire sequence, including the signal peptide and N-terminal region (Fig. 48). This high homology of *F. merguiensis* LGBP sequence to pattern recognition proteins (PRPs) of other decapod crustacean indicates an important role in recognition of pathogen (Fig. 48).

A phylogenetic analysis was also performed by using the multiple alignment amino acid sequences from LGBP or other BGBP and gram-negative binding protein (GNBP) from crustacean and insects by using the phylogenetic program, MEGA (version 4.1) (Kumar *et al.,* 2004). A phylogenetic tree was divided into two groups. The first group was comprised of three subgroups: (1) LGBP or BGBP from shrimp (*F. merguiensis*, *F. chinesis, L. stylirostris*, *P. monodon*, *M. japonicus*), and Common lobster *H. gammarus*; (2) GNBP from malaria mosquito (*Anopheles gambiae*) and purple sea urchin (*Strongylocentrotus purpuratus*), and LGBP from Scallop (*Chlamys farreri*); (3) LGBP from crayfish (*Pacifastacus*

leniusculus). The second group also contained silkworm BGBP (*B. mori*) and fruit fly GNBP (*Drosophila melanogaster*). Crustacean LGBPs or BGBPs, including that of banana shrimp, were included in subgroup 1.


1 60
F.M. LGBP (1) ----------MKGFIASVVLLACGALAADIVEPEDCTSFPCMIFEDNFDYLDNDIWEHEL
F.C. LGBP (1) ----------MKGFVASVVLLACGALAADIVEPEDCASFPCMIFEDNFDYLDNDIWEHEI
P.M. BGBP (1) ----------MKGFVASVVLLACGALAADIVEPEDCTSFPCMIFEDNFDYLDNDIWEHEI
L.S. LGBP (1) MKTVLLSLSRMKGFVASVVLLACGALAADIVQPEDCASFPCMIFEDNFDYLDNDVWEHEL
L.V. LGBP (1) ----------MKGFVASVLLLACGALAADIVEPEDCASFPCMIFEDNFDYLDNDVWEHEL
M.J. LGBP (1) ----------MKRFVPSVELLPCGPLPAHIVQPEVCASIACIVFQC--EICP--------
H.G. BGBP (1) ------------MMMLCLLLLACGVFAANVVDPKDCTAFPCMIFSDDFDYLDHDAWEHEI
P.L. LGBP (1) ------------MRALCFLLLACGALAVDVLDPGSCSSFPCLIFNDDFNDLNRNVWKPEV
Consensus (1) MKGFVASVVLLACGALAADIVEPEDCASFPCMIFEDNFDYLDNDIWEHEI
61 120
F.M. LGBP (51) TMSGGGNWEFQAYVNNRSISYTRDSTLFIKPELTADWKGEEFLSSGNLDLWGMNGRGDVC
F.C. LGBP (51) TMSGGGNWEFQAYVNNRSISYTRDSTLFIKPELTSTWKGEEFLSSGNLDLWGMNGRGDVC
P.M. BGBP (51) TMSGGGNWEFQAYVNNRSISYTRDSTLFIKPDLTSNWKGEDFLSSGTLDLWGMNGRGDVC
L.S. LGBP (61) TMSGGGNWEFQAYVNNRSISYTRDSTLFIKPELTANWKGDDFLTSGTLDLWGMNGRGDVC
L.V. LGBP (51) TMSGGGNWEFQAYVNNRSISYTRDSTLFIKPELTANWKGDDFLTSGTLDLWGMNGRGDVC
M.J. LGBP (41) --SGGGNWEFQAYVNNRSISYTRDSTLFIKPDLTSNWKGEEFLSSGNLDLWGMNGRGDVC
H.G. BGBP (49) TMSGGGNWEFQVYVNNRSVSYTRDSTLFIKPGLTSEWKSEEFLTSGDLNLWGMNGRGDVC
P.L. LGBP (49) TMSGGGNWEFQMYLNNPSLGYTRDSTLIIKPELTSKWYSEHFLFNDELNLG-----DKC
Consensus (61) TMSGGGNWEFQAYVNNRSISYTRDSTLFIKPELTSNWKGEEFLSSG LDLWGMNGRGDVC
121 180
F.M. LGBP (111) TGNSYYGCSRVGSSSNIVNPVLSARLRTMSNFAFRYGRIEIRAKMPRGDWLWPAIWMLPR
F.C. LGBP (111) TGNSYYGCSRVGSSSNIVNPVLSARLRTMSNFAFRYGRIEIRAKMPRGDWLWPAIWMLPR
P.M. BGBP (111) TGNSYYGCSRVGSSSNIINPVTSARLRTMSNFAFRYGRLEVRAKMPRGDWLWPAIWMLPR
L.S. LGBP (121) TGNSYYGCSRTGSSSNLVNPVLSARLRTMSNFAFRYGRIEIRAKMPRGDWLWPAIWMLPR
L.V. LGBP (111) TGNSYYGCSRTGSSSNLVNPVLSARLRTMSNFAFRYGRIEIRAKMPRGDWLWPAIWMLPR
M.J. LGBP (99) TGNSYYGCSRVGSSSNIINPITSARLRTMSDFAFRYGRLEIRAKMPRGDWLWPAIWMLPR
H.G. BGBP (109) TGNSYYGCKRVGTATNIVNPITSARLRTLSDFAFRYGCIEIRAKMPSGDWLWPAMWLLPK
P.L. LGBP (103) TDHRDYGCVRKGTSEHIINPIMSAKFTTHPSFAFRYGRVEVRAKMPRGDWLWPAIWLMPK
Consensus (121) TGNSYYGCSRVGSSSNIVNPVLSARLRTMSNFAFRYGRIEIRAKMPRGDWLWPAIWMLPR
181 240
F.M. LGBP (171) NWPYGSWPASGEIDILESRGNDDFGTLGNQYGGTTLHWGPFWPYNFFEKTHAEYQANTDS
F.C. LGBP (171) NWPYGSWPASGEIDILESRGNDDFGTLGNQYGGTTLHWGPFWPYNFFEKTHTEYSADTGS
P.M. BGBP (171) NWPYGLWPASGEIDILESRGNDDFGTLGNQYGGTTLHWGPFWPYNFFEKTHAEYSANTGS
L.S. LGBP (181) NWPYGAWPASGEIDILESRGNDNFGTLGNQYGGTTLHWGPFWPYNFFEKTHVEYSANEGS
L.V. LGBP (171) NWPYGAWPASGEIDILESRGNDNFGTLGNQYGGTTLHWGPFWPYNFFEKTHVEYSANQGS
M.J. LGBP (159) NWPYGLWPASGEIDILESRGNDDFGTLGNQYGGTTLHWGPFWPYNFFEKTHVEYSANTGS
H.G. BGBP (169) NWPYGPWPASGEIDIVESHGNSDYGTLGNQYGSTKLHWGPYWQQNMYQKTHADYQAPTGS
P.L. LGBP (163) DSRYGGWPASGEIDIVESRGNNDYGNLGHQHAGSTLHWGPNPQANMFLKTHKTYSANDGS
Consensus (181) NWPYGAWPASGEIDILESRGNDDFGTLGNQYGGTTLHWGPFWPYNFFEKTH EYSANTGS
241 300
F.M. LGBP (231) FADDFHVWRLDWTKENMEFYVDDVLQLTIDPGNSFWDFAGMGP-SFDNPWVAGSKMAPFD
F.C. LGBP (231) FADDFHVWRVDWTKENMEFYVDDVLQLTIDPGNSFWDFAGMGP-SFDNPWAAGDKMAPFD
P.M. BGBP (231) FADDFHVWRLDWTKDNMEFYVDDVLQLTVDPGTSFWDFAGMGP-FFDNPWAAGAKMAPFD
L.S. LGBP (241) FADDFHVWRLDWTKDGMEFYVDEVLQLTIDPGNSFWDFSGMDS-VYDNPWSAGSKMAPFD
L.V. LGBP (231) FADDFHVWRLDWTKDGMEFYVDEVLQLTIDPGSSFWDFAGMDSAVYDNPWAAGSKMAPFD
M.J. LGBP (219) FADDFHVWRLDWTKDKMEIYVDDVLQLTVDPGSSFWDFSGMDP-SFDNPWVAGSKMAPFD
H.G. BGBP (229) YADNFHTWRMNWTKDDMKFYVDDELKLAVDPGTNFWDFGGFGN-SYDNPWVAGGKMAPFD
P.L. LGBP (223) FANNFHIWRMDWTRDNMKFYVDDQLQLTVDPGSNFWDFGGLGN-SQNNPWRDGSKMAPFD
Consensus (241) FADDFHVWRLDWTKDNMEFYVDDVLQLTIDPGSSFWDFAGMGP SFDNPWAAGSKMAPFD
301 360
F.M. LGBP (290) QKFYLILNVAVGGTNGFFPDDIG---PKPWSNLSPTAFLDFWNARDEWLPSWKAGGDRIS
F.C. LGBP (290) QKFYLILNVAVGGTNGFFPDGIA---PKPWSNLSPTAFLDFWNARDEWLPSWKAGEDRIS
P.M. BGBP (290) QKFYLILNVAVGGTNGFFPDGIA---SKPWSNLSPTAFLDFWNARDEWLPSWKAGEDRIS
L.S. LGBP (300) QKFYLILNVAVGGTNGFFPDDVA---SKPWSNLSPTAFLDFWNARDEWLPSWQAGEGRIS
L.V. LGBP (291) QKFYLILNVAVGGTNGFFPDDVA---AKPWSNLSPTAFLDFWNARDEWLPSWQAGEDRIS
M.J. LGBP (278) QKFYLILNVAVGGTNGFFPDDIG---PKPWSNLSPTAFLDFWNARNEWLPTWQAGESRIS
H.G. BGBP (288) QKFYIVLNLAVGGTNGFFPDDVPSNPPKPWNNVSPQALLDFWNGHSSWLPTWEQGEGRIS
P.L. LGBP (282) QKFYLILNLAVGGTNGYFPDGVSSNPAKPWNNASPHASRDFWNARGSWLPSWEHGEGHIS
Consensus (301) QKFYLILNVAVGGTNGFFPDDIA PKPWSNLSPTAFLDFWNARDEWLPSW AGEDRIS

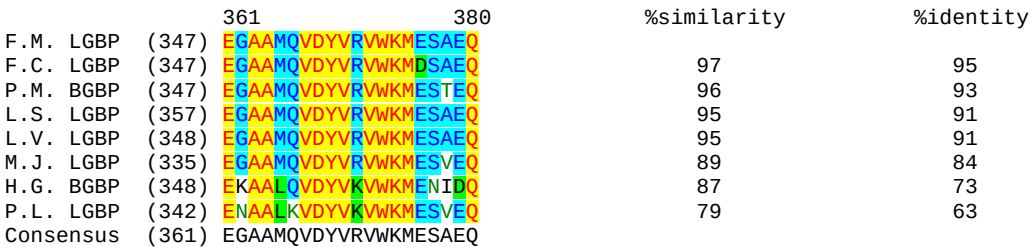


Fig. 48 Multiple alignment of amino acid sequences of *FmLGBP* with those of glucan binding proteins from other crustaceans.

Blank represented by dashes are inserted to maximally align the sequences. The highlighted regions indicate the identical amino acid residues. The identity and similarity of the entire protein in each species to that of *F .merguiensis* are shown at the ends of each LGBP or BGBP sequence. *F. merguiensis* (F.M. this study); *F. chinensis* (F.C. AAX63902); *P. monodon* (P.M. AAM21213); *L. stylirostris* (L.S. AAM73871); *L. vannamei* (L.V. ABU92557); *M. japonicas* (M.J. ABY89089); *H. gammarus* (H.G. CAE47485); *P. leniusculus* (P.L. CAB65353).

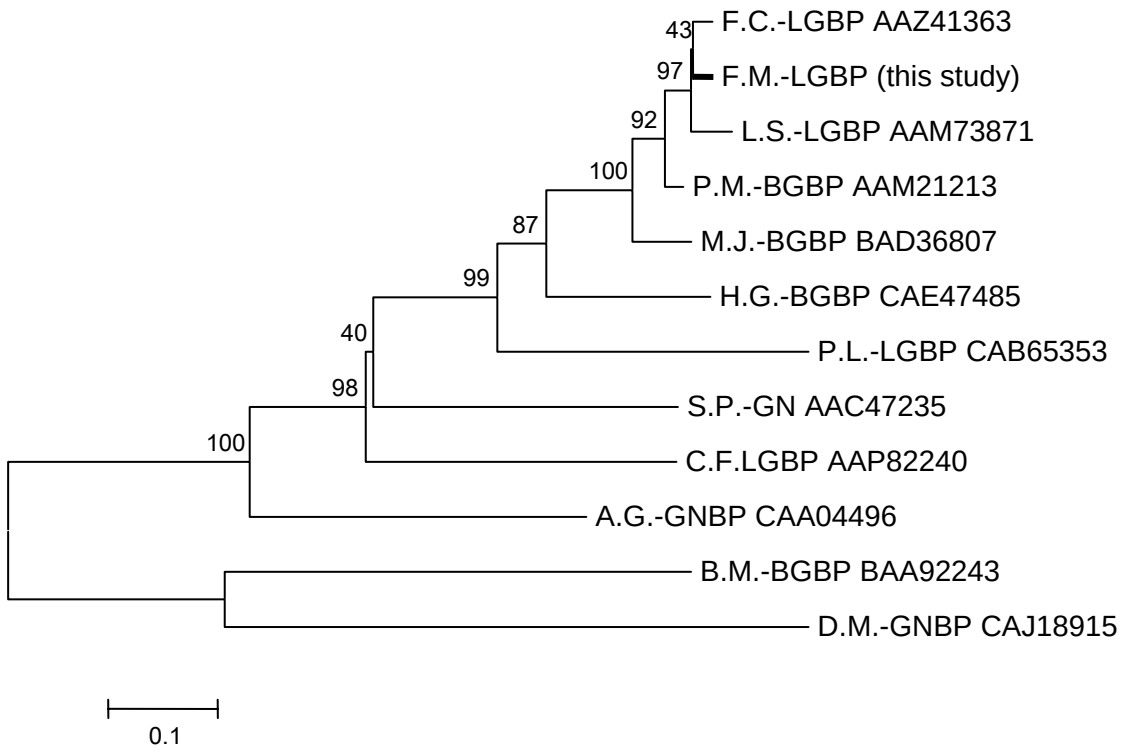


Fig. 49 Phylogenetic analysis of FmLGBP with other GNBP family members.

Neighbour-Joining tree produced by MEGA 4.1. One thousand bootstraps were performed for checking reproducibility of the results. LGBP stands for lipopolysaccharide and β-1,3-glucan binding proteins from *F. chinensis* (F.C. AAZ41363), *L. stylirostris* (L.S. AAM73871), *P. leniusculus* (P.L. CAB65353) and *Chlamys farreri* (C.F. AAP82240). BGBP represents β-1,3-glucan binding proteins from *P. monodon* (P.M. AAM21213), *M. japonicus* (M.J. BAD36807), *H. gammarus* (H.G. CAE47485), and *Bombyx mori* (B.M. BAA92243). GNBP stands for gram-negative binding protein from *Anopheles gambiae* (A.G. CAA04496) and *Drosophila melanogaster* (D.M. Q9NG98). GN stands for β-1,3-glucanase from *Strongylocentrotus purpuratus* (S.P. AAC47235).

6. Expression of FmL and FmLGBP gene in different tissues.

6.1 FmL gene

The site of *FmL* mRNA expression was determined by RT-PCR analysis of multiple tissues. Total RNA samples were extracted from various tissues from *F. merguiensis*. FmL specific primers (FmL-F2 and FmL-R2) were used to amplify *FmL* cDNA fragments. The expected size of the specific *FmL* PCR product is 531 bp. The specific *FmL* PCR fragment was detected only in the hepatopancreas of *F. merguiensis*. Hemocytes, heart, muscle, intestine, ovary, stomach and lymphoild of *F. merguiensis* did not have detectable levels of product (Fig. 50). β-actin was used as an internal control for each tissue sample to confirm that the same amount of total RNAs was used. The expected size of the β-actin PCR product is 500 bp. The specific band of β-actin indicated that roughly the same amount of total RNAs were loaded in the RT-PCR reactions for each tissue sample.

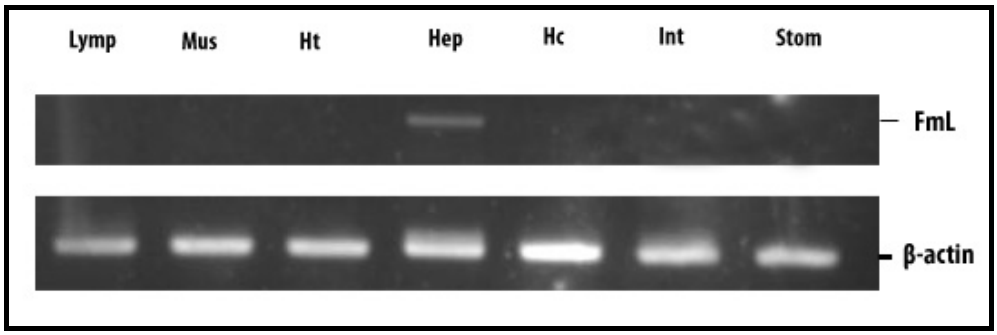


Fig. 50 RT-PCR analysis of the expression of *FmL* mRNA in different tissues.

Total RNA was extracted from the hepatopancreas (Hep), hemocyte (Hc), ovary (Ova), muscle (Mus), heart (Ht), lymphoid organ (Lymp), stomach (Stom) of *F. merguiensis*. The RT-PCR of β-actin transcript was an internal control.

6.2 FmLGBP gene

The site of *FmLGBP* mRNA expression was determined by RT-PCR analysis of multiple tissues. Total RNA samples were extracted from various tissues from *F. merguiensis*. FmLGBP specific primers (FmLGBP-F2 and FmLGBP-R2) were used to amplify *FmLGBP* cDNA fragments. The expected size of the specific *FmLGBP* PCR product is 729 bp. The specific *FmLGBP* PCR fragment was detected only in the hemocyte and hepatopancreas of *F. merguiensis*. Heart, muscle, intestine, ovary, stomach and lymphoild of *F. merguiensis* did not have detectable levels of product (Fig. 51). β-actin was used as an internal control for each tissue sample to confirm that the same amount of total RNAs was used. The expected size of the β-actin PCR product is 500 bp. The specific band of β-actin indicated that roughly the same amount of total RNAs were loaded in the RT-PCR reactions for each tissue sample.

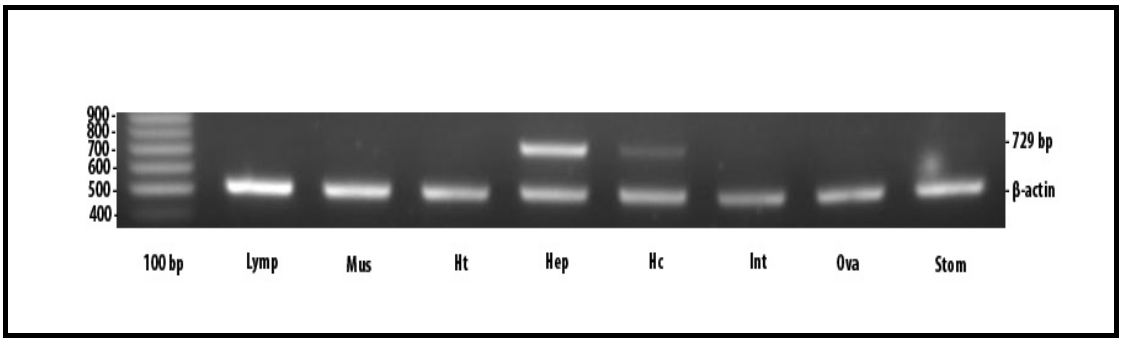


Fig. 51 RT-PCR analysis of the expression of *FmLGBP* mRNA in different tissues.

Total RNA was extracted from the hepatopancreas (Hep), hemocyte (Hc), ovary (Ova), muscle (Mus), heart (Ht), lymphoid organ (Lymp), stomach (Stom) of *F. merguiensis*. The RT-PCR of β-actin transcript was an internal control.

7. Recombinant protein expression of FmL in *E. coli*

7.1 Preparation of *FmL* cDNA fragment

The 5' termainal truncated FmL gene was constructed by PCR amplification using the plasmid containing full-length *FmL* (from method section 12.12) as template and oligonucleotide primers in co-operated with FmL Ex Forward and FmL Ex Reverse primer, which contain *Eco*RI and *Xho*I sites at the 5' and 3' ends. These allowed directional cloning into the expression vector. The resulting PCR products were run on 1.2% agarose gel and specific 1,017 bp fragment of expected size were observed (Fig. 52).

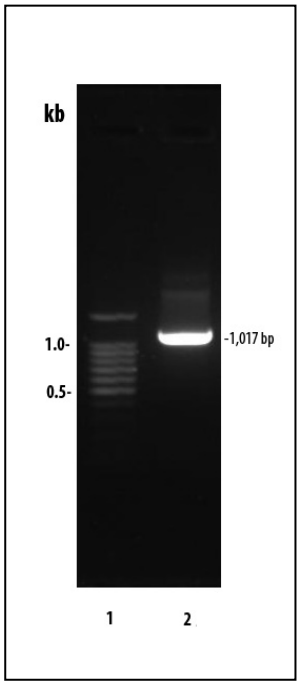


Fig. 52 Agarose gel electrophoresis of the amplified lectin gene.

The PCR product was separated on a 1.2% agarose gel and visualized by ethidium bromide staining. Lane 1, 100 bp DNA ladder markers; lane 2, PCR product.

7.2 Cloning of FmL gene into pDrive cloning vector

PCR products were purified and ligated into pDrive cloning vector. After ligation, the reaction mixture was transformed to *E. coli*. The recombinant clone was first selected with LB agar plate containing amplicillin, X-Gal, and IPTG. White colonies were randomly selected to culture in LB medium containing 50 μg/ml amplicillin. The plasmids were extracted from the selected colonies and cut with *Eco*R I and *Xho* I. The digested plasmids were analyzed in 1.2 % agarose gel electrophoresis shown in Fig. 53. The clone containing the inserted DNA in size of approximately 1.0 kb was selected for recombinant pDrive/FmL augmentation. After digestion and electrophoresis, the 1,017 bp inserted DNA was purified from the agarose gel for further ligation into pET32a(+) expression vector.

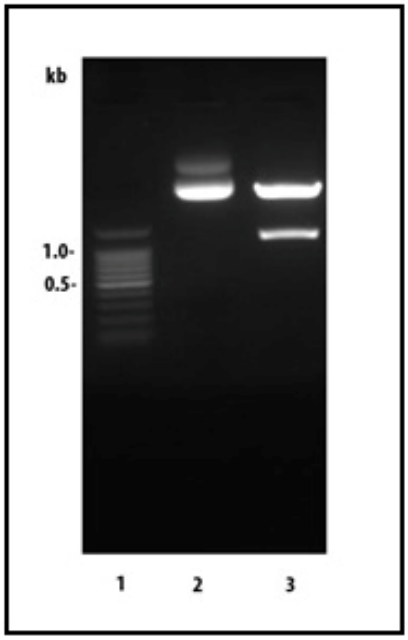


Fig. 53 Ethidium bromide staining of the recombinant pDrive/FmL digested with restriction enzyme *Eco*R I and *Xho* I.

The DNA was run on 1.2% agarose gel. Lane 1, 100 bp DNA ladder markers; lane 2, plasmid pDrive/FmL; lane 3, plasmid pDrive/FmL digested with *EcoR* I and *Xho* I.

7.3 Construction of recombinant plasmid pET32a(+)/FmL

*Eco*R I and *Xho* I digested pET32a(+) vectors were purified from agarose gel, and then insert was excised from the pDrive/FmL using *Eco*R I and *Xho* I. After that it was subsequently cloned into pET32a(+) expression vector at the *Eco*R I and *Xho* I sites. The *FmL* cDNA was subcloned into the pET32a(+) bacterial fusion protein expression system to produce an inframe fusion protein with thioredoxin- and His-tagged fusion protein. After ligation, the reaction mixture was transformed to *E. coli* strain DH5α. The recombinant clone was selected with LB agar plate containing 100 µg/ml amplicillin. The plasmids were extracted from the selected colonies and cut with *Eco*R I and *Xho* I. The digested plasmid was analyzed in 1.2% agarose gel electrophoresis as shown in Fig. 54. The correct construct was name pET32a(+)/FmL.

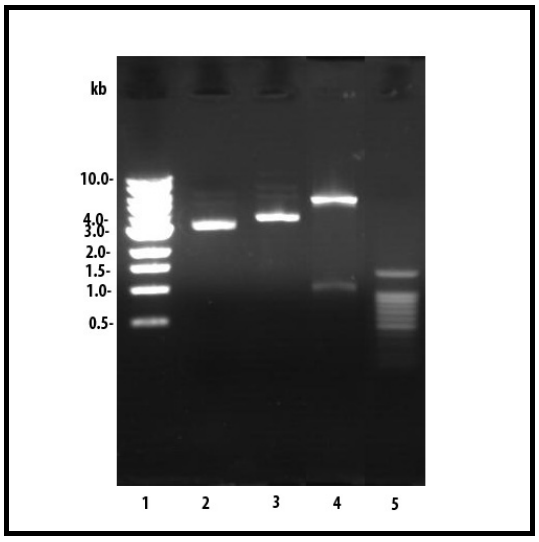


Fig. 54 Ethidium bromide staining of recombinant pET32a(+)/FmL plasmid digested with *Eco*R I and *Xho* I.

The DNA fragment was analyzed on 1.2% agarose gel. Lane 1, 1 kb DNA ladder markers; lane 2, uncut plasmid pET32a(+); lane 3, recombinant pET32a(+)/FmL; lane 4, *EcoR* I and *Xho* I digested recombinant pET32a(+)/FmL; lane 5, 100 bp DNA ladder markers.

7.4 Expression of fusion recombinant FmL in *E. coli*

The *FmL* was cloned into pET32a(+) vector and then expressed as a fusion protein with thioredoxin- and His-tagged fusion in *E. coli* BL21(DE3) star. The lysate of the *E. coli* after induction with 1 mM IPTG was separated in 12% SDS-PAGE. The molecular mass of about 55 kDa of the possible His Tag-FmL corresponds to the predicted molecular mass (Fig. 55).

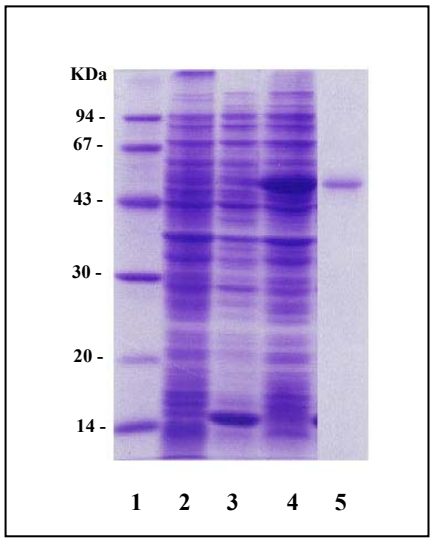


Fig. 55 SDS-PAGE analysis of recombinant expressed FmL in *E. coli*.

Lane 1, molecular weight markers, lane 2, total proteins of *E. coli* with pET32a(+)/FmL without induction; lane 3, soluble proteins; lane 4, insoluble proteins of *E. coli* with pET32a(+)/FmL inducted with IPTG; lane 5, purified recombinant FmL by His-bound resin chromatography.

Chapter 4

Discussions

1. Purification and characterization of lectin

Invertebrate animals lack an adaptive immune system; in these organisms lectins seem to play a relevant role in innate immunity (Vazquez *et al*., 1997; Iwanaga *et al*., 2002). The presence of naturally occurring lectins in the haemolymph of several crustaceans has been well known since the beginning of the 20th century. Most current research has emphasized the role of lectins as non-self recognition molecules in invertebrate immunity. The role of crustacean lectins in immune response has been reviewed by Marques and Barracco (2000). In addition, the role of lectins as pattern recognition proteins (PRPs) is now attracting a significant amount of attention (McGreal *et al.,* 2004; Holmskov *et al.,* 2003). *F. merguiensis* lectin (FmL) has been previously reported as an oligomeric protein of 316.2 kDa made up from two types of subunits of 32.3 and 30.9 kDa. Its agglutination pattern in the plasma with various untreated vertebrate erythrocytes demonstrated the specificity for rabbit and mouse erythrocytes, as well as some gram-negative bacterial strains, i.e. *V. harveyii*, *V. parahemolyticus* and *V. vulnificus*. However, when it was tested against human erythrocytes, it exhibited no specificity among blood groups; A, B, AB and O (Rittidach *et al*., 2007). By means of the hemagglutination inhibition assays, purified lectin has the highest specificity for NeuNAc. Three other N-acetyl aminosugars; GalNAc, GlcNAc and ManNAc, could inhibit FmL HA but with a lower efficiency. In this study, a lectin from the hemolymph of the banana shrimp *F. merguiensis* was achieved using repetitive Fetuin-agarose affinity column, it was purified up to 2,890-fold to apparent homogeneity. From 5 ml of hemolymph, 81 μg of pure lectin was obtained which corresponded to a yield of 27.5%. The fraction eluted from first Fetuin-agarose column gave rise two protein bands of around 300 and 200 kDa in non-denaturing PAGE. The latter, probably corresponds to a contaminated hemocyanin which is a copper-containing protein involving in the

transportation of oxygen and it is present at high concentration in crustaceans. After separating by second Fetuin- agarose column, purified lectin showed only one band in non-denaturing PAGE with an estimated M_r of 316.2 kDa corresponding to that reported by Rittidach *et al.* (2007). Analysis of the lectin in SDS-PAGE showed two bands of apparent molecular masses of 30.9 and 32.3 kDa, without any covalent disulfide linkages between the subunits, but a higher amount of the 30.9 kDa protein band was observed. These results agreed with those reported by Rittidach *et al*. (2007) on lectin isolated from the hemolymph of *F. merguiensis* which were 30.9 and 32.3 kDa. Lectins with various molecular masses and subunit organizations have been isolated from different species of penaeid shrimps. Most of the purified lectins are multimeric proteins, containing unique subunits ranging from 25 to 80 kDa (Marques and Barracco, 2000). For instance, the lectin purified from the hemolymph of *F. chinesis* is a 168 kDa protein, containing subunits of 34 and 40 kDa (Sun *et al*., 2008); that of *L. vannamei* (LvL) is 173 kDa, containing subunits of 32 and 38 kDa (Sun *et al*., 2007); that of *L. setiferus* (LsL) is 291 kDa, containing subunits of 52 and 80 kDa (Alpuche *et al*., 2005); whereas, the lectin isolated from *L. schmitti* is a 153 kDa protein, composed of 31 and 34 kDa subunits (Cominetti *et al*., 2002). Our results suggest that FmL is an apparently multimeric protein, consisting of different subunits of 30.9 and 32.3 kDa. Lectins from crustaceans, such as *P. japonicus* (Yang *et al.,* 2007), the horseshoe crab *T. tridentatus* (Inamori *et al*., 1999), or from the American lobster *H. americanus* (Hall and Rowlads, 1974), are also formed from different amounts of monomers. These results suggested that molecular masses, numbers and sizes of subunits of lectins in crustaceans are quite different, although they have similar sugar-binding preferences. The sizes of native and subunit forms of purified lectin comparing with other lectins from shrimps and crabs were shown in Table 15.

Table 15 Characteristics of lectins from the hemolymph of shrimps and some crabs

Species	Molecular mass (kDa)		References
	Intact	Subunit	
Fenneropenaeus merguiensis	316.2	30.9, 32.3	Rittidach *et al*., 2007; This study
Fenneropenaeus chinesis	168	34, 40	Sun *et al*., 2008
Penaeus japonicus	452	37	Yang *et al*., 2007
Litopenaeus vannamei	173	32, 38	Sun *et al*., 2007
Litopenaeus setiferus	291	52, 80	Alpuche *et al*., 2005
Litopenaeus schmitti	153	31, 34	Cominetti *et al*., 2002
Penaeus californiensis	175	41	Marques and Barracco, 2000
Penaeus monodon	420	27	Ratanapo and Chulavatnatol, 1990
Macrobrachium rosenbergii	19	9.6	Vazquez *et al*., 1993
Paratelphusa jacquemontii		34	Denis *et al*., 2003
Scylla serrata	55	30, 25	Mercy and Ravindranath, 1994

2. N-terminal amino acid sequence analysis of purified lectin subunits

The N-terminal amino acid sequences of purified lectin subunits from *F. merguiensis* were determined by automated Edman degradation. N-terminal amino acid sequencing of both lectin subunits yielded 10 amino acid residuals for each as shown in Table 16. The comparison of N-terminal amino acid sequence between 30.9 and 32.3 kDa subunits of purified FmL showed that they were identical as Leu-Ala-Gly-Ala-Gly-Thr-Val-Leu-Ala-Gly. Accordingly, the molecular mass difference between both subunits may cause by glycosylation heterogenity corresponding to previous study by Rittidach (2006) which found that after treatment by either PNGase F or trifluoromethanesulfonic acid to remove carbohydrate moieties, purified FmL showed only single 28 kDa band in SDS-PAGE. Lectins have been purified and characterized in several species of penaeid shrimp, but a few numbers of N-terminal amino acid sequences of these shrimp purified lectin were determined. Pereyra *et al.*

(2004) successfully determined N-terminal amino acid sequence of the main 81.4 kDa band identified in MrL-I (DVPLL/AXKQQQD) and 9.6 kDa MrL-III (DVPLL/A) isolated from freshwater prawn *M. rosenbergii*. Purified lectin from *L. setiferus* showed N-terminal amino acid sequence of 80 kDa subunit as DASNAQKQHDVNFLL but that of 52 kDa could not be determined since it was blocked (Alpuche *et al.,* 2005), whereas N-terminal amino acid sequences of 34 and 40 kDa subunits of purified *F. merguiensis* lectin were LAGXGTVLAGFPALG and TXERTDTVEPQQPGG, respectively (Sun *et al.,* 2008). By comparing N-terminal amino acid sequence of our purified lectin subunits with those available in the published paper, we found that those of 30.9 and 32.3 kDa subunits show only a high degree of similarity to 34 kDa subunit of *F. chinesis* lectin, i.e. Leu-Ala-Gly-Ala-Gly-Thr-Val-Leu-Ala-Gly-Phe-Pro-Ala-Leu-Gly. In contrast, no homology was found with N-terminal amino acid sequence of 40 kDa subunit of *F. chinesis* lectin (Table 16).

Table 16 The N-terminal amino acid sequences of purified lectin subunits.

Species	Subunits	N-terminal amino acid sequence	References
F. merguiensis	30.9 kDa 32.3 kDa	LAGAGTVLAG LAGAGTVLAG	This study This study
F. chinesis	34 kDa 40 kDa	LAGXGTVLAGFPALG TXERTDTVEPQQPGG	Sun *et al.,* 2008
L. setiferus	52 kDa 80 kDa	Blocked DASNAQKQHDVNFLL	Alpuche *et al.,* 2005
M. rosenbergii	9.6 kDa 81.4 kDa	DVPLL/A DVPLL/AXKQQQD	Pereyra *et al.,* 2004

3. Peptide analysis by MALDI-TOF-TOF MS

Mass spectrometry (MS) is an important and emerging method for the identification of gel separated proteins and makes use of the rapidly growing sequence databases. Generally, there are two MS methods that can be used to identify proteins: matrix-assisted laser desorption ionization-time of flight (MALDI-TOF) and tandem mass spectrometry (MS/MS). Both methods work best when trypstic digestion is used to break up each protein into multiple peptide fragments. Both types of spectra display fingerprints unique to specific proteins and peptide sequences. In this study, the good MS spectra were obtained for the two 30.9 and 32.3 kDa subunits of purified FmL but no matches were found by PMF/MS-MS analysis. The PMFs of the both subunits are almost identical. We can therefore conclude that FmL is a unique protein that has not been described and is not present in the database. However, the amino acid sequences obtained after tryptic treatment by MALDI-TOF of FmL were used as query sequence. The Blast search and alignment revealed that the predicted amino acid sequence of FmL fragment 3 (F3) is similar to conserved motifs present in fibrinogen-related domains (FReDs) of ficolin protein from human (GenBank # NP_001994), of fibrinogen-related proteins, tachylectins 5A and 5B from the horseshoe crab *T. tridentatus* (GenBank # BAA84188 and BAA84190) (Gokudan *et al*., 1999), and Dorin M from *Ornithodoros moubata* (GenBank #AAP93589) (Rego *et al*., 2006). Alignment of FmL F3 with those sequences as above mentioned is shown in Fig. 56.

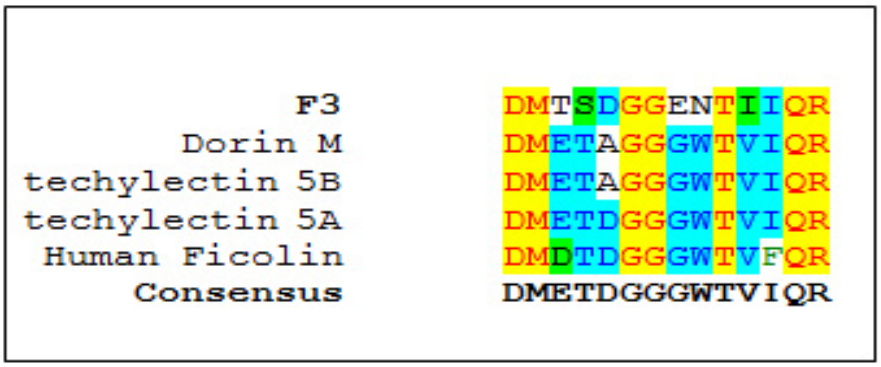


Fig. 56 Multiple alignments of predicted amino acid sequence of *F. merguiensis* lectin fragment 3 (F3) with partial sequences of fibrinogen-related domains (FReDs) of ficolin proteins.

Human ficolin from human (GenBank # NP_001994), tachylectins 5A and 5B from horseshoe crab *T. tridentatus* (GenBank # BAA84188 and BAA84190) and Dorin M from *Ornithodoros moubata* (GenBank #AAP93589). Identical amino acids are shaded in yellow. Other conserved, but not consensus amino acids, are shaded in blue and green.

Moreover, the predicted amino acid sequence of FmL fragment 4 (F4) is also similar to partial sequences of fibrinogen-related domains (FReDs) of two ficolin proteins from human (GenBank # NP_001994 and AAB50706), ficolin B from cattle *Bos taurus* (GenBank # AAW52550) (unpublished) and wild pig *Sus scrofa* (GenBank #AAC69641) (Ichijo *et al*., 1993). Alignment of F4 with these sequences as above mentioned is shown in Fig. 57. Based on these results, it indicated that in FmL molecule may contain fibrinogen-related domain. Likely, lectin from mosquito *Armigeres subalbatus* (aslectin) is fibrinogen-related protein (FREP) because its molecule contains a fibrinogen-like domain (Wang *et al*., 2004). Furthermore, many lectins from invertebrates had also a fibrinogen-like domain, for instance, tachylectins 5A and 5B from the horseshoe crab *T. tridentatus* (Gokudan *et al*., 1999). On the other hand, the predicted amino acid sequence of other FmL fragments (F1, F2, and F5) showed no similarity to any protein in the NCBI database.

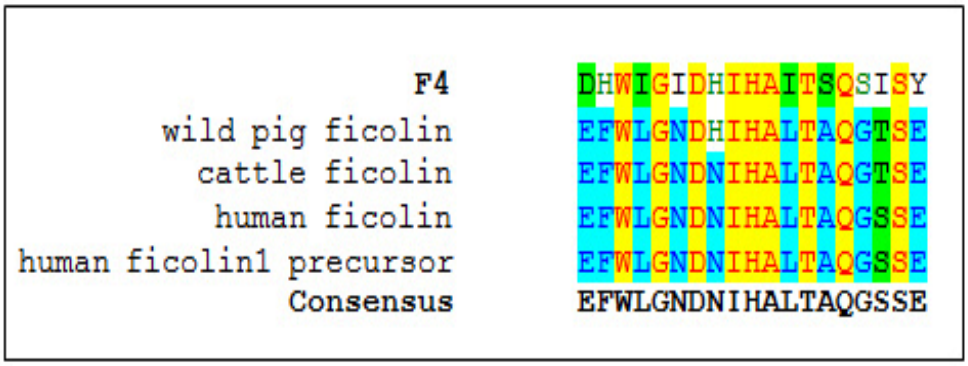


Fig. 57 Multiple alignments of predicted amino acid sequence of *F. merguiensis* lectin fragment 4 (F4) with partial sequences of fibrinogen-related domains (FReDs) of ficolin proteins.

Human ficolin1 precursor from human (GenBank # NP_001994), human ficolin from human (GenBank #AAB50706), cattle ficolin from cattle *Bos taurus* (GenBank #AAW52550) and wild pig ficolin from wild pig *Sus scrofa* (GenBank #AAC69641). Identical amino acids are shaded in yellow. Other conserved, but not consensus amino acids, are shaded in blue and green.

4. Molecular cloning and sequence analysis of cDNA encoding FmL

4.1 Cloning of FmL gene in *F. merguiensis*

PCR amplification of *FmL* cDNA was done with degenerate primers designed based on the highly conserved sequences of various shrimp lectins. Then, the missing sequences of *FmL* at 5' and 3' ends were identified by RACE method, and full-length ORF of FmL gene was re-amplified by DNA polymerase with proof reading ability to confirm the integrity and identity of the sequences. The nucleotide and deduced amino acid sequences of *FmL* cDNA were shown in Fig. 36. The full-length *FmL* cDNA composed of 1,118 bp with a 42 bp of 5' untranslated region and a 74 bp 3' of untranslated region. It contained an ORF of 1,002 bp encoding a putative polypeptide with 333 amino acids and with a predicted molecular mass of 36.3 kDa and pI of 4.14. The lengths of the nucleotide and amino acid sequences of *FmL* cDNA are agreed well with the previously reported sizes of other shrimp lectins that contained dual CRDs as shown in Table 17. The deduced amino acid sequence of FmL presented a common signal peptide of 20 residues. The cleavage of the signal sequence was predicted to occur between Gly^{20} and Arg^{21} residues (Bendtsen *et al*., 2004). C-type lectins have a characteristic carbohydrate recognition domain (CRD) with a well-defined structure stabilized by two or three pairs of disulfide bonds, which can bind to carbohydrate residues of foreign pathogens (Sun *et al*., 2008). The deduced amino acid sequence of FmL exhibited the typical structure of C-type lectin protein. Multiple alignments between FmL and other shrimp C-type lectins show the presence of several highly conserved regions, i.e. CRDs, QPD motif and EPN motif (Fig. 38).

Table 17 Properties of the predicted shrimp C-type lectins.

	ORF (bp)	Encoding Residues	M_r (kDa)	pI	References
F. merguiensis lectin	1,002	333	36.3	4.14	This study
P. monodon lectin	1,002	333	36.4	4.16	Ma *et al*., 2008
L. vannamei lectin	1,038	345	37.2	4.55	Ma *et al*., 2007
P. semisulcatus lectin	1,062	353	38.3	4.6	GenBank #ABI97372

The primary protein structures of four shrimp C-type lectins so far identified are aligned in Fig. 38. Consistent with other known C-type lectins, four residues of cysteine, which are important in formation of two disulfide bonds dictating the secondary/tertiary structure of protein, are strictly conserved in both CRDs (positions Cys^{36}, Cys^{47}, Cys^{65} and Cys^{144} in CRD1 and Cys^{197}, Cys^{208}, Cys^{226} and Cys^{307} in CRD2), In addition, another two cysteine residues were located at the N-terminus of each CRD (positions Cys^{160} and Cys^{168} in CRD1; and Cys^{321} and Cys^{329} in CRD2), indicating that both CRDs were of long-form. On the basis of known carbohydrate-binding specificity, the carbohydrate binding sites of C-type lectins could be classified as Gal-type or Man-type according to the presence of the sequence motif. It has been believed that C-type lectins contain a peptide segment that corresponds to the carbohydrate-binding specificity within CRD, and that the sequence of Gal- and GalNAc-binding, is Gln-Pro-Asp (QPD) motif. In contrast, the sequence in mannose, fucose, and GalNAc-specific lectins, such as serum mannose-binding proteins, was Glu-Pro-Asn (EPN) motif (Tsuiji *et al*., 2002). In this study, deduced amino acid sequence was analyzed by SMART (http://smart.embl-heidelberg.de/, accessed 9/04/07) to determine the motifs of FmL. Analysis of the deduced amino acid sequence for FmL indicated that it was a member of the C-type lectin family. Prediction of domain structures using the SMART system revealed that FmL is dual CRD C-type lectin which each CRD composed of 134 amino acid residues. Thus, it may bind to microbial cell surface carbohydrates through these domains. This structural organization of FmL is similar to the LPS-binding lectins (BmMBP) from *B. mori* (Watanabe *et al*., 2006); immulectin-2 of insect *Manduca sexta* (Yu and Kanost, 2000) and fall webworm *Hyphantria cunea* (Shin *et al*., 1998; 2000). In addition, this type of structure has also been found in various shrimp C-type lectins (LvL, PmL, PsL and FcLec) (Ma *et al*., 2007; 2008; Genbank # ABI97372; Luo *et al*., 2006), in which each consists of 2 CRDs that function as a pattern-recognition protein (PRP) in innate immunity. *M. sexta* immulectin-2 (IML-2) functions as a PRP in innate immunity. IML-2 caused aggregation of only gram-negative bacteria (Yu *et al*, 1999; Yu and Kanost, 2000). IML-2 was also found to bind granular cells and oenocytoids, as well as nematodes (Yu and Kanost, 2004). More recently IML-2 was shown to bind to a wide range of microbial patterns,

including lipoteichoic acid, laminarin (branched β-1, 3-glucan), mannose, and all three moieties of LPS (lipid A, core carbohydrate, and *O*-specific antigen) (Yu and Ma, 2006). Moreover, similar two tandem-domain C-type lectins, are found in lepidopterans and involved in bacteria and hemocyte aggregation (Koizumi *et al*, 1999, Shin *et al*, 2000). They have been suggested to have PRP functions. The number of CRD identified in each species may be variable. However, it has been reported that the binding of carbohydrates is weak for lectins those contain one CRD whereas lectins with multiple CRDs have a higher affinity (Zelensky and Gready, 2005). The existence of two CRDs may extend the protection of shrimp against a variety of microorganisms or facilitate the interaction of immunocytes with pathogens and subsequent induction of cellular defense. Further analysis of FmL CRDs revealed that a QPD motif, which is important for high-affinity galactose binding, is present in first CRD. Moreover, an EPN motif, which may involve in the binding capacity of mannose type carbohydrate (Weis *et al*, 1998), can be identified in the second CRD of FmL. Other penaeid shrimp lectins with one CRD (PmAV, PmLec and Fc-hsL) have either an EPN (PmAV and Fc-hsL) or QPD motif (PmLec). An incomplete WND motif was also found in the second CRD of FmL near the C-terminal region. This motif is exceedingly important in the binding of both carbohydrate and Ca^{2+} and highly conserved in vertebrate lectins (Iwanag and Lee, 2005). The evolutionary distinction of the invertebrates is highlighted by the lack of WND motif or their mutation to a similar motif (Ma *et al.,* 2008). For instance, all the vertebrates contain a WND (Trp-Asn-Asp) motif in the CRD but the insect and shrimp lectins compose of either FRD (Phe-Arg-Asp) or VND (Val-Asn-Asp) in their CRDs. The WND motif is substituted by VND in PmLec (Luo *et al*., 2007); FRD in PmL (Ma *et al*., 2008), LvL (Ma *et al.,* 2007), and PsL (GenBank # ABI97372) (Fig. 38). Likewise to FmL, the WND motif of IML-2 is replaced by LDD (Leu-Asp-Asp) (Yu *et al.,* 2002; 2005; Yu and Kanost, 2004).

4.2 Comparison of the deduced amino acid sequence of FmL

The overall homology of FmL with other vertebrate lectins was low, accepted for a few residues were highly conserved in CRDs (data not shown). For instance, FmL shared only 30-50% overall amino acid similarity with those of lectins from human, fish and planarian (worm) (Fig. 40) but six cysteine residues, which are

important in formation of three disulfide bonds, are completely conserved in each. Using the Vector NTI software (version 9.0, Invitrogen), comparison of the entire deduced amino acid sequences of FmL with those of other penaeid lectins showed that this sequence exhibits relatively high homology to other shrimp lectins, with the highest sequence identity to *P. monodon* (82% identity) (Table 18).

Table 18 Percent identity of the deduced amino acid sequences among FmL and other crustacean lectins.

	F.M Lectin	P.M. Lectin	L.V. Lectin	P.S. Lectin
F.M Lectin	100	**82**	65	68
P.M. Lectin		100	68	74
L.V. Lectin			100	68
P.S. Lectin				100

Pair-wise comparison was carried out using the Vector NTI software (version 9.0, Invitrogen). F.M., *F. merguiensis*; P.M., *P. monodon*; L.V., *L. vannamei*; P.S., *P. semisulcatus*.

A comparison of the deduced amino acid sequences of FmL along with those of other shrimps demonstrated the existence of two CRDs (CRD1 and CRD2) (Fig. 38). Based on the result of multiple amino acid sequence alignment, the two highly conserved motifs: the QPD and EPN motifs, both of which are characteristic of C-type lectins, are presented in the FmL protein. These motifs are considered as signature domains for protein belonging to C-type lectin family. The deduced amino acid sequence of FmL showed 81% identity and 89% similarity to PmL; 64% identity and 73% similarity to PsL; 62% identity and 75% similarity to LvL. Percentages of identity and similarity were also assessed between different conserved domains, i.e. CRD1 and CRD2 (Fig. 39). The result showed that FmL CRD1 had 83% identity and 90% similarity to PmL CRD1; 65% identity and 76% similarity to LvL CRD1; 64%

identity and 72% similarity to PsL CRD1, whereas FmL CRD2 had 80% identity and 88% similarity to PmL CRD2; 66% identity and 76% similarity to LvL CRD2; 78% identity and 86% similarity to PsL CRD2. This result suggested that they may also share similar carbohydrate (polysaccharide) binding functions. However, within the same lectin molecule, the two CRDs share less than 50% identity.

4.3 Phylogenetic analysis of the FmL sequence

To study the evolutionary relationship of FmL CRDs, a phylogenetic tree was constructed based on the amino acid sequence of CRDs of FmL and those of known C-type lectins from vertebrates and invertebrates. In the phylogenetic tree, CRD1 containing a QPD motif of FmL was clustered and had high similarity with CRD1 of those from PmL, LvL and PsL, indicating that all the shrimp CRDs1 were closely related to each other (Fig. 40). Furthermore, CRD2 containing an EPN motif of FmL was clustered and had high similarity with CRDs2 of those from PmL, LvL and PsL, indicating that all the shrimp CRDs2 were closely related to each other. The CRDs of C-type lectins from Japanese eel *A. japonica* (GenBank # BAC54021), human *Homo sapiens* (GenBank #AAG13815), worm *G. tigrina* (GenBank #AAL29933) and rainbow trout *O. mykiss* (GenBank # AAM21196) formed a separate cluster (Fig. 40).

4.4 Express of FmL gene in difference tissues

The hemocytes and hepatopancreas are considered as important tissues for immune defense in invertebrates (Sun *et al*., 2008). The circulating hemocytes play defense roles not only by direct sequestration and killing of infectious agents, but also by synthesis and exocytosis of a battery of bioactive molecules in immune defense (Liu *et al*., 2007). The hepatopancreas also contains highly specialized cells and phagocytes that also function in cellular immune responses (Gross *et al*., 2001). Several studies have been conducted to locate site of gene expression for lectin in shrimp. Some C-type lectins of invertebrates have been reported that were exclusively expressed in hemocytes, such as Fclectin from *F. chinensis* (Liu *et al*., 2006), PmAV and PmLec from *P. monodon* (Luo *et al*., 2006; 2007). Expression of *P. monodon* lectin in the hemocyte suggests its probable role to act as PRP of the innate immune system of shrimps. However, in other reports of shrimp C-type lectins, the hepatopancreas is the sole tissue of this gene expression in normal bodies, such as that

of LvL from *L. vannamei* (Ma *et al*., 2007) and Fc-hsL from *F. chinensis* (Sun *et al*., 2008). Moreove, in *L. vannamei*, *L. setiferus* and *L. stylirostris*, lectins form the largest group of immune-related cDNAs found in the hepatopancreas (Gross *et al*., 2001; Astrofsky *et al*., 2002) suggesting the importance of lectins in the initial detection of microorganism infection. It may have some functional implication in the different expressions of these lectins. In the present study, RT-PCR results showed that *FmL* mRNA was expressed only in the hepatopancreas of *F. merguiensis*, and not in other tissues (Fig. 50) indicating that FmL may have probable role to act as PRP of the innate immune system in this species of shrimp.

4.5 Expression of recombinant FmL protein

In order to express recombinant FmL protein in *E.coli,* the result showed that this protein was successfully expressed using the *E. coli* system, but the protein appears to be in inclusion bodies. To verify whether the recombinant FmL protein was expressed, the pET32a(+)/FmL transformant was cultured and induced with IPTG. The total cell pellet was then solubilized in SDS-PAGE sample buffer and subjected to 12% SDS-PAGE. An approximated 55 kDa FmL, in accordance with expected molecular mass, was detected after 3 h of induction in the cell lysate of *E.coli* containing pET32a(+)/FmL but absent in that containing the parental plasmid pET32a(+). At 3 h after induction, the intensity of 55 kDa protein band was enough for further purification. The bacterial expression system is particularly convenient for very large quantities of protein production under well-defined conditions. Unfortunately, expressed proteins by bacteria are often difficult to be purified because of their tendency to precipitate within the cells. The precipitated protein formed inclusion bodies, the dense, granular structures distributed throughout the cytoplasm (Marton, 1986; Krueger *et al*., 1989; Hockney, 1995). The formation of inclusion bodies is especially common for non-bacterial proteins. The recombinant FmL is no exception. The approximately 55 kDa protein was observed mainly in the insoluble fractions (Fig.55) which probably accounted for more than 50% of the insoluble proteins. The 55 kDa FmL protein was thus expressed and aggregated in the cells as inclusion bodies.

5. Molecular cloning and sequence analysis of cDNA encoding FmLGBP

A number of invertebrate PRPs (pattern recognition proteins) have been isolated and characterized. These PRPs, according to the molecules they recognize, are grouped as β-1,3-glucan binding proteins (BGBP); lipopolysaccharide (LPS)-binding proteins (LBP); LPS- and β-1,3-glucan-binding protein (LGBP) as well as peptidoglycan-binding proteins (PGBP). PRPs such as LGBP play an important role in activating the innate immune response of crustaceans and insects. LGBP as a vital component of the immune system in shrimp has been purified and cloned from several crustaceans. In crustacean, β-1, 3-glucan binding proteins (BGBPs) with 95-112 kDa have been isolated including several species of penaeids: yellowleg shrimp *F. californiensis* (Vargas-Albores *et al*, 1996; Yepiz-Plascencia *et al*., 1998), blue shrimp *L. stylirostris* (Roux *et al*., 2002), and white shrimp *L. vannamei* (Vargas-Albores *et al*, 1997; Yepiz-Plascencia *et al*., 1998). A LPS-binding protein with 175 kDa has been isolated from *F. californiensis* (Vargas-Albores *et al*, 1993). An LGBP with 40 kDa isolated from hemocytes of crayfish *P. leniusculus*, has binding activity to LPS as well as β-1,3-glucans such as curdlan and laminarin, but not to peptidoglycan (Lee *et al*., 2000). A BGBP has been cloned from crayfish *P. leniusculus* (Cerenius *et al*., 1994) and tiger shrimp *P. monodon* (Sritunyalucksana *et al*., 2002). An LGBP has been cloned from crayfish *P. leniusculus* (Lee *et al*., 2000) and blue shrimp *L. stylirostris* (Roux *et al*., 2002). Reported isolations and identifications of pattern recognition proteins including BGBP, LBP, and LGBP and their clones are shown in Table 5.

5.1 Cloning of FmLGBP gene in *F. merguiensis*

PCR amplification of *FmLGBP* cDNA was done with degenerate primers designed based on the highly conserved sequences of various shrimp LGBPs. Then, the missing sequences of *FmLGBP* at 5' and 3' ends were identified by RACE method, and full-length ORF of *FmLGBP* was re-amplified by DNA polymerase with proof reading ability to confirm the integrity and identity of the sequences. The nucleotide and deduced amino acid sequences of *FmLGBP* cDNA are shown in Fig. 47. *FmLGBP* composed of 1,280 bp with a 30 bp 5' untranslated region and a 152 bp 3' untranslated region. *FmLGBP* cDNA contained an ORF of 1,101 bp encoding a putative polypeptide with 366 amino acids and with a predicted molecular mass of

41.6 kDa and pI of 4.43. The lengths of nucleotide and amino acid sequences of *FmLGBP* cDNA are in good agreement with previously reported sizes of other shrimp LGBPs as shown in Table 18. The deduced amino acid sequence of FmLGBP presented a common signal peptide of 17 residues. The cleavage of the signal sequence was predicted to occur between Ala^{17} and Ala^{18} residues (Bendtsen *et al*., 2004). Two putative *N*-linked glycosylation sites, NRS and NLS, for *N*-linked carbohydrate chains are present in the mature protein at Asn^{66} and Asn^{318}, suggesting that *F. merguiensis* LGBP is a glycoprotein. In addition, two *N*-linked glycosylation sites (front NRS and rear NLS) were observed in other crustacean LGBPs and BGBPs, such as LGBP from fleshy prawn *F. chinensis* (Du *et al*., 2007), white shrimp *L. vannamei* (Cheng *et al*., 2005) as well as BGBP from tiger shrimp *P. monodon* (Sritunyalucksana *et al*., 2002). Characteristic potential polysaccharide binding, cell adhesion, and glucanase motifs, similar to invertebrate PRP motifs were also found in FmLGBP. A conserved potential recognition motif for polysaccharide β-1,3-linkage was observed in FmLGBP at the positions Phe^{235} to Asp^{253}. The amino acid sequence of this motif was conserved in *F.chinesis* (Du *et al*., 2007) as well as BGBP from tiger shrimp *P. monodon* (Sritunyalucksana *et al*., 2002). All four amino acid residues (positions Trp^{177}, Glu^{182}, Ile^{183}, and Asp^{184}) which are homologous to bacterial glucanase motif (GM), and considered to be necessary for the catalytic mechanism of bacterial glucanase are conserved in FmLGBP (this study). A similar glucanase motif has also been found in the LGBP and BGBP of crayfish *P. leniusculus* (Trp^{169}, Glu^{174}, Ile^{175} and Asp^{176}) (Lee *et al*., 2000; Cerenius *et al*., 1994), LGBP of blue shrimp *L. stylirostris* (Trp^{187}, Glu^{192}, Ile^{193} and Asp^{194}) (Roux *et al*., 2002), white shrimp *L. vannamei* (Trp^{177}, Glu^{182}, Ile^{183} and Asp^{184}) (Cheng *et al*., 2005), fleshy prawn *F.chinesis* (Trp^{177}, Glu^{182}, Ile^{183} and Asp^{184}) (Du *et al*., 2007) and LGBP or GNBP of earthworm *Eisenia foetida* (Bilej *et al*., 2001; Beschin *et al*., 1998; Lee *et al*., 1996) as well as BGBP from *M. sexta* (Jiang *et al.*, 2004). Although, until now, no glucanase activity have been identified in these proteins. It is hypothesized that LGBP genes evolved from a β-1,3-glucanase that lost its glucanase activity during evolution but maintained its glucan binding properties and therefore played a role in innate immunity (Lee *et al.*,2000). Another conserved motif, RGD (Arg-Gly-Asp), is the putative cell adhesion site (Johansson, 1999). RGD motif also presents in peroxinectin

molecule, whose adhesive function is likely to be mediated by the integrin-binding motif, KGD (Lys-Gly-Asp) or RGD (Arg-Gly-Asp) motifs (Ruoslahti, 1996). Holmblad *et al*. (1997) reported the presence of an integrin β-subunit on surfaces of the crayfish haemocytes. So it can be speculated that LGBP, a receptor protein present in plasma, could bind to haemocyte membrane through RGD site and induce a series of immune reactions. Not all LGBPs have this character. For example scallop *Chlamys farreri* LGBP, which was found to have no RGD motif, but a potential transmembrane domain was found (Su *et al*., 2004).

Table 19 Properties of the predicted shrimp LGBPs.

	ORF (bp)	Encoding Residues	M_r (kDa)	pI	References
F. merguiensis LGBP	1,101	366	41.6	4.43	**This study**
M. japonicas LGBP	1,065	354	40.15	4.47	Lin *et al*., 2008
F. chinesis LGBP	1,101	366	41.44	4.47	Du *et al*., 2007
L. vannamei LGBP	1,104	367	41.55	4.41	Cheng *et al*., 2005
P. stylirostris LGBP	1,131	376	42.61	4.52	Roux *et al*., 2002
P. leniusculus LGBP	1,086	361	41.04	5.94	Lee *et al*., 2000

5.2 Comparison and phylogenetic analysis of the deduced FmLGBP amino acid sequence

BLAST analysis demonstrated that LGBP isolated from banana shrimp had similarity to glucanase from bacteria, sea urchin and scallop (data not shown). Moreover, sequence data suggested that FmLGBP had also high similarity with other members of the glucan binding proteins in crustaceans and gram-negative binding protein (GNBP) family in mollusks, echinoderms and insects. The significant homology noted between different categorized groups clearly suggests that PRPs are conserved proteins in invertebrates, playing similar roles in different animals. Searching for sequence similarities of FmLGBP with known proteins revealed close matches with highly conserved amino acid residues commonly shared by LGBP of *F.chinensis* (Du *et al*., 2007). Using the Vector NTI software (version 9.0, Invitrogen), comparison of the deduced amino acid sequences of FmLGBP with those of other LGBPs showed that this sequence exhibits relatively high homology to other shrimp LGBPs, with the highest sequence identity to *F. chinensis* (96% identity) (Du *et al*., 2007) (Table 20). Based on the result of multiple amino acid sequence alignment, the highly conserved RGD motif which is the putative cell adhesion site, is present in FmLGBP protein. FmLGBP was found to be very similar to other LGBP genes. Conserved domains such as the LPS-binding site and the glucan-binding site were found in FmLGBP gene. These results indicated that FmLGBP was involved in the binding of both LPS and β-1,3-glucan. Phylogenetic analysis of the difference PRPs of 13 species revealed that β-1,3-glucanase of *S. purpuratus* was grouped with LGBP of *C. farreri* and GNBP of *A. gambiae*, which belonged to the first group which included FmLGBP. FmLGBP was grouped with *F. chinensis* LGBP, *L. stylirostris* LGBP, *P. monodon* BGBP, *M. japonicus* BGBP, *H. gammarus* with a bootstrap value of 87 and was closer to *F. chinensis* LGBP with a bootstrap value 43 (Fig. 49).

Table 20 Percent identity of the deduced amino acid sequence among FmLGBP and other crustacean LGBPs.

	F.M. LGBP	F. C. LGBP	H.G. BGBP	P.L. LGBP	M.J. LGBP	L.S. LGBP	L.V. LGBP	P.M. BGBP
F.M. LGBP	100	**96**	72	63	87	91	92	93
F. C. LGBP		100	72	64	87	91	92	94
H.G. BGBP			100	66	70	71	71	72
P.L. LGBP				100	63	63	64	65
M.J. LGBP					100	86	86	88
L.S. LGBP						100	98	91
L.V. LGBP							100	91
P.M. BGBP								100

Pair-wise comparison was carried out using the Vector NTI software (version 9.0, Invitrogen). F.M., *F. merguiensis*; F.C., *F. chinensis*; H.G., *H. gammarus*; P.L., *P. leniusculus*; M.J., *M. Japonicas*; L.S., *L. stylirostris*; L.V., *L. vannamei*; P.M., *P. monodon*.

5.3 Express of FmLGBP gene in difference tissues

In crustaceans, circulating haemocytes play an important role in the immune response against pathogens. Circulating haemocytes may synthesize various defense molecules, such as proPO, peroxinectin and LGBP (Lee *et al*., 2000). These defense molecules are secreted and activated by microbial cell wall components like lipopolysaccharides (LPS), β-1,3-glucans (BG) or peptidoglycans (PG). In crustaceans, previous studies have demonstrated that hepatopancreas is a key tissue involved in the immune response, and it is a major site for the synthesis of immune defense molecules and is involved in eliminating the pathogens or other particulate matters (Johnson, 1987; Gross *et al*., 2001; Vogt, 1996). In the crayfish *P. lenius-culus*, LGBP was detected in hemocytes by Northern blot analysis (Lee *et al*., 2000). In the black tiger shrimp *P. monodon*, BGBP transcripts were also detected mainly in hemocytes when Northern blotting was used (Sritunyalucksana *et al*., 2002). Moreover, in the fleshy prawn *F. chinenesis*, LGBP was detected only in hemocytes by RT-PCR analysis as well. In the blue shrimp *L. stylirostris*, LGBP gene was cloned from hepatopancreas cDNA library (Roux *et al*., 2002). On the other hand, in the Pacific white shrimp *L. vannamei*, RT-PCR analysis showed that LGBP gene were detected in both hemocyte and hepatopancreas (Cheng *et al*., 2005). Recently, a *LGBP* cDNA was detected in hemocyte of fleshy prawn *F. chinensis* (Du *et al*., 2007). In present study, RT-PCR was carried out to evaluate tissue specific mRNA expression in various tissues. The result showed that 729 bp *FmLGBP* cDNA fragment was detected in the haemocyte and hepatopancreas, suggesting that LGBP of *F. merguiensis* is synthesised in both haemocytes and hepatopancreas. However, the expression of *FmLGBP* in hemocytes was lower than in hepatopancreas (Fig. 51).

Chapter 5

Conclusions

1. Lectin from hemolymph of *F. merguiensis* was purified by repetitive affinity chromatography on Fetuin-agarose column. The corresponding protein showed a molecular mass of 316.2 kDa. It was an oligomeric protein made up from two types of distinct subunits of 32.3 and 30.9 kDa, without any covalent disulfide linkages between the subunits.

2. The N-terminal amino acid sequence of 30.9 and 32.3 kDa subunits of purified lectin are identical as LAGAGTVLAG.

3. The internal amino acid sequence of 30.9 and 32.3 kDa subunits of purified lectin were determined by MALDI-TOF-TOF MS and are identical for both subunits.

4. C-type lectin gene (designated FmL) was cloned and characterized from the hepatopacreas cDNA library of the banana shrimp *F. merguiensis*.

5. The *FmL* cDNA sequence preparing from hepatopancreas of *F. merguiensis* was submitted to NCBI and the GenBank accession number is FJ751773.

6. *FmL* cDNA consisted of 1,118 nucleotides including 1,002 bp of a single open reading frame, with a 5' untranslated region of 42 bp and a 3' untranslated region of 74 bp. *FmL* cDNA encoded a polypeptide of 333 amino acids with a putative signal peptide of 20 amino acid residues and a mature protein of 313 amino acids. FmL has an estimated molecular mass of 36.3 kDa and isoelectric point of 4.14.

7. FmL contains two CRDs (CRD1 and CRD2) which share 35% identity with each other. The first CRD contains a QPD motif with specificity for galectose binding, while the second CRD contains an EPN motif for mannose binding.

8. A sequence comparison showed that the deduced amino acids of *F. merguiensis* lectin has an overall similarity of 82%, 68% and 65% to those of tiger shrimp *P. monodon*, green tiger shrimp *P. semisulcatus*, and white shrimp *L. vannamei*, respectively.

9. RT-PCR showed that FmL mRNA was expressed only in hepatopancreas, with no detection in other tissues of the banana shrimp.

10. FmL protein was successfully expressed in the *E. coli* strain BL21(DE3)star expression system. Recombinant FmL protein purified under denaturing conditions using the nickel-based affinity chromatography showed a single band corresponding to the molecular mass of 55 kDa.

11. Lipopolysaccharide and β-1,3-glucan binding protein gene (designated FmLGBP) was cloned and characterized from the hemocyte cDNA library of *F. merguiensis*.

12. The *FmLGBP* cDNA sequence preparing from hemocyte was submitted to NCBI and the GenBank accession number is FJ751774.

13. *FmLGBP* cDNA consisted of 1,280 nucleotides including 1,101 bp of a single open reading frame, with a 5' untranslated region of 30 bp and a 3' untranslated region of 152 bp. *FmLGBP* cDNA encoded a polypeptide of 366 amino acids with a putative signal peptide of 17 amino acid residues.

14. The mature FmLGBP, therefore, consisted of 349 amino acid residues with a calculated molecular mass of 41.6 kDa and a predicted pI of 4.43.

15. The mature FmLGBP protein included two putative glycosylation sites (Asn-Xaa-Ser) for *N*-linked carbohydrate chains and a putative cell adhesion site (integrin-binding site), Arg-Gly-Asp (RGD). Additionally, a glycosyl hydrolase domain was identified.

16. A sequence comparison showed that the deduced amino acids of FmLGBP has an overall similarity of 96%, 93% and 91% to those of fleshy prawn *F. chinensis*, tiger shrimp *P. monodon*, and blue shrimp *L. stylirostris*, respectively.

17. RT-PCR showed that LGBP mRNA was expressed both in hemocyte and hepatopancreas, but not in other tissues of *F. merguiensis*.

References

Alabi, A. O., Latchford, J. W. and Jones, D. A. 2000. Demonstration of residual antibacterial activity in plasma of vaccinated *Penaeus vannamei*. Aquaculture 187: 15-34.

Adams, A. 1991. Response of penaeid shrimp to exposure to *Vibrio* species. Fish Shellfish Immunol. 1: 59-70.

Aderem, A and Ulevitch, R. 2000. Toll-like receptors in the induction of the innate immune response. Nature 406: 782-787.

Agrawal, B.B.L. and Goldstein, I.J. 1972. Concanavalin A, the Jack Bean (*Canavalia ensiformis*) Phytohemagglutinin. In: Methods in Enzymology, Ginsberg, V., Ed. Academic Press, New York., Vol. 28, pp. 314-316.

Alpuche, J., Pereyra, A., Agundis, C., Rosas, C., Pascual, C., Slomianny, M.C., Vazquez, L. and Zenteno, E. 2005. Purification and characterization of a lectin from the white shrimp *Litopenaeus setiferus* (Crustacea decapoda) hemolymph. Biochim. Biophys. Acta 1724: 86–93.

Amirante, G.A. 1986. Agglutinins and Lectins of Crustacea. In: Hemocytic and Humoral Immunity in Arthropods, Gupta, A.P., Ed. Wiley-Interscience, New York., pp. 359-380.

Arason, G.J. 1996. Lectins as defence molecules in vertebrates and invertebrates. Fish Shellfish Immunol. 6: 277-289.

Asgari, S., Theopold, U., Wellby, C. and Schmidt, O. 1998. A protein with protective properties against the cellular defense reactions in insects. Proc. Natl. Acad. Sci. USA. 95: 3690-3695.

Ashida, M. and Brey, P. 1998. Recent Advances in Research on the Insect Prophenoloxidase Cascade. In: Molecular Mechanisms of Immune Responses in Insects, Brey, P.T. and Hultmark, D. Eds. Chapman and Hall, New York., pp. 135-172.

Aspan, A., Hall, M. and Söderhäll, K. 1990. The effect of endogenous proteinase inhibitors on the prophenoloxidase activating enzyme, a serine proteinase from crayfish haemocytes. Insect Biochem. 20: 485-492.

Aspan, A., Huang, T.S., Cerenius, L. and Söderhäll, K. 1995. cDNA cloning of prophenol-oxidase from the freshwater crayfish *Pacifastacus leniusculus* and its activation. Proc. Natl. Acad. Sci. 192: 939-943.

Astrofsky, K.M., Roux, M.M., Klimpel, K.R., Fox, J.G. and Dhar, A.K. 2002. Isolation of differentially expressed genes from white spot virus (WSV) infected Pacific blue shrimp (*Penaeus stylirostris*). Arch. Virol. 147: 1799-1812.

Azumi, K., Yokosawa, H. and Ishii, S.-I. 1991. Lipopolysaccharide induces release of a metallo protease from hemocytes of the ascidian, *Halocynthia roretzi*. Dev. Comp. Immunol.15: 1-7.

Bachere, E., Gueguen, Y., Lorgeril, J. D., Garnier, J. and Romestand, B. 2004. Insights into the anti-microbial defense of marine invertebrates: the penaeid shrimps and the oyster *Crassostrea gigas*. Immunological Reviews 198: 149-168.

Bailey-Brock, J. and Moss, S. 1992. Penaeid Taxonomy, Biology and Zoogeography. In: Marine Shrimp Culture: Principles and Practices, Fast, A. and Lester, J. Eds. Elsevier, Amstedam., pp. 9-28.

Bartlett, T.C., Cuthbertson, B.J., Shepard, E.F., Chapman, R.W., Gross, P.S. and Warr, G.W. 2002. Crustins, homologues of an 11.5-kDa antibacterial peptide, from two species of penaeid shrimp, *Litopenaeus vannamei* and *Litopenaeus setiferus*, Mar. Biotechnol. 4: 278-293.

Bauchau, A.G. 1981. Crustaceans. In: Invertebrate Blood Cell, Ratcliffe, N.A. and Rowley, A.F. Ed. Academic Press, New York., pp. 386-420.

Bilej, M., De Baetselier, P., Dijck, E.V., Stijlemans, B. and Colige, A. 2001. Distinct carbohydrate recognition domains of an invertebrate defense molecule recognize gram-negative and gram-positive bacteria. J. Biol. Chem. 276: 45840-45847.

Bendtsen, J. D., Nielsen, H., Heijne, G. V. and Brunak, S. 2004. Improved prediction of signal peptides-Signal P 3.0. J. Mol. Biol. 340: 783-795.

Begum, N., Matsumoto, M., Tsuji, S., Toyoshima, K. and Seya, T. 2000. The primary host defense system across humans, flies and plants. Curr. Trends Immunol. 3: 59-74.

Beisel, H. G., Kawabata, S., Iwanaga, S., Huber, R. And Bode, W. 1999. Tachylectin-2: crystal structure of a specific GlcNAc/GalNAc-binding lectin involved in the innate immunity host defense of the Japanese horseshoe crab *Tacheypleus tridentatus*. EMBO J. 18: 2313-2322.

Bell, T.A. and Lightner, D.V. 1988. A Handbook of Normal Penaeid Shrimp Histology. The World Aquaculture Society, Baton Rouge., pp. 114.

Beschin, A., Bilej, M., Hanssens, F., Raymakers, J., Van Dyck, E., Revets, H., Brys, L., Gomez, J., De Baetselier, P. and Timmermans, M. 1998. Identification and cloning of a glucan- and lipopolysaccharide-binding protein from *Eisenia foetida* earthworm involved in the activation of prophenoloxidase cascade. J. Biol. Chem. 273: 24948-24954.

Bradford, M.M. 1976. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. Anal. Biochem. 72: 248-254.

Bulet, P., Hetru, C., Dimarcq, J.-L. and Hoffmann, D. 1999. Antimicrobial peptides in insects; structure and function. Dev. Comp. Immunol. 23: 329-344.

Burmester, T. and Scheller, K. 1996. Common origin of arthropod tyrosinase, arthropod hemocyanin, insect hexamerin, and Dipteran *Arylphorin* receptor. J. Mol. Evolution 42: 713-728.

Cerenius, L., Liang, Z., Duvic, B., Keyser, P., Hellman, U., Tapio-Palva, E. *et al*. 1994. Structure and biological activity of a 1,3-β-D-glucanbinding protein in crustacean blood. J. Biol. Chem. 269: 29462-29467.

Cerenius, L. and Söderhäll, K. 2002. Early events in crustacean innate immunity. Fish Shellfish Immunol. 12: 421-437.

Cerenius, L. and Söderhäll, K. 2004. The prophenoloxidase-activating system in invertebrates. Immunol. Rev. 198: 116-126.

Chen, C., Ratcliffe, N.A. and Rowley, A.F. 1993. Detection, isolation and characterization of multiple lectins from the haemolymph of the cockroach, *Blaberus discoidalis*. J. Biochem. 294: 181-190.

Cheng, W., Liu, C. H., Tsai, C. H. and Chen, J. C. 2005. Molecular cloning and characterization of a pattern recognition molecule, lipopolysaccharide- and β-1,3-glucan binding protein (LGBP) from the white shrimp *Litopenaeus vannamei*. Fish Shellfish Immunol. 18: 297-310.

Chattopadhyay, T., Guha, A. K. and Chatterjee, B. P. 1996. Novel antimicrobial activity of scyllin, a haemolymph lectin of the edible crab *Scylla serrata*. Biomed. Lett. 53: 29-40.

Chisholm, J. R. S. and Smith, V. J. 1995. Comparison of antibacterial activity in the haemocytes of different crustacean species. Comp. Biochem. Physiol, 110A: 39-45.

Cho, M.Y., Choi, H.W., Moon, G.Y., Kim, M.H., Kwon, T.H., Homma, K., Natori, S. and Lee, B. L. 1999b. An 86 kDa diapause protein 1-like protein is a component of early staged encapsulation-relating proteins in coleopteran insect, *Tenebrio molitor* larvae. FEBS Lett. 451: 303-307.

Chomczynski, P. and Sacchi, N. 1987. Single step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction. Anal. Biochem. 162: 156-159.

Cohen, E., Ilodi, G.H.U., Korytnyk, W. and Sharma, M. 1983. Inhibition of *Limulus* agglutinins with N-acetylneuraminic acid and derivatives. Dev. Comp. Immunol. 7: 189-192.

Cominetti, M.R., Marques, M.R.F., Lorenzini, D.M., Lofgrena, S.E., Daffrec, S. and Barracco, M.A. 2002. Characterization and partial purification of a lectin from the hemolymph of the white shrimp *Litopenaeus schmitti*. Dev. Comp. Immunol. 26: 715-721.

Cooper, E. L., Kauschke, E. and Cossarizza, A. 2002. Digging for innate immunity since Darwin and Metchnikoff. BioEssays 24: 319-333.

Cuthbertson, B. J., Shepare, E. F., Chapman, R. W. and Gross, P. S. 2002. Diversity of the penaeidin antimicrobial peptides in two shrimp species. Immunogenetics 54: 442-445.

Daquing, A. C., Sato, T., Koda, H., Takao, T., Fukuda, M., Shimonishi, Y. and Tsukamoto, T. 1999. A novel endogenous inhibitor of phenoloxidase from *Musca domestica* has a cystine motif commonly found in snail and spider toxins. Biochemistry 38: 2179-2188.

Davis, B.J. 1964. Disc electrophoresis: II. Method and application to human serum protein. Ann. N.Y. Acad. Sci. 121: 404-427.

Denis, M., Mercy, P. D., Renuka, B. N. and Jeya S. S., 2003. Purification and characterization of a sialic acid specific lectin from the hemolymph of the freshwater crab *Paratelphusa jacquemontii*. Eur. J. Biochem. 270: 4348-4355.

Destoumieux, D., Bulet, P., Loew, D., Alain Dorsselaer, V., Rodriguez, J. and Bachère E. 1997. Penaeidins, a new family of antimicrobial peptides isolated from the shrimp *Penaeus vannamei* (Decapoda). J. Biol. Chem. 272: 28398-28406.

Destoumieux, D., Munoz, M., Bulet, P. and Bachere, E. 2000. Penaeidins, a family of antimicrobial peptides from penaeid shrimp (Crustacea, Decapoda). Cell. Mol. Life Sci. 57: 1260-1271.

Destoumieux, D., Bulet, P., Strub, J.-M. and Bachere, E. 1999. Recombinant expression and range of activity of penaeidins, antimicrobial peptides from penaeid shrimp. Eur. J. Biochem. 266: 335-346.

Dodd, R.B. and Drickamer, K. 2001. Lectins-like protein in model organisms : Implications for evolution of carbohydrate binding activity. Glycobiology 11: 71-79.

Drickamer, K. 1988. Two distinct classes of carbohydrate-recognition domain in animal lectin. J. Biol. Chem. 263, 9557-9560.

Drickamer, K., and Taylor, M.E. 1998. Evolving views of protein glycosylation. Trends Biochem. Sci. 23: 321-324.

Du, X.-J., Zhao, X.-F. and Wang, J.-X. 2007. Molecular cloning and characterization of a lipopolysaccharide and β-1,3-glucan binding protein (LGBP) from fleshy prawn (*Fenneropenaeus chinensis*). Mol. Immunol. 44: 1085-1094.

Duvic, B. and Söderhäll, K. 1990. Purification and characterization of b-1,3-glucan-binding protein from the plasma of the crayfish *Pacifastacus leniusculus*. J. Biol. Chem. 265: 9332-9337.

Duvic, B. and Söderhäll, K. 1993. β-1,3-Glucan-binding proteins from plasma of the freshwater crayfishes *Astacus astacus* and *Procambarus clarkii*. J. Crusta. Biol. 13: 403-408.

Edman, P. and Begg, G. 1967. Automated equipment for sequence determinations. Eur. J. Biochem. 1: 80-91.

Elgavish, S. and Shaanan, B. 1997. Lectin-carbohydrate Interactions: Different Folds, Common Recognition Principles. Elsevier Science Ltd., The Hebrew University of Jerusalem, Jerusalem, Israel.

Elrod-Erickson, M., Mishra, S. and Schneider, D. 2000. Interaction between the cellular and humoral immune responses in *Drosophila*. Curr. Biol. 10: 781-784.

Expasy tools. http://www.expasy.ch/tools (accessed 9/04/07).

Evans, E. E., Painter, B., Evans, M. L., Weiheimer, P. and Acton, R. Y. 1968. An induced bactericidin in the spiny lobster *Panulirus argus*. Proceedings of the Society for Experimental Biology and Medicine 128: 394-398.

FAO (Food and Agriculture Organization of the United Nations). 2007. http://www.fao.org/fishery/species/2583/en, accessed 12/12/08.

FAO. 2008. http://www.fao.org/fishery/species/2583/en, accessed 12/12/08.

Fenouil, E. and Roch, P. 1991. Evidence and characterization of lysozyme in six species of freshwater crayfishes from Astacidae and Cambaridae families. Comp. Biochem. Physiol. 99B: 43-9.

Foukas, L.C., Katsoulas, H.L., Paraskevopoulou, N., Metheniti, A., Lambropoulou, M. and Marmaras, V.J. 1998. Phagocytosis of *Escherichia coli* by insect hemocytes requires both activation of the Ras/mitogen-activated protein kinase signal transduction pathway for attachment and beta 3 integrin for internalization. J. Biol. Chem. 273: 14813-14818.

Fountain, D.W. and Campbell, B.A. 1984. A lectin isolated from mucus of *Helix aspersa*. Comp. Biochem. Physiol. 77B: 419-429.

Fragkiadakis, G.A. and Stratakis, E.K. 1995. Characterization of hemolymph lectins in the prawn *Parapenaeus longirostris*. J. Invertbr. Pathol. 65: 111-117.

Fragkiadakis, G.A. and Stratakis, E.K. 1997. The lectin from the crustacean *Liocarcinus depurator* recognises *O*-acetylsialic acids. Comp. Biochem. Physiol. 117B: 545-552.

Fujita, T. 2002. Evolution of the lectin-complement pathway and its role in innate immunity. Nat. Rev. Immunol. 2: 346-353.

GenBank. http://www.ncbi.nlm.nih.gov/ (accessed 9/04/07).

Greenberg, S. and Grinstein, S. 2002. Phagocytosis and innate immunity. Curr. Opin. Immunol. 14: 136-145.

Giacometti, A., Cirioni, P., Del Prete, M. S., Barchiesi, F., Paggi, A. M., Petrelli, E. and Scalise, G. 2000. Comparative activities of polycationic peptides and clinically used antimicrobial agents against multidrug-resistant nosocomial isolates of *Acinetobacter baumannii*. J. Antimicrob. Chemotherapy 46: 807-810.

Goarant, C., Merien, F., Berthe, F., Mermoud, I. and Perolat, P. 1999. Arbitrarily primed PCR to type *Vibrio* spp. pathogenic for shrimp. App. Environ. Microbiol. 65: 1145-1151.

Gokudan, S., Muta, T., Tsuda, R., Koori, K., Kawahara, T., Seki, N., Mizunoe, Y., Wai, S.N., Iwanaga, S. and Kawabata, S.-I. 1999. Horseshoe crab acetyl group-recognizing lectins involved in innate immunity are structurally related to fibrinogen. Proc. Natl. Acad. Sci. USA. 96: 10086-10091.

Goldstein, I.J., Hughes, C.E., Monsigny, M., Osawa, T. and Sharon, N. 1980. What should be called a lectin? Nature 285: 66.

Graves, B.J., Crowther, R.L., Chandran, C., Rumberger, J.M., Li, S., Huang, K.S., Presky, D.H., Familletti, P.C., Wolitzky, B.A. and Burns, D.K. 1994. Insight into E-selectin/ligand interaction from the crystal structure and mutagenesis of the lec/EGF domains. Nature 367: 532-538.

Gross, P.S., Bartlett, T.C., Browdy, C.L., Chapm, R.W., Warr, G.W. 2001. Immune gene discovery by expressed sequence tag analysis of haemocytes and

hepatopancreas in the Pacific white shrimp, *Litopenaeus vannamei*, and the Atlantic white shrimp, *L. setiferus*. Dev. Comp. Immunol. 25: 565-577.

Hall, J.L. and Rowlands, D.T. 1974. Heterogeneity of lobster agglutinins. I. Purification and physiochemical characterization. Biochemistry 13: 821-827.

Hall, J.L. and Rowlands, D.T. 1974. Heterogeneity of lobster agglutinins: II. Specificity of agglutinin-erythrocyte binding. Biochemistry 13: 828-832.

Hall, M., Wang, R., Antwerpen, R., Sottrup-Jensen, L. and Söderhäll, K. 1999. The crayfish plasma clotting protein : A vitellogenin-related protein responsible for clot formation in crustacean blood. Proc. Natl. Acad. Sci. U.S.A. 96: 1965-1970.

Hardy, S.W., Grant, P. T. and Fletcher, T. C. 1977. A hemagglutinin in the tissue fluid of the Pacific oyster, *Crassostrea gigas* with specificity for sialic acid residues in glycoproteins. Experientia 33: 767-769.

Hatakeyama, T., Ouchi, K., Kuroki, M. and Yamasaki, N. 1995. Amino acid sequence of a C-type lectin CEL-IV from the marine invertebrate *Cucumaria echinata.* Biosci. Biotech. Biochem. 59: 1314-1317.

Hauton, C., Brockton, V. and Smith, V. J. 2007. Changes in immune gene expression and resistance to bacterial infection in lobster (*Homarus gammarus*) post-larval stage VI following acute or chronic exposure to immune stimulating compounds. Mol. Immunol. 44: 443-450.

Hikima, S., Hikima, J., Rojtinnakorn, J., Hirono, I. and Aoki, T. 2003. Characterization and function of Kuruma shrimp lysozyme possessing lytic activity against *Vibrio* species. Gene 316: 187-195.

Hoang, T. 2001. The banana shrimp-the right species for shrimp farming? World Aquaculture 32(4): 40-44.

Hockney, R. C. 1995. Development is heterologous protein production in *Escherichia coli*. Trends Biotech. 13: 456-463.

Hoebe, K., Jansen, E. and Beutler, B. 2004. The interface between innate and adaptive immunity. Nat. Immunol. 5: 971-974.

Hoffmann, J. A., Kafatos, F.C., Janeway, Jr. C. A. and Ezekowiz, R. A. B. 1999. Phylogenetic perspectives in innate immunity. Science 284: 1313-1318.

Holmblad, T., Thörnqvist, P.O., Söderhäll, K. and Johansson, M.W. 1997. Identification and cloning of an integrin beta subunit from hemocytes of the freshwater *Pacifastacus leniusculus*. J. Exp. Zool. 277: 255-261.

Holmskov, U., Malhotra, R., Sim, R.B. and Jensenius, J.C. 1994. Collectin: Colleagenous C-type lectins of the innate immune defense system. Immunol. Today 15: 67-74.

Holmskov, U., Thiel, S. and Jensenius, J.C.2003. Collectins and ficolins: Humoral lectins of the innate immune defense. Ann. Rev. Immunol. 21:547-578.

Holthuis, L.B. 1980. FAO species catalogue-shrimps and prawns of the world. An annotated catalogue of species of interest to fisheries. FAO Fish Synop, Roma 125: 1-271.

Huang, M.T.F., Enle, A.F. and Hummen, C.S. 1981. Immune response of the prawn, *Macrobrachium rosenbergii*, to bacterial infection. J. Invertbr. Pathol. 38: 213-219.

Huang, C. C., Sritunyalucksana, K., Söderhäll, K. and Song, Y. L. 2004. Molecular cloning and characterization of tiger shrimp (*Penaeus monodon*) transglutaminase. Dev. Comp. Immunol. 28: 279-294.

Hultmark, D. 1996. Insect lysozymes. In: Lysozymes: Model Enzymes in Biochemistry and Biology. Jolle's, P. Ed. Birkhäuser Verlag, Basel., pp. 87-102.

Ichijo, H., Hellman, U., Wernstedt, C., Gonez, L.J., Claesson-Welsh, L., Heldin, C.H. *et al*. 1993. Molecular cloning and characterization of ficolin, a multimeric protein with fibrinogen- and collagen-like domains. J. Biol. Chem. 268: 14505-14513.

Imuler, J. L. and Hoffmann, J. A. 2000. Signaling mechanisms in the antimicrobial host defense of *Drosophila*. Curr. Opin. Microbiol. 3: 16-22.

Inamori, K., Saito, T., Iwaki, D., Nagira, T., Iwanaga, S., Arisaka, F. and Kawabata, S. 1999. A newly identified horseshoe crab lectin with specificity for

blood group A antigen recognizes specific O-antigens of bacterial lipopolysaccharides. J. Biol. Chem. 274: 3272-3278.

Ito, Y., Yoshikawa, A., Hotani, T., Fukuda, S., Sugimura, K. and Imoto. T. 1999. Amino acid sequences of lysozymes newly purified from invertebrates imply wide distribution of a novel class in the lysozyme family. Eur. J. Biochem. 259: 456-461.

Iwanaga, S. 2002. The molecular basis of innate immunity in the horseshoe crab. *Curr. Opin. Immunol.* 14: 87-95.

Iwanaga, S. and Lee, B.L. 2005. Recent advances in the innate immunity of invertebrate animals. J. Biochem. Mol. Biol. 38: 128-150.

Iwanaga, S., Morita, T., Harada, T., Niwa, M., Takada, K., Kimura, T. and Sakakibara, S. 1978. Chromogenic substrates for horseshoe crab clotting enzyme. Its application for the assay of bacterial endotoxins. Haemostatis 7: 183-188.

Janeway, Jr., C.A. 1989. Approaching the asymptote? Evolution and revolution in immunology. Cold Spring Harb. Symp. Quant. Biol. 54: 1-13.

Janeway Jr., C.A. and Medzhitov, R., 2002. Innate immune recognition. Ann. Rev. Immunol. 20: 197-216.

Jiang, H., Wang, Y. and Kanost, M. R. 1998. Pro-phenol oxidase activating proteinase from an insect, *Manduca sexta*: bacteria-inducible protein similar to *Drosophila* easter. Proc. Natl. Acad. Sci. USA. 95: 12220-12225.

Jiravanichpaisal, P., Sricharoen, S., Söderhäll, I. and Söderhäll, K. 2006. White spot syndrome virus (WSSV) interaction with crayfish haemocytes. Fish Shellfish Immunol. 20: 718-727.

Jiravanichpaisal, P., Lee, S.Y., Kima, Y.A., Andrena, T. and Söderhäll, I. 2007. Antibacterial peptides in hemocytes and hemotopoietic tissue from freshwater crayfish *Pacifastacus leniusculus*: Charactertization and expression pattern. Dev. Comp. Immunol. 31: 441-455.

Johnson, P.T. 1980. Histology of the blue crab, *Callinectes sapidus*. A model for the Decapoda. Praeger, New York., pp. 440.

Johnson, P.T. 1987. A review of fixed phagocytic and pinocytic cells of decapod crustaceans with remark on hemocytes. Dev. Comp. Immunol. 19: 679-704.

Johansson, M. W., Holmblad, T., Thornqvist, P. O., Cammarata, M., Parrinello, N. and Söderhäll, K. 1999. A cell-surface peroxide dismutase is a binding protein for peroxinectin, a cell-adhesive peroxidase in crayfish. J. Cell Sci. 112: 917-925.

Johansson, M.W., Keyser, P., Sritunyalucksana, K. and Söderhäll, K. 2000. Crustacean haemocytes and haematopoiesis. Aquaculture 191: 45-52.

Jolles, P. Ed. 1996. Lysozymes: Model Enzymes in Biochemistry and Biology. Birkhäuser Verlag. Basel.

Jolles, J. and P. Jolles, P. 1975. The losozyme from *Asterias rubens*. Eur. J. Biochem. 54: 19-23.

Jolles, P. and Jolles. J. 1984. What's new in lysozyme research? Always a model system, today as yesterday. Mol. Cell. Biochem. 63: 165-189.

Kanost, M. R., Jiang, H. and Yu, X. Q. 2004. Innate immune responses of a lepidopteran insect, *Manduca sexta*. Immunol. Rev. 198: 97-105.

Kawabata, S. and Iwanaga, S. 1999. Role of lectins in the innate immunity of horseshoe crab. Dev. Comp. Immunol. 23: 391- 400.

Kim, Y., Ryu, J., Han, S., Choi, K., Nam, K., Jang, I., Lemaitre, B., Brey, P. T. and Lee, W. J. 2000. Gram-negative bacteria-binding protein, a pattern recognition receptor for lipopolysaccharide and β-1,3-glucan that mediates the signaling for the induction of innate immune genes in *Drosophila melanogaster* cells. J. Biol. Chem. 275: 32721-32727.

Kimura, T., Imai, Y. and Irimura, T. 1995. Calcium-dependent conformation of a mouse macrophage calcium-type lectin. Carbohydrate binding activity is stabilized by an antibody specific for a calcium-dependent epitope. J. boil. Chem. 270: 16056-16062.

Khoo, L., Robinette, D. W. and Noga, E.J. 1999. Callinectin, an antibacterial peptide from blue crab, *Callinectes sapidus*, hemocytes. Marine Biotechnol. 1: 44-51.

Kobayashi, M., Johansson, M. W. and Söderhäll, K. 1990. The 76 kD cell-adhesion factor from crayfish haemocytes promotes encapsulation in vitro. Cell and Tissue Research. 260: 13-18.

Koizumi, N., Imai, Y., Morozumi, A., Imamura, M., Kadotani, T., Yaoi, K., Iwahana, H., and Sato, R. 1999. Lipopolysaccharide-binding protein of *Bombyx mori* participates in a hemocyte mediated defense reaction against gram-negative bacteria. J. Insect Physiol. 45: 853-859.

Komatsu, M. and Ando, S. 1998. A very-high-density lipoprotein with clotting ability from hemolymph of sand crayfish, *Ibacus ciliatus*. Biosci. Biotechnol. Biochem. 62: 459-463.

Kondo, M., Itami, T. and Takahashi, Y. 1998. Preliminary characterization of lectins in the hemolymph of kuruma prawn. Fish Pathol. 33: 429-435.

Kondo, M., Matsuyama, H. and Yano, T. 1992. The opsonic effect of lectin on phagocytosis by hemocytes of Kuruma prawn, *Penaeus japonicus*. Fish. Pathol. 27: 217-222.

Kopacek, P., Grubhoffer, L. and Söderhäll, K. 1993. Isolation and characterization of a hemagglutinin with affinity for lipopolysaccharides from plasma of the crayfish *Pacifastacus leniusculus*. Dev. Comp. Immunol. 17: 407-418.

Kornfeld, R. and Kornfeld, S. 1985. Assembly of asparagine-linked oligosaccharides. Annu. Rev. Biochem. 54: 631-664.

Krueger, J.K., Kulke, M.H., Schutt, C. and Stock, J. 1989. Protein inclusion body formation and purification. Bio. Pharm. Manufact. 3: 40-45.

Krutzik, S. R., Sieling, P. A. and Modlin, R. L. 2001. The role of Toll-like receptors in host defense against microbial infection. Curr. Opin. Immunol. 13: 104-108.

Kumar, S., Tamura, K. and Nei, M. 2004. MEGA3: Integrated software for molecular evolutionary genetics analysis and sequence alignment. Brief Bioinform. 5: 150-163.

Laemmli, U.K. 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature 227: 680- 685.

Lai, C. Y., Cheng, W. and Kuo, C. M. 2005. Molecular cloning and characterization of prophenoloxidase from haemocytes of the white shrimp, *Litopenaeus vannamei*. Fish Shellfish Immunol. 18: 417-430.

Lavine, M. D. and Strand, M. R. 2002. Insect hemocytes and their role in immunity. Insect Biochem. Mol. Biol. 32: 1295-1309.

Lee, S.Y., Cho, M.Y., Hyun, J.H., Lee, K.M., Homma, K.I., Natari, S., Kawabata, S.-I., Iwanaga, S. and Lee, B.L. 1998a. Molecular cloning of cDNA from pro-phenoloxidase-activating factor I, a serine protease is induced by lipopolysaccharide or 1,3-β-glucan in coleopteran insect, *Holotrichia diomphalia* larvae. Eur. J. Biochem. 257: 615-621.

Lee, S.Y., Kwon, T.H., Hyun, J.H., Choi, J.S., Kawabata, S.-I., Iwanaga, S. and Lee, B.L. 1998b. In vitro activation of pro-phenol-oxidase by two kinds of pro-phenol-oxidase-activating factors isolated from hemolymph of coleopteran, *Holotrichia diomphalia* larvae. Eur. J. Biochem. 254: 50-57.

Lee, W.J., Lee, J.D., Kravchenko, V.V., Ulevitch, R.J. and Brey, P.T. 1996. Purification and molecular cloning of an inducible Gram negative bacteria-binding protein from the silkworm *Bombyx mori*. Proc. Natl. Acad. Sci. USA. 93: 7888-7893.

Lee, S.Y. and Söderhäll, K. 2002. Early events in crustacean innate immunity. Fish Shellfish Immunol. 12: 421-437.

Lee, S.Y., Wang, R. and Söderhäll, K. 2000. A lipopolysaccharide- and β-1,3-glucan-binding protein from hemocytes of the freshwater crayfish *Pacifastacus leninusculus*. Purification, characterization, and cDNA cloning. J. Biol. Chem. 275: 1337-1343.

Lee, S.Y., Wang, R. and Söderhäll, K. 2000. A lipopolysaccharide- and β-1,3-glucan-binding protein from hemocytes of the freshwater crayfish *Pacifastacus leninusculus*. Purification, characterization, and cDNA cloning. J. Biol. Chem. 275: 1337-1343.

Lemaitre, B., Nicolas, E., Michaut, L., Reichhart, J. M. and Hoffmann, J. A. 1996. The dorsoventral regulatory gene cassette spätzle/Toll/cactus controls the potent antifungal response in *Drosophila* adults. Cell 86: 973-983.

Liang, Z., Sottrup-Jensen, L. and Söderhäll, K. 1997. Pacifastin, a novel 155-kDa heterodimeric proteinase inhibitor containing a unique transferrin chain. Proc. Natl. Acad. Sci. USA. 94: 6682-6687.

Lightner, D.V. 1996. Epizootiology, distribution and the impact on international trade of two penaeid shrimp viruses in the Americas. Rev. Sci. Tech. Oie. 15: 579-601.

Lin, Y.-C., Vaseeharan, B. and Chen, J.-C. 2008. Identification and phylogenetic analysis on lipopolysaccharide and β-1,3-glucan binding protein (LGBP) of kuruma shrimp *Marsupenaeus japonicus*. Dev. Comp. Immunol. 32: 1260-1269.

Lis, H. and Sharon, N. 1998. Lectins: carbohydrate-specific proteins that mediate cellular recognition. Chem. Rev. 98: 637.

Liu, C.H., Tseng, D.Y., Lai, C.Y., Cheng, W. and Kuo, C.M. 2006. Molecular cloning and characterization of prophenoloxidase cDNA from haemocytes of the giant freshwater prawn, Macrobrachium rosenbergii, and its transcription in relation with the moult stage. Fish Shellfish Immunol. 21: 60-69.

Liu, Y.C., Li, F.H., Dong, B., Wang, B., Luan, W., Zhang, X.J., Zhang, L.S., Xiang, J.H. 2007. Molecular cloning, characterization and expression analysis of a putative C-type lectin (Fclectin) gene in Chinese shrimp *Fenneropenaeus chinensis*. Mol. Immunol. 44: 598-607.

Lodish, H. F. 1991. Recognition of complex oligosaccharides by the multi-subunit asialoglycoprotein receptor. Trends Biochem. Sci. 16: 374-377.

Lotan, R. and Nicolson, G. L. 1979. Purification of cell membrane glycoproteins by lectin affinity chromatography. Biochim. Biophys. Acta 559: 329-376.

Lu, J. 1997. Collectins: Collectors of microorganisms for the innate immune system. BioEssays 19: 509-518.

Luo, T., Li, F., Lei, K. and Xu, X. 2007. Genomic organization, promoter characterization and expression profiles of an antiviral gene PmAV from the shrimp *Penaeus monodon*. Mol. Immunol. 44: 1516-1523.

Luo, T., Yang, H., Li, F., Zhang, X. and Xu, X. 2006. Purification, characterization and cDNA cloning of a novel lipopolysaccharide-binding lectin from the shrimp *Penaeus monodon*. Dev. Comp. Immunol. 30: 607-617.

Ma, C. and Kanost, M. R. 2000. A beta 1,3-glucan recognition protein from an insect, *Manduca sexta*, agglutinates microorganisms and activates the phenoloxidase cascade. J. Biol. Chem. 275: 7505-7514.

Ma, T.H., Benzie, J.A.H., He, J.-G. and Chan, S.-M. 2008. PmLT, a C-type lectin specific to hepatopancreas is involved in the innate defense of the shrimp *Penaeus monodon*. J. Invert. Pathol. 99: 332-341.

Ma, T.H., Tiu, S.H., He, J.G. and Chan, S.-M. 2007. Molecular cloning of a C-type lectin (LvLT) from the shrimp *Litopenaeus vannamei*: early gene down-regulation after WSSV infection. Fish Shellfish Immunol. 23: 430-437.

Mai, W.-J. and Hu, C.-Q. 2008. Molecular cloning, characterization, expression and antibacterial analysis of a lysozyme homologue from *Fennerpenaeus merguiensis*. Mol. Biol. Rep. (In press, available on-line).

Maniatis, T., Fristsch, F. and Sambrook, J. 1989. Molecular Cloning: A Laboratory Manual 2nd Ed. Laboratory Press, Cold Spring Harbor, New York, USA.

Marques, M.R.F. and Barracco, M.A. 2000. Lectins, as non-self recognition factors, in crustaceans. Aquaculture 191: 23-44.

Martin, G. G., Poole, D., Poole, Ch., Hose, J. E., Arias, M., Reynolds, L. , McKerell, N. and Whang, A. 1993. Clearance of bacteria injected into the hemolymph of the penaid shrimp *Sicyonia ingentis*. J. Invert. Pathol. 62: 1-8.

Marton, A. O. 1986. The purification of eukaryotic polypeptides synthesized in *Escherichia coli*. Biochem. J. 240: 1-12.

Mashewari, R., Mullainadhan, P. and Arumugam, M. 1977. Characterization of a natural haemagglutinin with affinity for acetylated amino sugars in the serum of the marine prawn, *Penaeus indicus* (H Milne Edwards). Fish Shellfish Immunol. 7: 17-28.

Matsumoto, J., Nakamoto, C., Fujiwara, S., Yubisui, T. And Kawamura, K. 2001. A novel C-type lectin regulating cell growth, cell adhesion and cell differentiation of the multipotent epithelium in budding tunicates. Development 128: 3339-3347.

McGreal, E.P, Martinez-Pomares, L. and Gordon, S. 2004. Divergent roles for C-type lectins expressed by cells of the innate immune system. Mol. Immunol. 41: 1109-1121.

Medzhitov, R. and Janeway, Jr., C.A. 1997. Innate immunity: the virtues of a non-clonal system of recognition. Cell 91: 295.

Medzhitov, R. and Janeway, Jr. C.A. 2000. Innate immune recognition: mechanisms and pathway. Immunol. Rev. 173: 89-97.

Mercy, P. D. Sr. and Ravindranath, M.H. 1993. Purification and characterization of *N*-glycolylneuraminic-acid-specific lectin from *Scylla serrata*. Eur. J. Biochem. 215: 697-704.

Mercy, P. D. Sr. and Ravindranath, M.H. 1994. Hemolysis and clearance of erythrocytes in *Scylla serrata* are related to the agglutination by the native sialic acid-specific lectin. Comp. Biochem. Physiol. 109A: 1075-1083.

Millar, D.A. and Ratcliffe, N.A. 1994. Invertebrates. In: Immunology, a Comparative Approach, Turner, R. J. Ed. Wiley-Interscience, Chichester., pp.29-68.

Motoh, H. 1984. Biology and Ecology of *Penaeus monodon*. Proceedings of the 1st International Conference on the Culture of Penaeid Prawns/Shrimp. SEAFDEC. Aquaculture Department, Iloilo City., pp. 27-36.

Mullainadhan, P. and Renwrantz, L. 1986. Lectin-dependent recognition of foreign cells by hemocytes of the mussel, *Mytilus edulis*. Immunobiology 171: 263-273.

Murakami, T., Niwa, M., Tokunaga, F., Miyata, T. and Iwanaga, S. 1991. Direct virus inactivation of tachyplesin I and its isopeptides from horseshoe crab hemocytes. Chemotherapy 37: 327-334.

Muramoto, K. and Kamyia, H. 1989. The amino-acid sequence of multiple lectins of the acorn barnacle *Megabalanus rosa* and its homology with animal lectins. Biochim. Biophys. Acta 1039: 42-51.

Muramoto, K., Matsuda, T., Nakada, K. and Kamiya, H. 1995. Occurrence of multiple lectins in the hemolymph of kuruma prawn *Penaeus japonicus*. Fisheries Sci. 61: 131-135.

Nagai, T., Osaki, T. and Kawabata, S. 2001. Functional conservation of hemocyanin to phenoloxidase by horseshoe crab antimicrobial peptides. J. Biol. Chem. 276: 27166-27170.

Nalini M., Mullainadhan P. and Arumugam M. 1994. Characterization of a natural hemagglutinin in the serum of a fresh-water crab *Parathelphusa hydrodromus*. Arch. Int. Physiol., Biochim. Biophys. 102: 259-264.

Natori, S., Shiraishi, S., Hori, S. and Kobayashi, A. 1999. The roles of *Sarcophaga* defense molecules in immunity and metamorphosis. Dev. Comp. Immunol. 23: 317-328.

Nilsen, I. W. and Myrnes, B. 2001. The gene of chlamysin, a marine invertebrate-type lysozyme, is organized similar to vertebrate but different from invertebrate chicken-type lysozyme genes. Gene 269: 27-32.

Nishimura, T. S., Nakagawa, H. and Uchida, M. K. 2001. D-Galactose-specific sea urchin lectin sugar-specifically inhibited histamine release induced by *Datura stramonium* agglutinin: Differences between sugar-specific effects of sea urchin lectin and those of D-galactose- or L-fucose-sprcific plant lectins. Japan. J. Pharmacol. 85: 443-452.

Noga, E. J., Arroll, T. A., Bullis, R. A. and Khoo, L. 1996. Antibacterial activity in hemolymph of white shrimp *Penaeus setiferus*. J. Mar. Biotechnol. 4: 181-184.

Noga, E. J., Arroll, T. A. and Fan, Z. 1996b. Morphology of hemocyte lysis and clotting in ridgeback prawn *Sicyonia ingentis*. Cell Tissue Res. 255: 117-123.

Ochiai, M. and Ashida, M., 1988. Purification of a β-1,3-glucan recognition protein in the prophenoloxidase activating system from hemolymph of the silkworm, *Bombyx mori*. J. Biol. Chem. 263: 12056-12062.

Ochiai, M. and Ashida, M., 2000. A pattern-recognition protein for β-1,3-glucan. The binding domain and the cDNA cloning of β-1,3-glucan recognition

protein from the silkworm, *Bombyx mori*. J. Biol. Chem. 275: 4995-5002.

Omari, S.A., Martin, G.G. and Hose, J.E. 1989. Morphology of hemocyte lysis and clotting in the ridgeback prawn, *Sicyonia ingentis*. Cell Tissue Res. 255: 117-123.

Okino, N., Kawabata, S., Saito, T., Hirata, M., Takagi, T. and Iwanaga, S. 1995. Purification, characterization and cDNA cloning of a 27-kDa lectin L10 from horseshoe crab hemocytes. J. Biol. Chem. 270: 31008-31015.

Olafsen, J.A. 1986. Invertebrate lectins: Biochemical Heterogeneity as a Possible Key to Their Biological Function. In: Immunity in Invertebrates, Brehelin, M. Ed. Springer-Verlag, Berlin., pp. 94-111.

Olafsen, J. A. 1988. Role of lectins in invertebrate humoral defense. Am. Fish. Soc. Spec. Publ. 18: 189-205.

Olafsen, J.A., Fletcher, T. C. and Grant P. T. 1992. Agglutinin activity in Pacific oyster (*Crassostrea gigas*) hemolymph following in vivo *Vibrio anguillarum* Challenge. Dev. Com. Immunol. 16:123-138.

Park, D.S., Shin, S.W., Hong, S.D. and Park, H.Y. 2000. Immunological detection of serpin in the fall webworm, Hyphantria cunea and its inhibitory activity on the prophenoloxidase system. Molecules and Cells 10: 186-192.

Pereyra, A., Zenteno, R., Vazquez, L., Martinez-Cairo, S., Rodriguez, A., Mendoza-Hernandez, G., Zenteno, E. and Agundis, C. 2004. Characterization of lectin aggregates in the hemolymph of freshwater prawn *Macrobrachium rosenbergii*. Biochim. Biophys. Acta 1673: 122-130.

Perez-Farfante, I. and Kensley, B. 1997. Penaeoid and Sergestoid shrimps and prawns of the world: Keys and diagnoses for the families and genera. Memoirs du museum nationale d'histoire naturelle, tome. 175: 233.

Persson, M., Cerenius, L. and Söderhäll, K. 1987. The influence of haemocyte number on the resistance of the freshwater crayfish *Pacifastacus leniusculus* Dana, to the parasite fungus, *Aphanomyces astaci*. J. Fish Dis. 10: 471-477.

Pless, D., Aguilar, M., Falcon, A., Lozano-Alvarez, E. and Heimer de la Cotera, E. 2003. Latent phenoloxidase activity and N-terminal amino acid sequence of hemocyanin from *Bathynomus giganteus*, a primitive crustacean. Arch. Biochem. Biophys. 409: 402-410.

Primavera, J. H. 1988. Biology and culture of Penaeus monodon. Aquaculture Extension Manual 7, 3rd Ed.

Rattanachai, A., Hirono, I., Ohira, T., Takahashi, Y. and Aoki, T. 2004. Cloning of kuruma prawn *Marsupenaeus* japonicus crustin-like peptide cDNA and analysis of its expression. Fish. Sci. 70: 765-771.

Ratanapo, S. and Chulavatnatol, M. 1990. Monodin, a new sialic acid-specific lectin from black tiger prawn (*Penaeus monodon*). Comp. Biochem. Physiol. 97B: 515-520.

Ratanapo, S. and Chulavatnatol, M. 1992. Monodin-induced agglutination of *Vibrio vulnificus*, a major infective bacterium in black tiger prawn (*Penaeus monodon*). Comp. Biochem. Physiol. 102B: 855-859.

Ratcliffe, N.A., Rowley, A.F., Fitzgerald, S.W. and Rhodes, C.P. 1985. Invertebrate immunity: Basic concepts and recent advances. Int. Rev. Cytol. 97: 183-350.

Ravindranath, M.H., Higa, H.H., Cooper, E.L. and Paulson, J.C. 1985. Purification and characterization of an *O*-acetylsialic acid-specific lectin from a marine crab *Cancer antennarius*. J. Biol. Chem. 260: 8850-8856.

Rego, R.O., Kovar, V., Kopacek, P., Weise, C., Man, P., Sauman, I. *et al*. 2006. The tick plasma lectin, Dorin M, is a fibrinogen-related molecule. Insect Biochem. Mol. Biol. 36: 291-299.

Relf, J. M., Chisholm, J. R. S., Kemp, G. D. and Smith, V. J. 1999. Purification and characterization of a cysteine-rich 11.5 kDa antibacterial protein from the granular hemocytes of the shore crab, *Carcinus maenas*. Eur. J. Biochem. 264: 350-357.

Renwrantz, L. 1986. Lectins in molluscs and arthropods: their occurrence, origin and roles in immunity. Symp. Zool. Soc. Lond. 56: 81-93.

Richards, E.H. and Renwrantz, L.R. 1991. Two lectins on the surface of *Helix pomatia* haemocytes : Ca^2-dependent, GalNac-specific lectin and a Ca^2

-independent, mannose-6-phosphate-specific lectin which recognises activated homologous opsonins. J. Comp. Physiol. 161B: 43-54.

Richman, A. M., Dimopoulos, G., Seeley, D. and Kafatos, F.C. 1997. Plasmodium activates the innate immune response of *Anopheles gambiae* mosquitoes. EMBO J. 16: 6114-6119.

Rittidach, W. 2006. Characterization of sialic acid-specific lectin in the hemolymph of Banana shrimp (*Penaeus merguiensis*). Ph.D. Thesis, Prince of Songkla University, Songkhla, Thailand.

Rittidach, W., Paijit, N. and Utarabhand, P. 2007. Purification and characterization of a lectin from the banana shrimp *Fenneropenaeus merguiensis* hemolymph. Biochim. Biophys. Acta 1770: 106-114.

Rojtinnakorn, J., Hirono, I., Itami, T., Takahashi, Y. and Aoki, T. 2002. Gene Expression in haemocytes of Kuruma prawn, *Penaeus japonicus*, in response to infection with WSSV by EST approach. Fish Shellfish Immunol. 13: 69-83.

Rosenberry, B. 1997. World Shrimp Farming, 1997. Shrimp News International, San Diego.

Romeo-Figueroa, M.G., Vargas-Requena, C., Sotelo-Mundo, R.R., Vargas-Albores, F., Higuera-Ciapara, I., Söderhäll, K. *et al.* 2004. Molecular cloning of a β-glucan pattern-recognition lipoprotein from the white shrimp *Penaeus (Litopenaeus) vannamei*: correlations between the deduced amino acid sequence and the native protein structure. Dev. Comp. Immunol. 28: 713-26

Roux, M.M., Pain, A., Klimper, K.R., Dhar, A.Y. 2002. The lipopolysaccharide and β-1,3-glucan binding protein gene is upregulated in white spot virus-infected shrimp (*Penaeus stylirostris*). J. Virol. 76: 7140-7149.

Ruoslahti, E. 1996. RGD and other recognition sequences for integrins. Ann. Rev. Cell Dev. Biol. 12:.697-715.

Salzet, M. 2001. Vertebrate innate immunity resembles a mosaic of invertebrate immune responses. Trends Immunol. 22: 285-288.

Sambrook, J., Fritsch, E. F. and Maniatis, T. 1989. Molecular Cloning: A Laboratory Manual, 2nd Ed. Cold Spring Harbor Laboratory, Cold Spring Harbor, New York.

Sanger, F., Nicklen, S. and Coulson, A. R. 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Aced. Sci. 74: 5463-5467.

Sansonetti, P. J. 2006. The innate signaling of dangers and the dangers of innate signaling. Nat. Immunol. 7: 1237-1242.

Satoh, D., Horii, A., Ochiai, M. and Ashida, M. 1999. Prophenoloxidase-activating enzyme of the silkworm, *Bombyx mori*. J. Biol. Chem. 274: 7441-7453.

Schnapp, D., Kemp, G. D. and Smith, V. J. 1996. Purification and characterization of a proline-rich antibacterial peptide, with sequence similarity to bactenecin-7 from the hemocytes of the shor crab, Carcinus maenas. Eur. J. Biochem. 240: 532-539.

Sharon, N. and Lis, H. 1972. Lectin: Cell-agglutination and sugar-specific protein. Science 177: 949-958.

Sharon, N. and Lis, H. 1989. Lectins as cell recognition molecules. Science. 246: 227-246.

Sharon, N and Lis, H. 1995. Lectins, protein with a sweet tooth function in cell recognition. Essays in Biochemistry 30: 59-75.

Sharon, N. and Lis, H. 2004. History of lectins: From hemagglutinins to biological recognition molecules. Glycobiology 14: 53-62.

Shin, S.W., Park, D.-S., Kim, S.C. and Park, H.-Y. 2000. Two carbohydrate recognition domains of *Hyphantria cunea* lectin bind to bacterial lipopolysac-charides through O-specific chain. FEBS Lett. 467: 70-74.

Shin, S.W., Park, S.-S., Park, D.-S., Kim, M.G., Kim, S.C., Brey, P.T. and Park, H.-Y. 1998. Isolation and characterization of immune-related genes from the fall webworm, *Hyphantria cunea*, using PCR-based differential display and subtractive cloning. Insect Biochem. Mol. Biol. 28: 826-837.

Signal P. http:// www.cbs.dtu.dk/services/SignalP/ (accessed 9/04/07).

SMART. http://smart.embl-heidelberg.de/ (accessed 9/04/07).

Smith, V.J. and Chisholm, R.S., 1992. Non-cellular immunity in crustaceans. Fish Shellfish Immunol. 2: 1-31.

Söderhäll, K. and Cerenius, L. 1992. Crustacean immunity. Annu. Rev. Fish Dis. 2: 3-23.

Söderhäll, K., Cerenius, L. and Johansson, M. W. 1996. The Prophenoloxidase Activating System in Invertebrates. In: New Directions in Invertebrate Immunology, Söderhäll, K., Iwanaga, S. and Vasta, G. R. Eds. SOS Publications, Fair Haven., pp. 229-253.

Söderhäll, K. and Cerenius, L. 1998. Role of prophenoloxidase-activating system in invertebrate immunity. Curr. Opin. Immunol. 10: 23-28.

Söderhäll, K., Smith, V. J. and Johansson, M. W. 1986. Exocytosis and uptake of bacteria by isolated haemocyte populations of crustaceans: evidence for cellular co-operation in the defence reactions of arthropods. Cell Tissue Res. 24: 43-49.

Solis, N. B. 1988. Biology and Ecology. In: Biology and Culture of *Penaeus monodon*, Taki, Y. Premavara J. H. and Lobrera J. Eds. Aquaculture Department, Southeast Asian Fisheries Development Center.

Sotelo-Mundo, P.R., Islas-Osuna, M. A., de-la-Re-Vega, E., Hernandez-Lopez, J., Vargas-Albores, F. and Yepiz-Plascencia, G. 2003. cDNA cloning of the lysozyme of the white shrimp *Penaeus vannamei*. Fish Shellfish Immunol. 15: 325-331.

Sritunyalucksana, K., Lee, S.Y., Söderhäll, K. 2002. A β-1,3-glucan binding protein from the black tiger shrimp, *Penaeus monodon*. Dev. Comp. Immunol. 26: 237-45.

Sritunyalucksana, K., Sithisarn, P., Withayachumnarnkul, B. and Flegel, T.W. 1999. Activation of prophenoloxidase, agglutinin and antibacterial activity in haemolymph of the black tiger prawn, *Penaeus monodon*, by immunostimulants. Fish Shellfish Immunol. 9: 21-30.

Sritunyalucksana, K. and Söderhäll, K. 2000. The proPO and clotting system in crustaceans. Aquaculture 191: 53-69.

Stabili, L., Miglietta, A. M. and Belmonte, G. 1999. Lysozyme-like and trypsin-like activities in the cysts of *Artemia franciscana* Kellog, 1906. Is there a passive immunity in a resting stage? J. Experim. Mar. Biol. Ecol. 237: 291-303.

Stahl, P. D. 1992. The mannose receptor and other macrophage lectins. Curr. Opin. Immunol. 4: 49-52.

Stoss, T., Nickel, M., Hardin, D., Derby, C. and McClintock, T. 2003. Inducible transcript expressed by reactive epithelial cells at sites of olfactory sensory neuron proliferation. J. Neurobiol. 58: 355-368.

Su, J.G., Song, L.S., Xu, W., Wu, L.T., Li, H.L. and Xiang, J.H. 2004. cDNA cloning and mRNA expression of the lipopolysaccharide-and β-1,3-glucan-binding protein gene from scallop *Chlamys farreri*. Aquaculture 239: 69-80.

Sugumaran, M. 1996. Roles of the Insect Cuticle in Host Defense Reactions. In: Parasites and Pathogens of Insects, Söderhöll, K., Iwanaga, S., Vasta, G.R. and Federici, B.A. Eds. pp. 317-342.

Sugumaran, M. 2002. Comparative biochemistry of eumelanogenesis and the protective roles of phenoloxidase and melanin in insects. Pigment Cell Res. 15: 2-9.

Sugumaran, M. and Nellaiappan, K. 2000. Characterization of new phenoloxidase inhibitor from the cuticle of *Manduca sexta*. Biochem. Biophys. Res. Commun. 268: 379-383.

Sun, J., Wang, L., Wang, B., Guo, Z., Liu, M., Jiang, K. and Luo, Z. 2007. Purification and characterization of a natural lectin from the serum of the shrimp *Litopenaeus vannamei*. Fish Shellfish Immunol. 23: 292-299.

Sun, J., Wang, L., Wang, B., Guo, Z., Liu, M., Jiang, K., Tao, R. and Zhang, G. 2008. Purification and characterization of a natural lectin from the plasma of the shrimp *Fenneropenaeus chinensis*. Fish Shellfish Immunol. 25: 290-297.

Sun, Y.D., Fu, L.D., Jia, Y.P., Du, X.J., Wang, Q., Wang, Y.H. *et al*. 2008. A hepato-pancreas-specific C-type lectin from the Chinese shrimp *Fennero-*

penaeus chinensis exhibits antimicrobial activity. Mol Immunol. 45: 348-361.

Supungul, P., Klinbunga, S., Pichyanghura, R., Jitrapakdee, S., Hirono, I., Aoki, T. and Tassanakajon, A. 2002. Identification of immune related genes in hemocytes of black tiger shrimp (*Penaeus monodon*). Mar. Biotechnol. 4: 487-494.

Suzuki, T. and Mori, K. 1991. Immunolocalization and in vitro secretion of hemolymph lectin of the pearl oyster, *Pinctada fucata martensii*. Zool. Sci. 8: 23-29.

Takagi, T., Nakamura, A, Deguchi, R., Kyozuka, K. 1994. Isolation, characterization, and primary structure of three major proteins obtained from *Mytilus edulis* sperm. J. Biochem. 116: 598-605.

Takahashi T., Itoh M. and Shimabayashi Y. 1980. Aggregation and fragmentation of *Phaseolus vulgaris* lectin. Agric. Biol. Chem. 44: 1655-1657.

Thornqvist, P. O., Johansson, M. W. and Söderhäll, K. 1994. Opsonic activity of cell adhesion proteins and β-1,3-glucan-binding proteins from two crustaceans. Dev. Comp. Immunol. 18: 3-12.

Toda, M., Jimbo, M., Muramoto, K., Sakai, R. and Kamiya, H. 1998. Isolation and characterization of a D-galactose-binding lectin from the acorn barnacle *Balanus rostratus*. Fish Sci. 64: 638-642.

Tokunaga, F., Muta, T., Iwanaga, S., Ichinose, A., Davie, E.W., Kuma, K. and Miyata, T. 1993a. *Limulus* hemocyte trangulaminase. cDNA cloning, amino acid sequence and tissue localization. J. Biol. Chem. 268: 262-268.

Tokunaga, F., Yamada, M., Miyata, T., Ding, Y.L., Hiranaga-Kawabata, M., Muta, T., Iwanaga, S., Ichinose, A. and Davie, E.W., 1993b. *Limulus* hemocyte trangulaminase. Its purification and characterization and identification of the intracellular substrates. J. Biol. Chem. 268: 252-261.

Tsuiji, M., Fujimori, M., Ohashi, Y., Higashi, N., Onami, T.M., Hedrick, S.M. and Irimura, T. 2002. Molecular cloning and characterization of a novel mouse macrophage C-type lectin, mMGL2, which has a distinct

carbohydrate specificity from mMGL1. J. Biol. Chem. 277: 28892-28901.

Tunkijianukij, S. and Olafsen, J.A. 1998. Sialic acid-binding lectin with antibacterial activity from the horse mussel: Further characterization and immunolocalization. Dev. Comp. Immunol. 22: 139-150.

Underhill, D. M. and Orinsky, A. 2002. Toll-like receptors: key mediators of microbe detection. Curr. Opin. Immunol. 14: 103-110.

Vargas-Albores, F. 1995. The defense system of brown shrimp *Penaeus californiensis* : humoral recognition and cellular responses. J. Mar. Biotechnol. 3: 153–156.

Vargas-Albores, F., Guzman-Murillo, M.A. and Ochoa, J.L. 1993. A lipopolysaccharide-binding agglutinin isolated from brown shrimp (*Penaeus californiensis* Holmes) hemolymph. Comp. Biochem. Physiol. 104A: 407-413.

Vargas-Albores, F., Jimenez-Vega, F. and Söderhäll, K. 1996. A plasma protein isolated from brown shrimp (*Penaeus californiensis*) which enhances the activation of prophenoloxidase system by β-1,3-glucan. Dev. Comp. Immunol. 20: 299-306.

Vargas-Albores, F., Jimenez-Vega, F. and Yepiz-Plascencia, G. 1997. Purification and comparison of β-1,3-glucan binding protein from white shrimp (*Penaeus vannamei*). Comp. Biochem. Physiol. 116B: 453-458.

Vargas-Albores, F., Baltazar, P. H., Clark, G. P. and Barajas, F. M. 1998. Influence of temperature and salinity on the yellowlegs shrimp, *Penaeus californiensis* Holmes, prophenoloxidase system. Aquacult. Res. 29: 549-553.

Vargas-Albores, F., Yepiz-Plascencia, G. Jimenez-Vega, F. and Avila-Villa, A. 2004. Structural and functional differences of *Litopenaeus vannamei* crustins. Comp. Biochem. Physiol. 138B: 415-422.

Vargas-Albores, F. and Yepiz-Plascencia, G. 2000. Beta glucan binding protein and its role in shrimp immune response. Aquaculture 191: 13-21.

Vasta, G.R. 1991. The Multiple Biological Roles of Invertebrate Lectins : Their Participation in Non-self Recognition Mechanisms. In: Phylogenesis

of Immune Functions, Warr, G.R. and Cohen, N. Eds. CRC Press, Boca Raton, FL., pp. 73-101.

Vasta, G.R. 1992. Invertebrate Lectins: Distribution, Synthesis, Molecular Biology, and Function. In: Glycoconjugates, Allen, H. J. and Kisalius, E. C. Eds. Marcel Dekker, New York., pp. 593-634.

Vasta, G.R., Ahmed, H. and Quesenberry, M.S. 1996. Invertebrate C-types Lectins and Pentraxins as Non-self Recognition Molecules. In: New Directions in Invertebrate Immunology, Söderhäll, K., Iwanaga, S. and Vasta, G.R. Eds. SOS Publications, Fair Haven, NJ., pp.189-227.

Vazquez, L., Jaramillo, L., Lascurain, R., Cooper, E.L., Rosas, P. and Zenteno, E. 1996. Bacterial agglutination by the sialic acid-specific serum lectin from *Macrobrachium rosenbergii*. Comp. Biochem. Physiol. 113B: 355-359.

Vazquez, L., Maldonado, G., Agundis, C., Perez, A., Cooper, E.L. and Zenteno, E. 1997. Participation of a sialic acid specific lectin from freshwater prawn *Macrobrachium rosenbergii* hemocytes in recognition of non-self cells. J Exp. Zoo. 279: 265-272.

Vazquez, L., Masso, F., Rosas, P., Montano, L. and Zenteno, E. 1993. Purification and characterization of a lectin from *Macrobrachium rosenbergii* (Crustacea, Decapoda) hemolymph. Comp. Biochem. Physiol. 105B: 617-623.

Vazquez, L., Perez, A., Millan, D., Agundis, C., Martin, G., Cooper, E.L., Lascurain, R. and Zenteno, E. 1997. Morphology of hemocytes from the freshwater prawn *Macrobrachium rosenbergii*. J. Morphol. 234: 147-153.

Vogt, G. 1996. Cytopathology of Bay of Piran shrimp virus (BPSV), a new crustacean virus from the Mediterranean Sea. J. Invertebr. Pathol. 68: 239-245.

Wang, R., Liang, Z., Hall, M. and Söderhäll, K. 2001. A transglutaminase involved in the coagulation system of the freshwater crayfish, *Pacifastacus leniusculus*. Tissue localization and cDNA cloning. Fish Shellfish Immunol. 11: 623-637.

Wang, X., Rocheleau, T.A., Fuchs, J.F., Hillyer, J.F., Chen, C.C. and Christensen, B.M. 2004. A novel lectin with a fibrinogen-like domain and its potential involvement in the innate immune response of *Armigeres subalbatus* against bacteria. Insect Mol. Biol. 13: 273-282.

Watanabe, A., Miyazawa, S., Kitami, M., Tabunoki, H., Ueda, K. and Sato, R. 2006. Characterization of a novel C-type lectin, *Bombyx mori* multibinding protein, from the *B. mori* hemolymph: Mechanism of wide-range microorganism recognition and role in immunity. J. Immunol. 177: 4594-4604.

World Aquaculture Society (WAS). 2006. http://www.was.org (accessed 13/12/08)

Weis, W.I. and Drickamer, K. 1996. Structural basis of lectin-carbohydrate recognition. Annu. Rev. Biochem. 65: 441-473.

Weis, W. I., Drickamer, K. and Hendrickson, W. A. 1992. Structure of a C-type mannose-binding protein complexed with an oligosaccharide. Nature 360: 127-134.

Weis, W.I., Taylor, M.E. and Drickamer, K. 1998. The C-type lectin superfamily in the immune system. Immunol. Rev. 163: 19-34.

Wilson, R., Chen, C. and Ratcliffe, N.A. 1999. Innate immunity in insects: the role of multiple, endogenous serum lectins in the recognition of foreign invaders in the cockroach, *Blaberus discoidalis*. J. Immunol. 162: 1590-1596.

Yang, H., Luo, T., Li, F., Li, S. and Xu, X. 2007. Purification and characterization of a calcium-independent lectin (PjLec) from the haemolymph of the shrimp *Penaeus japonicus*. Fish Shellfish Immunol. 22: 88-97.

Yeaton, R.W. 1981. Invertebrate lectins: I. Occurrence. Dev. Comp. Immunol. 5: 391-402.

Yeh, M. S., Chen, Y. L. and Tsai, I. H. 1998. The hemolymph clottable proteins of tiger shrimp, *Penaeus monodon*, and related species. Comp. Biochem. Physiol. 121B: 169-176.

Yepiz-Plascencia, G., Vargas-Albore, F., Jimenz-Vega, F., Ruiz-Verdugo, L.M., Romo-Figueroa, G. 1998. Shrimp plasma HDL and β-glucan binding

protein (BGBP): Comparison of biochemical characteristics. Comp. Biochem. Physiol. 121B: 309-314.

Yu. X.Q., Gan, H. and Kanost, M.R. 1999. Immulectin, an inducible C-type lectin from an insect, *Manduca sexta*, stimulates activation of plasma prophenol oxidase. Insect Biochem. Mol. Biol. 29: 585-597.

Yu. X.Q. and Kanost, M.R. 2000. Immulectin-2, a lipopolysaccharide-specific lectin from an insect, *Manduca sexta*, is induced in response to gram-negative bacteria. J. Biol. Chem. 275: 37373-37381.

Yu. X.Q. and Kanost, M.R. 2004. Immulectin-2, a pattern recognition receptor that stimulates hemocyte encapsulation and melanization in the tobacco hornworm, *Manduca sexta*. Dev. Comp. Immunol. 28: 891-900.

Yu. X.Q. and Ma, Y. 2006. Calcium is not required for immulectin-2 binding, but protects the protein from proteinase digestion. Insect Biochem. Mol. Biol. 36: 505-516.

Yu, X.Q., Tracy, M.E., Ling, E., Scholz, F.R. and Trencze, T. 2005. A novel C-type immulectin-3 from *Manduca sexta* is translocated from hemolymph into the cytoplasm of hemocytes. Insect Biochem. Mol. Biol. 35: 285-295.

Yu, X.-Q., Zhu, Y.-F., Ma, C., Fabrick, J. A. and Kanost, M. R. 2002. Pattern recognition proteins in *Manduca sexta* plasma. Insect Biochem. Mol. Biol. 32: 1287-1293.

Zelensky, A.N. and Gready, J.E. 2005. The C-type lectin-like domain superfamily. FEBS J. 272: 6179-6217.

Zhang, J., Li, F., Wang, Z. and Xiang, J. 2007. Cloning and recombinant expression of a crustin-like gene from Chinese shrimp, Fenneropenaeus chinensis. J. Biotechnol. 127: 605-614.

Zipris, D., Gilboa-Garber, N. and Susswein, A. J. 1986. Interaction of lectins from gonads and haemolymph of the sea hare *Aplysia* with bacteria. Microbios. 46: 193-198.

http://www.shrimpnews.com/FarmedSpecies.html (accessed 13/12/08).

Appendix

Amino acids and their abbreviations

Amino acid	Three letter code	Single letter code
Glycine	Gly	G
Alanine	Ala	A
Valine	Val	V
Leucine	Leu	L
Isoleucine	Ile	I
Methionine	Met	M
Phenylalanine	Phe	F
Tryptophan	Trp	W
Proline	Pro	P
Serine	Ser	S
Threonine	Thr	T
Cysteine	Cys	C
Tyrosine	Tyr	Y
Asparagine	Asn	N
Glutamine	Gln	Q
Aspartic acid	Asp	D
Glutamic acid	Glu	E
Lysine	Lys	K
Arginine	Arg	R
Histidine	His	H

Amino acids classifications

Physiochemistry	Amino acids
Hydrophobic aliphatic R groups	G A V L I M* C* P
Hydrophobic aromatic R groups	F Y W
Polar charged R group	R** K** H D*** E***
Polar uncharged R group	S T N Q

* = Sulphur R groups, ** = Acidic R groups and *** = Basic R groups

IUB codes

A = adenine
C = cytosine
G = guanosine
T = thymidine
U = uracil
K = G or T (Keto)
M = A or C (aMino)
R = A or G (puRine)
S = G or C (Strong-3H bonds)
W = A or T (Weak-2H bonds)
Y = C or T (pYrimidine)
B = C, G or T
D = A, G or T
H = A, C or T
V = A, C or G
N = any base

Vitae

Name Miss Onnicha Rattanaporn

Student ID 4523016

Educational Attainment

Degree	Name of Institution	Year of Graduation
B.Sc. (Biotechnology)	Prince of Songkla University	2002

Scholarship Awards during Enrolment

The Royal Jubilee Graduate Program from the Thailand Research Fund.

List of Publication and Proceeding

1. **Rattanaporn, O.**, Hedrick, J. L. and Utarabhand, P. 2008. Molecular cloning of lipopolysaccharide and β-1,3-glucan-binding protein (LGBP) from banana prawn *Fenneropepaeus merguiensis*. The 9th National Grad Research Conference. 11-12 September 2006, Bangkok, Thailand.
2. **Rattanaporn, O.**, Barisone, G. A., Hedrick, J. L. and Utarabhand, P. 2006. Molecular cloning of a lectin from the banana prawn *Fenneropepaeus merguiensis*. Reproductive Biology Colloquium, 2006. 27 october 2006, University of California, Davis.
3. **Rattanaporn, O.** and Utarabhand, P. 2009. Molecular cloning of a C-type lectin from the banana prawn *Fenneropepaeus merguiensis* (in preparation).
4. **Rattanaporn, O.**, Hedrick, J. L. and Utarabhand, P. 2009. Molecular cloning of lipopolysaccharide and β-1,3-glucan-binding protein (LGBP) from banana prawn *Fenneropepaeus merguiensis* (in preparation).