CLONING AND CHARACTERISATION OF CrSARM, A NOVEL SIGNALING MOLECULE RESPONSIVE TO *PSEUDOMONAS* INFECTION

BUI THI HONG HANH

(B.Sc (Hons), University of New South Wales)

A THESIS SUBMITTED

FOR THE DEGREE OF MASTER OF SCIENCE

DEPARTMENT OF BIOLOGICAL SCIENCES

NATIONAL UNIVERSITY OF SINGAPORE

2006

Acknowledgements

I would like to express my deepest gratitude to Prof. Ding and Prof Ho for giving me the opportunity to work on this project. They have provided excellent guidance and support throughout my years in the lab. From them, I have learnt many invaluable skills that are essential for my future career.

I would like to especially thank my mentor, Dr. Thangamani Saravanan, for giving me countless advices and suggestions, and for being so patient and understanding.

I also would like to express my gratefulness to LiYue, Patricia, Xiaolei, Nicole, Agnes, Belinda and Songyu for always sharing their experiences, giving their hands and being so supportive. Without them, I would not have such an enjoyable time in the lab.

Many thank also go to Shuba and Bee Ling for the efficient technical supports and to all other lab mates, Siou Ting, Xiaowei, Li Peng, Naxin, Cui Fang, Shijia, Bao Zhen, Jianmin, Diana, Derick, and Lihui for the help from time to time.

Most importantly, I would like to thank my family for their love, understanding and encouragement, which make the lonely time studying oversea bearable.

Table of Contents

Summary

Toll-like receptors (TLRs) play key roles in innate immunity. The hallmark of TLR signaling is the activation of distinct patterns of immune-related gene expression and optimization of innate immune response. The specificity is achieved by: (i) recognition and differentiation of different pathogen associated molecular patterns (PAMPs) by different TLRs and (ii) recruitment of distinct Toll/Interleukin-1 receptor (TIR) domain-containing adaptor proteins (via the dimerization of TIR domains) after ligation of PAMPs to respective TLRs. This leads to the activation of different downstream signal transduction pathways. Sterile alpha and Armadillo motif containing protein (SARM) is the most recently described TIR domain-containing adaptor protein. SARM homologs from different organisms share common domain architecture of two SAM motifs, an ARM motif and a TIR domain. All the individual domains of SARM are well-known for their ability to promote protein-protein interaction. Unlike other TIR domain-containing adaptors, the role of SARM remains unclear albeit its crucial contribution in *C. elegans* host defense against infection. Our investigation of proteins in the horseshoe crab, *Carcinoscopius rotundicauda*, that are responsive to *Pseudomonas aeruginosa* infection led to the isolation of an EST clone homologous to human SARM. Sequence alignment of CrSARM cDNA showed that it is highly homologous to SARM from other organisms, especially those from the arthropods. Although not ubiquitously expressed, CrSARM transcripts were detected in various tissues including the amebocytes and hepatopancreas (immune-responsive), heart and muscle. Transcript profiling indicated that CrSARM expression is induced rapidly in response to *P. aeruginosa* infection, hence its involvement in innate immunity. Yeast two hybrid screening identified the potential interaction partners of CrSARM to be: CaMKI, SUMO-1, HAX-1, proteasome-α subunit, and Hsp40. We

propose that CrSARM participates in calcium-dependent signaling cascade leading to the activation of p38 kinase and/or HAX-1 mediated signaling pathway that controls cytoskeletal remodeling and/or apoptosis. Activation of p38, apoptosis and cytoskeleton remodeling have been demonstrated to play significant roles in innate immunity. Further studies on CrSARM will give insights on this TIR domain-containing protein, as well as its implications in TLR signaling.

List of Tables

List of Figures

List of Abbreviations

aa	Amino acid
AD	Transcription activation domain
ARM	Armadillo
ASK1	Apoptosis signal regulated kinase1
BD	DNA-binding domain
BSA	Bovine serum albumin
CaMKII	Calcium-calmodulin-dependent protein kinase II
CDD	Conserved domain database
CFU	Colony forming unit
dATP	Deoxyadenosine triphosphate
dCTP	Deoxycytosine triphosphate
dGTP	Deoxyguanosine triphosphate
dNTP	Deoxynucleoside triphosphate
dTTP	Deoxythymidine triphosphate
DEPC	Diethylpyrocarbonate
DMSO	Dimethyl sulfoxide
DO	Drop out
EDTA	Ethylenediaminetetraacetic acid
EST	Express sequence tag
H	Hour
hpi	Hour post infection
IKK	Inhibitory κB kinase
IL	Interleukin
IOD	Intergrated optical density
IRF3	Interferon regulatory factor-3
IPTG	Isopropyl-1-thio-β-D-galactopyranoside
IRAK	IL-1 receptor-associated kinase
LB	Luria-Bertani
LPS	Lipopolysaccharide
LRR	Leucine rich repeat
Mal	MyD88-adaptor-like
MAPK	Mitogen-activated protein kinase
MBL	Mannan-binding lectin
MMLV	Moloney Murine Leukemia Virus
NF-κB	Nuclear factor-κB
OD	Optical density
ORF	Open reading frame
QDO	Quadruple drop out
PAMP	Pathogen-associated molecular pattern
PBS	Phosphate-buffered saline
PCR	Polymerase chain reaction
PMSF	Phenylmethylsulphonylfluoride
PRR	Pattern recognition receptors
RACE	Rapid amplification of cDNA ends
RbL3	Ribosomal protein L3

SAM	Sterile-α
SARM	SAM and ARM-containing protein
SDS	Sodium Dodecyl Sulfate
TAK1	Transforming growth factor-β-activating kinase
TBK1	TANK binding kinase-1
TICAM-1	TIR-containing adaptor molecule-1
TIR	Toll/Interleukin-1 receptor
TIRAP	TIR domain-containing adaptor protein
TLR	Toll-like receptor
TNF	Tumour necrosis factor
TRAM	Trif-related adaptor molecule
TRAP-6	Tumour necrosis factor receptor-associated factor-6
TRIF	TIR domain-containing adaptor inducting interferon-β
TIR-1	Toll and interleukin 1 receptor domain protein
TSA	Tryptone soy agar
TSB	Tryptone soy broth
U	Weiss Units
v/v	Volume/volume
w/v	Weight/volume

List of primers

Name	Sequence (5' to 3')	Purpose
CrSARM GENE SPECIFIC PRIMERS		
TIR-5RACE	CCACAGACCTCTTCCAGTTGGTC TTCCG	Reverse primer for 5'RACE of CrSARM
TIR-17-F1	CTGAAGTCAGACAGAAG	Forward primer for sequencing of 5'RACE product
TIR-17-R1	GGATCATGACTAGCAA	Reverse primer for sequencing of 5'RACE product
TIR-F1	AGACGTAGAGAGGCTCGAAGC	Forward primer for accessing tissue distribution and transcription profile of CrSARM
TIR-R1	TGTTCCCAGGGTCTTTCTTGT	Reverse primer for accessing tissue distribution and transcription profile of CrSARM
CrSARM GENE SPECIFIC PRIMERS WITH RESTRICTION SITES		
ARM-F-Nde	CATATGAATAGAGCTTACGTTGT GGA	Forward primer for cloning CrSARM-AST and ARM domain into pGBKT7
ARM-R-Bam	GGATCCTTAAGCTACTAAAGTAG TGATA	Reverse primer for cloning CrSARM-AST and ARM domain and into pGBKT7
SAM-F-Eco	GAATTCAACAAAGAGATAGAATT TGCAG	Forward primer for cloning SAM domain into pGBKT7
SAM-R-Bam	GGATCCTTATTTGTTGCAACTAAT ATCAC	Reverse primer for cloning SAM domain into pGBKT7
SARM-F-Nde	TACAGGACATATGGAAAATGGAT TCGCCC	Forward primer for cloning CrSARM-ORF into pGBKT7
SARM-R-Bam	GGATCCTTAAAGTTCCACCGAAC AAG	Reverse primer for cloning CrSARM-ORF and CrSARM-AST into pGBKT7
TIR-F-Nde	CAACAAACATATGACATTAGATG TCTTC	Forward primer for cloning TIR domain into pGBKT7
TIR-R-Eco	GGAATTCTCACTCCCCGCGCATG AACCT	Reverse primer for cloning TIR domain into pGBKT7
Carcinoscopius Ribosomal protein L3 GENE SPECIFIC PRIMERS		
RiboF:	TGTTTCTTCAGAGGACCCA	Forward primer for positive control of RT-PCR to analyze expression and transcription profile of CrSARM
RiboR:	CACCAAGAAGTTGCCTCG	Reverse primer for positive control of RT-PCR to analyze expression and transcription profile of CrSARM
VECTOR OR ANCHOR PRIMERS		
5'-RACE CDS primer	$(T)_{25}VN$	Reverse primer for synthesis of 5'-RACE-ready cDNA
SMART II A Oligo	AAGCAGTGGTATCAACGCAGAGT ACGCGGG	Forward primer for synthesis of 5'-RACE-ready cDNA
Universal primer A mix	*Long primer:* CTAATACGACTCACTATAGGGCA AGCAGTGGTATCAACGCAGAGT *Short primer:* CTAATACGACTCACTATAGGGC	Forward primer for RACE-PCR amplification of 5'-end of CrSARM
CDS III primer	ATTCTAGAGGCCGAGGCGGCCGA CATG-$d(T)_{30}VN$	Reverse primer for cDNA synthesis for Y2H
SMART III™ Oligo	AAGCAGTGGTATCAACGCAGAGT GGCCATTATGGCCGGG	Forward primer for cDNA synthesis for Y2H

pGAD-nt2025*f*	TTCCACCCAAGCAGTGGTATCAA CGCAGAGTGG	Forward primer for colony screening of pGADT7-cDNA clones
pGAD-nt2049*r*	GTATCGATGCCCACCCTCTAGAG GCCGAGGCGGCCGACA	Reverse primer for colony screening of pGADT7-cDNA clones
T7	ATACGACTCACTATAGGG	Sequencing primer

The coding for degenerate bases are: N = A, C, G, or T; V = A, G, or C

Chapter 1 Introduction

1.1 The innate immunity

Innate immunity is the host defense mechanism that is evolutionarily conserved in all metazoans. It serves as the powerful first-line defense against a wide variety of pathogens in both invertebrates, in which innate immunity is the exclusive host defense system, and vertebrates, which are also armed with the adaptive immunity. Indeed, it takes three to five days for the adaptive immune system to produce sufficient number of effector cells from the extremely diverse reservoir of naïve lymphocytes through the process of clonal selection and amplification. The delay in the generation of the adaptive immune responses would give the pathogens enough time to invade the host. Fortunately, upon the recognition of invading pathogens by the germ-line encoded receptors, the innate immune system rapidly mounts various responses including phagocytosis; synthesis and release of antimicrobial peptides; production of reactive oxygen and nitrogen; and activation of the alternative complement pathway to contain the proliferation of the infective pathogen until the adaptive immunity is ready to execute effective immune responses (Janeway and Medzhitov, 2002; Medzhitov and Janeway, 2000).

Not only providing immediately available defense mechanisms, in the vertebrates, the innate immune system also control and instruct the adaptive immune responses through the regulation of the expression of co-stimulators, chemokines, and cytokines upon pathogen infection (Fearon and Locksley, 1996; Janeway and Medzhitov, 2002). For example, the activation of T cells requires two signals presented on the surface of the antigen-presenting cells: (1) a peptide, which can be of

either self- or non-self origin, bound to MHC class II molecule; and (2) the presence of co-stimulatory molecules CD80 and CD86. The expression of CD80 and CDE86, however, can only be induced upon the recognition of infectious microbes by the receptor of the innate immune system (Banchereau and Steinman, 1998). This mechanism helps to avoid the generation of adaptive immune responses against the host itself once the self-antigen is presented on the surface of the antigen-presenting cells.

1.1.1 Basic paradigm of innate immunity: the recognition of pathogen associated molecular patterns (PAMP) by pattern recognition receptors (PRRs)

The first step in innate immune responses is the recognition of microbial components by the germ-line encoded receptors, called pattern recognition receptors (PRR) (Medzhitov and Janeway, 2000). In contrast to the amazing diversity of the T and B-cell receptors of adaptive immunity, which are generated by somatic recombination and hypermutation, the repertoire of PRRs is much more restricted due to the limited number of genes encoded in the genome of every organism. To overcome this limitation, the PRRs have evolved to recognize invariant molecular motifs common for the large groups of microorganisms, called pathogen-associated molecular patterns (PAMPs), rather than detect every possible antigen like the receptors of the adaptive immunity. Medzhitov and Janeway (Medzhitov and Janeway, 2000) have defined the common properties shared by all PAMPs to be: (1) expressed exclusively by the microbes but not the host, allowing the discrimination between self and non-self; (2) fundamental for the survival of microbes to prevent the generation of mutants that can escape the detection by PRRs; and (3) highly conserved in the entire array of microorganisms, and thus serves as indicators of the

classes of microbes. This feature not only allows the detection of a wide variety of microorganisms by a restricted repertoire of PRRs but also ensures that the innate immune system mounts the most appropriate responses at the critical time of infection. Examples of PAMPs are lipopolysaccharide (LPS) of the Gram-negative bacteria, lipoteichoic acid of Gram-positive bacteria; zymosan, mannan and β-glucan of fungi (Aderem and Ulevitch, 2000).

The introduction of PRR-PAMP paradigm has a great implication to the field of innate immunity since, for the first time, it provides a logical explanation of how this ancient host defense system is able to recognize virtually all microorganisms and to optimize the defense responses depending to the type of invading microbes.

1.1.2 Activation of intracellular signaling by PRRs

PRRs can be divided into three functional classes: opsonizing, endocytic and signaling (Medzhitov and Janeway, 2000). Opsonizing PRRs are plasma proteins that can bind to PAMPs on the surface of pathogens and activate the complement system and phagocytes to clear the invading microbes from the circulation. An example of a PRR belonging to this class of PRRs is mannan-binding lectin (MBL), a secretory product of the liver that can recognize equatorial 3-hydroxyl and 4-hydroxyl groups on the terminal sugar of the carbohydrates on the cell wall of gram-positive, gram negative, yeast, some viruses and parasites (Holmskov et al., 2003). The ligation of MBL to its correspondent PAMP leads to the activation of MBL-associated proteases 1 and 2 (MASP-1 and -2), ultimately resulting in the activation of lectin-dependent complement cascade (Epstein et al., 1996).

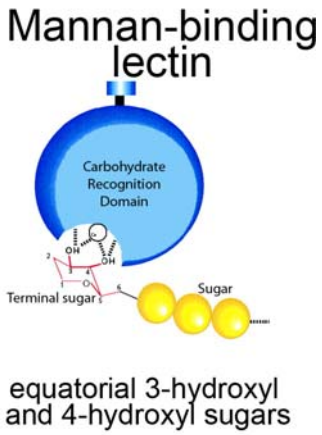


Figure 1.1 Mannan binding lectin recognizes equatorial 3-hydroxyl and 4-hydroxyl sugars. Figure is adapted from (Ng, 2005)

Members of the second group of PRRs are those expressed on the surface of phagocytes and mediate the endocytosis of the pathogens. Scavenger receptors, mannose receptors and β-glucan receptors of the macrophages are among the members of the endocytic PRR group (Mukhopadhyay and Gordon, 2004). The binding of these receptors to microbial ligands initiates the killing of the pathogen by phagocytosis. Phagocytosis also results in the generation of pathogen-derived peptides that can be presented on the surface of the macrophage by the major histocompatibility-complex for the activation of cells of the adaptive immune system (Fraser et al., 1998; Thomas et al., 2000).

Signaling PRRs are referred to those receptors that recognize PAMPs and activate intracellular signaling pathways resulting in the activation of immune-related genes. Toll-like receptors (TLRs), the best-characterized member of this group of PRRs, play the key role in the detection and elimination of pathogens. Signaling pathways activated by TLRs are conserved from invertebrates to vertebrates. Within the scope of this thesis, the following sections will focus on the significance of TLR-mediated signaling pathways to the host defense against infection.

1.2 Toll-like receptor signaling

1.2.1 Toll-like receptors and their ligands

Toll of *Drosophila* is the first member of Toll-like receptor family that was identified. At first, Toll was only known as a transmembrane receptor that functions in the establishment of dorsoventral polarity in the fly during embryogenesis (Hashimoto et al., 1988). Latter, the involvement of Toll in the innate immune responses was proposed based on the similarity between Toll and the mammalian interleukin-1 (IL-1) receptor. Indeed, genetic analysis of Toll revealed that its cytoplasmic domain is highly similar to that of IL-1 receptor. This motif was hence defined as Toll/IL-1 receptor (TIR) domain. In addition, signaling transductions induced by both receptors ultimately result in the activation of the same transcription factor, nuclear factor (NF)-κB, that is known to be involved in the regulation of many immune related genes (Belvin and Anderson, 1996). The role of Toll in host defense was experimentally confirmed by Lemaitre *et al* (Lemaitre et al., 1996) who showed that flies harboring a mutation in the *toll* gene were more susceptible to fungal infection. Subsequently, Toll was reported to be essential in immune response against Gram-positive bacteria (Rutschmann et al., 2002). However, not as expected, microbial component is not the direct ligand of *Drosophila* Toll. Rather, Toll was shown to recognize Spätzle, an endogenous protein that is cleaved from the precursor protein pro-Spätzle, upon infection by Gram-positive bacteria or fungi (Levashina et al., 1999). Molecular mechanism responsible for the cleavage of Spätzle, however, remains unclear. It was reported that peptidoglycan recognition protein PGRP-SA (Werner et al., 2000) and a serine protease encoded by *Persephone* gene (Ligoxygakis et al., 2002) are involved in the activation of Toll in response to the invasion of gram-positive bacteria and fungi, respectively. The former is a PRR that recognizes Gram-

positive lysine-type peptidoglycan whereas the latter does not contain any pattern recognition motif (Figure 1.2).

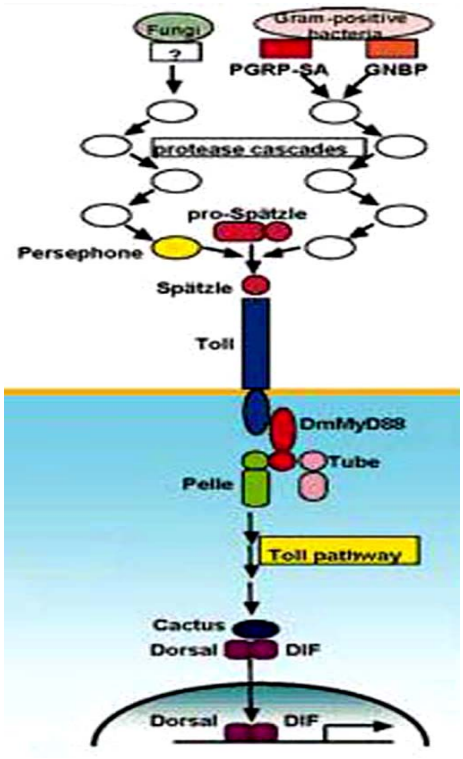


Figure 1.2 *Drosophila* Toll signaling pathway

Toll signaling pathway is essential for immune responses against Gram-positive bacterial and fungal infection in *Drosophila*. Gram-positive bacterial or fungi challenges initiate protease cascades leading to the cleavage of pro-Späetzle to Späetzle, which then binds and triggers the recruitment of DmMyD88 to TIR domain of Toll. The death domain of DmMyD88, on the other hand, interacts with the death domains of Tube and Pelle, which possesses a serine-threonine kinase domain. The formation of this receptor-adaptor complex induces signal transduction that leads to the dissociation of the ankyrin-repeat inhibitory protein, Cactus, a homolog of mammalian IκB, from the Dorsal-related immunity factor (Dorsal/Dif), which is a homolog of mammalian NF-κB. Nuclear translocation of the Dorsal-related immunity factor ultimately results in the activation of antimicrobial proteins. Figure is adapted from Takeda and Akira (Takeda and Akira, 2005)

Since the elucidation of the role of *Drosophila* Toll in innate immunity, 11 homologs of Toll, which are collectively named as Toll-like receptors (TLRs), have been identified in mammals (Takeda and Akira, 2005, and Table 1.1). The presence of a TLR in horseshoe crab, *Tachypleus tridentatus*, was reported recently by Inamori et al (Inamori et al., 2004). In our lab, the TIR domain of *Carcinoscopius rotund*icauda

has been recently cloned (Loh *et al*, unpublished). Members of TLR family share a common structural feature with a transmembrane portion that is flanked with an extracellular region containing leucine-rich repeat (LRR) motif and an intracellular region which contains a TIR domain9 (Figure 1.3).

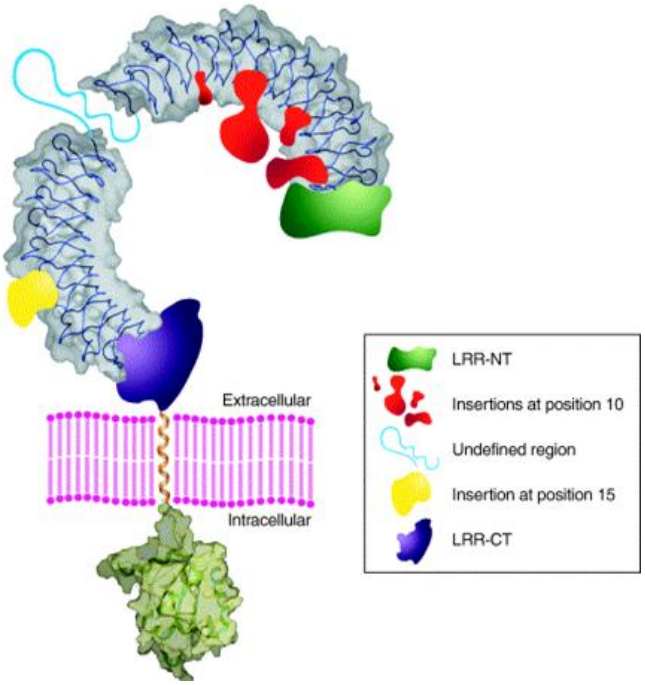


Figure 1.3 Predicted three dimensional structure of TLR.

Each TLR is a transmembrane protein with the extracellular LRR motif responsible for ligand binding and the intracellular TIR domain for signal transduction, which are in grey and cyan, respectively. Mammalian TLRs have large insertions at position 10 or 15 within the LRRs that may provide the binding site for PAMPs. LRR-NT & -CT: N- and C-terminal flanking regions of the LRR motif, respectively. Figure is adapted from Bell et al (Bell et al., 2003).

LRR domain of TLRs, which contains up to 25 tandem repeats of a conserved leucine-rich repeat sequence of 24-29 amino acids, is responsible for ligand binding. Based on the three dimensional structures of other LRR-containing proteins, it was predicted that LRR motif of the TLR form a horseshoe structure with the concave surface is likely to be involved in ligand binding. Unlike *Drosophila* Toll, extracellular domains of individual mammalian TLRs can directly recognize specific

pathogen-associated molecular patterns from bacteria, fungi, virus, and protozoa (Akira and Takeda, 2004, and Table 1.1). This difference was proposed to be related to the presence of large insertions at residue 10 or 15 within the LRRs of mammalian TLRs that are absent from LRRs of Toll. It was hypothesized that these insertions provided the binding sites for the PRRs of mammalian TLRs (Bell et al., 2003, and Figure 1.3).

Table 1.1 Ligands of mammalian TLRs

Receptor	Ligand	Origin of ligand
TLR1	Triacyl lipopeptides Soluble factors	Bacteria and mycobacteria *Neisseria meningitidis*
TLR2	Lipoprotein/lipopeptides Peptidoglycan Lipoteichoic acid Lipoarabinomannan Phenol-soluble modulin Glycoinositolphospholipids Glycolipids Porins Atypical lipopolysaccharide Atypical lipopolysaccharide Zymosan Heat-shock protein 70*	Various pathogens Gram-positive bacteria Gram-positive bacteria Mycobacteria *Staphylococcus epidermidis* *Trypanosoma cruzi* *Treponema maltophilum* *Neisseria* *Leptospira interrogans* *Porphyromonas gingivalis* Fungi Host
TLR3	Double-stranded RNA	Viruses
TLR4	Lipopolysaccharide Taxol Fusion protein Envelope protein Heat-shock protein 60* Heat-shock protein 70* Type III repeat extra domain A of fibronectin* Oligosaccharides of hyaluronic acid* Polysaccharide fragments of heparan sulphate* Fibrinogen*	Gram-negative bacteria Plants Respiratory syncytial virus Mouse mammary-tumour virus *Chlamydia pneumoniae* Host Host Host Host Host
TLR5	Flagellin	Bacteria
TLR6	Diacyl lipopeptides Lipoteichoic acid Zymosan	*Mycoplasma* Gram-positive bacteria Fungi
TLR7	Imidazoquinoline Loxoribine Bropirimine Single-stranded RNA	Synthetic compounds Synthetic compounds Synthetic compounds Viruses
TLR8	Imidazoquinoline Single-stranded RNA	Synthetic compounds Viruses
TLR9	CpG-containing DNA	Bacteria and viruses
TLR10	N.D.	N.D.
TLR11	N.D.	Uropathogenic bacteria

The preparation of ligands marked with star (*) may be contaminated with LPS. Further investigation is thus needed to confirm the connection between these ligands with TLRs. Table is adapted from Akira and Takeda (Akira and Takeda, 2004)

While the extracellular domain of TLRs functions in the detection of pathogens, the **TIR domain** at their cytoplasmic tail mediates signal transduction across the plasma membrane to the downstream components of TLR signaling pathway within the cell. TIR domain is a protein motif of approximately 200 amino acids with relatively low sequence homology (20 – 30 %). However, within the TIR domain, there are three highly conserved regions, named boxes 1, 2 and 3. These boxes play a critical role in mediating the interaction between the TIR domains of two TLRs or of the TLRs and their respective downstream adaptive proteins (Slack et al., 2000). Homodimerization of TLR4 and heterodimerization between TLR2 with either TLR1 and TLR6 that are mediated by the interaction of TIR domains of the TLRs are crucial for the recruitment of downstream adaptor proteins. For other TLRs, however, there is no evidence of dimerization of the receptor in response to ligand binding. The recruitment of adaptor molecule just downstream of TLR to the activated TLR is also based on the interaction between the TIR domains present in both molecules. In addition to the TLRs and IL-1 receptors, these cytoplasmic adaptor proteins form the third group of proteins that possess TIR domain (Akira, 2000). They are thus collectively referred as TIR domain-containing adaptors.

Research has shown that, depending on the nature of microbial challenge, different TLRs are stimulated to recruit the respective TIR domain-containing adaptor molecule, leading to the activation of distinct sets of genes. For example, the stimulation of TLR3 and TLR4 lead to the activation of genes under the control of the transcription factor interferon regulatory factor-3 (IRF-3) whereas activation of TLR2, TLR7, and TLR9 result in the induction of genes under control of transcription factor NF-κB. In addition to IRF-3, signal transduction mediated by TLR4 can activate NF-κB through two distinct mechanisms (Figure 1.4). Therefore, it was proposed that

differential usage of the TIR domain-containing adaptors by TLRs allows the induction of appropriate immune responses against a particular pathogen (O'Neill et al., 2003). This will be elaborated further in the following sections.

1.2.2 TIR domain-containing adaptors

The TIR domain-containing adaptor family is referred to as a group of cytoplasmic proteins with the presence of the signature TIR domain in their structure. To date, five members of this protein family have been described:

(a) MyD88, the product of the myeloid differentiation primary response gene 88

(b) TIR domain-containing adaptor protein (TIRAP) that is also known as MyD88-adaptor-like (Mal)

(c) TIR domain-containing adaptor inducting interferon-β (TRIF) or TIR-containing adaptor molecule-1 (TICAM-1)

(d) TRIF-related adaptor molecule (TRAM) or TICAM-2

(e) Sterile alpha and Armadillo motif-containing protein (SARM).

Except for MyD88, which is a "universal" adaptor recruited by all TLR homologs apart from TLR3, other TIR domain-containing adaptors are involved in the distinct TLR mediated signaling pathways (Takeda and Akira, 2005, and Figure 1.4). The functions of each TIR domain-containing adaptor in TLR signaling in response to infection and inflammation are being uncovered as a result of extensive research on TLR signaling, especially those based on mice that are deficient in individual or combination of the adaptor molecules. The following sections will review current understanding on the role of each TIR domain-containing adaptor in TLR signaling.

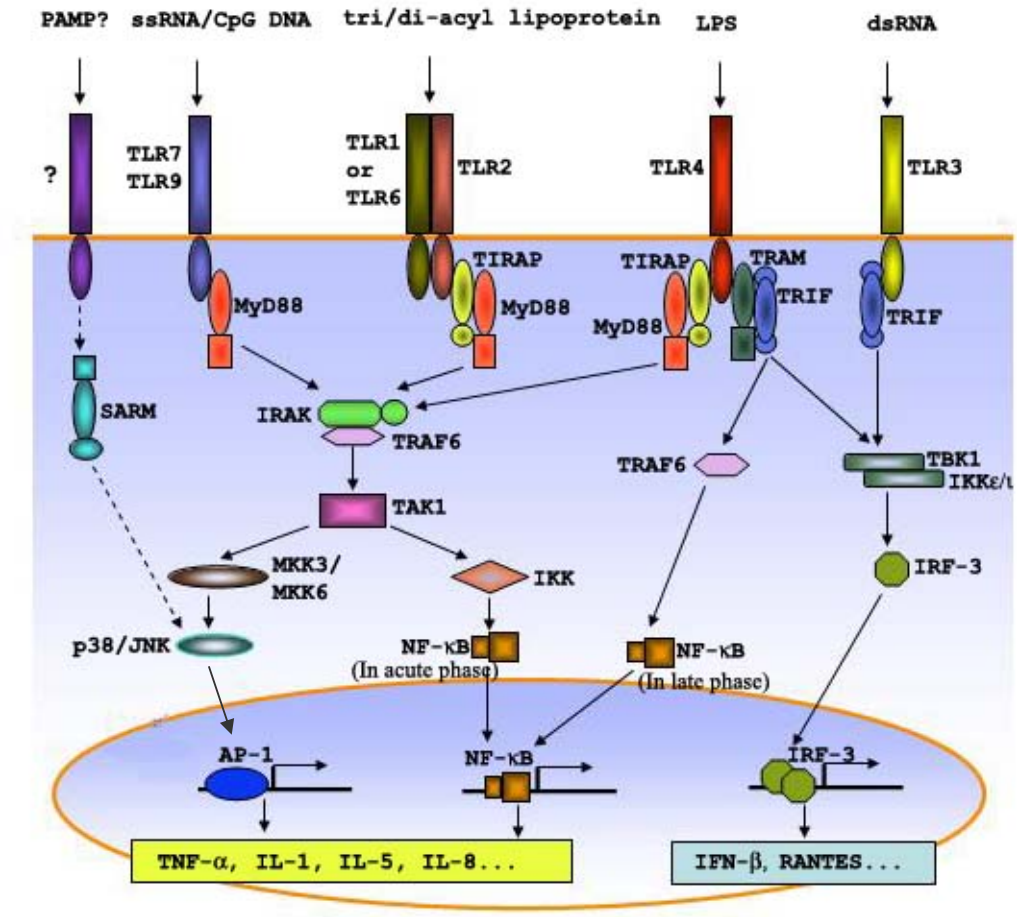


Figure 1.4 TLRs differentially use TIR domain-containing adaptors

Five TIR domain-containing adaptors, MyD88, TIRAP, TRAM, TRIF and SARM, are differently used by TLRs in order to mediate distinct gene expression profile according to the nature of immune challenge. Refer to the text for detail description of the pathways. Figure is adapted with modification from Takeda and Akira (Takeda and Akira, 2005). AP-1, activator protein 1; IKK, inhibitory κB kinase; IRAK, IL-1 receptor-associated kinase; IRF-3, interferon regulatory factor-3; MKK, mitogen-activated kinase kinase; NF-κB, nuclear factor-κB; SARM, SAM and ARM-containing protein; TIRAP, TIR domain-containing adaptor protein; TAK1, transforming growth factor-β-activating kinase; TBK1, TANK-binding kinase-1;TLR, Toll-like receptor; TRAM, TRIF-related adaptor molecule; TRAP-6, receptor-associated factor-6; TRIF, TIR domain-containing adaptor inducing interferon-β.

(a) MyD88

MyD88 is the first TIR domain-containing adaptor protein to be discovered and is also the most extensively studied. In addition to the C-terminal TIR domain, MyD88 harbors a death domain at its N-terminus. Researches have shown that all TLRs except for TLR3 are able to recruit MyD88 to their cytoplasmic TIR domain, leading to the activation of a signaling pathway that is analogous to IL-1 receptor-mediated pathway (Janssens and Beyaert, 2002). Following the engagement of MyD88 to TLR, MyD88 recruits IL-1 receptor-associated kinase-4 (IRAK-4) via the interaction of the death domains of both molecules. IRAK-4 then mediates the phosphorylation of IRAK-1, which then recruits and activates tumour necrosis factor (TNF) receptor-associated factor-6 (TRAP-6). Signaling downstream of TRAP-6 finally results in the activation and nuclear translocation of AP-1 and NF-κB transcription factors through the activation of p38 and c-Jun N-terminal kinase (JNK) and the inhibitory κB kinase (IKK) complex (McGettrick and O'Neill, 2004, and Figure 1.5). This MyD88 dependent signaling cascade is thus important for the production of inflammatory cytokines, whose expression is controlled by NF-κB and AP-1 transcription factor, such as interleukin (IL)-1, IL-6, IL-8, and TNF-α. In addition, the activation of p38 kinase by TLR can induce phagocytosis of the invading microbes probably through upregulating the expression of scavenger receptors (Doyle et al., 2004).

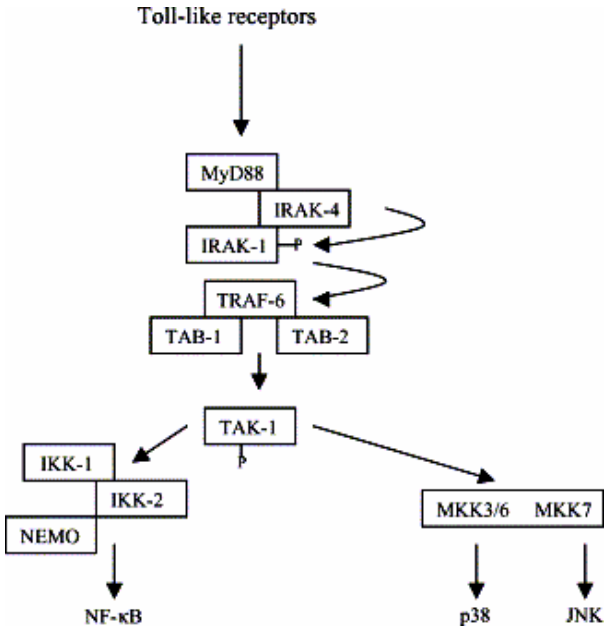


Figure 1.5 MyD88-dependent signaling pathway

TLRs except for TLR3 recruit MyD88 to activate the transcription factor NF-κB and MAPK signaling cascade that leading to the phosphorylation of p38 and JNK. The detail of this MyD88-dependent signaling pathway can be found in the text. Figure is adapted from McGettrick and O'neil (McGettrick and O'Neill, 2004)

(b) TIRAP/Mal

TIRAP/Mal was identified as a result of the database search for proteins which are structurally related to MyD88. Using mice deficient in either Mal or MyD88, researcher found that TIRAP and MyD88 work together to mediate signal transduction within the TLR2- and TLR4- signaling pathways (Horng et al., 2002; Yamamoto et al., 2002a). Subsequently, Dunne et al (Dunne et al., 2003) reported that these two adaptors can physically interact with each other and individually interact with either TLR2 or TLR4.

TRIF/TICAM-1

TRIF was described by Yamamoto et al (Yamamoto et al., 2002b) and Oshiumi et al (Oshiumi et al., 2003a), who named it as TICAM-1, as the TIR-containing adaptor molecule that is responsible for the MyD88-independent activations of IRF3 and NF-κB by TLR3. Indeed, TRIF was found to associate with TLR3 in coimmunoprecipitation (Yamamoto et al., 2002b) and yeast two hybrid studies (Oshiumi et al., 2003a). Mutagenesis of the gene encoding TRIF inhibits the activation of IRF3, and thus the induction IFN-β and IFN-inducible promoters, and NF-κB by TLR3 (Hoebe et al., 2003; Yamamoto et al., 2003a). In addition, TRIF was found to be involved in the MyD88-independent delay activation of NF-κB (Kawai et al., 2001) and the activation of IRF3 (Hoebe et al., 2003) by LPS, the ligand of the TLR4. TRIF-dependent activations of IRF3 and NF-κB were found to be associated with the recruitment of IKKs including TBK1 and IKK*i*/IKKε complex and TRAP-6 to the N-terminal of TRIF, respectively (Sato et al., 2003). In brief, TRIF is engaged in two distinct signaling pathways that are mediated by TLR-4 or TLR-3: one leading to the activation of IRF-3 transcription factor; and the other resulting in MyD88-independent activation of NF-κB.

(c) TRAM

Although TRIF was found to be a component of TLR4 signaling pathway that activates IRF3, its interaction with TLR4 was not observed (Yamamoto et al., 2002b). The discovery of TRAM provided insights to the missing link between TLR4 and TRIF. Indeed, TRAM was demonstrated to function in the activation of IRF3 and TRIF-dependent activation of NF-κB by acting as a bridge between TLR4 and TRIF (Fitzgerald et al., 2003; Oshiumi et al., 2003b; Yamamoto et al., 2003b).

(d) SARM

Although sterile alpha and Armadillo motif containing protein (SARM) was first described in human in 2001 (Mink et al., 2001), it has just been included into the TIR domain-containing adaptor family upon the discovery of a TIR domain within this molecule (O'Neill et al., 2003). SARM is an evolutionary conserved protein. Homologs of human SARM have been found in mouse, zebra fish, *Drosophila*, and *C. elegans* (Couillault et al., 2004; Jault et al., 2004; Meijer et al., 2004; Mink et al., 2001). Although varied in length, SARM homologs share a common domain architecture of an N-terminal Armadillo repeat (ARM) motif, followed by two sterile alpha (SAM) motifs, and a TIR domain located just C-terminal to SAM motifs (Couillault et al., 2004). The domain organization of SARM is shown in Figure 1.6. Interestingly, all three protein domains made of SARM function in the mediation of protein-protein interaction.

Figure 1.6 Schematic diagram of the domain organization of SARM

Figure is not to scale. ARM = Armadillo motif, SAM = Sterile alpha motif, TIR = Toll/Interleukin-1 domain.

The **Armadillo motif** is characterized by the tandem repeats of a conserved 42 amino-acid-long sequence (Peifer et al., 1994). Each repeat folds into three α helices. Bundles of helices of the multiple repeats composing ARM motif in turn, fold further, creating a regular structure of a right-handed superhelix (Coates, 2003). This is clearly demonstrated by the crystal structures of β-catenin (Huber et al., 1997) and importin-α (Conti and Kuriyan, 2000), which contain 12 and 10 tandem ARM repeats, respectively (Figure 1.7). ARM motif is found in proteins from wide variety of

eukaryotes ranging from uni- to multi-cellular animal. Examples of such proteins are listed in Table 1.2. ARM repeat-containing proteins have diverse functions including regulation of cytoskeleton; transportation of proteins between the cytosol and the nucleus; acting as the guanine nucleotide exchange factor; controlling of gene expression; and signaling (Table 1.2). The versatile functions of ARM repeat-containing proteins are explained by the fact that the superhelix structure of ARM motif provides surface for multiple protein interactions, promoting complex formation (Conti and Kuriyan, 2000; Huber et al., 1997). For example, the interaction of ARM repeats of β-catenin with cadherin is involved in the regulation of cytoskeletal functions (Aberle et al., 1994) while the recognition of nuclear localization signals by the ARM motif of the importin-α is essential for the transport of proteins through the nuclear pores (Conti and Izaurralde, 2001).

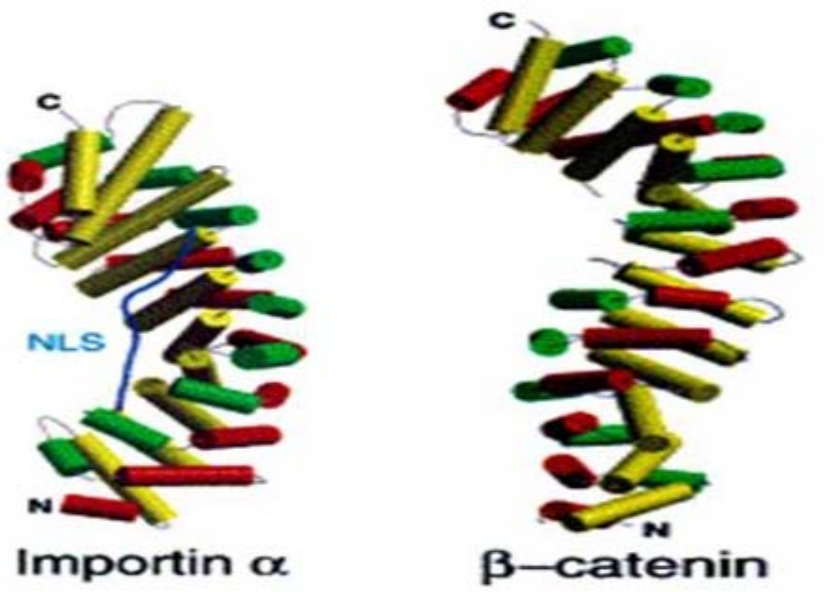
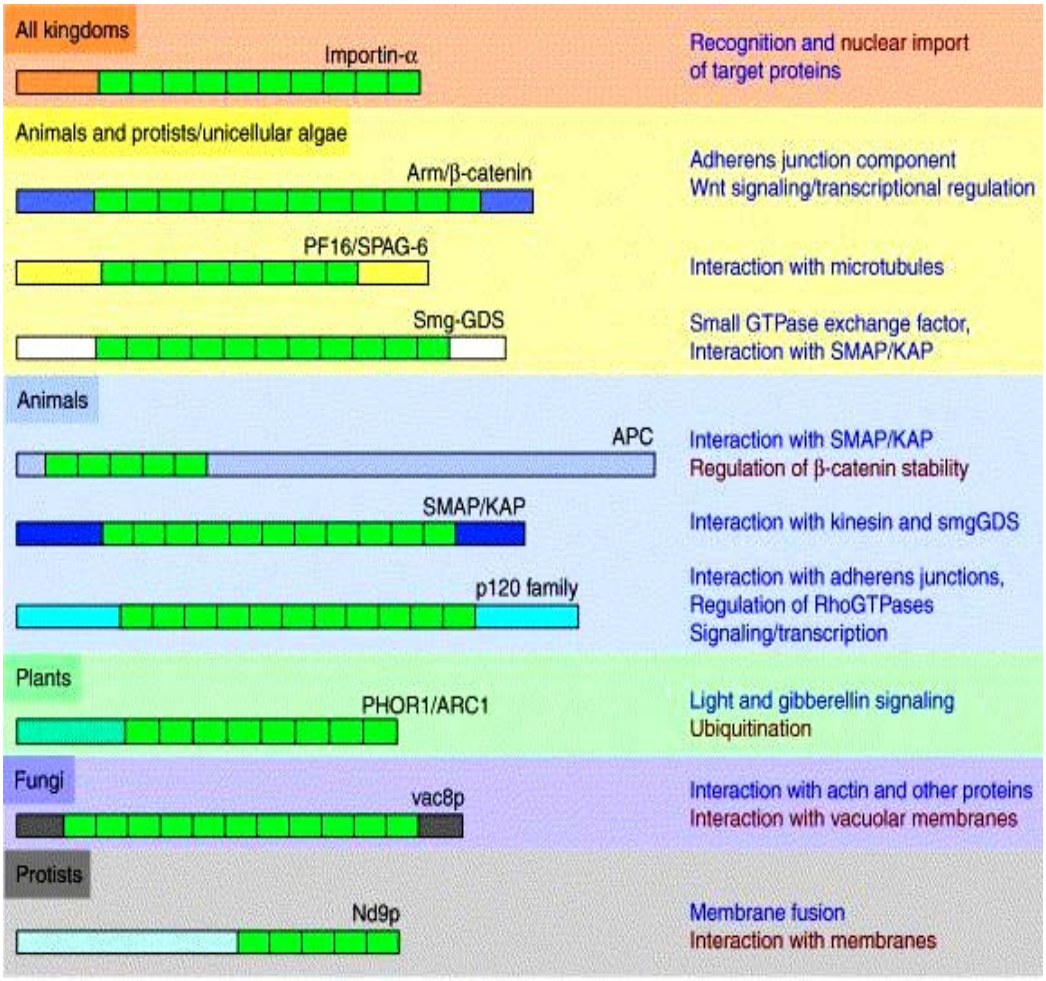


Figure 1.7 Three-dimensional structure of ARM repeat motif of importin-α and β-catenin.

The ARM repeat-motifs of *Saccharomyces cerevisiae* importin-α and mouse β-catenin have superhelix three dimensional structure. The three helices, termed H1, H2 and H3, of each ARM repeat are colored in green, red and yellow, respectively. Importin-α is shown as complex with the nuclear localization signal (NLS) of nucleoplasmin. Figure is adapted from Andrade et al (Andrade et al., 2001).

Table 1.2 ARM motif proteins have diverse origins and functions



Examples of ARM motif-containing protein families. Each ARM repeat is represented by one green box. Known functions of each protein are listed in the right with those functions associated with ARM motif written in blue and those are not are in red. Table is adapted from Coates (Coates, 2003)

The sterile alpha (SAM) domain is a conserved domain of approximately 70 amino acids (Ponting, 1995) that is present in almost 1000 proteins from a wide variety of eukaryotes and even some bacteria (Schultz et al., 1998). The functions of SAM domain-harboring proteins are extremely diverse, ranging from transcriptional/translational regulation, apoptosis to signal transduction (Kim and Bowie, 2003 and Table 1.3). The hallmark of this huge group of proteins is their ability to form polymeric complexes via: (1) the oligomerization of SAM domains of

the same type of protein as in the case of transcriptional repressor TEL (Kim et al., 2001); (2) the polymerization of SAM domains of different types of proteins for example the complex of Mae (modulator of the activity of Ets), and transcriptional regulators Yan and Pnt is formed due to the interaction between SAM domains of each of the proteins; (3) the interaction with proteins do not contain SAM domain such as those listed in Table 1.3.

Table 1.3 Example of interaction partners of SAM motif-containing proteins

SAM domain-containing protein	Binding partner	Biological role
Smaug	*nos* mRNA TCE	Translational repression
BAR	Bcl-2, Bcl-X_L	Apoptosis
Ets2	Cdk10	Transactivation
Ets2	Brg-1	Chromatin remodelling
LIP	LARPTP	Focal adhesion localization
EphB1	Grb10	RTK signaling
EphB1	LMPTP	RTK signaling
Eph (various)	AF6	RTK signaling
TEL	Ubc9	SUMOylation, localization

SAM domain-containing proteins bind variety of proteins and performed diverse functions. Table is adapted from Kim and Bowie (Kim and Bowie, 2003)

TIR domain, as described previously, is a conserved ~ 200 residue-long protein motif that is found in three protein families that play a role in immune response, namely IL-1 receptor family, Toll-like receptor family, and TIR domain-containing protein family. Interactions between the TIR domains of TLRs or between TIR domains of TLRs and the downstream TIR domain-containing adapters are critical for signal transduction leading to the activation of immune-related genes.

The conservation of SARM across the animal kingdom and the fact that mutation in the gene encoding the SARM homolog in *Drosophila* was lethal (Mink et al., 2001) suggest functional significance of this protein. However, unlike the other four TIR domain-containing adaptor proteins, the role of SARM in TLR signaling remains to be determined. Nevertheless, research in *C. elegans* strongly demonstrated

that SARM is essential for innate immune response against microbial infection (Couillault et al., 2004; Liberati et al., 2004). Both Couillault et al (Couillault et al., 2004) and Liberati et al (Liberati et al., 2004) reported that RNA-interference suppression of *tir-1* gene, which encodes TIR-1, the SARM homolog of *C. elegans* rendered the worm more sensitive to bacterial and fungal infection. Although the exact causes are still unknown, the increased susceptibility was speculated to be partially related to the reduction in the expression of NLP-31 antimicrobial peptide (Couillault et al., 2004) and/or the inhibition of activation of p38 mitogen activated protein kinase (MAPK) PMK-1 (Liberati et al., 2004). The linkage between TIR-1 with the p38 MAP kinase was also recently reported by Chuang and Bargamann (Chuang and Bargmann, 2005). The authors demonstrated that TIR-1 physically interacted with the *C. elegans* calcium-calmodulin-dependent protein kinase II (CaMKII), UNC-3, and probably its downstream target, the MAP kinase kinase kinase NSY-1, to regulate the expression of odorant receptors during the differentiation of olfactory neurons. NSY-1 is the homolog of mammalian apoptosis signal regulated kinase1 (ASK1) that is known to be able to activate p38 and JNK. As a result, it was hypothesized that the role of TIR-1 in neuronal development is to mediate signal transduction through the Ca^{+2}/MAPK cascade, which finally leads to the activation of the p38/JNK kinases. Nevertheless, further efforts should be undertaken to determine the implication of this signaling pathway in innate immunity. In addition, the unique combination of three protein-protein interaction modules in SARM suggests that it may be engaged in the more complicated signaling pathways that are distinct from that mediated by other TIR-containing adaptors.

1.3 Horseshoe crab is an excellent model for innate immunity research

Invertebrates serve as good models for the study of innate immunity for following reasons. First of all, the invertebrates rely solely on the innate immune system for protection against pathogen invasion. In the absence of adaptive immunity, the interpretation of experimental results is uninterrupted since the influence from the adaptive immune system to the innate immune responses is totally absent. Secondly, due to the evolutionary conservation of innate immune-related molecules, knowledge of the innate immunity in the invertebrates is very useful for the understanding of molecular mechanisms underlying the innate immune responses in the vertebrates.

Over the last two decades, a wide variety of invertebrates have been used for the studies of the innate immunity, examples of which are the threadworm, *Caenorhabditis elegans;* the tobacco hornworm, *Manduca sexta;* the silkworm *Bombyx mori*; the fruit fly, *Drosophila melanogaster*; the mosquito, *Anopheles gambiae*; the horseshoe crabs, *Tachypleus tridentatus, Limulus polyphemus* and *Carcinoscorpius rotundicauda*; and the Pacific oysters, *Crassostrea gigas*. Amongst these species, the *Drosophila* and *C. elegans* are the animal models of choice due to the availability of genome sequences that allows the high throughput genomic and proteomic analysis and ease of genetic manipulation (Royet, 2004). Indeed, studies in these organisms have greatly contributed to the understanding of innate immunity, especially the discovery of Toll and Toll signaling pathway in *Drosophila*. Horseshoe crab, however, is also a good model for innate immune study since it has much larger volume of blood and bigger tissues compared with most of other invertebrate models, allowing convenient physiological and molecular manipulations. In addition, this organism habours a very sophisticated innate immune system that ensures its survival for over 200 million years. This point will be elaborated further.

Horseshoe crab is a "living fossil"

The horseshoe crab belongs to the order Xiphosura that has more than 500 million year of evolutionary history. Evolution leading to the formation of modern horseshoe crab took place for about 200 million years from the Silurian period (~ 420 million years ago) to the Jurassic period (~ 200 million years ago). After that, it has remained largely unchanged until now (Stormer, 1952). Due to its long history of evolution, horseshoe crab is often referred as a "living fossil". Today, there are four species of horseshoe crabs in different habitats around the world: *Limulus polyphemus* in the East coast of USA; *Tachypleus tridentatus* in China and Japan and *Tachypleus gigas* and *Carcinoscorpius rotundicauda* in South Asia (Ding et al., 2005).

Horseshoe crab possesses a powerful innate immune system

In order to survive for more than 200 million years, the horseshoe crab has developed a powerful innate immune system to combat the pathogenic microorganisms, especially Gram-negative bacteria, the main infective agents in the marine environment. Indeed, Ng et al (Ng et al., 2004) demonstrated that the horseshoe crab, *C. rotundicauda*, survived an infection of 2 x 10^6 CFU of *Pseudomonas aeruginosa* / 100 g of body weight, a dose that was shown to be lethal to mice. The immune response was so fast and efficient that majority of the bacteria were cleared from the plasma after three hours of infection and the rest was completely removed finally (Figure 1.8)

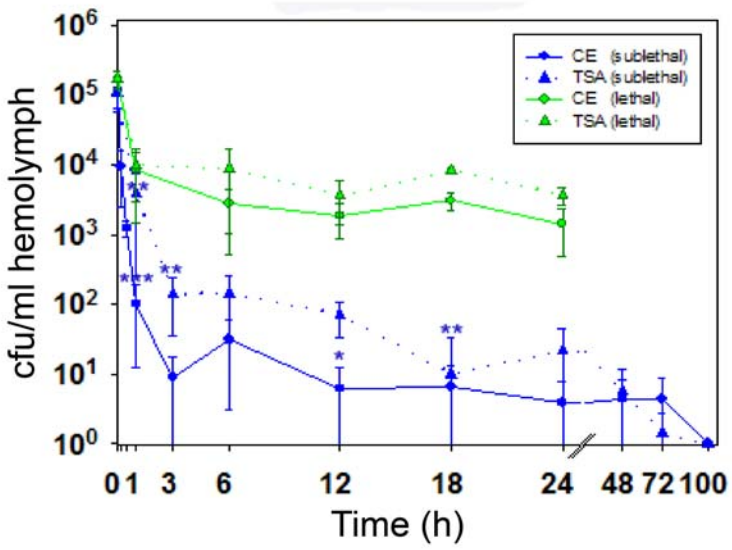


Figure 1.8 Horseshoe crab mounts a powerful immune response against *P. aeruginosa*

Horseshoe crabs were infected with a sublethal (2 x 10^6 CFU / 100 g of body weight) and a lethal dose (2 x 10^8 CFU / 100 g of body weight) of .*P. aeruginosa*. At different time points, crabs were bled and the hemolymph was plated onto TSA (dotted lines) and *Pseudomonas* selective cetrimide (bold lines) agar media to determine level of viable bacteria for the level of live bacteria. The significance difference between certain time point to the previous one indicates by the asterisk. One, two and three asterisk, represent significant differences with $P<0.05$, $P<0.01$ and $P<0.001$, respectively. Figure is adapted from Ng et al (Ng et al., 2004)

Research over the last 20 years has revealed that this organism harbours a large repertoire of firstline defense molecules including antimicrobial peptides, lectins, clotting factors, serine proteases, and protease inhibitors (Iwanaga, 2002; Ng et al., 2004; Yau et al., 2001; Zhu et al., 2005, and Table 1.4). Most of these proteins are identified either in the hemolymph or in the granules of the amebocyte, which is the major type of blood cells in the horseshoe crab. Many of these firstline defense molecules are involved in the Gram-negative endotoxin (LPS)-induced coagulation cascade leading to the formation of the clots, into which the intruders are entrapped and thus destroyed. The gelation of the blood starts with the activation of the protease cascade by the LPS, ultimately leading to the cleavage of coagulogen into coagulin, which are crosslinked to form the clot (Ding et al., 1993; Iwanaga and Kawabata, 1998). Pathogen killing is also achieved by the release of many potent antimicrobial

peptides from the amebocytes into the hemolymph (Iwanaga, 2002 and Figure 1.9). Recently, Ding *et al* (Ding et al., 2005) reported that not only the firstline defense proteins are essential for in immune responses against Gram-negative bacteria in the horseshoe crab, proteins functioning in apoptosis,stress responses and cell signaling are also involved. In addition, Ariki et al (Ariki et al., 2004) reported that factor C, in addition to play a role in coagulation, also serve as a PRR for Gram-negative bacteria endotoxin through the interaction with lipid A of the LPS. The presence of such broad spectrum of immune responsive molecules suggests that the molecular mechanisms of innate immune responses in the horseshoe crab are more complicated than what has been known so far. The existence of effector mechanisms other than what have been identified as well as cellular signaling pathways that are critical for innate immunity in the horseshoe crab remain to be uncovered and exploited.

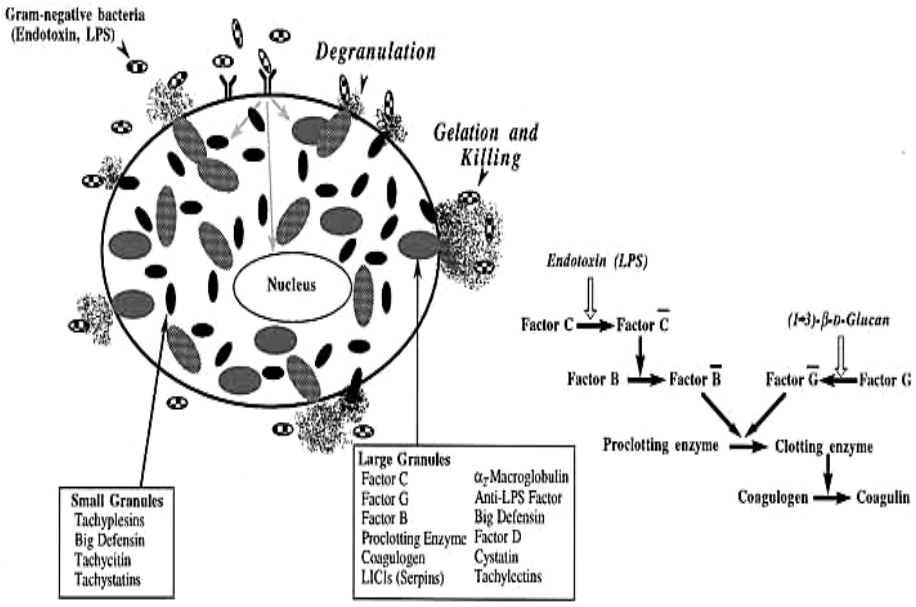


Figure 1.9 Amebocyte mediated immune responses against Gram-negative bacteria.

Upon contact with Gram-negative bacteria, amebocytes are degranulated, releasing into the hemolymph the contents of the large and the small granules, majority of which are antimicrobial peptides and clotting factors. Invading bacteria are trapped in the clots that are formed as the result of a cascade activated by Factor C upon its binding to the endotoxin and killed by the antimicrobial peptides. The figure is adapted from Iwanaga and Kawabata (Iwanaga and Kawabata, 1998)

Table 1.4 Innate immune molecules of the horseshoe crab

Proteins and peptides	Mass (kDa)	Function/specificity	Localization
Coagulation factors			
Factor C	123	Serine protease	L-granule
Factor B	64	Serine protease	L-granule
Factor G	110	Serine protease	L-granule
Proclotting enzyme	54	Serine protease	L-granule
Coagulogen	20	Gelation	L-granule
Protease inhibitors			
LICI-1 (*Limulus* intracellular coagulation inhibitor)	48	Serpin/factor C	L-granule
LICI-2	42	Serpin/clotting enzyme	L-granule
LICI-3	53	Serpin/factor G	L-granule
Trypsin inhibitor	6.8	Kunitz-type	ND
Limulus trypsin inhibitor	16	New type	ND
Limulus endotoxin-binding protein-protease inhibitor	12	New type	L-granule
Limulus cystatin	12.6	Cystatin family 2	L-granule
α2-Macroglobulin	180	Complement	Plasma & L-granule
Chymotrypsin inhibitor	10	ND	Plasma
Antimicrobial substances			
Anti-LPS factor	12	GNB	L-granule
Tachyplesins	2.3	GNB, GPB, Fungus	S-granule
Polyphemusins	2.3	GNB, GPB, Fungus	S-granule
Big defensin	8.6	GNB, GPB, Fungus	L & S-granule
Tachycitin	8.3	GNB, GPB, Fungus	S-granule
Tachystatins	6.5	GNB, GPB, Fungus	S-granule
Factor D	42	GNB	L-granule
Lectins			
Tachylectin-1	27	LPS (KDO), LTA	L-granule
Tachylectin-2	27	GlcNAc, LTA	L-granule
Tachylectin-3	15	LPS (O-antigen)	L-granule
Tachylectin-4	470	LPS (O-antigen), LTA	ND
Tachylectin-5	380-440	*N-acetyl group*	Plasma
Limunectin	54	PC	L-granule
*Limulus*18-kDa agglutination-aggregation factor	18	Hemocyte aggregation	L-granule
Limulin	300	Hemolytic/PC, PE, SA, KDO	Plasma
Limulus C-reactive protein (CRP)	300	PC, PE	Plasma
Tachypleus CRP-1	300	PE	Plasma
Tachypleus CRP-2	330	Hemolytic/PE, SA	Plasma
Tachypleus CRP-3	340	Hemolytic/SA, KDO	Plasma
Polyphemin	ND	LTA, GlcNAc	Plasma
Tachypleus tridentatus agglutinin	ND	SA, GlcNAc, GalNAc	Plasma
Liphemin	400-500	SA	Hemolymph
Carcinoscorpin	420	SA, KDO	Hemolymph
galactose-binding protein	40	Gal	Hemolymph
Protein A binding protein	40	Protein A	Hemolymph
(1 → 3)β-D-glucan binding protein	168	Pachyman, cardlan	Hemocyte
Others			
Transglutaminase (TGase)	86	Cross-linking	Cytosol
8.6 kDa protein	8.6	TGase substrate	L-granule
Pro-rich proteins (Proxins)	80	TGase substrate	L-granule
Limulus kexin	70	Precursor processing	ND
Hemocyanin	3600	O2 transporter/ Phenoloxidase	Plasma
Toll-like receptor (tToll)	110	ND	Hemocyte
L	11	Unknown	L-granule
L4	11	Unknown	L-granule

GNB, Gram-negative bacteria; GPB, Gram-positive bacteria; LPS, lipopolysaccharide; KDO, 2-keto-3-deoxyoctonic acid; PC, phosphorylcholine; PE, phosphorylethanolamine; SA, sialic acid;, LTA, lipoteichoic acid; ND, not determined. Talbe is adapted from Iwanaga and Lee (Iwanaga and Lee, 2005).

1.4 Aims and rationale of the project

The aim for the first part of this project was to identify proteins of the horseshoe crab, *C. rotundicauda,* that are differentially expressed in response to *P. aeruginosa* infection. With the isolation of the EST clone that is homologous to human SARM, a potential TIR domain-containing adaptor protein of the TLR signaling pathways, subsequent study focussed on the cloning and functional characterization of *C. rotundicauda* SARM, CrSARM.

P. aeruginosa is a ubiquitous Gram-negative bacterium that can be found (Hardalo and Edberg, 1997), in the coastal marine environment, the natural habitat of *C. rotundicauda* the animal model used in this project. *P. aeruginosa* has a wide range of host, from plants, insects to mammals. In addition, this bacterium has developed multi-resistance against various antibiotics (Stover et al., 2000). Altogether, these factors make *P. aeruginosa* the major cause of nosocomial infection and a significant opportunistic human pathogen. Study of the regulation of the expression of immune-related genes in response to *P. aeruginosa* infection, is thus beneficial not only for the understanding of the host-pathogen interactions but also for theurapeutic intervention.

Since the discovery of human SARM in 2001 (Mink et al., 2001) and subsequent identification of its homologs in other organisms including *Drosophila,* zebra fish and *C. elegans* (Couillault et al., 2004; Liberati et al., 2004; Meijer et al., 2004; Mink et al., 2001), the function of this signaling molecule has remained largely unknown. The study of SARM has been restricted to TIR-1, SARM homolog of *C. elegans*, which was demonstrated to be essential for immune response against bacterial and fungal infections (Couillault et al., 2004; Liberati et al., 2004) in addition to its role in neuronal development in the worm (Chuang and Bargmann,

2005). The immune-related signaling pathway that TIR-1 is involved in, however, is still undefined except for the fact that this molecule works upstream of the p38 kinase (Liberati et al., 2004). Our work on CrSARM, would thus provide insights into the biological relevance of this novel signaling protein, especially in the context of innate immunity against *P. aeruginosa* infection. In addition, the findings of this project would enrich the limited information about intracellular signaling that allows this "living fossil" to mount a very powerful frontline immune response against *P. aeruginosa* infection (Ng et al., 2004). Last but not least, we aim to contribute to the current knowledge on the TLR signaling, the crucial component of infection, host-pathogen interaction and the resulting immune response.

With the limitation of literature support, our approach to functionally characterize CrSARM was, firstly, to search for its interaction partners since the success of this would allow the speculation of potential signaling pathways that CrSARM is involved in. Further experiments were then designed accordingly to validate these hypotheses. To this end, the spatiotemporal expression of CrSARM under various hours post-infection by *P. aeruginosa* was investigated. Subsequently, yeast-2-hybrid screens were carried out to determine the interaction partners of CrSARM. The putative partners of CrSARM yielded very exciting possible explanations on the molecular mechanisms of signaling action of SARM in general.

Chapter 2 Materials and Methods

2.1 Materials

2.1.1 Organisms

Horseshoe crabs, *C. rotundicauda*, were collected from the Kranji estuary of Singapore. The specimens were washed to remove mud and acclimated overnight in 30% (v/v) sea water/fresh water before being used for experiments.

Bacteria strains used for infection experiments was *P. aeruginosa* ATCC 27853. *Escherichia coli* XL1-Blue was used as host bacteria for propagation of the phage cDNA library. *E. coli* BM25.8, which expresses Cre recombinase, was used as the host bacteria to facilitate the conversion of λTriplEx2 phagemid to pTriplEx2 plasmid.

2.1.2 Biochemicals and enzymes

Glutathione Sepharose^{TM} 4B, Protein A Sepharose 4 Fast Flow, Redivue [^{32}P]-dCTP, [^{35}S]-methionine, Hybond-N^{+} nylon membrane, and Hybond PDVF membrane were products of GH Healthcare. Deoxyadenosine triphosphate (dATP), deoxyguanosine triphosphate (dGTP), deoxycytosine triphosphate (dCTP) and deoxythymidine triphosphate (dTTP) were from Promega. X-α-gal was from BD Biosciences Clontech. Colony/Plaque Screen^{TM} nylon membrane was a product of Perkin-Elmer. Isopropyl-1-thio-β-D-galactopyranoside (IPTG) was from Biorad. Zymolyase was from Seikagaku Corporation. All restriction enzymes were from New England Biolabs or Fermentas. Mouse monoclonal anti-HA antibody and Complete Cocktail Inhibitors were from Roche. SuperSignal® West Pico Chemiluminescent Substrate was from Pierce. Mouse monoclonal anti-c-myc antibody was a product of Invitrogen. Phenylmethylsufonylfluoride (PMSF), bovine serum albumin fraction V,

calf thymus DNA, Ficoll 400 were from Sigma. Super RX X-ray film was from Fuji. Common chemicals were from Sigma and Merk.

2.1.3 Medium and agar

Luria Bertani (LB) broth was prepared with 1.0 % sodium chloride, 1 % (w/v) Tryptone (Difco) and 0.5 % (w/v) Yeast extract (Difco) at pH 7.0. For LB agar, 1.5 % (w/v) agar (Difco) was added.

Tryptone soy broth (TSB) was made by dissolving premix formula (Oxoid) consisting of 0.5 % (w/v) sodium chloride, 1.7% (w/v) pancreatic digest of casein, 0.3 % (w/v) papain digest of soybean meal, 0.25 % (w/v) di-potassium hydrogen phosphate and 0.25 % (w/v) glucose to 3 % (w/v) with deionised water. 1.5 % (w/v) agar (Difco) was added for preparation of tryptic soy agar (TSA).

SD medium was prepared with 0.67 % (w/v) yeast nitrogen base without amino acid (Difco), 2 % (w/v) glucose and appropriate premix amino acid supplements (BD Biosciences) at pH 5.8. SD medium without tryptophan (SD/-Trp) was made with 0.074 % (w/v) Trp Drop Out (DO) supplement. Similarly, SD/-Leu, SD /-Trp/-Leu, and SD/-Trp/Leu/His/Ade media were made with 0.069 % (w/v) Leu, 0.064 % (w/v) Trp/Leu and 0.06 % (w/v) Trp/Leu/His/Ade DO supplement, respectively. Whenever needed, SD medium was also supplemented with an additional 0.003 % (w/v) adenine hemisulfate. For SD agar, 2 % (w/v) agar (Difco) was added.

YPD medium was made with 2 % (w/v) peptone, 1 % (w/v) Yeast extract, 2 % (w/v) glucose at pH 6.5. YPDA was prepared by adding 0.003 % (w/v) adenine hemisulfate to YPD medium. For YPD and YPDA agar, 2 % (w/v) agar was added.

Whenever required, ampicillin and kanamycin were included in the medium or agar at concentrations of 100 mg/l and 30 mg/l, respectively, unless otherwise stated. For IPTG/X-gal plates, 0.05 % (w/v) IPTG and 0.04 % (w/v) X-gal were included.

For X-α-gal plates, 100 µl of a 2 mg/ml stock, which was made by reconstituting the powder in dimethylformamide, was spread per plate.

2.2 Challenging horseshoe crabs with *Pseudomonas aeruginosa*

2.2.1 Preparation of *P. aeruginosa* for infection of horseshoe crabs

A single clone of *P. aeruginosa* was inoculated into 10 ml of TSB and cultured overnight at 37 °C with shaking at 230 rpm. Bacteria was collected by centrifugation at 5000 x *g* for 10 min at 4 °C, washed with 10 ml of saline (0.9 % NaCl) and resuspended to the original culture volume in saline. An aliquot of 100 µl of this *Pseudomonas* suspension was serially diluted for enumeration and the rest was stored at 4 ^{o}C as stock culture for injection of the horseshoe crabs. Aliquots of 100 µl of 10^6, 10^7, and 10^8 dilutions were plated on TSA agar and incubated at 37 ^{o}C. The next day, bacteria was enumerated and the cell density of the 4 ^{o}C stock culture was calculated accordingly.

2.2.2 Challenging horseshoe crabs with bacteria

Horseshoe crabs were infected with *P. aeruginosa* by intracardial injection. Before that, crabs were cleaned and weighed. The septum, which was the site of injection was cleaned gently with detergent and disinfected with 70 % ethanol. To ensure that the horseshoe crab was free of infection before the treatment, a few drops of hemolymph was collected by cardiac puncture prior to bacterial challenge. The hemolymph was then plated on TSA plates. Only samples collected from microbial-free crabs were used for subsequent analysis. A sub-lethal dose, as established by Ng (Ng, 2005), of $1.2x10^6$ cfu *P. aeruginosa*/100 g body weight was used to challenged each horseshoe crabs. The volume of bacteria that was administered to each crab was

adjusted to 200 µl with saline before injection. Samples collection was performed at 1, 3, 6, 9, 12, and 24 hour post infection (hpi).

2.2.3 Collection of amebocytes and other tissues

Uninfected, which is called naïve hereafter, or infected horseshoe crabs were bled for the collection of the amebocytes. The hemolymph was diluted with the same volume of anti-coagulation solution (3 % NaCl, 0.125 % N-ethylmaleimide) that was prewarmed to 42 ^{o}C. The mixture was then incubated at 42 ^{o}C for 8 min and centrifuged at 150 x g for 10 min. The supernatant was discarded and the amebocytes were snap-frozen in liquid nitrogen before storage at -80 ^{o}C until use.

Hepatopancreas, heart, muscle, intestine, and stomach were obtained by dissection. Once removed from the horseshoe crabs, the tissues were immediately frozen in liquild nitrogen and stored at - 80 ^{o}C until use.

2.3 Isolation of RNA

2.3.1 Preparation for RNA purification

RNA was required for the synthesis of cDNA that was used for cloning CrSARM, analysis of tissue distribution and transcript profiling of CrSARM by RT-PCR. Since RNA is extremely sensitive to RNase contamination, maintenance of RNase free condition is very critical for the extraction of good quality RNA. Water and solutions were treated with 0.01% of diethylpyrocarbonate (DEPC) overnight at 37 °C and autoclaved for 2 h to remove residual DEPC. Metal and glass apparatus were baked at 200 ^{o}C for 4 h. Plastic-wares that were not disposable, were soaked in 3 % hydrogen peroxide overnight, rinsed thoroughly with DEPC treated water and autoclaved for 2 h.

2.3.2 Extraction of total RNA

Tissues were homogenized in Trizol® reagent (Invitrogen) at the ratio of 1 ml of Trizol per 100 mg of tissue. Homogenized samples were left at room temperature for 5 min so that the nucleoprotein complexes dissociated completely. Chloroform (1/5 volume of the homogenate) was then added and the mixture was shaken vigorously for 15 s, and incubated at room temperature for 2 min. Next, the mixture was centrifuged 1000 x g for 15 min at 4 ^{o}C to facilitate phase separation. The top colorless aqueous phase, which contained RNA, was carefully transferred to a fresh RNase-free tube. RNA was precipitated by incubation with isopropanol (0.5 ml/ 1ml of Trizol) for 10 min at room temperature. RNA precipitate was collected by centrifugation at 12, 000 x g for 10 min at 4 ^{o}C. The gel-like RNA precipitate was washed with 75 % ethanol, and pelleted by spinning at 14, 000 x g for 5 min at 4 ^{o}C. RNA pellet was air-dried for 30 min and stored at – 70 ^{o}C.

The RNA pellet was redissolved in 20 µl of water and pipetted gently for 3 times and incubated at 60 ^{o}C for 10 min. The concentration of the isolated RNA was determined by OD reading at 260 nm. To check for protein contamination, RNA samples were buffered with 10 mM Tris-HCl (pH 7.0) since stable pH is essential for an accurate measurement of A_{260}/A_{280} ratio. The RNA preparations had acceptable quality when the A_{260}/A_{280} ratio is in the range of 1.9 to 2.1.

2.3.3 Isolation of messenger RNA

mRNA was isolated from total RNA using Oligotex® kit (Qiagen). Approximately 250 µg of total RNA was used for each preparation. The volume of total RNA was adjusted to 250 µl with water. After this, 250 µl of binding buffer (20 mM Tris-HCl, pH 7.5, 1 M sodium chloride, 2 mM EDTA, 0.2 % SDS) and 15 µl of oligotex suspension, which was pre-warmed at 37 °C, were added to the total RNA

solution. The sample was incubated at 70 °C to denature the secondary structure of the RNA. mRNAs were then allowed to anneal to the Oligotex beads by incubation at room temperature for 10 min. Next, the oligo-dT beads were collected by centrifugation for 2 min at 14,000 x g and resuspended in 400 µl of wash buffer (20 mM Tris-HCl, pH 7.5, 1 M sodium chloride, 2 mM EDTA, and 0.2 % SDS). The suspension was then loaded onto a spin column, which was placed inside a collection tube. The solution was removed by centrifugation at 14,000 x g for 1 min. The column was transferred to a new collection tube and the beads were washed one more time with 400 µl of wash buffer. The column was transferred to another fresh tube for elution of the mRNA. Elution was done twice in order to dissociate maximum amount of mRNA, which was bound to the oligodT beads. Each time, 50 µl of 70 °C-warmed elution buffer (5 mM Tris-HCl, pH 7.5) was added to the column. The beads were quickly resuspended by repeated pipetting for four times and the eluted mRNA was collected by centrifugation at 14,000 x g for 1 min. The second eluate was collected in the same tube with the first one. Since the concentration of mRNA in the eluate was low, mRNA was next precipitated with 2.5 volume of 95 % ethanol and 1/10 volume of 3M sodium acetate. Precipitation was done at -20 ^{o}C overnight. mRNA pellets were recovered by centrifugation at 14, 000 x g for 30 min at 4 ^{o}C, washed with 500 µl of 70 % ethanol, and air-dried. mRNA was redissolved in 10 µl of water and stored at – 20 °C until use.

2.4 Cloning of full-length CrSARM cDNA

2.4.1 Isolation of expressed sequence tag (EST) that encodes CrSARM from the amebocyte subtractive cDNA library

In order to identify and isolate genes from *C. rotundicauda* that are differentially expressed upon microbial infection, our lab had constructed four subtractive cDNA libraries from the two major immune responsive tissues in *C. rotundicauda*: amebocyte and hepatopancreas. These cDNA libraries were generated using suppression subtractive cDNA hybridization technique. The experiment was performed as described in Ding et al (Ding et al., 2005). Briefly, horseshoe crabs were challenged with *P. aeruginosa* or mock-infected with saline. for three or six hours. These time points were chosen based on the report that the acute-phase expression of antimicrobial genes occurred in 3-6 hpi in *Drosophila* (Lemaitre et al., 1997). Tissues were isolated from both treatments i.e. *Pseudomonas* infection and mock-infection. For each treatment, samples of the two time points were combined. mRNAs were then isolated and used for double stranded cDNA preparation using SMART cDNA synthesis kit (Clontech). The double stranded cDNAs were then digested with *Rsa*I, which yields blunt-ended cDNA fragments. Next, suppression subtractive cDNA hybridization was carried out using PCR-select cDNA subtraction kit (Clontech, Figure 2.1). The tester cDNA, which contained differentially expressed transcripts, was divided into two equal potions. Each aliquot was ligated to either adaptor-1 or adaptor-2, briefly denatured, and separately hybridized with the denatured driver cDNA, which serves as the reference for the subtraction, with the ratio of 1 tester : 40 driver. The two hybridization mixtures were then combined before freshly denatured driver cDNA was added to enrich the differential expressed transcripts. Subsequently, secondary hybridization mixture was subjected to PCR using primers that are complement to adaptor-1 or adaptor-2. Since only the cDNA of the differential

transcripts harboured different adaptors on each end, they were to selectively amplify, and cloned into pGEM-T Easy vector (Promega).

Using the above principle, genes that were differentially upregulated or downregulated in *C. rotundicauda* amebocyte and hepatopancreas in response to *P. aeruginosa* infection were clustered into amebocyte-/hepatopancreas-forward and amebocyte-/hepatopancreas-reverse libraries, respectively. By comparing sequences of cDNA clones from these substractive libraries with known gene sequences deposited in the GeneBank, a broad spectrum (more than 400 non-redundant ESTs) of microbial infection responsive genes from the horseshoe crabs was revealed.

In this study, an EST clone (AmeR209) that was homologous to human SARM (GeneBank accession number: AAR17520), was identified and isolated from the amebocyte reverse library. It is henceforth referred to as the *C. rotundicauda* SARM (CrSARM). Conserved Domain Database Blast identified a TIR domain within the amino acid squence encoded by AmeR209. The CrSARM TIR domain showed high homology to TIR domain located at the C-terminal of SARM homologs from other organisms. Subsequently, this EST, which was 617-nucleotide long, was used as a probe to screen the phage amebocyte cDNA library and as a template to design primer for rapid amplification of cDNA ends (RACE) PCR in order to obtain the full-length cDNA sequence of CrSARM.

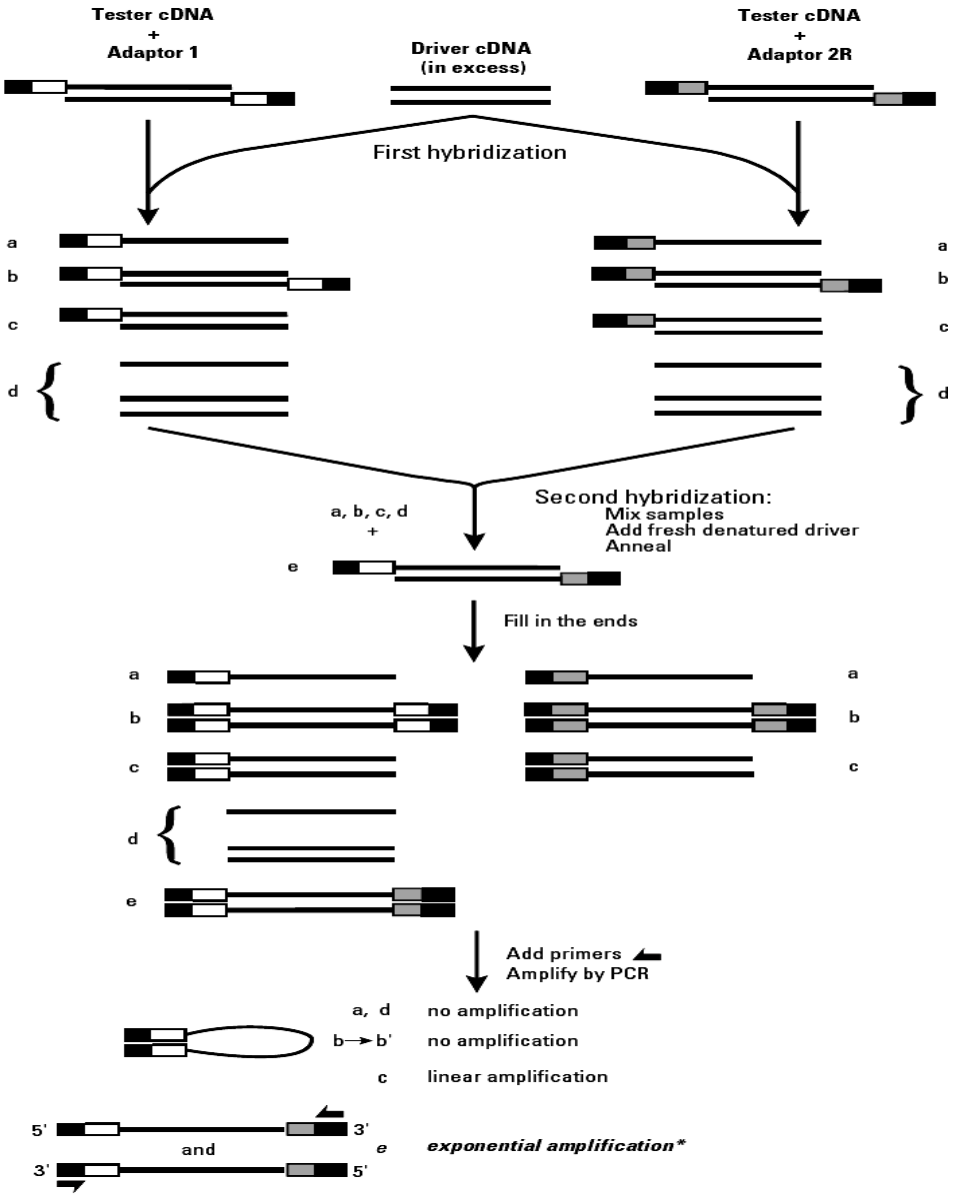


Figure 2.1 Supression subtraction cDNA hybridization

Figure is adapted from the user manual for PCR-select cDNA subtraction kit (Clontech)

2.4.2 Cloning of 3' end of CrSARM by phage cDNA library screening

Screening of phage amebocyte cDNA library using EST clone AmeR209 as a probe was one approach to clone full-length cDNA of CrSARM. This library was constructed using SMART cDNA Library Construction kit (Clontech). The cDNA was synthesized from the RNA of the amebocytes collected from the uninfected

horseshoe crabs, cloned into λ TriplEx2™ phagemid vector, and packed in λ phages using Gigapack® III Gold Packaging Extract (Stratagene). The titer of this library was determined to be 0.5 x 10^9 pfu/ml. Screening process is summarized in Figure 2.2.

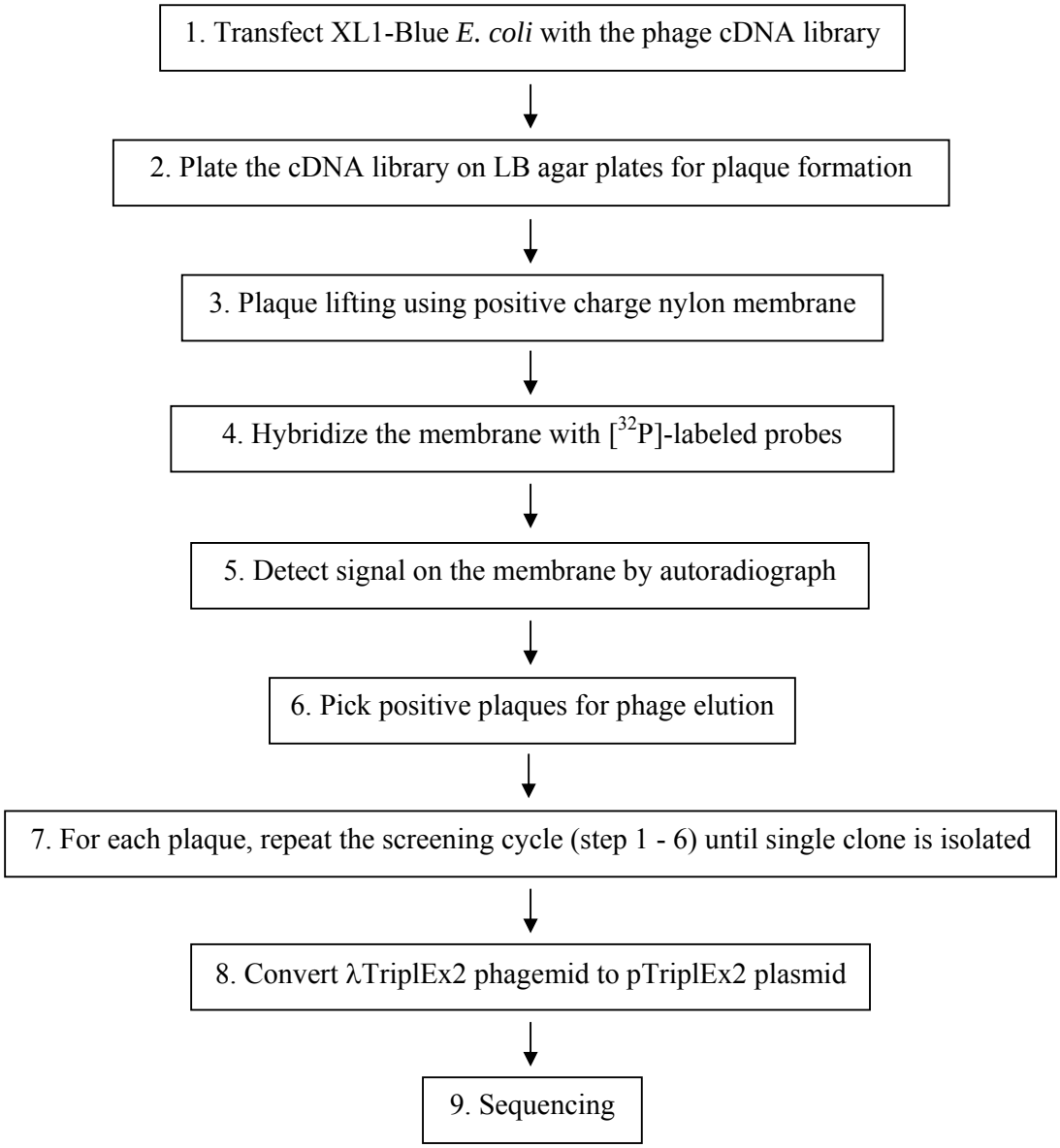


Figure 2.2 Screening of the phage cDNA library

Preparation of [^{32}P]-labeled probe

AmeR209 was released from pGEM-T Easy plasmid by restriction digest with restriction enzyme *Not*I and *Eag*I. The released fragment was separated from the linearised vector by eletrophoresis on 1.2 % agarose gel and purified using QIAquick gel extraction kit (Qiagen). The DNA fragment was then used as template to synthesize [^{32}P]-labeled probes using Rediprime II Random Prime Labelling system (Amersham Biosciences). In brief, 25 ng of DNA was diluted to 45 µl in TE buffer (10 mM Tris-HCl, 1 mM EDTA, pH 8.0), denatured for 5 min at 95 ^{o}C and cooled on ice for 5 min. After this, denatured DNA and 50 µCi of Redivue [^{32}P] dCTP were added to the reaction mixure that contained exonuclease free Klenow enzyme, dATP, dGTP, dTTP and random primers. The tube was then incubated at 37 ^{o}C for 1.5 h. When labeling reaction was completed, free nucleotides were removed by Qiaquick nucleotide removal kit (Qiagen). The radioactive labeled probe was stored at 4 ^{o}C until use.

Screening the phage amebocyte cDNA library

(a) Preparation of host bacteria

E. coli XL1-Blue was used as host bacteria to which the phage amebocyte cDNA library was transfected. To prepare for transfection, a single colony of *E. coli* XL1-Blue was inoculated into 25 ml of LB-broth containing 0.2 % maltose and 10 mM $MgSO_4$. The culture was incubated at 37 ^{o}C with 200 rpm shaking until OD_{600} reached 1.0. Cells were collected by centrifugation at 500 x g for 10 min, resuspended in 10 mM $MgSO_4$ to OD_{600} of 0.5. These cells were stored at 4 ^{o}C and used for transfection within 2-3 days.

(b) Plating the phage amebocyte cDNA library for primary screening

One million phages were plated on twenty 150 mm LB plates for primary screening. Each plate was prepared by incubating 50,000 phages with 600 μl of *E. coli* XL1-Blue at 37 ^{o}C for 15 min. Next, 5 ml of the top agarose (LB + 0.07 % agarose), which was prewarmed to 50 ^{o}C, was added and the tube was inverted several times for mixing. The phage/bacteria mixture was then quickly poured onto 37 ^{o}C-pre-warmed LB plate. The plate was swirled gently so that the top agarose spread evenly. The plates were incubated at 37 ^{o}C until the plaques were visible (around 12 h), then chilled to 4 ^{o}C for 2 h before the plaques were transferred onto nylon filters.

(c) Plaque lifting

Colony/Plaque Screen^{TM} (Perkin-Elmer) membranes, a positively-charged nylon filter, were used for plaque lifting. The membrane was carefully overlaid onto the plaque surface for 1 min. The position of the membrane was marked by stabbing a needle through it and the agar at three asymmetric positions. After that, the membrane was removed from the plate. Then, with the plaque side up, the membrane was immersed successively in denaturation solution (0.5 N NaOH, 1.5 M NaCl) for 2 min, neutralization solution (1.5 M NaCl, 0.5 M Tris, pH 7.4) for 5 min, washed twice, each time with 2X SSC (300 mM NaCl, 30 mM Na_3-citrate) for 30 s, then dried on 3MM Whatman paper. Subsequently, the DNA was fixed to the filter by UV-crosslinking.

(d) Hybridization of [^{32}P]-labeled probe to the filter

The membrane was moistened in 2 x SSC for 5 min then placed in the hybridization bottle (Hybaid). For prehybridization, the membrane was rotated in 50 ml of prehybridization buffer (5 x SSC, 1 % SDS, 100 μg/ml calf thymus DNA, 0.1 % Ficoll 400, 0.1 % bovine serum albumin (BSA), and 0.1 % polyvinylpyrrolidone) at

65 ^{o}C for 2 h. The buffer was then replaced by 10 ml of prewarmed, fresh prehybridization buffer. Next, [^{32}P]-labeled probe, which was denatured for 10 min at 100 ^{o}C and then snap-chilled on ice for 5 min, was added. Hybridization was carried out at 65 ^{o}C overnight. The membrane was then subjected to successive stringent wash with 2 x SSC, 0.1 % SDS at room temperature for 20 min; 1 x SSC, 0.1 % SDS at 37 ^{o}C for 20 min; 0.2 x SSC, 0.1 % SDS at 42 ^{o}C for 15 min; and 2 x SSC at room temperature for 15 min. After washing, the membrane was exposed to X-ray film (Fuji) and developed using Kodak X-Omat 3000 RA autoradiographic processor.

(e) Screening for and procurement of positive plaques

The location of positive plaques was determined by lining up the film with the respective plate. The putative plaques were removed from the plate by using the large end of the yellow pipette tip to cut through the agar. Next, the agar piece was placed in 1 ml of SM buffer (50 mM Tris-HCl, pH 7.5, 100mM NaCl, 8mM $MgSO_4$, 0.01% gelatin) and 20 µl of chloroform, vortexed vigorously, and incubated at 4 ^{o}C overnight for phage elution.

(f) Secondary screening

Due to the high density of plaque forming units on the primary screening plates, it was necessary to perform a second round of screening in order to authenticate the positive plaques as the single clones. Phage elution from primary screening was diluted 1000 times in SM buffer to ensure the plaques would be well-separated. An aliquot of 3 µl of the diluted phage was used to infect 600 µl of the host bacteria as described before. Secondary screening was carried out similarly as the primary screening. Putative plaques were picked using the small end of the yellow pipett tip and eluted in 500 µl of SM buffer and 20 µl of chloroform.

(g) Conversion of λTriplEx2 phagemid to pTriplEx2 plasmid

In order to determine the sequence of the cDNA harboring by the putative plaques, conversion of λTriplEx2 phagemid to pTriplEx2 plasmid was necessary to facilitate the stable propagation in *E. coli*. This conversion involved *in vivo* recombination of the loxP sites that flanked the pTriplEx2 plasmid, which was embedded inside the λTriplEx2 phagemid, resulting in the release of the plasmid. *E. coli* BM25.8 was used as the host bacteria since it expressed Cre recombinase that was essential for the recombination event (Figure 2.3). The single clone of *E. coli* BM25.8 was cultured at 31 ^{o}C in 10 ml of LB broth with shaking at 150 rpm. When the OD_{600} reached 1.1-1.4, 10 mM of $MgCl_2$ was added to the culture. Next, 150 µl of eluted phages were combined with 200 µl of host cell culture. The phage/host mix was incubated at 31 ^{o}C for 30 min without shaking. Another aliquot of 400 µl of LB broth was then added before the mixture was incubated for 1 h at 31 ^{o}C with shaking at 230 rpm. After that, 10 µl of the infected cell suspension was spread on an LB-ampicillin (100 µg/ml). The culture was incubated overnight at 31 ^{o}C and subsequently, a single colony was isolated from the plate, grown overnight in LB-ampicillin broth, and subjected to plasmid isolation using *Wizard plus* SV miniprep (Promega). The insert cDNA was sequenced using Lamda TriplEx2 sequencing primers (Clontech) in BigDye Terminator v3.1 (Applied Biosience) sequencing reaction with an ABI Prism 3100A model sequencer. The sequence was analyzed with DNAMAN version 4.15 software. Vector sequence was manually removed from the insert for alignment with the AmeR209 EST sequence.

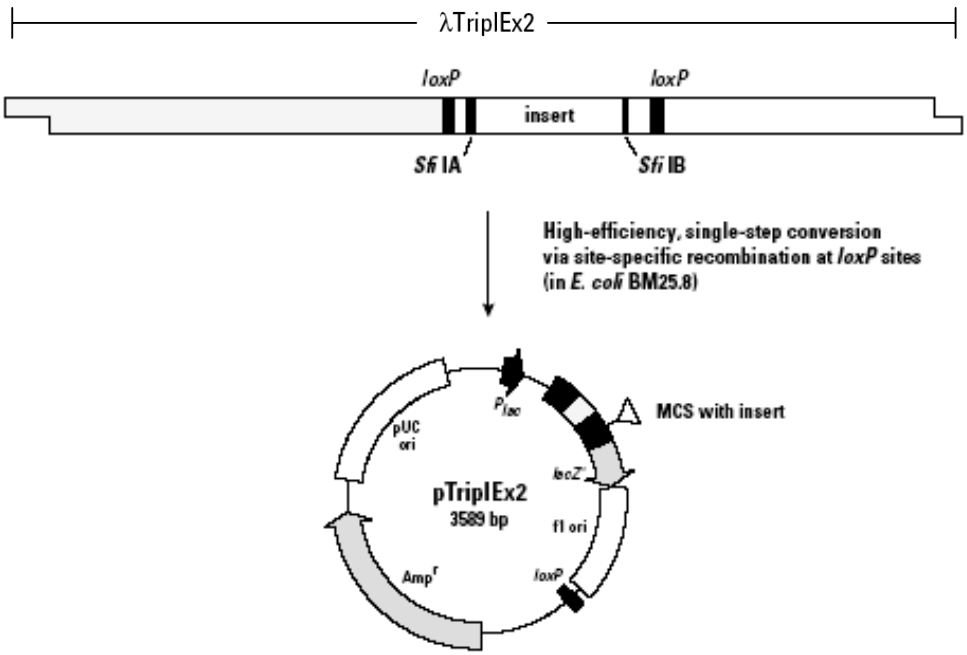


Figure 2.3 Conversion of λTriplEx2 phagemid to pTriplEx2 plasmid.

Figure is adapted from instruction manual for SMART cDNA library construction kit (Clontech)

2.4.3 Cloning of 5' end cDNA by RACE PCR

The 5'end of CrSARM was obtained using BD SMART™ RACE cDNA amplification kit (Clontech). To generate first strand cDNA for 5'-RACE PCR reaction, a 10 μl reaction mixture was set up with 2.0 μg of total amebocyte naive RNA, 12 pmols of 5'-CDS primer (5'-$(T)_{25}$VN-3'), 12 pmols of BD SMART II A Oligo (5'-AAGCAGTGGTATCAACGCAGAGTACGCGGG-3'), BD PowerScript reverse transcriptase, 2 mM dithiothreitol (DTT) and 1 mM dNTP mix in 1 x First strand buffer (50 mM Tris-HCl, pH 8.3, 6 mM MgCl, 75 mM KCl). cDNA synthesis was carried out at 42 ^{o}C for 1.5 h, followed by 72 ^{o}C for 7 min. The reaction mixture was then diluted in 50 μl of water and stored at – 20 ^{o}C until use.

RACE PCR was performed with the gene-specific reverse primer, TIR-5RACE (5'-CGGAAGACCAACTGGAAGAGGTCTGTGG-3'), which was complementary with the sequence close to the 5' end of the AmeR209 clone (Figure 3.1-B), and the universal primer mix (UPM), which was the mixture of the long (5'-CTAATACGACTCACTATAGGGCAAGCAGTGGTATCAACGCAGAGT-3') and the short (5'-CTAATACGACTCACTATAGGGC-3') oligonucleotides. The underlined sequence of the long oligonucleotide is complementary to the sequence introduced at the 5'-end of the first strand cDNA product by BD SMART II A Oligo primer. An aliquot of 50 µl of the 5'-RACE PCR was set up with 2.5 µl of first strand cDNA, 1 µl of 10 mM TIR-5RACE primer, 5 µl of 10 x UPM (0.4 mM of long and 2.0 mM of short oligonucleotide), 0.2 mM dNTP mix, and 1 µl of 50 x BD Advantage 2 Polymerase Mix (BD TITANIUM Taq, 1 µg/ml TaqStart Antibody, and small amount of proof reading polymerase) in 1 x Advantage 2 PCR Buffer (40 mM Tricine-KOH, pH 8.7, 15 mM KOAc, 3.5 mM $Mg(OAc)_2$, 3.75 µg/ml BSA, 0.005% Tween 20, 0.005% Nonidet-P40). Standard PCR program of one cycle at 95 °C for 30 s, followed by 30 cycles of denaturation at 95 °C for 10 s, annealing at 65 °C for 30 s and extension at 72 °C for 3 min, and termination with one cycle at 72 °C for 10 min, was performed. The PCR product was loaded on a 1.2 % agarose gel. A band at ~ 2.5 kb was isolated from the gel using QIAquick gel extraction kit (Qiagen). Before being cloned into pGEM-T Easy (Promega), the DNA fragment, which was blunt-ended due to the presence of proof reading polymerase in the Advantage polymerase mix, was subjected to A-tailing. Gel-purified DNA was concentrated to 7 µl using a speedvac, then, combined with 0.2 mM dATP, 5 U of Taq DNA polymerase, and 1 x Thermopol buffer (NEB) in 10 µl reaction mix. The tube was incubated at 70 ^{o}C for 30 min. Subsequently, 3.0 µl of A-tailing products was cloned into 50 ng of pGEM-T

Easy vector using 3 U of T4 DNA ligase and Rapid Ligation buffer (Promega). Ligation was carried out at 4 ^{o}C, overnight. The ligation mixture was then transformed into 100 μl of competent *E. coli* Top10 and spread on LB-Xgal-IPTG-ampicillin plates for blue/white colony selection. White colonies were selected for colony PCR with T7 and SP6 primers, which are complement to sequences flanking the multiple cloning sites of pGEM-T Easy vector. Subsequently, the plasmids were isolated from positive clones and sequenced with T7 and SP6 primers. Based on these sequences, two internal primers, TIR-17-F1 (5'-CTGAAGTCAGACAGAAG-3') and TIR-17-R1 (5'-GGATCATGACTAGCAAT-3'), were designed to obtain the sequence in the middle of the insert. Finally, the sequence of the 5'RACE product was determined by multiple alignment of all of these fragmented sequences.

2.4.4 Reconstitution of full-length cDNA of CrSARM by *in silico* assembly of the partial sequences

The sequences of clone AmeR209, H3-2.1 and TIR-5RACE#17, which were obtained from the characterization of the EST library, phage library screening and 5'RACE PCR, respectively, were aligned using DNAMAN software. These three sequences were combined based on the overlapping regions. The contiguous sequence was then conceptionally translated in six reading frames, three for sense and three for antisense strand. The full-length open reading frame (ORF) was defined as the longest ORF with the poly-adenine tail downstream of its stop codon. There should be a stop codon located just upstream from its real start codon. In addition, Blastp search was performed to further validate that the amino acid sequence encoded by the full length ORF was indeed homologous to SARM homologues from other organisms. Next, the domain architecture of the full length CrSARM protein was predicted by CDD search and domain boundaries were determined by Globplot software.

2.5 Phylogenetic analysis of CrSARM

To investigate the evolutionary relationship between CrSARM and its homologues, multiple sequence alignment was performed to compare the degree of sequence homology between them. The amino acid sequences were obtained from the NCBI database. Their annotations and accession numbers are listed in Table 3.1. Whenever the sequences were supported by a publication the annotations were according to the published reference. Sequence alignment was performed using Clustal X (version 1.81) with the default parameters. Based on the alignments, unrooted phylogenetic trees were constructed using neighbor-joining method (Saitou and Nei, 1987). Bootstrap tests at 1000 replicates were carried out to examine the validity of the branching topologies. The "Correction for multiple substitution" option was selected to optimize the calculation of the evolutionary distances.

2.6 Characterization of tissue distribution of SARM

CrSARM was initially cloned from the amebocyte. In order to determine whether CrSARM is expressed in tissues other than the amebocytes, the presence of CrSARM transcript in hepatopancreas, heart, stomach, intestine and muscle was studied by RT-PCR.

2.6.1 Synthesis of first strand cDNA from mRNA

First strand cDNAs of various tissues were synthesized using SuperScript™ First-Strand Synthesis System (Invitrogen). For each tissue, 20 μl reaction mixture was set up with 300 ng of mRNA, 0.5 mM dNTP mix, 500 ng of oligo$(dT)_{12-18}$ primers, 5 mM of $MgCl_2$, 10 mM DTT, 40 units of RNaseOUT™ recombinant RNase inhibitor and 50 units of SuperScript™ II reverse transcriptase in 1X RT buffer (20 mM Tris-HCl, pH 8.4, 50 mM KCl). Reverse transcription reaction was carried out for 50 min at 42 ^{o}C and terminated by incubation at 70 ^{o}C for 15 min. Next, RNA was

removed by incubation the reaction mixture with 2 units of RNase H at 37 ^{o}C for 20 min. First strand cDNA was stored at - 20 ^{o}C until use.

2.6.2 Analysis of tissue distribution of CrSARM

Expression of CrSARM in various tissues was assessed by PCR using CrSARM gene specific primers, TIR-F1 (5'-AGACGTAGAGAGGCTCGAAGC-3') and TIR-R1 (5'-TGTTCCCAGGGTCTTTCTTGT-3'). For each tissue, 10 µl reaction mixture consisting of 0.5 µl of first strand cDNA, 200 nM of dNTP mix, 600 nM of each primer, 0.35 units of Taq DNA polymerase (NEB) in 1 X Thermopol buffer (20 mM Tris-HCl, pH 8.8, 10 mM $(NH_4)_2SO_4$, 10 mM KCl, 2 mM $MgSO_4$, 0.1 % Triton X-100) was set up. PCR was performed under the following conditions: initial denaturation at 94 ^{o}C for 5 min, followed by 35 cycles of denaturation at 94 ^{o}C for 30 s, annealing at 60 ^{o}C for 30 s, and extension at 72 ^{o}C for 1 min. For positive control, PCR was done with ribosomal protein L3 (RbL3) specific primers: RiboF (5'-TGTTTCTTCAGAGGACCCA3') and RiboR (5'- CACCAAGAAGTTGCCTCG -3') using the same conditions as for CrSARM. PCR products were resolved by electroporation on a 1.2 % agarose gel. Amplicons of 540 and 450 bp in size were expected for CrSARM and RbL3, respectively.

2.7 Trascriptional profiling of CrSARM in the amebocyte upon *Pseudomonas* infection

The identification of CrSARM EST in the amebocyte subtractive cDNA library (as described in section 2.4.1) suggested that CrSARM gene was differentially regulated in response to pathogen invasion. Monitoring the level of CrSARM transcript during infection would reveal the kinetics of this regulation, thus, assisting the functional study of this molecule.

Changes in the transcription level of CrSARM were studied by semi-quantitative RT-PCR. RbL3 was used as an internal normalization standard to eliminate sample-to-sample variations in the initial cDNA concentrations. Firstly, mRNAs were isolated from naïve, 1, 3, 6, 9, 12, 24 hpi amebocytes (section 2.2.3 & 2.3). For each time point, mRNAs from three individual animals were then pooled together before first strand cDNAs were prepared. Pooling was to neutralize the individual variation between the crabs. Next, PCR amplification of CrSARM and RbL3 were done as described section 2.6 except that the number of PCR cycles was reduced to 28 for the former and 23 for the later to avoid the saturation of PCR product. PCR products were resolved on 1.2 % agarose gel, stained with ethidium bromide. Gel image was acquired and analyzed by Image Master VDS version 2.0 software (Pharmacia Biotech). Band intensities were quantified and expressed as intergrated optical density (IOD). For normalization, ratio between CrSARM and RbL3 IODs was calculated for each time point. The change in transcription level during the course of infection was depicted as ratio of the infected sample to the naïve sample. Three independent experiments were performed. Mean and standard deviation were determined after normalization to RbL3.

2.8 Identification of protein interaction partners of CrSARM by yeast two hybrid assay

Yeast two hybrid assay is a widely used technique for the identification of novel interaction partners of a protein. This method overcomes some problems associated with the *in vitro* pull-down assay such as the difficulty of obtaining cDNA encoding the protein interaction partners, the low abundance of protein involved in the interaction, and the lack of glycosylation of recombinant proteins expressed in a bacterial system.

GAL4-based yeast two hybrid system (BD Matchmaker™ Construction and Screening Kit, Clonetech, Figure 2.4) was used to discover interaction partners of SARM. GAL4 is a yeast transcription factor consisting of two structurally and functionally distinct domains, namely DNA-binding domain (BD) and transcription activation domain (AD). For yeast two hybrid assay, the bait gene was cloned in frame with the GAL4 BD in pGBKT7 vector and transformed into yeast reporter strain Y187. A cDNA library was created in yeast reporter strain AH109 as a recombinant construct with pGAD7-Rec vector, which expresses protein as a fusion with GAL4 AD. Hybrids formed as a result of the library and the bait mating were selected on SD medium that lack of Ade, Trp, His, and Leu (Quadruple Dropout Medium or QDO). Only hybrids in which GAL4 DNA-BD and GAL4 AD was brought to proximity by the interaction between the bait and the prey grew on QO medium due to the activation of reporter genes. Positive clones were further verified by streaking on to QDO medium containing X-α-gal (Figure 2.4). Nucleotide sequence of the prey was determined by sequencing pGAD7-Rec recombinant plasmid rescued from the hybrid. Figure 2.5 is the overview of the yeast two hybrid screening process, the detail of which will be described in the following sections.

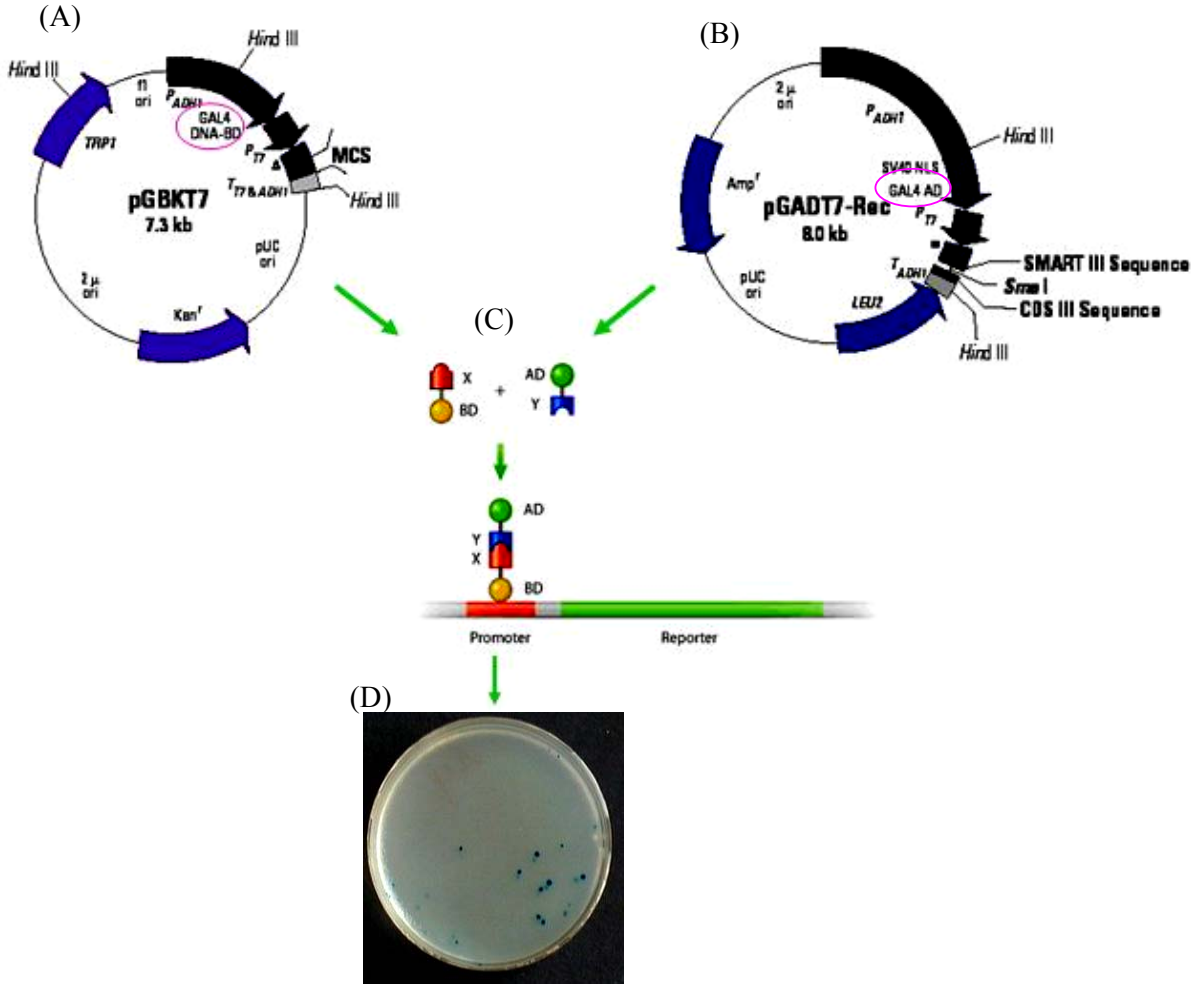


Figure 2.4 Principle of GAL4-based yeast two hybrid system

(A) Gene encoding the protein of which the interaction partners are unknown is cloned in frame with GAL4-DNA binding domain (BD) of the pGBKT7 vector. (B) cDNA library is cloned downstream of GAL4-activation domain (AD) of the pGADT7-Rec vector. Screening can be performed by either mating or co-transformation. (C) Interaction between the bait (X) and the putative prey (Y) brings GAL4-AD and GAL4–BD in close proximity, resulting in the activation of the reporter genes, and thus (D) the growing of positive clones on selective medium (SD/-Ade/-His/-Leu/-Trp/X-gal). Refer to the text for the detail description.

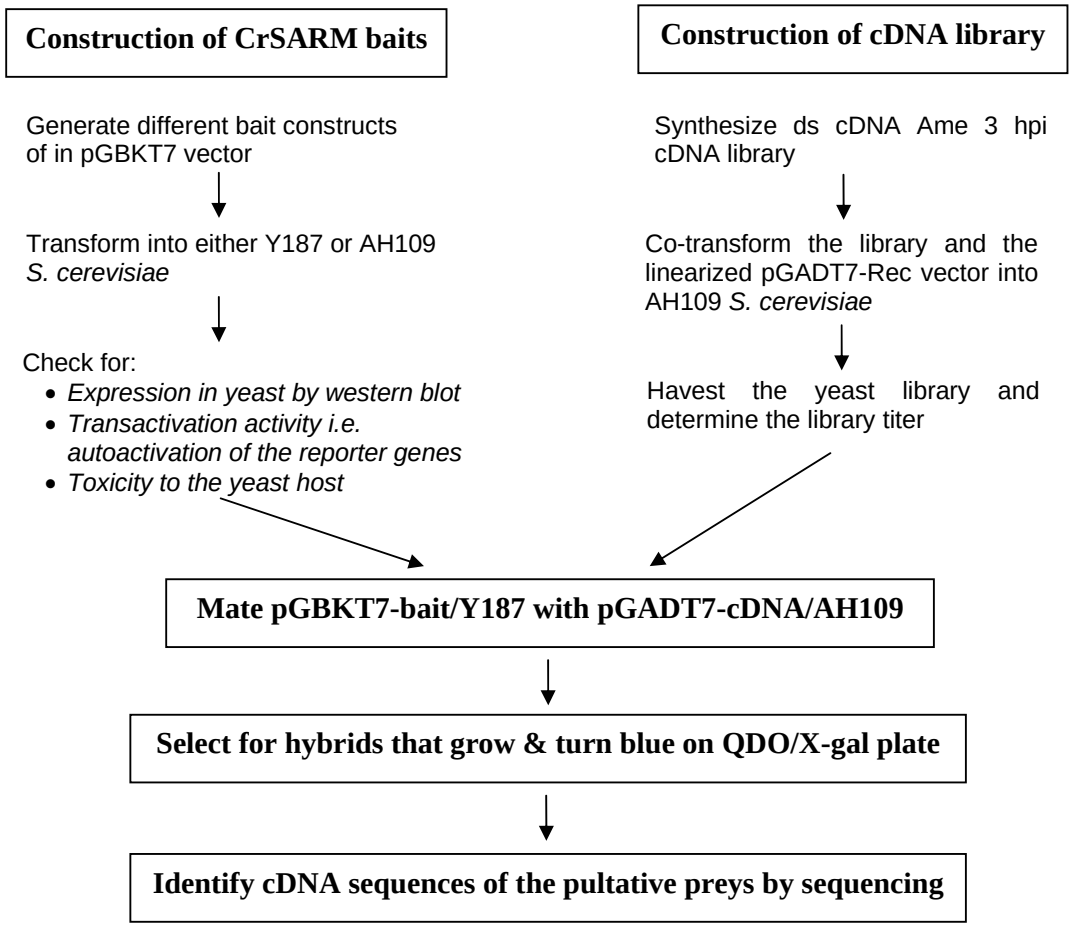


Figure 2.5 Overview of yeast two hybrid screening assay for identification of putative interaction partners of CrSARM

2.8.1 Synthesis of amebocyte 3 hpi double stranded cDNA library

Double stranded cDNA library for yeast two hybrid was generated from mRNA of 3 hpi amebocyte. This source of mRNA was selected since the transcription profile showed an upregulation of CrSARM in the amebocyte shortly after *Pseudomonas* infection. This cDNA library was anticipated to harbour a possibility for the isolation of interaction partners of CrSARM .

Synthesis of first-strand cDNA

A mixture of 1.0 μg of mRNA and 10 pmols of CDS III oligo (dT) primer (5'-ATTCTAGAGGCCGAGGCGGCCGACATG-d(T)30VN-3') was prepared. Then 2.0 μl of 5X First-Strand Buffer (250 mM Tris, pH 8.3; 30 mM MgCL; 375 mM KCl), 1.0 μl of 20 mM DTT, 1.0 μl of 10 mM dNTP Mix and 1.0 of Moloney Murine Leukemia Virus (MMLV) Reverse Transcriptase were added. The reaction mixture was incubated for 10 min at 42 ^{o}C before 10 pmol of SMART III^{TM} Oligonucleotide (5'-AAGCAGTGGTATCAACGCAGAGTGGCCATTATGGCCGGG-3') was added. The reaction was then subjected to PCR in a hot-lid thermal cycler (MJ Research) for 1 hr at 42 ^{o}C followed by 15 min at 75 ^{o}C. After cooling to room temperature, the parental mRNA was removed by incubating the reaction mixture with 2 U of RNase H at 37 ^{o}C for 20 min. The first-strand cDNA was stored at – 20 ^{o}C until use.

Generation of double-stranded cDNA library

The second strand of cDNA was synthesized from the first-stranded template in two of 100 μl reactions. Each reaction consisted of 2 μl of first-strand cDNA, 1 x Advantage 2 PCR Buffer, 0.2 mM dNTP, 0.2 μM of 5' PCR primer (5'-TTCCACCCAAGCAGTGGTATCAACGCAGAGTGG-3'), 0.2 μM and 3' PCR primer (5'-GTATCGATGCCCACCCTCTAGAGGCC GAGGCGGCCGACA-3'), 1X GC-melt solution and 2 μl of 50 x BD Advantage 2 Polymerase Mix. The reaction mixtures were subjected to a PCR program of 1 cycle of 95 °C for 30 s, 21 cycles of 95 °C for 10 s, 68 °C for 6 min (with additional 5 s added per successive cycle), 1 cycle of 68 °C for 5 min. Upon completion, the PCR product was pooled and 7 μl was analyzed on an agarose gel while the rest was stored at – 20 °C until purification.

Purification of double-stranded cDNA

The double-stranded cDNA library was purified and size-fractionated using BD Chroma Spin™ TE-400 column, which was pre-packed with resins that facilitated the selection of DNA molecules longer than 200 bp. The gel matrix in the spin column was resuspended thoroughly before the column was centrifuged in a swinging bucket centrifuge for 5 min at 700 x g to re-establish the gel bed. Ninety five µl of PCR product was carefully applied to the center of each column. The columns were then spun at 700 x g for 5 min. The flow-through, which was purified cDNA library, was collected and pooled in a single tube for ethanol precipitation with 2.5 volume of 95 % ethanol (- 20 °C) and 1/10 volume of sodium acetate (3 M; pH 4.8). The cDNA was precipitated at - 20 °C overnight and recovered by centrifugation for 20 min at 14, 000 x g at room temperature. The supernatant was removed and the pellet was air dried for 30 min before reconstitution in 20 µl of deionized water. The purified cDNA was stored at - 20 °C until being used for *in vivo* recombination with pGADT7-Rec plasmid in the yeast.

2.8.2 Generation of GAL4 fusion library in *Saccharomyces cereviseae* AH109 strain

Using strain AH109, an yeast recombinant library was generated by co-transformation of double-stranded cDNA library and *Sma* I-linearized pGADT7-Rec plasmid. *In vivo* recombination between homologous sequences at the ends of cDNAs and the linear plasmid would yield a functional GAL4 AD expression plasmid.

Yeast competent cells were prepared using YEASTMAKER™ Yeast Transfomation System 2 (Clontech). Three ml of YPDA medium was inoculated with one fresh colony of AH109 yeast strain and incubated with shaking at 230 rpm for 8 h at 30 °C. Five µl of the overnight culture was transferred to 50 µl of YPDA. The

culture was then grown with shaking at 230 rpm at 30 °C until the OD_{600} reached 0.2 (~18 h). After that, cells were collected by centrifugation at 700 x g for 5 min at room temperature and resuspended in 100 ml of fresh YPDA. Yeast was incubated for another 4 h until the OD_{600} of 0.49. The cells were collected by centrifugation and washed with 60 ml of sterile water before resuspension in 3 ml of 1.1 X TE/LiAc solution, which was freshly prepared by mixing 1.1 ml of 10 X TE (0.1 M Tris-HCl, 10 mM EDTA, pH 7.5) with 1.1 ml of 1 M LiAc. Next, the cell suspension was centrifuged at 14, 000 x g for 15 s. The cell pellets were resuspended in 600 µl of 1.1 X TE/LiAc solution. Competent yeast was now ready for transformation.

Cotransformation was performed by combining 20 µl of double stranded cDNA, 6 µl of *Sma* I-linearized pGADT7-Rec plasmid, 20 µl of herring testes carrier DNA, which was pre-denatured by heating at 100 ^{o}C two times for 5 min each, snap-chilled, and added to 600 µl of competent cells in a prechilled 15-ml tube. The tube was gently vortexed and 2.5 ml of PEG/LiAc solution, which was freshly prepared by combining 8 volumes of 50 % PEG with 1 volume of 10 X TE buffer and 1 volume of 1M LiAc, was added. The mixture was briefly vortexed before incubating at 30 ^{o}C for 45 min with gentle mixing every 15 min. To the mixture, 160 µl of DMSO was added and the cells were left in a 42 ^{o}C water bath for 20 min with gentle votexing at every 10 min interval. Cells were collected by centrifugation at 700 x g for 5 min then resuspended in 3 ml of YPD Plus liquid medium (Clonetech). Yeast was then incubated at 30 ^{o}C, 230 rpm for 90 min. After that, the cells were collected by centrifugation and resuspended in 30 ml of 0.9 % NaCl. A 150 µl of this suspension were plated on each of 150-mm SD/-Leu agar plates for selection of GAL4 fusion transformants. The plates were incubated at 30 ^{o}C for 5 days and stored at 4 ^{o}C for 5 h before the transformants were collected.

Harvesting of the library was done by adding 5 ml of Freezing Medium (YPD medium, 25% (v/v) glycerol) to each of SD/-Leu plate. Yeast was scraped from the agar surface by gently shaking the plates with sterile glassbeads. After that, the yeast suspensions were collected and combined. Next, the cell density of the yeast library was determined by using Neubauer hemacytometer counting chamber to be 1.5 x 10^8 cells/ml. One ml aliquots of the library were then prepared for storage at – 80 ^oC.

2.8.3 Generation of pGBKT8-bait constructs

Cloning of the baits into pGBKT7 vector

In order to identify protein partners that associated with CrSARM, cDNAs encoding the ORF of full length CrSASM and individual domains of CrSARM (ARM, SAM and TIR) were sub-cloned into pGBKT7 vector. Since the full length construct was failed to expressed in yeast, shorter cDNA sequence encoding for three putative domains made of CrSARMs but not 443 residues at the N-terminal, was also cloned into pGBKT vector. This construct is, thus, named as CrSARM-AST (Figure 2.6).

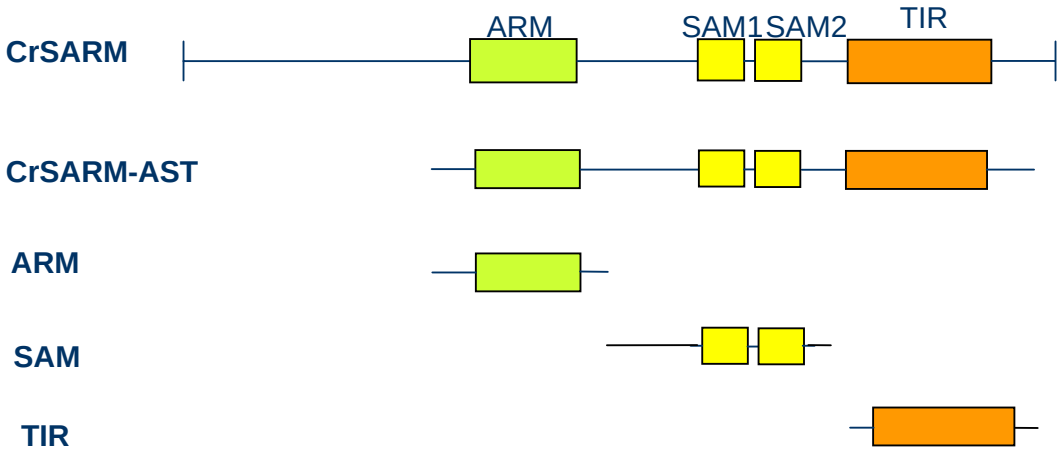


Figure 2.6 Schematic diagram of cDNA constructs of the baits

On the left are the names of the constructs that were cloned in pGBKT7 vector. Drawing is not to scale.

Nucleotide sequences encoding for the baits were amplified from first strand cDNA that was synthesized from the naïve amebocyte mRNA (section 2.6.1) using iProofTM high-fidelity DNA polymerase (Bio-RAD) and primers with appropriate restriction enzyme sites (Table 2.1). For each fragment, 20 μl reaction mixture was set up with 0.2 mM dNTP, 0.4 U of iProof DNA polymerase, 100 ng of DNA template and 0.5 μM of each of the forward and reverse primers in 1 X iProof HF buffer (Bio-RAD). The standard PCR program of one cycle of 98 ^{o}C for 30 s followed with 35 cycles of denaturing at 98 °C for 10 s, annealing for 30 s and extension at 72 °C (refer to Table 2.1 for detail of annealing time and extension time of each fragment) and final extension at 72 °C for 5 min was ran. The PCR products were resolved on agarose gel and the band of interest was extracted from the gel by QIAquick gel extraction kit (Qiagen). After gel purification, PCR products were cloned into pGEMT-Easy vector before they were digested with appropriate restriction enzymes and cloned into pGBKT vector. Due to the presence of proof reading enzyme in iProof polymerase mix, possible blunt-ended PCR products were subjected to A-tailing procedure, followed by TA cloning into pGEM-T Easy vector and transformation into *E. coli* Top10. The DNA sequence of the insert was confirmed as describe before (2.4.3). Next, the excision of the inserts from pGEM-T Easy vector was performed by restriction enzyme digestion with *Nde*I and *Bam*HI or *Nde*I and *Eco*RI to isolate CrSARM-ORF, CrSARM-AST, ARM and TIR, respectively. SAM was released by simultaneous digestion with *Bam*HI and *Eco*RI. pGBKT7 vector was also linearized with the compatible enzymes in order to received the respective insert. After restriction digestion, the linearized inserts and vectors were gel-purified and ligated at 4 ^{o}C, overnight using T4 DNA ligase (Roche). The ligation reactions were set up with the insert : vector molar ratio of 8 : 1 for CrSARM-ORF; 5 : 1 for

CrSARM-AST, SAM and TIR; 3 : 1 for ARM. Subsequently, the ligation mixtures were transformed into *E. coli* Top10 and plated onto LB-Kanamycin (30 μg/ml) for selection of positive clone. Colony PCR and sequencing were performed to confirm the success of the ligation.

Table 2.1 PCR amplification for cloning of bait fragments to pGBKT7

Bait fragment	Primer	Annealing temperature (^{o}C)	Extension time (s)
CrSARM-ORF (aa 1 to 1039)	**Sense: SARM-F-Nde** *(5'-TACAGGACATATGGAAAATGGATTCGCCC-3')* **Antisense: SARM-R-Bam** *(5'-CGGGATCCAAGTTCCACCGAACAAGA-3')*	58.0	120
CrSARM-AST (aa. 444 to 1039)	**Sense: ARM-F-Nde** *(5'-CATATGAATAGAGCTTACGTTGTGGA-3')* **Antisense: SARM-R-Bam** *(5'-GGATCCTTAAAGTTCCACCGAACAAG-3')*	56.0	90
ARM (aa 444 to 570)	**Sense: ARM-F-Nde** *(5'-CATATGAATAGAGCTTACGTTGTGGA-3')* **Antisense: ARM-R-Bam** *(5'-GGATCCTTAAGCTACTAAAGTAGTGATA-3')*	52.0	20
SAM (aa 570 to 846)	**Sense: SAM-F-Eco** *(5'-GAATTCAACAAAGAGATAGAATTTGCAG-3')* **Antisense: SAM-R-Bam** *(5'-GGATCCTTATTTGTTGCAACTAATATCAC-3')*	58.0	30
TIR (aa 847 to 988)	**Sense: TIR-F-Nde** *(5'-CAACAAACATATGACATTAGATGTCTTC-3')* **Antisense: TIR-R-Eco** *(5'-GGAATTCTCACTCCCCGCGCATGAACCT-3')*	56.0	20

Restriction sites are underlined.

Checking the baits for the expression in yeast, transcriptional activation and potential toxicity

Before proceeding with the yeast two-hybrid library screening, one must ensure that that the baits were expressed properly in the yeast and their expression does not hamper the growth of the host strain, which is Y187. In addition, it is important to check that baits themselves are not able to activate the expression of *ADE3* and *HIS2* reporter genes since the putative interactions are isolated based on the transcription activation of these nutrient markers.

Firstly, pGBKT7-CrSARM, -AST, -ARM, -SAM, -TIR and the pGBKT7 vector backbone as a negative control for transcriptional activation test, were transformed into strain Y187 and AH109 using small-scale transformation protocol. This was done by incubating 50 µl competent yeast cells with 1.0 µg of either empty pGBKT7 or recombinant pGBKT7 plasmids, 50 µg of denatured herring testes carrier DNA and 100 µl of PEG/LiAc solution at 30 ^{o}C for 30 min with gentle vortexing every 10 min. 20 µl of DMSO was added into each sample before the cells were heat-shocked at 42 ^{o}C for 15 min. The cells were collected by centrifugation at 14, 000 x g for 15 s, resuspended in 1.0 ml of YPD Plus liquid medium (Clontech), and incubated at 30 ^{o}C with 230 rpm shaking for 90 min and spun down and resuspended in 1.0 ml of 0.9 % NaCl. The transformants were selected on SD minimal medium lacking of tryptophan (SD/-Trp) agar plates.

The expression of the baits in Y187 was detected by western blotting. To prepare for the yeast cell lysate, a single colony of Y187 transformant was picked from SD/-Trp plate and inoculated 5 ml of SD/-Trp broth. The cultures were grown at 30 ^{o}C with shaking at 250 rpm until the OD_{600} reached 1.0. The cells were collected by centrifugation at 10 000 x g for 5 min, washed with 5 ml of water and transferred to an eppendorf tube. Cell pellets were resuspended in 1 volume of lysis buffer (20 mM MOPS, pH 7.3, 0.1 mM EDTA, 1 mM PMSF, 1 x cocktail protease inhibitor (Roche) and 1 volume of glass beads. Cell lysis were achieved by vortexing the mixtures six times for 30 s each. The tubes were chilled on ice for 30 s in between each vortexing.

The crude cell lysates were used for western blotting. One volume of 2 x SDS loading buffer was added to the cell lysates. The mixtures were boiled for 5 min, centrifuged for 10 min at 14 000 x g. The supernatants were resolved on 12 % SDS-

PAGE. After the eletrophoresis, the proteins were transferred to PDVF membrane (GE Healthcare) by electro-transferred. The membrane was prewet in methanol before putting on top of the gel. The transfer was done in a Trans-Blot Electrophoretic Transfer Cell (BioRad) at a constant voltage of 70 V for 1 h with the transfer buffer containing 25 mM Tris-base, 192 mM glycine, 20 % methanol. After transfer, the membranes were blocked with the blocking buffer(3 % w/v skimmed milk in 0.1% Tween 20 in Phosphate buffer saline, pH 7.2 (PBST)) for 1 h, followed by overnight incubation with mouse monoclonal anti-myc antibody (Invitrogen), which was diluted 5000 times in blocking buffer. The membrane was then washed four times with PBST for 15 min each, incubated with HRP-conjugated rabbit anti-mouse IgG (1:10,000 dilution with blocking buffer), for 1 h at room tem incubation, and again washed for times like before. The blot was developed with SuperSignal® West Pico Chemiluminescent Substrate (Pierce) and exposed to Fuji Super RX film.

To test for transcriptional activation of the baits, 15 well-isolated colonies of each transformation were restreacked to fresh SD/-Trp plates, then replicated on SD minimal medium lacking of histidine and tryptophan (SD/-His/-Trp), and adenine and tryptophan (SD/-Ade/-Trp) agar plates. All plates were incubated at 30 ^{o}C for 4 days. Growth on different selection media was then recorded.

Once it was confirmed that the yeast transformants only grew on SD/-Trp but not SD/-His/-Trp and SD/-Ade/-Trp agar plates, the baits was further tested for their toxicity. This can be done by comparing the growth rate of Y187 cells transformed with the pGBKT7-bait plasmid with that transformed with pGBKT7 empty vector in liquid culture. Cultures of 50 ml of SD/-Trp/Kanamycin (20 μg/ml) liquid medium inoculated with one fresh colony of Y187 transformant were set up and grown at 30 ^{o}C, 250 rpm for 24 h. After that, OD_{600} of the cultures were measured and compared.

If OD_{600} of Y187 transformed with the baits were higher or equal to 0.8 and not significantly different from that of the empty pGBKT7, the baits were considered as not toxic. Cells were then collected from the culture by centrifugation at 600 x g for 5 min. Cells pellets were resuspended in 5 ml of SD/-Trp/Kanamycin (20 μg/ml) broth and cell count was performed using the Neubauer hemacytometer. These cells were then used for mating with the amebocyte 3 hpi library if the cell density was ≥ 1 x 10^9 cells/ml.

2.8.4 Screening of two-hybrid Ame 3 hpi cDNA library by yeast mating

Y187 baits and AH109 library mating

In order to identify interaction partners of CrSARM, the Y187 clones harboring pGBKT-baits constructs were mated with the amebocyte 3 hpi cDNA library/AH109. For each mating, 5 ml of Y187 bait cells collected by the end of toxicity test were combined with 1 ml of the Ame 3 hpi AH109 library and 45 ml of 2 X YPDA/Kanamycin (50 μg/ml) liquid medium in a 3-L culture flask. Vial that held the library was rinsed twice, each time with 1.0 ml of 2 X YPDA/Kanamycin (50 μg/ml). The rinsing solutions were also added into the mating flask, which then was incubated at 30 ^oC with gentle shaking at 40 rpm. After incubation for 20 hr, the mating culture was checked under microscope for the presence of the three-lobed shaped zygotes. If zygotes were observed, mating was continued for another 4 hr before cells were collected by 10 min centrifugation at 100 x g. Mating flask was rinsed with two aliquots of 50 ml of 0.5 X YPDA/ Kanamycin (50 μg/ml). The rinsed solutions were combined to resuspend the cell pellets. Centrifugation was repeated and this time the cell pellets was resuspended in 10 ml of 0.5 X YPDA/ Kanamycin

(50 μg/ml). The total volume of this cell suspension was recorded in order to estimate the number of hybrids that had been screened.

Screening for positive hybrids

SD/-Ade/-His/-Leu/-Trp (Quadruple Dropout Medium-QDO) medium was used to select for yeast diploid expressed interaction partners of CrSARM. The absence of the four amino acid supplements made the selection highly stringent. Aliquots of 200 μl of the mating mixture were spread on QDO agar plates. In addition, serial dilution of the mating mixture was done and 100 μl aliquots of 1:10, 1:100, 1:1000, and 1:10000 dilutions were plated on SD/-Leu, SD/-Trp and SD/-Leu/-Trp agar for determination of mating efficiency. All plates were incubated at 30 ^{o}C until the colonies appeared.

Hybrids that grew on QDO plates were further selected for the activation of *MEL1* reporter gene, which encodes for α-galactosidase, by restreaking on fresh QDO containing X-α-Gal. Only clones that grew and turned blue on this agar medium were further analyzed for protein interaction partners.

2.8.5 Identification of cDNA sequences of the putative preys

Sequencing of the pGAD recombinant plasmids harboring in the positive yeast diploids would allow the identification of the cDNA library inserts that encoded the interacting proteins.

Firstly, pGAD plasmids were isolated from the positive hybrid. For each hybrid, 2 ml of QDO/Kanamycin (30 μg/ml) inoculated with one fresh yeast colony was cultured overnight at 30 ^{o}C, 230 rpm. Cells were collected by centrifugation at 18,000 x g for 5 min. Most of supernatant were poured off and cell pellets were resuspended in left-over liquid. Next, the yeast cell wall was disrupted by incubating the cells with 100 U of zymolyase (Seikagaku) at 37 ^{o}C with shaking at 230 rpm for

1.5 h. Subsequently, the cell membrane was broken by vigorously vortexing the cells with 20 µl of 20 % SDS followed by one freeze (at – 20 ^{o}C)/thaw cycle. The volume of the cell lysate was then adjusted to 400 µl using TE buffer pH 7.0. After that, phenol/choloroform extraction was performed to isolate the DNA by vortexing the cell lysate vigorously for 5 min with 400 µl of phenol : chloroform : isoamyl alcohol (25:24:1) mixture. Next, phase separation was facilitated by centrifugation at 14, 000 x g for 10 min. After that, the aqueous phase was carefully transferred to a fresh tube and 400 µl of chloroform : isoamyl alcohol (24:1) was added into each tube and the same procedure was followed to remove the residual phenol from the DNA extract. Next, ethanol precipitation of the plasmid DNA was done by incubating the DNA extract with 2.5 volume of 95 % ethanol and 1/25 volume of 10 M ammonium acetate for 1 h at – 80^{o}C. DNA precipitates were collected by centrifuging the tube at 14, 000 x g for 30 min at 4 ^{o}C and washed with 500 µl of 70 % ethanol. After drying by speedvac, the DNA pellet was dissolved in 20 µl of water.

Plasmid DNA isolated from the positive yeast diploid is a mixture of pGBKT7-bait and pGAD-prey plasmids. Hence, to segregate them, 10 µl of the isolated plasmids were transformed into 100 µl of *E. coli* Top10 and the transformants were divided into two equal aliquots, which were plated on either LB/Ampicilin (100 µg/ml) or LB/Kanamycin (30 µg/ml) agar medium for the selection of pGAD-prey and pGBKT7-bait transformants, respectively. Plasmids isolated from clones grown on LB selection media were sequenced using T7 primer in order to identify the gene encoding the prey and to verify the sequence of the bait.

2.9 Verification of novel protein-protein interactions identified by yeast two hybrid screening

Although multiple selection markers are included, they may not totally eliminate the occurrence of false positive yeast diploids. It was therefore essential to retest the interactions identified by yeast two hybrid screening by using other experimental approaches. Co-transformation of the bait and the prey into AH109 *S. cereviseae* is one of the methods to confirm that the protein-protein interactions identified by yeast two hybrid screening were not an artifact but a reproducible event.

AH109 yeast competent cells were prepared as described before (section 2.8.2). Cotransformation of 1.0 μg of pGBKT-bait and 1.0 μg of pGAD-prey plasmid into 50 μl of competent AH109 cells were done using small scale transformation protocol (section 2.8.4). At the same time, cotransformation of pGBKT-bait and pGADT7-Rec empty vector and of pGBKT7 empty vector with pGADT7-prey plasmid were included as negative controls. The transformants were selected on QDO agar plates, which were incubated at 30 ^{o}C until the colonies appeared.

Chapter 3 Results

3.1 An EST clone encoding CrSARM was identified from the reverse amebocyte substractive cDNA library

Extensive investigation on the innate immune system of the horseshoe crab in the last two decades has revealed a large number of proteins essential for host protection against microbial infection. Coordination of the immune related genes and proteins is critical for the successful control of pathogen invasion. Therefore, in collaboration with other lab members, the first part of this project was aimed to investigate how genes were differentially regulated in the amebocyte and hepatopancreas of *C. rotundicauda*, in response to *P. aeruginosa* infection. Using suppression subtractive cDNA hybridization technique, up- and down-regulated genes in each tissue were selectively cloned into forward- and reverse-cDNA library, respectively. Sequencing of 776 randomly selected clones from these libraries resulted in the characterization of 447 non-redundant ESTs allowing the discovery of many novel immune related proteins from the horseshoe crab (Ding et al., 2005).

Of particular interest is the finding for the first time of a homolog of the mammalian SARM in the *C. rotundicauda* species of horseshoe crab, which was, thus, named as CrSARM. Indeed, EST AmeR209 (dbEST accession: CK086823) isolated from the amebocyte reverse cDNA library showed significant BLASTx match to SARM from human, mouse and SARM homologs in *Drosophila* and *C. elegans* (Figure 3.1 A). Conserved Domain Database Blast identified a TIR domain and a partial SAM motif within this 617 residue-long EST (Figure 3.1 B&C).

(A)

Putative match	Organism	Putative match Accession	E-value
CG7915-PA & CG7915-PB	*Drosophila melanogaster*	AAF50491 & AAN12011	1e-74
SARM	*Mus musculus*	NP_766383	1e-60
SARM	*Homo sapiens*	NP_055892	4e-60
TIR-1	*Caenorhabditis elegans*	AAV91314	2e-53

(B)

```
1         GTACACCTACATTATGCTCCAGGCTGGACTGTCCAGAGACGTCTTTCCCTTTCTGACGGA
1           Y  T  Y  I  M  L  Q  A  G  L  S  R  D  V  F  P  F  L  T  E

61        AGACCAACTGGAAGAGGTCTGTGGTATCAGTAACAGTATTCACAGACGAAGAATTTGGCA
21          D  Q  L  E  E  V  C  G  I  S  N  S  I  H  R  R  R  I  W  Q

121       AGCACTACTAGGTTCTACCCATTCATACACCAGTTGCGATGGTGATATTAGTTGCAACAA
41          A  L  L  G  S  T  H  S  Y  T  S  C  D  G  D  I  S  C  N  K

181       AACATTAGATGTCTTCATCAGTTATCGTAGGCCAACTGGCTCTCAGCTTGCTAGCCTTCT
61          T  L  D  V  F  I  S  Y  R  R  P  T  G  S  Q  L  A  S  L  L

241       GAAGGTTCACCTCCAACTTCGAGGCTTTTCAGTTTTTATAGACGTAGAGAGGCTCGAAGC
81          K  V  H  L  Q  L  R  G  F  S  V  F  I  D  V  E  R  L  E  A

301       TGGAAAATTTGACAATAATCTCGTCAACAGCATACACCAAGCAAAACATTTTATTTTAGT
101         G  K  F  D  N  N  L  V  N  S  I  H  Q  A  K  H  F  I  L  V

361       TCTAACGTCAGACGCTCTCGCTCGTTGCATACGGGACACTGAATGCAAAGATTGGGTGCA
121         L  T  S  D  A  L  A  R  C  I  R  D  T  E  C  K  D  W  V  H

421       CAAGGAAATAGTGGAAGCACTCCAAAGCCAGTGTAACATCATTCCAATTCTTGACAATTT
141         K  E  I  V  E  A  L  Q  S  Q  C  N  I  I  P  I  L  D  N  F

481       TCAATGGCCTGAGCCTGAAAGTTTACCAGAAGATATGAGGGCAGTGTGTTACTTTAATGG
161         Q  W  P  E  P  E  S  L  P  E  D  M  R  A  V  C  Y  F  N  G

541       TGTGAGATGGATTCACGACTACCAAGACGCCTGCGTCGACAAGCTCGAGAGGTTCATGCG
181         V  R  W  I  H  D  Y  Q  D  A  C  V  D  K  L  E  R  F  M  R

601       CGGGGAGCTGAACGTAC
201         G  E  L  N  V
```

Figure 3.1 EST clone AmeR209 is homologous to SARMs from other organisms

(A) BLASTx search showed significant match (E-value ≤ 10^{-3}) between EST AmeR209 (dbEST accession: CK086823) with human, mouse SARM and CG7915 and TIR-1 (Toll and interleukin 1 receptor domain protein), which are homologs of mammalian in *Drosophila*, *C. elegans*. (B) Nucleotide sequence of EST AmeR209 and the correspondent amino acid sequence, which was generated by conceptional translation. Regions that show homology to TIR domain and SAM motif are shaded and underlined, respectively. Nucleotide sequence complementary to TIR-5RACE primer is in bold.

Although, the function of SARM is largely unknown, the presense of TIR domain in its structure strongly suggests its potential function as a signaling molecule of the TLR signaling pathways, the key component of the the innate immunity.

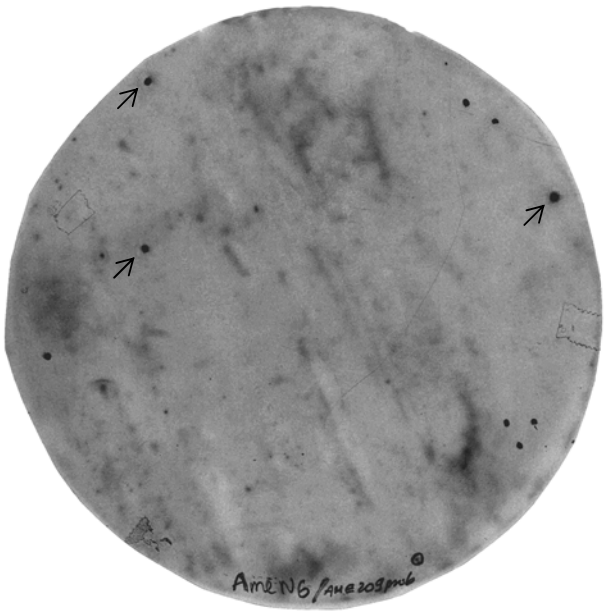
Studying of CrSARM, hence, would provide insights into the signaling events leading to the generation of potent immune responses against microbial infection in the horseshoe crab. Moreover, since SARM is an evolutionary conserved protein, understanding of CrSARM would help to uncover the significance of this immune-related molecule in host defense in higher organisms, particularly in human. Those were the reasons accounting for the selection of CrSARM as a molecule of research interest.

Alignment of the AmeR209 sequence with full length SARM homologs from other organisms indicated that CrSARM aligns with the sequence near the C-terminal of those molecules. Screening of phage amebocyte cDNA library, was, thus, performed to isolate the full length CrSARM.

3.2 The 3'-end of CrSARM cDNA was isolated from the phagemid library

AmeR209, which codes for CrSARM, was used as a probe to screen phage amebocyte cDNA library. Forty seven putative clones, which gave rise to clear but not very sharp dots on X-ray film (Figure 3.2 A), were isolated after two rounds of screening of one million phage clones. Sequencing showed that 40 out of 47 clones harbored DNA fragments whose 5'-ends overlapped with the 3'-end of the probe. Other clones, however, contained irrelevant sequences, which had been isolated due to unspecific hybridization. All of the 40 positive clones contained the poly-adenine sequence at their 3'-terminal, though varied in length, indicating that the 3'-end of CrSARM cDNA was successfully isolated. The sequence of the longest fragment, which was 1150 bp in size, harbored in clone H3-2.1 is shown in Figure 3.2 B. None of the positive clones, however, extended beyond the 5'end of AmeR209. Therefore, 5'RACE PCR was carried out to obtain the 5'-end of CrSARM.

(A)



(B)

```
1       ACACCAAGCA AAACATTTTA TTTTAGTTCT AACGTCAGAC GCTCTCGCTC GTTGCATACG
61      GGACACTGAA TGCAAAGATT GGGTGCACAA GGAAATAGTG GAAGCACTCC AAAGCCAGTG
121     TAACATCATT CCAATTCTTG ACAATTTTCA ATGGCCTGAG CCTGAAAGTT TACCAGAAGA
181     TATGAGGGCA GTGTGTTACT TTAATGGTGT GAGATGGATT CACGACTACC AAGACGCCTG
241     CGTCGACAAG CTCGAGAGGT TCATGCGCGG GGAGCTGAAC GTACGGCCGG ATGGTCCCCT
301     TGGACGACAC ATCGGAATGG GGGGACCTGG AACACCTGGA ACCCCCAGTA CACTTAGTGG
361     ACGTCAGTTG TTCCAAAGAA GCTCGAGTTA CGACAGTGCT AAGGGAAGTT CTTGTTCGGT
421     GGAACTTTAA CGTCTAATTT TAAGGACACG TCTTTTGTGA TGTTAACAAG AAAGACCCTG
481     GGAACACGGT AGCGTAGACT CATGGTCTTC GTCTCTCTAT TAAGGCCTAG TACCTGATCT
541     ATGAATAAGC AAAGTCGAGT TTAGGTTTTT ATATTTCCAC GAACTTGTAA GACAACGTGC
601     GGTAGTAACG GTTCAGACAC GCGCTCATCA CTGTCTTTAT CATGTAACAC ATTGTATAAC
661     AATCGCAACC TGGTGTTTTG TTTGGTTTTT TTATTTAGGT TTGGTCTCTC TTAAATATTT
721     TTACGTATCA CATGAAATCT CCAGTCATTG AGAAGTTTAG GTTGTCCTAA CCATCAGTAA
781     ATCGCAGTCA TTTTCATTAG TAAATATATT TAGCGCTCTT AATCCCGAAT TACGTAAAAA
841     CTAAGCAGAA GGTCCTTTCA GCAGGCACTT TATTGTTAGA GNGAAATAAA AATTACGAGA
901     AAATAATAAT TTCTGATTTG CTAGCTTCAA AGTTTACGAA TTTGAATTAA TTGCATATTT
961     TAATGAAAAG GCTTAAGTTT AATATAACAG CATTTTACTG CGAAGCAGAC GTCAAGTTAT
1021    TTGTACGCAA TGATATTGAA TTCCGATAGG TTTCATTTCC ATGTTTTATC TATGAACCTT
1081    TTATCGATCA GTAATATAAT TAAATATCAA TCAAGGCCGC AAAAAAAAAA AAAAAAAAAA
1141    AAAAAAAAAA
```

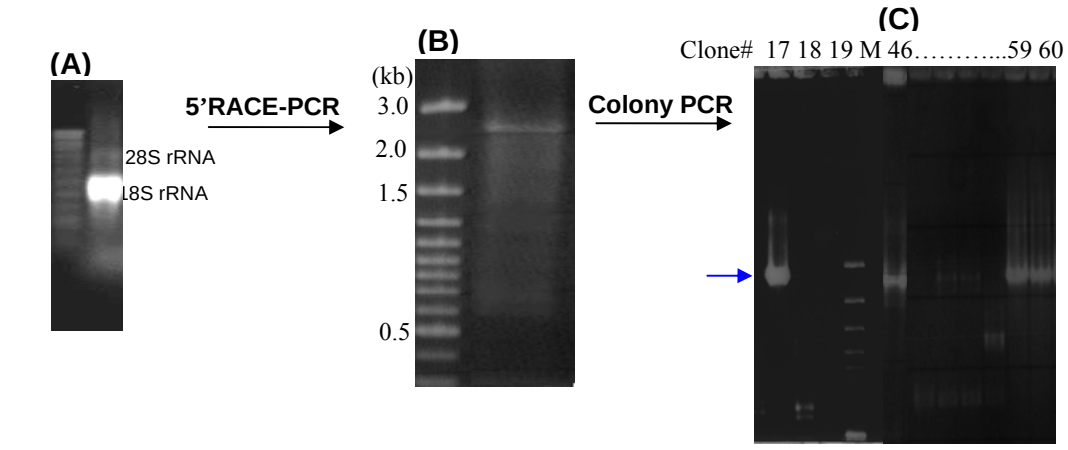
Figure 3.2 3'end of CrSARM was obtained from phage cDNA library screening

(A) *E. coli* XL1-Blue was transfected with phage amebocyte cDNA library and plated on LB agar for plaque formation. The plaques were transferred to positive-charged nylon membrane and hybridized with [^{32}P]-labeled probe. The membrane was subjected to autoradiography. Dots with fuzzy edges, which are marked with arrows, on the X-ray film indicate position of putative clones. Sharp dots were created by the tips of forceps/needle used to handle/mark the membrane. (B) Nucleotide sequence of clone H3-2.1, the longest clone that was isolated from the screening of phage amebocyte cDNA library using AmeR209 cDNA as a probe. Poly-adenine tail that marked the extreme 3'-end of CrSARM cDNA is underlined. Sequence overlapping with EST AmeR209 is shaded.

3.3 5'-end of CrSARM was obtained by 5'-RACE

First strand cDNA, which was synthesized from the mRNA of naïve amebocyte, was used as a template for 5'-RACE. Before isolation of the mRNA, the quality of the total RNA was checked on agarose gel. 18S and 28S ribosomal RNAs were clearly seen indicating that the total RNA was intact (Figure 3.3 A). A single band of ~ 2.5 kb was obtained by 5'RACE PCR (Figure 3.3 B). This fragment was cloned into pGEM-T Easy vector to facilitate its sequencing. Three positive pGEM-T Easy transformants, TIR-5RACE#17, 46, 59 and 60 were selected after colony PCR. Sequencing of these clones showed that they all contained an identical fragment of 2563 nucleotides and their 3'-end aligned with the 5'-end of AmeR209 (Figure 3.3 C & D). Hence, TIR-5RACE#17 was used for analysis hereafter.

CDD Blast identified an Armadillo/Heat motif (ARM) and two consecutive SAM domains at the 5'- and 3'-ends of TIR-5RACE#17, respectively. In addition, a putative start codon that was preceded by a stop codon (TAG) was found upstream of ARM (Figure 3.3 D). Indeed, this arrangement was similar to the sequence at the 5'-end of SARM homologs from other organisms. It therefore can be said that the 5'-terminal of CrSARM cDNA was successfully isolated.



(D)

```
1       GGAAGGCAGTGCTTTCAACGCAGAGTACGCGGGGACTTCTTAGAGTGACTGTTGGCGGACGTTCACTCAT
71      TCACGATGATCTAGACGTCTGCAAGGCGTAGTCTTTTGTTACAGGATTCATGGAAAATGGATTCGCCCCC
141     CCCCCCCTGGTTCCTGATTTTTCAGAAGATTATAGCCAGAAGACTAACTCCGATACTAACATCTTGATCC
211     AAAGTTTCGATAACAATCCTTCTTTAAATGGTAAAATGAAGAACAACGCATCTTCAAAGACCAAAGGAAC
281     AAATATAACTGCTAGGTATCATGCTACGAAGTCAGCCGATAATGCCACCACCGCCACATCATCAACGTCG
351     GCTATATGCAAGAAGTCAGAGGTCAAAGTTTGTCAGAAGCTCAAGCAGGAACAAAATGAAAGTTTCCAAG
421     AAAAACAAAACTTTTCCAGAAGTTTTAGCGTATCTTCGAGCGCAGGCAGTAATACCAAAACGAAACTGAA
491     GTCAGACAGAAGAGGAAACTCCCTCGACACTGTTCTGGACACCACAAAAACTCCCCACCAGAGAGATTTC
561     ACTTTTGAACAGAAACAAGTAATGTCTTCTTCCCAAACCAAAACCACCACAGATAGTTTTTCCTGCGAAA
631     AAGCCAAGATAGATAACAAAGAATTGAAGGAGCTCCAAACAGAAGAGCTGAGTTATGCTGAAAAACATTC
701     TTCTTCAATAGTCAATGCTCGACTAAAAGCTGATGACTTCTCTGCTGAGAAGACAAACGTTATTAATAAA
771     GATCAGGTACAGTTGAAGTGTGAAGAAACTATCCATCAACAGCAAAGGGTTGCTACGACCACTTCTCAAC
841     GAGTAGCAACAGAAAAATTTACAACAGAAAAAATATCCACGACAGCTCAAGAAGAACGTCATACTTTTTC
911     CCAATCATCGCTCCAACAAGAAACAAACGTTTCCTCCAAGCTAAAGATTTCCAATAATAAGATGACTTCC
981     ATTCTCAGTGGTCATGACTTCCATGAACTAGATATTGGAGGACTGATGTCGCCATCGGGTGGAAGTCCAA
1051    TCATTACTTTTCCCGAGGAAGAAGTGTTTCACCAAAGACGTCAAAATCAAGATTTAGCTATCACCGAAGG
1121    CATTCAAGATGACATACAACTTCTGCAGTCGCATACCTCATCTTCTGAAGTGGATAAAACGATTACCAGG
1191    TATTGTTCTAGAATGGAAAACTGTATAAACCGTTTGAAGACATCTTCAACTCCGCAGGCTGTAGAACTGT
1261    TGTCTCTTATGGCAGATATGATAATAAAAGCTTGGTCTTTATCATCTCATGGGCATGAACTTGGTCAAAA
1331    ATTATGTGAAACCATCCACAATAAAGGCGGACTCGACATTCTCATAAGTAACTGTGCAACGACAGACAAT
1401    GACCTGCAGTTCGCCAGCGCTAAACTTCTAGAACATTGCCTCAACGATAACAATAGAGCTTACGTTGTGG
1471    AAAGAGGTCTGATGAATATCGTGACGATGGCGTGTTCCTTCAGCAAAGCGAATTCTGTGTGTCATTCCAG
1541    GACTGGAACAGCAATACTGGAAAATCTCTTCAGACACTCAGAGAAGACATGCAGTGACGTTGTTAGGCTT
1611    GGAGGACTAAAATCAATTCTCTATGACTGCAGACACATAGACGTACGAACTCTTCGTCACTGTGCTATAG
1681    CTTTGGCGAATCTCTCTCTTTATGGTGGTCCTGAGAACCAGCAAGCTATGATCAAGGAAAAAGTTCCGGT
1751    CTGGCTGTTTCCGCTAGCTTTTAGCCAGGATGACATCATTAAGTACTTCGCCTGCTTGGCTATCACTACT
1821    TTAGTAGCTAACAAAGAGATAGAATTTGCAGTCTTGCGCTCTGGAACCTTGGACCTTGTTGAACCATTTA
1891    TTGCTAGTCATGATCCAGAAGAGTTTGCCAGCAAAGCTCCAGCTCATATCCACAGACAATCCAAAGAATG
1961    GCTTTTGAGACTGGTTCCTGTCTTAGACAGTAACAGGGAAGAAGCTAGAAGCTTAGCAGCTTTTCACTTC
2031    GCCCTAGAGGCCAGCATAAAAAAAAAACTTGGAAACACAGAGATATTTCACGACATCGGTGCAATAGAAG
2101    TCTTGAAGAGAGTTGCCAGTTCTCCAAATGCCATTGCATCAAAGTACGCCGCTCAAGCTCTCCAGCTGAT
2171    TGGTGAAGAGATACCACACAAACTCTCTCAACAAGTTCCGCTGTGGAGTGCGGAAGACGTCACACAGTGG
2241    GTTAAACAGATAGGGTTTTCGACTTACTGTCAAGAATTTATGAATAGCCGGGTGGACGGAGACTTACTGC
2311    TACAGCTCAATGAAGATACCTTAAAAAATGATATTGGTATCAGTAATGGTATTCTGAGGAAAAGGTTTCT
2381    TCGAGAACTGGTCCAACTAAAACGTCTTGCTGATTACTCGTCTTGCGACACAATCTCTCTTAACATGATA
2451    CTACAAGAACTGGGTTCAGAATATAGCCAGTACACCTACATTATGCTCCAGGCTGGACTGTCCAGAGACG
2521    TCTTTCCCTTTCTGACGGAAGACCAACTGGAAGAGGTCTGTGG
```

ARM

SAMS

Figure 3.3 Cloning of 5'-end cDNA of CrSARM

(A) Total RNA from naïve amebocyte showed 18S and 28S ribosomal RNA bands. mRNA was isolated from total RNA and used for synthesizing cDNA, which in turn served as template for (B) 5'RACE PCR. ~ 2.5 kb fragment was gel purified and cloned into pGEM-T Easy vector. (C) Colony PCR using T7 and SP6 primers was performed to screen the pGEM transformant. Clones TIR-5RACE#17, 46, 59 & 60 have inserts with the correct size (indicated by the blue arrow). M = 100bp DNA ladder plus marker (Fermentas). (D) TIR-5RACE#17 contains 2562 nucleotides. It encodes for a putative ARM motif and two SAM motifs, which are written in *blue* and *red,* respectively. Putative start codon (ATG) and its preceding stop codon (TAG) are *circled*. The sequence overlapping with that of AmeR209 is *underlined*. The annealing position for TIR-5Race primer is *double-underlined*.

3.4 *In silico* assembly of full length CrSARM cDNA sequence

Full length CrSARM (GeneBank accession number: DQ295792), which was determined by careful assembling of AmeR209, H3-2.1 and TIR-5RACE#17 sequences using DNAMAN software, encompassed a total of 3936 nucleotides with a poly-adenine tail at the extreme 3'-end. This is the first horseshoe crab SARM deposited to the GenBank. Horseshoe crab SARM has a single ORF of 3123 nucleotides encoding for a polypeptide of 1040 amino-acid residues containing 720 bp 3' untranslated region. Conserved domain search revealed a mammalian SARM-characteristic modular architecture of CrSARM which contains two central SAM domains (SAM1 aa 696-763 and SAM2 aa 766-835) that are flanked by an N-terminal ARM motif from aa 441-570) and an extreme C-terminal TIR domain (aa 849 - 987) (Figure 3.4 & Figure 3.5).

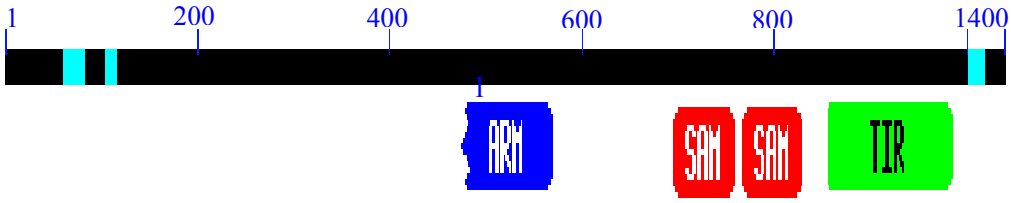


Figure 3.4 CrSARM has a standard modular structure of the vertebrate SARM

The schematic domain arrangement of CrSARM as shown by conserved domain search at http://www.ncbi.nlm.nih.gov.libproxy1.nus.edu.sg/Structure/cdd/wrpsb.cgi. Conserved domain database references: ARM (cd00020), Armadillo/beta-catenin-like repeat; SAM (smart00454, cd00166), sterile-alpha motif; TIR (smart00255), Toll/interleukin-1 receptor domain.

3.5 CrSARM is evolutionarily conserved

Searching the database with full length amino acid sequence of CrSARM led to the retrieval of its homologs from worm, insects, bird and mammals. High homology between SARM sequences from different organism (Figure 3.5), especially at the C-terminal region where ARM, TIR and SAM are located, indicates that the unique combination of these three domains bears critical significance to the function of this protein. Sequence at the N-terminus of the invertebrates homologs, however, are significantly longer than those of the vertebrates homologs (Figure 3.5 and Appendix A).

Phylogenetic study showed that CrSARM is grouped of with other SARM homologs of insects like *Drosophila* and *Anopheles* whereas the vertebrate SARMs and the homolog of vertebrate SARM in *C. elegans*, TIR-1, form separate groups (Figure 3.6).

Table 3.1 SARM homologs that were used for phylogenetic analysis

Protein	Organisms	Accession number
Toll and interleukin 1 receptor domain protein isoform a (TIR-1a)	*Caenorhabditis elegans*	AAV91312
Toll and interleukin 1 receptor domain protein isoform c (TIR-1c)	*Caenorhabditis elegans*	AAV91313
Drosophila Ect4 isoform a	*Drosophila melanogaster*	NP_648173
Drosophila Ect4 isoform b	*Drosophila melanogaster*	NP_729327
Cow SARM	*Bos taurus*	XP_616381
Dog SARM	*Canis familiaris*	XP_548285
Horseshoe crab SARM	*Carcinoscopius rotundicauda*	DQ295792
Human SARM	*Homo sapiens*	NP_055892
Mouse SARM	*Mus musculus*	NP_766383
Mosquitoes ENSANGP00000013511	*Anopheles gambiae*	EAA11071

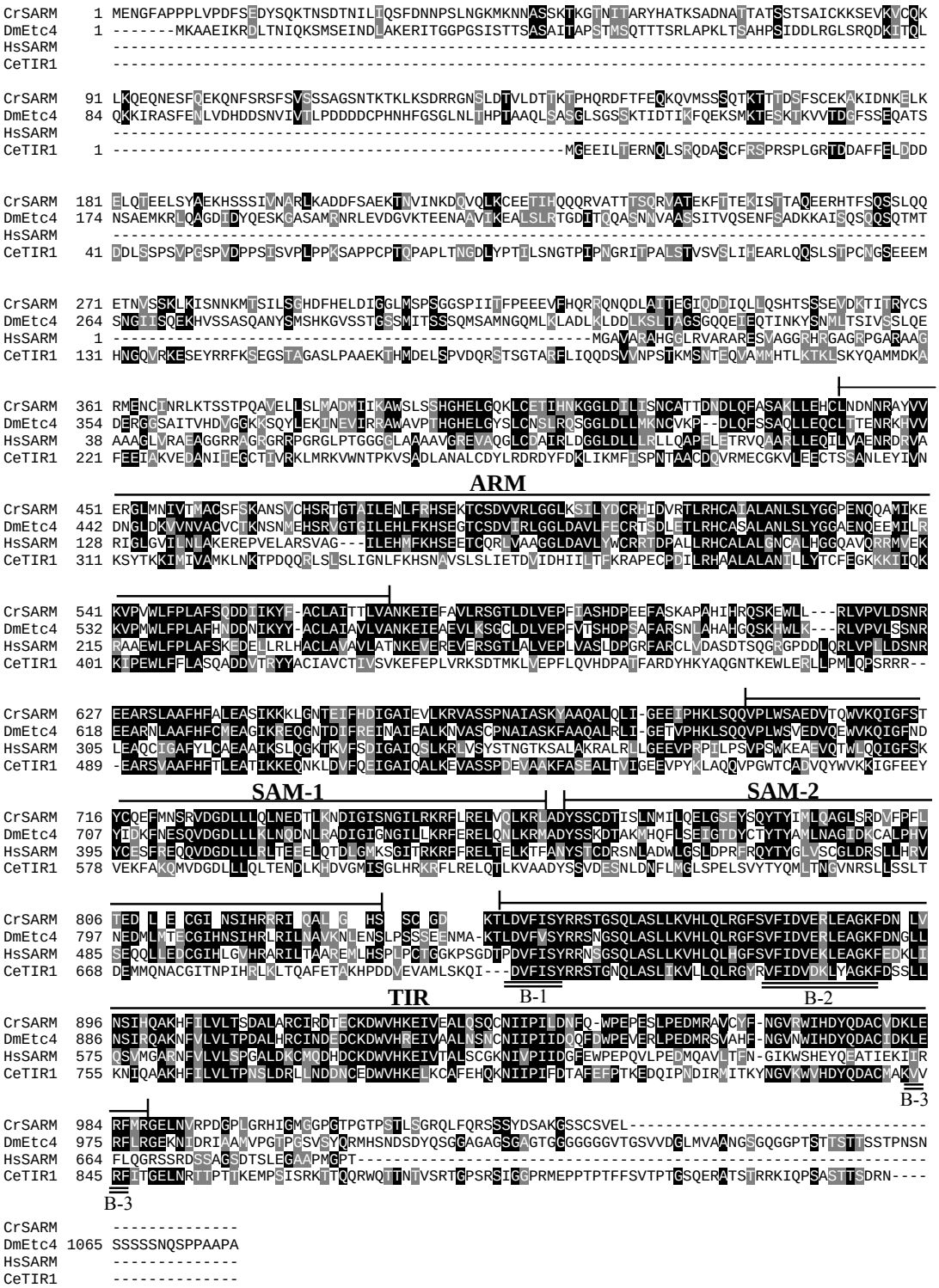


Figure 3.5 CrSARM has high homology to the orthologs

CrSARM amino acid sequence was aligned with the SARM homologs from *C. elegans* (CeTIR1), *D. melanogaster* (DmEtc4), and human (HsSARM). GeneBank accession numbers for these proteins are listed in Table 3.1. ARM repeat motif, SAM motifs, and TIR domain of CrSARM are indicated by the line above the alignment. The three conserved boxes (B-1, B-2 & B-3), which are important for the dimerization of TIR domains, are double underline.

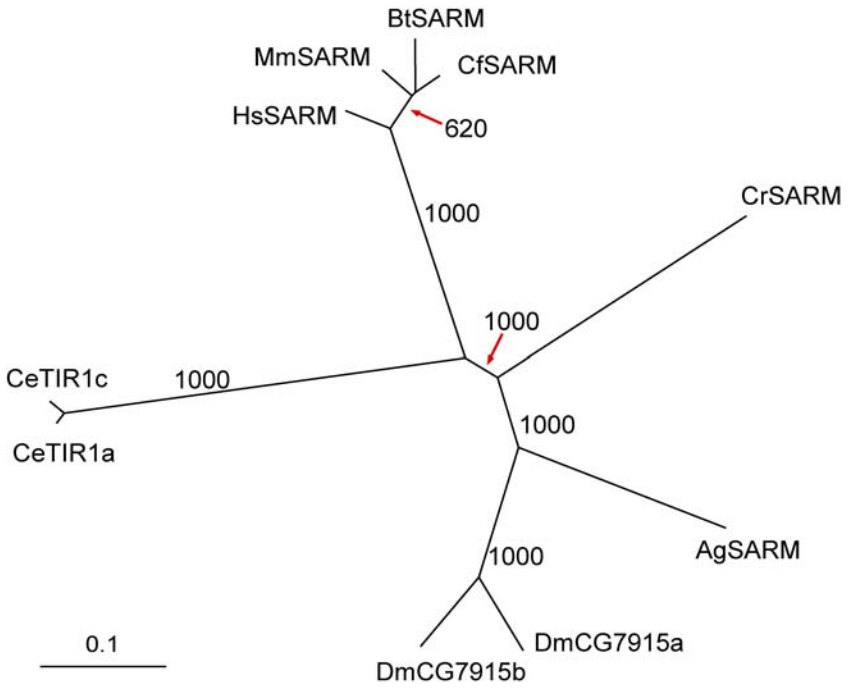


Figure 3.6 Phylogenetic tree of CrSARM and SARM from other

Selected proteins and the corresponding source of origin of the organisms are listed in Table 3.1. AgEAA11071, mosquitoes EAA11071; BtSARM, cow SARM; CeTIR-1a & CeTIR-1c, *C. elegans* TIR1 isoform a and c; CfSARM, dog SARM, CrSARM, horseshoe crab SARM, DmCG7915a/b, Drosophila Etc4 isoform a and b; HsSARM, human SARM; MmSARM, mouse SARM. The un-rooted tree was constructed using neighbor-joining method based on the multiple sequence alignment produced using ClustalX (1.81) (Appendix A).The scale bar corresponds to 0.1 estimated amino-acid substitutions per site. The confidence scores (in %) of a bootstrap test of 1000 replicates are indicated for major branching nodes.

3.6 CrSARM shows tissue specific expression

Tissue distribution of CrSARM was studied by RT-PCR using CrSARM gene specific primers. The result showed that this protein is not ubiquitously expressed. CrSARM transcripts were detected in the two immune related tissues, hepatopancreas and amebocyte, muscle and heart but not in intestine and stomach (Figure 3.7). Unlike CrSARM, human SARM was found exclusively in the liver and kidney due to the presence of a liver- and kidney- specific nuclear factor binding sequences upstream of *SARM* gene (Mink et al., 2001). The difference in tissue expression is another indication that this signaling molecule may perform different roles in invertebrates and vertebrates.

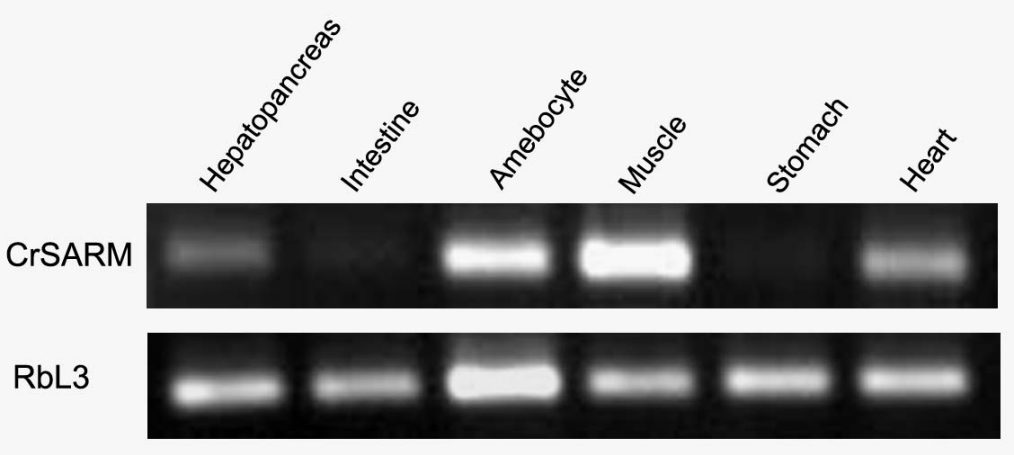


Figure 3.7 Expression of CrSARM in different tissues

RT-PCR detection of CrSARM mRNA from various tissues. Amplifications of CrSARM cDNA fragments were performed using equal number of PCR cycles. The horseshoe crab ribosomal protein L3 (RpL3) was used as the control.

3.7 Infection with *P. aeruginosa* up-regulated CrSARM gene

The EST encoding CrSARM was isolated from the 3+6 hpi amebocyte reverse library indicating its expression was differentially regulated by *Pseudomonas* infection (Ding et al., 2005). Studying the kinetics of this regulation was essential for uncovering the function of this novel signaling molecule. Transcription profiling was performed by semi-quantitative PCR. The relative transcriptional level of CrSARM at different time points post infection was compared to that of the naïve (uninfected) sample after normalization against RpL3. The bar-chart in Figure 3.8 shows that the amount of CrSARM transcript was almost double at 3 hpi. After that, however, the transcription was repressed significantly before gradually returning to the basal level after 12 hpi. The increase at 3 hpi was unexpected since CrSARM was found in a reversed amebocyte cDNA library. However, this contradiction can be explained by the fact that the drastic reduction at 6 hpi had compensated the increase at 3 hpi, leading to the net repression effect when the cDNAs of these two time points were combined for the preparation of the amebocyte subtractive cDNA library

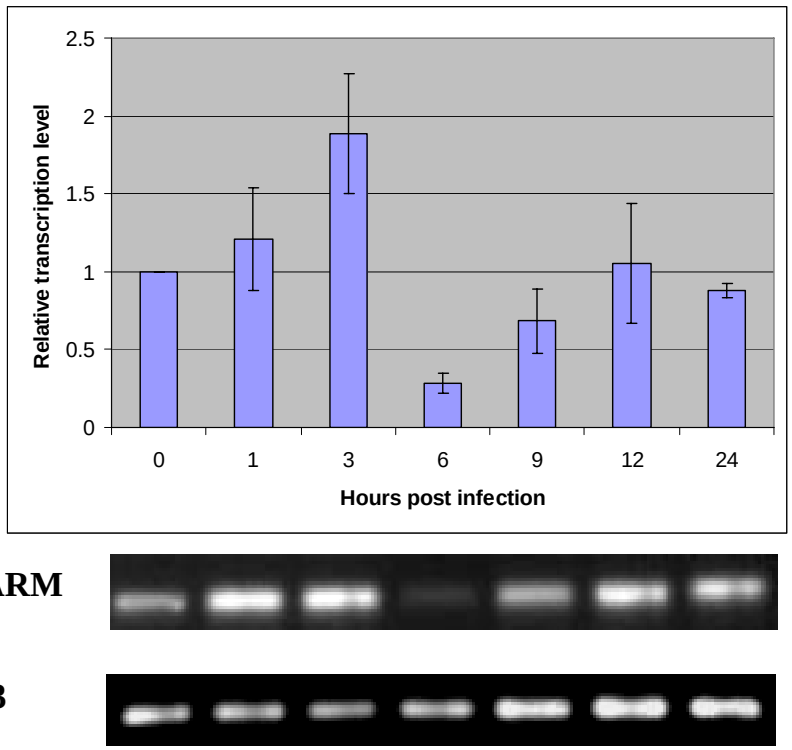


Figure 3.8 Transcription profile of CrSARM upon *Pseudomonas* infection

Horseshoe crabs were infected with sub-lethal dosage of *P. aeruginosa.* Amebocytes were collected at certain time points after infection (0, 1, 3, 6, 9, 12, and 24 hpi) and used for mRNA isolation. RT-PCR was performed with gene specific primers for CrSARM or RbL3. cDNAs were resolved on 1.2 % agarose gel, quantified as intergrated optical density (IOD) by Image Master VDS version 2.0 software. For normalization, the ratio of CrSARM to RpL3 was calculated for each time point. Relative transcription level was determined by comparing the infected sample with the naïve sample.

3.8 Putative interaction partners of CrSARM were isolated by yeast two hybrid screening

As previously mentioned, all of the three domains that made up SARM are well known for their ability to mediate protein-protein interaction (section 1.2.2 e). The unique combination of these domains in SARM highly suggests that this intracellular signaling molecule may function as an adaptor protein. By assembling on this adaptor, signaling molecules of TLR or a novel signaling pathway may be brought into proximity, allowing their interaction, and thus signal transduction to

occur. Identification of proteins that interact with CrSARM is hence a feasible approach to test this hypothesis.

Yeast two hybrid was a method of choice for identification interaction of partners of CrSARM since it has several advantages over other techniques that can also be used for studying protein-protein interaction, for example, GST pull-down. Indeed, post-transcriptional modification of recombinant proteins, which may be essential for the interaction, is possible in an eukaryotic expression system like yeast. Moreover, low expression level of target proteins (typically that of signaling proteins) does not pose a hindrance for the detection of potential interaction protein partners since expression of both bait and preys are constitutively expressed under the control of a strong promoter (P_{ADH1}).

The cDNA library for yeast two hybrid screening was prepared from mRNA of 3 hpi amebocytes because amebocytes are immune-responsive and the expression of CrSARM as well as many signaling proteins were found to be upregulated in the first few hours post infection (Figure 3.8 and Ding et al., 2005)). Double-stranded cDNA were generated from mRNA. Salts and molecules smaller than 200 bp were removed by passing the cDNA through TE column. The size of the purified cDNAs was in the range of 300 bp, which was good enough for downstream works (Figure 3.9).

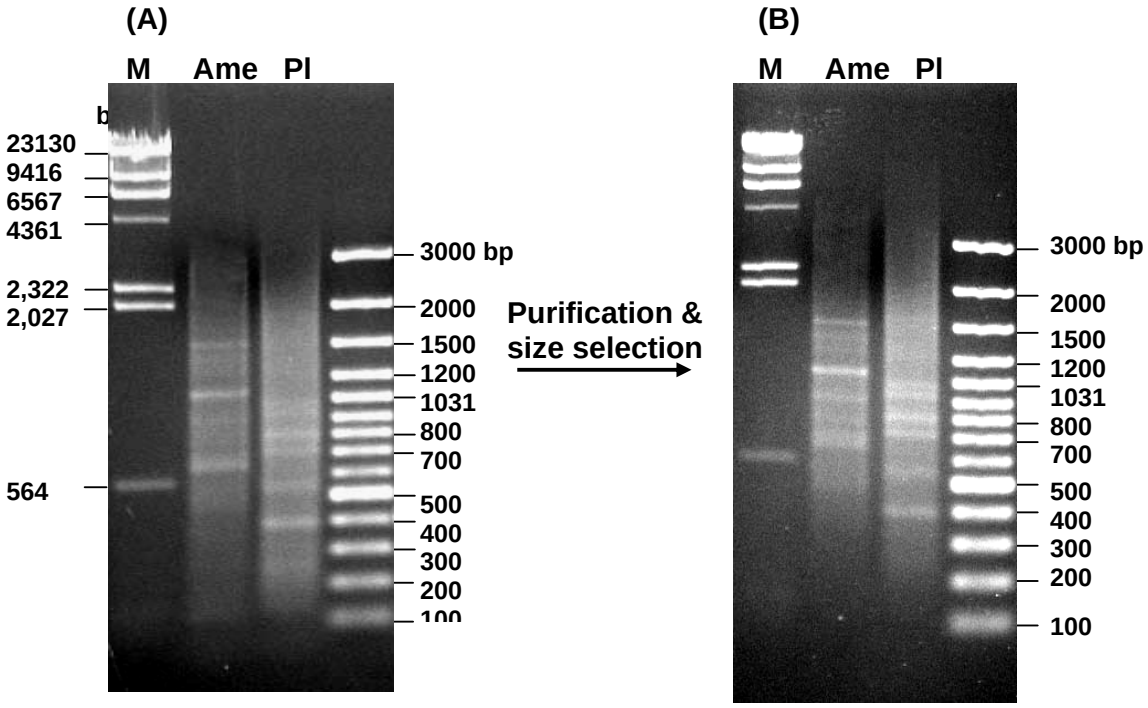


Figure 3.9 The synthesis of double stranded cDNA for the yeast library construction

(A) Double stranded cDNA was synthesized by PCR amplification of single stranded cDNA. 7 µl of PCR product was checked on agarose gel, the remainder, 193 µl, was loaded into two TE columns for size selection and salt removal. Purified cDNA was concentrated by ethanol precipitation to 20 µl and (B) 1 µl was analyzed on agarose gel. Positive control was prepared with human placenta mRNA provided by the manufacturer using the same protocol. M, marker with molecular mass in bp; Ame, amebocyte cDNA; Pl, human placenta cDNA as control.

Double stranded cDNAs and the linearized pGAD-Rec vector were co-transformed into *S. cereviseae* strain AH109. Circular recombinant plasmids were generated as a result of *in vivo* recombination. Transformation efficiency was determined to be 4.8 x 10^6 cfu/3µg of vector, which was higher than that of the positive control, SV40 DNA fragment provided by the manufacturer. In addition, transformation of the linearized vector alone did not yield any transformant, indicating that the vector was not able to re-circularized by itself *in vivo* (Table 3.2). Hence, there were sufficient representations of different cDNAs in the amebocyte 3 hpi cDNA library.

Table 3.2 Transformation efficiency of the yeast amebocyte 3 hpi cDNA library

Double stranded DNA	Vector	Efficiency (cfu/3 μg of vector)
Ame 3 hpi cDNA	Linearized pGAD-Rec	4.8×10^6
SV40	Linearized pGAD-Rec	3.4×10^6
None	Linearized pGAD-Rec	0

AH109 was transformed with circular pGAD-Rec or co-transformed with double stranded cDNA and linearized pGAD-Rec. Serial dilution was carried out and the transformants were selected on SD/-Leu agar. Transformation efficiency was determined as number of colonies grew on selective medium per 3 μg of vector transformed.

The transformed yeasts were plated on SD/-Leu agar and harvested from nearly 200 plates. The cDNA library titer was determined to be 1.2×10^8 cfu/ml. For an effective screening, a cell density higher than 10^7 cfu/ml is required. The amebocyte 3 hpi cDNA library, therefore, harbors a satisfactory titer.

3.8.1 Potential interaction partners of CrSARM were retrieved from yeast two hybrid screening of 3 hpi amebocyte cDNA library

Construction of the baits

For screening of the yeast library, five constructs of baits were sub-cloned into pGBKT7 vector. They are the full-length CrSARM; individual domain of CrSARM: ARM, SAM, and TIR; and the C-terminal sequence (CrSARM-AST) encoding for all three domains of CrSARMs (aa 444-1039) (Figure 2.6). These baits were PCR amplified from single stranded cDNA that was synthesized from naïve amebocyte mRNA. The PCR products with the expected sizes (Table 3.3) were obtained as shown in Figure 3.10. The PCR fragments were cloned into pGEMT-Easy vector and sequenced to ensure that they were free of mutation. After that, the inserts were released from pGEMT-Easy vector by digestion with the respective restriction enzymes and cloned into pGBKT7 vectors. Sequencing of recombinant pGBKT7 vectors were performed to make sure that the inserts were in-frame with cDNA sequence encoding for GAL4-BD domain, and thus c-myc epitope tag in pGBKT7.

Bait constructs were then transformed into both yeast strain AH109 and Y187. Transformants that grew on SD/-Trp agar were selected for toxicity and transcriptional activation tests.

Table 3.3 pGBKT7-bait constructs

Clone ID	Size of cDNA insert (bp)	Expected size of BD-fusion protein (kDa) *
pGBKT7-CrSARM	3123	135.6
pGBKT7-CrSARM-AST	1791	86.2
pGBKT7-ARM	378	33.2
pGBKT7-SAM	831	50.5
pGBKT7-TIR	426	35.7

* Molecular of GAL4-BD is ~ 20 kDa

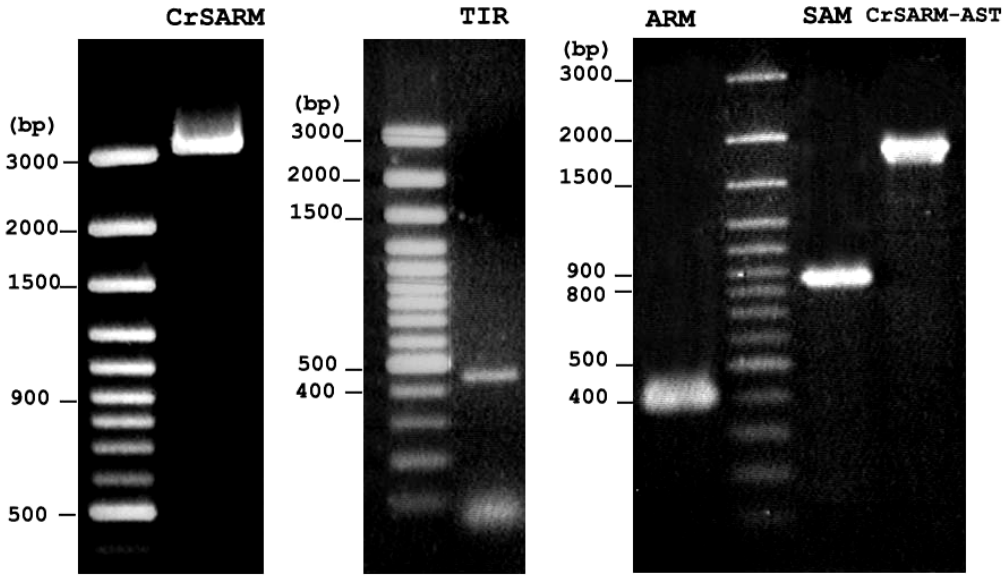


Figure 3.10 PCR amplification of cDNAs encoding the baits

Full-length CrSARM, its C-terminal fragment (CrSARM-AST), ARM, SAM and TIR domain (referred to Figure 2.6 schematic diagram of each fragment) were amplified by PCR and cloned into pGBKT7 vector. Expected size of each fragment can be found in Table 3.3. 100 bp DNA Ladder Plus (Fermentas) was used as molecular marker.

Testing of the baits for the expression, toxicity and transcriptional activation

The expression of BD-fusion proteins in Y187, which was later used for mating with the cDNA library in AH109, was verified before proceeding with the screening process. This was achieved by Western blotting using anti-myc antibody, which recognized the c-myc epitope tag in the fusion proteins, to detect the presence

of the fusion proteins in the crude yeast cell extract. Bands of expected sizes were observed in all of the bait constructs except for the full length CrSARM (Figure 3.11) indicating that the BD-bait fusions except the full length CrSARM were properly expressed. It was unable to detect BD-CrSARM protein even though the Western blotting was repeated with higher amount of cell lysate. Due to this problem, pGBKT7-CrSARM was not included in subsequent experiments.

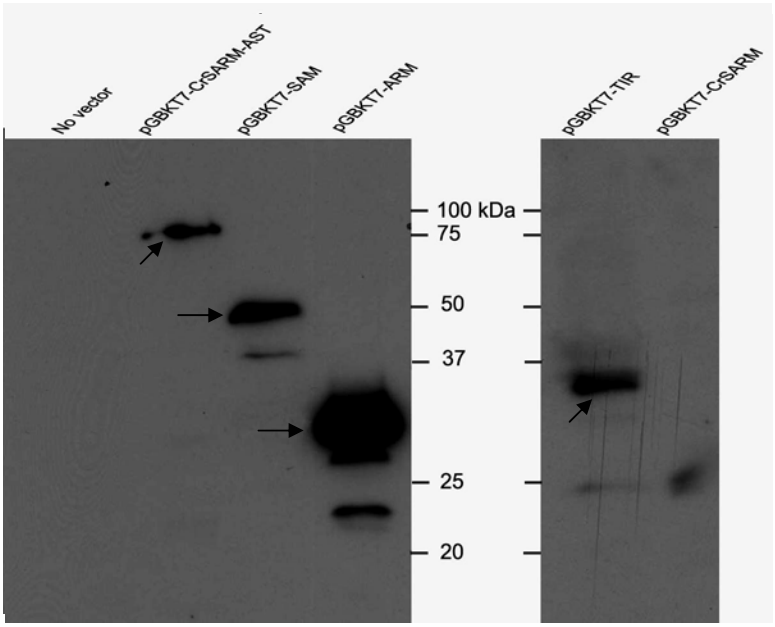


Figure 3.11 Expression of BD-bait fusion proteins in Y187 yeast

pGBKT7-bait/Y187 and the non-transformed Y187 clones were cultured overnight in SD/-Trp and YPD liquid media, respectively. Crude cell lysates were prepared and resolved on SDS-PAGE. Proteins were transfer to PDVF membrane for Western blotting with anti-myc antibody. Lysate from non-transformed Y187 *S. cerevisiae* was used as the negative control. Arrows indicate bands with the expected sizes for the BD-bait fusions as shown in Table 3.3. Fainter bands below the BD-SAM, BD-ARM and BD-TIR bands are degradation products.

Before mating with the library, it was also important to make sure that the BD-bait fusion proteins were not toxic to the yeast and do not activate the expression of marker gene *ADE2*, *HIS3* and *lacZ*. To check for toxicity, the growth of Y187 transformed with the bait constructs in SD/-Trp broth was compared with those transformed with empty pGBKT7 vector. Table 3.4 showed that expression of the baits did not hamper the growth of Y187 yeast. The transcriptional activity of the

baits was tested by streaking AH109 and Y187 transformed with the bait constructs on SD/-Ade/-Trp and SD/-His/-Trip agar. The typical transcriptional activity test for one of the baits is shown in Figure 3.12. All of the baits did not grow on medium lacking adenine or histidine. However, AH109 transformed with GBKT7-ARM and pGBKT7-CrSARM-AST showed weak growth on SD/-His/-Trp. Nevertheless, the interference of the low background growth on the –His medium was eliminated by using minimal medium lacking both adenine and histidine (SD/-Ade/-His/-Leu/-Trp or QDO) to select for the positive hybrids. Therefore, pGBKT7-CrSARM-AST, -ARM, -SAM, -TIR constructs were used as baits to screen the amebocyte 3 hpi cDNA library.

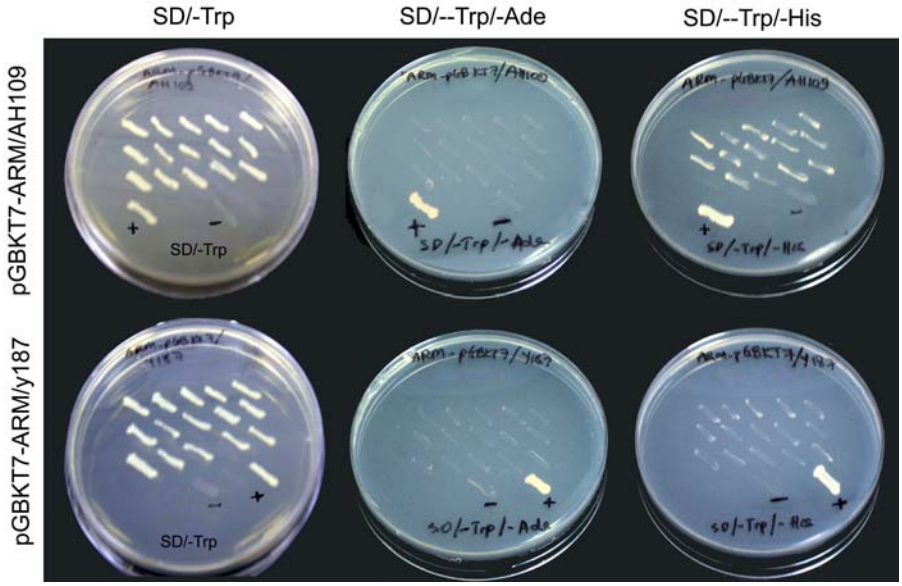


Figure 3.12 Checking for transcriptional activity of pGBKT7-ARM

S. cereviseae Y187 and AH109 were transformed with pGBKT7-ARM plasmid. The transformed cells were selected on SD/-Trp agar. For each yeast strain, 15 clones were selected for transcriptional activity test. Those clones were firstly streaked to SD/-Trp agar plates then replicated to SD/-Ade/-Trp and SD/-His/-Trp agar plates. Y187 and AH109 transformed with pGBKT7-53 provided by the manufacturer were included as positive controls (+). Negative controls (-) were the non-transformed Y187 or AH109. All plates were incubated for 4 days at 30 ^{o}C.

Table 3.4 Testing the baits for toxicity

Y187 Clone	OD_{600}
pGBKT7	1.57
pGBKT7-CrSARM-AST	1.38
pGBKT7-ARM	1.52
pGBKT7-SAM	1.48
pGBKT7-TIR	1.63

An aliquot of 50 ml of SD/-Trp/Kan (20 μg/ml) liquid medium was inoculated with one colony of Y187 transformed with empty pGBKT7 or different pGBKT7-bait constructs. The culture was incubated at 30°C with 250 rpm shaking for 24 h before OD_{600} was measured.

Screening of the amebocyte 3 hpi cDNA library by yeast mating

The amebocyte 3 hpi cDNA/AH109 library was screened with the CrSARM-AST, ARM, SAM and TIR baits using the yeast mating method. To obtain an effective screening, at least 5 x 10^9 bait clones were used in each mating with 1.2 x 10^8 clones of preys (Table 3.5). In all of the mating, the three-lobed shaped zygotes (Figure 3.13 A) were observed after 20 h of incubation, indicating that the budding diploid cell had not detached from the two haploid parental cells. Yeast mating was hence carried out for another 4 h. A small aliquot of the mating mixture was used for determination of mating efficiency, the rest was spreaded on SD/-Ade/-His/-Leu/-Trp (QDO) agar plates for the selection of positive hybrids. Table 3.5 shows the mating efficiencies of the four baits. Although the mating efficiencies of SAM and TIR baits were not very high, they were in the acceptable range for a successful identification of novel protein-protein interaction. Indeed, the mating efficiencies of 6 % and above gave more than 7 x 10^6 diploids for the selection of positive hybrids considering 1.2 x 10^8 library clones were screened. The mating efficiency of CrSARM-AST and ARM baits, however, were much lower though comparable amount of baits were used for each mating. In addition, the efficiency did not improved when the mating was repeated with ten times higher amount of ARM bait. Low mating efficiencies of these two baits were therefore not due to either the toxicity of the BD-bait fusion proteins (Table 3.4) or the insufficient number of the bait clones. This leaves the possibility

that the ARM domain, which was present in both of the baits, interferes with the mating process. This hindrance may be overcome if library screening is done by transforming the pGBKT7-bait plasmids into the cDNA library in AH109 cells. The transformants in which the bait and the prey interact can be selected on QDO medium similarly to the selection for the positive hybrids. This could be a future approach to solve this problem.

Table 3.5 Screening of amebocyte 3 hpi cDNA library by the yeast mating method

Bait clone	Cell count of baits as determined using the hemocytometer (cell / ml)	Mating efficiency (%)	Number of hybrids grew on QDO plates	Number of hybrids grew and turn blue on QDO-Xgal plates
pGBKT7-CrSARM-AST	1.2×10^9	2.6	22	0
pGBKT7-ARM	1.4×10^9	0.75	18	2
pGBKT7-SAM	1.3×10^9	6.0	31	16
pGBKT7-TIR	1.7×10^9	6.9	85	43

pGBKT7-bait/Y187 clone was cultured in 50 ml of SD/-Trp/Kan (20 μg/ml) liquid medium for 24 h at 30 ^{o}C with 250 rpm shaking. Cells were collected and resuspended in 5 ml of 2 x YPDA/Kan (50 μg/ml) broth. Cell count was performed with the hemocytometer. The bait suspension was combined with 1 ml of the amebocyte 3 hpi cDNA library for mating for 24 h at 30 ^{o}C. Viabilities (cfu/ml) of the bait, the library and the diploids were determined by spreading the mating mixture on SD/-Trp, SD/-Leu and SD/-Leu/-Trp agar, respectively. Mating efficiency was calculated as the percentage of the diploid's viability to the limiting partner's viability, which was the library in this case. Putative hybrids were first selected on QDO plates, and further validated by restreaking to QDO-Xgal plates.

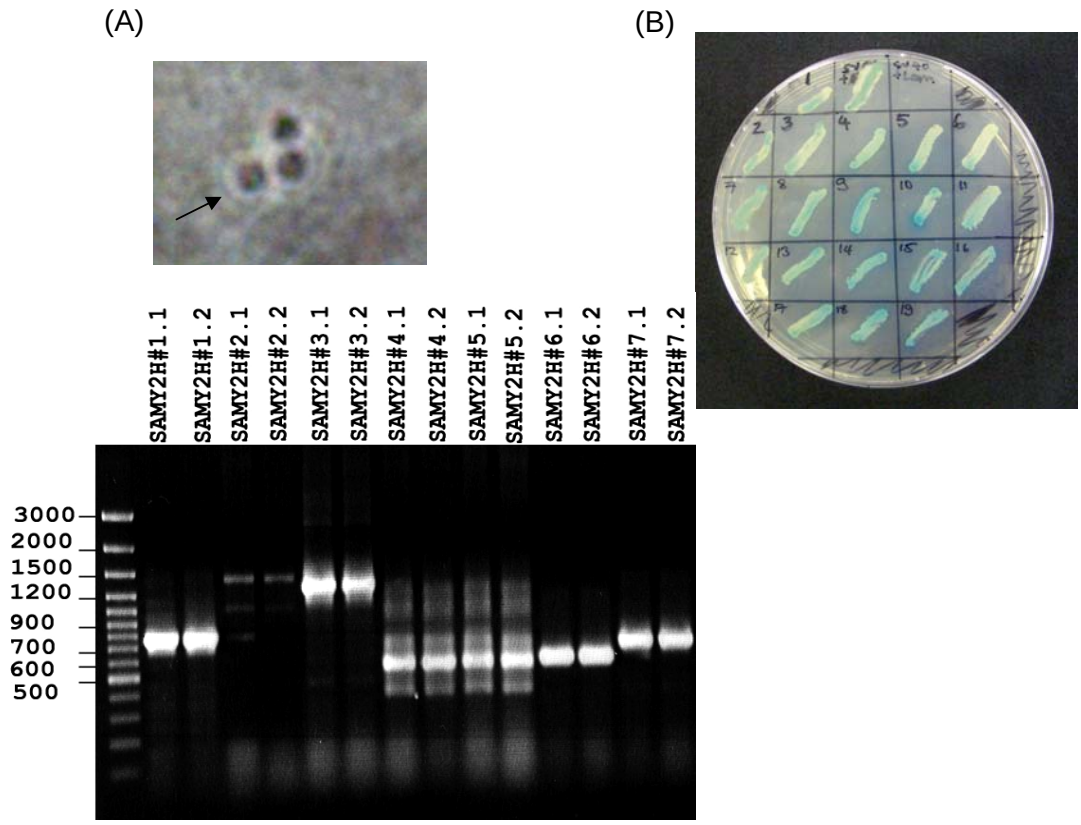


Figure 3.13 Screening of the amebocyte 3 hpi cDNA library by yeast mating

(A) The zygote with the tri-lobe shape is indicated by the arrow. (B) Checking the diploids on QDO-Xgal plates. Positive diploid should turn blue due to the activation of *LacZ* reporter gene. (C) Colony PCR of *E. coli* clones transformed with DNA isolated from the positive hybrids of SAM bait using primers flanking cDNA inserts in the pGADT7 vector.

The numbers of diploid cells that grew on QDO plates were consistent with the mating efficiencies (Table 3.5). Approximately half of the diploids isolated from the screenings with TIR and SAM baits grew to the diameter of > 1 mm within 14 days of incubation at 30 ^{o}C; the rest and all the diploids of the screenings with CrSARM-AST and ARM baits took at least 16 days to grow to > 1 mm in diameter. Most of the slow growing colonies were actually the false positives since they did not grow or grew but did not turn blue when being restreaked to fresh QDO agar plates containing X-α-gal (Table 3.5 & Figure 3.13 B).

pGADT7-cDNA plasmids were rescued from the positive diploids, which were 43 for the TIR-bait, 16 for the SAM bait, and 2 for the ARM bait, and transformed into *E. coli* Top10. Two bacterial clones were randomly selected from each transformation for colony PCR with the primers flanking the respective cDNA insert. Results showed that clones transformed with DNA from the same yeast colony have inserts of the same size while those transformed with DNA from different yeast colonies have inserts of different sizes (Figure 3.13 C), indicating that each yeast hybrid seemed to harbor one type of pGADT7-cDNA plasmid and each bait may have more than one putative interaction partner.

The isolated preys were identified by sequencing followed by blast search. The result is summarized in Table 3.6. Since some of the preys were expressed in more than one hybrid, the numbers of non-redundant potential interaction candidates of TIR, SAM, and ARM domains were 26, 11 and 2, respectively. Although the yeast two hybrid system with three reporter genes (*Ade1, His3* and *LacZ*) should have given lower number of false positives than the older system with only one reporter gene, the isolation of false positives is still the major problem associated with this application. The next step after obtaining the sequencing data of the preys was therefore to identify the interactions of potential interest for subsequent analysis of interaction specificity (Van Criekinge and Beyaert, 1999). Putative interaction partners of each bait were categorized into two groups: the ones of potential interests and the others (Table 3.6). Genes that are included in the later group are those encoding commonly found false positives reported in literature such as ribosomal proteins (Van Criekinge and Beyaert, 1999); those repeatedly isolated by our lab from the yeast two hybrid screening using unrelated baits (cathepsin-L, coagulogen, ring finger proteins 36); those encoding biologically irrelevant proteins such as proteins involved in

metabolism (fumarylacetoacetate hydrolase, quinine reductase, alcohol dehydrogenase) or proteins of different localization from CrSARM such as glycoprotein hormone alpha 2 and factor C, which are secretory proteins. In addition, genes that were not inframe with the AD domain or gave no significant match were also omitted from further analysis. Therefore, further characterization was performed with the clones belong to the "Potential interest" group to verify the specificity of the interactions.

3.8.2 Yeast co-transformation confirms the interaction partners of CrSARM and eliminates the false positives

The putative interactions were further verified by co-transformation. Individual pGADT7-prey was co-transformed with the correspondent pGBKT7-bait into AH109 yeast to accessed the reproducibility of check whether the interaction. At the same time, the specificity of the interaction was analyzed by co-transformations of the pGADT7-prey with the empty pGBKT7 vector, and the pGBKT7-baits with the empty pGADT7 vector. Analysis shows that co-transformation allowed the identification of some false positives among the potential interaction candidates. Indeed, co-transformation of some preys with their correspondent baits did not result in the transformants that could grow on the QDO plate (Table 3.6) although the presence of the two plasmids in those transformants was ensured by the growth of yeast on SD/-Leu/-Trp agar before transferring to the QDO plates (Figure 3.14). Other group of false positives was the ones that activated the expression of reporter genes in the absence of the baits as shown by the growth of yeast harboring the prey and the empty pGBKT7 plasmids on QDO medium (Figure 3.14 & Table 3.6). In Table 3.6, the interactions that can be re-established and shown to be specific by co-transformation are those with all of the three last columns shaded.

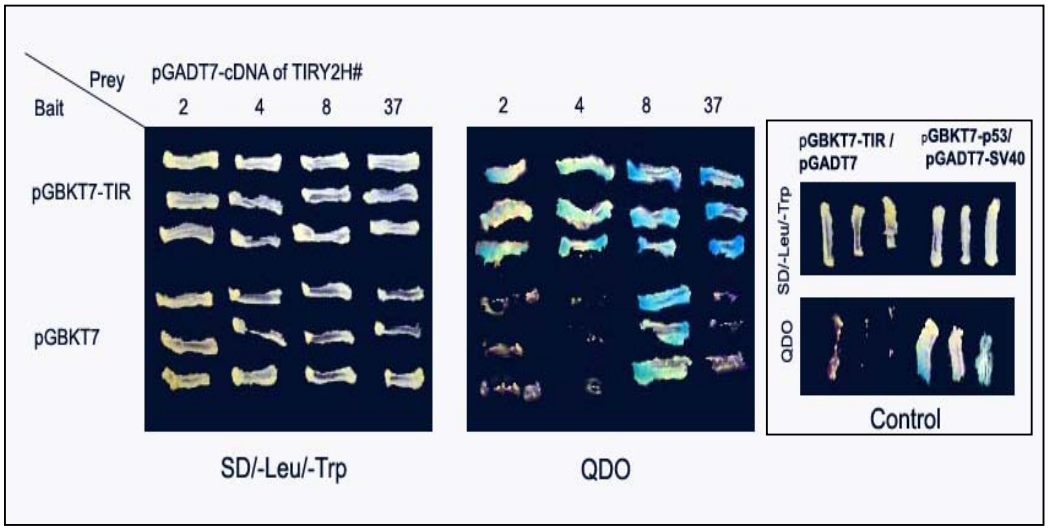


Figure 3.14 Analysis of interaction specificity by yeast co-transformation

The pGADT7-cDNA plasmids that were isolate from the putative diploids of the TIR baits. were co-transformed with either pGBKT7-TIR or the pGBKT7 vector backbone. into AH109 yeast. Co-transformations of pGBKT7-TIR and pGADT7 vector backbone, and pGBKT7-p53 and pGADT7-SV40 (provided by the manufacturer) were set up as the negative and positive controls, respectively. The transformants were selected on SD/-Leu/-Trp for the expression of *Leu2* and *Trp1* genes, which are the reporter genes harbored on pGADT7 and pGBKT7 vectors, respectively. Three transformants were then restreaked to fresh SD/-Leu/-Trp and QDO plate for testing of interaction specificity. The plates were incubated at 30 ^{o}C for four days before the growth of yeast were recorded. Result of co-transformation assay is summarized in Table 3.6

Result of the co-transformation analysis (Table 3.6) showed that only six "preys" were found to interact specifically with TIR domain. One of them was HS1-associated protein X-1 (HAX1). Hax1 got its name from the interaction with HS-1, which is expressed exclusively in the cells of heamatopoetic lineage and potentially functions in signal transduction from the antigen receptors in B cells (Suzuki et al., 1997). Though the exact function is unknown, it was speculated that HAX1 could involved in the regulation of apoptosis since it was found to interact with the anti-apototic proteins such as Bcl2 and the viral homolog of Bcl2: BHRF1 (Matsuda et al., 2003). In addition, HAX1 may play a role in cytoskeleton remodeling as suggested by its interaction with cortactin, which is an F-actin-associated protein that is closely related to HS1 (Gallagher et al., 2000). Another interesting putative interaction

partner of TIR domain is SUMO-1, which was shown to regulate wide variety of cellular processes by covalent attachment to the lysine residue of the target proteins or physical interaction with the target proteins. SUMO-1 modification, which is also known as sumoylation, very often results in the repression of the cellular process due to the alteration in the molecular interaction properties and/or the subcellular localization of the target proteins. Though, majority of proteins regulated by sumoylation is transcription factors, sumoylation of signal transducing molecules such as Fas, tumor necrosis factor receptor 1, apoptosis signal regulated kinase (ASK1) had been reported (Lee et al., 2005). Proteasome alpha subunit and heat shock protein 40 (Hsp40), which are structural components of the proteasome and chaperone machinery, respectively, are also amogst the putative interaction partners of the TIR domain. This may suggest the potential roles of CrSARM in protein degradation and modification. In addition, two out of six proteins showed specific interaction with TIR domain are hypothetical proteins. Further study is needed to clarify the identities as well as the functions of these proteins.

Amogst nine positive clones isolated from yeast two hybrid screening with the SARM domain bait, only calcium/calmodulin-dependent protein kinase I (CaMKI) showed specific interaction with SAM domain in the co-transformation test. Proteins of CaMK family are activated in response to the change in intracellular Ca^{2+} concentration as a result of the activation of Ca^{2+}-binding proteins such as Ca^{2+}/calmodulin. With a broad range of substrates, CaMKs can mediate signal transduction leading to the activation of a wide variety of cellular processes (Fujisawa, 2001). The fact that infection by some pathogens leads to the change of cytosolic Ca^{2+} concentration (Brunet et al., 2000; Haque et al., 1998; Low et al., 1992) suggests the potential function of CaMKs in host defense. Indeed, recently, Verploegen *et al*

reported that CKliK, a novel CaMKI exclusively found in the granulocytes (Verploegen et al., 2000), plays a critical role in cytokine-responsive functions of the granulocytes such as phagocytosis, cell migration and adhesion (Verploegen et al., 2005).

Table 3.6 List of putative interaction candidates of CrSARM

(A)

Clone ID TIRY2H#	Blast match + e-value	E-value	Accession number	Prey + Bait	Prey + pGBKT7	Bait + pGAD7Rec
	Potential interest					
3, 2, 9, 18	Proteasome alpha subunit [*Ornithodoros moubata*]	7e-37	AAS01024	√	X	X
4	HS1-associated protein X-1 (HAX1) [*Homo sapiens*]	5e-7	AAB51196	√	X	X
37	Small ubiquitin-like modifier, SUMO-1 (aa.7-44/98) [*Apis mellifera*]	7e-10	XP_392826	√	X	X
10, 16, 17, 38,42	DnaJ (Hsp40) homolog, subfamily A, member 1 [*Homo Sapiens*]	4e-117	NP_001530	√	X	X
21	Hypothetical protein [*Drosophila pseudoobscura*]	0.066	EAL32171	√	X	X
30	Hypothetical protein [*Yarrowia lipolytica*]	0.67	XP_502158	√	X	X
8	Tyrosine kinase-interacting protein [*Saimiriine herpesvirus 2*]	0.04	CAC84292	√	√	X
28**,** 48	Adenylate cyclase-associated protein 1[*Danio rerio*]	5e-41	NP_956203	√	√	X
49	Activated protein kinase C receptor; RACK1 [*Xenopus laevis*]	2e-108	AAD42045	√	√	X
11	Novel ubiquitin-like protein [*Danio rerio*]	4e-17	CAD60858	√	√	X
20, 24, 36	Hypothetical protein [*Strongylocentrotus purpuratus*]	1e-27	XP_800816	√	√	X
40	Similar to protein tyrosine phosphatase [*Danio rerio*]	0.003	XP_698134	X	X	X
13	Hypothetical protein LOC68201 [*Mus musculus*]	9e-7	NP_080889	X	X	X
	Others					
50	Zinc finger protein 36 [*Rattus norvegicus*]	0.13	NP_001031703			
15	U2 small nuclear ribonucleoprotein B [*Gallus gallus*]	3e-38	XP_419331			
3,19, 46, 52	46 Ribosomal protein S20 [*Crassostrea gigas*]	3e-32	CAD91428			
14	Glycoprotein hormone alpha 2 [*Drosophila melanogaster*]	8e-30	AAX38184			
1	Flocculin-like protein [*Candida albicans*]	0.018	XP_712591			
29	Methionine sulfoxide reductase B2b [*Schistosoma mansoni*]	3e-40	AAT77264			
44	Cystallin, zeta (quinine reductase) [*Xenopus tropicalis*]	2e-51	NP_00100568			
27	Fumarylacetoacetate hydrolase domain containing [*Mus musculus*]	2e-29	XP_547188			
31	Cathepsin-L [*Strongylocentrotus purpuratus*]	2e-67	XP_801784			
25	Coagulogen type II [*Tachypleus tridentatus*]	3e-51	CAA27780			
34	Alcohol dehydrogenase [*Gryllotalpa orientalis*]	4e-23	AAR84629			
43, 45, 54,55	No ORF found in frame with AD domain					
23, 33	No significant match					

(B)

Clone ID SAMY2H#	Blast match	E-value	Accession number	Prey + Bait	Prey + pGBKT7	Bait + pGAD7Rec
	Potential interest					
3	Calcium/calmodulin-dependent protein kinase 1D [*Rattus norvegicus*]	8e-5	XP_225567	√	X	X
4, 5, 9	Hypothetical protein [*Burkholderia cenocepacia*]	4e-7	ZP_00464021	√	√	X
16	Hypothetical protein LOC496712 [*Xenopus tropicalis*]	7e-28	AAH88681	X	X	X
11	Sulfatase domain protein 3 [*Caenorhabditis elegans*]	1e-24	AAA79198	X	X	X
	Others					
1, 15	Coagulogen type II [*Tachypleus tridentatus*]	3e-97	CAA27780			
17	Factor C [*Carcinoscorpius rotundicauda*]	8e-40	AAB34362			
2	Bahma associated protein 60 kDa [Drosophila melanogaster].	1e-132	AAC28455			
13	Mitochondria ribosomal protein L50 [*Homo sapiens*]	4e-15	CAI14610			
7, 12,14	Zinc finger protein 36 [*Rattus norvegicus*]	0.13	NP_001031703			
8	ribosomal protein S16 [*Heteropneustes fossilis*]	6e-37	AAK11731			
6	No significant match					

(C)

Clone ID ARMY2H#	Blast match	E-value	Accession number	Prey + Bait	Prey + pGBKT7	Bait + pGAD7Rec
2	8.6 kDa Transglutaminase substrate [*Tachypleus tridentatus*]	7e-41	AAB24710	X	X	X
1	Zinc finger protein 36 [*Rattus norvegicus*]	0.14	NP_001031703			

The preys isolated by the yeast two hybrid screening of TIR domain (A), SAM domain (B) and ARM motif (C) were sequenced. ORFs that were inframe with the AD domain encoding by pGADT7 vector were used for protein-protein blast (Blastp) search to identify the homologous proteins. The preys were divided into two categories, namely "the potential interest" and "the others" based on the criteria that described in pages 83-84. Potentially biological relevant interactions were further confirmed and validated for interaction specificity by co-transformation with either the correspondent pGBKT7-bait plasmid or empty pGBKT7 vector into AH109 yeast. Co-transformation of the bait with empty pGADT7 vector was included as another negative control. The co-transformation yeasts were first selected on SD/-Leu/-Trp agar. Three colonies were then replicated on fresh SD/-Leu/-Trp and QDO plates. All plates were incubated for 4 days at 30 ^{o}C before the growth on QDO plates were recorded. (√): growth; (X): no growth; gray shading: expected outcome if the interaction is specific.

Chapter 4 Discussion

4.1 CrSARM – a signaling molecule responsive to *P. aeruginosa* infection

SARM was first identified in human and its homologs were found in mouse, *Drosophila* and *C. elegans* (Mink et al., 2001). Since then, SARM quickly became a molecule of research interest because of its unique domain structure composing three domains: ARM, SAM and TIR, all of which are well-known for their ability to mediate protein interaction. In addition, in SARM due to the presence of TIR domain, which is present exclusively in molecules involved in IL-1 and TLR signaling pathways, it is believed that SARM may be a novel immune-related signaling protein (McGettrick and O'Neill, 2004; O'Neill et al., 2003). Indeed, Liberati et al (Liberati et al., 2004) and Couillault et al. (Couillault et al., 2004) showed the involvement of *C. elegans* SARM homolog, TIR-1, in immune response against bacterial and fungal infections in the worm. Nevertheless, the potential role that SARM plays in embryonic development was also reported (Chuang and Bargmann, 2005).

Our work further confirms the participation of this novel signaling molecule in the innate immune response. Indeed, the horseshoe crab SARM was identified as one of the signaling molecules whose expressions were differentially regulated in response to *Pseudomonas* infection (Ding et al., 2005). The kinetics of this regulation, which was studied by RT-PCR, reveals that the transcription of CrSARM was induced during the first 3 h after *Pseudomonas* infection, followed by a suppression for the next 6 h before returning to the basal level at 12 hpi (Figure 3.8). On the contrary, Meijer *et al* reported that the zebra fish SARM (zSARM) transcription level did not change upon the infection with *Mycobacterium marinum*, a gram positive

bacteria that is closely related to *M. tuberculosis* (Meijer et al., 2004). This deviation may be due to either of the following reasons. Firstly, the transcript level of zSARM may revert to the basal level at the eighth day after infection, which was the time point reported by Meijer *et al* (Meijer et al., 2004). Other possibility could be that SARM may not be involved in immune response against gram positive bacteria, but is only restricted to host defense against gram negative bacteria and fungi as shown by Couillault *et al* (Couillault et al., 2004) and our work. Biological significance of the change in transcription level of CrSARM upon *Pseudomonas* infection will be elaborated further in sections 4.4.1 and 4.4.2.

4.2 What can be derived from the comparison of sequence homology between CrSARM and SARM homologs of other organisms?

Multiple sequence alignment shows that the invertebrate SARMs, including CrSARM, share high homology with the vertebrate homologs (Figure 3.5). In addition, all SARM homologs that have been identified so far harbour the same modular structure: a combination of an N-terminal ARM motif and a C-terminal TIR domain that flanks the two consecutive SAM motifs (Jault et al., 2004; Liberati et al., 2004; Meijer et al., 2004; Mink et al., 2001). The conservations of protein sequence and domain architecture of SARM homologs across the animal kingdom strongly indicates that some if not all the functions of SARM have been retained through evolution, justifying the extrapolation of the findings in the ancient invertebrate like the horseshoe crab to the mammals.

Interestingly, the invertebrate SARM homologs harbour the N-terminal sequences that are missing in their vertebrate counterparts. Indeed, the N-terminal sequences of CrSARM and the *Drosophila* SARM homolog, Ect4, are almost 200

amino acid longer than that of human and mouse SARMs. TIR-1, SARM homolog in the *C. elegans*, also has around 50 residues extra at its N-terminus (Figure 3.5). Consistently, the phylogenetic analysis shows that SARM protein family can be divided into three subgroups: the ones from the *C. elegans*, the ones from other invertebrates other than *C. elegans*, and those from the vertebrates (Figure 3.6). The significance of the N-terminal sequences, which do not encode any functional domain, in the invertebrate SARMs is unknown. However, based on the above observations, one can speculate that the invertebrate SARMs may have additional function besides that conserved from worm to mammals. An example of proteins that are involved in more functions in the invertebrates than in the vertebrates is the Toll like receptors (TLR). The *Drosophila* Toll functions in both embryonic dorso-ventral patterning process (Morisato and Anderson, 1995) and in the immune responses against bacterial and fungal infection (Lemaitre et al., 1996; Rutschmann et al., 2002). Its homolog in vertebrates, however, exclusively involved in host defense (reviewed in Akira and Takeda, 2004; Takeda and Akira, 2005)

4.3 Difference in tissue specific expression may reflect differences in function between the invertebrate and vertebrate SARM homologs

Expression of the horseshoe crab SARM analyzed by RT-PCR was restricted to the hepatopancreas, amebocyte, muscle and heart (Figure 3.7). Human SARM, however, was detected only in the kidney and liver, which is equivalent to the hepatopancreas of the horseshoe crab. The lack of expression of human SARM in other tissues such as the brain and some cancer cell lines is possibly due to the silencing effect caused by the antisense RNA of SARM detected at high level in these tissues (Mink et al., 2001). It is likely that a similar mechanism of repression regulates

the expression of CrSARM in different tissues of the horseshoe crab. In *C. elegans*, SARM homolog was found in the embryo and the hypodermis, where it involved in the regulation of the differentiation of olfactory neurons (Chuang and Bargmann, 2005) and the expression of antimicrobial peptides (Couillault et al., 2004), respectively. It can be seen that the expression of SARM in immune-responsive tissues has been conserved from the worm to human, suggesting that the immune-related function of this protein has been maintained through evolution. Whether SARM functions in the embryogenesis in organisms higher than *C. elegans* and why it is expressed in the horseshoe crab's muscle and heart remain to be solved.

4.4 SARM, a distinct TIR domain containing adaptor

SARM is the newest member of the TIR domain-containing adaptor proteins. The presence of TIR domain in this protein suggests its potential role in the TLR signaling. However, until now, the link between SARM and the TLR signaling pathways has not been established. Studies in *C. elegans* showed that TIR-1, the mammalian SARM homolog, regulates the expression of antimicrobial peptides in Tol-1 independently manner although TIR-1 and Tol-1 are the only two TIR domain-containing proteins predicted to be present in *C. elegans* (Couillault et al., 2004; Liberati et al., 2004). However, while Tol-1 does not exhibit immune-related function (Couillault et al., 2004), *Drosophila* Toll and TLRs in higher organisms do. Therefore, it is a worthwhile effort to verify the potential role of SARM in TLR signaling in higher organisms.

The very unique domain architecture of SARM with three protein-interaction domains (ARM, SAM and TIR) strongly suggests that SARM is a component of novel signaling pathways. Indeed, this is reflected from our study of CrSARM. A spectrum of potential protein interaction partners of CrSARM was isolated by yeast

two hybrid screening. Subsequently, co-transformation confirmed that only calcium/calmodulin dependent protein kinase I (CaMKI), HS1-associated protein X-1 (HAX1), small ubiquitin-like modifier (SUMO-1), Hsp40, proteasome alpha-subunit, and two hypothetical proteins interact specifically with CrSARM (Table 3.6). The diverse cellular functions of these SARM-interaction protein partners suggest that CrSARM may mediate signal transductions within more than one signaling pathway. The following sections will discuss the potential biological significance of these interactions based on our current understanding of SARM; the biological functions of the interaction partners; as well as the innate immune response of the horseshoe crab.

4.4.1 CrSARM may be involved in Ca^{2+}-dependent regulation of p38 MAPK

Chuang and Bargmann (Chuang and Bargmann, 2005) reported that during embryonic development in *C. elegans*, TIR-1, the SARM homolog of *C. elegans*, provides surface for the signaling complex downstream of a voltage-gated calcium channel to assemble and to regulate the expression of the odorant receptor, thus, establishing the asymmetry of the olfactory AWC neurons. UNC-43, a homolog of CaMKII, and NSY-1, a homolog of human ASK1, were found to belong to the complex. However, physical interaction was only confirmed between TIR-1 and UNC-43 though the co-localization of TIR-1, UNC-43 and NSY-1 to the post-synaptic regions was observed (Chuang and Bargmann, 2005). Since ASK1 is a mitogen-activated protein kinase (MAPK) kinase kinase that can activate both JNK and p38 MAPK signaling pathways, it was thought that TIR-1 is involved in the regulation of the odorant receptor expression via the UNC-43/NSY-1/p38 or JNK cascade. Consistent with Chuang and Bargmann (Chuang and Bargmann, 2005), Leberati et al (Liberati et al., 2004) reported that the suppression of TIR-1 expression

by RNA interference repressed the activation of p38 MAPK, reducing resistance of *C. elegans* to pathogen infection.

In the yeast two hybrid screening of the amebocyte 3 hpi cDNA library using the SAM domain of CrSARM as the bait, we isolated **calcium/calmodulin dependent protein kinase I (CaMKI)** (B). Although the *C. elegans* SARM homolog, TIR-1, was shown to interact with CaMKII to modulate the expression of olfactory receptors in neuronal cells (Chuang and Bargmann, 2005), there are data in support of the interaction between CrSARM and CaMKI. Indeed, while CaMKII is mainly expressed in the brain, CaMKI is found in the cytosol of a wide variety of tissues (Tobimatsu and Fujisawa, 1989). Moreover, a CaMKI-like kinase expressed exclusively in human granulocytes, which are the functional equivalence of the horseshoe crab amebocytes, was isolated several years ago (Verploegen et al., 2000). It is therefore likely that in non-neuronal cells including the amebocytes, interaction between CrSARM and CaMKI would mediate various cellular functions in response to the change in intracellular Ca^{2+} level. It has been known that infection can trigger the alteration in free cytosolic Ca^{2+} concentration, resulting in the attenuation of a variety of immune-related processes including degranulation, phagocytosis, and cytoskeleton remodeling (Brunet et al., 2000; Haque et al., 1998; Low et al., 1992). Nevertheless, further study is required to determine how CrSARM-CaMKI complex may contribute to the immune responses against bacterial infection.

Although, the downstream target of the horseshoe crab CaMKI remains unknown, the finding of Chuang and Bargmann (Chuang and Bargmann, 2005) suggests that ASK1 could be a potential candidate. Interestingly, **SUMO-1,** one of the potential interaction partners of the TIR domain of CrSARM (Table 3.6 A), was shown in a recent research to physically interact with ASK1 and inhibit its activation

(Lee et al., 2005). Interaction between SUMO-1 and CrSARM, therefore, would be essential to down-regulate the Ca^{2+}/CaMKI/ASK1/p38 or JNK signaling cascade to prevent the damage of the host tissue and the wasting of cellular resources due to the overwhelming immune responses activated by this signaling cascade. Besides SUMO-1 mediated down-regulation of the downstream signaling, suppression of CrSARM expression at the later stage of infection (Figure 3.8) could be an additional mechanism to suppress this signaling pathway. The presence of two negative check-points would assure that the downstream signaling is repressed quickly and efficiently. Taken together, we speculate that upon infection, CrSARM may function as a scaffold protein supporting the assembly and signal transduction through the CaMKI/ASK1/(p38/JNK) cascade, ultimately resulting in the activation of immune related genes that are under the control of AP-1 transcription factor, which is the downstream target of p38 MAPK. When the invading pathogen is under control, down-regulation of this signaling would be achieved by the recruitment of SUMO-1, a negative regulator of ASK1, to the TIR-domain of CrSARM and the inhibition of CrSARM expression itself. This hypothesis is schematically illustrated in (Figure 4.1). This speculation can be validated by further investigation: (a) to confirm the link between ASK1 with CaMKI and SUMO-1; (b) to determine whether the binding of CaMKI, SUMO-1 and possibly ASK1 to CrSARM is constitutive or temporally regulated; (c) to define the signaling that controls the switching between activation and repression decision; and (d) to identify the ultimate immune responses mediated by this signaling cascade.

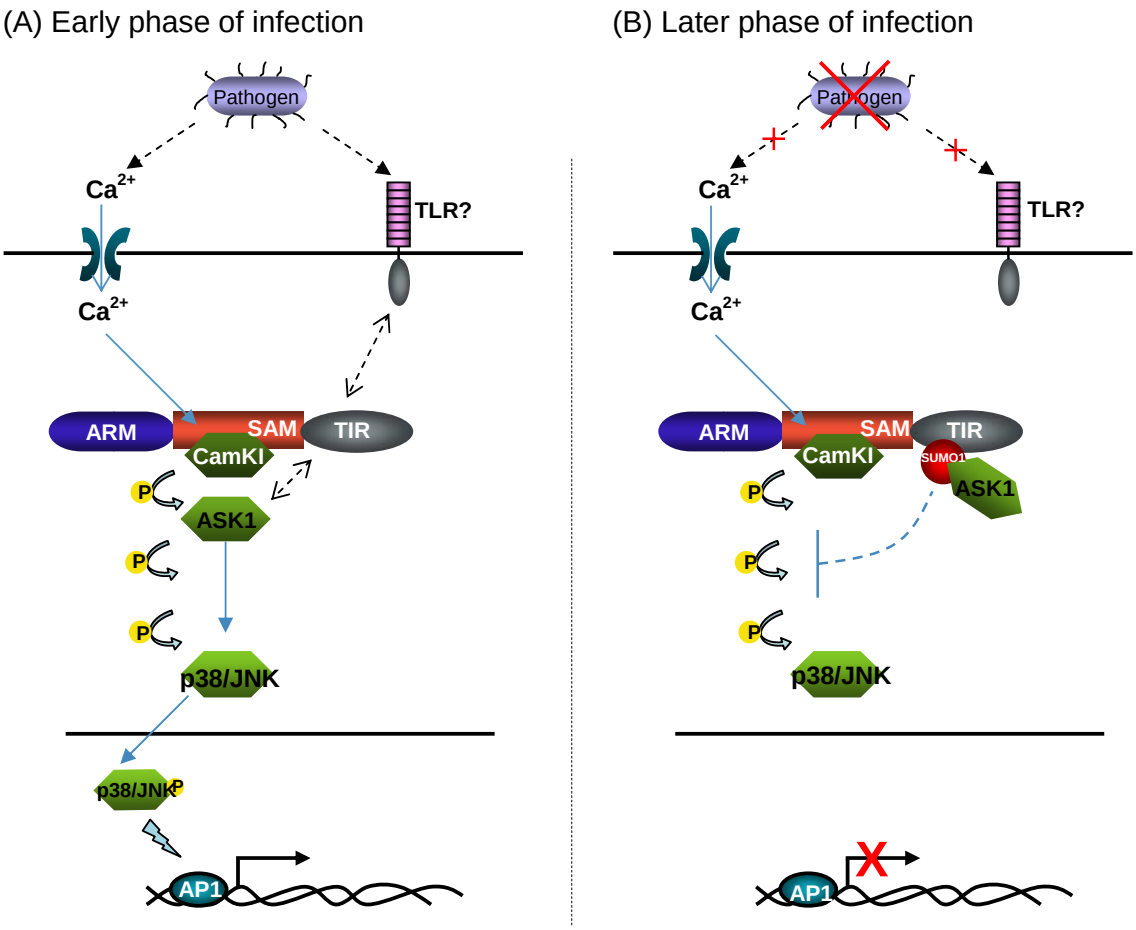


Figure 4.1 CrSARM regulates Ca^{2+}-dependent activation of p38/JNK kinases in response to pathogen infection via the interactions with CaMKI and SUMO-1

(A) Infection causes the influx of extracellular Ca^{2+} into the cell via the ion channels and simultaneously activation of TLR or an unknown PRR upstream of CrSARM, promoting the interaction between CrSARM with CaMKI and ASK1. At the same time, the increase in intracellular Ca^{2+} concentration leads to the activation of CaMKI, which in turn activates ASK1, an upstream component of p38 and JNK kinase signaling cascades. As a result, p38 and/or JNK are phosphorylated and translocate into the nucleus to activate the transcription factor AP1, resulting in the expression of AP1 dependent immune-related genes. (B) To avoid tissue damage and exhaustion of cellular resources due to overwhelming immune response, the signaling cascade is repressed by two mechanisms: (1) downregulation of CrSARM expression, and (2) inhibition of ASK1 activation by SUMO-1, which was shown to interact with the TIR domain of CrSARM. Solid arrows represent processes that are confirmed experimentally. Dotted arrows and dotted line represent interactions and processes that require further validation.

4.4.2 Does CrSARM play a role in the regulation of apoptosis or remodeling of cytoskeleton?

We identified **HAX1** amongst putative interaction partners of CrSARM TIR domain. As mentioned previously, HAX1 was thought to be involved in the regulation of apoptosis and cytoskeleton remodeling signaling. This is based on its ability to interact with a prosurvival protein, Bcl2, and the F-actin associated protein cortactin, respectively (Gallagher et al., 2000; Matsuda et al., 2003). Both apoptosis and cytoskeleton remodeling highly relate to pathogenesis as well as host defense.

CrSARM-HAX1 interaction may be essential for the host resistance against pathogen-induced apoptosis

Apoptosis of the host cells that are infected by pathogens is a sacrificial immune response to protect the non-infected cells from pathogen invasion (DosReis and Barcinski, 2001; Koyama et al., 1998). Host induced apoptosis is particularly important for an effective control of viral and obligate intracellular parasite infection. However, several pathogens including *P. aeruginosa* exploit the host signaling pathways to induce the death of immune related cells, and thus impair the host defense. Indeed, researchers have shown that several proteins produced by *P. aeruginosa* such as quorum sensing molecule N-3-oxododecanoyl homoserine lactone, pyocyanin and exotoxin A toxin, are able to induce apoptosis of macrophages and/or neutrophils (Allen et al., 2005; Jenkins et al., 2004; Tateda et al., 2003). Exotoxin A was shown to trigger a caspase dependent signaling pathway, which can be suppressed by Bcl-2, whereas molecular mechanisms associated with proapoptotic activity of the two former *Pseudomonas* proteins remain unclear.

In the horseshoe crabs, the amebocytes play a very active role in the clearance of the infected pathogen (Ding et al., 2005; Ng et al., 2004). Hence, induction of the

amebocyte cell death would favor the survival of the *P. aeruginosa* inside the horseshoe crab. It is thus tempting to speculate that CrSARM/HAX1 complex, through the interaction with Bcl2, would inhibit the cell death signaling initiated in the amebocytes by the *Pseudomonas*. In this context, CrSARM would serve as an adaptor to couple the undefined upstream signaling pathway, which is triggered by the pathogen, with the downstream anti-apototic signaling cascade. In line with this hypothesis, we found that the change in the transcript level of CrSARM in response to *Pseudomonas* infection (Figure 3.8) synchronizes with the profile of bacterial clearance from the infected horseshoe crabs, which was shown in the previous study in the lab (Ng et al., 2004), and Figure 1.8). It seems that the upregulation of CrSARM during the first three hours of infection occured to protect of the cells from apoptosis induced by the great burden of the invading bacteria. After 3 hpi, the majority of the pathogens was cleared from the infected horseshoe crabs, thus lifting the proapoptosis driving force, and hence the anti-apoptotic activity is no longer required. This is coincident with the rapid reduction in transcription level of CrSARM, which would allow the level of CrSARM protein return to the basal level.

The potential role of CrSARM in cytoskeleton remodeling

We should not rule out the possibility that the interaction between CrSARM and HAX1 may play a role in the remodeling of actin cytoskeleton due to the possible interaction between the latter with cortactin, which is an F-actin binding protein. The expression of CrSARM in tissues with actin-based contractile capacity like muscle and heart (Figure 3.7) and the presence of cortactin in human skeletal and cardiac muscles as shown by Wu and Montone (Wu and Montone, 1998) support this argument. In addition to muscle contraction, actin dynamics play a significant role in many cellular processes including cell migration, membrane dynamics, and cellular

transport. If the interaction between CrSARM, HAX1, and cortactin truly happens *in vivo*, CrSARM would serve as a bridge to link the upstream signaling, which can be a physiological and/or immune-related, to the regulation of actin cytoskeleton, which would effect various actin-dependent cellular processes.

Among the essential host defense responses, phagocytosis and receptor-mediated endocytosis are the two processes that are tightly associated with cytoskeleton remodeling. The phagocytosis-related function of CrSARM is supported by the following lines of evidence. A recent research has revealed that activation of TLR may enhance the antigen-endocytosis of dendritic cells via the modulation of podosome turn-over. Podosome is an F-actin rich structure associated with the invasive potential of the cells (West et al., 2004). The formation of podosome was reported to be regulated by cortactin (Kaverina et al., 2003). Therefore, through the interaction with HAX1, which is a cortactin binding protein, CrSARM may play a role in TLR signal transduction leading to the phagocytosis of the pathogen. Proof of this hypothesis will shed light on the molecular mechanism of the phagocytosis of bacterial pathogen by the amebocytes that was reported previously by Zhu *et al* (Zhu et al., 2005). Alternatively, CrSARM/HAX1/cortactin complex may participate in the receptor-mediated endocytosis, which is essential for the downregulation of some signaling cascades. Indeed, cortactin was shown to be involved in the endocytosis of the activated epidermal growth factor receptors, a process required for the downregulation of EGF signaling (Daly, 2004; van Rossum et al., 2005). Other TIR domain-containing adaptor molecule that was found to act as a negative regulator of the TLR signaling pathways is Triad3A, although the downregulation of signaling is achieved by ubiquitination-dependent degradation of the activated TLRs.

In order to determine the biological significance of CrSARM/HAX1 interaction, future work should focus on the identification of signaling cascades up- and down-stream of CrSARM/HAX1 complex.

4.4.3 Implication of other potential interactions

Effort should be taken to evaluate the validity and the implication of the interactions between TIR domain of CrSARM with **proteasome alpha-subunit**, **Hsp40**, which are structural components of proteasome and chaperone machineries, respectively, and the two **hypothetical proteins** (Table 3.6). Though there is no definite indication how such interactions are essential for the cells, interactions between proteasome alpha subunit, a structural component of proteasome, with nonubiquitinated proteins and between heat shock proteins that compose of the chaperone machinery with the signaling proteins were documented (Cho et al., 2001; Dong et al., 2004; Pratt and Toft, 2003). Indeed, it was speculated that the recruitment of proteasome subunit to unubiquitinated proteins would direct the proteins to the ubiquitin-independent proteasomal degradation pathway (Cho et al., 2001; Dong et al., 2004). In the interaction between the signaling molecules with the chaperon machinery, it is possible that the latter might severe as a temporary carrier to shuttle the former to the site of action in the cells. Alternatively, by interaction with the chaperon, signaling molecules could acquire the conformational change that favors its interaction with the ligand or other signaling molecules. An example is the binding of the glucocorticoid receptor to the hsp90/hsp70 chaperone machinery converts steroid-binding cleft on the receptor from the closed to open conformation, allowing the access of the steroid (Pratt and Toft, 2003). It is thus worthwhile to investigate the biological functions of the interactions between CrSARM and the components of proteasome and chaperone machinery.

Chapter 5 Conclusion and Future Perspectives

Differential usage of TIR domain-containing adaptors by TLRs tailors the innate immune responses to match the nature of the immune challenges. As this class of signaling molecule is expanding, the more we understand of the functions of these adaptor proteins, the greater is our appreciation of the sophisticated contribution of TLR signaling in the ancient host defense component of immunity. Among the TIR domain-containing adaptor proteins, SARM is the newest to be described and, thus, also the least to be functionally characterized. Research on this signaling molecule is hence, constrained by the limitation of literature support. However, any new findings will more than offset the challenge of the lack of information, which will offer significant insights into not only TLR signaling but also innate immunity, a field that is still rudimentary compared with adaptive immunity.

We have successfully obtained the full-length cDNA of *C. rotundicauda* SARM and showed that the expression of CrSARM is responsive to *P. aeruginosa* infection, confirming the immune-related function of this protein. The search for interaction partners of CrSARM led to our identification of CaMKI, SUMO-1, HAX-1, proteasome-α subunit, Hsp40 and two hypothetical proteins, thus revealing the potential signaling events that CrSARM may participate in. Based on our results and the scientific literature on the *C. elegans* SARM homolog, we propose that this TIR domain-containing adaptor may promote and regulate signal transduction leading to the activation of p38 MAPK through the interaction with CaMKI, SUMO-1 and ASK1. The interaction between CrSARM with HAX1, on the other hand, indicates the potential role of CrSARM as an adaptor protein coupling the upstream signaling, which is possibly TLR-related, to the downstream signaling that either results in anti-apoptotic activity or actin remodeling. Thus, SARM may be regarded as a

cytoplasmic bridging molecule for the upstream TLR signaling to the downstream p38 MAPK pathways. Although more work is needed to fully decipher and confirm this hypothesis, the interaction between CrSARM and proteasome-α subunit and Hsp40 suggests the possible connections between CrSARM with proteasome degradation and chaperone-mediated protein modification, respectively. Needless to say, the activation of p38 MAPK, the induction of anti-apoptosis, actin remodeling, proteasome degradation, and protein folding by the chaperone machinery have been known to play essential roles in various physiological as well as immune related functions such as the regulation of antimicrobial peptide expression, promotion of host cell survival upon pathogen infection, pathogen clearance by phagocytosis, and down-regulation of signal transduction by endocytosis of activated receptors

Future study should firstly aim to validate of the interaction of CrSARM with its putative partners using different approaches such as immunoprecipitation and/or GST pull-down assays. The functional study of each individual interaction may be performed to uncover the biological relevance of these interactions in the context of global intracellular signaling in response to pathogen invasion. In addition, experiments could be designed to investigate whether CrSARM functions downstream of CrTLR that is currently being isolated by our lab. Last but not least, further work should also address the following questions: (1) what receptors/ligands activate the signaling cascades through CrSARM; (2) how CrSARM co-ordinates with various signaling pathways; and (3) what is the significance of that co-ordination under naïve physiological conditions and during infection.

Bibliography

Aberle, H., Butz, S., Stappert, J., Weissig, H., Kemler, R. and Hoschuetzky, H. (1994) Assembly of the cadherin-catenin complex in vitro with recombinant proteins. *J Cell Sci*, **107** , 3655-3663.

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

Akira, S. (2000) Toll-like receptors: lessons from knockout mice. *Biochem Soc Trans*, **28**, 551-556.

Akira, S. and Takeda, K. (2004) Toll-like receptor signalling. *Nat Rev Immunol*, **4**, 499-511.

Allen, L., Dockrell, D.H., Pattery, T., Lee, D.G., Cornelis, P., Hellewell, P.G. and Whyte, M.K. (2005) Pyocyanin production by *Pseudomonas aeruginosa* induces neutrophil apoptosis and impairs neutrophil-mediated host defenses in vivo. *J Immunol*, **174**, 3643-3649.

Andrade, M.A., Petosa, C., O'Donoghue, S.I., Muller, C.W. and Bork, P. (2001) Comparison of ARM and HEAT protein repeats. *J Mol Biol*, **309**, 1-18.

Ariki, S., Koori, K., Osaki, T., Motoyama, K., Inamori, K. and Kawabata, S. (2004) A serine protease zymogen functions as a pattern-recognition receptor for lipopolysaccharides. *Proc Natl Acad Sci U S A*, **101**, 953-958.

Banchereau, J. and Steinman, R.M. (1998) Dendritic cells and the control of immunity. *Nature*, **392**, 245-252.

Bell, J.K., Mullen, G.E., Leifer, C.A., Mazzoni, A., Davies, D.R. and Segal, D.M. (2003) Leucine-rich repeats and pathogen recognition in Toll-like receptors. *Trends Immunol*, **24**, 528-533.

Belvin, M.P. and Anderson, K.V. (1996) A conserved signaling pathway: the Drosophila toll-dorsal pathway. *Annu Rev Cell Dev Biol*, **12**, 393-416.

Brunet, J.P., Cotte-Laffitte, J., Linxe, C., Quero, A.M., Geniteau-Legendre, M. and Servin, A. (2000) Rotavirus infection induces an increase in intracellular calcium concentration in human intestinal epithelial cells: role in microvillar actin alteration. *J Virol*, **74**, 2323-2332.

Cho, S., Choi, Y.J., Kim, J.M., Jeong, S.T., Kim, J.H., Kim, S.H. and Ryu, S.E. (2001) Binding and regulation of HIF-1alpha by a subunit of the proteasome complex, PSMA7. *FEBS Lett*, **498**, 62-66.

Chuang, C.F. and Bargmann, C.I. (2005) A Toll-interleukin 1 repeat protein at the synapse specifies asymmetric odorant receptor expression via ASK1 MAPKKK signaling. *Genes Dev*, **19**, 270-281.

Coates, J.C. (2003) Armadillo repeat proteins: beyond the animal kingdom. *Trends Cell Biol*, **13**, 463-471.

Conti, E. and Izaurralde, E. (2001) Nucleocytoplasmic transport enters the atomic age. *Curr Opin Cell Biol*, **13**, 310-319.

Conti, E. and Kuriyan, J. (2000) Crystallographic analysis of the specific yet versatile recognition of distinct nuclear localization signals by karyopherin alpha. *Structure*, **8**, 329-338.

Couillault, C., Pujol, N., Reboul, J., Sabatier, L., Guichou, J.F., Kohara, Y. and Ewbank, J.J. (2004) TLR-independent control of innate immunity in *Caenorhabditis elegans* by the TIR domain adaptor protein TIR-1, an ortholog of human SARM. *Nat Immunol*, **5**, 488-494.

Daly, R.J. (2004) Cortactin signalling and dynamic actin networks. *Biochem J*, **382**, 13-25.

Ding, J.L., Navas, M.A., 3rd and Ho, B. (1993) Two forms of factor C from the amoebocytes of *Carcinoscorpius rotundicauda*: purification and characterisation. *Biochim Biophys Acta*, **1202**, 149-156.

Ding, J.L., Tan, K.C., Thangamani, S., Kusuma, N., Seow, W.K., Bui, T.H., Wang, J. and Ho, B. (2005) Spatial and temporal coordination of expression of immune response genes during *Pseudomonas* infection of horseshoe crab, *Carcinoscorpius rotundicauda*. *Genes Immun*.

Dong, J., Chen, W., Welford, A. and Wandinger-Ness, A. (2004) The proteasome alpha-subunit XAPC7 interacts specifically with Rab7 and late endosomes. *J Biol Chem*, **279**, 21334-21342.

DosReis, G.A. and Barcinski, M.A. (2001) Apoptosis and parasitism: from the parasite to the host immune response. *Adv Parasitol*, **49**, 133-161.

Doyle, S.E., O'Connell, R.M., Miranda, G.A., Vaidya, S.A., Chow, E.K., Liu, P.T., Suzuki, S., Suzuki, N., Modlin, R.L., Yeh, W.C., Lane, T.F. and Cheng, G. (2004) Toll-like receptors induce a phagocytic gene program through p38. *J Exp Med*, **199**, 81-90.

Dunne, A., Ejdeback, M., Ludidi, P.L., O'Neill, L.A. and Gay, N.J. (2003) Structural complementarity of Toll/interleukin-1 receptor domains in Toll-like receptors and the adaptors Mal and MyD88. *J Biol Chem*, **278**, 41443-41451.

Epstein, J., Eichbaum, Q., Sheriff, S. and Ezekowitz, R.A. (1996) The collectins in innate immunity. *Curr Opin Immunol*, **8**, 29-35.

Fearon, D.T. and Locksley, R.M. (1996) The instructive role of innate immunity in the acquired immune response. *Science*, **272**, 50-53.

Fitzgerald, K.A., Rowe, D.C., Barnes, B.J., Caffrey, D.R., Visintin, A., Latz, E., Monks, B., Pitha, P.M. and Golenbock, D.T. (2003) LPS-TLR4 signaling to IRF-3/7 and NF-kappaB involves the toll adapters TRAM and TRIF. *J Exp Med*, **198**, 1043-1055.

Fraser, I.P., Koziel, H. and Ezekowitz, R.A. (1998) The serum mannose-binding protein and the macrophage mannose receptor are pattern recognition molecules that link innate and adaptive immunity. *Semin Immunol*, **10**, 363-372.

Fujisawa, H. (2001) Regulation of the activities of multifunctional Ca2+/calmodulin-dependent protein kinases. *J Biochem (Tokyo)*, **129**, 193-199.

Gallagher, A.R., Cedzich, A., Gretz, N., Somlo, S. and Witzgall, R. (2000) The polycystic kidney disease protein PKD2 interacts with Hax-1, a protein associated with the actin cytoskeleton. *Proc Natl Acad Sci U S A*, **97**, 4017-4022.

Haque, S., Dumon, H., Haque, A. and Kasper, L.H. (1998) Alteration of intracellular calcium flux and impairment of nuclear factor-AT translocation in T cells during acute Toxoplasma gondii infection in mice. *J Immunol*, **161**, 6812-6818.

Hardalo, C. and Edberg, S.C. (1997) *Pseudomonas aeruginosa*: assessment of risk from drinking water. *Crit Rev Microbiol*, **23**, 47-75.

Hashimoto, C., Hudson, K.L. and Anderson, K.V. (1988) The Toll gene of *Drosophila*, required for dorsal-ventral embryonic polarity, appears to encode a transmembrane protein. *Cell*, **52**, 269-279.

Hoebe, K., Du, X., Georgel, P., Janssen, E., Tabeta, K., Kim, S.O., Goode, J., Lin, P., Mann, N., Mudd, S., Crozat, K., Sovath, S., Han, J. and Beutler, B. (2003) Identification of Lps2 as a key transducer of MyD88-independent TIR signalling. *Nature*, **424**, 743-748.

Holmskov, U., Thiel, S. and Jensenius, J.C. (2003) Collections and ficolins: humoral lectins of the innate immune defense. *Annu Rev Immunol*, **21**, 547-578.
Horng, T., Barton, G.M., Flavell, R.A. and Medzhitov, R. (2002) The adaptor molecule TIRAP provides signalling specificity for Toll-like receptors. *Nature*, **420**, 329-333.
Huber, A.H., Nelson, W.J. and Weis, W.I. (1997) Three-dimensional structure of the armadillo repeat region of beta-catenin. *Cell*, **90**, 871-882.
Inamori, K., Ariki, S. and Kawabata, S. (2004) A Toll-like receptor in horseshoe crabs. *Immunol Rev*, **198**, 106-115.
Iwanaga, S. (2002) The molecular basis of innate immunity in the horseshoe crab. *Curr Opin Immunol*, **14**, 87-95.
Iwanaga, S. and Kawabata, S. (1998) Evolution and phylogeny of defense molecules associated with innate immunity in horseshoe crab. *Front Biosci*, **3**, D973-984.
Iwanaga, S. and Lee, B.L. (2005) Recent advances in the innate immunity of invertebrate animals. *J Biochem Mol Biol*, **38**, 128-150.
Janeway, C.A., Jr. and Medzhitov, R. (2002) Innate immune recognition. *Annu Rev Immunol*, **20**, 197-216.
Janssens, S. and Beyaert, R. (2002) A universal role for MyD88 in TLR/IL-1R-mediated signaling. *Trends Biochem Sci*, **27**, 474-482.
Jault, C., Pichon, L. and Chluba, J. (2004) Toll-like receptor gene family and TIR-domain adapters in Danio rerio. *Mol Immunol*, **40**, 759-771.
Jenkins, C.E., Swiatoniowski, A., Issekutz, A.C. and Lin, T.J. (2004) Pseudomonas aeruginosa exotoxin A induces human mast cell apoptosis by a caspase-8 and -3-dependent mechanism. *J Biol Chem*, **279**, 37201-37207.
Kaverina, I., Stradal, T.E. and Gimona, M. (2003) Podosome formation in cultured A7r5 vascular smooth muscle cells requires Arp2/3-dependent de-novo actin polymerization at discrete microdomains. *J Cell Sci*, **116**, 4915-4924.
Kawai, T., Takeuchi, O., Fujita, T., Inoue, J., Muhlradt, P.F., Sato, S., Hoshino, K. and Akira, S. (2001) Lipopolysaccharide stimulates the MyD88-independent pathway and results in activation of IFN-regulatory factor 3 and the expression of a subset of lipopolysaccharide-inducible genes. *J Immunol*, **167**, 5887-5894.
Kim, C.A. and Bowie, J.U. (2003) SAM domains: uniform structure, diversity of function. *Trends Biochem Sci*, **28**, 625-628.
Kim, C.A., Phillips, M.L., Kim, W., Gingery, M., Tran, H.H., Robinson, M.A., Faham, S. and Bowie, J.U. (2001) Polymerization of the SAM domain of TEL in leukemogenesis and transcriptional repression. *Embo J*, **20**, 4173-4182.
Koyama, A.H., Irie, H., Fukumori, T., Hata, S., Iida, S., Akari, H. and Adachi, A. (1998) Role of virus-induced apoptosis in a host defense mechanism against virus infection. *J Med Invest*, **45**, 37-45.
Lee, Y.S., Jang, M.S., Lee, J.S., Choi, E.J. and Kim, E. (2005) SUMO-1 represses apoptosis signal-regulating kinase 1 activation through physical interaction and not through covalent modification. *EMBO Rep*, **6**, 949-955.
Lemaitre, B., Nicolas, E., Michaut, L., Reichhart, J.M. and Hoffmann, J.A. (1996) The dorsoventral regulatory gene cassette spatzle/Toll/cactus controls the potent antifungal response in Drosophila adults. *Cell*, **86**, 973-983.
Lemaitre, B., Reichhart, J.M. and Hoffmann, J.A. (1997) *Drosophila* host defense: differential induction of antimicrobial peptide genes after infection by various classes of microorganisms. *Proc Natl Acad Sci U S A*, **94**, 14614-14619.

Levashina, E.A., Langley, E., Green, C., Gubb, D., Ashburner, M., Hoffmann, J.A. and Reichhart, J.M. (1999) Constitutive activation of toll-mediated antifungal defense in serpin-deficient *Drosophila*. *Science*, **285**, 1917-1919.

Liberati, N.T., Fitzgerald, K.A., Kim, D.H., Feinbaum, R., Golenbock, D.T. and Ausubel, F.M. (2004) Requirement for a conserved Toll/interleukin-1 resistance domain protein in the Caenorhabditis elegans immune response. *Proc Natl Acad Sci U S A*, **101**, 6593-6598.

Ligoxygakis, P., Pelte, N., Hoffmann, J.A. and Reichhart, J.M. (2002) Activation of Drosophila Toll during fungal infection by a blood serine protease. *Science*, **297**, 114-116.

Low, H.P., Paulin, J.J. and Keith, C.H. (1992) Trypanosoma cruzi infection of BSC-1 fibroblast cells causes cytoskeletal disruption and changes in intracellular calcium levels. *J Protozool*, **39**, 463-470.

Matsuda, G., Nakajima, K., Kawaguchi, Y., Yamanashi, Y. and Hirai, K. (2003) Epstein-Barr virus (EBV) nuclear antigen leader protein (EBNA-LP) forms complexes with a cellular anti-apoptosis protein Bcl-2 or its EBV counterpart BHRF1 through HS1-associated protein X-1. *Microbiol Immunol*, **47**, 91-99.

McGettrick, A.F. and O'Neill, L.A. (2004) The expanding family of MyD88-like adaptors in Toll-like receptor signal transduction. *Mol Immunol*, **41**, 577-582.

Medzhitov, R. and Janeway, C., Jr. (2000) Innate immunity. *N Engl J Med*, **343**, 338-344.

Meijer, A.H., Gabby Krens, S.F., Medina Rodriguez, I.A., He, S., Bitter, W., Ewa Snaar-Jagalska, B. and Spaink, H.P. (2004) Expression analysis of the Toll-like receptor and TIR domain adaptor families of zebrafish. *Mol Immunol*, **40**, 773-783.

Mink, M., Fogelgren, B., Olszewski, K., Maroy, P. and Csiszar, K. (2001) A novel human gene (SARM) at chromosome 17q11 encodes a protein with a SAM motif and structural similarity to Armadillo/beta-catenin that is conserved in mouse, *Drosophila,* and *Caenorhabditis elegans*. *Genomics*, **74**, 234-244.

Morisato, D. and Anderson, K.V. (1995) Signaling pathways that establish the dorsal-ventral pattern of the *Drosophila* embryo. *Annu Rev Genet*, **29**, 371-399.

Mukhopadhyay, S. and Gordon, S. (2004) The role of scavenger receptors in pathogen recognition and innate immunity. *Immunobiology*, **209**, 39-49.

Ng, M.L. (2005) The dynamic role of lectins in bacterial infection: a model for the formation of LPS-recognition complex by pathogen recognition receptors. *Department of Biological Sciences*. National University of Singapore, Singapore, p. 9.

Ng, P.M., Jin, Z., Tan, S.S., Ho, B. and Ding, J.L. (2004) C-reactive protein: a predominant LPS-binding acute phase protein responsive to Pseudomonas infection. *J Endotoxin Res*, **10**, 163-174.

O'Neill, L.A., Fitzgerald, K.A. and Bowie, A.G. (2003) The Toll-IL-1 receptor adaptor family grows to five members. *Trends Immunol*, **24**, 286-290.

Oshiumi, H., Matsumoto, M., Funami, K., Akazawa, T. and Seya, T. (2003a) TICAM-1, an adaptor molecule that participates in Toll-like receptor 3-mediated interferon-beta induction. *Nat Immunol*, **4**, 161-167.

Oshiumi, H., Sasai, M., Shida, K., Fujita, T., Matsumoto, M. and Seya, T. (2003b) TIR-containing adapter molecule (TICAM)-2, a bridging adapter recruiting to toll-like receptor 4 TICAM-1 that induces interferon-beta. *J Biol Chem*, **278**, 49751-49762.

Peifer, M., Berg, S. and Reynolds, A.B. (1994) A repeating amino acid motif shared by proteins with diverse cellular roles. *Cell*, **76**, 789-791.
Ponting, C.P. (1995) SAM: a novel motif in yeast sterile and Drosophila polyhomeotic proteins. *Protein Sci*, **4**, 1928-1930.
Pratt, W.B. and Toft, D.O. (2003) Regulation of signaling protein function and trafficking by the hsp90/hsp70-based chaperone machinery. *Exp Biol Med (Maywood)*, **228**, 111-133.
Royet, J. (2004) Infectious non-self recognition in invertebrates: lessons from Drosophila and other insect models. *Mol Immunol*, **41**, 1063-1075.
Rutschmann, S., Kilinc, A. and Ferrandon, D. (2002) Cutting edge: the toll pathway is required for resistance to gram-positive bacterial infections in *Drosophila*. *J Immunol*, **168**, 1542-1546.
Saitou, N. and Nei, M. (1987) The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol Biol Evol*, **4**, 406-425.
Sato, S., Sugiyama, M., Yamamoto, M., Watanabe, Y., Kawai, T., Takeda, K. and Akira, S. (2003) Toll/IL-1 receptor domain-containing adaptor inducing IFN-beta (TRIF) associates with TNF receptor-associated factor 6 and TANK-binding kinase 1, and activates two distinct transcription factors, NF-kappa B and IFN-regulatory factor-3, in the Toll-like receptor signaling. *J Immunol*, **171**, 4304-4310.
Schultz, J., Milpetz, F., Bork, P. and Ponting, C.P. (1998) SMART, a simple modular architecture research tool: identification of signaling domains. *Proc Natl Acad Sci U S A*, **95**, 5857-5864.
Slack, J.L., Schooley, K., Bonnert, T.P., Mitcham, J.L., Qwarnstrom, E.E., Sims, J.E. and Dower, S.K. (2000) Identification of two major sites in the type I interleukin-1 receptor cytoplasmic region responsible for coupling to pro-inflammatory signaling pathways. *J Biol Chem*, **275**, 4670-4678.
Stormer, L. (1952) Phylogeny and taxonomy of fossil horseshoe crabs. *J. Paleont*, **26**, 630-639.
Stover, C.K., Pham, X.Q., Erwin, A.L., Mizoguchi, S.D., Warrener, P., Hickey, M.J., Brinkman, F.S., Hufnagle, W.O., Kowalik, D.J., Lagrou, M., Garber, R.L., Goltry, L., Tolentino, E., Westbrock-Wadman, S., Yuan, Y., Brody, L.L., Coulter, S.N., Folger, K.R., Kas, A., Larbig, K., Lim, R., Smith, K., Spencer, D., Wong, G.K., Wu, Z., Paulsen, I.T., Reizer, J., Saier, M.H., Hancock, R.E., Lory, S. and Olson, M.V. (2000) Complete genome sequence of *Pseudomonas aeruginosa* PA01, an opportunistic pathogen. *Nature*, **406**, 959-964.
Suzuki, Y., Demoliere, C., Kitamura, D., Takeshita, H., Deuschle, U. and Watanabe, T. (1997) HAX-1, a novel intracellular protein, localized on mitochondria, directly associates with HS1, a substrate of Src family tyrosine kinases. *J Immunol*, **158**, 2736-2744.
Takeda, K. and Akira, S. (2005) Toll-like receptors in innate immunity. *Int Immunol*, **17**, 1-14.
Tateda, K., Ishii, Y., Horikawa, M., Matsumoto, T., Miyairi, S., Pechere, J.C., Standiford, T.J., Ishiguro, M. and Yamaguchi, K. (2003) The *Pseudomonas aeruginosa* autoinducer N-3-oxododecanoyl homoserine lactone accelerates apoptosis in macrophages and neutrophils. *Infect Immun*, **71**, 5785-5793.
Thomas, C.A., Li, Y., Kodama, T., Suzuki, H., Silverstein, S.C. and El Khoury, J. (2000) Protection from lethal gram-positive infection by macrophage scavenger receptor-dependent phagocytosis. *J Exp Med*, **191**, 147-156.

Tobimatsu, T. and Fujisawa, H. (1989) Tissue-specific expression of four types of rat calmodulin-dependent protein kinase II mRNAs. *J Biol Chem*, **264**, 17907-17912.

Van Criekinge, W. and Beyaert, R. (1999) Yeast Two-Hybrid: State of the Art. *Biol Proced Online*, **2**, 1-38.

van Rossum, A.G., Gibcus, J., van der Wal, J. and Schuuring, E. (2005) Cortactin overexpression results in sustained epidermal growth factor receptor signaling by preventing ligand-induced receptor degradation in human carcinoma cells. *Breast Cancer Res*, **7**, 235-237.

Verploegen, S., Lammers, J.W., Koenderman, L. and Coffer, P.J. (2000) Identification and characterization of CKLiK, a novel granulocyte Ca(++)/calmodulin-dependent kinase. *Blood*, **96**, 3215-3223.

Verploegen, S., Ulfman, L., van Deutekom, H.W., van Aalst, C., Honing, H., Lammers, J.W., Koenderman, L. and Coffer, P.J. (2005) Characterization of the role of CaMKI-like kinase (CKLiK) in human granulocyte function. *Blood*, **106**, 1076-1083.

Werner, T., Liu, G., Kang, D., Ekengren, S., Steiner, H. and Hultmark, D. (2000) A family of peptidoglycan recognition proteins in the fruit fly *Drosophila melanogaster*. *Proc Natl Acad Sci U S A*, **97**, 13772-13777.

West, M.A., Wallin, R.P., Matthews, S.P., Svensson, H.G., Zaru, R., Ljunggren, H.G., Prescott, A.R. and Watts, C. (2004) Enhanced dendritic cell antigen capture via toll-like receptor-induced actin remodeling. *Science*, **305**, 1153-1157.

Wu, H. and Montone, K.T. (1998) Cortactin localization in actin-containing adult and fetal tissues. *J Histochem Cytochem*, **46**, 1189-1191.

Yamamoto, M., Sato, S., Hemmi, H., Hoshino, K., Kaisho, T., Sanjo, H., Takeuchi, O., Sugiyama, M., Okabe, M., Takeda, K. and Akira, S. (2003a) Role of adaptor TRIF in the MyD88-independent toll-like receptor signaling pathway. *Science*, **301**, 640-643.

Yamamoto, M., Sato, S., Hemmi, H., Sanjo, H., Uematsu, S., Kaisho, T., Hoshino, K., Takeuchi, O., Kobayashi, M., Fujita, T., Takeda, K. and Akira, S. (2002a) Essential role for TIRAP in activation of the signalling cascade shared by TLR2 and TLR4. *Nature*, **420**, 324-329.

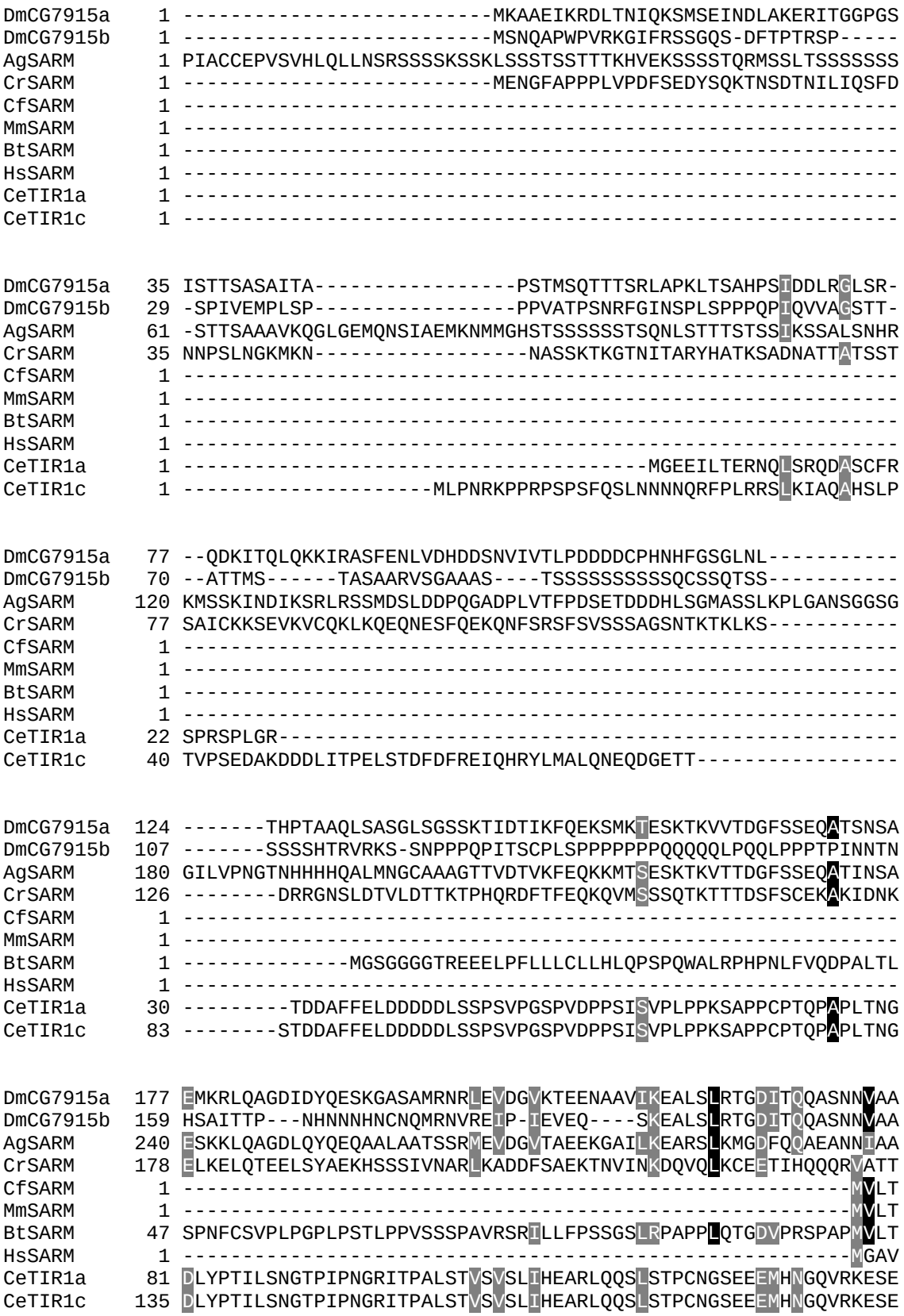
Yamamoto, M., Sato, S., Hemmi, H., Uematsu, S., Hoshino, K., Kaisho, T., Takeuchi, O., Takeda, K. and Akira, S. (2003b) TRAM is specifically involved in the Toll-like receptor 4-mediated MyD88-independent signaling pathway. *Nat Immunol*, **4**, 1144-1150.

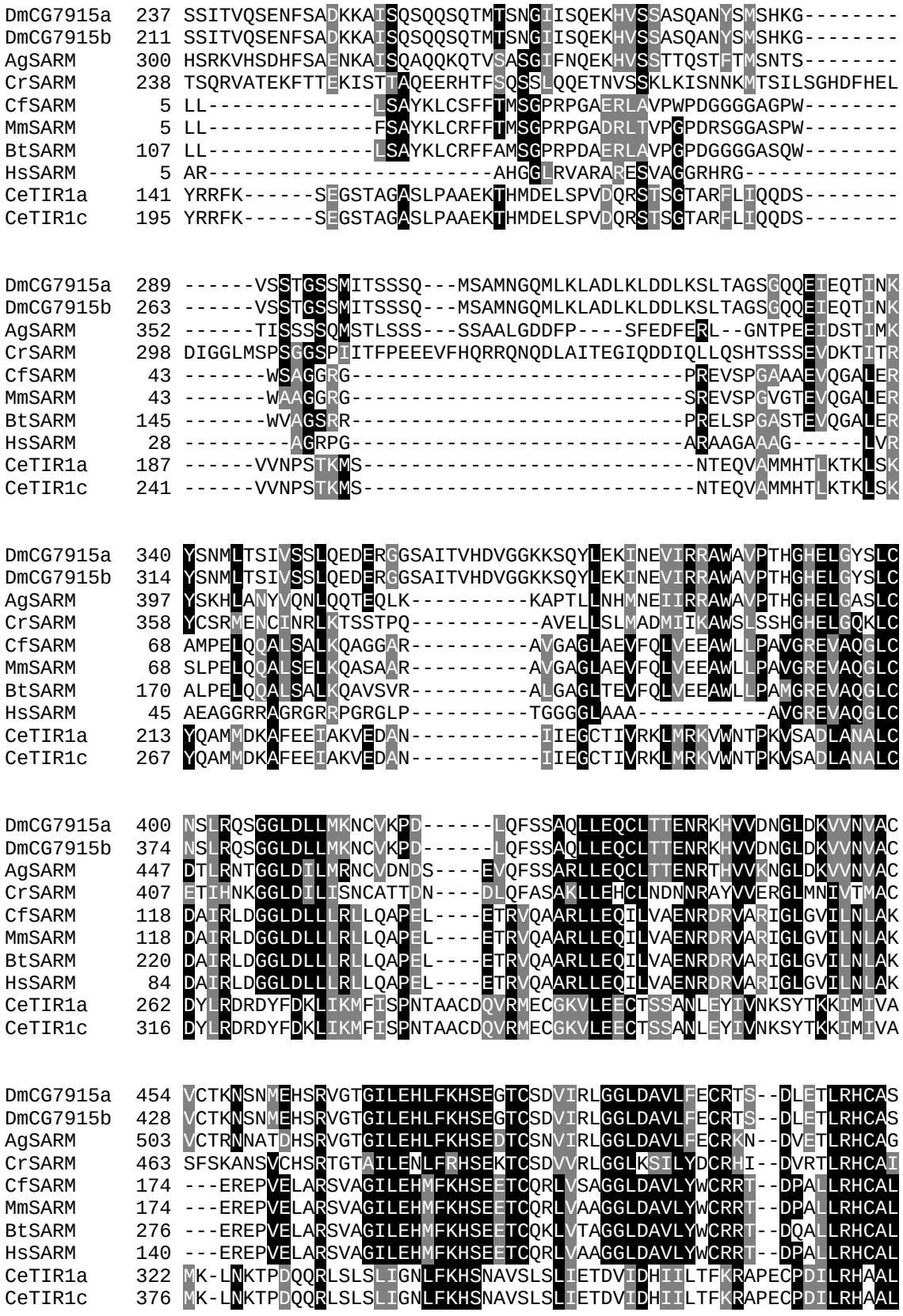
Yamamoto, M., Sato, S., Mori, K., Hoshino, K., Takeuchi, O., Takeda, K. and Akira, S. (2002b) Cutting edge: a novel Toll/IL-1 receptor domain-containing adapter that preferentially activates the IFN-beta promoter in the Toll-like receptor signaling. *J Immunol*, **169**, 6668-6672.

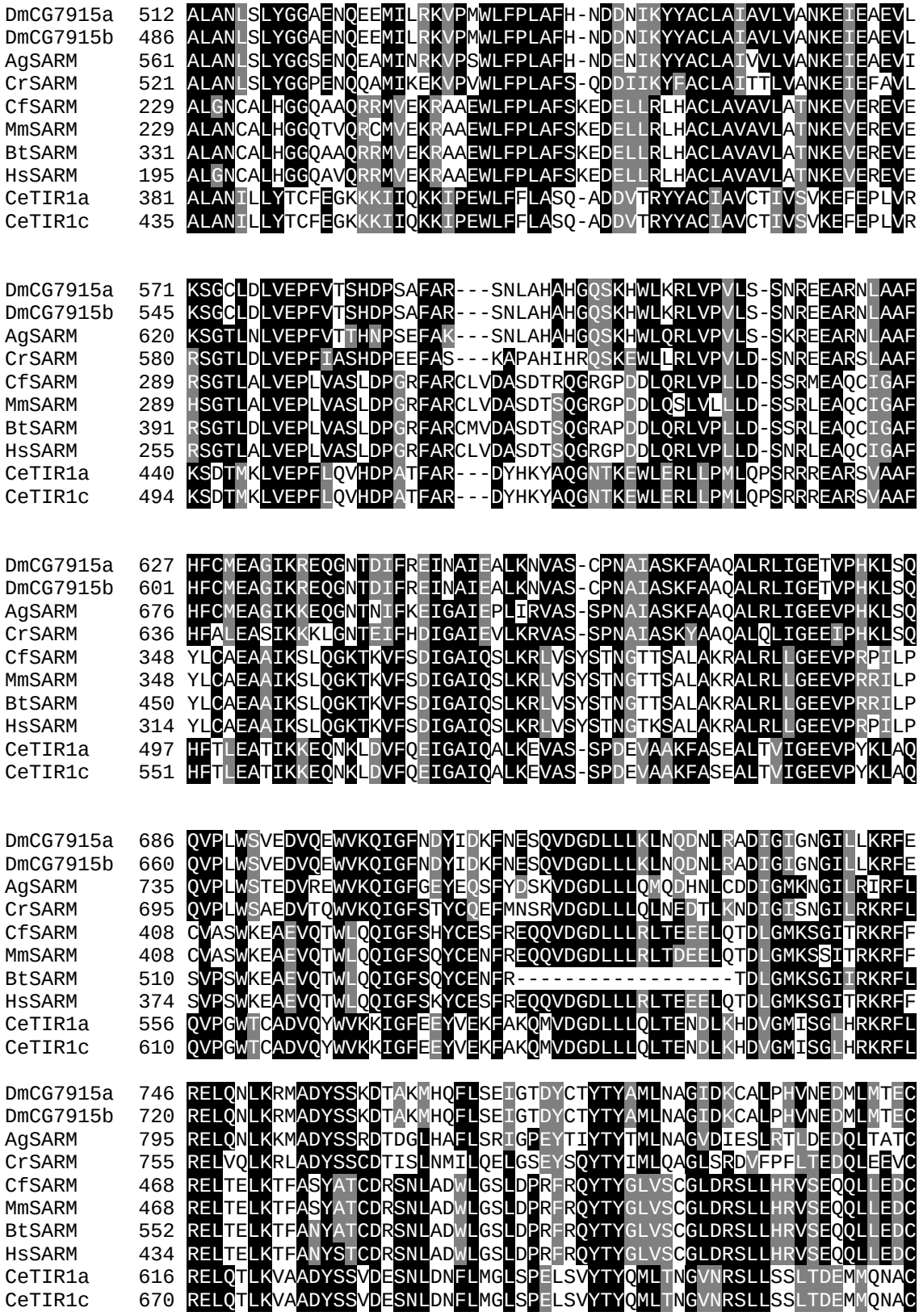
Yau, Y.H., Ho, B., Tan, N.S., Ng, M.L. and Ding, J.L. (2001) High therapeutic index of factor C Sushi peptides: potent antimicrobials against *Pseudomonas aeruginosa*. *Antimicrob Agents Chemother*, **45**, 2820-2825.

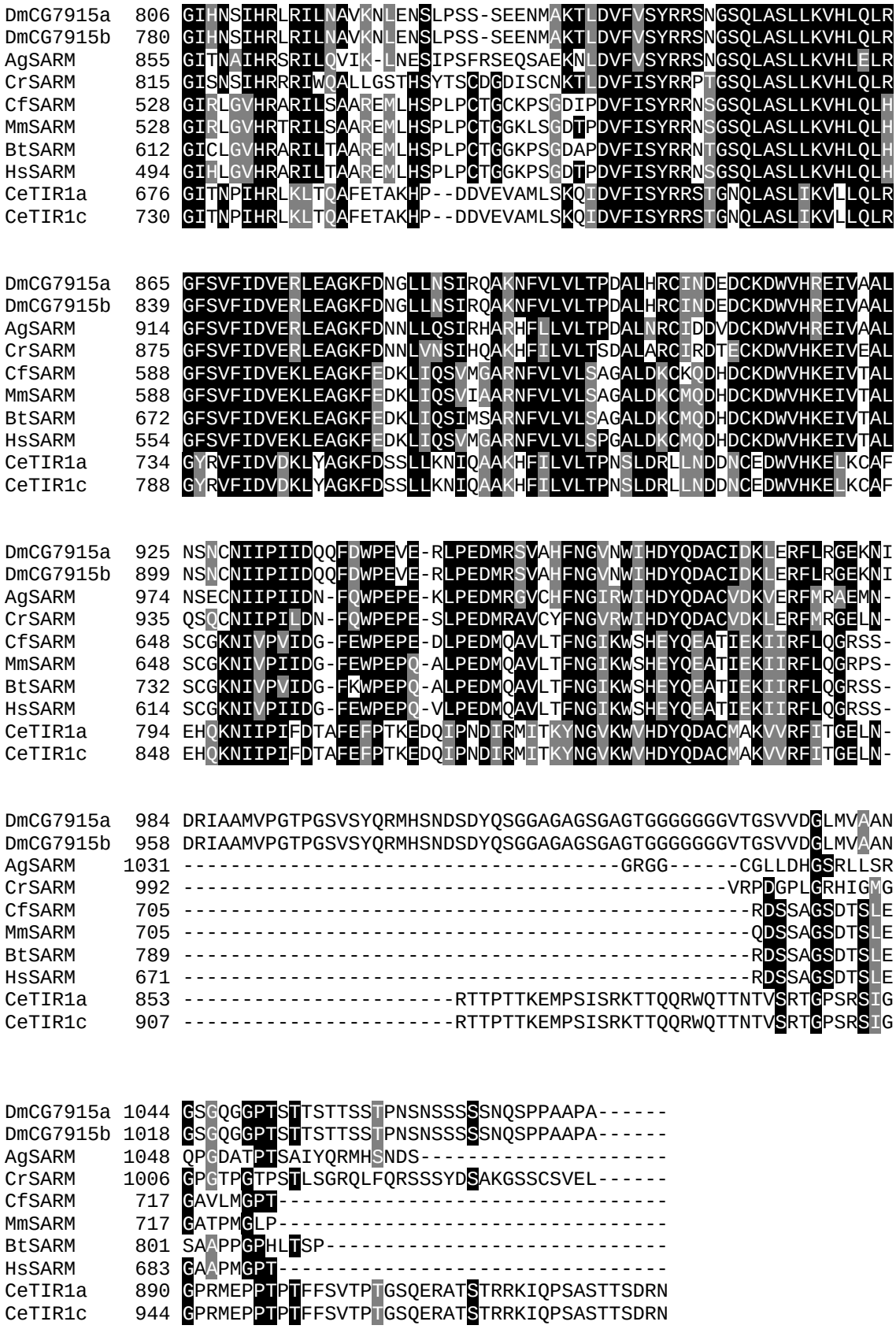
Zhu, Y., Thangamani, S., Ho, B. and Ding, J.L. (2005) The ancient origin of the complement system. *Embo J*, **24**, 382-394.

Appendix A


DmCG7915a 1 -------------------------MKAAEIKRDLTNIQKSMSEINDLAKERITGGPGS
DmCG7915b 1 -------------------------MSNQAPWPVRKGIFRSSGQS-DFTPTRSP-----
AgSARM 1 PIACCEPVSVHLQLLNSRSSSSKSSKLSSSTSSTTTKHVEKSSSSTQRMSSLTSSSSSSS
CrSARM 1 -------------------------MENGFAPPPLVPDFSEDYSQKTNSDTNILIQSFD
CfSARM 1 --
MmSARM 1 --
BtSARM 1 --
HsSARM 1 --
CeTIR1a 1 --
CeTIR1c 1 --
DmCG7915a 35 ISTTSASAITA-----------------PSTMSQTTTSRLAPKLTSAHPSIDDLRGLSR-
DmCG7915b 29 -SPIVEMPLSP-----------------PPVATPSNRFGINSPLSPPPQPIQVVAGSTT-
AgSARM 61 -STTSAAAVKQGLGEMQNSIAEMKNMMGHSTSSSSSSTSQNLSTTTSTSSIKSSALSNHR
CrSARM 35 NNPSLNGKMKN-----------------NASSKTKGTNITARYHATKSADNATTATSST
CfSARM 1 --
MmSARM 1 --
BtSARM 1 --
HsSARM 1 --
CeTIR1a 1 ---------------------------------------MGEEILTERNQLSRQDASCFR
CeTIR1c 1 --------------------MLPNRKPPRPSPSFQSLNNNNQRFPLRRSLKIAQAHSLP
DmCG7915a 77 --QDKITQLQKKIRASFENLVDHDDSNVIVTLPDDDDCPHNHFGSGLNL-----------
DmCG7915b 70 --ATTMS------TASAARVSGAAAS----TSSSSSSSSSQCSSQTSS-----------
AgSARM 120 KMSSKINDIKSRLRSSMDSLDDPQGADPLVTFPDSETDDDHLSGMASSLKPLGANSGGSG
CrSARM 77 SAICKKSEVKVCQKLKQEQNESFQEKQNFSRSFSVSSSAGSNTKTKLKS-----------
CfSARM 1 --
MmSARM 1 --
BtSARM 1 --
HsSARM 1 --
CeTIR1a 22 SPRSPLGR--
CeTIR1c 40 TVPSEDAKDDDLITPELSTDFDFREIQHRYLMALQNEQDGETT-----------------
DmCG7915a 124 -------THPTAAQLSASGLSGSSKTIDTIKFQEKSMKTESKTKVVTDGFSSEQATSNSA
DmCG7915b 107 -------SSSSHTRVRKS-SNPPPQPITSCPLSPPPPPPQQQQQLPQQLPPPTPINNTN
AgSARM 180 GILVPNGTNHHHHQALMNGCAAAGTTVDTVKFEQKKMTSESKTKVTTDGFSSEQATINSA
CrSARM 126 --------DRRGNSLDTVLDTTKTPHQRDFTFEQKQVMSSSQTKTTTDSFSCEKAKIDNK
CfSARM 1 --
MmSARM 1 --
BtSARM 1 -------------MGSGGGGTREEELPFLLLCLLHLQPSPQWALRPHPNLFVQDPALTL
HsSARM 1 --
CeTIR1a 30 ---------TDDAFFELDDDDDLSSPSVPGSPVDPPSISVPLPPKSAPPCPTQPAPLTNG
CeTIR1c 83 --------STDDAFFELDDDDDLSSPSVPGSPVDPPSISVPLPPKSAPPCPTQPAPLTNG
DmCG7915a 177 EMKRLQAGDIDYQESKGASAMRNRLEVDGVKTEENAAVIKEALSLRTGDITQQASNNVAA
DmCG7915b 159 HSAITTP---NHNNNHNCNQMRNVREIP-IEVEQ----SKEALSLRTGDITQQASNNVAA
AgSARM 240 ESKKLQAGDLQYQEQAALAATSSRMEVDGVTAEEKGAILKEARSLKMGDFQQAEANNIAA
CrSARM 178 ELKELQTEELSYAEKHSSSIVNARLKADDFSAEKTNVINKDQVQLKCEETIHQQQRVATT
CfSARM 1 --MVLT
MmSARM 1 --MVLT
BtSARM 47 SPNFCSVPLPGPLPSTLPPVSSSPAVRSRILLFPSSGSLRPAPPLQTGDVPRSPAPMVLT
HsSARM 1 --MGAV
CeTIR1a 81 DLYPTILSNGTPIPNGRITPALSTVSVSLIHEARLQQSLSTPCNGSEEEMHNGQVRKESE
CeTIR1c 135 DLYPTILSNGTPIPNGRITPALSTVSVSLIHEARLQQSLSTPCNGSEEEMHNGQVRKESE


DmCG7915a 237 SSITVQSENFSADKKAISQSQQSQTMTSNGIISQEKHVSSASQANYSMSHKG--------
DmCG7915b 211 SSITVQSENFSADKKAISQSQQSQTMTSNGIISQEKHVSSASQANYSMSHKG--------
AgSARM 300 HSRKVHSDHFSAENKAISQAQQKQTVSASGIFNQEKHVSSTTQSTFTMSNTS--------
CrSARM 238 TSQRVATEKFTTEKISTTAQEERHTFSQSSLQQETNVSSKLKISNNKMTSILSGHDFHEL
CfSARM 5 LL--------------LSAYKLCSFFTMSGPRPGAERLAVPWPDGGGGAGPW--------
MmSARM 5 LL--------------FSAYKLCRFFTMSGPRPGADRLTVPGPDRSGGASPW--------
BtSARM 107 LL--------------LSAYKLCRFFAMSGPRPDAERLAVPGPDGGGGASQW--------
HsSARM 5 AR-----------------------AHGGLRVARARESVAGGRHRG------------
CeTIR1a 141 YRRFK------SEGSTAGASLPAAEKTHMDELSPVDQRSTSGTARFLIQQDS--------
CeTIR1c 195 YRRFK------SEGSTAGASLPAAEKTHMDELSPVDQRSTSGTARFLIQQDS--------
DmCG7915a 289 ------VSSTGSSMITSSSQ---MSAMNGQMLKLADLKLDDLKSLTAGSGQQEIEQTINK
DmCG7915b 263 ------VSSTGSSMITSSSQ---MSAMNGQMLKLADLKLDDLKSLTAGSGQQEIEQTINK
AgSARM 352 ------TISSSSQMSTLSSS---SSAALGDDFP----SFEDFERL--GNTPEEIDSTIMK
CrSARM 298 DIGGLMSPSGGSPIITFPEEEVFHQRRQNQDLAITEGIQDDIQLLQSHTSSSEVDKTITR
CfSARM 43 -------WSAGGRG---------------------------PREVSPGAAAEVQGALER
MmSARM 43 -------WAAGGRG---------------------------SREVSPGVGTEVQGALER
BtSARM 145 -------WVAGSRR---------------------------PRELSPGASTEVQGALER
HsSARM 28 ---------AGRPG---------------------------ARAAGAAAG------LVR
CeTIR1a 187 ------VVNPSTKMS---------------------------NTEQVAMMHTLKTKLSK
CeTIR1c 241 ------VVNPSTKMS---------------------------NTEQVAMMHTLKTKLSK
DmCG7915a 340 YSNMLTSIVSSLQEDERGGSAITVHDVGGKKSQYLEKINEVIRRAWAVPTHGHELGYSLC
DmCG7915b 314 YSNMLTSIVSSLQEDERGGSAITVHDVGGKKSQYLEKINEVIRRAWAVPTHGHELGYSLC
AgSARM 397 YSKHLANYVQNLQQTEQLK----------KAPTLLNHMNEIIRRAWAVPTHGHELGASLC
CrSARM 358 YCSRMENCINRLKTSSTPQ----------AVELLSLMADMIIKAWSLSSHGHELGQKLC
CfSARM 68 AMPELQQALSALKQAGGAR----------AVGAGLAEVFQLVEEAWLLPAVGREVAQGLC
MmSARM 68 SLPELQQALSELKQASAAR----------AVGAGLAEVFQLVEEAWLLPAVGREVAQGLC
BtSARM 170 ALPELQQALSALKQAVSVR----------ALGAGLTEVFQLVEEAWLLPAMGREVAQGLC
HsSARM 45 AEAGGRRAGRGRRPGRGLP----------TGGGGLAAA----------AVGREVAQGLC
CeTIR1a 213 YQAMMDKAFEEIAKVEDAN----------IIEGCTIVRKLMRKVWNTPKVSADLANALC
CeTIR1c 267 YQAMMDKAFEEIAKVEDAN----------IIEGCTIVRKLMRKVWNTPKVSADLANALC
DmCG7915a 400 NSLRQSGGLDLLMKNCVKPD------LQFSSAQLLEQCLTTENRKHVVDNGLDKVVNVAC
DmCG7915b 374 NSLRQSGGLDLLMKNCVKPD------LQFSSAQLLEQCLTTENRKHVVDNGLDKVVNVAC
AgSARM 447 DTLRNTGGLDILMRNCVDNDS----EVQFSSARLLEQCLTTENRTHVVKNGLDKVVNVAC
CrSARM 407 ETIHNKGGLDILISNCATTDN----DLQFASAKLLEHCLNDNNRAYVVERGLMNIVTMAC
CfSARM 118 DAIRLDGGLDLLLRLLQAPEL----ETRVQAARLLEQILVAENRDRVARIGLGVILNLAK
MmSARM 118 DAIRLDGGLDLLLRLLQAPEL----ETRVQAARLLEQILVAENRDRVARIGLGVILNLAK
BtSARM 220 DAIRLDGGLDLLLRLLQAPEL----ETRVQAARLLEQILVAENRDRVARIGLGVILNLAK
HsSARM 84 DAIRLDGGLDLLLRLLQAPEL----ETRVQAARLLEQILVAENRDRVARIGLGVILNLAK
CeTIR1a 262 DYLRDRDYFDKLIKMFISPNTAACDQVRMECGKVLEECTSSANLEYIVNKSYTKKIMIVA
CeTIR1c 316 DYLRDRDYFDKLIKMFISPNTAACDQVRMECGKVLEECTSSANLEYIVNKSYTKKIMIVA
DmCG7915a 454 VCTKNSNMEHSRVGTGILEHLFKHSEGTCSDVIRLGGLDAVLFECRTS--DLETLRHCAS
DmCG7915b 428 VCTKNSNMEHSRVGTGILEHLFKHSEGTCSDVIRLGGLDAVLFECRTS--DLETLRHCAS
AgSARM 503 VCTRNNATDHSRVGTGILEHLFKHSEDTCSNVIRLGGLDAVLFECRKN--DVETLRHCAG
CrSARM 463 SFSKANSVCHSRTGTAILENLFRHSEKTCSDVVRLGGLKSILYDCRHI--DVRTLRHCAI
CfSARM 174 ---EREPVELARSVAGILEHMFKHSEETCQRLVSAGGLDAVLYWCRRT--DPALLRHCAL
MmSARM 174 ---EREPVELARSVAGILEHMFKHSEETCQRLVAAGGLDAVLYWCRRT--DPALLRHCAL
BtSARM 276 ---EREPVELARSVAGILEHMFKHSEETCQKLVTAGGLDAVLYWCRRT--DQALLRHCAL
HsSARM 140 ---EREPVELARSVAGILEHMFKHSEETCQRLVAAGGLDAVLYWCRRT--DPALLRHCAL
CeTIR1a 322 MK-LNKTPDQQRLSLSLIGNLFKHSNAVSLSLIETDVIDHIILTFKRAPECPDILRHAAL
CeTIR1c 376 MK-LNKTPDQQRLSLSLIGNLFKHSNAVSLSLIETDVIDHIILTFKRAPECPDILRHAAL


DmCG7915a 512 ALANLSLYGGAENQEEMILRKVPMWLFPLAFH-NDDNIKYYACLAIAVLVANKEIEAEVL
DmCG7915b 486 ALANLSLYGGAENQEEMILRKVPMWLFPLAFH-NDDNIKYYACLAIAVLVANKEIEAEVL
AgSARM 561 ALANLSLYGGSENQEAMINRKVPSWLFPLAFH-NDENIKYYACLAIVVLVANKEIEAEVI
CrSARM 521 ALANLSLYGGPENQQAMIKEKVPVWLFPLAFS-QDDIIKYFACLAITTLVANKEIEFAVL
CfSARM 229 ALGNCALHGGQAAQRRMVEKRAAEWLFPLAFSKEDELLRLHACLAVAVLATNKEVEREVE
MmSARM 229 ALANCALHGGQTVQRCMVEKRAAEWLFPLAFSKEDELLRLHACLAVAVLATNKEVEREVE
BtSARM 331 ALANCALHGGQAAQRRMVEKRAAEWLFPLAFSKEDELLRLHACLAVAVLATNKEVEREVE
HsSARM 195 ALGNCALHGGQAVQRRMVEKRAAEWLFPLAFSKEDELLRLHACLAVAVLATNKEVEREVE
CeTIR1a 381 ALANILLYTCFEGKKKIIQKKIPEWLFFLASQ-ADDVTRYYACIAVCTIVSVKEFEPLVR
CeTIR1c 435 ALANILLYTCFEGKKKIIQKKIPEWLFFLASQ-ADDVTRYYACIAVCTIVSVKEFEPLVR
DmCG7915a 571 KSGCLDLVEPFVTSHDPSAFAR---SNLAHAHGQSKHWLKRLVPVLS-SNREEARNLAAF
DmCG7915b 545 KSGCLDLVEPFVTSHDPSAFAR---SNLAHAHGQSKHWLKRLVPVLS-SNREEARNLAAF
AgSARM 620 KSGTLNLVEPFVTTHNPSEFAK---SNLAHAHGQSKHWLQRLVPVLS-SKREEARNLAAF
CrSARM 580 RSGTLDLVEPFIASHDPEEFAS---KAPAHIHRQSKEWLLRLVPVLD-SNREEARSLAAF
CfSARM 289 RSGTLALVEPLVASLDPGRFARCLVDASDTRQGRGPDDLQRLVPLLD-SSRMEAQCIGAF
MmSARM 289 HSGTLALVEPLVASLDPGRFARCLVDASDTSQGRGPDDLQSLVLLLD-SSRLEAQCIGAF
BtSARM 391 RSGTLDLVEPLVASLDPGRFARCMVDASDTSQGRAPDDLQRLVPLLD-SSRLEAQCIGAF
HsSARM 255 RSGTLALVEPLVASLDPGRFARCLVDASDTSQGRGPDDLQRLVPLLD-SNRLEAQCIGAF
CeTIR1a 440 KSDTMKLVEPFLQVHDPATFAR---DYHKYAQGNTKEWLERLLPMLQPSRRREARSVAAF
CeTIR1c 494 KSDTMKLVEPFLQVHDPATFAR---DYHKYAQGNTKEWLERLLPMLQPSRRREARSVAAF
DmCG7915a 627 HFCMEAGIKREQGNTDIFREINAIEALKNVAS-CPNAIASKFAAQALRLIGETVPHKLSQ
DmCG7915b 601 HFCMEAGIKREQGNTDIFREINAIEALKNVAS-CPNAIASKFAAQALRLIGETVPHKLSQ
AgSARM 676 HFCMEAGIKKEQGNTNIFKEIGAIEPLIRVAS-SPNAIASKFAAQALRLIGEEVPHKLSQ
CrSARM 636 HFALEASIKKKLGNTEIFHDIGAIEVLKRVAS-SPNAIASKYAAQALQLIGEEIPHKLSQ
CfSARM 348 YLCAEAAIKSLQGKTKVFSDIGAIQSLKRLVSYSTNGTTSALAKRALRLLGEEVPRPILP
MmSARM 348 YLCAEAAIKSLQGKTKVFSDIGAIQSLKRLVSYSTNGTTSALAKRALRLLGEEVPRRILP
BtSARM 450 YLCAEAAIKSLQGKTKVFSDIGAIQSLKRLVSYSTNGTTSALAKRALRLLGEEVPRRILP
HsSARM 314 YLCAEAAIKSLQGKTKVFSDIGAIQSLKRLVSYSTNGTKSALAKRALRLLGEEVPRPILP
CeTIR1a 497 HFTLEATIKKEQNKLDVFQEIGAIQALKEVAS-SPDEVAAKFASEALTVIGEEVPYKLAQ
CeTIR1c 551 HFTLEATIKKEQNKLDVFQEIGAIQALKEVAS-SPDEVAAKFASEALTVIGEEVPYKLAQ
DmCG7915a 686 QVPLWSVEDVQEWVKQIGFNDYIDKFNESQVDGDLLLKLNQDNLRADIGIGNGILLKRFE
DmCG7915b 660 QVPLWSVEDVQEWVKQIGFNDYIDKFNESQVDGDLLLKLNQDNLRADIGIGNGILLKRFE
AgSARM 735 QVPLWSTEDVREWVKQIGFGEYEQSFYDSKVDGDLLLQMQDHNLCDDIGMKNGILRIRFL
CrSARM 695 QVPLWSAEDVTQWVKQIGFSTYCQEFMNSRVDGDLLLQLNEDTLKNDIGISNGILRKRFL
CfSARM 408 CVASWKEAEVQTWLQQIGFSHYCESFREQQVDGDLLLRLTEEELQTDLGMKSGITRKRFF
MmSARM 408 CVASWKEAEVQTWLQQIGFSQYCENFREQQVDGDLLLRLTDEELQTDLGMKSSITRKRFF
BtSARM 510 SVPSWKEAEVQTWLQQIGFSQYCENFR----------TDLGMKSGIIRKRFL
HsSARM 374 SVPSWKEAEVQTWLQQIGFSKYCESFREQQVDGDLLLRLTEEELQTDLGMKSGITRKRFF
CeTIR1a 556 QVPGWTCADVQYWVKKIGFEEYVEKFAKQMVDGDLLLQLTENDLKHDVGMISGLHRKRFL
CeTIR1c 610 QVPGWTCADVQYWVKKIGFEEYVEKFAKQMVDGDLLLQLTENDLKHDVGMISGLHRKRFL
DmCG7915a 746 RELQNLKRMADYSSKDTAKMHQFLSEIGTDYCTYTYAMLNAGIDKCALPHVNEDMLMTEC
DmCG7915b 720 RELQNLKRMADYSSKDTAKMHQFLSEIGTDYCTYTYAMLNAGIDKCALPHVNEDMLMTEC
AgSARM 795 RELQNLKKMADYSSRDTDGLHAFLSRIGPEYTIYTYTMLNAGVDIESLRTLDEDQLTATC
CrSARM 755 RELVQLKRLADYSSCDTISLNMILQELGSEYSQYTYIMLQAGLSRDVFPFLTEDQLEEVC
CfSARM 468 RELTELKTFASYATCDRSNLADWLGSLDPRFRQYTYGLVSCGLDRSLLHRVSEQQLLEDC
MmSARM 468 RELTELKTFASYATCDRSNLADWLGSLDPRFRQYTYGLVSCGLDRSLLHRVSEQQLLEDC
BtSARM 552 RELTELKTFANYATCDRSNLADWLGSLDPRFRQYTYGLVSCGLDRSLLHRVSEQQLLEDC
HsSARM 434 RELTELKTFANYSTCDRSNLADWLGSLDPRFRQYTYGLVSCGLDRSLLHRVSEQQLLEDC
CeTIR1a 616 RELQTLKVAADYSSVDESNLDNFLMGLSPELSVYTYQMLTNGVNRSLLSSLTDEMMQNAC
CeTIR1c 670 RELQTLKVAADYSSVDESNLDNFLMGLSPELSVYTYQMLTNGVNRSLLSSLTDEMMQNAC


DmCG7915a 806 GIHNSIHRLRILNAVKNLENSLPSS-SEENMAKTLDVFVSYRRSNGSQLASLLKVHLQLR
DmCG7915b 780 GIHNSIHRLRILNAVKNLENSLPSS-SEENMAKTLDVFVSYRRSNGSQLASLLKVHLQLR
AgSARM 855 GITNAIHRSRILQVIK-LNESIPSFRSEQSAEKNLDVFVSYRRSNGSQLASLLKVHLELR
CrSARM 815 GISNSIHRRRIWQALLGSTHSYTSCDGDISCNKTLDVFISYRRPTGSQLASLLKVHLQLR
CfSARM 528 GIRLGVHRARILSAAREMLHSPLPCTGCKPSGDIPDVFISYRRNSGSQLASLLKVHLQLH
MmSARM 528 GIRLGVHRTRILSAAREMLHSPLPCTGGKLSGDTPDVFISYRRNSGSQLASLLKVHLQLH
BtSARM 612 GICLGVHRARILTAAREMLHSPLPCTGGKPSGDAPDVFISYRRNTGSQLASLLKVHLQLH
HsSARM 494 GIHLGVHRARILTAAREMLHSPLPCTGGKPSGDTPDVFISYRRNSGSQLASLLKVHLQLH
CeTIR1a 676 GITNPIHRLKLTQAFETAKHP--DDVEVAMLSKQIDVFISYRRSTGNQLASLIKVLLQLR
CeTIR1c 730 GITNPIHRLKLTQAFETAKHP--DDVEVAMLSKQIDVFISYRRSTGNQLASLIKVLLQLR
DmCG7915a 865 GFSVFIDVERLEAGKFDNGLLNSIRQAKNFVLVLTPDALHRCINDEDCKDWVHREIVAAL
DmCG7915b 839 GFSVFIDVERLEAGKFDNGLLNSIRQAKNFVLVLTPDALHRCINDEDCKDWVHREIVAAL
AgSARM 914 GFSVFIDVERLEAGKFDNNLLQSIRHARHFLLVLTPDALNRCIDDVDCKDWVHREIVAAL
CrSARM 875 GFSVFIDVERLEAGKFDNNLVNSIHQAKHFILVLTSDALARCIRDTECKDWVHKEIVEAL
CfSARM 588 GFSVFIDVEKLEAGKFEDKLIQSVMGARNFVLVLSAGALDKCKQDHDCKDWVHKEIVTAL
MmSARM 588 GFSVFIDVEKLEAGKFEDKLIQSVIAARNFVLVLSAGALDKCMQDHDCKDWVHKEIVTAL
BtSARM 672 GFSVFIDVEKLEAGKFEDKLIQSIMSARNFVLVLSAGALDKCMQDHDCKDWVHKEIVTAL
HsSARM 554 GFSVFIDVEKLEAGKFEDKLIQSVMGARNFVLVLSPGALDKCMQDHDCKDWVHKEIVTAL
CeTIR1a 734 GYRVFIDVDKLYAGKFDSSLLKNIQAAKHFILVLTPNSLDRLLNDDNCEDWVHKELKCAF
CeTIR1c 788 GYRVFIDVDKLYAGKFDSSLLKNIQAAKHFILVLTPNSLDRLLNDDNCEDWVHKELKCAF
DmCG7915a 925 NSNCNIIPIIDQQFDWPEVE-RLPEDMRSVAHFNGVNWIHDYQDACIDKLERFLRGEKNI
DmCG7915b 899 NSNCNIIPIIDQQFDWPEVE-RLPEDMRSVAHFNGVNWIHDYQDACIDKLERFLRGEKNI
AgSARM 974 NSECNIIPIIDN-FQWPEPE-KLPEDMRGVCHFNGIRWIHDYQDACVDKVERFMRAEMN-
CrSARM 935 QSQCNIIPILDN-FQWPEPE-SLPEDMRAVCYFNGVRWIHDYQDACVDKLERFMRGELN-
CfSARM 648 SCGKNIVPVIDG-FEWPEPE-DLPEDMQAVLTFNGIKWSHEYQEATIEKIIRFLQGRSS-
MmSARM 648 SCGKNIVPIIDG-FEWPEPQ-ALPEDMQAVLTFNGIKWSHEYQEATIEKIIRFLQGRPS-
BtSARM 732 SCGKNIVPVIDG-FKWPEPQ-ALPEDMQAVLTFNGIKWSHEYQEATIEKIIRFLQGRSS-
HsSARM 614 SCGKNIVPIIDG-FEWPEPQ-VLPEDMQAVLTFNGIKWSHEYQEATIEKIIRFLQGRSS-
CeTIR1a 794 EHQKNIIPIFDTAFEFPTKEDQIPNDIRMITKYNGVKWVHDYQDACMAKVVRFITGELN-
CeTIR1c 848 EHQKNIIPIFDTAFEFPTKEDQIPNDIRMITKYNGVKWVHDYQDACMAKVVRFITGELN-
DmCG7915a 984 DRIAAMVPGTPGSVSYQRMHSNDSDYQSGGAGAGSGAGTGGGGGGGVTGSVVDGLMVAAN
DmCG7915b 958 DRIAAMVPGTPGSVSYQRMHSNDSDYQSGGAGAGSGAGTGGGGGGGVTGSVVDGLMVAAN
AgSARM 1031 ----------GRGG------CGLLDHGSRLLSR
CrSARM 992 ----------VRPDGPLGRHIGMG
CfSARM 705 ----------RDSSAGSDTSLE
MmSARM 705 ----------QDSSAGSDTSLE
BtSARM 789 ----------RDSSAGSDTSLE
HsSARM 671 ----------RDSSAGSDTSLE
CeTIR1a 853 ----------RTTPTTKEMPSISRKTTQQRWQTTNTVSRTGPSRSIG
CeTIR1c 907 ----------RTTPTTKEMPSISRKTTQQRWQTTNTVSRTGPSRSIG
DmCG7915a 1044 GSGQGGPTSTTSTTSSTPNSNSSSSSNQSPPAAPA------
DmCG7915b 1018 GSGQGGPTSTTSTTSSTPNSNSSSSSNQSPPAAPA------
AgSARM 1048 QPGDATPTSAIYQRMHSNDS----------
CrSARM 1006 GPGTPGTPSTLSGRQLFQRSSSYDSAKGSSCSVEL------
CfSARM 717 GAVLMGPT----------
MmSARM 717 GATPMGLP----------
BtSARM 801 SAAPPGPHLTSP----------
HsSARM 683 GAAPMGPT----------
CeTIR1a 890 GPRMEPPTPTFFSVTPTGSQERATSTRRKIQPSASTTSDRN
CeTIR1c 944 GPRMEPPTPTFFSVTPTGSQERATSTRRKIQPSASTTSDRN