REGULATION OF ANTIMICROBIAL PEPTIDE GENES IN INSECTS BY THE TOLL PATHWAY AND NF-KB TRANSCRIPTION FACTORS

A DISSERTATION IN

Cell Biology and Biophysics
and
Molecular Biology and Biochemistry

Presented to the Faculty of the University
of Missouri-Kansas City in partial fulfillment of
the requirements for the degree

DOCTOR OF PHILOSOPHY

by

MUNMUN CHOWDHURY

M.S., University of Missouri-Kansas City, 2013
M. Phil., Calcutta University, 2007
M.Sc., Durgavati University, 2005
B. Sc., Rani Durgavati University, 2001

Kansas City, Missouri
2017

REGULATION OF ANTIMICROBIAL PEPTIDE GENES IN INSECTS BY THE TOLL PATHWAY AND NF-KB TRANSCRIPTION FACTORS

Munmun Chowdhury, Candidate for the Doctor of Philosophy Degree
University of Missouri-Kansas City, 2017

ABSTRACT

Insects lack adaptive immunity and solely rely on innate immunity to combat pathogens. Innate immunity in dipteran and lepidopteran insects mainly depends on two canonical pathways, the Toll pathway that provides protection against Gram-positive bacteria, fungi and viruses, and the Immune Deficiency (IMD) pathway, which provides protection against Gram-negative bacteria. Insect innate immune responses can be categorized broadly into two types: cellular and humoral. The cellular responses include insect blood cells known as hemocytes that are involved in phagocytosis, nodulation, encapsulation, and wound healing. The humoral immune responses, on the other hand include production of antimicrobial peptides (AMPs) by fat body, which is equivalent to mammalian liver, prophenoloxidase, lectins and compliment-like factors. Chapter 1 talks about the important features of insect innate immunity. Chapter 2 discusses the role of the Toll-7 receptor in *Drosophila melanogaster* against bacterial and viral infections. Chapter 3 depicts a new constitutively active short relish isoform lacking inhibitory domain in *D. melanogaster* that functions towards providing protection to low load of infection. Chapter 4 illustrates the role of a transcription factor in *Manduca sexta* known as Forkhead (MsFkh) that is involved in innate immunity. Lastly, Chapter 5 will summarize all data and foretell future perspectives.

APPROVAL PAGE

The faculty listed below, appointed by the Dean of the School of Graduate Studies, have examined a dissertation titled “Regulation of Antimicrobial Peptide Genes in Insects by the Toll Pathway and NF-kB Transcription Factors” presented by Munmun Chowdhury, candidate for the Doctor of Philosophy degree, and certify that in their opinion it is worthy of acceptance.

Supervisory Committee

Xiao-Qiang Yu, Ph.D., Committee Chair
Cell Biology and Biophysics

Jeffrey Price, Ph.D.
Molecular Biology and Biochemistry

Leonard Dobens, Ph.D.
Molecular Biology and Biochemistry

Alexander Idnurm, Ph.D.
Cell Biology and Biophysics

Ryan Mohan, Ph.D.
Cell Biology and Biophysics

CONTENTS

ILLUSTRATIONS

Figures Page

TABLES

Tables Page

ABBREVIATIONS

AMP	Antimicrobial peptide
ATCC	American Type Culture Collection
BGRP	beta-1,3-glucan recognition protein
bp	Base Pair
cDNA	Complementary DNA
DAP	Diaminopimelic acid
dH_2O	Deionized water
DIF	Dorsal-related immunity factor
DNA	Deoxyribonucleic acid
GNBP	Gram-negative binding protein
IL-1R	Interleukin-1 receptor
IRF	Interferon regulatory factor
IκB	Inhibitor of κB
IKK	IκB kinase
IMD	Immune deficiency
IRAK	IL-1R-associated kinase
ModSP	Modular Serine Protease
MPAE	Moricin Promoter Activating Element
mRNA	Messenger Ribonucleic Acid
NF-κB	Nuclear factor kappa B
NLR	NOD-like receptor
NLS	Nuclear localization signal

NOD	Nucleotide-binding oligomerization domain
ORF	Open Reading Frame
PAGE	Polyacrylamide Gel Electrophoresis
PAMP	Pathogen-associated molecular pattern
PCR	Polymerase chain reaction
PG	Peptidoglycan
PGRP	Peptidoglycan recognition protein
PRR	Pattern recognition receptor
RACE	Rapid amplification of cDNA ends
RHD	Rel homology domain
RNA	Ribonucleic acid
S2	Schnieder 2
SDS	Sodium dodecyl sulfate
Serpin	Serine protease inhibitor
SP	Serine protease
SPE	Spätzle processing enzyme
SPH	Serine protease homolog
Spz	Spätzle
TAB2	TAK1-binding protein 2
TAK1	TGFβ-activated kinase 1
TBS	Tris buffered saline
TCT	Tracheal cytotoxin
TGF	Transforming growth factor

TIR	Toll/Interleukin-1 receptor
TISC	Toll-induced signaling complex
TLR	Toll-like receptor

ACKNOWLEDGEMENTS

I would like to thank my advisor and committee chair Dr. Xiao-Qiang Yu for his constant guidance, help and encouragement to support me throughout my research and graduate career. I would also like to thank Dr. Jeffrey Price, Dr. Alexander Idnurm, Dr. Leonard Dobens and Dr. Ryan Mohan for their helpful comments and suggestions while serving on my graduate advising committee.

I would like to thank all former and present members of the Yu lab for kindly sharing thoughts and ideas. I want to thank Dr. Chungfeng Li for cloning all the genes encoding the *Drosophila* Toll receptors and spätzle proteins. Lastly, I thank my parents for their support in these years.

CHAPTER 1

INTRODUCTION TO INSECT INNATE IMMUNITY

Innate Immunity in Insects

Insects originated about 480 million years ago on Earth, well before the existence of mankind and other mammals. They survived this long evolutionary period depending on their remarkable innate immune system. Unlike vertebrates, insects lack an adaptive immune system, which develops in response to a challenge by an antigen and produces antibody molecules and memory cells. The innate immune system is the first line of defense that comes into action as soon as it recognizes a foreign invader, and is conserved from insects to mammals (Ferrandon et al., 2007b; Lemaitre et al., 2007; Muller et al., 2008). Insect innate immune system identify pathogens by pattern recognition receptors (PRRs) that can recognize and bind to pathogen-associated molecular patterns (PAMPs), such as lipopolysaccharide (LPS), peptidoglycan (PG), lipoteichoic acid (LTA), β-1,3-glucan, and mannan present on the invading microorganisms (Bischoff et al., 2004). Insect innate immune reactions can be categorized into cellular and humoral events. Cellular events are mediated by insect blood cells known as hemocytes, which consists of plasmatocytes (phagocytic cells), lamellocytes (for nodulation/encapsulation), and crystal cells (for melanization and clotting) in the fruit fly *Drosophila melanogaster*. The humoral events include the production of antimicrobial peptides (AMPs), prophenoloxidase cascade, lectins and a compliment like factors (Medzhitov et al., 2000). In this chapter, I will discuss the two canonical innate immune pathways of insects known as the Toll and IMD pathways that regulate the synthesis of AMPs. Then I will explain the structure of Toll receptor and Spätzle ligand. This will be followed by the description of transcription factors known as

NF-κB that regulates AMPs in insect. Lastly, I will talk about another transcription factor known as Forkhead that belongs to the Fox family of proteins and how it regulates AMPs in insects.

Recognition of Pathogens

In insects, the Lys-type peptidoglycan (PG) from Gram-positive bacteria is recognized by pattern recognition receptors such as Gram negative binding protein 1 (GNBP1) and peptidoglycan recognition proteins PGRP-SA and PGRP-SD (Choe et al., 2002; Lee et al., 1996; Pili-Floury et al., 2004). In *D. melanogaster*, so far three GNBPs have been identified, and functions of DmGNBP-1 and DmGNBP-3 have been characterized. During infection, DmGNBP-1 can recognize LPS or β-1, 3-glucans (Kim et al., 2000), and interacts with PGRP-SA to recognize the diaminopimelic acid (DAP)-type PG (Ferrandon et al., 2004). DmGNBP-3 recognizes yeast (Buchon et al., 2009; Gottar et al., 2006). Both DmGNBP-1 and -3 can activate the Toll pathway (Ferrandon et al., 2004) (Figure 1). Peptidoglycan recognition proteins (PGRPs) play important roles in activating the Toll and immune deficiency (IMD) pathways. Until now, 13 PGRP genes have been identified in *Drosophila* that encode 17 PGRP proteins by alternative splicing. Some PGRPs are short (PGRP-SA, -SB1, -SB2, -SC1A, -SC1B, -SC2 and -SD), whereas, others are long forms (PGRP-LA, -LB, -LC, -LD, -LE and -LF) (Werner et al., 2000). PGRPs are also classified based on their cellular localizations. That is, some PGRPs associated with the signal peptides could be secreted out of the cell, some are transmembrane proteins, and others remain in the cytosol (Werner et al., 2000). PGRP-SA and PGRP-SD are involved in activation of the Toll pathway (Bischoff et al., 2004; Michel et al., 2001b).

The recognition of Gram-negative bacteria is mediated by the IMD pathway and requires two transmembrane receptors PGRP-LC and PGRP-LE (Figure 2). PGRP-LC has three splicing variants: PGRP-LCx, PGRP-LCy and PGRP-LCa (Werner et al., 2003). PGRP-LCx can form oligomers on the plasma membrane to mediate the recognition of polymeric DAP-type PG (Kaneko et al., 2004; Werner et al., 2003). PGRP-LCx and PGRP-LCa dimerize and interact with monomeric PG and tracheal cytotoxin (TCT) to activate the intracellular signaling pathway (Choe et al., 2005; Kaneko et al., 2006; Neyen et al., 2012) (Figure 3).

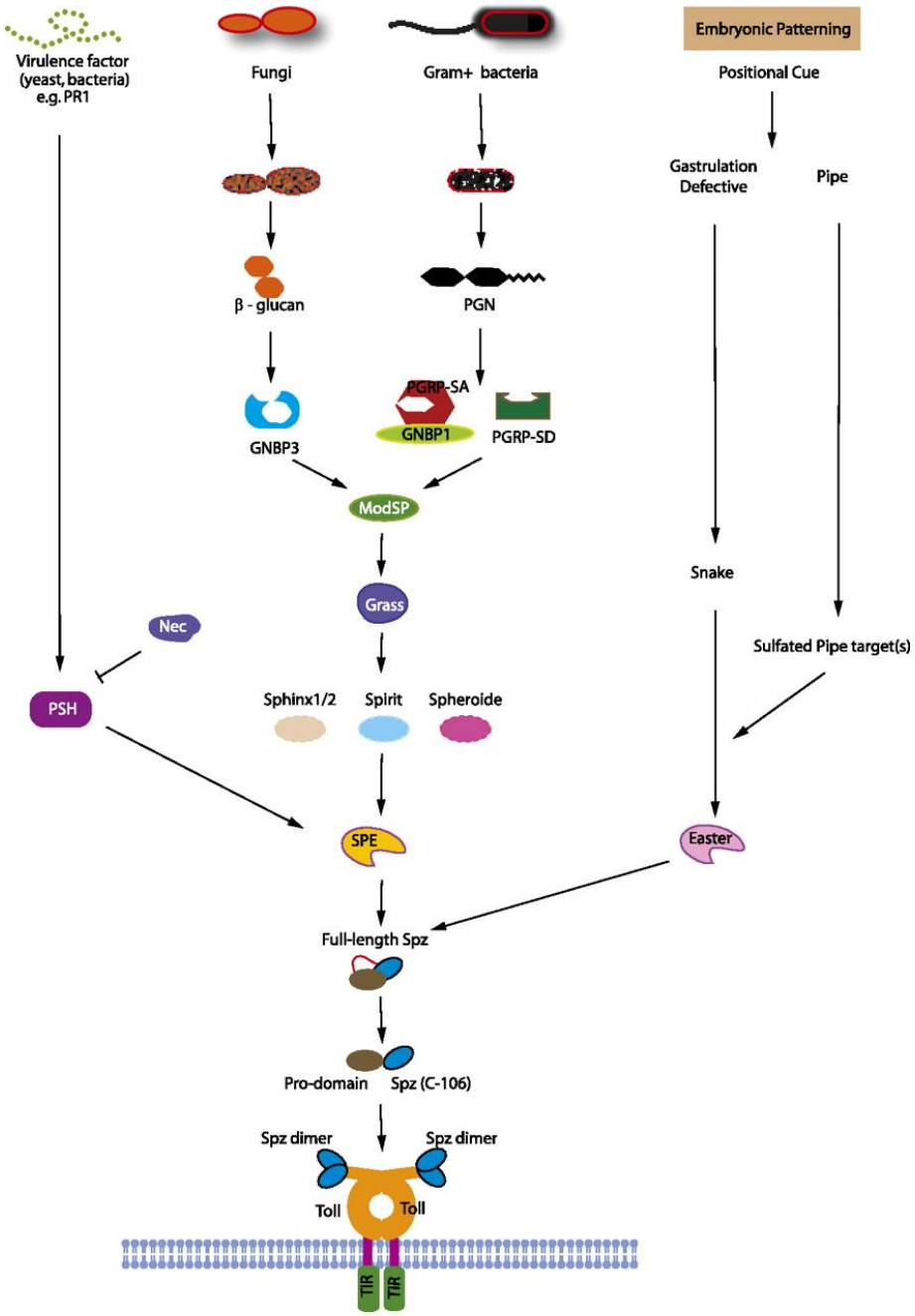


Figure 1. *Drosophila* Toll Pathway

Both early embryogenesis and immune recognition activate the Toll-Spz pathway via proteolytic processing. The GNBP3 recognizes β-glucan of fungi whereas, GNBP1, PGRP-SA and PGRP-SD recognize Gram-positive bacteria. In the development, the protease Easter cleaves full-length Spz; however, in immune response, SPE cleaves full-length Spz, exposing the C-terminal Spz fragment, which is critical for the binding to Toll.

(Valanne, S., Wang, J. H., Ramet, M., 2011. The *Drosophila* Toll signaling pathway. J Immunol. 186**,** 649-56.)

Spätzle Activation

Drosophila Toll pathway is activated either during embryonic development or during the immune challenge when extracellular recognition factors initiate protease cascades leading to the activation of the Toll receptor ligand, Spätzle (spz) (Morisato et al., 1994; Schneider et al., 1994). During the inactive condition, the N'-terminal prodomain of spz masks a predominantly hydrophobic C'-terminal spz region (Arnot et al., 2010b). Activation induces proteolysis, which causes a conformational change to expose determinants that are critical for binding to the Toll receptor (Arnot et al., 2010b). Interestingly, the prodomain remains associated with the C-terminus and is only released when the Toll extracellular domain binds to the cleaved spz (Weber et al., 2007). Two active spz dimers, each binding to the N-terminus of one of the two Toll receptors, trigger a conformational change in the Toll receptors to activate the downstream signaling (Gangloff et al., 2008) (Fig. 1). During early embryogenesis, the protease cascade Gastrulation Defective-Snake activates the protease Easter, which cleaves full-length spz (Han et al., 2000). In immune responses, three protease cascades lead to activation of Spätzle processing enzyme (SPE) to cleave full-length spz. In the Persephone (PSH) cascade, virulence factors from live Gram-positive bacteria and fungi activate PSH (Buchon et al., 2009; Ming et al., 2014). The other two cascades are activated by pattern recognition receptors that bind to cell wall components from Gram-positive bacteria and fungi, respectively. All three cascades converge at ModSP-Grass to activate SPE (Buchon et al., 2009). Upon proteolytic processing, the spz prodomain is cleaved, exposing the C-terminal spz region critical for binding to Toll (Weber et al., 2007). Spz binding to the Toll receptor initiates intracellular signaling (Weber et al., 2003). In *Drosophila*, six spz proteins have been identified so far (Parker et al., 2001). All six spz proteins contain six

conserved cysteine residues in the C'-terminal regions and hence known as Cysteine-knot domains. Structural predictions have confirmed that spz -1, 2 and -5 belongs to the neurotrophin (NT) superfamily (Weber et al., 2007) that are known to be involved in the formation of the nervous system. A search of the *Drosophila* genome using spz-1 as the query has identified a distant spz paralogs (Parker et al., 2001) known as DNT1, which is spätzle 2 (spz2) and DNT2, which is spätzle 5 (spz-5). DNT1 is more closely related to spz-1 and spz-5 (CG9972) than to other paralogs (Parker et al., 2001). Structure-based alignment while comparing to human NTs reveals that DNT1 (spz-2), spz -1 and spz-5 are more closely related to each other, whereas spz-3 (CG7104) and spz-6 (CG9196) are less closely related to vertebrates NTs (Zhu et al., 2008). Interestingly, spz-5 is highly conserved amongst insects. The Cys-knots of spz-3, spz-4 (GC14928) and spz-6 differ from the canonical NT Cys-knot: spz-3 and spz-4 have two extra cysteines and the spz-6 Cys-knot lacks two of the conserved cysteines and has three extra ones in unusual locations. Comparision of the genomic sequence shows that the Cys-knots of spz-6 and spz-4 also differ from the rest as that they lack a conserved intron (Weber et al., 2007). Additionally, sequence alignment shows that spz-4 is more similar to spz-3 (51% identity) than to other paralogs and is closest to coagulogen (29% identity) (Weber et al., 2007). Thus, the six spz paralogs can be categorized into two groups: one formed by DNT1 (spz-2), spz-1 and spz-5, and the other formed by spz-3, spz-4, and spz-6 (Zhu et al., 2008).

Toll Pathway

The activated Toll dimer interacts with the adaptor protein Myeloid differentiation primary response gene 88 (MyD88) through their intracellular Toll/Interleukin-1 receptor

(TIR) domains (Horng et al., 2001; Sun et al., 2002). MyD88 is a death domain (DD)-containing adaptor, which comes into contact with another DD adaptor Tube that in turn recruits kinase Pelle. MyD88, Tube and Pelle associate to form a heterotrimeric complex through the death domains. MyD88 does not directly interact with Pelle, but the two distinct DD surfaces of Tube bind to MyD88 and Pelle (Sun et al., 2002). The MyD88-Tube-Pelle heterotrimer is critical for proceeding to the phosphorylation and degradation of *Drosophila* IκB factor Cactus. In non-active conditions, Cactus associates with the transcription factors of the Rel/NF-κB family Dorsal and Dorsal-related immunity gene (DIF), retaining them in the cytosol and inhibiting their nuclear localization (Wu et al., 1998). During the immune challenge, the active kinase Pelle phosphorylates and degrades Cactus (Towb et al., 2001). Dorsal or DIF are released from Cactus-Dorsal or Cactus-DIF complex and translocated into the nucleus to activate the transcription of several target genes by binding to the NF-κB sites of promoters (Figure 2).

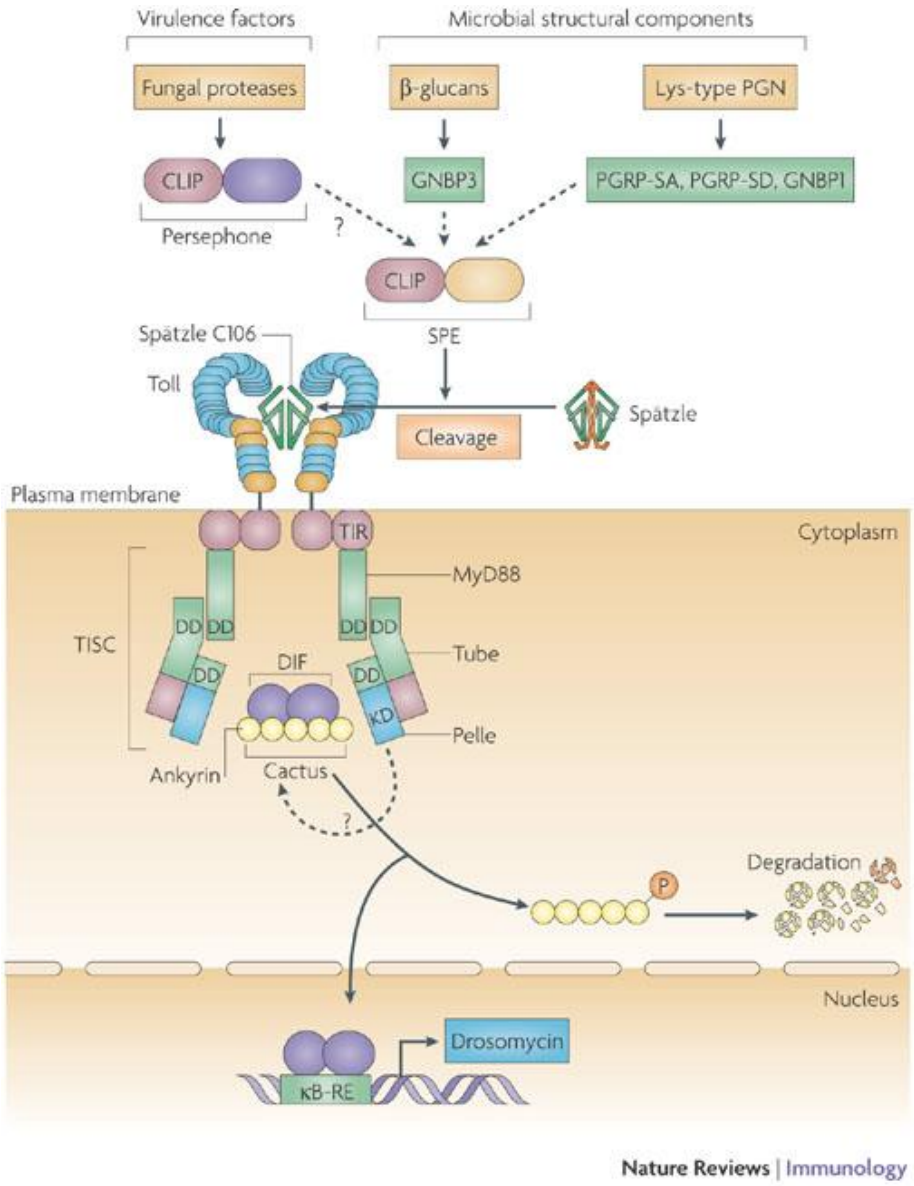


Figure 2. *Drosophila* Spätzle Activation.

The Toll pathway is activated upon dimeric Spätzle-C106 binds with Toll receptor to induce Toll-induced signaling complex (TISC), which is composed of three death-domain (DD)-containing proteins, MyD88 (myeloid differentiation primary-response gene 88), Tube and Pelle. Cactus may be phosphorylated by Pelle, phosphorylated Cactus is rapidly polyubiquitylated and degraded, allowing for the nuclear translocation of DIF, and binding to NF-κB response elements (κB-RE), which in turn induces the expression of genes encoding antimicrobial peptides, such as *Drosomycin* (Ferrandon et al., 2007a).

(Ferrandon D., Imler J.-L., Hetru C., Hoffmann J. A. 2007a. The Drosophila systemic immune response: sensing and signaling during bacterial and fungal infections. Nat Rev Immunol 7: 862-874).

Toll Receptors

There are 11 Toll-like-Receptors (TLRs) in mice and 9 in humans, 9 Toll receptors in *Drosophila* and 16 in *M. sexta*. Both Toll and TLR contain a highly conserved extracellular domain made up of Leucine-Rich Repeats (LRRs) and Cysteine-Rich motifs followed by a transmembrane region and trailed by an intracellular domain known as the Toll/interleukin-1 receptor (TIR) homology domain (Figure 3). *Drosophila* Toll-9 is closest to mammalian TLR-4 (Figure 4). In mammals, TLRs directly recognize and bind to different Pathogen-associated molecular patterns (PAMPs) with the help of co-receptor proteins. These bindings trigger a conformation change, which induces the formation of TLR homo- or heterodimers to initiate the intracellular signal. In contrast, insect Toll does not directly detect or bind to PAMPs but plays a crucial role in activating the downstream cascades. It has been shown that *Drosophila* Toll-1 is necessary for the Toll-Spätzle pathway in both immune recognition and embryogenesis (Lemaitre et al., 1996). Research has shown that *drosomycin* (an antimicrobial peptide gene) is upregulated by Toll-9 in *D. melanogaster* Schneider 2 (S2) cell lines during infection (Narbonne-Reveau et al., 2011; Ooi et al., 2002). The active *Drosophila* Toll receptors form dimers, which bind to the intracellular adaptor protein MyD88 via a TIR domain, to initiate the downstream signaling pathway (Horng et al., 2001). A Toll receptor in *M. sexta*, which can be induced by some Gram-positive/negative bacteria or fungi, was reported, and the TIR domain of *Ms*Toll has high similarity to that of vertebrate TLR4 (Ao et al., 2008). Recently, researchers found that *Drosophila* Toll-5 and 9 could activate *drosomycin* promoter (Ooi et al., 2002; Tauszig et al., 2000). Toll-6 and Toll-7 are expressed in central nervous system (CNS) and help in the formation of CNS (Mcilroy et al., 2013a). It has also been reported that *Drosophila* Toll-7 is involved in recognition of

vesicular stomatitis virus infection (Nakamoto et al., 2012). In *Drosophila*, research shows that Toll-7 could reduce the Rift Valley Fever Virus (RVFV) replication and mortality through activation of autophagy (Moy et al., 2014). Conversely, another research group reported contradictory results where they found that during VSV infection, hemocytes use autophagy to recognize the virus but this recognition does not depend on the Toll-7 receptor. (Lamiable et al., 2016b). Similarly, in other lepidopteran insects like *M.sexta* Toll interacts with *Ms*Spz-C108 and co-expression of MsToll-MsSpz-C108, could up-regulate expression of *drosomycin* gene in *Drosophila* S2 cells, indicating that MsToll-MsSpz-C108 complex can activate the Toll signaling pathway (Zhong et al., 2012b).

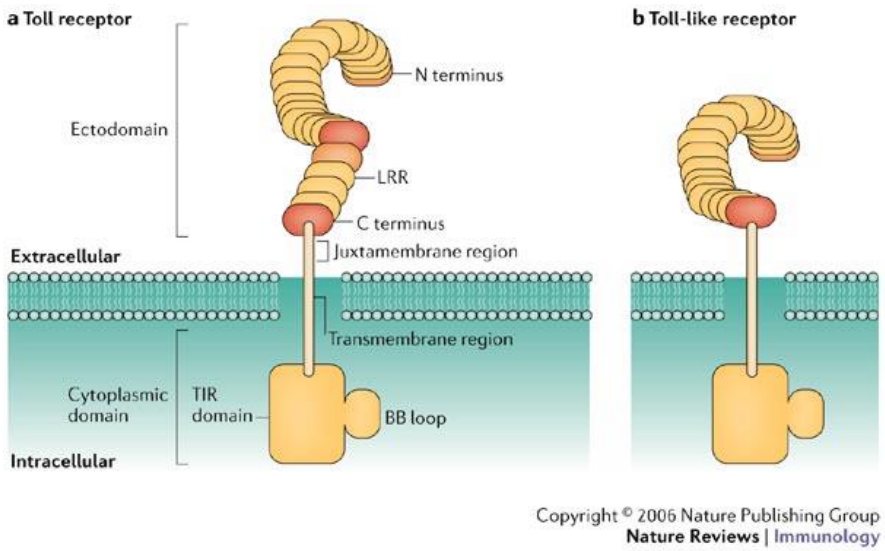


Figure 3. Schematic Diagram of Toll Receptors

Insect Toll receptor (a) and mammalian Toll-like receptors (b), showing the extracellular, transmembrane and cytoplasmic domains of the receptors. Note that *Drosophila* Toll has two independent blocks of leucine-rich repeats (LRRs) in the extracellular domain (ectodomain). C terminus, carboxyl terminus; N terminus, amino terminus; TIR domain, Toll/interleukin-1–receptor-containing domain (Gay et al., 2006).

(Gay N. J., Gangloff M., Weber A. N. R. 2006. Toll-like receptors as molecular switches. Nat Rev Immunol 6: 693-698).

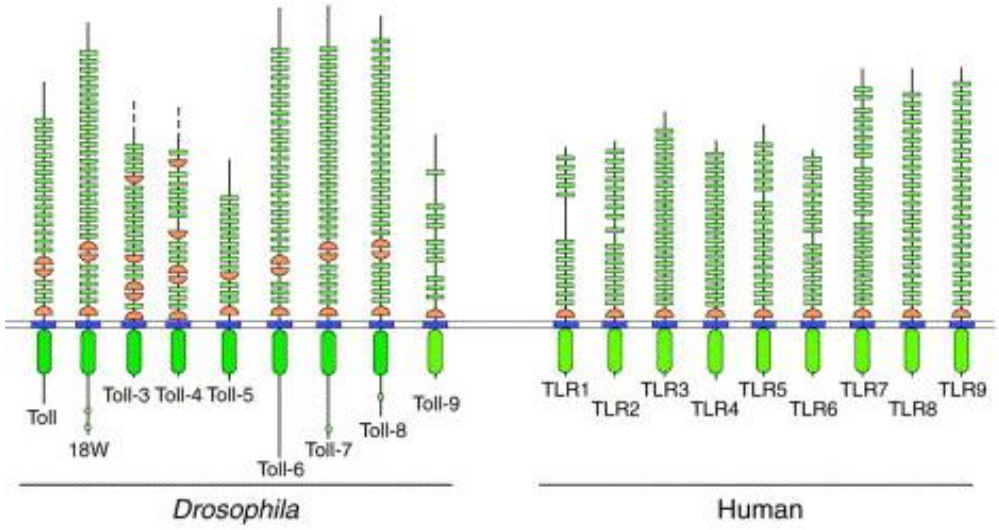


Figure 4. Comparison of *Drosophila* Toll and Human Toll-like Receptors.

The green rectangles are lucine-rich repeats, red semi-circles are cysteine-rich motifs and oblong bright green is the TIR domain (Imler et al., 2001).

(Imler J. L., Hoffmann J. A. 2001. Toll receptors in innate immunity. Trends Cell Biol 11: 304-311).

Insect IMD Pathway

In insects, Gram-negative bacterial infections are recognized by the IMD pathway. IMD signaling is initiated by DAP-type PG from Gram-negative and some Gram-positive bacteria. PGRP-LC has three alternative splice variants PGRP-LCa, PGRP-LCx and PGRP-LCy, which are critical for the initiation of the IMD pathway (Werner et al., 2003). PGRP-LCa and PGRP-LCx proteins contain a transmembrane domain and can recognize tracheal cytotoxin (TCT) in DAP-type PG (Chang et al., 2005a; Chang et al., 2006; Chang et al., 2005b; Kaneko et al., 2006; Mellroth et al., 2005). The function of PGRP-LCy is somewhat redundant. DAP-type PG triggers the formation of a heterodimer of PGRP-LCa and PGRP-LCx or the homodimer of PGRP-LCx (Werner et al., 2000). PGRP-LE lacks the transmembrane domain. Inside the cell, PGRP-LE recognizes PG; however, outside the cell, PGRP-LE is the co-receptor for PGRP-LC (Kaneko et al., 2006; Yano et al., 2008). Other PGRP family members play either negative or positive roles (Stenbak et al., 2004). For example, PGRP-SC1b inhibits the activation of AMPs in *Drosophila* blood cell lines (Mellroth et al., 2003).

Binding of PG to the extracellular domains triggers the formation of PGRP-LC hetero- or homodimers and the cytoplasmic domains interact with IMD. IMD is a DD-containing protein, which recruits an adaptor protein Fas-associated death domain-containing protein (FADD) (Georgel et al., 2001). FADD is connected to a caspase Death-related ced-3/Nedd2-like protein (DREDD). DREDD proteolytically cleaves IMD and the transcription factor NF-κB protein Relish (Leulier et al., 2000; Paquette et al., 2010; Stoven et al., 2000). Both IKK complex and DREDD are required for the cleavage of transcription factor Relish (Silverman et al., 2000a). Relish contains an N-terminal Rel homology domain (RHD) and a

C-terminal inhibitor of NF-κB (IκB) motif. In non-infectious condition, the IκB part inhibits RHD activity and nuclear localization. During infection, Relish is phosphorylated and cleaved by kinase immune response deficient 5 (Ird5) and caspase DREDD. The IκB region is degraded and the RHD is released and, translocates to the nucleus to activates the transcription of AMPs (Stoven et al., 2000; Stoven et al., 2003) (Figure 5).

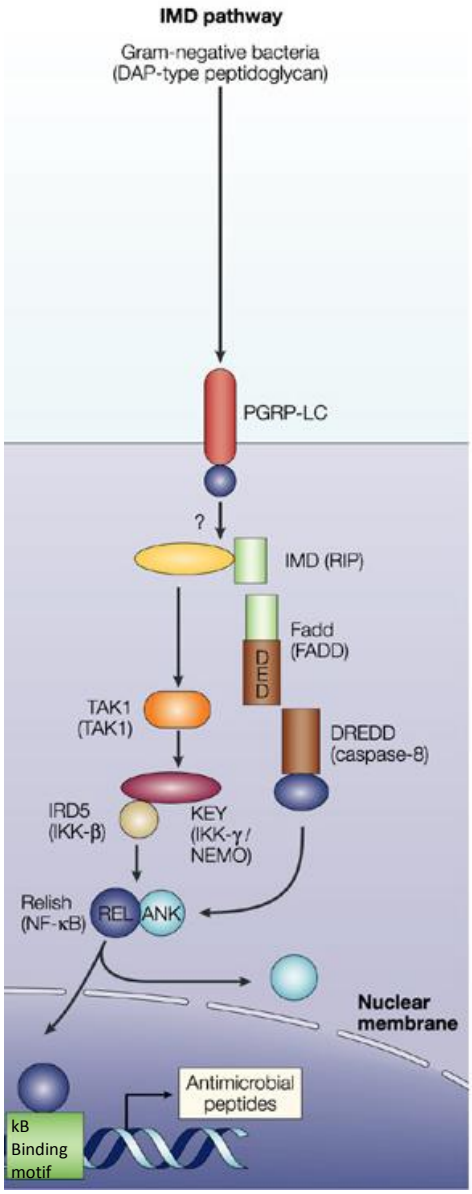


Figure 5. *Drosophila* IMD pathway.

The IMD pathway is triggered by an interaction between the transmembrane receptor PGRP-LC and peptidoglycan from Gram-negative bacteria (diaminopimelate (DAP)-type peptidoglycan). Following PGRP-LC activation, the death-domain (DD) adaptor protein, IMD, is recruited and binds to Fadd, which interacts with the caspase DREDD (Death-related ced-3/Nedd2-like protein). DREDD associates with Relish after Relish has been phosphorylated by the *Drosophila* IKK (inhibitor of NF-κB (IκB)-kinase) complex, which comprises Immune response deficient 5 (IRD5) and Kenny (KEY). The IKK complex is itself activated by TAK1 (Transforming-growth-factor-beta-activated kinase 1, a mitogen-activated protein kinase kinase kinase) in an IMD-dependent manner. After cleavage, the RHD domain of Relish moves to the nucleus, where it regulates the transcription of genes with immune function (Lemaitre 2004).

(Lemaitre, 2004. The road to Toll. Nat Rev Immunol 4: 521-527).

NF-κB Transcription Factors

Nuclear factor kappa B (NF-κB), which consists of a family of transcription factors in insects and mammals, plays a crucial role in apoptosis, inflammation, immunity, cell proliferation, differentiation, and survival. Yet, it is absent in bacteria, fungi, plants, and *C. elegans.* In humans, there are five NF-κB members (Figure 6); and the *Drosophila* genome encodes three NF-κB homologs (Figure 6). Based on different regulatory mechanisms, NF-κB transcription factors are divided into four groups.

Class I NF-κB factors contain human RelA (p65), RelB, c-Rel, *Drosophila* Dorsal, and DIF. They share a conserved 300 amino acids Rel-homology domain (RHD) at the N-terminal region, which is critical for the formation of homo- or heterodimers, binding to promoters, interaction with IκB, and nuclear translocation. They also contain a C-terminal transactivation domain (TAD) (Baldwin 1996; Ghosh et al., 1998) (Figures 5 and 6). In this group, only RelB includes an N-terminal Leucine zipper motif (Dobrzanski et al., 1993). The primary mechanism for regulating gene expression in class I NF-κB factors is phosphorylation-induced proteolysis. In non-infectious conditions, RHD is sequestered in the cell cytoplasm by ankyrin (ANK) repeats of IκB. Whereas, during infection, the IKK complex phosphorylates, ubiquitinates, and degrades IκB to release NF-κB. The active NF-κB translocates into the nucleus, binds to the κB sites of AMP gene promoters to activate the target genes (Brown et al., 1995). *Drosophila* Dorsal and DIF are 45% identical to mammalian c-Rel, RelA, and RelB. Dorsal and DIF are retained in the cytosol via binding to Cactus, which is homologous to mammalian IκBa (Geisler et al., 1992; Huguet et al., 1997). Cactus contains a conserved IκB domain, ANK repeats, an N-terminal acidic domain (Ac),

and C-terminal proline, glutamic acid, serine and threonine-rich segment (PEST) (Oeckinghaus et al., 2009a) (Figure 6).

The class II NF-κB factors include human p100 and p105, which are the precursors of p52 and p50, respectively. Proteolysis is required for activation of the class II NF-κB factors. They contain an RHD at the N-terminal region, followed by a glycine-rich region (GRR) and multiple copies of ANK, and a DD at the C-terminal region. C-terminal ANK is characteristic of IκB, which inhibits nuclear translocation and transcriptional activity of the N-terminal RHD domain. Under immune challenge, p100 and p105 are proteolytically cleaved after phosphorylation and ubiquitination to remove ANK repeats, then the active p52 and p50 are released (Huguet et al., 1997; Oeckinghaus et al., 2009b).

The class III NF-κB factor is *Drosophila* Relish, which contains an N-terminal RHD and C-terminal ANK repeats typical of mammalian IκB and *Drosophila* Cactus. *Drosophila* Relish requires caspase-mediated cleavage to degrade ANK repeats and release RHD the region, which translocates to the nucleus and activates the transcription of target genes (Stoven et al., 2003). In Relish, other than RHD and ANK domains, the serine-rich region (SRR) and the PEST region seem to regulate Relish activities (Stoven et al., 2003).

Class IV NF-κB factors consist of a family of nuclear factors of activated T-cells (NF-AT) transcription factors and induce the downstream genes in the adaptive immune system.

Drosophila has 3 NF-κB transcription factors, Dorsal, DIF, and Relish (Fig.7). *Drosophila* Dorsal was originally identified as a critical component that determines the correct dorsal-ventral pattern of *Drosophila* embryos along with other maternal effector genes (Anderson et al., 1985). Later, κB-binding sites were identified in the regulatory

regions of many immune-related genes, suggesting the involvement of the Toll pathway in the regulation of immune genes (Engstrom et al., 1993; Kappler et al., 1993; Reichhart et al., 1993). The other two *Drosophila* NF-κB factors DIF and Relish were identified subsequently (Dushay et al., 1996b; Ip et al., 1993). In larvae, Dorsal and DIF play redundant roles in regulating AMP gene expression and promote blood cell survival (Manfruelli et al., 1999a; Matova et al., 2006). In adults, DIF is the major factor (Rutschmann et al., 2000a; Rutschmann et al., 2000b). DIF and Dorsal are the NF-kB factors of Toll pathway.

Once inside the nucleus, dimers of DIF/DIF, Dorsal/Dorsal, DIF/Dorsal, DIF/Relish, Dorsal/Relish or Relish/Relish binds to specific sequences in the promoter/regulatory region of some immune genes to induce gene expression. Mammalian NF-κB factors bind to a consensus sequence GGG(G/A)NN(T/C)(T/C)CC (Baeuerle 1991). One study in *Drosophila* shows that DIF prefers binding to GGGAAA(A/T/G)(C/T)CC, while Relish prefers binding to GGGGATT(T/C)(T/C)(T/C) (Busse et al., 2007). Some binding sites such as GGGAATTCCC can bind to both Relish and DIF (Busse et al., 2007). However, another report based on a different assay showed different results (Senger et al., 2004), indicating the ambiguity about consensus binding sequence. One particularly interesting example that shows the complexity of this is the *Drosophila* drosomycin promoter. There are two functional NF-κB sites in the promoter; site1 binds to DIF and Dorsal, while site2 binds to Relish, but DIF/Relish heterodimers bind to site2 (Tanji et al., 2010c). Simultaneous activation of the Toll and IMD pathways may cause a crosstalk between the two pathways, and the promoter activity is significantly higher when both pathways are activated than either pathway is activated individually (Tanji et al., 2007b; Tanji et al., 2010a). But whether and how the synergic effect is due to the formation of heterodimers is still not clear.

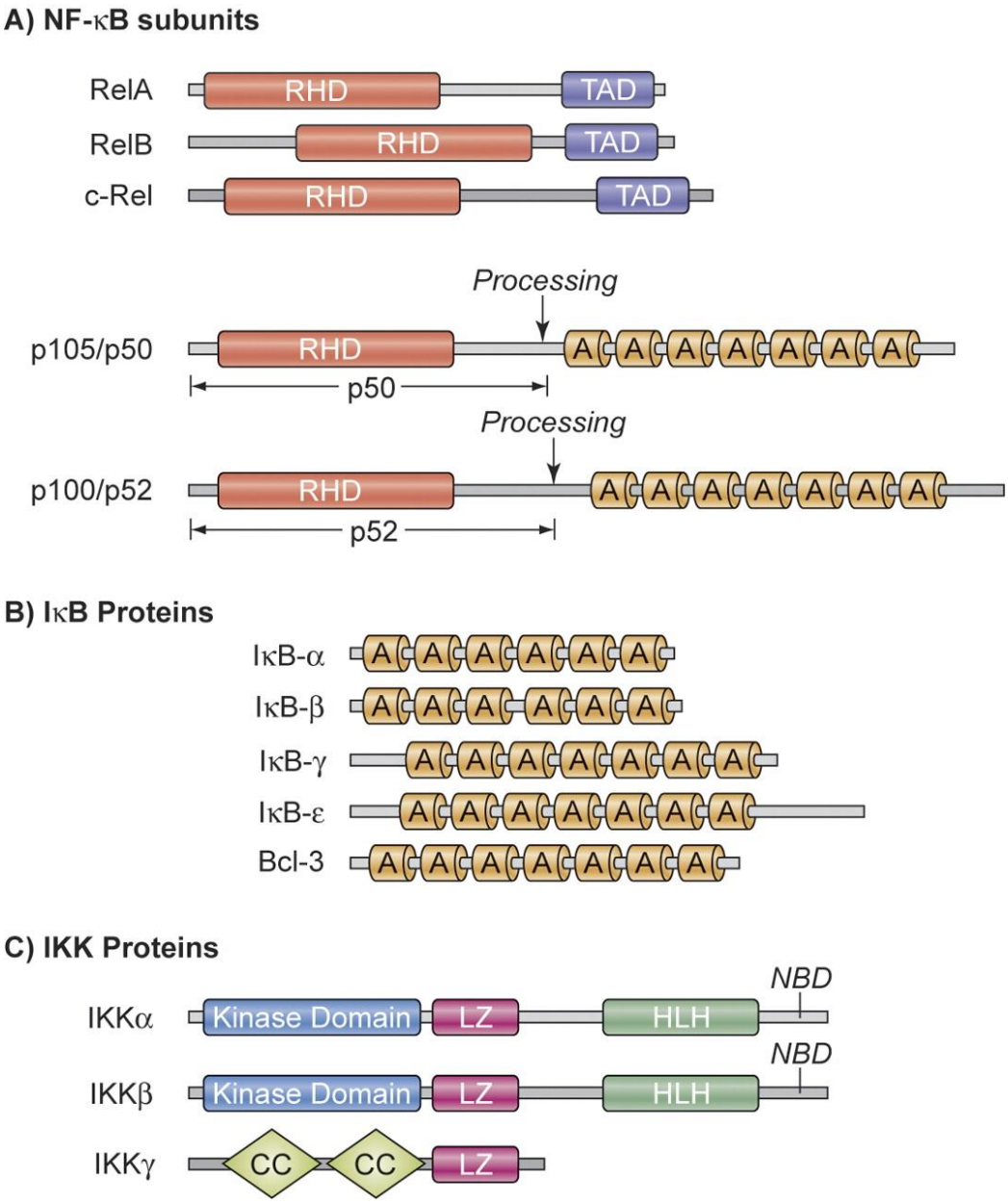


Figure 6. Mammalian Members of the NF-κB, IκB and IKK Protein Families

ANK, ankyrin repeats; CC, coiled-coil; DD, death domain; GRR, glycine-rich region; HLH, helix-loop-helix; IKK, IκB kinase; LZ, leucine-zipper; NBD, NF-κB essential modulator (NEMO)-binding domain; PEST, proline-, glutamic acid-, serine- and threonine-rich region; TAD, transactivation domain; ZF, zinc finger (Jost et al., 2007).

(Jost P. J., Ruland J. 2007. Aberrant NF-κB signaling in lymphoma: mechanisms, consequences, and therapeutic implications. Blood 109: 2700-2707).

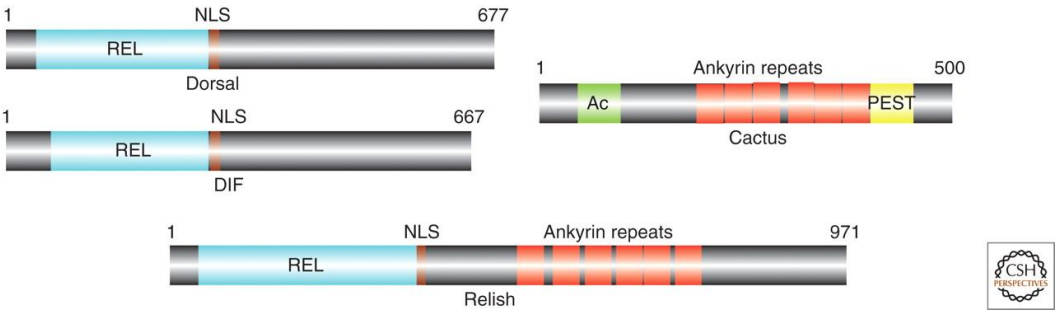


Figure 7. NF-κB and IκB Proteins in *Drosophila*.

REL, Rel-homology domain; NLS, nuclear localization sequence; PEST, proline, glutamic acid, serine and threonine-rich segment; Ac, acidic domain.

(Hetru, C., Hoffmann, J. A., 2009. NF-kappaB in the immune response of Drosophila. Cold Spring Harb Perspect Biol. 1, a000232.)

Forkhead Box Transcription Factors

The forkhead box, or Fox, gene family of transcriptional regulators is an evolutionarily ancient gene family that is named after *D. melanogaster* gene *fork head* (*fkh*). Mutations in *fkh* gene cause defects in head fold involution during embryogenesis, resulting in a characteristic spiked head appearance in adult flies (Weigel et al., 1989). Hundreds of Fox genes have been identified in species ranging from yeasts to humans (Hannenhalli et al., 2009; Jackson et al., 2010). Genetic analyses have shown that many of these genes have important biological functions in multiple species, from the control of the cell cycle to differentiation of epithelia, and from placental development to formation of the inner ear (Hannenhalli et al., 2009). The evolutionary conservation of the crucial DNA-binding domain between orthologous members of the Fox gene family is remarkable; for example, there is 90% amino acid similarity between *D. melanogaster* Forkhead and human FOXA1 protein. The unifying feature of Fox proteins is the ~100-residue forkhead (FKH) DNA-binding domain, which is highly conserved across all members of the Fox family. In 1990, Weigle et al (Weigel et al., 1989) noticed the astonishing similarity of the central 110 amino acids of *D. melanogaster* and mammalian Fox proteins. These 110 amino acids are inside the DNA-binding domains identified by Lai and colleagues (Lai et al., 1990), and thus suggested the name forkhead domain for this characteristic DNA-binding motif. The canonical FKH domain consists of three α-helices, three β-sheets and two 'wing' regions that flank the third β-sheet. Because of the butterfly-like winged structure adopted by the DNA-bound Fox proteins, the FKH domain has also been termed as the winged-helix domain (Lehmann et al., 2003a), yet the winged-helix structure is not unique to Fox proteins. Most Fox proteins bind to DNA as monomers, contacting their target sequences by the third α-helix, the flanking

residues, and the two wings. So far, 50 Fox proteins have been discovered in humans (Jackson et al., 2010), and they are categorized from Fox A to Fox Q (Figure 8). There are at least 4 Fox genes in yeast, 16 in *D. melanogaster*, 15 in *C. elegans* and 44 in mice (Mazet et al., 2003; Tuteja et al., 2007a; b). The genome of *Aedes aegypti* contains eighteen loci that encode putative Fox factors and six of them are involved in reproduction (Hansen et al., 2007). In *D. melanogaster*, several Fox transcription factors have been described. These include Fox A to D, Fox F to G, Fox K, Fox L, Fox N, Fox O, Fox P, Fox Q, and one orphan Fox gene CG32006 (Figure 10) (Carlsson et al., 2002; Mazet et al., 2003). Forkhead (Fkh) is the founding member of the FoxO family and it is activated upon TOR inhibition by rapamycin (Varma et al., 2014). In Lepidopteran insect *Bombyx mori*, a transcription factor, silk gland factor-1 (SGF-1), is homologous to *D. melanogaster* forkhead (Fkh). SGF-1 can bind *in vitro* to a *cis*-element located in the promoter region of *sericin-1*, which encodes a silk protein in the middle silk gland (MSG) cells (Hui et al., 1990; Mach et al., 1995; Xu et al., 1994). However, existence and functions of Fox proteins in other Lepidopteran insects are yet to be revealed.

The structural organization of all forkhead transcription factors shares a highly conserved forkhead DNA-binding domain (FHD), which binds to the conserved sequences in the target genes. The FHD can be located to the amino‑terminal and carboxyl‑terminal regions (Lam et al., 2013). Among the 16 *Drosophila* forkhead proteins, Fork head (Fkh) is the only one belonging to the FoxA subfamily (Bulow et al., 2010). *Drosophila* Fkh domain consists of 114 amino acids, having a globular structure with 37% alpha-helix content. Specific interaction with DNA is mediated by two contact regions, separated by one turn of

DNA (Kaufmann et al., 1994) (Figure 9). Fkh binds to the consensus sequence of (T/C)(G/A)AAACAA (Li et al., 2004).

It has been shown that forkhead protein plays an important role in development, insulin signaling, immunity, metabolism, cancer and so on. Study the role of forkhead protein in immunity in beneficial or harmful insects could eventually benefit human society.

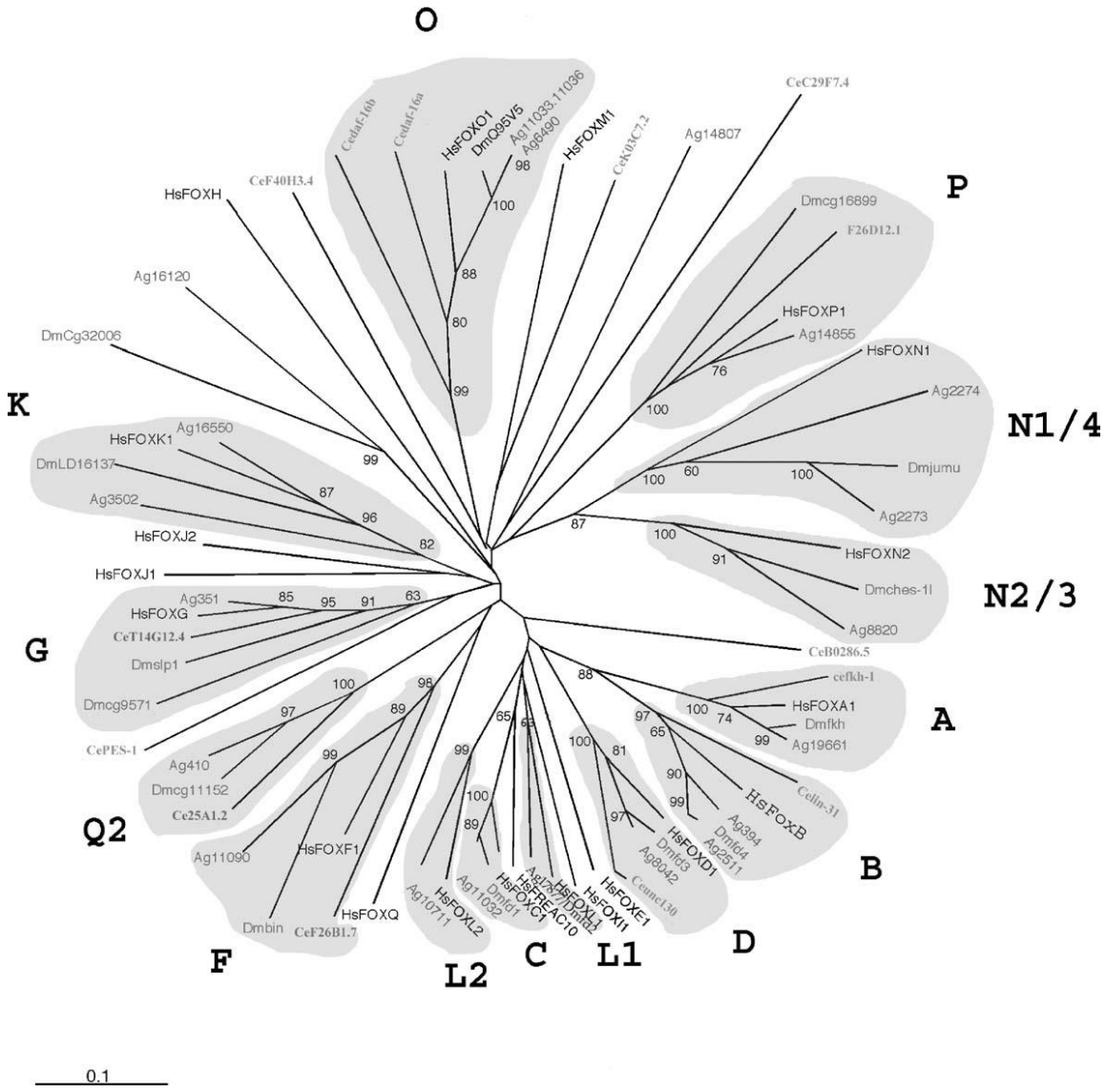


Figure 8. Phylogenetic Tree of Fox Gene Family from *C. elegans* (Ce), *D. melanogaster* (Dm), *A. gambiae* (Ag) and *H. sapiens* (Hs).

Only one gene from each subclass of human fox genes has been selected. Subclasses with genes present in vertebrates plus at least one ecdysozoan are shaded (Mazet et al., 2003).

(Mazet F., et al. 2003. Phylogenetic relationships of the Fox (Forkhead) gene family in the Bilateria. Gene 316: 79-89).

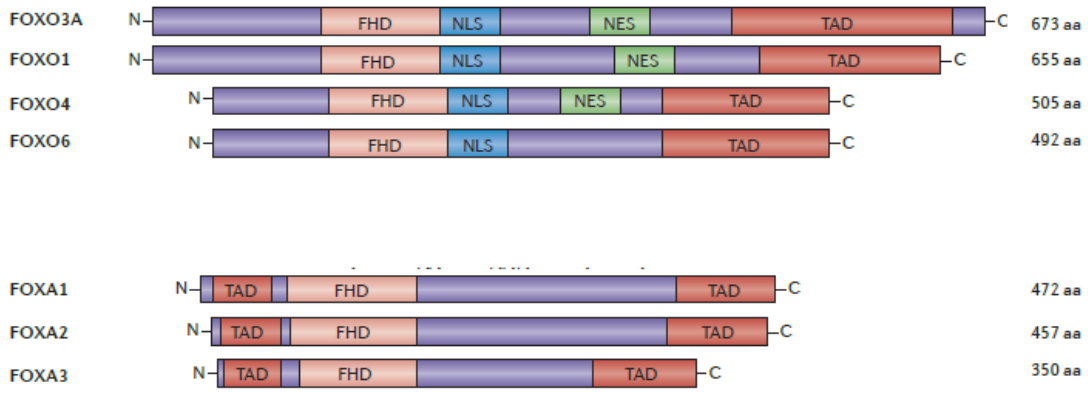


Figure 9. Structural Organization of Human FOXO and FOXA Subfamilies.

FDH, Forkhead DNA binding domain; NES, nuclear export sequence; NLS, nuclear localization sequence; TAD, transactivation domain (Lam et al., 2013).

(Lam E. W. F., Brosens J. J., Gomes A. R., Koo C.-Y. 2013. Forkhead box proteins: tuning forks for transcriptional harmony. Nat Rev Cancer 13: 482-495).

Fox Subclass	*C. elegans*	*D. melanogaster*	*A. gambiae*	*C. intestinalis*	*B. floridae*	*H. Sapiens*
A						
B						
C						
D						
E						
F						
G						
H						
I						
J						
K						
L1						
L2						
M						
N1/4						
N2/3						
O						
P						
Q1						
Q2						

Figure 10. Fox Gene Distribution by Subclass in Different Taxa.

Filled boxes indicate that a member of that subclass has been detected in the indicated species. For simplicity, only human genes are shown as representative of vertebrates.

(Mazet F., et al. 2003. Phylogenetic relationships of the Fox (Forkhead) gene family in the Bilateria. Gene 316: 79-89).

Insect Antimicrobial Peptides (AMPs)

AMPs are small cationic proteins (12 – 50 amino acids) that are either stored in granules or rapidly produced and secreted into insect hemolymph during infection. AMPs are produced mainly in fat body (an equivalent of the mammalian liver) and are also synthesized in hemocytes and epithelial cells. Most AMPs are amphiphilic molecules that can specifically disrupt the bacterial membrane by forming pores (Dathe et al., 1999; Hancock et al., 1998; Oren et al., 1998; Pieters et al., 2009; Wang 2015). In *Drosophila*, there are seven families of AMPs, as summarized in Table 1. Diptericin (Dimarcq et al., 1990; Wicker et al., 1990), Attacin (Carlsson et al., 1998; Kockum et al., 1984), Drosocin (Bulet et al., 1996) and Cecropin (Boman et al., 1989) have antibacterial activities against Gram-negative bacteria. Defensin (Hoffmann et al., 1992) has activity against Gram-positive bacteria. Drosomycin (Fehlbaum et al., 1994) and Metchnikowin (Levashina et al., 1995) are anti-fungal peptides (Table 1). AMPs can also be classified based on structures, some have intramolecular disulfide bonds, while some other are linear peptides.

Table 1 Antimicrobial Peptides in *D. melanogaster*

AMPs	Main Function	Reference
Diptericin	Gram-negative bacteria	(Dimarcq et al., 1988; Wicker et al., 1990)
Attacin	Gram-negative bacteria	(Kockum et al., 1984)
Drosocin	Gram-negative bacteria	(Bulet et al., 1993)
Cecropin	Gram-positive and Gram-negative bacteria, Fungi	(Boman et al., 1991;Ekengren and Hultmark, 1999)
Defensin	Gram-positive bacteria	(Hoffmann and Hetru, 1992; Imler and Bulet, 2005)
Drosomycin	Fungi	(Fehlbaum et al.,1994; Landon et al., 1997)
Metchnikowin	Fungi	(Levashina et al., 1995)

Questions and Hypotheses

The Toll and Imd pathways are the two conserved canonical innate immune pathways in invertebrates. The Toll pathway was initially identified for its role in the early dorsoventral (DV) patterning of *Drosophila* embryo. In all, nine toll receptors have been identified in *D. melanogaster* and only the functions of Toll (Toll-1) have been well studied so far. Similarly, six spätzles have been identified in *D. melanogaster*. Although multiple Toll and Spz genes have been identified *D. melanogaster,* interactions between a Toll receptor and six Spz ligands in eliciting innate immune responses have not been demonstrated in *D. melanogaster*. In Chapter 2, I will propose the existence of multiple Toll-Spz pathways in *D. melanogaster* in innate immunity. The Toll pathway is activated by Gram-positive bacteria, fungi, and viruses, whereas, the IMD pathway is activated by Gram-negative bacteria in *D. melanogaster*. Activation of the Toll pathway eventually activates NF-κB factor called DIF and Dorsal, whereas, the IMD pathway leads to activation of NF-κB factor called Relish. These NF-κB factors then translocate into the nucleus and bind to κB sites of AMP gene and other immune-related gene promoters and activate these genes. Relish has two domains, the N-terminal RHD domain that binds to DNA and C-terminal inhibitory domain that contains 6 ankyrin repeats. As long as the inhibitory domain is bound to Relish, it is inactive and once the inhibitory domain is cleaved and free from the C-terminal domain, the N-terminal RHD domain moves into the nucleus and activates DNA transcription. In several insect species, it has been revealed that there is a short Relish isoform that does not contain the C-terminal inhibitory domain and thus is always active. However, in *D. melanogaster*, such a relish short isoform has not been

reported. In Chapter 3, I hypothesize that *D. melanogaster* also has the relish short isoform that is always active to provide the basal level of activity and protection to the organism.

Antimicrobial peptides (AMPs) play an important role in defense against microbial infections in insects. Expression of AMPs is regulated mainly by NF-κB factors Dorsal, DIF and Relish. Studies in *M. sexta* showed that both NF-κB and GATA-1 factors are required for activation of moricin gene promoter, and there may be additional transcription factors that are involved in regulation of AMP expression. In Chapter 4, I hypothesize the presence of additional transcription factor binding sites, Forkhead (Fkh) binding sites, in AMP gene promoters of *M. sexta*. These Fkh binding sites are required for activation of AMP gene promoters by Fkh factors. Lastly, in Chapter 5, I will summarize my research and provide future directions for my research work.

CHAPTER 2

MULTIPLE TOLL- SPÄTZLE PATHWAYS IN *DROSOPHILA MELANOGASTER* INNATE IMMUNITY

Abstract

In *Drosophila melanogaster*, Toll receptor is involved in dorsal-ventral patterning during embryonic development and also in regulation of antimicrobial peptide (AMP) genes in larvae and adult flies. Since the discovery of *Drosophila* Toll, over ten Toll-like receptors (TLRs) have been identified in vertebrates, which can recognize bacteria, fungi, and viruses to trigger different signaling pathways. *Drosophila* genome encodes nine Toll receptors and six Spätzle (spz) proteins; however, only *Drosophila* Toll (Toll-1) and spz (spz1) have been well characterized in the canonical Toll pathway. Recently, it has been reported that *Drosophila* Toll-7 can recognize vesicular stomatitis virus (VSV), Toll-6 and Toll-7 function as neurotrophin receptors. Our results showed that *Drosophila* Toll-1 and Toll-7 were functionally active receptors that could activate *drosomycin* (an AMP gene regulated by the Toll pathway) promoter activity. The Co-IP result shows that Toll-1 and Toll-7 could interact with spz-1, -2 and -5 and additionally Toll-7 could interact with spz-6. Our result suggests that both Toll-1 and Toll-7 functions in the regulation of AMPs in *Drosophila* and they may work in concert or respond to different microbial infections. Septic infection with *Candida albicans* shows that Toll-7 mutant males are more susceptible than any other tested group. We showed that in *Drosophila* multiple Toll-Spatzle pathway functions in innate immunity. Determination of such pathways will

greatly help to understand insect defense mechanisms against microbial infections and could provide insight into the evolution of Toll and TLR signaling pathways from insects to humans.

Introduction

Toll receptors are a group of proteins that play essential roles in the first line of defense against invading pathogens. Toll receptors also play a significant role in inflammation, immune cell regulation, survival, and proliferation. Toll receptor superfamily consists of Toll and Toll-like receptor (TLR) and is present in all metazoan (Leulier et al., 2008). Toll receptors are single pass transmembrane pattern-recognition receptors (PRRs) that recognize various pathogens (Pasare et al., 2005) and activate NF-κB pathways to initiate innate or adaptive immunity (Iwasaki et al., 2004; Wasserman 1993). In humans, there are nine TLRs that directly bind to pathogens to trigger the releasing of pro-inflammatory substances (Akira et al., 2006; Takeuchi O 2001). Mammalian TLR is the functional ortholog of *Drosophila* Toll (Taguchi 1996). In *D. melanogaster*, the Toll pathway was discovered as an essential part of dorsoventral patterning during early embryogenesis (Nüsslein-Volhard 1980). Later, genetic screens established the function of the Toll pathway and NF-κB in mammalian innate immunity (Belvin 1996; Sun et al., 1992). It became more evident that interconnected signaling pathways exist in *Drosophila* for embryonic development and eliciting innate immune responses. *Drosophila* mounts two types of immune responses upon infection: humoral and cellular responses (Ferrandon et al., 2007a). Humoral responses include the production of Anti-Microbial Peptides (AMPs) from the fat body cells (equivalent to the mammalian liver) into the circulatory system composed of hemolymph (Diamond et al., 2009; Hetru 2009). Cellular immunity, on the other hand, is mediated by hemocytes involved in phagocytosis and encapsulation of invading pathogens (Honti et al., 2014; Hultmark 2003b; Zettervall et al., 2004). Upon infection

from Gram-positive bacteria or fungi, the Toll pathway is activated, which leads to the production of principle anti-fungal target gene called *Drosomycin* (Manfruelli 1999; Michel et al., 2001a).

So far, nine Toll receptors (Toll-1 to Toll-9) have been discovered in *Drosophila* (Tauszig S 2000). Toll-1 was the first receptor identified to be involved in the production of AMPs. All Toll receptors share a common molecular structure with a conserved ectodomain composed of leucin rich repeats (LRRs) flanked by cysteine rich motifs (Parthier et al., 2014). However, the arrangement of cysteine motifs differs between *Drosophila* and mammalian Tolls. This ectodomain is followed by a transmembrane region, which is tailed by a cytosolic Toll-Interlukin-1R homology (TIR) domain. TIR domain is the most evolutionary conserved region of the receptor and is shared by other molecules involved in innate immune signaling, such as MYD88, IL-1 and IL-18 (Servane Tauszig 2000).

The activation of Toll receptor is mediated by binding of active spätzle to the ectodomain. There are six spz in *Drosophila* (spz-1 to spz-6). Spätzle is a member of the neurotrophin family of growth factors (Lewis et al., 2013; Yvonne Delotto 1998). All spätzle proteins are synthesized and secreted as inactive proproteins with an N-terminal prodomain and an active C-terminal cysteine knot region (Arnot et al., 2010a; Yvonne Delotto 1998). During dorsoventral patterning, spz is activated by serine protease cascade (Chasan 1992; Takeuchi et al., 2001). However, during microbial infection, spz-processing enzyme (SPE) cleaves the inactive prodomain, which mask the active C-terminal domain and exposes the region needed for Toll binding (Jang et al., 2006). Two spz dimers bind to the two Toll ectodomains and trigger the

intracellular signaling that ultimately leads to the nuclear translocation of NF-κB transcription factors, Dorsal and DIF (Dorsal related immunity factor), to activate AMP genes (Gangloff 2008).

Drosophila Toll has been well known for its function in development and immunity. Toll-3 and Toll-4 are expressed in low levels, where as Toll-6, Toll-7 and Toll-8 are expressed in high levels during embryogenesis, suggesting that they may also function in development. Toll-5 and Toll-9 have been found to activate *drosomycin* promoter in transfected cells (Servane Tauszig 2000), (James Y. Ooi 2002). Recently, Toll-6 and Toll-7 have been discovered to be *Drosophila* neurotrophin receptors in central nervous system (CNS) that can genetically interact with DNT1 (spz-2) and DNT2 (spz-5), respectively (Graham Mcilroy 2013). It was reported that Toll-7 alone functions as virus recognition receptor that activates autophagy in *Drosophila* (Margaret Nakamoto 2012), and Toll-7 can also limit Rift Valley Fever Virus (RVFV) replication and mortality through activation of autophagy (Moy et al., 2014). However, a recent article published contradictory result showing, Toll-7 does not recognize VSV but autophagy and phagocytosis is required for anti-viral immunity in insects (Lamiable et al., 2016a). Nevertheless, no other spz have been studied so far that could bind with Toll receptors and play roles in *Drosophila* innate immunity.

In this research, we aim to investigate other spätzle proteins that can interact with any Toll receptors to activate AMP genes. Based on our findings, we proposed the existence of multiple Toll-Spätzle pathways in *Drosophila* host defense against microbial pathogens.

Materials and Methods

Gene Cloning

All nine Toll-TIR domains, full-length Toll -1, Toll-7 and ectodomains of Toll-1 and Toll-7 were PCR amplified using forward and reverse primers listed in Table-1 and cloned into pMT/BiP/V5-His A vector (V413020, Invitrogen) having a C-terminal V5 tag. Active C-terminal domain of spaetzles was cloned into recombinant pMT/Bip-A vector with N-terminal Flag tag. PCR reactions were performed with following conditions: 94°C for 3 min, 35 cycles of 94°C for 30s, Tm-5°C for 30s, 72°C for 30s to 4min, followed by a final extension at 72°C for 10min. The PCR products were recovered by agarose gel electrophoresis-Wizard® SV Gel and PCR Clean-Up System (A9285, Promega) and subcloned into T-Easy vectors (A1360, Promega). Plasmid DNAs in T-vectors were purified using PureYield™ Plasmid Miniprep System (A1222, Promega) according to the manufacturer's instructions and digested with respective restriction enzymes, and DNA fragments were recovered and inserted into digested pMT/BiP/V5-His A vector (V413020, Invitrogen) using T4 DNA ligase (M0202L, NEB). Recombinant expression vectors were then purified and sequenced at University of Missouri-Columbia sequencing facility for further experiments.

Dual Luciferase Assay

Dual luciferase assay was performed as described in the previous publication (Zhong et al., 2017). Briefly, S2 cells were plated in 24-well culture plates (3×10^5 cells/well) overnight in normal growth media. These S2 cells were then transiently co-transfected with recombinant pMT/BiP/V5-His A expression plasmid (500ng), pGL3B,

pGL3B-drosomycin, *pGL3B-diptericin* firefly luciferase reporter plasmid (250ng) and renilla luciferase reporter plasmid (25ng) (as an internal standard) (pRL-TK, Promega) with lipid carrier (Gencarrier-2, Epoch Life Sc.) after washing with serum free media. After overnight transfection, serum-free medium was replaced with complete growth medium containing 500μM copper sulfate (final concentration) for protein expression, and firefly luciferase and renilla luciferase activities were measured at 36 h after protein expression using the Dual-Luciferase Reporter Assay System (E1980, Promega) in the GloMax® Multi Microplate Luminometer (Promega). Relative luciferase activity (RLA) was obtained as the ratio of firefly luciferase activity to renilla luciferase activity. RLA from S2 cells cotransfected with empty pMT/BiP/V5-His A and pGL3B (empty reporter vector) plasmids were used as the calibrator. These experiments were repeated at least three times (three independent biological samples, or three independent cell cultures), and a representative set of data was used to make figures.

Cell Culture and Establishment Stable S2 of Cell Lines

D. melanogaster Schneider S2 cells were maintained at 27°C in Insect Cell Culture Media (SH30610.02, Hyclone), supplemented with 10% heat-inactivated fetal bovine serum (#10082063, Invitrogen) containing 1% penicillin-streptomycin solution (G6784, Sigma-Aldrich). Transfection procedure and stable S2 cell line were established following our previous publication (Zhong et al., 2012a). Concisely, for DNA transfection, cells were seeded overnight in normal growth medium and washed with serum free medium (SH30278.01, Hyclone) just before transfection. GenCarrier-1™ transfection reagent (#31-00110, Epoch Biolabs) was used for transient transfection based on the manufacturer's instructions. Briefly, cells in culture dishes or plates were

grown to 70-80% confluency prior to transfection. DES®–Inducible/Secreted Kit with pCoBlast (K5130-01, Invitrogen) was used to establish stable S2 cell lines. To select stable S2 cells expressing recombinant proteins, pCoBlast (Invitrogen) was co-transfected with recombinant pMT/BiP/V5-His A vectors. After 48h transfection, S2 cells were centrifuged and resuspended in complete growth medium containing 25µg/ml Blasticidine S hydrochloride (No.15205, Sigma-Aldrich). Resistant colonies appeared 1week later.

Fly Stocks and Viral Titer Maintenance

Toll-7^{g1-5}/CyO gift from (Yagi et al., 2010), Toll^{1-RXA}, Tlr^{632} gift from (Lionakis et al., 2012), *w*^{1118} (gift from Dobens), whereas, Df(2R)BSC22 lines were bought from Bloomington Stock center, Indiana. All flies were cultured on corn-meal diet and transferred to fresh food at least 24 h prior to injection/infection. 5-7 days old flies in a batch of 20-30 adult male and females were infected with gram positive *Enterococcus faecalis V583*, gram negative *Pseudomonas aeruginosa PA-14, Candida albicans*, and *Vesicular Stomatitis Virus Indiana strain* (*VSV).* Briefly, the overnight bacterial culture was diluted to OD_{600}= 0.5, washed with 1X PBS and resuspended in PBS for injection. Adult males and females were anesthetized with CO_2 (for no longer than 15min at a time). 50 nl of resuspended bacteria and 10,000 pfu/50nl of VSV per fly were injected using Drummond nano injectors and pulled glass capillary needles into the left intra-thoracic region. Following injection, flies were maintained in clean bottles with fresh cornmeal diet that was changed every day throughout the course of the experiment. Flies dying within 3 h of infection were excluded from the study due to death by injury. Flies were monitored every hour.

Vesicular stomatitis virus, Indiana strain with GFP inserted between 3' leader and N gene was a kind gift from Whelan lab (Harvard University). VSV-GFP was cultured and maintained in HEK293 cells in DMEM medium with 10% FBS, 1% pen-strep. Viral titer was measured by plaque assay done in HEK293 cells. For infection assay in the S2 cells 10,000 pfu/ml was used and for infection assay, in adult male/ female flies 10,000 pfu/50nl were used.

Western Blot Analysis

For Western blot analysis, copper sulfate (final concentration of 250μM) was added to the stable S2 cell lines (5×10^6cells/well) in 6-well plates, and protein expression was induced for 48h. Cells and culture medium (2 ml each) were collected, and stable S2 cells were homogenized in 400 μl lysis buffer [50 mM Tris-HCl, pH 7.4, 150 mM NaCl, 5 mM EDTA, 1% NP-40, 0.5mM PMSF, protease inhibitor cocktail (P8340, Sigma-Aldrich)] following our previously published protocol (Yi et al., 2013). The cell homogenates were incubated on ice for 15 min and sonicated briefly several times, and then centrifuged at 12,000 g for 15 min at 4°C. The supernatants were collected for western blot analysis. The cell culture media (10μl each from 2 ml total) and cell lysate (10μl each from 400 μl total, equivalent to ~5×10^4 cells) were separated on 8%, 12%, or 15% SDS-PAGE and proteins were transferred onto nitrocellulose membranes (162-0097, Bio-Rad). The membrane was blocked with 5% BSA in Tris-buffered saline (25mM Tris-HCl, pH7.6, 150 mM NaCl) containing 0.05% Tween-20 (TBS-T) at room temperature for at least 3h and then incubated overnight with primary antibody at 4°C in 5% BSA in TBS-T with gentle rocking. Then, the membrane was washed four times with TBS-T and incubated with secondary antibody in 5% BSA in

TBS-T for 2h at room temperature. After washing three times with TBS-T (10min each time), the signal was developed by using alkaline phosphatase (AP)-conjugate color development Kit (#170-6432, Bio-Rad). Anti-Flag M2 antibody (F-1804, Sigma-Aldrich, 1:5000 dilution) and anti-V5 antibody (V-8012, Sigma-Aldrich, 1:5000 dilution) were used as primary antibodies, alkaline phosphatase-conjugated anti-mouse antibody (A4312, Sigma-Aldrich, 1:10,000) was used as secondary antibody for color development.

Co-immunoprecipitation (Co-IP) Assay

Immunoprecipitation (Co-IP) assay was performed following our previously published protocols (Zhong et al., 2016a) by using 300μl of cell extract, which is equivalent to approximately 10^6 cells, or equivalent cell culture medium containing recombinant proteins. The cell extracts or cell culture media were precleared for 1 hour with 20μl of Protein G Sepharose (50% slurry, No.17-0618-01, GE Healthcare) in a total volume of 1ml. After centrifugation, the supernatant was incubated with 2 μl of anti-Flag M2 or anti-V5 antibody (final concentration of 1μg/μl) at 4°C for 10h with gentle rocking. Then, 30μl Protein G Sepharose (50% slurry) in lysis buffer was added to the protein-antibody mixture and incubated at 4°C overnight with gentle rocking. The Sepharose beads containing immunoprecipitated proteins were collected after centrifugation, washed three times with lysis buffer, resuspended in 50μl of 1×SDS sample buffer, boiled at 95°C for 5 min, and used for subsequent immunoblotting analysis.

RNA Extraction and qRT-PCR

To quantitate transcript levels RNA were extracted and target gene expressions were determined following protocols describes in our previous papers (Rao et al., 2014b; Xu et al., 2012). Flies were anesthetized on CO_2 bed, placed in 1.5ml tubes and homogenized with disposable pestles in 1ml of TRIzol® Reagent (T9424, Sigma-Aldrich). Homogenates were transferred to fresh tubes. 300µl of chloroform was added and the mixture was shaken for 15s. The samples were then centrifuged at 12000 g for 10min at 4^0C, and the aqueous phase was transferred to clean 1.5ml tubes. 500 µl of isopropanol was added and RNA was precipitated at room temperature for 10min then centrifuged at 12000 g for 10min. The RNA pellet was washed with 1ml of 70% ethanol and samples were centrifuged at 7.4 g for 5min. The 70% ethanol was decanted and the pellet was allowed to air dry and resuspended in 50µl of nuclease-free water. The RNA quality was determined by nano drop UV-Vis spectrophotometer (ND-1000, Thermo). For subsequent experiments, 2 µg of RNA was used for all samples.

For making cDNA, residual genomic DNA was digested with RQ1 RNase-free DNase (M6101, Promega). cDNA was prepared from 2µg total RNA in a 25µl reaction using moloney murine leukemia virus (M-MLV) reverse transcriptase (M1701, Promega) an anchor-oligo(dT)18 primer following the manufacturer's instructions. Each cDNA sample (diluted 1:50) was used as template for quantitative real-time PCR analysis. The *Drosophila ribosomal protein 49* (*rp49*) gene was used as an internal standard to normalize the amount of RNA template. The primer pairs (Table 2) were designed based on the sequences of *rp49*, *drosomycin* and *diptericin*. The real-time PCR was performed in 20µl reactions containing 10µl 2×SYBR® GreenER™ qPCR SuperMix Universal (No. 204141, Qiagen), 4µl H_2O, 4µl diluted cDNA template, and

1μl (10pmol) each of the forward and reverse primers. Real-time PCR program was 2 min at 50°C, 10 min at 95°C, followed by 40 cycles of 95°C for 15s, 60°C for 1 min and the dissociation curve analysis. Data from three replicas of each sample were analyzed by the ABI 7500 SDS software (Applied Biosystems) using a comparative method ($2^{-\Delta\Delta CT}$). The baseline was set automatically by the software to maintain the consistency. cDNA sample from S2 cells transfected with empty pMT/BiP/V5-His A plasmid or wild type flies (w^{1118}) were used as the calibrator control. The expression levels of *drosomycin* and *diptericin* transcripts in other cDNA samples were calculated by the $2^{-\Delta\Delta CT}$ method (Livak et al., 2001), which stands for the n-fold difference in relative expression to the calibrator. All the data were presented as relative mRNA expression. All these experiments were repeated at least three times.

Data Analysis

All the experiments were performed in 3-4 replicates and repeated with three to four independent biological samples. The means of a typical set of data were used to prepare the figures by GraphPad Prism (GraphPad, San Diego, CA). Statistical significance was calculated by one way ANOVA followed by a Tukey's multiple comparison tests using GraphPad Prism and identical letters are not significant difference ($p>0.05$) while different letters indicate significant difference ($p<0.05$). The significance of difference was also determined by an unpaired t-test with the GraphpadInStat software (*, $p<0.05$; **, $p<0.01$).

Genes for cloning	Primers (5'-3')
DmToll-1-KpnI-N	ATGGTACCAATGAGTCGACTAAAGGCC
DmToll-1-ApaI-C	ATGGGCCCTACGTCGCTCTGTTTGGC
DmToll-7-KpnI-N	CGGGGTACCAATGGCGGCAATCCTGCTGCT
DmToll-7-NotI-C	GAATGCGGCCGCTTCACCAGATACGCCTGAACAT
DmToll-1^{ecto}-ApaI-C	ATGGGCCCGAACACGCCCTTTTCCGCCGG
DmToll-7^{ecto}-NotI-C	GAATGCGGCCGCTTATACGATTCTGGGATACCATGCT
MsTIR-KpnI-N	ATGGTACCACCGTACGACGCGTTTGTGTCTTTCGCAC A
MsTIR-ApaI-C	ATGGGCCCTTTGTAGCAAGGACTCGCGCCCGGCGCTG G
DmTIR-1-KpnI-N	ATGGTACCAAAGTTCGATGCCTTCATCTCG
DmTIR-1-ApaI-C	ATGGGCCCTACGTCGCTCTGTTTGGCA
DmTIR-2-SpeI-N	CTGGACTAGTATGATCATCCTGCACTCGGAGAA
DmTIR-2-NotI-C	GAATGCGGCCGCTTGACCAGGAAAGCTTGGCCGTT
DmTIR-3-SpeI-N	CTGGACTAGTATGAGGTTCGATGCCTTTCTGGC
DmTIR-3-NotI-C	GAATGCGGCCGCTTGTCAACGTAGCTTGGTAGTAG
DmTIR-4-SpeI-N	CTGGACTAGTATGAAATACGATGCATTCCTATC
DmTIR-4-ApaI-C	AATGGGCCCTACCTTTGTTTCTGCATCTGA
DmTIR-5-SpeI-N	CTGGACTAGTATGACCTACGATGCCTTCATCTC
DmTIR-5-NotI-C	GAATGCGGCCGCTTGATTAGCGGCCCCGCATGCTT
DmTIR-6-SpeI-N	CTGGACTAGTATGGATGCCTACTTCGCCTACAG
DmTIR-6-NotI-C	GAATGCGGCCGCTTCGCCCACAGGTTCTTCTGCTGA
DmTIR-7-SpeI-N	CTGGACTAGTATGGTGCTCCTGCATTCCGCCAA
DmTIR-7-NotI-C	GAATGCGGCCGCTTCACCAGATACGCCTGAACAT
DmTIR-8-SpeI-N	CTGGACTAGTATGTTCGACGCCTTCGTTTCGTA
DmTIR-8-NotI-C	GAATGCGGCCGCTTCATGTGCAGATTTCTAGACG
DmTIR-9-SpeI-N	CTGGACTAGTATGTTCATCAGCTACTGCCAGAA
DmTIR-9-ApaI-C	AATGGGCCCAACACTGATCTCTCTGGAGT
Dmspz-1-KpnI-N	ATGGTACCCATGGACTACAAGGACGACGATGACAAGG CGGCCGCTGTTGGTGGCTCAGACGA
Dmspz-1-PmeI-C	CGGTTTAAACTCACCCAGTCTTCAACGC
Dmspz-2-KpnI-N	GAATGCGGCCGCTCTCGATGCCTGCGAGTCGAAG
Dmspz-2-PmeI-C	CGGTTTAAACCTAGCGATAACCATCCACTTGGC
Dmspz-3-KpnI-N	GAATGCGGCCGCTCTCGATGCCTGCGAGTCGAAG
Dmspz-3-PmeI-C	CGGTTTAAACCTAGGGATTACATCTACAGA
Dmspz-4-KpnI-N	GAATGCGGCCGCTGGAGTAAATGCCTGTCCCGT
Dmspz-4-PmeI-C	CTTTTTAAACTTAGTCCTCCAAGAAATCGA
Dmspz-5-KpnI-N	GAATGCGGCCGCTCAAAGTCCGGGGCGCTCCAC
Dmspz-5-PmeI-C	CGGTTTAAACTTAATTGGCGGCTATCGTGC
Dmspz-6-KpnI-N	GAATGCGGCCGCTTGTCACTACCTGGACGGCGG
Dmspz-6-KpnI-C	CGGTTTAAACTCACAACTCGGCCACCGACT
pGL3B-Dpt-N	GGGGTACCAGTAACTTTACTGATAAGACTTGGATTCT C
pGL3B-Dpt-C	GAAGATCTCTCAGTTGTTCTCAATTGAAGAACTG
pGL3B-Drs-N	GGGGTACCCAATGAAAGTGATAATACGAATTGACC
pGL3B-Drs-C	GAAGATCTATTGGAAAAGGTTCTCACGGAGC
Genes for qPCR	**Primers (5'-3')**
DmDpt-N	ATGACCATGAAGCCCACTCC

DmDpt-C	ATTCAGTCCAATCTCCGGGC
DmDrs-N	TACTTGTTCGCCCTCTTCGC
DmDrs-C	GGAGCGTCCCTCCTCCTTGC
DmRP49-N	GCCCAAGGGTATCGACAACA
DmRP49-C	ACCTCCAGCTCGCGCACGTT
DmToll-1-N	TCCAGACCCAGATCAACTCC
DmToll-1-C	TAGCCCAGCGAGCTAATGTT
DmToll-7-N	AGTTCGAGTGCGAGTGCC
DmToll-7-C	TTGCATTGTTCGCTGGCG

Table 1. Primers used for Toll, Spätzle Gene Cloning and qPCR. N represents forward primer and C represents reverse primer.

Result and analysis

Toll-4 and Toll-7 TIR Domain could Activate *Drosomycin* Promoter in S2 Cells

In *Drosophila* Toll-Spätzle pathway is well known to activate the *Drosomycin* gene. However, the role of other Toll receptor in innate immunity has not been well studied. Toll receptors dimerize to activate the intracellular signals where the TIR domain binds with MYD88 adaptor protein to initiate the intracellular cascade. Since, there are nine toll receptors in *Drosophila* we cloned the TIR domain of all nine toll receptors in pMT/BiP/V5-His A vector and expressed the protein in S2 cells (Fig. 1 A and B). For detecting the AMP gene promoter activity, all nine TIR were transfected into S2 cells along with either *drosomycin* or *diptericin* promoter (Fig. 1 C). Overexpression of Toll TIR domain in S2 cells is sufficient to form homo or hetero dimers, which then binds with MYD88 adaptor proteins to activate the intracellular signals. Dual luciferase activity suggests that, over expression of DmToll-4 TIR and Dm Toll-7 TIR domain activate *drosomycin* promoter along with DmToll-1 and MsToll.

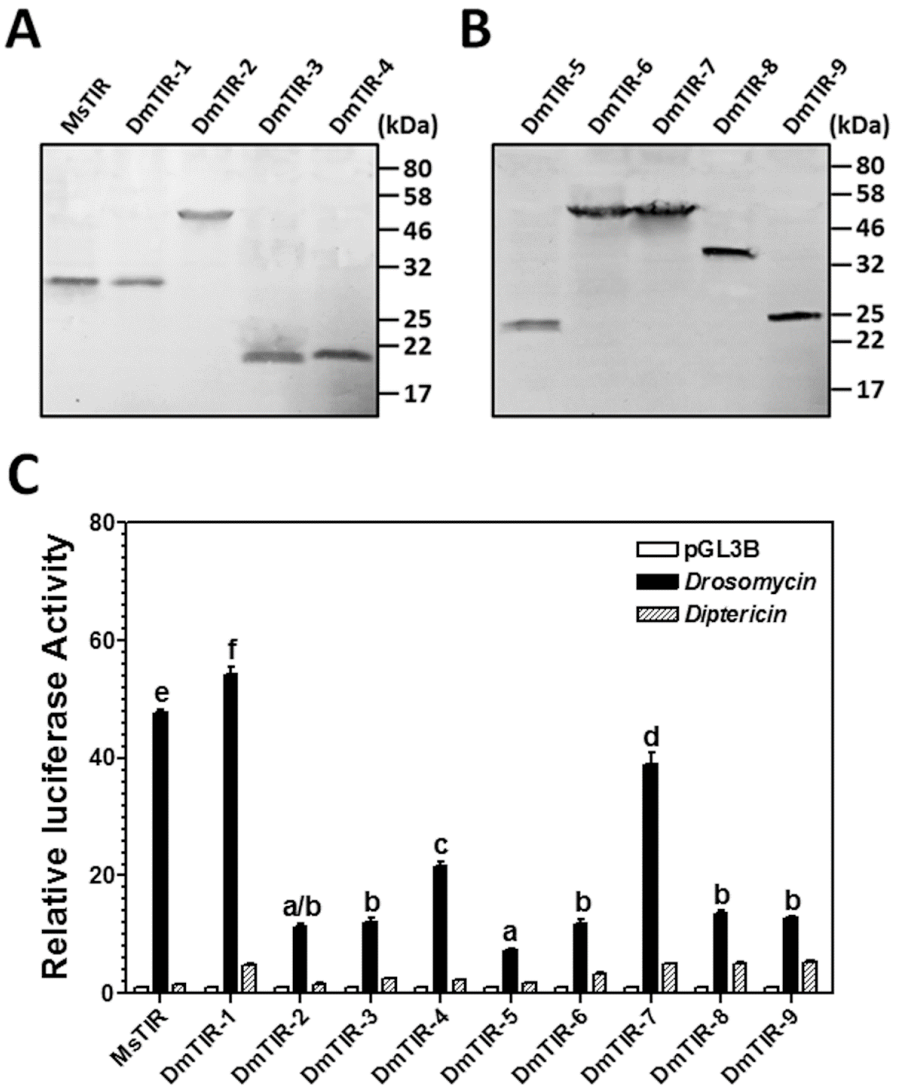


Fig. 1. Western Blot Analysis of Cloned and Expressed Toll TIR Domains.

(A) and (B) are Toll TIR domain expression in S2 cells stained with anti-V5 antibody. MsTIR and DmTIR-1 is used as a positive control. (C) The relative luciferase activities of the Toll-TIR domains and *drosomycin* or *dipterecin* promoters expressed in S2 cells were determined by Dual-Luciferase® Reporter Assay System as described in the Materials and Methods. Bars represent the mean of three independent measurements ± SEM. For the activity among different promoters activated by transcription factors (comparing solid bars by drosomycin promoters, stripe bars by *dipterecin* across the promoters, empty white box as pGL3B), identical letters (capital letters for solid bars and small letters for stripe bars) or identical numerical numbers (dotted bars) are not significant difference ($p > 0.05$) while different letters or different numerical numbers indicate significant difference ($p < 0.05$). The significance of difference was also determined by an unpaired t-test (*$p < 0.05$; **$p < 0.01$), and "n" indicates not significant.

Interaction of Toll Receptors with Spätzles in S2 cells

Upon microbial infection spätzle processing enzyme cleave the full length spätzles into C'-terminal 106-residues long active spätzle, which then dimerize and binds to the ecto-domain of the Toll receptor to activate the intracellular signal. Since drosophila has 6 spätzles and 9 toll receptors, we wanted to check the interaction between Toll-1 and Toll-7 with all six active spätzles. For this, we cloned the Toll-1 and Toll-7 ecto-domain into pMT/Bip/V5-His-A vector with C' terminal V5 tag and all six active spätzles into recombinant pMT/Bip/Flag-A vector with N' terminal FLAG tag. Proteins were over expressed in S2 cells for doing co-immunoprecipitation (Co-IP) assay. Co-IP of Toll receptor ecto-domain with six active Spätzles shows that Toll-1 could interact with spz-1, spz-2 and spz-5 (Fig. 2 A to D) but not with spz-3, -4 and -6 whereas; Toll-7 could interact with spz-1, spz-2, spz-5, and spz-6 but not with spz-3, and -4 (Fig. 2 E to H). Spätzle-1, -2 and -5 are phylogenetically more closely related than spätzle-3, -4 and -6 and thus could have similar functions.

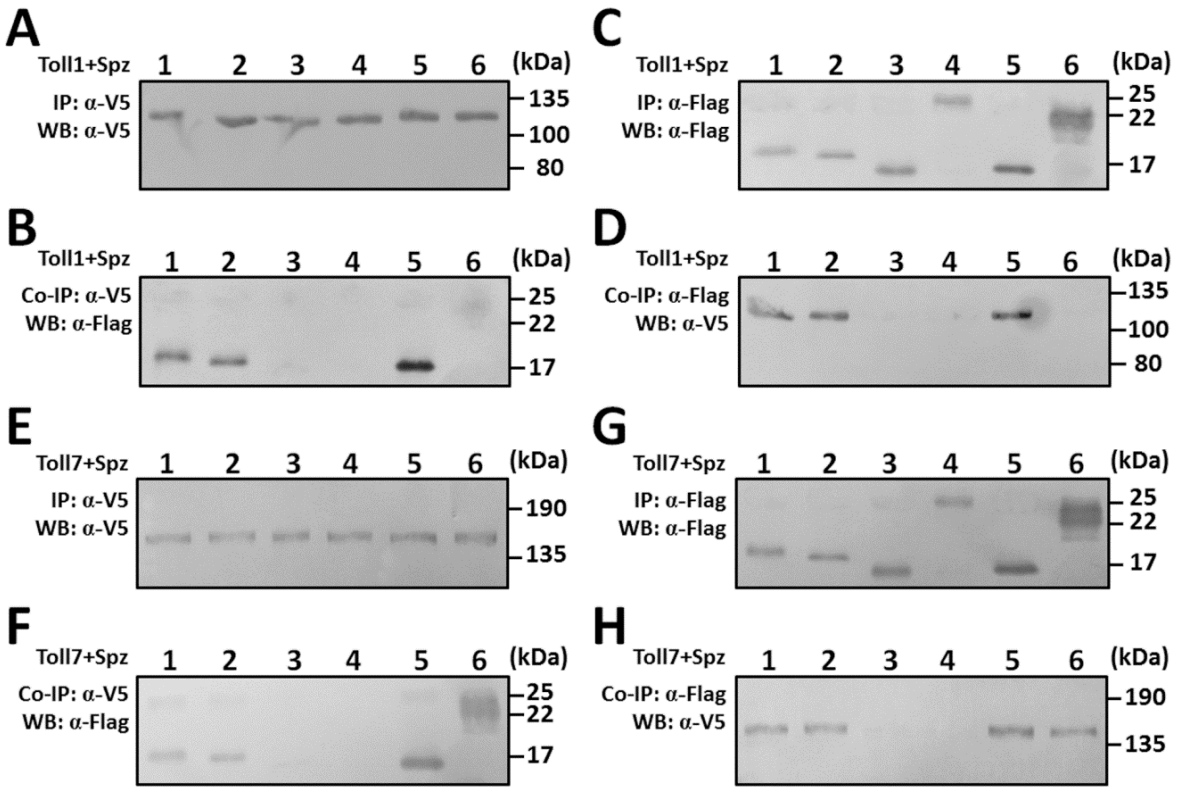


Figure 2. Interaction of *Drosophila* Toll-1 and Toll-7 with Six Active Spätzles.

Recombinant V5-tagged DmToll-1 and DmToll-7 ecto domains, and Flag-tagged six Dm spätzles were expressed in S2 cells separately, and cell lysates from two samples were mixed for co-immunoprecipitation (Co-IP) assays as described in the Materials and Methods. Immunoprecipitated (IP) proteins or Co-IP proteins were detected by immunoblotting using anti-Flag or anti-V5 monoclonal primary antibody. Lanes 1–6 were S2 cells overexpressing V5-tagged Toll-1 or Toll-7 ecto domain in media and S2 cell lysate and media by overexpression of Flag-tagged Dm Spatzles-1 to -6. A,C, E, G are IP and B,D,F,H are Co-IP proteins. V5-tagged Dm Toll was co-immunoprecipitated with Flag-tagged Dm Spz (B, D, F & H).

Function of Toll-1, Toll-7 and Six Spätzles in Innate Immunity

Based on our co-immunoprecipitation assay, we tried to look for the function of these pairs in activation of AMP genes. We cloned *Drosophila* full length Toll-1 and Toll-7, six active spätzles, and *Drosomycin* and *Diptercin* promoters and expressed them in S2 cells. Relative luciferase activity shows that Toll-1 could activate *Drosomycin* promoter to the highest level when co-transfected with spz-1, -2 and -5 but not with spz-3, -4 and -6, (Fig. 3 A) whereas, Toll-7 could activate *Drosomycin* promoter when co-transfected with spz-1, -2, -5 but not with spz-3, -4 or -6 (Fig. 3 B). This could be explained by the fact that spz-1, -2, and -5 are phylogenetically more closely related than Spz-3, -4, and -6. All this data showed for the first time that multiple Toll-Spätzle pathways exist in *Drosophila* and could play role in innate immunity. To further confirm that Toll-1 and Toll-7 could activate AMP gene only when co-transfected with active spätzles, we overexpressed only receptor, or only ligand and compared the *Drosomycin* promoter activity with Toll-spätzle pairs when co-transfected together. Our result suggests that only receptor and only ligand cannot activate *Drosomycin* gene promoter (Fig. 3 C) and both receptor and ligand pair interaction is needed for activation of intracellular signal.

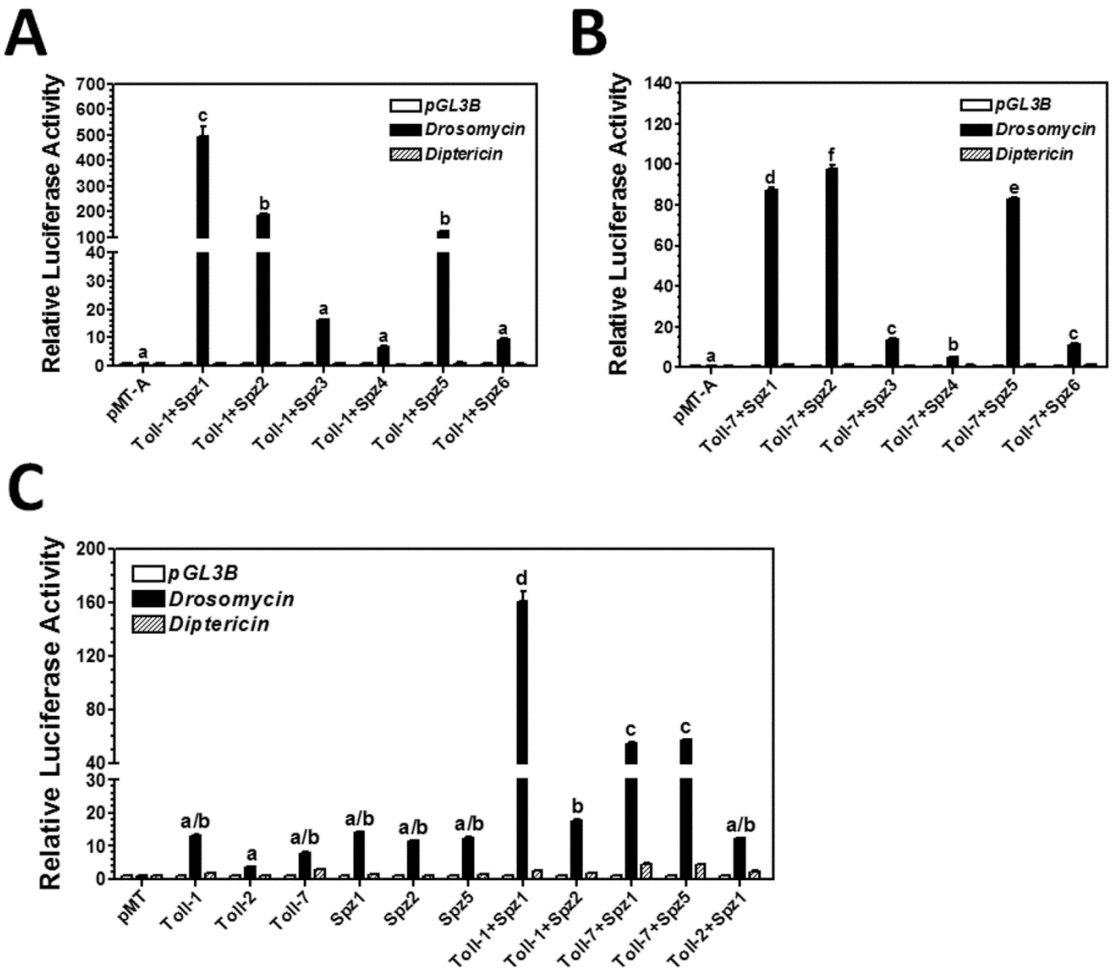


Figure 3. Activation of AMP Gene Promoters by *Drosophila* Toll-1 or Toll-7 and Six Active Spätzles.

(A) The relative luciferase activities of Toll-1 full length and six spätzles co-expressed with either *drosomycin* or *diptericin* promoters (B). The relative luciferase activities of Toll-7 full length and six active spätzles co-expressed with either *drosomycin* or *diptericin* promoters. All proteins are expressed in S2 cells were determined by Dual-Luciferase® Reporter Assay System as described in the Materials and Methods. Bars represent the mean of three independent measurements ± SEM. For relative luciferase activity (A-B) among different promoters activated by Toll- Spätzle pairs (comparing striped bars or solid bars) or (C) comparing receptor only or ligand only activation of AMP promoters compared with Toll-Spätzle pairs when co-transfected. Identical letters are not significant difference ($p > 0.05$) while different letters indicate significant difference ($p < 0.05$).The significance of difference was also determined by an unpaired t-test (*$p < 0.05$; **$p < 0.01$).

VSV-G Protein Interacts with Toll Receptor and could Activate AMP Genes

Recently, a paper published showing that Toll-7 confer resistance to VSV infection and Toll-7 mutant flies are more susceptible to VSV infection (Nakamoto et al., 2012). Later, in 2016 another paper was published suggesting that VSV recognition does not depends on Toll-7 receptor and autophagy does not play a major role in antiviral immunity in *Drosophila* (Lamiable et al., 2016a). Due to this controversial data, we decided to test if VSV could interact with Toll receptor and if it could activate AMP genes. Firstly, we overexpressed Toll-1 and Toll-7 ectodomain with the V5 tag in S2 cells and collected the media with expressed proteins to do the Co-IP. DMEM media with VSV were collected and mixed with S2 media with expressed Toll proteins. This mixture was pulled down with anti-V5 antibody and stained with either anti-V5 or anti-VSV-G antibody. The Co-IP assay shows that Toll-1 and Toll-7 both could interact with VSV-G proteins (Fig. 4 A-C). Next, we wanted to test if VSV infection in S2 cells could activate AMP genes. For this, we used cells lines expressing full-length Toll-1 and Toll-7, transfected them with *Drosomycin* or *Attacin* or *Diptericin* promoters. After 48 hrs of protein expression, S2 cells were infected with 10,000 pfu/ml of VSV and incubated for another 24 hrs. Dual luciferase assay shows that both Toll-1 and Toll-7 cell lines infected with VSV could mainly activate *Drosomycin* and *Attacin* promoters (Fig. 4 E and D). However, the overall induction was low as predicted because AMPs do not confer adequate resistance against VSV. However, we did see some activation of AMPs during VSV infection, which could be explained by body's natural process to combat infection by activating the major immune pathway but this is definitely not the key mechanism to combat viral infection.

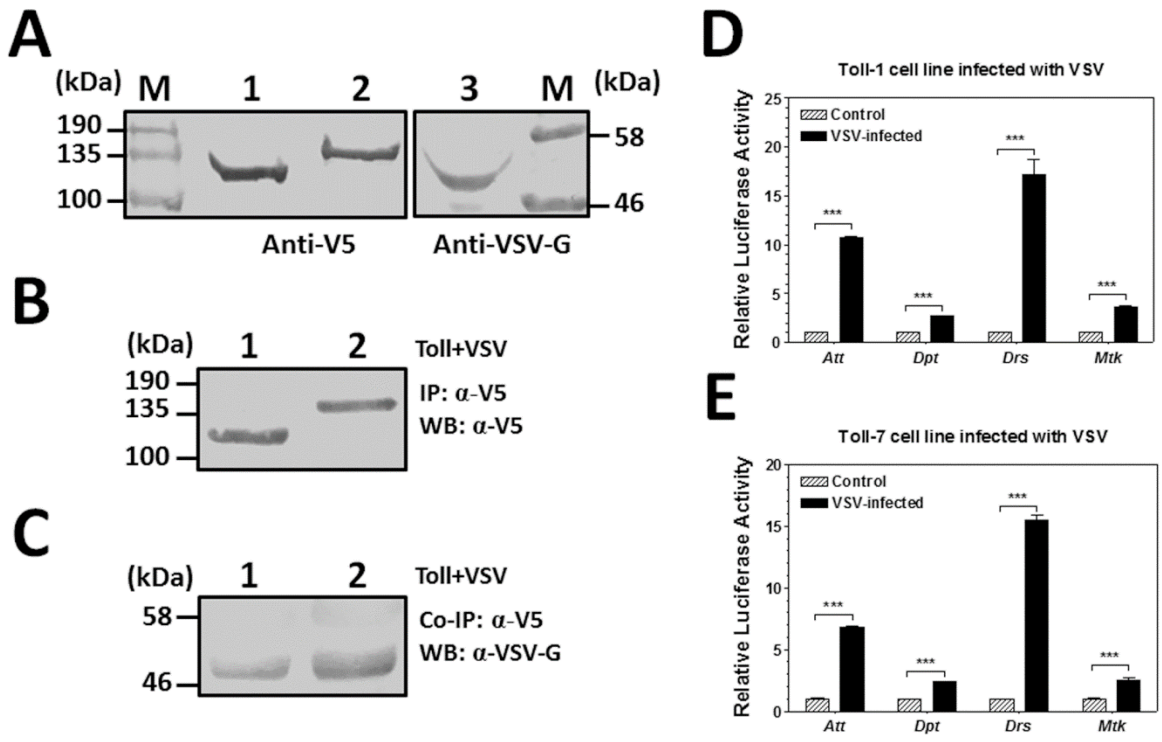


Figure 4. Interaction of *Drosophila* Toll-1 or Toll-7 Ectodomain with VSV-G Protein and Activation of AMP genes.

(A) Western blot of recombinant V5-tagged Toll-1 ectodomain (Lane-1) and Toll-7 ectodomain (Lane-2) expressed in S2 cells separately; cell media from two samples and VSV-G proteins in media (Lane-3) were assayed for protein expression. (B) Toll-1 ectodomain media (Lane-1) and Toll-7 ectodomain media (Lane-2) were mixed with VSV-G proteins expressed in media and were immunoprecipitated (IP) using the anti-VSV-G antibody as the primary antibody, which was later detected by immunoblotting. (C) Toll-1 ectodomain media (Lane-1) and Toll-7 ectodomain media (Lane-2) mixed with VSV-G proteins expressed in media were Co-immunoprecipitated (Co-IP) by the anti-V5 antibody as the primary antibody and detected by immunoblotting using anti-VSV-G antibody. (D-E) The relative luciferase activities of the AMP promoters activated by VSV. Recombinant full-length *Dm*Toll-1 or *Dm*Toll-7 were co-expressed with AMP gene promoters in S2 cells and later infected by VSV. These promoter activities were determined by Dual-Luciferase® Reporter Assay System as described in the Materials and Methods. Bars represent the mean of three independent measurements ± SEM. Solid bars represent S2 cells infected by VSV whereas, stripped bars represents non-infected cells. Identical letters (capital letters for solid bars and small letters for striped bars) or identical numerical numbers (dotted bars) are not significant difference ($p > 0.05$) while different letters or different numerical numbers indicate significant difference ($p < 0.05$). The significance of difference was also

determined by an unpaired t-test (*p < 0.05; **p < 0.01), and "n" indicates not significant.

Toll-7 Confers Resistance to *Candida albicans* Infection in Males Only

Upon analysis of the Toll-1 and Toll-7 gene expression pattern in adult males and females, it is clear that Toll-1 is expressed in both male and female but females have 10-fold higher expressions than males. On the other hand, Toll-7 is expressed in males only. During gram positive bacterial or fungal infection, both Toll-1 and Toll-7 is expressed in males but females express only Toll-1 (Fig. 6). Infection with *Candida albicans* shows that Toll-1 mutant females, mutant male and Toll-7 mutant males were more susceptible compare to Toll-7 mutant females (Fig. 5 C and G). Infection with gram-positive *Enterococcus faecalis V583* shows that Toll-1 mutants females and Toll-7 mutant males are more susceptible to infection (Fig. 5 A and E). Whereas, Toll-7 mutant males are more susceptible to *Pseudomonas aeruginosa PA14* infection; however, the mortality rate at the end of the study are similar in all the genotypes (Fig. 5 B and F). Lastly, we also challenged the Toll mutant flies to VSV and found that even with a high dose of VSV (10,000 pfu/ml) fly population has no effect on the mortality rate and all the population behave the similarly (Fig. 5 D and H). To further confirm the Toll-1 and Toll-7 expression in mutant flies, we performed qPCR to check the expression levels of *Toll-1, Toll-7*, *Drosomycin* and *Diptericin* gene. The *Toll-1* or *Toll-7* genes are expressed less in mutants and are not induced in mutants even during infection (Fig. 6 A- H). Similarly, we tested the Toll mutants for the induced expression of *Drosomycin* and *Diptericin* in and found that these mutants express low levels of

Drosomycin and *Diptericin*, which suggests that Toll pathway plays a major role in combating infection against these microbes but not during VSV infection (Fig. 7 A – H).

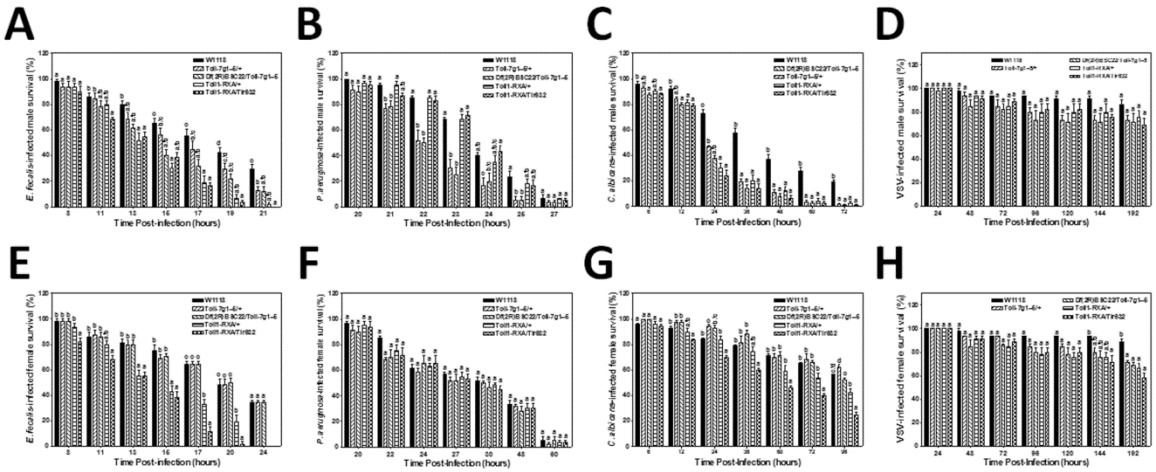


Figure 5. Toll-7 Confers Resistance to *C. albicans* Infection in Adult Males Only.

(A-H) *Toll7g1-5/Df(2R)BSC22* mutant male flies are more sensitive to *C.albicans* infection. Survival of *Toll-7*$^{g1-5}$ deletion mutant, Toll-7^{g1-5} deletion mutant genetic background (*yw; Toll-7g1-5*) were crossed with a line containing a deficiency covering the Toll-7 gene (*Df(2R)BSC22*) and flies from the progeny were selected. Similarly, *Toll1-*^{RXA} and *Tlr*^{632} flies were crossed together to get a strong phenotype of Toll-1 mutant (*Toll-1*^{RXA}*/Tlr*^{632}) and were used as positive control. *W*^{1118} served as experimental control flies. All genotype of flies were injected with either PBS or conditioned medium (CM) or microbes diluted in PBS. (A and E) Male and female flies infected with *E. fecalis*, (B and F) flies infected with *P.aeruginosa*, (C and G) infected with *C. albicans* (D and H) infected with *VSV*. Solid bar represents the control flies whereas, patterned bars are the mutant flies infected with different microbes. Data represent the mean ± SE or SD of 4 independent experiments, each containing 30 flies per group, ns not significant, * $p< 0.05$, ** $p< 0.01$, *** $p< 0.001$ using log-rank or unpaired t-test.

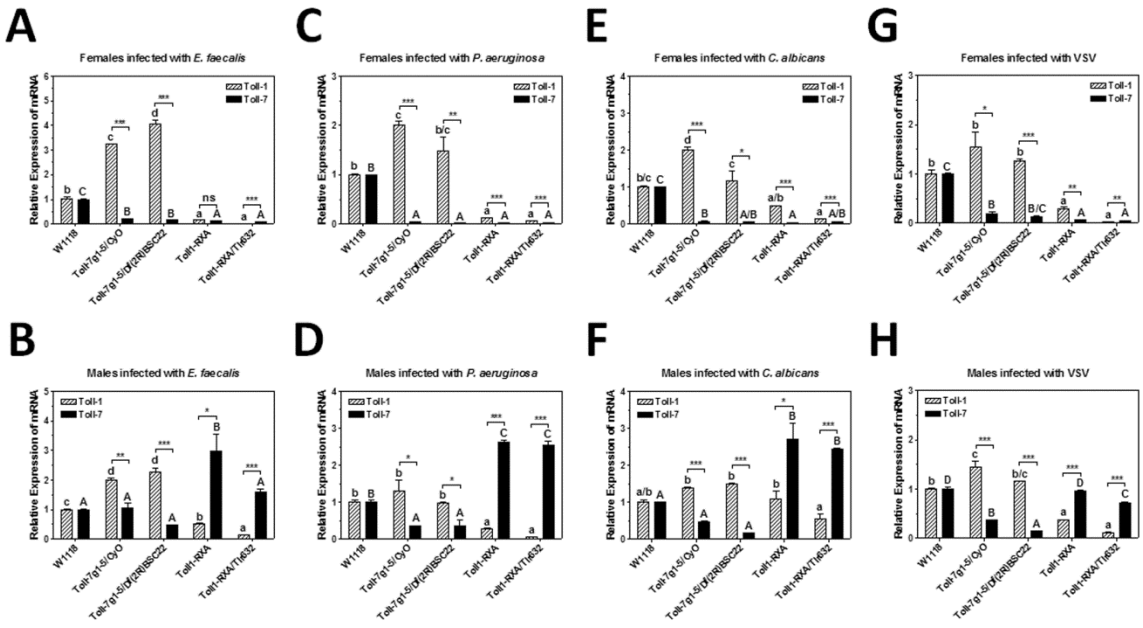


Figure 6. (A-H) Quantitative RT-PCR Analysis of Toll Genes in Toll Mutant Flies Infected by *E.fecalis*, *P.aeruginosa*, *C.albicans*, and *VSV*.

Toll-1 mutants represented by stripped bar whereas, Toll-7 mutant is represented by a solid bar. Bars represent the mean of three independent measurements ± SEM. Comparing different genotypes or different injection conditions, identical letters are not significant difference ($p>0.05$) while different letters indicate significant difference ($p<0.05$).

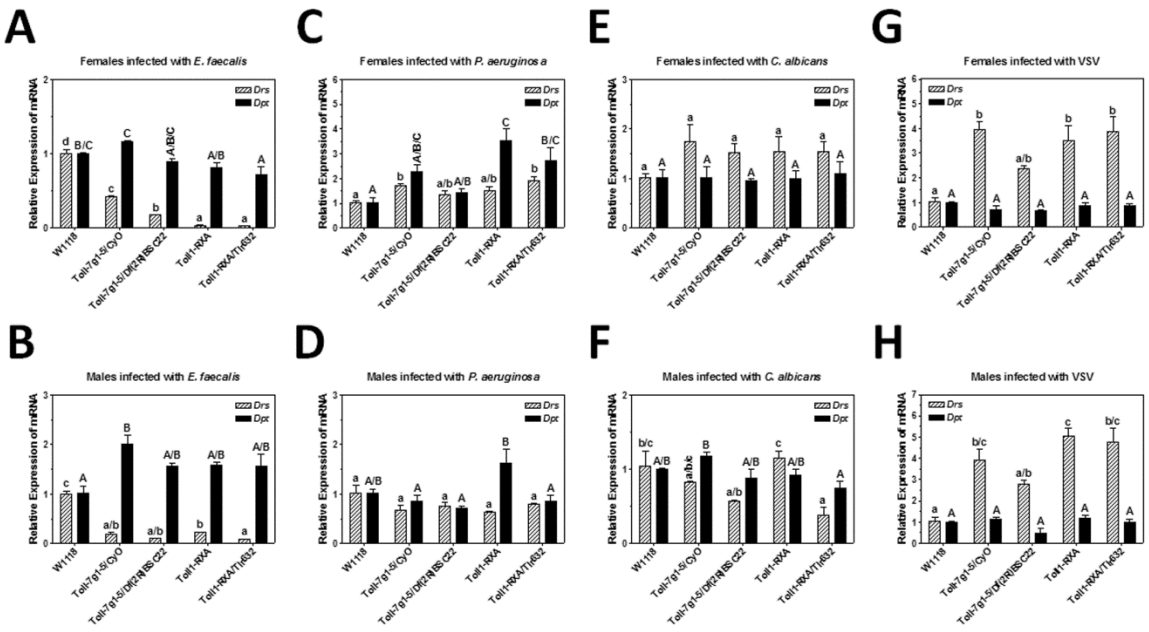


Figure 7. (A-H) Quantitative RT-PCR Analysis of *Drosomycin* and *Diptericin* genes in Toll Mutant Flies Infected by *E.fecalis*, *P.aeruginosa*, *C.albicans*, and *VSV*.

Solid bar represents *diptericin* and stripped bar is *drosomycin*. Bars represent the mean of three independent measurements ± SEM. Comparing different genotypes or different injection conditions, identical letters are not significant difference ($p>0.05$) while different letters indicate significant difference ($p<0.05$).

Discussion

Toll is a transmembrane receptor that is responsible for dorsoventral polarity during embryogenesis (Gay et al., 1992) and for triggering AMPs during infection (Takeuchi et al., 2001). Toll receptor is activated by binding with Spätzle proteins and receptor dimerization to activate the intracellular signal (Arnot et al., 2010a; Lewis et al., 2013). So far, nine Toll receptor and six spätzles have been identified in *Drosophila* and functions of few Toll receptors have been delineated. For example, Toll-5 can induce *Drosomycin* and *Metchnikowin* expression (Imler et al., 2001; Imler et al., 2000; Luo et al., 2001). Toll-9 has been reported to activate the constitutive expression of *Drosomycin* (Ooi et al., 2002), using Toll signaling pathway components (Bettencourt et al., 2004). However, the role of these six spätzle proteins as ligand activating the Toll pathway for immunity have not been well studied. To better understand the function of Toll receptor and spätzle ligand in activating the innate immune pathway we focused on studying Toll-1, Toll-7 receptors and six spätzles in *Drosophila* during microbial infection. During activation of Toll pathway, the cytoplasmic domain of the Toll receptor, known as Toll-Interlukin domain (TIR) dimerizes and binds to the adaptor protein MYD88 to activate the intracellular cascade to trigger AMP genes (Janssens et al., 2002). The cloning and overexpression of all full-length of Toll receptor were time-consuming, so we cloned the TIR domains and expressed them in S2 cells to perform the dual luciferase assay to check the activation of *drosomycin* and *diptericin* promoter reporter. The result showed that along with Toll-1, Toll-7 and Toll-4 could activate the *drosomycin* promoter reporter. Based on the high activity of Toll-7 compared to Toll-4

we selected Toll-7 for our further studies. Next, we cloned the full-length Toll-1, Toll-7, and all six active spätzles and expressed them in S2 cells to do co-immunoprecipitation and dual luciferase assays. The co-immunoprecipitation assays showed that Toll-1 and Toll-7 both could interact with spz-1, -2 and -5, whereas, Toll-7 could additionally interact with spz-6. Based on the article published in 2013 (Mcilroy et al., 2013b), these data show that spz-1, -2, -5 are closely related than spz-3, -4 and -6 and could have a similar function. Sutcliffe et al. (Sutcliffe et al., 2013) showed that spz-2 and -5 are involved in nervous system formation. However, since spz-1 is involved in innate immunity, so we wanted to verify if all six spätzles could interact with Toll-1 or Toll-7 receptors and later if the receptor-ligand pair could activate the *drosomycin* promoter. Dual luciferase assay shows that Toll-1 and Toll-7 both could activate *drosomycin* promoters when cotransfected with spz-1, -2 or -5 but not with other spz. This led us to verify further the role of Toll-7 in innate immunity. To do this, we performed several septic infection experiments with different microbial agents and found that Toll-7 males are more susceptible to *C. albicans*, and *P. aeruginosa* infection than females. Whereas, infection with *E. fecalis* shows that Toll-1 females were more susceptible than males. This could be due to the high expression of Toll-1 in females and very low expression of Toll-7 gene in males. In addition, during microbial infection, no susceptible mutant flies could express respective Toll genes and *drosomycin* gene. This also confirms our hypothesis of Toll-7 being involved against *C. albicans* infection in males only. However, VSV infection in Toll-1 or Toll-7 mutant flies had no effect on the mortality rate of the respective group. This indicates that Toll-1 or Toll-7 receptor does not participate in providing immunity against VSV infection.

Since Toll-7 and spätzle overexpression could activate the *drosomycin* promoter that is involved against fungal infection, it is more likely that Toll-7 could be involved against fungal infection rather than viral infection, which we confirmed by our experiments.

CHAPTER 3

NF-KB TRANSCRIPTION FACTORS COOPERATIVELY REGULATE ANTIMICROBIAL PEPTIDE GENES IN *DROSOPHILA MELANOGASTER*

Abstract

The innate immune response in *Drosophila* involves the inducible expression of antimicrobial peptide genes mediated by the Toll and IMD signaling pathways. Dorsal, DIF and Relish are NF-kappaB-related transcription factors and functions by forming dimers. Dorsal and DIF act downstream of Toll, whereas Relish acts downstream of IMD to regulate target gene expression. Relish was identified to be involved in the *Drosophila* immunity during infection. Unlike the single domain proteins Dorsal, DIF and Cactus, Relish contains both a Rel homology domain and an IkappaB-like domain with six ankyrin repeats. Relish is a dual domain protein with an inhibitory domain. Several studies in insects has revealed the presence of Relish short isoforms without this inhibitory domain that are involved in innate immunity. Therefore, we searched for the presence of Relish short (Rel-S) isoform in *Drosophila* by 3' RACE PCR and found a unique Relish short isoform that is expressed highly in the L2, and L3 stages and in adult males. This Relish short isoform differs from full-length relish by only 20bp at the 3' end. Relish short isoform dimerizes with DIF and Dorsal and could activate the *Drosomycin* promoter to its highest level. qPCR result suggests that during microbial infection short relish expression increases 10-fold in females suggesting its role in immunity. To activate the AMP genes, the NF-kB transcription factors must to bind with kB sites. Therefore, to determine the key bases for kB site activity we performed

several substitution mutations and found that in *Drosomycin* kB site 2, 6^{th} and 10^{th} position bases are important for its activity and in *Metchnikovin* site, position 8^{th} base is crucial for its activity. Altogether, our data reveal the existence of a new Relish short isoform and the enhancer element important for initiating the downstream gene transcription.

Introduction

In the fruit fly *Drosophila melanogaster*, innate immune-related genes, including antimicrobial peptide (AMP) genes, are regulated mainly by the Toll and immune deficiency (IMD) pathways. The Toll pathway is activated by Gram-positive bacteria, viruses and fungi (De Gregorio et al., 2001; Habayeb et al., 2009; Merkling et al., 2013; Michel et al., 2001a; Sabin et al., 2010), whereas the IMD pathway is triggered by Gram-negative bacteria (Choe et al., 2002; Kaneko et al., 2004; Myllymaki et al., 2014). Several experimental pieces of evidence suggest that the two signaling pathways function independently and display specificity for recognizing various microorganisms. Both signaling pathways eventually activate nuclear factor κB, i.e., nuclear factor kappa-light-chain-enhancer of activated B cells (NF-κB) transcription factors that play a major role in host defense (Ganesan et al., 2011; Hetru et al., 2009; Meng et al., 1999).

The *Drosophila* genome encodes three NF-κB family members: two 70 kDa proteins, Dorsal and DIF (Dorsal-related immunity factor), and a 100 kDa protein Relish. Upon microbial infection, the Toll pathway relays its signal via MyD88-Tube-Pelle to phosphorylate and degrade the Inhibitor of κB (IκB) factor Cactus, resulting in the release of DIF and Dorsal, which then translocate into the nucleus (Wu et al., 1998) to activate target genes. Dorsal was originally identified as a gene involved in dorsoventral polarity during embryogenesis (Mohier 1993; Stein et al., 1998) and also functions in innate immunity in larval stage but not in adults (Hedengren-Olcott et al., 2004). DIF was identified as Dorsal-related immune response gene in *Drosophila* larval fat body (Ip et al., 1993) and as a predominant transactivator that works in conjunction

with Dorsal. Both Dorsal and DIF contain a Rel-homology domain (RHD) required for DNA binding and nuclear localization, and a C-terminal domain (CTD) (Hetru et al., 2009). Both DIF and Dorsal share 45% identity to mammalian NF-κB factors, c-Rel, Rel A and Rel B. DIF/Dorsal can activate AMP genes like *Drosomycin*, *Mechnikowin* and *Cecropin* in response to fungal infections, and upregulate *Defensin* and *Mechnikowin* genes in response to Gram-positive bacterial infection (Imler et al., 2005; Lemaitre et al., 1995; Levashina et al., 1998; Manfruelli et al., 1999b).

The IMD pathway is triggered via PGRP-LC-IMD-Dredd to finally activate NF-κB factor Relish (Dushay et al., 1996a; Hedengren et al., 1999; Silverman et al., 2000b). Relish contains an RHD and a C-terminal IκB-like domain consisting of six ankyrin repeats (Hetru et al., 2009). Relish is similar to mammalian p100 and p105 factors and its activation requires proteolytic cleavage to remove the ankyrin repeats, a process similar to activation of p100 and p105 to p52 and p50, respectively, in mammals (Hultmark 2003a). Relish controls the expression of *Attacin*, *Cecropin* and *Diptericin* genes (Dushay et al., 1996a; Hedengren-Olcott et al., 2004; Imler et al., 2005; Stoven et al., 2003). Conserved Toll and IMD signaling pathways, as well as downstream signaling proteins have been identified in mosquitoes (Christophides et al., 2002), bees (Evans et al., 2006), and lepidopteran insects (Xia et al., 2015; Zhong et al., 2012a).

Although different AMP genes are regulated predominantly by one or the other signaling pathway, it has been reported that AMPs can be activated synergistically by the Toll and IMD pathways (Tanji et al., 2007c), and DIF, Dorsal, and Relish may form homo- and heterodimers (Tanji et al., 2010b). In mammals, the five NF-κB factors (c-

Rel, Rel A, Rel B, p50 and p52) can form 15 different combinations of homo- and heterodimers that bind to specific DNA sequences to activate target genes (Hoffmann et al., 2006). However, in order to form DIF-Relish or Dorsal-Relish heterodimers, both the Toll and IMD pathways must be activated simultaneously to release DIF and Dorsal from the IκB factor Cactus, and to cleave the ankyrin repeats of the Relish to produce Relish-N (N-terminus of Relish without inhibitory ankyrin repeats). Thus, it is not clear how the three insect NF-κB factors (DIF, Dorsal, and Relish) cooperatively regulate expression of AMPs and other immune-related genes. Recently, we cloned two short isoforms of Relish, named Rel2A and Rel2B, in the tobacco hornworm *Manduca sexta*, which contains only RHDs but no IκB-like domains, and Rel2-RHD can form heterodimers with Dorsal-RHD to negatively regulate AMP gene expression (Zhong et al., 2016b). In *Drosophila*, an IκB protein Pickle, which is similar to the IκB-like domain of Relish, can selectively inhibit Relish homo-dimers (Morris et al., 2016a). These results suggest the existence of Relish alternatively spliced isoforms in insects. Indeed, in the mosquito *Aedes aegypti*, three alternatively spliced transcripts of Relish have been identified, with the predominant Relish transcript containing both the RHD and IκB-like domain, and the other two Relish isoforms either containing only the RHD or only the IκB-like domain (Shin et al., 2002). In *Anopheles gambiae*, REL2 (Relish) gene produces two spliced isoforms: a full-length REL2-F and a shorter REL2-S that lacks the ankyrin repeats (Meister et al., 2005). In the silkworm *Bombyx mori*, two Relish isoforms have also been identified, with BmRelish1 containing both the RHD and IκB-like domain, whereas BmRelish2 lacks the ankyrin repeats and functions as a dominant negative regulator of the active BmRelish1 (Tanaka et al., 2007).

The existence of alternatively spliced short isoform of Relish containing only RHD may account for the formation of heterodimers between DIF or Dorsal and Relish, but Relish short isoform in *Drosophila* has not been identified. This study aims to investigate direct interaction among the three NF-κB factors and how NF-κB homo- and heterodimers regulate AMP gene expression, to identify key nucleotides in the NF-κB binding sites that account for activation of NF-κB homo- and/or heterodimers, and to clone and characterize *D. melanogaster* short isoform of Relish.

Materials and Methods

Fly and Cell Line

D. melanogaster (wild-type w^{1118}) were a gift from Professor Leonard Dobens, School of Biological Sciences at University of Missouri – Kansas City. These flies were reared on artificial diets made with sterile cornmeal-molasses, maintained at 25°C, which were later used for RNA isolation and infection experiments. *D. melanogaster* Schneider 2 (S2) cells were purchased from American Type Culture Collection (ATCC). Cells were maintained at 25°C in Insect Cell Culture Media (SH30610.02, Hyclone) supplemented with 1% penicillin-streptomycin solution (G6784, Sigma-Aldrich), 1% L-Glutamine (25030081,Gibco™) and 10% heat-inactivated fetal bovine serum (#10082063, Invitrogen).

Cloning of *D. melanogaster* Short Isoform of Relish (Rel-S)

Total RNA from the embryo (0-7 h), larvae (1st, 2nd and 3rd instar), wandering larvae, pupae (5 days old) and adults (females and males) of wild type *D. melanogaster* W^{1118} strain were isolated using TRIzol® Reagent (T9424, Sigma-Aldrich) according to the manufacturer's protocol. RNA integrity was confirmed using the NanoDrop™ 2000 (Thermo Scientific, USA). Primers (Table 1) were designed based on the sequence of Relish (Genbank accession number: NM_057746). 5' and 3' cDNA were synthesized from total RNA using SMARTer® RACE 5'/3' Kit (634858, Clontech, USA) following manufacturer's instruction. The 5' and 3' cDNA were used to do 5' and 3' RACE PCR following manufacturer's instruction. 5' RACE PCR yielded three bands of ~400bp,

~700bp and ~1.2Kb, whereas, 3' RACE PCR yielded three bands of ~800bp, ~1.2kb and ~2.2Kb. All the PCR products were cloned into pUC19 vector and sequenced with M13 primers. ~800bp PCR product showed the very short 3' UTR followed by a poly A tail and 5' 700bp product shows unique 5'UTR region before ORF. Hence, another set of primers were designed based on this 5' and 3' UTR regions and used to do long distance PCR with LongAmp® Taq DNA Polymerase (M0323, NEB). PCR condition used was 94^0C for 30", 94^0C for 30", 60^0C for 30", 65^0C for 2 min for 35 cycles, followed by 65^0C for 10 min and final hold at 4^0C. PCR product of ~1.5 Kb was obtained and sequenced, and this cDNA clone was named short isoform of Relish (Rel-S).

Developmental and Induced Expression of Rel-S mRNA

Total RNA from embryo (0-7 h), larvae (1^{st}, 2^{nd} and 3^{rd} instar), wandering larvae, pupae (5 day old) and adults (females and males) of wild type *D. melanogaster* W^{1118} strain were isolated using TRIzol® Reagent (T9424, Sigma-Aldrich) according to the manufacturer's protocol for analysis of Rel-S mRNA expression at different developmental stages. To determine induced expression of Rel-S transcript in adult flies, the following microorganisms were used for septic infection: *Escherichia coli DH5α* (TIANGEN, China), *Staphylococcus aureus* (a gift from Professor Brian Geisbrecht, School of Biological Sciences at University of Missouri – Kansas City), *Saccharomyces cerevisiae* (BY4741), *Cryptococcus neoformans (alpha)* (gifts from Professor Alexander Idnurm, School of Biological Sciences at University of Missouri – Kansas City), *Enterococcus faecalis* V583 strain (gift from Professor Michael Gilmore,

Harvard Medical School), and *Pseudomonas aeruginosa* PA-14 strain (gift from Professor Kalai Mathee, Florida International University). Yeast cells were cultured in YPD medium (0.5% yeast extract, 1% peptone, 1% dextrose, 2% agar) at 30°C for 18-24 h with agitation**,** *E. faecalis* cells were grown in Brain Heart Infusion medium at 37°C with agitation, and all other bacteria were cultured in LB medium at 37°C with agitation. Overnight cultured microbial cells were collected by centrifugation, washed with PBS twice and diluted in PBS to OD_{600} = 0.5 for injection experiments. Adult male and female flies (2-7 days old) (10 male or female flies per group) were anesthetized on CO_2 bed briefly, injected with 18.4 nl of different microbes by Nanoject II Auto-Nanoliter Injector (3-000-204, Drummond Scientific) and capillary needles into the lateral thoracic region below the wings, and the flies were quickly transferred to the food for recovery (Khalil et al., 2015). These flies were collected 6 h post-injection for RNA isolation and cDNA synthesis as described above.

Real-time PCR was performed in 15-µl reactions containing 7.5 µl 2× FastStart Universal SYBR Green Master (Rox) (04913914001, Roche), 3 µl H_2O, 3 µl diluted (1:50) cDNA, and 0.75 µl each reverse and forward diluted primer (10 pmol/µl), using the following program: 2 min at 50°C, 10 min at 95°C, followed by 40 cycles of 95°C for 15 s, 60°C for 1 min and the dissociation curve analysis. Data from three replicates of each sample was analyzed with SDS software (ABI) using a comparative method ($2^{-\Delta\Delta Ct}$) (Livak et al., 2001).

Construction of Expression Vectors

To express Relish-N (N-terminus of Relish, residues 1- 475), Relish-RHD (residues 152 - 335), DIF-RHD (residues 80 – 251), Dorsal-RHD (residues 49–219) and Rel-S (residues 94 – 567), gene specific primers (Table 1) were designed based on the sequences acquired from http://flybase.org/. Total RNAs from 12 h white pre-pupa and 0-7 h embryo were used for cDNA synthesis. Contaminated genomic DNA was removed by RQ1 RNase-free DNase I (Promega). cDNA was prepared from total RNA (2 μg for each sample) using moloney murine leukemia virus (M-MLV) reverse transcriptase (M1701, Promega) with an anchor-oligo$(dT)_{18}$ primer following the manufacturer's instructions as described previously (Zhong et al., 2016b). Touchdown PCR was performed as described above. PCR products were analyzed by 1% agarose gel, recovered using agarose gel electrophoresis-Wizard® SV Gel and PCR Clean-Up System (A9285, Promega), digested with Kpn I/Xho I or Kpn I/Not I restriction enzymes (#R3142, R0146, R3189, NEB), and ligated into the digested pAC5.1/V5-His A or pAC5.1/FLAG A vector (V413020, Invitrogen) respectively using Rapid DNA Ligation Kit (K1422, ThermoFisher Scientific) according to the manufacturer's protocol. Recombinant expression vectors were transformed into competent *E. coli* DH5α cells. Positive bacterial colonies were selected, plasmid DNAs were extracted and sequenced by Sanger Sequencing using Applied Biosystems 3730xl 96-capillary DNA Analyzer in the DNA Core Facility at University of Missouri – Columbia, and then used for protein expression in S2 cells.

Construction of Luciferase Reporter Plasmids

Luciferase reporter plasmids for the seven *Drosophila* AMP gene promoters are available in our laboratory (Rao et al., 2011). To construct mutated *Diptericin*, *Attacin* A, *Drosomycin*, *Metchnikowin* (*Mtk*) promoters, site-directed mutagenesis was performed using the wild-type reporter plasmids as templates and primers (Table 1) designed using the online tool at http://nebasechanger.neb.com/. Mutagenesis reactions were performed using Q5® Site-Directed Mutagenesis Kit (E0554, NEB) following the manufacturer's protocol. The mutagenesis products were transformed into competent *E. coli* DH5α cells, mutant reporter plasmids were purified and sequenced and then used for transient transfection in S2 cells.

Dual Luciferase Assays

For transient transfection experiments, 3×10^5 S2 cells were plated in 24-well plates and transfected with 500 ng/well of pAC5.1/V5-His A (empty vector) or recombinant pAC5.1 A expression plasmid, 250 ng/well of pGL3B empty vector or pGL3B-promoter reporter plasmid and 25 ng/well of pRL-TK renilla luciferase plasmid (internal control) (Promega, USA) using GenCarrier-1™ (31-00110, Epoch Life Science) in serum free HyClone™ SFX-Insect cell culture media (SH30278.01, GE Healthcare Life Sciences). Twenty-four hours after transfection, media were replaced with normal growth media for protein expression up to 48 h. Dual-luciferase reporter assays were performed in 96-well plates using Dual-Luciferase® Reporter 1000 Assay System (E1980, Promega) following the manufacturer's instructions in GloMax®-Multi Microplate Reader with dual injector system (E7031, Promega). Relative

luciferase activity (RLA) from S2 cells cotransfected with empty pAC5.1/V5-His A and pGL3B-promoter reporter plasmid was used as the calibrator control (Rao et al., 2011).

Western Blot and Co-immunoprecipitation (Co-IP) Assay

Expression of recombinant Relish-N, Relish-RHD, Rel-S, DIF-RHD and Dorsal-RHD proteins was analyzed by Western blot. S2 cells (2×10^6 cells/well) transfected with recombinant expression plasmids were allowed for protein expression for 48 h, harvested by centrifugation, and lysed in NP-40 Lysis Buffer (BP-119, Boston BioProducts) containing 1:1000 dilution of protease inhibitor cocktail (1860, Sigma-Aldrich). The lysates were centrifuged at 12,000g for 5 min at 4°C to remove cell debris and the clear cell lysates were transferred to fresh tubes. Cell lysates (10 µl from each sample, equivalent to ~5×10^4 cells) were separated on 12% SDS-PAGE and proteins were transferred to nitrocellulose membranes (162–0097, Bio-Rad). The membrane was blocked with 5% BSA in Tris-buffered saline (TBS) (100 mM Tris-HCl, pH 7.6, 150 mM NaCl) containing 0.1% Tween-20 (TBS-T) at room temperature for at least 3 h, and incubated overnight with Anti-Flag M2 antibody (F-1804, Sigma-Aldrich, 1:3000 dilution) or anti-V5 antibody (V-8012, Sigma-Aldrich, 1:3000 dilution) as primary antibody at 4°C in TBS-T containing 5% BSA with gentle rocking. Then, the membrane was washed three times with TBS-T and incubated with the alkaline phosphatase-conjugated anti-mouse antibody (A4312, Sigma-Aldrich, 1:10,000) in TBS-T containing 5% BSA for 2 h at room temperature. After washing three times with

TBS-T (10 min each time), the signal was developed by using alkaline phosphatase (AP)-conjugated color development Kit (170-6432, Bio-Rad) (Yi et al., 2013).

Co-immunoprecipitation (Co-IP) assays were performed the same as described previously (Zhong et al., 2012a) using S2 cells expressing recombinant proteins with a Flag-tag (DIF-RHD, Dorsal-RDH) or a V5-tag (Relish-RHD, short-Relish). Briefly, S2 cells were lysed in NP-40 lysis buffer containing 1:1000 dilution of protease inhibitor cocktail and centrifuged, clear cell lysates were transferred to fresh tubes, and protein concentrations were determined by NanoDrop using BSA as the standard. Lysates were pre-cleared with 10 μl of protein-G agarose beads for 2 h with agitation. Equal amounts of bait and pray proteins were mixed (in a total volume of 1 ml) and incubated overnight at 4°C with 2 μl of either anti-Flag M2 or anti-V5 antibody. Then 20 μl of protein-G agarose beads were added to the mixture and incubated overnight at 4°C with agitation. The protein-bead complex was then washed three times with PBS and collected by centrifugation, mixed with 2 × SDS loading buffer and heated at 95°C for 5 min. The samples were then subjected to Western blot analysis the same as described above.

Data Analysis

At least three replicates of each sample were analyzed for each experiment, and experiments were repeated with three independent biological samples (or three independent cell cultures), and a typical set of data was used to make figures. Figures were made from means of three independent biological replicates with the GraphPad Prism (GraphPad, San Diego, CA). The significance of difference was determined by

one way ANOVA followed by a Tukey's multiple comparison tests using GraphPad Prism. Identical letters are not a significant difference ($p>0.05$) while different letters indicate significant difference ($p<0.05$). The significance of difference was also determined by an unpaired t-test with the GraphPad Prism (*, $p<0.05$; **, $p<0.01$; ***, $p<0.001$).

Genes for Cloning	Primers (5’ to 3’)
DmRelishRHD-KpnI-N’	AATGGTACCGCCGCCACCATGCTGCGGATCGTTGAGCAACC
DmRelishRHD-XhoI-C’	CCGCTCGAGGCGGTTATTGATTGCATTGCT
RelN-KpnI-N’	AATGGTACCATGAACATGAATCAGTACTACGA
DmDIF-RHD-KpnI-N’	AATGGTACCGCCGCCACCATGCTGCGTATCGTGGAGGAGCC
DmDIF-RHD-NotI-C’	CAAGCGGCCGCCTTGCCGTAAATCGGTGAGGA
DmDorsal-RHD-KpnI-N’	AATGGTACCGCCGCCACCATGGTAAAGATCACCGAACAACCGGC
DmDorsal-RHD-NotI-C’	CAAGCGGCCGCTGAGTCCATTGGCACGTACTC
pGL3B-Drs-N’	GGGGTACCCAATGAAAGTGATAATACGAATTGACC
pGL3B-Drs-C’	GAAGATCTATTGGAAAAGGTTCTCACGGAGC
pGL3B-Dpt-N’	GGGGTACCAGTAACTTTACTGATAAGACTTGGATTCTC
pGL3B-Dpt-C’	GAAGATCTCTCAGTTGTTCTCAATTGAAGAACTG
pGL3B-Att-N’	GGGGTACCATACTTGCTCAAAACAAAACCACA
pGL3B-Att-C’	GAAGATCTGTTGCTGAACTGGATTGCTGG
pGL3B-Dro-N’	CGGGGTACCGCCCAAGAGTTTTCCCAAAGAGC
pGL3B-Dro-C’	GGAAGATCTTGTGCTTGGAGTGGTGGACAAATCG
pGL3B-Cec-N’	GGGGTACCACTAAGTTACTAACGCAAGACTTTTAGTTAAG
pGL3B-Cec-C’	GAAGATCTGGTGATATTTTCTTGATTTTTTCTTAGG
pGL3B-Mtk-N’	CGGGGTACCTTTCTTAGCCCAGTTCTTAGTTCTG
pGL3B-Mtk-C’	GGAAGATCTCTTAGCTCGGTGGCGGGAATTGATTG
pGL3B-Def-N’	GGGGTACCGAAAACGTGAGCCGTCGAG
pGL3B-Def-C’	GAAGATCTCTTGGAATACAACTGGAGAGATAC
DmRel-S-KpnI-N’	AATGGTACCATGAACATGAATCAGTACTACGACC
DmRel-S-XhoI-C’	CCGCTCGAGAGGCATCCAAGTATTAGAGCTTTCTGTTC
Site directed mutagenesis	**Primers (5’ to 3’)**
Drs2-Mtkmut-N’	GTCCCCTACCGAAGGCCTATAAATG
Drs2-Mtkmut-C’	TTCCCTGTACGCTTATCGAAAAG
Dpt1-Mtkmut-N’	GTCCCGTCTTTTCCGGTGGACCTTC
Dpt1-Mtkmut-C’	TTCCCTGGGGCAATGATGACCAG
Mtk-Drs2mut-N’	TTCCCCTTTGGGTGGTGCTGGCT
Mtk-Drs2mut-C’	CTACTACGGTCCCGGATTATCAG
Mtk-Dpt1mut-N’	ATTCCCTTTGGGTGGTGCTGGCT
Mtk-Dpt1mut-C’	TTCCCACGGTCCCGGATTATCAG
Atta-Mtk-N’	GTCCCGCTTTGATAAGGCATCCAG
Atta-Mtk-C’	TTCCCGCTCAGATGTGATGGTGG
Mtk-Atta-N’	ATTTCCTTTGGGTGGTGCTGGCT
Mtk-Atta-C’	TCCCCACGGTCCCGGATTATCAG
Mtk-G-A-N’	ATCCCCTTTGGGTGGTGCTGGCT

Mtk-G-A-C’	TTCCCACGGTCCCGGATTATCAG
Mtk-C-T-N’	GTTCCCTTTGGGTGGTGCTGGCT
Mtk-C-T-C’	CTGCCCTTTGGGTGGTGCTGGCT
Mtk-G-C-N’	CTCCCCTTTGGGTGGTGCTGGCT
Mtk-G-C-C’	TTCCCACGGTCCCGGATTATCAG
Mtk-C-G-N’	GTGCCCTTTGGGTGGTGCTGGCT
Dpt-A-G-N’	GTTCCGTCTTTTCCGGTGGACCTTC
Dpt-A-G-C’	TTCCCTGGGGCAATGATGACCAG
Dpt-T-C-N’	ATCCCGTCTTTTCCGGTGGACCTTC
Dpt-T-C-C’	TTCCCTGGGGCAATGATGACCAG
Drs-C-G-N’	GTACTTGTACGCTTATCGAAAAG
Drs-C-G-C’	TTCCCCTACCGAAGGCCTATAAATG
Drs-A-C-N’	CTCCTTGTACGCTTATCGAAAAG
Drs-A-C-C’	TTCCCCTACCGAAGGCCTATAAATG
Drs-T-C-N’	CTACCTGTACGCTTATCGAAAAG
Drs-T-C-C’	TTCCCCTACCGAAGGCCTATAAATG
Rel-RHD 3’RACE	**Primers (5’ to 3’)**
3’Rel-GSP2-N’	GATTACGCCAAGCTTGCGCGGCTATGTGGCGCAATTTATCAAC
3’Rel-NGSP2- N’	GATTACGCCAAGCTTGAACCAGGTGCGGCTCTGCTTTGAGGCC
5’Rel-GSP1-R’	GATTACGCCAAGCTTCAAATCATGCGGATCGCACACATCCCG
5’Rel-NGSP1-R’	GATTACGCCAAGCTTAGTTGCACAGTGTAACCTCCGGGAAGG
3’UPM	CTAATACGACTCACTATAGGGCAAGCAGTGGTATCAACGCAGAGT
3’UPS	CTAATACGACTCACTATAGGGC

Table 1. Primers for Cloning *Drosophila* NF-kB Transcription Factors, AMP gene promoters, AMP Promoter Mutation and 3’RACE of Relish-RHD. N’ represents forward primer and C’ represents reverse primer.

Results

cDNA Cloning of *Drosophila* Relish-Short Isoform

Relish is a member of the Rel/NF-κB transcription factors with a Rel-homology domain (RHD) and an inhibitory IκB-like domain that is composed of six ankyrin repeats. In Drosophila, there are four full-length Relish transcript variants containing both RHD domain and ankyrin repeats, and an IκB-like Relish short isoform has been reported recently (Morris et al., 2016b). To obtain the RHD-like Relish short isoform, 3' and 5' RACE were performed using gene specific primers (GSP) designed from the Rel-homology domain (Table. 1). Three DNA fragments of ~800 bp, 1.2 kb and 2.2 kb were obtained from the 3'-RACE, and sequence analysis revealed that the 800-bp fragment contains a TGA stop codon followed by a short 3'-UTR region with a poly A tail (Fig. 1 A). All four *Drosophila* full-length Relish transcripts have the identical 41 bp (5'-TTA TGT ATT GAA TGT TGA TCA ATA AAG TAC CTT AGT TTT AC-3') at the 3'-UTR (www.flybase.org), and the last 12 bp (underlined) match the sequence of our 800-bp RACE product. Comparison of the DNA sequences between the 800-bp fragment and the full-length Relish reveals that the 800-bp fragment encodes RHD and Pest domains and differs only by 20 bp at the 3' end. We believe the 800-bp fragment belongs to the RHD-like Relish short isoform, and thus designated as Relish-S (Rel-S). To obtain the full-length cDNA of Rel-S, 5'-RACE was performed using the primers from the 3'-UTR, and three fragments of ~400 bp, 700 bp and 1.2 kb were obtained. The 700 bp sequence contains the unique 5' UTR. The sequence of full-length Rel-S was eventually assembled by overlapping the sequences from 5'- and 3'-RACE. The ORF of Rel-S is 1434 bp encoding a protein of 477 amino acids, which is identical to

the first 477 residues of Relish, with a calculated molecular weight of 54 kDa and *p*I of 8.1. Sequence comparison between Rel-S and Long isoform revels the 20 bp sequence difference at the 3' end followed by poly A tail of the Rel-S isoform (Fig. 1 B).

ATGAACATGAATCAGTACTACGACCTGGACAATGGGAAAAATGTGATGTTTATGAACGATGCATCCAG
CACCAGTGGCTATAGCAGCAGTACATCACCCAACTCCACCAACCGATCCTTTTCACCAGCCCACTCCCC
GAAAACCATGGAACTGCAAACGGACTTCGCCAATCTTAATTTGCCCGGCGGCAATTCACCACACCAGC
CGCCAATGGCAAACTCACCATACCAGAATCAGCTGTTGAACAACGGCGGAATTTGCCAATTGGGTGCT
ACCAATTTAATAAACTCCACTGGCGTTAGTTTCGGCGTTGCTAATGTCACCAGTTTTGGCAACATGTACA
TGGATCACCAGTACTTTGTGCCCGCTCCGGCCACTGTGCCACCGTCCCAAAACTTTGGATACCATCAAA
ATGGCCTGGCATCTGACGGCGATATCAAACACGTGCCGCAGCTGCGGATCGTTGAGCAACCGGTGGA
GAAGTTCCGCTTTCGGTACAAGAGCGAGATGCATGGAACACATGGATCGCTAAATGGCGCCAATTCGA
AGCGGACGCCCAAAACCTTCCCGGAGGTTACACTGTGCAACTACGATGGACCCGCCGTCATCCGGTGC
AGTTTGTTCCAAACTAACTTGGACAGCCCACATTCCCATCAGCTGGTCGTGCGCAAGGACGATCGGGA
TGTGTGCGATCCGCATGATTTGCACGTGTCCAAGGAGCGCGGCTATGTGGCGCAATTTATCAACATGG
GCATCATACACACCGCCAAGAAGTACATATTCGAGGAACTGTGCAAGAAGAAGCAGGATCGCCTGGT
CTTTCAGATGAACCGCCGCGAGTTGTCCCACAAACAGCTACAGGAACTGCATCAGGAGACAGAGCGTG
AGGCCAAGGACATGAACTTGAACCAGGTGCGGCTCTGCTTTGAGGCCTTTAAAATTGAGGACAACGGC
GCATGGGTTCCACTTGCACCGCCTGTATACAGCAATGCAATCAATAACCGCAAGTCGGCACAAACTGG
AGAGCTTCGCATCGTCCGCCTGAGCAAACCCACTGGCGGGGTGATGGGCAACGATGAGCTTATTCTAC
TGGTGGAAAAGGTCAGCAAGAAGAACATTAAAGTGAGGTTCTTCGAGGAGGATGAGGACGGGGAAA
CCGTGTGGGAGGCATACGCAAAGTTCCGCGAATCGGATGTACACCACCAATATGCCATTGTGTGCCAG
ACGCCTCCGTACAAAGATAAGGACGTGGACCGCGAGGTCAACGTGTACATCGAGCTCATTCGTCCCTC
CGACGATGAGCGCTCATTCCCGGCGCTGCCCTTCCGCTACAAGCCACGGAGCGTAATTGTGTCGAGGA
AACGCCGACGAACCGGGTCCTCTGCCAATAGCAGCAGTTCGGGAACAGAAAGCTCTAATA**CTTGGAT**
GCCTTAGTTTTACAAAAAAAAAAAAAAAAAAAA

Figure 1. A. Cloning of *Drosophila* Relish-Short Isoform.

Open reading frame (ORF) of Relish-short isoform. ATG is the start of translation site and TAG is the translation termination site. The bases boxed are the 20 bases difference between Relish full-length and Relish-short isoform. These 20 bases are followed by 3' poly A tail. Unique feature of this Relish-short isoform is the presence of a very short 3' UTR before the poly A tail.

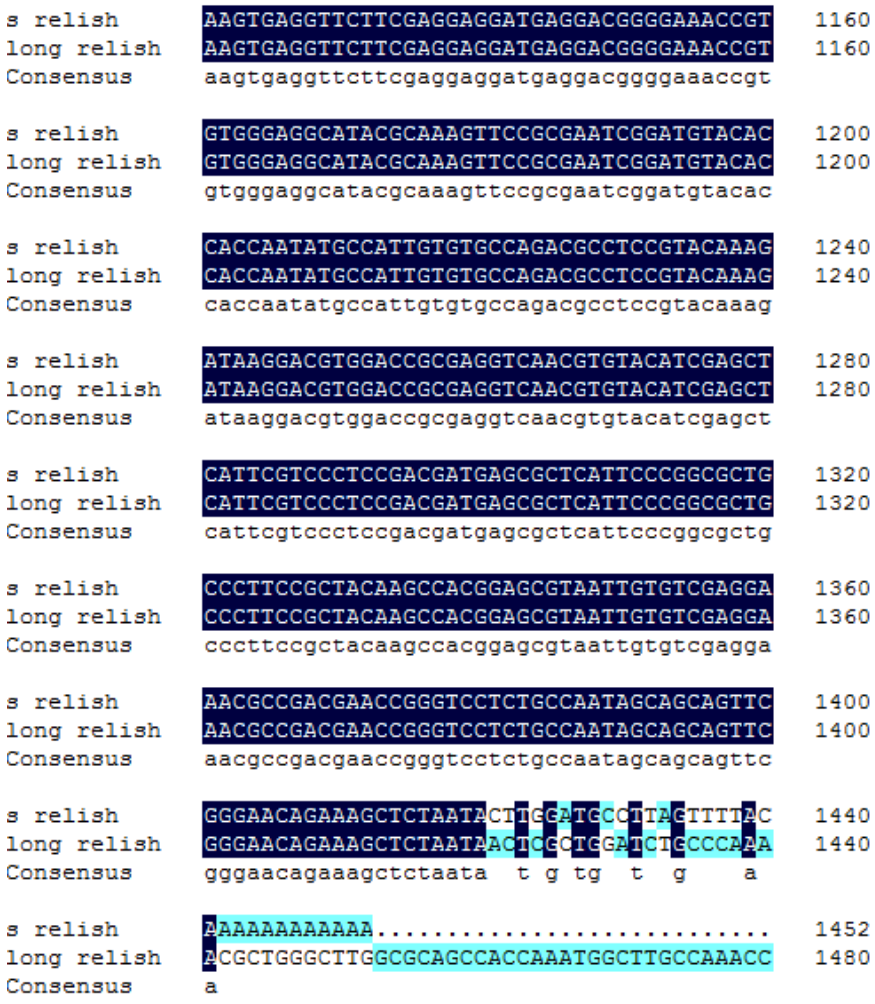


Figure 1. B. Sequence Comparison of Relish Long and Short Isoforms.

DNA sequence comparison between Relish Long and Short isoform using DNAman Multiple Sequence Alignment program. The dark blue are the identical sequence whereas, the light blue represents the differences in bases. After 1420bp the two isoforms differ, and short relish end with poly A tail after 1441 base. Comparison of amino acid sequence shows the absence of ankyrin repeats in both the sequences (Fig.8 A) and suggest the result of alternative splicing. Lower case represents consensus sequences.

Overexpression of Rel-S, Relish-RHD, Dorsal-RHD and DIF-RHD in *Drosophila* S2 cells

The full-length cDNA of Rel-S was obtained by PCR and cloned into the expression vector pAC5.1A/V5-His. Recombinant Rel-S contains a V5-tag at the C-terminus and was overexpressed in *Drosophila* S2 cells. *Drosophila* Relish (Rel)-RHD-V5, Relish-N (Rel-N-V5) (N-terminal region of Relish), Dorsal (Dl)-RHD-Flag and DIF-RHD-Flag domains were cloned also overexpressed in S2 cells. Western blot analysis showed that recombinant V5-tagged Rel-S, Rel-RHD and Rel-N were recognized by anti-V5 antibody (Fig. 2 A and B), while Flag-tagged Dl-RHD and DIF-RHD were recognized by anti-Flag antibody (Fig. 2 C). For Rel-S, at least three protein bands were recognized by anti-V5 antibody (Fig. 2A, lane 2), but only one Rel-RHD band was detected (Fig. 2B, lane 1), suggesting that Rel-S may be post-translationally modified. For Rel-N, two protein bands at ~50 kDa and over 100 kDa were detected (Fig. 2A, lane 3), indicating the existence of both Rel-N monomers and stable homodimers. When Rel-S was co-expressed with Rel-N, multiple protein bands were detected (Fig. 2A, lane 4), including at least two bands over 100 kDa, suggesting that Rel-S may form heterodimers with Rel-N.

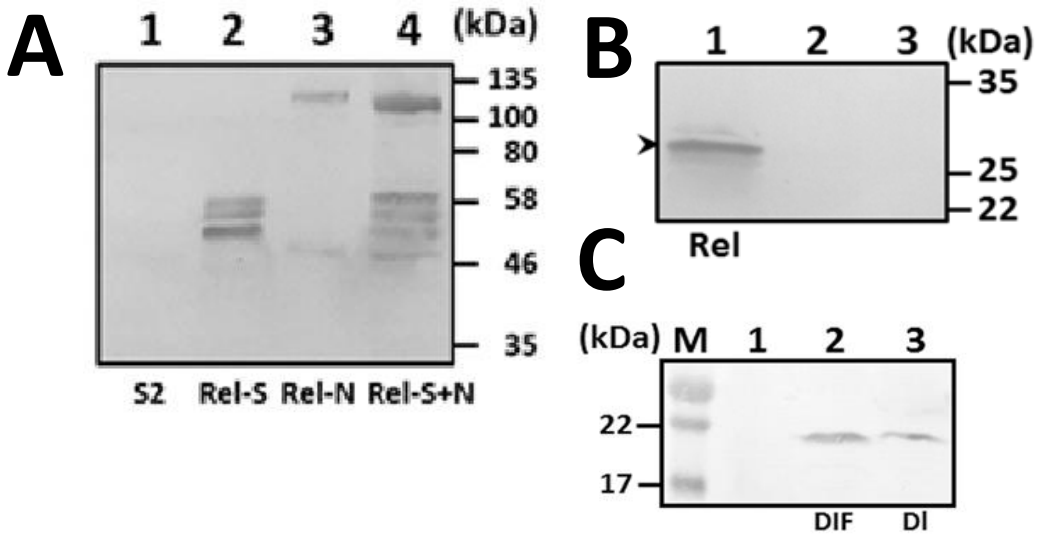


Figure 2. Protein Expression of Rel-S, Relish-RHD, Dorsal-RHD and DIF-RHD in *Drosophila* S2 Cells.

A. Overexpression of Rel-S (55.4 KDa), Rel-N (63 KDa) and WB detected by anti-V5 mouse antibody. B. Overexpression of Rel-RHD and WB detected by anti-V5 mouse monoclonal antibody. C. Overexpression of DIF-RHD (19.65 KDa) , Dorsal-RHD (19.01) and WB detected by anti-FLAG mouse monoclonal antibody.

Interaction of Relish-Short Isoform with NF-κB Transcription Factors

It has been suggested that heterodimers of NF-κB transcription factors can form in *D. melanogaster* with different efficiencies, and DIF-Relish heterodimers can regulate antimicrobial peptide genes (Tanji et al., 2010b). To test interaction of Rel-S with Dl-RHD and DIF-RHD, co-immunoprecipitation (Co-IP) assay was performed. Co-IP results showed that Flag-tagged DIF-RHD and Dl-RHD were co-precipitated with V5-tagged Rel-S (Fig. 3E), and vice versa, V5-tagged Rel-S was co-precipitated with Flag-tagged DIF-RHD and Dl-RHD (Fig. 3F), suggesting that Rel-S can interact with both DIF-RHD and Dl-RHD. Co-IP assays were also performed for Relish-RHD (Rel-RHD) with DIF-RHD and Dl-RHD, and the results showed that V5-tagged Rel-RHD co-precipitated with Flag-tagged DIF-RHD and Dorsal-RHD (Fig. 4 E and F), further confirming that RHD-like short isoform (Rel-S) or RHD domain of Relish can interact with both DIF-RHD and Dl-RHD.

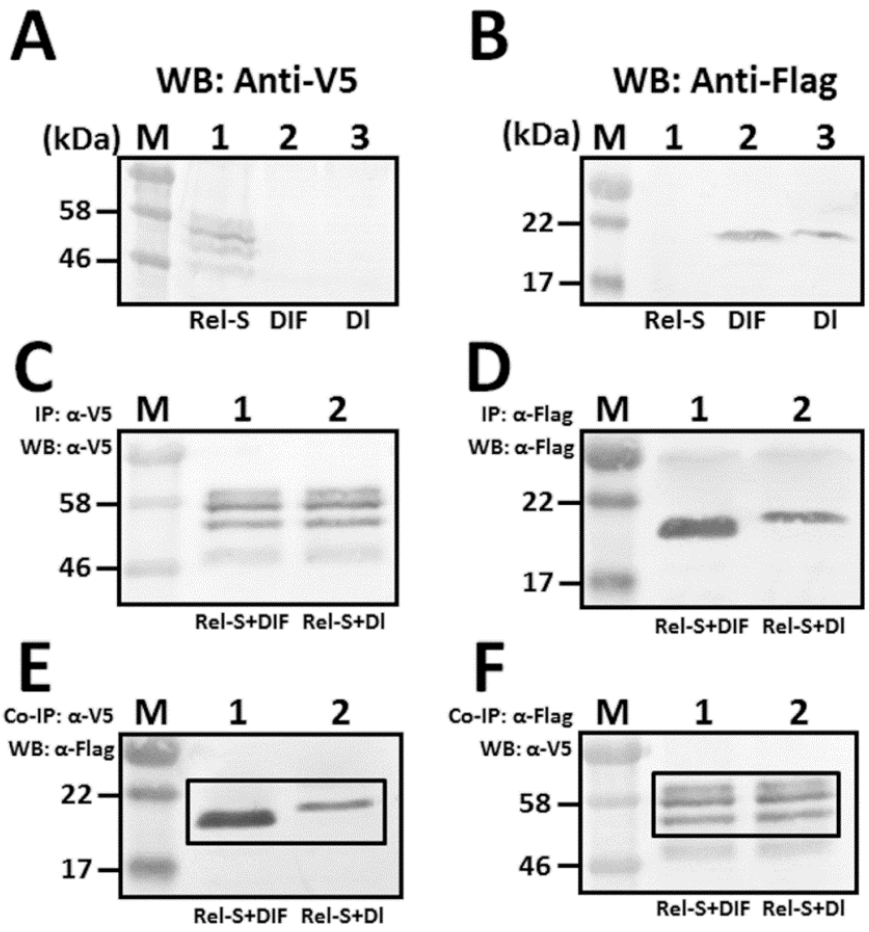


Figure 3. Interaction of Relish-Short Isoform with NF-κB Transcription Factors.

Recombinant V5-tagged DmRel-S, Flag-tagged DmDIF-RHD, and Flag-tagged DmDorsal-RHD were expressed in S2 cells separately, and cell lysates from two samples were mixed for co-immunoprecipitation (Co-IP) assays as described in the Materials and Methods. Immunoprecipitated (IP) proteins or Co-IP proteins were detected by immunoblotting using anti-Flag or anti-V5 monoclonal antibody as the primary antibody. (A-B) Lanes 1–3 are western blot of cell lysates (protein inputs) from S2 cells overexpressing V5-tagged DmRel-S (Lane 1), Flag-tagged DmDIF-RHD (Lane 2) and Flag-tagged DmDorsal-RHD (lane 3). (C-D) is IP of V5-tagged DmRel-S mixed with DmDIF-RHD mix (Lane 1) and V5-tagged DmRel-S mixed with DmDorsal-RHD mix (Lane 2). (E-F) V5-tagged DmRel-S was co-immunoprecipitated with Flag-tagged Flag-tagged DmDIF-RHD, DmDorsal-RHD mix (boxed bands).

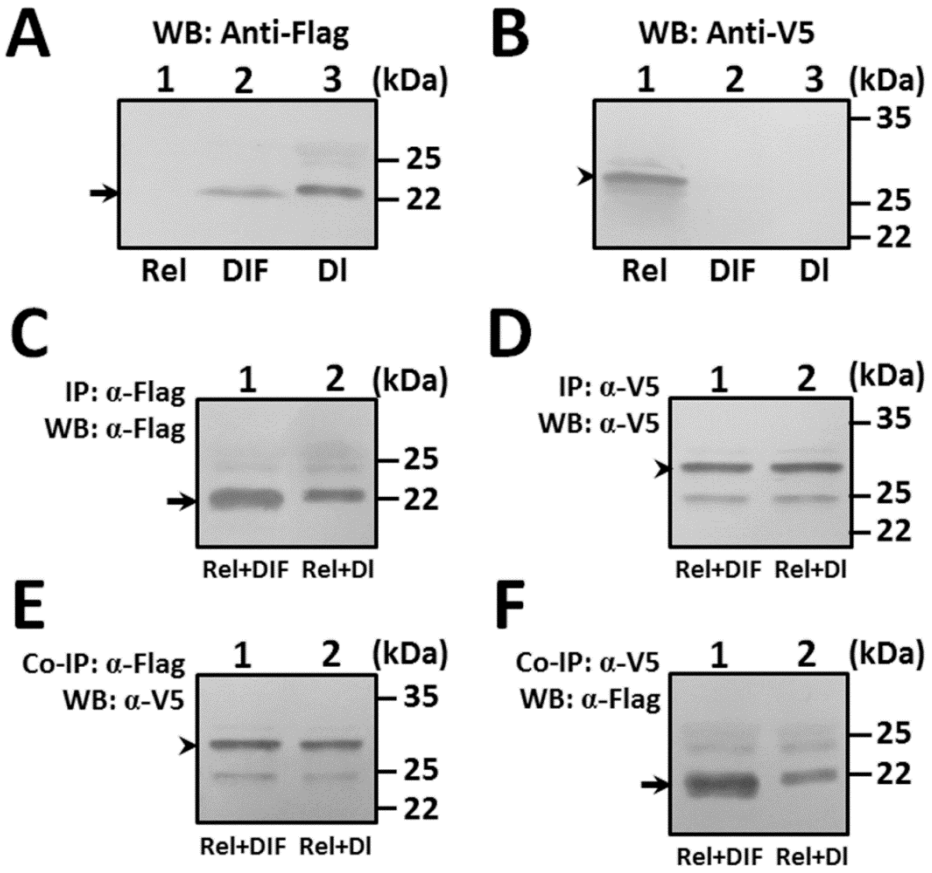


Figure 4. Interaction of Rel-RHD with NF-κB Transcription Factors.

Recombinant V5-tagged DmRel-RHD, Flag-tagged DmDIF-RHD, and Flag-tagged DmDorsal-RHD were expressed in S2 cells separately, and cell lysates from two samples were mixed for co-immunoprecipitation (Co-IP) assays as described in the Materials and Methods. Immunoprecipitated (IP) proteins or Co-IP proteins were detected by immunoblotting using anti-Flag or anti-V5 monoclonal antibody as the primary antibody. (A-B) Lanes 1–3 are western blot of cell lysates (protein inputs) from S2 cells overexpressing V5-tagged DmRel-RHD (Lane 1), Flag-tagged DmDIF-RHD (Lane 2) and Flag-tagged DmDorsal-RHD (lane 3). (C-D) is IP of V5-tagged DmRel-RHD mixed with DmDIF-RHD mix (Lane 1) and V5-tagged DmRel-RHD mixed with DmDorsal-RHD mix (Lane 2). (E-F) V5-tagged DmRel-RHD was co-immunoprecipitated with Flag-tagged Flag-tagged DmDIF-RHD, DmDorsal-RHD mix.

Activation of AMP Gene Promoters by Relish, Dorsal and DIF

To determine activation of AMP gene promoters by Relish, Dorsal and DIF, Rel-S, Rel-RHD, Dl-RHD or DIF-RHD was overexpressed in S2 cells with seven *Drosophila* AMP gene promoter-luciferase reporters separately for dual luciferase assays. The results showed that *drosomycin*, *cecropin*, *diptericin*, *drosocin* and *metchnikowin* promoters were all activated by Rel-S, Rel-RHD, Dl-RHD and DIF-RHD, but *attacin* and *defensin* promoters were not activated by any of the four NF-κB factors (Fig. 5). Among the five AMP gene promoters, Rel-RHD, Dl-RHD and DIF-RHD activated *metchnikowin* promoter to the highest level, while Rel-S activated *cecropin* promoter to the highest level (Fig. 5).

Since Rel-RHD and Rel-S can interact with Dl-RHD and DIF-RHD, we next tested AMP gene promoter activity by overexpression of different pairs of NF-κB transcription factors in S2 cells. Our results showed that *diptericin*, *drosomycin* and *metchnikowin* promoters were all activated to various levels by co-expression of Rel-RHD and DIF-RHD, Rel-RHD and Dl-RHD, Dl-RHD and DIF-RHD, Rel-S and DIF-RHD, and Rel-S and Dl-RHD (Fig. 6). Interestingly, co-expression of Rel-RHD with DIF-RHD or Dl-RHD activated the activity of the three promoters to a significantly lower level than overexpression of Rel-RHD, DIF-RHD or Dl-RHD alone (Fig. 6A-C, except for *drosomycin* promoter with Dl-RHD), while co-expression of DIF-RHD and Dl-RHD did not have such a significant effect on the activity of the three promoters compared to overexpression of DIF-RHD or Dl-RHD alone (Fig. 6A-C). But co-expression of Rel-S with DIF-RHD or Dl-RHD activated the three promoters to a similarly high activity as did overexpression of Rel-S alone (Fig. 6D-F), suggesting that

Rel-S can regulate AMP genes equally well as homodimers or heterodimers with Dorsal or DIF.

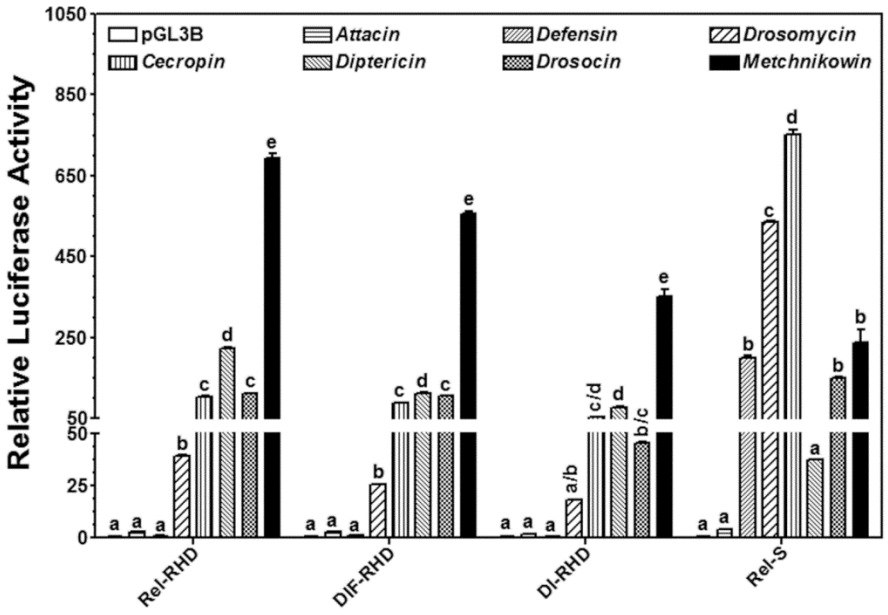


Figure 5. Activation of AMP Gene Promoters by *Drosophila* Rel-RHD, DIF-RHD, Dl-RHD and Rel-S Independently.

The relative luciferase activities of the promoters activated by recombinant *Drosophila* Rel-RHD, DIF-RHD, Dl-RHD, and Rel-S alone, were determined by Dual-Luciferase® Reporter Assay System as described in the Materials and Methods. Bars represent the mean of three independent measurements ± SEM. For the activity among different promoters activated by transcription factors, identical letters (capital letters for solid bars and small letters for striped bars) or identical numerical numbers (dotted bars) are not a significant difference ($p>0.05$) while different letters or different numerical numbers indicate significant difference ($p<0.05$). Comparing the activity of the same promoter stimulated by different transcription factors (between solid and striped bars, solid and dotted bars, as well as stripe and dotted bars for each promoter), the significance of difference was also determined by an unpaired t-test (*, $p<0.05$; **, $p<0.01$), and "n" indicates not significant.

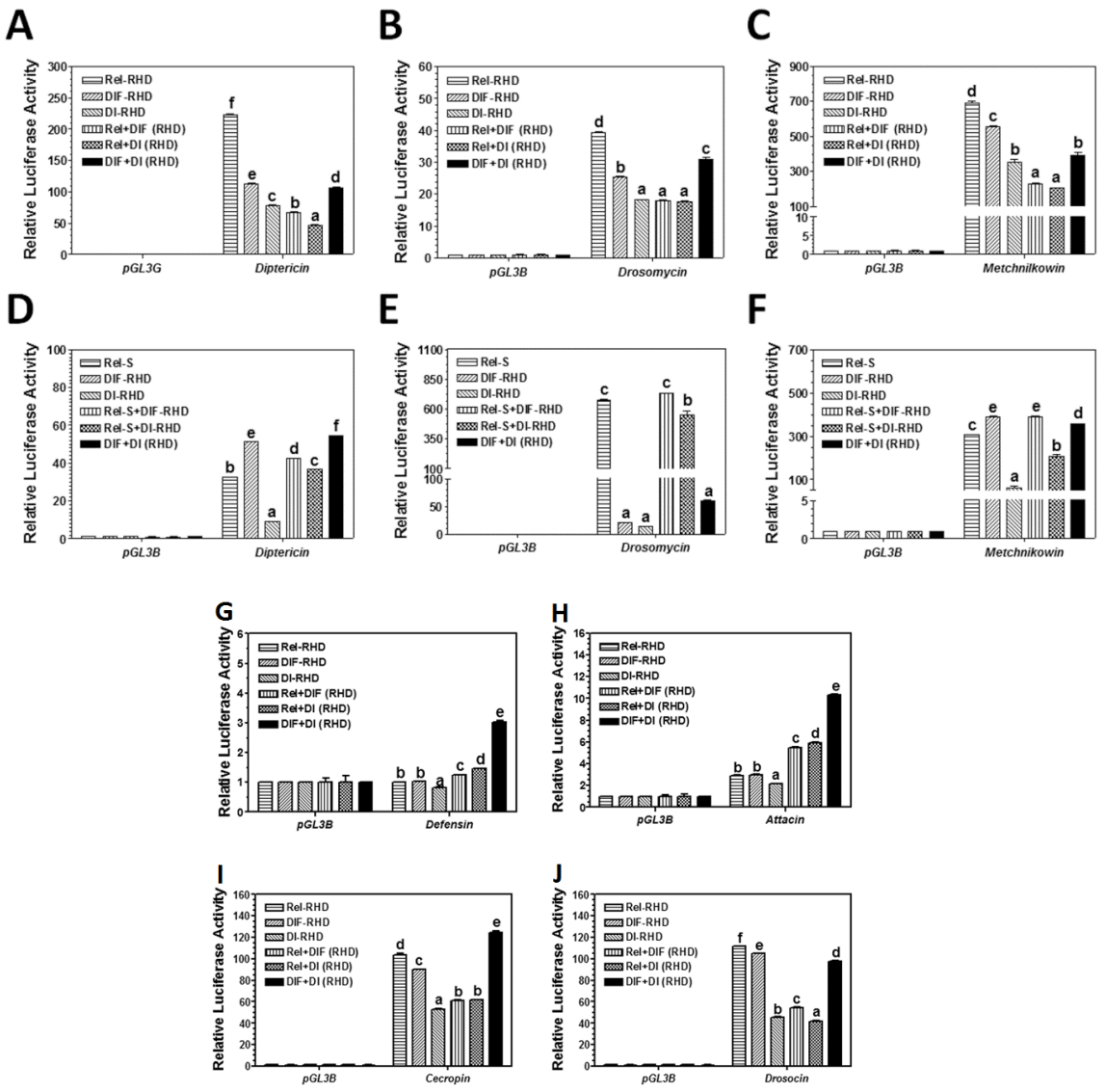


Figure 6. (A-J) Activation of AMP Gene Promoters by *Drosophila* NF-kB pairs.

The relative luciferase activities of the promoters activated by recombinant *Drosophila* Rel-RHD, DIF-RHD, Dl-RHD and Rel-S pair combinations, were determined by Dual-Luciferase® Reporter Assay System as described in the Materials and Methods. Bars represent the mean of three independent measurements ± SEM. For the activity among different promoters activated by transcription factors, identical letters (capital letters for solid bars and small letters for striped bars) or identical numerical numbers (dotted bars) are not a significant difference ($p>0.05$) while different letters or different numerical numbers indicate significant difference ($p<0.05$). Comparing the activity of the same promoter stimulated by different transcription factors (between solid and striped bars, solid and dotted bars, as well as stripe and dotted bars for each promoter), the significance of difference was also determined by an unpaired t-test (*, $p<0.05$; **, $p<0.01$), and "n" indicates not significant.

Developmental and Induced Expression of *Relish* and *Rel-S* in *Drosophila*

To determine expression of *Relish* and *Rel-S* mRNAs in *Drosophila*, we designed two pairs of gene-specific primers for PCR analysis. The result showed that both *Relish* and *Rel-S* transcripts were detected in *Drosophila* females and males, with higher expression level of *Relish* than *Rel-S* mRNA in males (Fig. 4A). Then, developmental and induced expression of *Relish* and *Rel-S* mRNAs in *Drosophila* was determined by real-time PCR. The results showed that *Relish* mRNA was expressed at a low level in the embryos, 1st-instar (L1) and 3rd-instar (L3) larvae, but at a higher level in the 2nd-instar (L2) larvae, pupae, females and males, while *Rel-S* mRNA was expressed at a significantly higher level in the 2nd-instar larvae than any other developmental stages (Fig. 7A). Comparing expression of *Relish* with *Rel-S, Relish* transcript was expressed at a significantly higher level than *Rel*-S in females, whereas *Rel-S* mRNA was expressed at a significantly higher level than *Relish* in the 2nd-instar and 3rd-instar larvae (Fig. 7A).

Induced expression of *Relish* and *Rel-S* mRNAs in *Drosophila* females and males after injection of microorganisms was also determined by real-time PCR. Compared to the control (PBS-injection) group, expression of *Relish* transcript in females was induced only by *C. neoformans* and *S. cerevisiae*, but inhibited by *E. coli* and *P. aeruginosa*; however, expression of *Rel-S* mRNA in females was inhibited by all the microorganisms tested, particularly by *E. fecalis, P. aeruginosa*, *S. aureus* and *S. cerevisiae* (Fig. 7B). Expression of *Relish* mRNA in males was induced by all the microorganisms tested but *E. coli*, especially by *S. aureus*, while expression of *Rel-S* transcript in males was induced by *C. neoformans*, *E. fecalis* and *S. aureus* (Fig. 7C). In

addition, *Relish* transcript was expressed at significantly higher levels than *Rel-S* in females after injection of all the microorganisms tested (except *E.* coli) (Fig. 7B) and in males after injection of *P. aeruginosa*, *S. aureus* and *S. cerevisiae* (Fig. 7C).

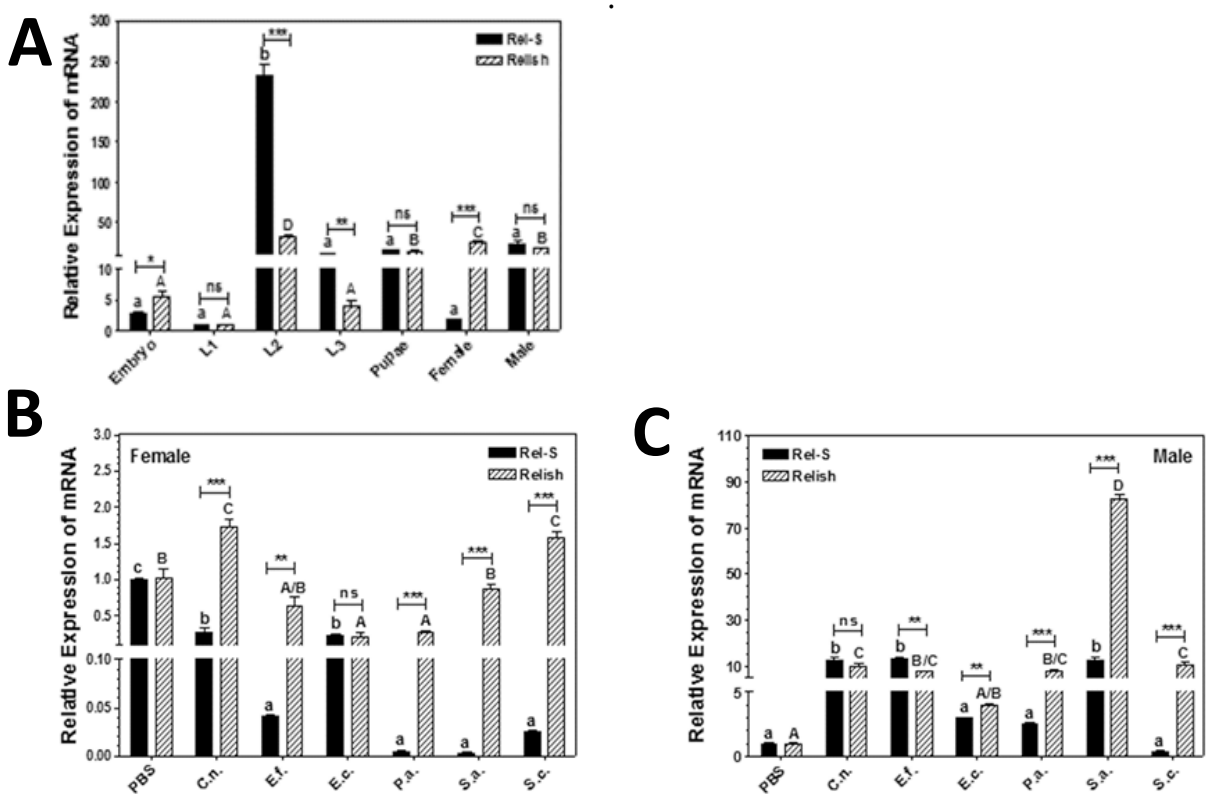


Figure 7. Expression and Function of Relish-Short Isoform in *Drosophila*

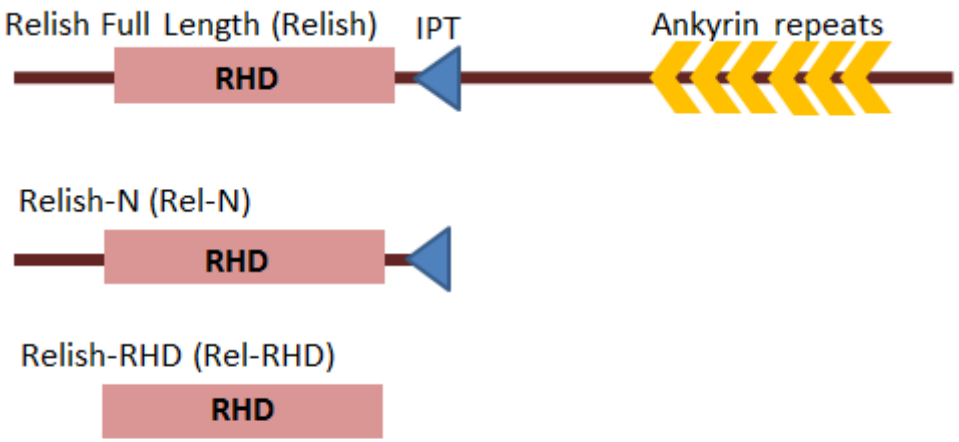
(A) qPCR of Rel-S (solid bar) and Relish (striped bar) expression in different developmental stages. (B) qPCR of Rel-S (solid bar) and Relish (striped bar) expression in females.(C) qPCR of Rel-S (solid bar) and Relish (striped bar) expression in males.

Activity of Rel-S in Activation of AMP Gene Promoters

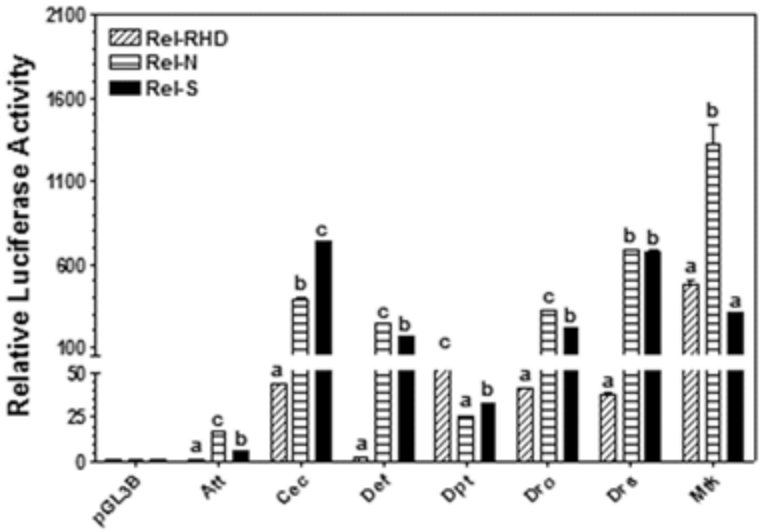
The amino acid sequence of Rel-S is almost identical to the N-terminal fragment of Relish (Rel-N) except a few residues at the C-terminus (Fig. 8 A). To compare the activity of Rel-S, Rel-N and Rel-RHD in activation of AMP genes, Rel-S,

Rel-N and Rel-RHD were overexpressed in S2 cells, and dual luciferase assays were performed. Rel-S and Rel-N activated *attacin*, *cecropin*, *defensin*, *drosocin* and *drosomycin* promoter activities to significantly higher levels than Rel-RHD did, but Rel-RHD showed significantly higher activity than Rel-N and Rel-S in activation of *diptericin*, and Rel-N activated *metchnikowin* promoter activity to a significantly higher level than Rel-S and Rel-RHD did (Fig. 8B). Overall, the activity of Rel-S is comparable to that of Rel-N in activation of AMP gene promoters (except for *metchnikowin*). Since Rel-S may form hetero-dimers with Rel-N (Fig. 2A, lane 4), Rel-S and Rel-N were also co-expressed, and the results showed that co-expression of Rel-S with Rel-N activated *metchnikowin* promoter to a significantly lower level than overexpression of Rel-N alone, but to a similarly high level as overexpression of Rel-S alone (Fig. 8C). For *diptericin* and *drosomycin* promoters, co-expression of Rel-S with Rel-N activated the activity to a comparable level as did by overexpression of Rel-S or Rel-N alone (Fig. 8C).

A


Relish Full Length (Relish)
IPT
Ankyrin repeats
RHD
Relish-N (Rel-N)
RHD
Relish-RHD (Rel-RHD)
RHD

B


Rel-RHD
Rel-N
Rel-S
Relative Luciferase Activity
2100
1600
1100
600
100
50
25
0
pGL3B
Att
Cec
Def
Dpt
Dro
Drs
Mtk

C

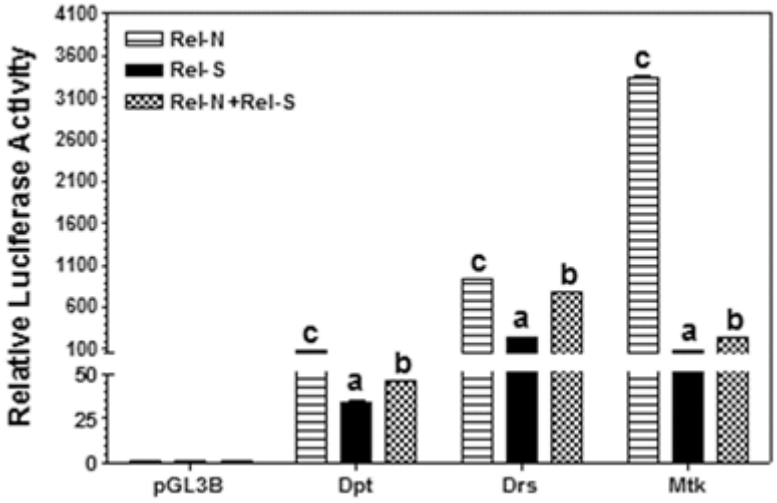

Rel-N
Rel-S
Rel-N+Rel-S
Relative Luciferase Activity
4100
3600
3100
2600
2100
1600
1100
600
100
50
25
0
pGL3B
Dpt
Drs
Mtk

Figure 8. Activity of Rel-S in Activation of AMP Gene Promoters.

(A) Comparison of all four types of Relish used in this study. (B-C) The relative luciferase activities of the promoters activated by recombinant *Drosophila* Rel-RHD, Relish and Rel-S pair combinations, were determined by Dual-Luciferase® Reporter Assay System as described in the Materials and Methods. Bars represent the mean of three independent measurements ± SEM. For the activity among different promoters activated by transcription factors, identical letters (capital letters for solid bars and small letters for striped bars) or identical numerical numbers (dotted bars) are not a significant difference ($p>0.05$) while different letters or different numerical numbers indicate significant difference ($p<0.05$). Comparing the activity of the same promoter stimulated by different transcription factors (between solid and striped bars, solid and dotted bars, as well as stripe and dotted bars for each promoter), the significance of difference was also determined by an unpaired t-test (*, $p<0.05$; **, $p<0.01$), and "n" indicates not significant.

Identification and Verification of NF-κB Binding Sites in AMP Gene Promoters

Since one NF-κB factor (Rel-RHD, DIF-RHD, Dl-RHD or Rel-S) can activate different AMP gene promoters to various levels (Fig. 5) and different NF-κB factors can also activate one AMP gene promoter to various levels (Fig. 6), the nucleotide sequence in the NF-κB binding sites must be important for the binding of NF-κB factors. Analysis of NF-κB binding sites in the seven AMP gene promoters used in our study showed that *metchnikowin*, *drosocin*, *defensin*, *cecropin* and *attacin* promoters contain only one NF-κB binding site, *diptericin* has three NF-κB sites and site 1 and 2 are identical (Kappler et al., 1993), and *drosomycin* has four NF-κB sites (Fig. 9A and B). *Drosomycin* NF-κB site 3 is inactive (Tanji et al., 2007), NF-κB site 4 is not required for its activity (Rao et al, 2011), NF-κB site 2 is activated by Gram-negative peptidoglycan via the IMD pathway (Tanji et al., 2007; Tanji et al., 2010), and mutation of the NF-κB site 2 can significantly decrease promoter activity in S2 cells (Rao et al, 2011). Thus, we chose NF-κB site 2 in *drosomycin* promoter and site 3 in *diptericin* promoter (sites 1 and 2 are identical) as well as the only one NF-κB site in *attacin* and *metchnikowin* promoters for further study (Fig. 9B). We first swapped the NF-κB site in *attacin*, *diptericin* and *drosomycin* promoters with the one in *metchnikowin* promoter and vice versa, since *metchnikowin* promoter showed the highest activity activated by Rel-RHD, DIF-RHD and Dorsal-RHD (Fig. 5). Dual luciferase results showed that after swapping the NF-κB sites in *attacin*, *diptericin* (site 3) and *drosomycin* (site 2) with *metchnikowin* NF-κB site, the overall activity of these three promoters was significantly higher than the original promoters after overexpression of Rel-RHD, DIF-RHD, Dorsal-RHD or co-expression of any of the two transcription factors (Fig. 9C-E).

On the contrary, swapping *metchnikowin* NF-κB site with the one from *attacin*, *diptericin* (site 3) or *drosomycin* (site 2) significantly decreased the activity of *metchnikowin* promoters after overexpression of NF-κB transcription factors (Fig. 9F-H). These results indicate the importance of NF-κB sites for transcription factor binding and transcriptional activity.

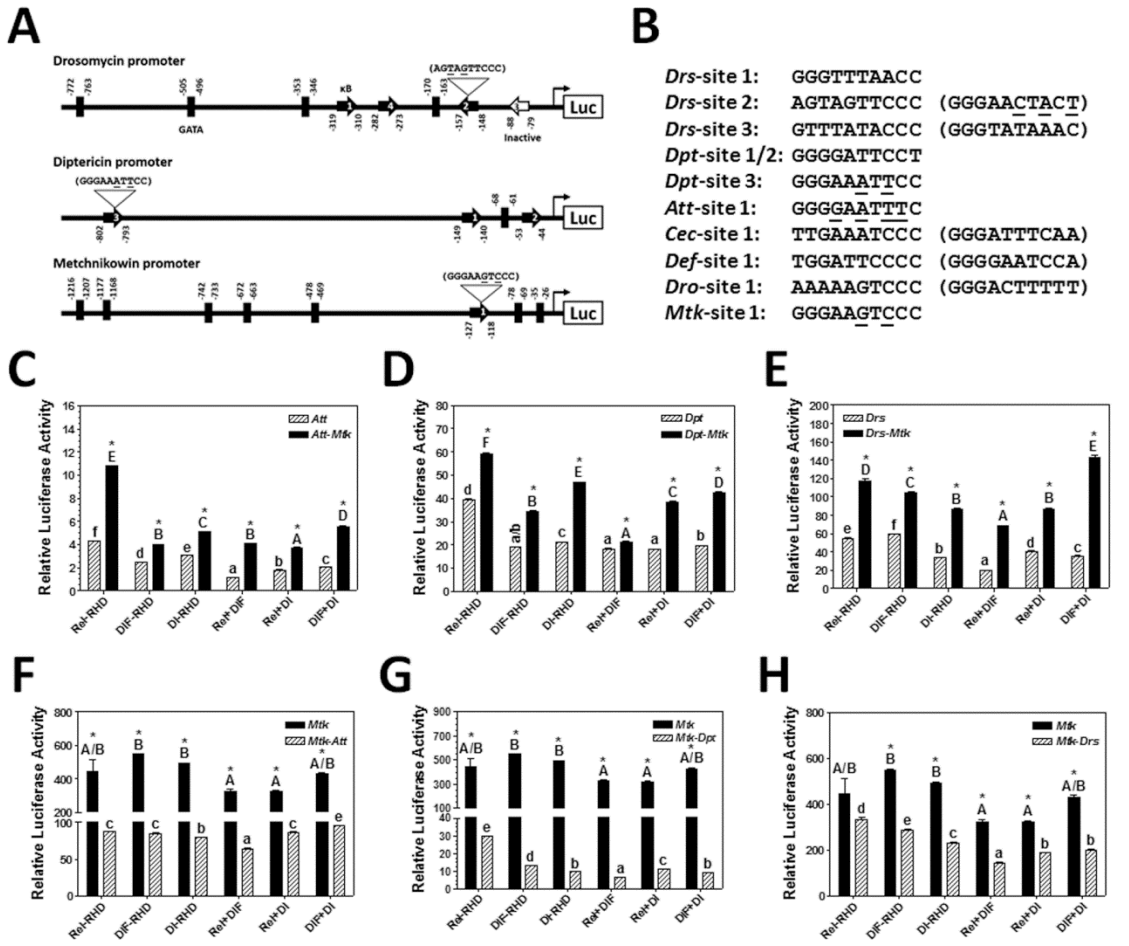


Figure 9. Identification and Verification of NF-κB Binding Sites in AMP Gene Promoters.

(A) Schematic diagrams of the *Drosomycin, Dipterecin* and *Metchnikovin* promoters. (B) Drosophila AMP gene promoter NF-kB site sequences, bracket represents reverse order, underlined alphabets are the bases that were changed. (C-H) Mutational analysis of AMP gene promoters. (C) Solid bars represent *Attacin* kB site changed to *Metchnikovin* kB site; (D) *Dipterecin* kB site changed to *Metchnikovin* kB site; (E) *Drosomycin* kB site changed to *Metchnikovin* kB site; (F) *Metchnikovin* kB site changed to *Attacin* kB site; (G) *Metchnikovin* kB site changed to *Dipterecin* kB site; *Metchnikovin* kB site changed to *Drosomycin* kB site. Promoter substitution mutation promoters are activated by recombinant DmRel-RHD, DmDIF-RHD, DmDl-RHD singly or in combination expressed in S2 cells were determined by Dual-Luciferase® Reporter Assay System as described in the Materials and Methods. Bars represent the mean of three independent measurements ± SEM. For the relative luciferase activity of different promoters activated by NF-kB transcription factors, identical letters are not significant difference ($p>0.05$) while different letters indicate significant difference ($p<0.05$). The significance of difference was also determined by an unpaired t-test (*, $p<0.05$; **, $p<0.01$).

Identification of Nucleotides in NF-κB Binding Sites Important for Transcriptional Activity

NF-κB factors bind to κB DNA elements with a consensus sequence of 5'-GGGRNYYYCC-3' (R is a purine, Y is a pyrimidine, and N is any nucleotide) (Fig. 6A and B). Comparing *metchnikovin* κB site (GGGAAGTCCC) with *diptericin* κB site 3 (GGGAAATTCC) and *drosomycin* κB site 2 (GGGAACTACT) showed that they differ at the 6th and 8th nucleotides and the 6th, 8th and 10th nucleotides (underlined bases), respectively. To investigate whether these bases are critical for transcriptional activity activated by Relish-RHD, DIF-RHD, Dorsal-RHD homo- and hetero-dimers, single or double point mutations were generated in the *metchnikovin* κB site (changed bases to those in the *drosomycin* or *diptericin* promoter) (Fig. 10 B), and dual luciferase assays were performed after overexpression of NF-κB factors. The results showed that when C at the 8th position was changed to G, G and C at the 6th and 8th positions were changed to C and G or to A and T (this double mutations changed the *metchnikovin* κB site to *diptericin* κB site 3), the relative luciferase activity was significantly reduced after overexpression of any NF-κB homo- or hetero-dimers (Fig. 10A). However, when G at the 6th position was changed to C or A, and C at the 8th position was changed to T, the relative luciferase activity was not reduced as significantly as that of the above mutations (Fig. 10 A). We then investigated single base mutations in *drosomycin* κB site 2 and *diptericin* κB site 3 for their involvement in binding to NF-κB factors (Fig. 11 B and D). When C, A and T at the 6th, 8th and 10th positions in *drosomycin* κB site 2 were changed to G, C and C in the *metchnikovin* κB site, respectively, only C to G mutation at the 6th position significantly increased the relative luciferase activity of the

promoter after overexpression of any NF-κB homo- or hetero-dimers (Fig. 11A). Similarly, when A and T at the 6^{th} and 8^{th} positions in the *diptericin* κB site 3 were changed to G and C in the *metchnikovin* κB site, respectively, relative luciferase activity was significantly increased only after overexpression of DIF-RHD/Dorsal-RHD and Relish-RHD/DIF-RHD hetero-dimers, but not after overexpression of other NF-κB homo- or hetero-dimers (Fig. 11C). Together, these results suggest that other than the invariant GGG at the first three nucleotides, other individual bases in the κB sites are also important for binding to NF-κB homo- and hetero-dimers.

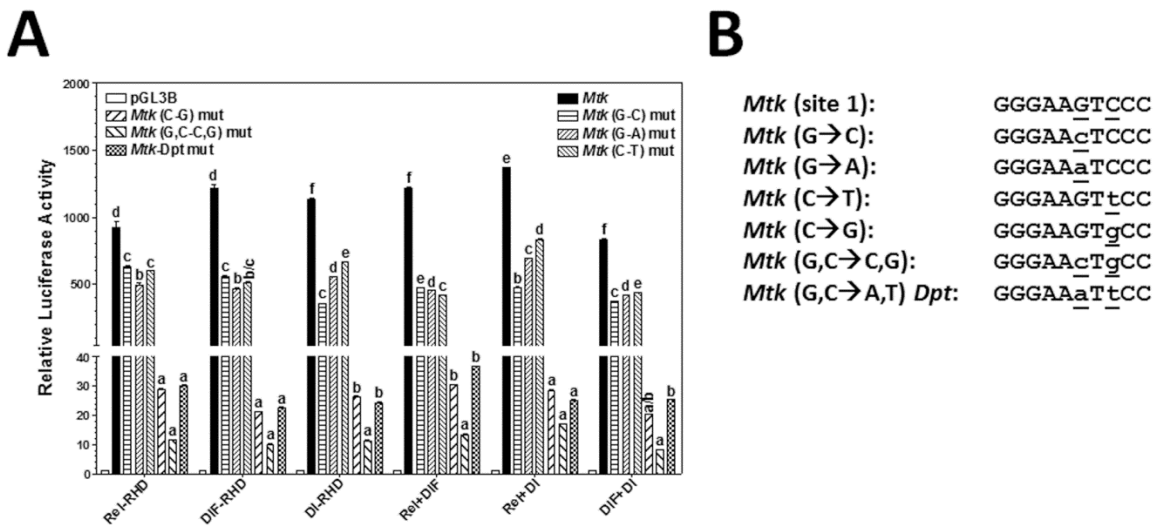


Figure 10. Substitution Mutation of *Metchnikovin* kB Site.

(A) *Metchnikovin* promoter substitution mutations are activated by recombinant DmRel-RHD, DmDIF-RHD, DmDl-RHD singly or in combination expressed in S2 cells were determined by Dual-Luciferase® Reporter Assay System as described in the Materials and Methods. Bars represent the mean of three independent measurements ± SEM. For the relative luciferase activity of different promoters activated by NF-kB transcription factors, identical letters are not significant difference ($p>0.05$) while different letters indicate significant difference ($p<0.05$). The significance of difference was also determined by an unpaired t-test (*, $p<0.05$; **, $p<0.01$). (B) *Metchnikovin* promoter kB site with underlined bases representing substituted bases.

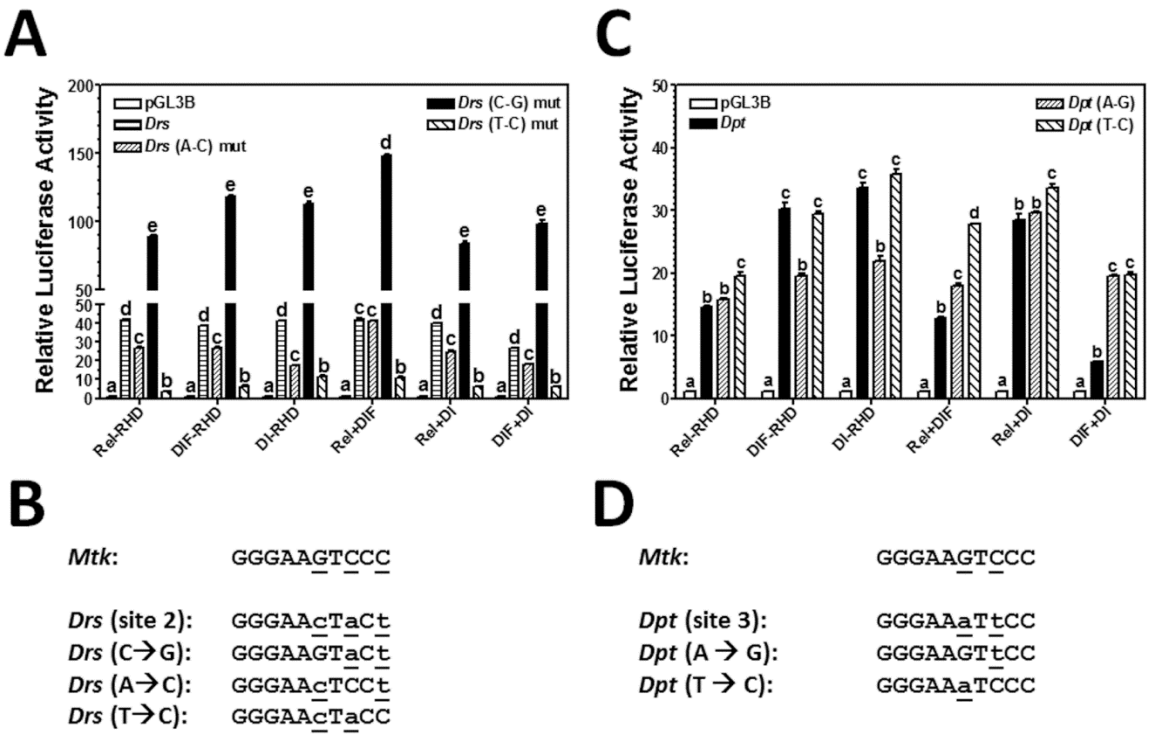


Figure 11. Substitution Mutation of *Drosomycin* and *Diptericin* kB Site.

(A) *Drosomycin* kB site 2 promoter substitution mutations (C) *Diptericin* kB site 3 promoter substitution mutations. Both *Drosomycin* and *Diptericin* kB site mutant promoters are activated by recombinant DmRel-RHD, DmDIF-RHD, DmDl-RHD singly or in combination expressed in S2 cells were determined by Dual-Luciferase® Reporter Assay System as described in the Materials and Methods. Bars represent the mean of three independent measurements ± SEM. For the relative luciferase activity of different promoters activated by NF-kB transcription factors, identical letters are not significant difference ($p>0.05$) while different letters indicate significant difference ($p<0.05$). The significance of difference was also determined by an unpaired t-test (*, $p<0.05$; **, $p<0.01$). (B) *Drosomycin* kB site 2 promoter; (D) *Diptericin* kB site 3 promoter with underlined bases representing substituted bases.

Discussion

In order to form heterodimers between Relish and DIF or Dorsal in vivo, Relish has to be cleaved to release the active N-terminal part (Relish-N or Rel-N) simultaneously with the release of DIF or Dorsal from the inhibitor Cactus. In this case, the Toll and Imd pathways have to be activated at the same time. Toll and IMD are two major signaling pathways that are activated during humoral responses, and they regulate the activation of *Drosophila* NF-κB proteins Dorsal, DIF, and Relish (Ganesan et al., 2011). There are seven known classes of inducible antimicrobial peptides (AMPs) in *Drosophila*: *Attacin, Cecropin, Defensin, Diptericin, Drosocin, Drosomycin, Metchnikowin* (Bulet et al., 1991; Hetru et al., 2003; Yi et al., 2014). Several of the antimicrobial peptides work by disrupting bacterial membranes (Almeida et al., 2009; Wimley et al., 2011). Antimicrobial peptides are released by the fat body of the flies, a functional analog of the mammalian liver, within hours of an immune challenge. The AMPs are then secreted into the hemocoel where they block the proliferation of microorganisms (Bulet et al., 1999). NF-κB transcription factors play a central role in both mammalian and insect innate immune responses. Because of the significant conservation in the signaling pathways responsible for NF-κB activation, the fruit fly *Drosophila melanogaster* has been an attractive model to study these transcription factors (Hoffmann 2003; Khush et al.,). Signaling pathways are complicated and this complexity is more pronounced with the interplay of isoforms as means of cross-regulators. Isoforms are produced when one gene makes several proteins by alternative splicing or different proteins with similar amino acid sequence and function comes into existence. In our study, we discovered a new relish isoform in *Drosophila*, which is

always activated due to the absence of inhibitory ankyrin repeats. Cloning and protein expression of this short-relish in S2 cells also indicate the absence of homo-dimerization unlike relish-N'. However, due to the presence of similar rel homology domain, it could interact similarly with both dif and dorsal transcription factors. The peculiarity of this short-relish isoform lies in its highest activation of anti-fungal genes, which clearly indicates that it might play a major role against fungal infection and might act as a cross-regulatory protein between both Toll and IMD pathways. Recently, a new IkB protein pickle has been discovered that inhibits genes activated by homodimerization of relish. This opens a further research on regulators for short-relish isoform. It would be interesting to check if the relish/IMD pathway negative regulators could inhibit this short relish activity. Sequence comparison suggests 99% similarity with relish-N' except the last 20 bases in C-terminal region. Furthermore, tissue distribution and induced expression shows that short-relish is present in all the developmental stages and could be induced by both bacteria and fungi. Nevertheless, to activate AMP genes NF-kB transcription factor must bind to a variation of consensus DNA sequence of 5′-GGGRNYY YCC-3′ (in which R is a purine, Y is a pyrimidine, and N is any nucleotide) known as κB sites (Chen et al., 1998) on the AMP gene promoter to recruit RNA polymerase and initiate transcription. Research from different labs has predicted the NF-kB binding sites on all the seven Drosophila AMP genes (Busse et al., 2007; Engstrom et al., 1993; Kappler et al., 1993; Levashina et al., 1998; Tanji et al., 2007a). Because our luciferase results gave very high activity of anti-fungal gene *Metchnikovin*, we switched the kB site of *Attacin, Diptericin* and *Drosomycin* genes with single kB site of *Metchnikovin* and found an expected increase in all genes.

So, in later experiments we tried to figure out the key bases that are responsible for the gene expression upon NF-kB binding. Sequence analysis shows that *Diptericin* kB site1 differ in 6^{th} and 8^{th} positions and *Drosomycin* kB site2 differ in 6^{th}, 8^{th} and 10^{th} positions compared to *Metchnikovin*. Mutational analysis found that in *Drosomycin* kB binding site2, the 6^{th} and 10^{th} position bases are important, whereas in *Dipterecin* kB site 1, the 6^{th} position base is crucial for AMP regulation.

In conclusion, our research discovered a new *Drosophila* short-relish isoform that lacks the inhibitory ankyrin repeats and thus is an active NF-kB transcription factor that could be involved in antifungal and antibacterial infection. Presence of such active relish that doesn't form homo-dimers opens the possibility of negative regulators that could exist for regulating heterodimers for maintaining beneficial symbionts inside insect body.

CHAPTER 4

TRANSCRIPTION FACTOR FORKHEAD REGULATES EXPRESSION OF ANTIMICROBIAL PEPTIDES IN THE TOBACCO HORNWORM, *MANDUCA SEXTA*

Abstract

Antimicrobial peptides (AMPs) play an important role in defense against microbial infections in insects. Expression of AMPs is regulated mainly by NF-κB factors Dorsal, Dif and Relish. Our previous study showed that both NF-κB and GATA-1 factors are required for activation of moricin promoter in the tobacco hornworm, *Manduca sexta*, and a 140-bp region in the moricin promoter contains binding sites for additional transcription factors. In this study, we identified three forkhead (Fkh)-binding sites in the 140-bp region of the moricin promoter and several Fkh-binding sites in the lysozyme promoter, and demonstrated that Fkh-binding sites are required for activation of both moricin and lysozyme promoters by Fkh factors. In addition, we found that Fkh mRNA was undetectable in *Drosophila* S2 cells, and *M. sexta* Fkh (MsFkh) interacted with Relish-Rel-homology domain (RHD) but not with Dorsal-RHD. Dual luciferase assays with moricin mutant promoters showed that co-expression of MsFkh with Relish-RHD did not have an additive effect on the activity of moricin promoter, suggesting that MsFkh and Relish regulate moricin activation independently. Our results suggest that insect AMPs can be activated by Fkh factors under non-infectious conditions, which may be important for protection of insects from microbial infection during molting and metamorphosis.

Introduction

In insects, innate immunity is the first line of defense against pathogens and it is composed of both humoral and cellular immune responses (Bangham et al., 2006; Ferrandon et al., 2007b; Hultmark et al., 2007; Jiang et al., 2010; Lemaitre et al., 2007). Synthesis of small cationic antimicrobial peptides (AMPs) in the fat body of insects is an important defense mechanism of humoral immune responses against microbial infection (Al Souhail et al., 2016; Cao et al., 2016; Hanson et al., 2016; Zhang et al., 2015). Expression of insect AMPs is regulated mainly by the evolutionarily conserved Toll and immune deficiency (IMD) pathways (Hou et al., 2014; Imler et al., 2005; Shi et al., 2012; Zhong et al., 2012a). The Toll pathway defends against Gram-positive bacteria, fungi and viruses (Michel et al., 2001b; Zambon et al., 2005), whereas the IMD pathway acts against Gram-negative bacteria (Gottar et al., 2002). Both the Toll and IMD pathways activate the Rel/NF-κB family of transcription factors, including Dorsal, Dif (Dorsal-related immunity factor) (Busse et al., 2007) and Relish (Choe et al., 2002). These NF-κB factors contain N-terminal Rel-homology domain (RHD) that is required for DNA binding and dimerization (Hetru et al., 2009).

In our previous study with the moricin promoter in the tobacco hornworm *Manduca sexta*, we found that both NF-κB and GATA-1 binding sites are crucial for activation of moricin promoter, and we also identified a 140-bp region (between -240 and -100 bp) in the moricin promoter, designated as moricin promoter activating element (MPAE), which contains binding sites for additional transcription factors required for activation of moricin promoter (Rao et al., 2011). The purpose of this study is to identify the additional transcription factor(s) that can bind to the MPAE region,

investigate whether the new factor(s) can also activate other AMP genes, and whether the new factor(s) acts independently or cooperatively with NF-κB factors. Sequence analysis of the 140-bp MPAE region showed a binding site for silk gland factor-1 (SGF-1), a transcription factor in the silk worm *Bombyx mori* that is homologous to *Drosophila melanogaster* forkhead (Fkh). SGF-1 can bind *in vitro* to a *cis*-element located in the promoter region of *sericin-1*, which encodes a silk protein in the middle silk gland (MSG) cells (Hui et al., 1990; Mach et al., 1995; Xu et al., 1994). Further analysis showed that there are three Fkh-binding sites in the 140-bp MPAE region (Fig. 1A).

Forkhead transcription factors belong to the Fox (Forkhead box) superfamily proteins (Kaestner et al., 2000). Members of the Fox family proteins consist of a conserved long DNA binding domain connected to a pair of loops or "wings" via a small β-sheet (Clark et al., 1993). The long DNA binding domain, also known as winged-helix or forkhead domain, is about 90 residues and composed of three α-helices with a helix–turn–helix core, thus, the Fox family proteins are also known as winged helix transcription factors (Hannenhalli et al., 2009). Fox proteins have been identified in eukaryotic organisms from yeast to human, and they play important roles in various biological processes, including cellular differentiation, development, metabolism, insulin signaling, immune regulation, cancer development, and aging (Carlsson et al., 2002; Coffer et al., 2004; Jonsson et al., 2005; Jünger et al., 2003; Kaufmann et al., 1996; Lehmann et al., 2003b; Takahashi et al., 1997). Forkhead gene A (*FoxA*) was first discovered in *D. melanogaster* since mutation in this gene resulted in fork-headed structure in the embryos with defects in the anterior and posterior gut formation

(Weigel et al., 1989). Later, cDNA encoding hepatocyte nuclear factor 3α (HNF3α), also known as FoxA1, was cloned in rat (Lai et al., 1990) . The central 90-residue DNA-binding domain (forkhead domain) of D. melanogaster FoxA and human FoxA1 show shockingly high (>85%) identity (Weigel et al., 1990). Since the discovery of *Drosophila FoxA*, hundreds of Fox genes have been identified in various species. There are at least 4 Fox genes in yeast, 16 in *D. melanogaster*, 15 in *Caenorhabditis elegans*, 44 in mice, and 50 in humans (Mazet et al., 2003; Tuteja et al., 2007a; b). The genome of *Aedes aegypti* contains eighteen loci that encode putative Fox factors and six of them are involved in reproduction (Hansen et al., 2007). Based on phylogenetic analysis, metazoan Fox family is divided into 19 subfamilies (FoxA-S) (Kaestner et al., 2000). Human Fox subfamilies are classified into two classes based on the sequence homology within and beyond the forkhead domain: Fox A-G, I-L and Q are class 1, Fox H and M-P are class 2 Fox proteins (Kaestner et al., 2000).

In *D. melanogaster*, several Fox transcription factors have been described. Crocodile (FoxC) is involved in early patterning in embryo (Hacker et al., 1995), Biniou (FoxF) plays a role in visceral mesoderm (Zaffran et al., 2001), sloppy paired 1 & 2 (FoxG) are involved in early embryo segmentation (Cadigan et al., 1994; Grossniklaus et al., 1992), domina (FoxN) is involved in vitality and fertility (Strodicke et al., 2000), and FoxO regulates the insulin signaling in the brain and fat body (Hwangbo et al., 2004). *Drosophila* FoxO is also involved in innate immunity and expression of AMPs (Becker et al., 2010; Varma et al., 2014). Forkhead (Fkh) is the founding member of the FoxO family and it is activated upon TOR inhibition by

rapamycin (Varma et al., 2014). However, functions of Fox transcription factors in regulation of AMPs in other insect species have yet to be reported.

In this study, we identified Fkh-binding sites in the promoters of *M. sexta* moricin and lysozyme genes, cloned Fkh genes from both *M. sexta* and *D. melanogaster*, and demonstrated activation of AMP gene promoters by Fkh factors. We also found that Fkh mRNA was undetectable in *Drosophila* S2 cells, which may account for low activity of some AMP promoters in S2 cells. In addition, co-expression study showed that *M. sexta* Fkh (MsFkh) interacted with Relish (Rel2)-RHD, but did not show any interaction with Dorsal-RHD. However, co-expression of MsFkh and MsRelish-RHD did not have an additive effect on the activity of moricin promoter compared to expression of MsFkh or MsRelish-RHD alone, suggesting that MsFkh and MsRelish regulate AMP gene expression independently. Activation of AMPs by Fkh factors under non-infectious conditions is particularly important for insects during molting and metamorphosis to protect them from microbial infections.

Methods

Manduca sexta and Insect Cell Lines

We purchased *M. sexta* eggs from Carolina Biological Supplies (Burlington, NC, USA) and reared larvae to fifth-instar on an artificial diet at 25°C (Dunn et al., 1983) for all the experiments. *D. melanogaster* Schneider S2 cells were obtained from American Type Culture Collection (ATCC), and *Spodoptera frugiperda* Sf9 cells were from Invitrogen (12552-014, Invitrogen). Cells were maintained at 27°C in Insect Cell Culture Media (SH30610.02, Hyclone) supplemented with 1% penicillin-streptomycin solution (G6784, Sigma-Aldrich) and 10% heat-inactivated fetal bovine serum (#10082063, Invitrogen).

Electrophoretic Mobility Shift Assay (EMSA)

Cytosolic and nuclear proteins were isolated from S2 and Sf9 cells using the Nuclear Extraction Kit (2900, EMD Millipore) following the manufacturer's instructions. Briefly, S2 and Sf9 cells with 70-80% confluency were collected (2×10^8 cells) and homogenized in 500 µl of 1× Cytoplasmic Lysis Buffer with 27-gauge needle and centrifuged at 8,000×g for 20 min at 4°C. The supernatants containing the cytosolic proteins were transferred to fresh tubes and stored at -80°C for later use, whereas the pellets were resuspended in 100 µl of ice-cold Nuclear Extraction Buffer containing 0.5 mM DTT and 1/1000 Protease Inhibitor Cocktail. The nuclei were disrupted using 27-gauge needle, the mixture was incubated at 4°C for 60 min with gentle agitation, and

then centrifuged at 16,000×g for 5 min at 4°C. These supernatants containing nuclear proteins were removed to fresh tubes and stored at -80°C for later use.

Electrophoretic mobility shift assay (EMSA) was performed using the LightShift® Chemiluminescent EMSA Kit (20148, Thermo Scientific) following the manufacturer's instructions. Genomic DNA fragments from the 140-bp MPAE region of moricin promoter (MPAE-1, -2 and -3, as well as MPAE-3a, -3b and -3c) were synthesized with or without covalently linked biotins by Integrated DNA Technologies (IDT) (San Diego, CA). EMSA was performed in 20 μl reactions containing 2 μl of biotinylated DNA (20 fmole) with 3 μg of cytosolic or nuclear proteins from S2 or Sf9 cells, in the absence or presence of unlabeled DNAs (200-fold, or from 1-, 10-, 100- to 200-fold molar excess of the biotinylated DNA). The mixtures were incubated at room temperature for 30 min, separated on 6% polyacrylamide gel (EC6365BOX, Invitrogen) pre-run in 0.5× TBE at 100V for 60 min. The gels were then transferred to nylon membrane pre-soaked in 0.5 × TBE at 380 mA for 30 min in ice-cold 0.5 × TBE, and the membranes were crosslinked for 1 min using a UV-light crosslinking instrument. The biotinylated DNA was detected by Streptavidin-Horseradish Peroxidase conjugated anti-mouse antibody (SC-2005, Santa Cruz Biotechnology, 1:10,000) by chemiluminescence using ECL Chemiluminescence Detection Kit (RPN2134, GE Healthcare), and the membranes were exposed to films and scanned by Typhoon FLA7000 (GE healthcare).

Analyses of Transcripts of *M. sexta* Forkhead (MsFkh) in Larvae and *D. melanogaster* Fox Genes in S2 Cells

There are two cDNA sequences for *D. melanogaster* forkhead (transcript variant A) (DmFkh) in the database (Genbank accession numbers: J03177.1 and NM_079818.3) with identical open reading frame (ORF) of 1530 bp, which encodes 510 amino acid of DmFkh. To identify *M. sexta* forkhead, we used the ORF of DmFkh to blast the *M. sexta* database (http://agripestbase.org/manduca/?q=blast) and obtained one sequence [Msex2.13928-RA, scaffold01234:21957-23009(-)], which is significantly similar to DmFkh (E-value = e^{-22}). This *M. sexta* cDNA sequence contains an ORF of 1065 bp, encoding a protein of 355 amino acids with a forkhead domain. We named this protein *M. sexta* Fkh (MsFkh).

To determine tissue distribution of *MsFkh* mRNA, day 2 fifth-instar *M. sexta* naïve larvae were dissected. Hemocytes, fat body, midgut, epidermis and testis were collected and washed 3 times in anti-coagulant (AC) saline (4 mM NaCl, 40 mM KCl, 8 mM EDTA, 9.5 mM citric acid-monohydrate, 27 mM sodium citrate, 5% sucrose, 0.1% polyvinylpyrrolidone, 1.7 mM PIPES). Total RNAs were extracted from these tissues with TRIzol® Reagent (T9424, Sigma-Aldrich) and cDNAs were prepared from total RNAs (1 μg for each sample) using moloney murine leukemia virus (M-MLV) reverse transcriptase (M1701, Promega) with an anchor-oligo$(dT)_{18}$ primer following the manufacturer's instructions as described previously (Zhong et al., 2016b).

To determine induced expression of *MsFkh* transcript in *M. sexta* larvae, day 2 fifth-instar naïve larvae were injected with heat-killed *Staphylococcus aureus*, *Bacillus subtilis*, *Escherichia coli* strain XL1-blue, *Serratia marcescens* (each at 5×10^7

cells/larva), or *Saccharomyces cerevisiae* (10^7 cells/larva), or with water as a control. Hemocytes, fat body and midgut were collected separately at 24 h post-injection for total RNA extraction and cDNA preparation as described previously (Rao et al., 2014a). Total RNA and cDNA were also prepared from *Drosophila* S2 cells. Briefly, S2 cells (5×10^6 cells) were collected in 1 ml of TRIzol® Reagent (T9424, Sigma-Aldrich) and homogenized using hand held pestle and mixer (Argos Technologies, Elgin, IL). Then, 200 µl of chloroform were added, and the mixture was vortexed and centrifuged at 12000×g at 4°C for 15 min. The top aqueous phase (200 µl) was transferred to fresh Eppendorf tubes and isopropanol (500 µl) was added. RNA was precipitated by centrifugation at 12,000×g at 4°C for 10min. The RNA pellets were washed with chilled 70% ethanol, air dried and re-suspended in 50 µl of nuclease free water and stored at -80 °C for later use.

Real-time PCR was performed for *MsFkh* (primers MsFkh-N and MsFkh-C, amplicon size of 151 bp) in different tissues of *M. sexta* larvae and several *Drosophila* Fox genes in S2 cells, including *DmFkh* (Genbank accession no. NP_524542.1) (primers DmFkh-N and DmFkh-C, 150 bp), forkhead box K (*FoxK*) long and short isoforms (Genbank accession no. AY787838) (primers DmFoxK-(L+S)-N and DmFoxK-(L+S)-R for both isoforms, primers DmFoxK-(L)-N and DmFoxK-(L)-R for long isoform only, amplicon sizes of 150 bp), which are also known as forkhead transcription factor long isoform (mnf-l) and short isoform (mnf-s), *jumu* (Genbank accession no. NM_079578.3) (primers DmJumu-N and DmJumu-R, 150 bp) and dFoxO (Genbank accession no. NM_206483.3) (primers dFoxO-N and dFoxO-R, 150 bp), in 20 µl reactions containing 10 µl 2×SYBR® GreenER™ qPCR SuperMix

Universal (No. 204141, Qiagen), 4 μl H_2O, 4 μl diluted cDNA template, and 1 μl forward and reverse primers (10 pmol each), and *M. sexta* ribosomal protein S3 (*rpS3*) (primers rpS3-N and rpS3-C, 150 bp) or *D. melanogaster* ribosomal protein 49 (*rp49*) (primers rp49-N and rp49-C, 150 bp) gene was used as an internal standard to normalize the amount of RNA template. Real-time PCR program was 50°C for 2 min, 95°C for 10 min, followed by 40 cycles of 95°C for 15 s, 60°C for 1 min and the dissociation curve analysis. Data from three replicas of each sample were analyzed by the ABI 7500 SDS software (Applied Biosystems) using a comparative method ($2^{-\Delta\Delta CT}$) (Livak et al., 2001; Pfaffl 2001). These experiments were repeated with three different biological samples.

Construction of Recombinant MsFkh and DmFkh pAC5.1/V5-His A Expression Vectors

Recombinant MsRelish-RHD (MsRel2-RHD) and MsDorsal-RHD (MsDl-RHD) with a Flag-tag in pAC5.1/V5-His A expression vectors were already constructed as described previously (Zhong et al., 2016b). To construct V5-tagged Fkh into pAC5.1/V5-His A expression vector, cDNA fragments encoding MsFkh (residues 1-355) and DmFkh (residues 1-510) were amplified by PCR using forward and reverse primers (Table S1). Forward primers for MsFkh and DmFkh (MsFkh-F-*Kpn* I and DmFkh-F-*Kpn* I) contain a *Kpn* I site and the reverse primers (MsFkh-R-*Not* I and DmFkh-R-*Not* I) contain a *Not* I site in-frame fused to V5 tag and stop codon of pAC5.1/V5-His A expression vector. PCR reactions were performed with the following conditions: 94°C for 3 min, 35 cycles of 94°C for 30 s, Tm-5°C for 30 s, 72°C for 30 s

to 4 min, followed by a final extension at 72°C for 10 min. The PCR products were recovered by agarose gel electrophoresis-Wizard® SV Gel and PCR Clean-Up System (A9285, Promega) and subcloned into T-Easy vectors (A1360, Promega). Recombinant T-vectors were purified using PureYield™ Plasmid Miniprep System (A1222, Promega) according to the manufacturer's instruction and digested with *Kpn* I/*Not* I restriction enzymes, and DNA fragments were recovered and inserted into *Kpn* I/*Not* I digested pAC5.1/V5-His A expression vector using T4 DNA ligase (M0202L, NEB). Recombinant plasmids were then purified and sequenced by an Applied Biosystems 3730 DNA Analyzer in the DNA Sequencing and Genotyping Facility at University of Missouri – Kansas City, and used to transfect S2 cells.

Construction of Luciferase Reporter Plasmids

The moricin and lysozyme truncated promoters were constructed as described previously (Rao et al., 2011; Zhong et al., 2016b). To construct Fkh-binding site mutation promoters, site-directed mutagenesis was performed using the truncated *M. sexta* moricin-242 (242 bp) and lysozyme-345 (345 bp) promoters as templates. Primers with specific mutation sites were designed for each mutated promoter and listed in Table S1. There are three and four predicted Fkh-binding sites in moricin-242 and lysozyme-345 promoters, respectively. Primers Lyz-D3-Fkh-1, Lyz-D3-Fkh-2, Lyz-D3-Fkh-3 and Lyz-D3-Fkh-4 (Table S1) were used to generate mutations of Fkh-binding site 1, 2, 3 and 4, respectively, in lysozyme-345 promoter. Whereas primers MPAE-Fkh-1, MPAE-Fkh-2 and MPAE-Fkh-3 were used to generate mutations of Fkh-binding site 1, 2 and 3, respectively, in moricin-242 promoter. To pre-screen

positive colonies prior to DNA sequencing, restriction enzyme cleavage sites of *Eco*R I, *Nde* I and *Bam* HI were engineered in the mutant Fkh-binding site 1, 2 and 3 of the moricin-242 promoter, respectively. To generate $Mor^{mut-1\&2}$, $Mor^{mut-1\&3}$, $Mor^{mut-2\&3}$ and $Mor^{mut-1,2\&3}$ mutant promoters, site-directed mutagenesis was performed using the mutant promoter as the template with the second pairs of primers. For example, to generate $Mor^{mut-1\&2}$ promoter, Mor^{mut-1} was used as the template with MPAE-Fkh-2 primers, and to obtain $Mor^{mut-1,2\&3}$ promoter, $Mor^{mut-1\&2}$ was used as the template with MPAE-Fkh-3 primers. PCR program was 95°C for 3 min, and then 17 cycles of 95°C for 1 min, 55°C for 2 min, 68°C for 15 min, followed by a final extension of 68°C for 30 min. The PCR products were recovered, digested with *Dpn* I, and then transformed into competent *E. coli* XL1 Blue cells. The mutant reporter plasmids were then purified and sequenced by an Applied Biosystems 3730 DNA Analyzer in the DNA Sequencing and Genotyping Facility at University of Missouri – Kansas City, and used for transient transfection in S2 cells.

Dual-Luciferase Reporter Assays

For DNA transfection, S2 cells were placed overnight to 70% confluence prior to transfection in serum-free medium (SH30278.01, Hyclone). GenCarrier-1™ transfection reagent (#31-00110, Epoch Biolabs) was used for transient transfection according to the manufacturer's instructions. After overnight transfection, S2 cells were centrifuged and resuspended in complete growth medium to induce protein expression for 48 h. Protein expression in cell culture media and cell extracts were analyzed by immunoblotting.

Dual-luciferase reporter assays in S2 cells were performed in 96-well culture plates with recombinant pAC5.1/V5-His A expression plasmid (0.3 μg), pGL3B (empty vector) or different pGL3B firefly luciferase reporter plasmids from the promoters of *M. sexta* moricin, lysozyme, cecropin, defensin-1, defensin-2, defensin-3, attacin-1, and attacin-2 (See Fig. S1 for the promoter sequences), as well as several mutant moricin or lysozyme promoters (0.15 μg), and renilla luciferase reporter plasmid (0.015μg) (as an internal standard) (pRL-TK, Promega) as described previously (Rao et al., 2011). Firefly luciferase and renilla luciferase activities were measured at 48 h after protein expression using the Dual-Luciferase Reporter Assay System (E1980, Promega) in the GloMax® Multi Microplate Luminometer (Promega). Relative luciferase activity (RLA) from S2 cells co-transfected with empty pAC5.1/V5-His A and pGL3B (empty reporter vector) was used as the calibrator. Relative luciferase activity (RLA) was obtained as the ratio of firefly luciferase activity to renilla luciferase activity (Agrawal et al., 2013). These assays were performed in quadruplet and three independent experiments were repeated.

Immunoblotting Analysis and Co-immunoprecipitation (Co-IP) Assay

For immunoblotting analysis, cell extracts from S2 cells (2×10^6 cells/well) expressing MsFkh, MsRel2-RHD and MsDorsal-RHD proteins were prepared as described previously (Zhong et al., 2012a). Cell culture media (10 μl each) and cell extracts (10 μl each, equivalent to ~5×10^4 cells) were separated on 10%, 12%, or 15% SDS-PAGE and proteins were transferred to nitrocellulose membranes (162-0097, Bio-Rad). Anti-Flag M2 antibody (F-1804, Sigma-Aldrich, 1:5000 dilution) and anti-V5

antibody (V-8012, Sigma-Aldrich, 1:5000 dilution) were used as primary antibodies, horseradish peroxidase-conjugate anti-mouse antibody (SC-2005, Santa Cruz Biotechnology, 1:10,000) was used as secondary antibody for chemiluminescence using ECL Chemiluminescence Detection Kit (RPN2134, GE Healthcare), and membranes were exposed to films and scanned by Typhoon FLA7000 (GE healthcare), while alkaline phosphatase (AP)-conjugate anti-mouse antibody (A4312, Sigma-Aldrich, 1:10,000) was used as secondary antibody for color development using AP-conjugate color development Kit (#170-6432, Bio-Rad).

Co-immunoprecipitation (Co-IP) assays were performed using cell extracts from S2 cells overexpressing MsFkh, MsRel2-RHD and MsDorsal-RHD proteins. Cell extracts were mixed and Co-IP was performed as described previously (Zhong et al., 2012a). Proteins immunoprecipitated with anti-Flag M2 or anti-V5 primary antibody were captured by protein G Sepharose pre-swollen beads (#17-0618-01, GE Healthcare). Captured proteins were eluted with 30 μl of sample buffer mixed with 0.1% bromphenol blue, heated to 95ºC for 3 minutes, centrifuged at 12,000×g for 1 min, and the supernatants were loaded onto SDS-PAGE for immunoblotting analysis as described previously (Zhong et al., 2012a).

Data Analysis

All the experiments were performed in 3-4 replicates and repeated with three independent biological samples. The means of a typical set of data were used to prepare the figures by GraphPad Prism (GraphPad, San Diego, CA). Statistical significance was calculated by one way ANOVA followed by a Tukey's multiple comparison tests using

GraphPad Prism for comparisons of *MsFkh* mRNA in different tissues, or *MsFkh* mRNA in hemocytes, fat body or midgut by different treatments (Fig. 3), expression levels of *D. melanogaster* Fox genes in S2 cells (Fig. 4A), or relative luciferase activity across different promoters by overexpression of DmFkh or MsFkh (Figs. 4, 5 and 7), and identical letters are not significant difference ($p>0.05$) while different letters indicate significant difference ($p<0.05$). The significance of difference was also determined by an unpaired t-test with the GraphpadInStat software (*, $p<0.05$; **, $p<0.01$) to compare the activity of a promoter stimulated by different transcription factors.

Primers	Forward Primer (5'-3')
Cloning	
MsFkh-F-Kpn I	**ACGGGGTACCATGATCTCGCAGAAGTTATCGTACGGCG**
DmFkh-F-Kpn I	**ACGGGGTACCATGACCAGACCACCATTGACATCATGCAGAAG CTCTACGCGG**
Lyz-D3-Fkh-1-N	**CTACGATATTTTAGCGGCCGCTAATTGATTTAATCATTAGC**
Lyz-D3-Fkh-2-N	**AGTATTTAAAAACGCGGCCGCAAGACATTAATGTAATCTGG**
Lyz-D3-Fkh-3-N	**AAGCGCATCACTGGCGGCCGCTAAACAAACAGATACGCAGG**
Lyz-D3-Fkh-4-N	**CTGAAACATAAACGCGGCCGCGATACGCAGGAACCCCCTTT**
MPAE-Fkh-1-F	**TGTATATGTATAGAATCCTGGCAGATTATAATATGAATG**
MPAE-Fkh-2-F	**TAGGAATAGGTACATATGTCGACTAGTATACGTGTTACG**
MPAE-Fkh-3-F	**CGCGTTTAATAAGGATCCTAATATAAACGCTTATGACAA**
Cloning	**Reverse Primer (5'-3')**
MsFkh-R-Not I	**AAGAATGCGGCCGCGAATGCGGCCGCTTCAAGGGCGGCTGC**
DmFkh-R-Not I	**AAGAATGCGGCCGCTTCAAGCTCGTGGTTCCGGCGG**
Lyz-D3-Fkh-1-C	**AATCAATTAGCGGCCGCTAAAATATCGTAGGCTTACTAGC**
Lyz-D3-Fkh-2-C	**TAATGTCTTGCGGCCGCGTTTTTAAATACTGTAGCTATGG**
Lyz-D3-Fkh-3-C	**GTTTGTTTAGCGGCCGCCAGTGATGCGCTTTGTTAATAAC**
Lyz-D3-Fkh-4-C	**CTGCGTATCGCGGCCGCGTTTATGTTTCAGTGATGCGCTT**
MPAE-Fkh-1-R	**ATAATCTGCCAGAATTCTATACATATACATTTTAATTTAAAA G**
MPAE-Fkh-2-R	**GTATACTAGTCGACATATGTACCTATTCCTAATCATTCATAT**
MPAE-Fkh-3-R	**GCGTTTATATTAGGATCCTTATTAAACGCGACCCGTAACAC**
Real-time PCR	
MsFkh-N	**GAGCCAAGCGGGTACGCG**
DmFkh-N	**ATGTCGGCAGCGAGTATGTC**
DmFoxK (L+S)-N	**GCTCACCGCTCCAAGATA**
DmFoxK (L)-N	**CAGCATCGTTGTGGCCCC**
DmFoxN-N	**ATGCCAACGGAAACCAGG**
dFoxO-N	**GGCCACGGTCAACACGAA**
rp49-N	**GCCCAAGGGTATCGACAACA**

(*Drosophila*)	
rpS3-N (*M. sexta*)	**CCCGTGGGACCAACAGGG**
MsFkh-C	**GGTACAGGGGCGACTGGT**
DmFkh-C	**ATGCGTGTAGCTCCTTCTGT**
DmFoxK (L+S)-R	**TCTGCTGCTGCTGCTGCT**
DmFoxK (L)-R	**TTGGCAGCGCTTATTGTA**
DmFoxN-R	**TTCGAGATAATTCGACGC**
dFoxO-R	**TGCTGCTGCTGCTGCTGT**
rp49-C	**ACCTCCAGCTCGCGCACGTT**
rpS3-C	**GCGGCGACCGGCTGCGGC**

Table 1. Primers Used to Clone *Manduca* Fkh, *Drosophila* Fkh and for qPCR.

Results

Binding of Nuclear Proteins in Sf9 Cells to MPAE of *M. sexta* Moricin Promoter

In our previous study, we identified a 140-bp MPAE region in the *M. sexta* moricin, which was responsible for activation of moricin and *D. melanogaster* drosomycin promoters in *S. frugiperda* Sf9 cells (Rao et al., 2011). To identify nuclear factor(s) in Sf9 cells that can bind to the MPAE region, we divided the 140-bp MPAE region into three fragments (MPAE-1 to 3) (Fig. 1A), and performed electrophoretic mobility shift assay (EMSA). EMSA results showed that proteins from nuclear extracts of Sf9 cells, but not S2 cells, bound to all three MPAE fragments (Fig. 1B, lanes 5, 10 and 14, arrow). We then analyzed transcription factor binding sites (AliBaba 2.1 program, http://www.gene-regulation.com) in the three MPAE fragments, and found that there is an SGF-1 binding site in MPAE-3 (Fig. 1A). SGF-1 is a transcription factor in the silkworm *B. mori* that regulates expression of silk protein Sericin-1 in the middle silk gland cells (Mach et al., 1995; Xu et al., 1994), and it is a homolog of *D. melanogaster* forkhead (Fkh). Fkh is a member of the FoxO family proteins and FoxO binds to the consensus sequence of (T/C)(G/A)AAACAA (Li et al., 2004). We found an ATAAACA sequence in both MPAE-2 and MPAE-3, and an ATAAAGA sequence in MPAE-1 (Fig. 1A). Thus, we speculate that Fkh transcription factor in Sf9 nuclear extract bound to all three MPAE fragments.

To further determine DNA binding motif in the MPAE region, we selected MPAE-3 fragment and divided it into three smaller fragments (22-24 bp) with 10-bp overlapping region (MPAE-3a to 3c) (Fig. 2A) for competitive EMSA assays. The

results showed that only MPAE-3b fragment, which contains the ATAAACA sequence, competed with binding of the labeled MPAE-3 to nuclear proteins of Sf9 cells and the competition by MPAE-3b was dose-dependent (Fig. 2B, lane 4 and Fig. 2C, lanes 3-6, arrows). These results suggest that Fkh transcription factor in Sf9 cells binds to the ATAAACA sequence in the moricin promoter.

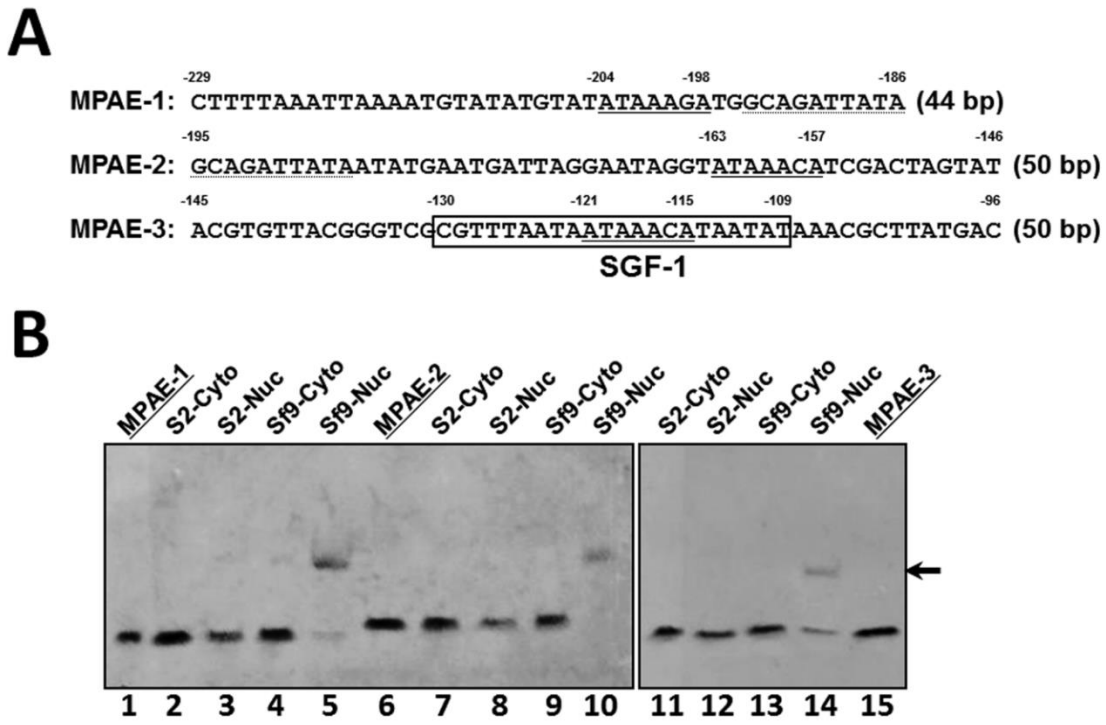


Fig. 1. Binding of Nuclear Proteins from Sf9 Cells to MPAE Fragments.

(A) DNA sequences of MPAE-1, -2 and -3 fragments of *M. sexta* moricin promoter. Predicted Fkh-binding sites are underlined and SGF-1 binding site is boxed. The 10-bp overlapping region in MPAE-1 and MPAE-2 fragments was dotted underlined. The numbers above the sequences indicate the positions upstream of the predicted transcription initiation site. (B): Electrophoretic mobility shift assay (EMSA). Cytosolic (Cyto) and nuclear (Nuc) proteins were prepared from Sf9 and S2 cells and incubated with biotinylated MPAE-1, -2 or -3 fragments for EMSA assays as described in the Materials and Methods. Lanes 1, 6 and 15 contained only biotinylated MPAE-1, -2 or -3; lanes 2-5, lanes 7-10 and lanes 11-14 contained biotinylated MPAE-1, MPAE-2 and MPAE-3, respectively, with cytosolic or nuclear proteins. Only nuclear proteins from Sf9 cells (Sf9-Nuc) bound to all three biotinylated MPAE fragments and caused mobility shift of the DNA fragments (lanes 5, 10 and 14, arrow).

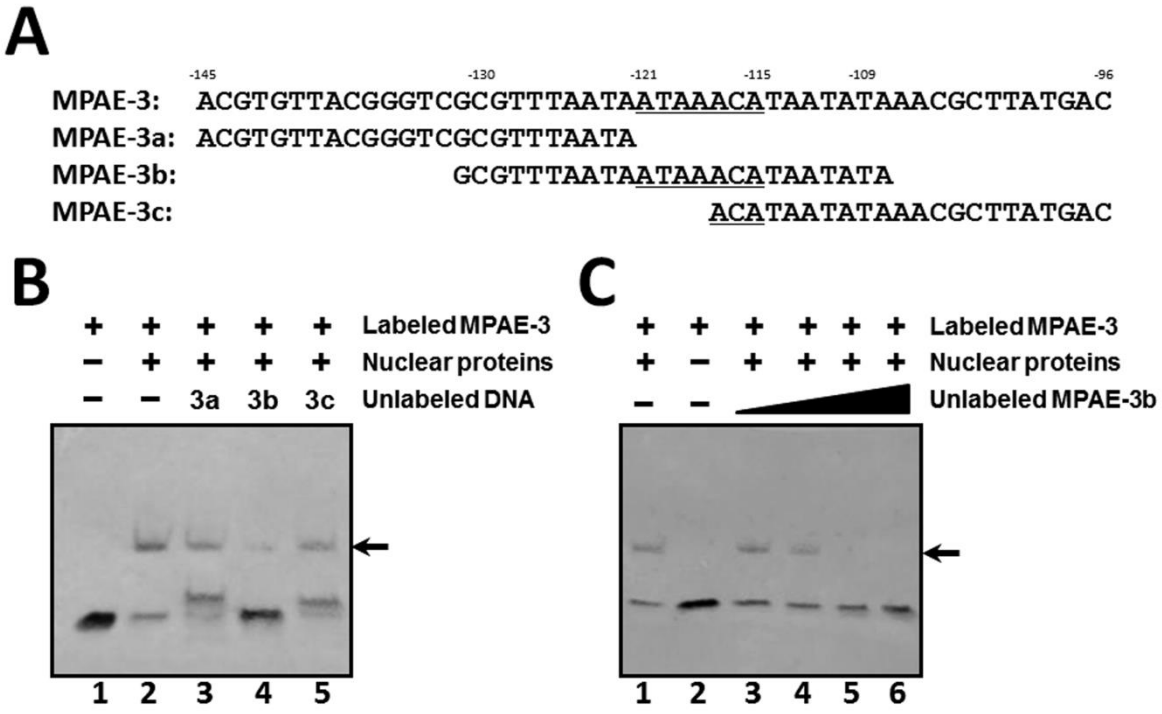


Fig. 2. Binding of Nuclear Proteins from Sf9 Cells to Fkh-Binding Site in MPAE-3.

(A): DNA sequences of MPAE-3, MPAE-3a, -3b and -3c. Fkh-binding site is underlined. The numbers above the sequence indicate the positions upstream of the predicted transcription initiation site. (B and C): MPAE-3b competed binding of nuclear proteins to MPAE-3 fragment. Biotinylated MPAE-3 fragment alone (lane 1 in panel B and lane 2 in panel C) or incubated with nuclear proteins from Sf9 cells in the absence (lane 2 in panel B and lane 1 in panel C) or presence of excess (200-fold) unlabeled MPAE-3a, -3b, or -3c (lanes 3-5 in panel B), or in the presence of increasing amount of excess (1-, 10-, 100- and 200-fold) unlabeled MPAE-3b (lanes 3-6 in panel C), and EMSA assays were performed the same as in Fig. 1. Arrows indicated the mobility-shifted bands.

Activation of *M. sexta* Moricin and Lysozyme Promoters by Fkh Factors

To investigate functions of *M. sexta* Fkh (MsFkh), we first cloned *MsFkh* cDNA and analyzed tissue distribution and induced expression of *MsFkh* mRNA. Real-time PCR results showed that *MsFkh* mRNA was highly expressed in the hemocytes of the 5th instar *M. sexta* naïve larvae compared to fat body, midgut and other tissues (Fig. 3A), and expression of *MsFkh* transcript was induced in hemocytes (Fig. 3B) and fat body (Fig. 3C) by some bacteria (*B. subtilis* and *E. coli*), but was not induced in the midgut by any of the microorganisms tested (Fig. 3D).

Since MPAE did not bind to nuclear proteins in S2 cells (Fig. 1B), and *Drosophila Fkh* (*DmFkh*, CG10002) is not expressed in S2 cells (Flybase), we performed real-time PCR to determined transcripts of several *Drosophila* Fox genes in S2 cells, including *DmFkh*, *FoxK* long and short isoforms, *jumeau* (*jumu*) and *dFoxO*. The results showed that mRNAs of *FoxK*, *jumu* and *dFoxO*, but not *DmFkh*, were detected in S2 cells, with higher transcript level of *FoxK* and *jumu* than *dFoxO* (Fig. 4A), confirming that *DmFkh* was not expressed (or was expressed at an undetectable level) in S2 cells. To test whether overexpression of DmFkh in S2 cells can activate expression of AMP genes, *DmFkh* was also cloned.

To investigate activation of *M. sexta* moricin and lysozyme promoters by forkhead factors, MsFkh and DmFkh were over-expressed in S2 cells, and dual luciferase assays were performed with several truncated moricin and lysozyme promoters. MsFkh is 355 residues long with a theoretical molecular weight of 39.5 kDa and *p*I of 9, while DmFkh (Forkhead isoform A) is 510 residues with molecular weight

of 54.3 kDa and *p*I of 8.7. The full length MsFkh shows 96% identity in amino acid sequence to *B. mori* SGF-1 (Genbank accession no. NP_001037329.1), 95% identity to *Helicoverpa armigera* (Genbank accession no. AAW56613.1) and *Spodoptera exigua* (Genbank accession no. ACA30303.1) forkhead domain transcription factors, but only 42% identity to the full length DmFkh. However, the forkhead domains of MsFkh and DmFkh proteins share 98% identity. Dual luciferase results showed that both MsFkh and DmFkh activated *M. sexta* moricin, lysozyme, defensin-1, defensin-3 and attacin-2 promoters, but did not activate *M. sexta* cecropin, attacin-1, or defensin-2 promoter (Fig. 4B-D). DmFkh stimulated the activity of AMP promoters to a similarly high level as or to a higher level than MsFkh. Comparing different truncated promoters of *M. sexta* moricin and lysozyme promoters, the 1.4-kb moricin promoter (Mor-1400) was activated by Fkh factors to a similarly high level as the 242-bp truncated moricin promoter (Mor-242), but further truncation of the 242-bp moricin promoter significantly decreased the activity of the truncated promoters (Mor-190, Mor-134, Mor-99 and Mor-80) (Fig. 4B), suggesting that the Fkh-binding sites are within the MPAE region. Similarly, the 1.2-kb lysozyme promoter (Lyz-1200) and the 345-bp truncated promoter (Lyz-345) showed a similarly high activity in S2 cells after overexpression of MsFkh or DmFkh, and further truncation of the Lyz-345 promoter significantly decreased the activity of the truncated lysozyme promoters Lyz-279, Lyz-230, Lyz-139 and Lyz-67 (Fig. 4C), indicating that the Fkh-binding sites are within the 345-bp region of the lysozyme promoter.

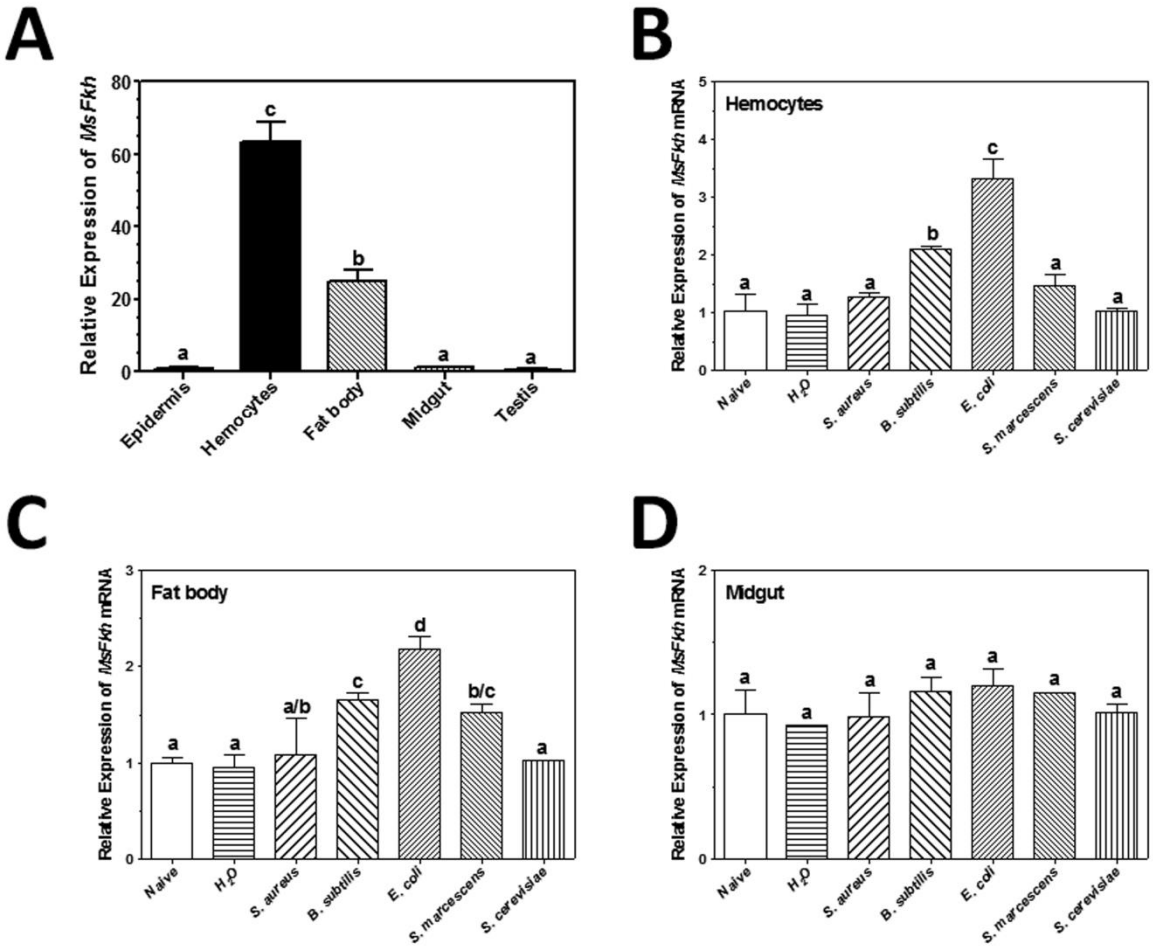


Fig. 3. Tissue Distribution and Induced Expression of *M. sexta* Fkh (*MsFkh*) mRNA.

Fat body, hemocytes, midgut, epidermis and testis were collected from *M. sexta* naïve fifth-instar larvae (A). Hemocytes (B), fat body (C) and midgut (D) were also collected from *M. sexta* larvae at 24 h post-injection of bacteria or yeast for preparation of total RNAs and cDNAs as described in the Materials and Methods. Real-time PCRs were performed with these RNA samples using *M. sexta* ribosomal protein S3 (*rpS3*) gene as an internal standard. Expression of *MsFkh* transcript in the epidermis of naïve larvae (A), or in the hemocytes (B), fat body (C) and midgut (D) of naïve larvae was arbitrarily set as 1. Bars represent the mean of three independent measurements ± SEM. Comparing different tissues (A) or different injection conditions (B-D), identical letters are not significant difference ($p>0.05$) while different letters indicate significant difference ($p<0.05$).

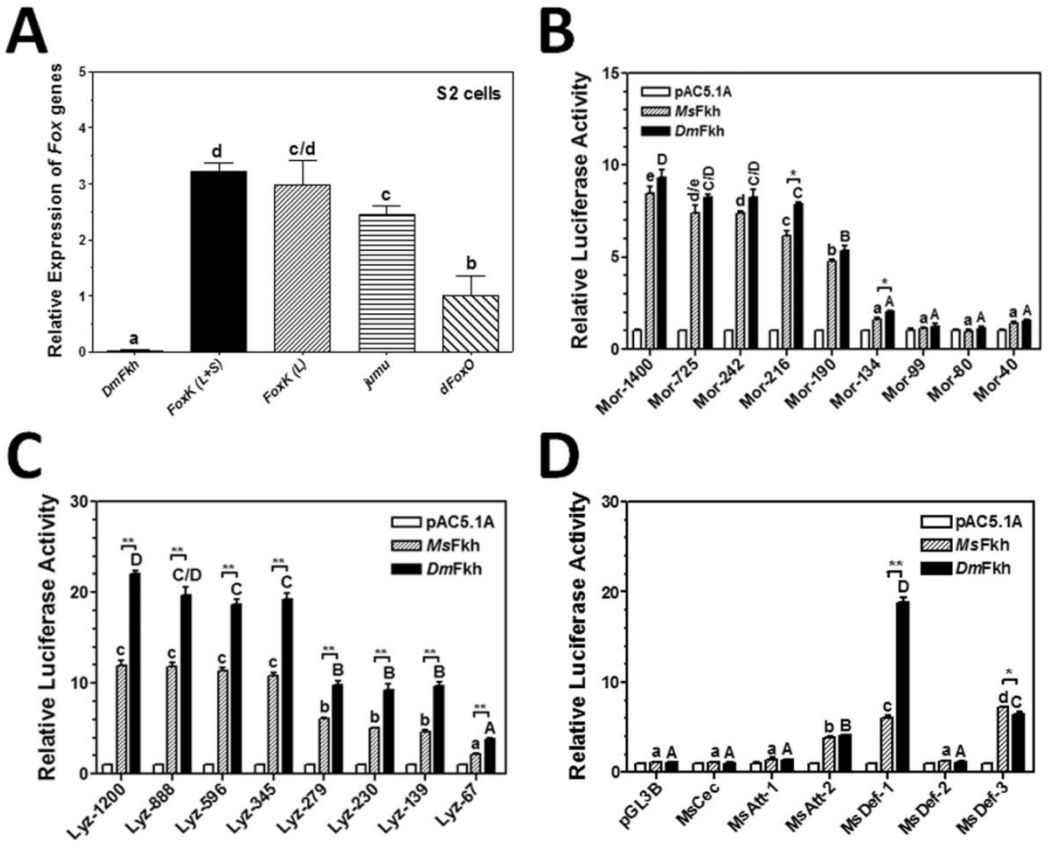


Fig. 4. Activation of AMP Gene Promoters by *M. sexta* and *D. melanogaster* Fkh Factors.

(A): Real-time PCR analysis of transcripts of *Fkh* (*DmFkh*), *FoxK* long and short isoforms (*FoxK* (L+S)), *FoxK* long isoform (*FoxK* (L)), *jumu* and *FoxO* (*dFoxO*) in S2 cells. *Drosophila* ribosomal protein 49 (*rp49*) gene was used as an internal standard. (B-D): Activation of AMP promoters by overexpression of MsFkh and DmFkh. The relative luciferase activities of truncated moricin promoters (B), truncated lysozyme promoters (C), and different *M. sexta* AMP gene promoters (D) activated by recombinant MsFkh or DmFkh in S2 cells were determined by Dual-Luciferase® Reporter Assay System as described in the Materials and Methods. Bars represent the mean of three independent measurements ± SEM. For transcription levels of Fox genes in S2 cells (A), or relative luciferase activity (B-D) among different promoters activated by one transcription factor (comparing striped bars or solid bars), identical letters are not significant difference ($p>0.05$) while different letters indicate significant difference ($p<0.05$). For the activity of the same promoter stimulated by different transcription factors (between MsFkh and DmFkh), the significance of difference was also determined by an unpaired t-test (*, $p<0.05$; **, $p<0.01$). MsCec, MsAtt-1, MsAtt-2, MsDef-1, MsDef-2 and MsDef-3 are *M. sexta* cecropin, attacin-1, attacin-2, defensin-1, defensin-2 and defensin-3 promoters (See Fig. S1 for the promoter sequences).

Identification of Fkh Binding Sites in *M. sexta* Moricin and Lysozyme Promoters

To identify Fkh-binding sites in the MPAE region of *M. sexta* moricin promoter and the 345-bp region of lysozyme promoter that are responsible for activation by Fkh factors, we first analyzed the Fkh-binding sites in the 140-bp MPAE and the 345-bp region of lysozyme promoter. Three and four Fkh-binding sites with the core sequence of AAACA were predicted within the 140 bp MPAE of moricin and 345 bp of lysozyme promoters, respectively (Figs. 1A, 5A, 5B and S1). To determine the active Fkh-binding sites in moricin and lysozyme promoters, the truncated Mor-242 and Lyz-345 promoters were selected for mutation of each Fkh-binding site for dual luciferase assays (Fig. 5A and B). Mutation of Fkh-binding site 2 or 3 alone in the Mor-242 promoter significantly decreased the activity of the mutant promoter by more than 60% compared to Mor-242 promoter, and mutation of both sites 2 and 3 together completely abolished activation of the mutant Mor-242 promoter by DmFkh (Fig. 5C). Mutation of Fkh-binding site 1 decreased the activity of the mutant promoter by ~20% compared to Mor-242 promoter, while mutation of both sites 1 and 2 or sites 1 and 3 together did not significantly decrease the activity further compared to mutation of site 2 or site 3 alone (Fig. 5C). These results suggest that Fkh-binding sites 2 and 3 in the MPAE region of moricin promoter play an equally important role in activation of moricin, while Fkh-binding site 1 may also contribute to activation of moricin. Similarly, mutation of the Fkh-binding site 1 or 2 alone in the Lyz-345 promoter significantly decreased the activity by more than 50%, and mutation of the Fkh-binding site 3 or 4 alone completely abolished DmFkh-activated Lyz-345 promoter activity (Fig. 5D), suggesting that Fkh-binding sites 3 and 4 play an important role in activation of

lysozyme by Fkh factor whereas Fkh-binding sites 1 and 2 also contribute to regulation of lysozyme.

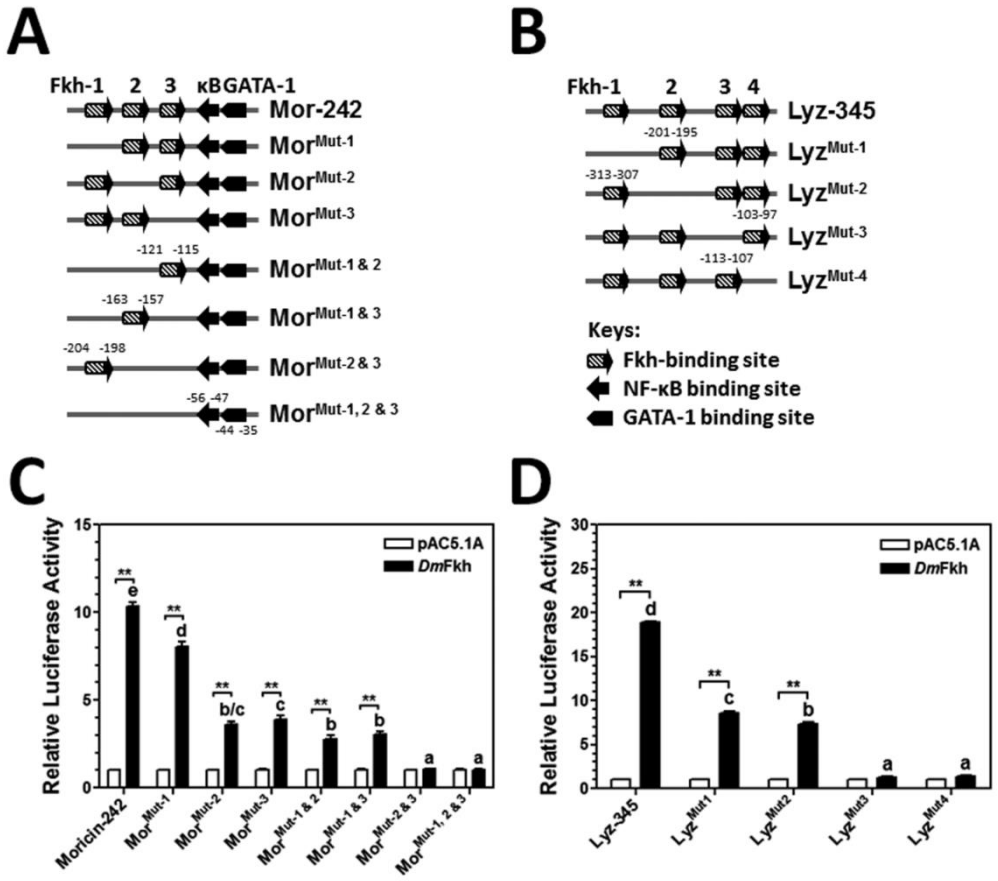


Fig. 5. Identification of Active Fkh-Binding Sites in *M. sexta* Moricin and Lysozyme Promoters.

(A, B): Schematic diagrams of the truncated Mor-242 (A) and Lyz-345 (B) promoters. Fkh-1, 2, 3, 4 indicate the predicted Fkh-binding sites 1, 2, 3, 4; κB and GATA-1 indicate the NF-κB and GATA-1 binding sites in the moricin promoter. Mut-1, -2, -3, -4 indicate the mutation of Fkh-binding sites 1, 2, 3, 4; Mut-1 & 2, Mut-1 & 3, Mut-2 & 3, and Mut-1, 2 & 3 indicate the mutations of Fkh-binding sites 1 and 2, 1 and 3, 2 and 3, as well as 1, 2 and 3, respectively. (C, D): Activation of Mor-242 and Lyz-345 promoters by recombinant DmFkh. The relative luciferase activities of Mor-242 and its Fkh-binding site mutant promoters (C), Lyz-345 and its Fkh-binding site mutant promoters (D) activated by recombinant DmFkh in S2 cells were determined by Dual-Luciferase® Reporter Assay System as described in the Materials and Methods. Bars represent the mean of three independent measurements ± SEM. For the relative luciferase activity among different promoters activated by DmFkh (comparing the solid bars), identical letters are not significant difference ($p>0.05$) while different letters

indicate significant difference ($p<0.05$). For the activity of the same promoter after overexpression of DmFkh (comparing the solid and open bars for each promoter), the significance of difference was also determined by an unpaired t-test (*, $p<0.05$; **, $p<0.01$).

Interaction of *M. sexta* Fkh with Relish-RHD

In the Mor-242 promoter, an NF-κB and a GATA-1 binding sites, which were both required for activation of moricin by immune signaling pathway (Rao et al., 2011), were still present near the transcription initiation site (Fig. 5A), while in the Lyz-345 promoter, the NF-κB site was absent (the only NF-κB site in the 1.2-kb lysozyme promoter was between -1191 and -1182 bp, Fig. S1). In order to test whether NF-κB and Fkh factors regulate AMP genes independently or cooperatively, we first determined interaction of Fkh factor with NF-κB factors Dorsal and/or Relish (Rel2). Co-immunoprecipitation (Co-IP) assays showed that V5-tagged MsFkh co-precipitated with Flag-tagged *M. sexta* Rel2-RHD (Fig. 6B and D, lane 4), but did not co-precipitate with Flag-tagged Dorsal-RHD (Dl-RHD) (Fig. 6F and H, lane 4), suggesting that MsFkh interacts with MsRelish but did not interact with MsDorsal.

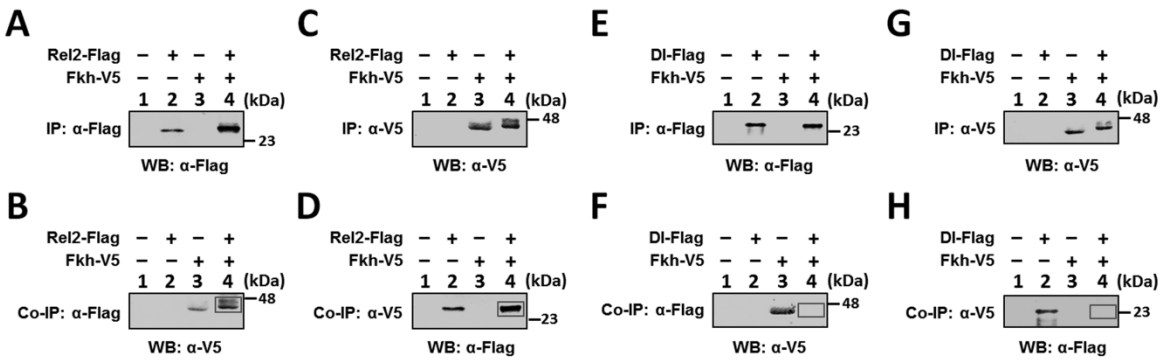


Figure 6. Interaction of *M. sexta* Fkh with Relish-RHD.

Recombinant V5-tagged MsFkh, Flag-tagged MsRelish-RHD (MsRel2-RHD) and Flag-tagged MsDorsal-RHD (MsDl-RHD) were expressed in S2 cells separately, and cell lysates from two samples were mixed for co-immunoprecipitation (Co-IP) assays as described in the Materials and Methods. Immunoprecipitated (IP) proteins or Co-IP proteins were detected by immunoblotting using anti-Flag or anti-V5 monoclonal antibody as the primary antibody. Lanes 1-3 were cell lysates (protein inputs) from control S2 cells (lane 1), S2 cells overexpressing Flag-tagged MsRel2-RHD or Flag-tagged MsDl-RHD (lane 2), and S2 cells overexpressing V5-tagged MsFkh (lane 3), and lane 4 was IP or Co-IP proteins. V5-tagged MsFkh was co-immunoprecipitated with Flag-tagged MsRel2-RHD (B and D, lane 4, boxed bands), but was not co-immunoprecipitated with Flag-tagged MsDl-RHD (F and H, lane 4, boxes).

Regulation of AMP Genes Independently by *M. sexta* Fkh and Relish

To determine whether interaction of MsFkh with MsRelish has an impact on regulation of moricin promoter by MsFkh and MsRelish, we used Mor-242 and its seven Fkh mutant promoters (Fig. 5A) for the dual luciferase assays. For MsFkh-stimulated activity (Fig. 7, solid bars), mutations of individual Fkh-binding sites alone (site 1, 2 or 3) or combination of Fkh-binding sites (sites 1 and 2, 1 and 3, 2 and 3, or all three sites) significantly decreased the activity of Fkh mutant promoters, a result similar to that in Fig. 5C, indicating that Fkh-binding sites indeed play a role in activation of moricin promoter by Fkh factor. Co-expression of MsFkh and MsRel2-RHD had similar effect as overexpression of MsRel2-RHD alone (Fig. 7, comparing stripe bars and dotted bars across different promoters, as well as between the stripe and dotted bars in each promoter), suggesting that presence/absence of MsFkh and Fkh-binding sites does not have an impact on activation of moricin promoter by MsRel2-RHD, and the overall activity of these moricin promoters is due to MsRel2-RHD binding to NF-κB site in the moricin promoter. Co-expression of MsFkh and MsRel2-RHD activated the activity of moricin promoters to either significantly higher or significantly lower level than that activated by overexpression of MsFkh alone, depending on the presence/absence of Fkh-binding sites (Fig. 7, comparing between the solid and dotted bars in each promoter), further confirming that when MsFkh and MsRel2-RHD are expressed together, MsRel2-RHD tends to bind NF-κB sites regardless of MsFkh and Fkh-binding sites. Together, these results suggest that even though MsFkh can interact with MsRelish, formation of MsRelish homodimers may be

predominant, and MsFkh and MsRelish regulate moricin promoter activation independently.

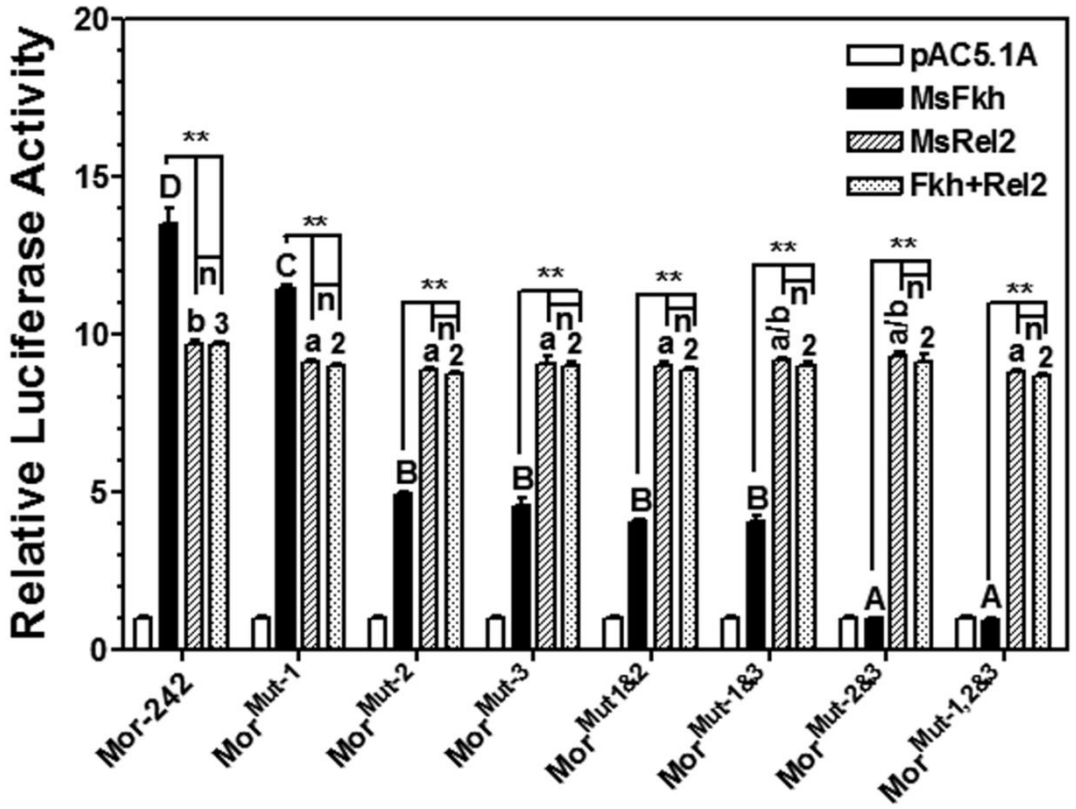


Fig. 7. Activation of Moricin Promoters by MsFkh and MsRel2-RHD Independently.

The relative luciferase activities of the truncated Mor-242 and its Fkh-binding site mutant promoters activated by recombinant MsFkh or MsRel2-RHD alone, or by co-expression of MsFkh and MsRel2-RHD in S2 cells were determined by Dual-Luciferase® Reporter Assay System as described in the Materials and Methods. Bars represent the mean of three independent measurements ± SEM. For the activity among different promoters activated by transcription factors (comparing solid bars by MsFkh, stripe bars by MsRel2-RHD, or dotted bars by MsFkh/MsRel2-RHD across the promoters), identical letters (capital letters for solid bars and small letters for stripe bars) or identical numerical numbers (dotted bars) are not significant difference ($p>0.05$) while different letters or different numerical numbers indicate significant difference ($p<0.05$). Comparing the activity of the same promoter stimulated by different transcription factors (between solid and stripe bars, solid and dotted bars, as well as stripe and dotted bars for each promoter), the significance of difference was also determined by an unpaired t-test (*, $p<0.05$; **, $p<0.01$), and "n" indicates not significant.

Discussion

Synthesis of AMPs is a major defense mechanism against infection in insects (Haine et al., 2008; Silverman et al., 2001; Strand 2008; Xiao et al., 2014; Yi et al., 2014), and expression of AMPs is regulated by the Toll and IMD pathways via activation of NF-κB transcription factors Dorsal, DIF and Relish (Ganesan et al., 2011; Matova et al., 2006). Other proteins and factors that can modulate the Toll and/or IMD pathways have been identified. For example, a Zn finger homeodomain 1 (zfh1) transcription factor has been reported as a negative regulator of *Drosophila* IMD pathway downstream of, or parallel to Relish (Myllymaki et al., 2013); Dorsal interacting protein 3 (Dip3) can bind to the RHD of Dorsal and Relish via its BESS domain, and it functions in both dorsoventral patterning and immune response (Ratnaparkhi et al., 2008). In *Drosophila*, it has been reported that activation of AMPs can be achieved by the transcription factor FoxO independent of the immune signaling pathways (Becker et al., 2010), and induction of two AMPs, diptericin and metchnikowin, after downregulation of TOR by rapamycin is regulated by the transcription factor forkhead (Fkh) (Varma et al., 2014). FoxO is an important regulator of stress, metabolism and aging (Haeusler et al., 2010; Maiese et al., 2009; Miyamoto 2008; Webb et al., 2014), and a key transcription factor in the insulin signaling pathway (Kitamura et al., 2002; Nakae et al., 2002); whereas Fkh is the founding member of the FoxO family. However, very little is known about regulation of AMPs in other insect species by the FoxO family transcription factors.

Insects can synthesize a variety of AMPs, some AMPs are common to most insects, whereas some other AMPs are found only in certain insect species (Haine et al., 2008). For example, moricin and gloverin have been identified only in the lepidopteran insects. We have previously shown that *M. sexta* moricin is regulated by NF-κB factor Relish and GATA-1 factor (Rao et al., 2011; Zhong et al., 2016b), and it can also be activated by unidentified nuclear factor(s) that bind to the MPAE region of moricin promoter (Rao et al., 2011). In this study, we identified three Fkh-binding sites in the MPAE region and demonstrated that Fkh-binding sites 2 and 3 played an equally important role in activation of moricin promoter by Fkh factor. We also identified four Fkh-binding sites in *M. sexta* lysozyme promoter and all four Fkh-binding sites are important for activation of lysozyme promoter by Fkh factor. Since the consensus sequence for FoxO is about 8 bp [(T/C)(G/A)AAACAA] (Li et al., 2004) with a core sequence of AAACA, there are many predicted/potential Fkh-binding sites in promoters. Thus, we used truncated promoters first to narrow the length of promoters and then focused on potential active Fkh-binding sites in the moricin and lysozyme promoters with the core sequence of AAACA, as the core sequence may be crucial for binding of Fkh factor. Among the three and four Fkh-binding sites in the Mor-242 and Lyz-345 truncated promoters, all six Fkh-binding sites but the binding site 1 in the Mor-242 contains the AAACA core sequence (Figs. 1A and 5B). Indeed, mutation of Fkh-binding site 1 (with the core sequence of AAAGA) did not decrease the activity of Mor-242 promoter as significantly as did mutation of Fkh-binding site 2 or 3, suggesting that the core sequence of AAACA is crucial for activation of genes by Fkh factor. In addition to moricin and lysozyme promoters, Fkh factor also activated *M.*

sexta defensin-1, defensin-3 and attacin-2 promoters. To our knowledge, this is the first report about activation of AMPs by Fkh factors in a lepidopteran insect. Activation of AMPs by Fkh factor under non-infectious conditions may be important for insects during molting and metamorphosis, since insects at these particular developmental stages are vulnerable to infection, and induced expression of AMPs by Fkh factor may protect insects from microbial infection. We observed that DmFkh stimulated the activity of some AMP promoters to a higher level than MsFkh did in S2 cells. This may be because MsFkh is not completely compatible in *Drosophila* S2 cells or the longer DmFkh (510 residues) may contain some other domains/motifs that can help Fkh to activate AMP gene promoters.

Different transcription factors may regulate gene expression independently or cooperatively. For *M. sexta* moricin, we previously found that both NF-κB and GATA-1 factors are required for activation of moricin promoter (Rao et al., 2011). We also found that *M. sexta* Dorsal can interact with Relish (Rel2), and Dorsal/Relish heterodimers serve as negative regulators to prevent over-activation of *M. sexta* AMP genes (Zhong et al., 2016b). We found that MsFkh interacted with MsRel2-RHD. In the Mor-242 promoter, which contains both NF-κB and Fkh binding sites, overexpression of MsFkh factor alone activated the promoter activity to a significantly higher level compared to overexpression of MsRel2-RHD alone, suggesting that Fkh factor plays an important role in activation of moricin under non-infectious conditions. In the mutant Mor-242 promoters in which the Fkh-binding sites were mutated, co-expression of MsFkh and MsRel2-RHD activated the activity of Fkh-binding site mutant promoters to a similar high level as that activated by MsRel2-RHD alone, indicating that MsFkh and

MsRel2 regulate moricin activation independently. This result also suggests that although MsFkh can interact with MsRelish to form heterodimers, homodimers of MsRelish may be predominant. This may be because the distance between Fkh and NF-κB binding sites in the moricin promoter (~58 bp between NF-κB site and Fkh-binding site 3, Fig. 5A) is too far away for the two factors to form heterodimers. For NF-κB and GATA-1 sites in the moricin promoter, the two sites are separated by only 2 bp (Fig. 5A), and we showed that both NF-κB and GATA-1 sites are required for activation of moricin promoter (Rao et al., 2011).

We also confirmed that *Fkh* mRNA was undetectable in *Drosophila* S2 cells, but the transcripts for *FoxK* (long and short isoforms), *jumu* and *dFoxO* were detected. Nuclear proteins from S2 cells did not bind to MPAE region, further supporting that *DmFkh* is not expressed (or was expressed at an undetectable level) in S2 cells. This information is important when performing promoter reporter assays for Fkh factor in S2 cells, as stress conditions may not activate the promoters in S2 cells due to lack of endogenous Fkh factor. In future, we will investigate activation of AMPs under non-infectious conditions via Fkh factor and maybe other members of the FoxO family and whether enhanced expression of AMPs by FoxO family members can protect insects from microbial infection during molting and/or metamorphosis.

CHAPTER 5

RESEARCH SUMMARY AND FUTURE DIRECTION

The innate immune system is conserved from insects to humans. In insects, the Toll and IMD signaling pathways, Dorsal, DIF and Relish NF-κB factors in the regulation of immune genes have been well studied in the *D. melanogaster*. So far, functions of few Toll receptors in *Drosophila* have been characterized. Similarly, the function of spätzle ligand that could pair up with these Toll receptors to activate the Toll pathway has not been well studied in *D. melanogaster*.

In my research, I have confirmed the presence of multiple Toll-Spz pathways in *D. melanogaster*. We showed direct interaction between DmToll-1 with spz -1, -2, -5 and DmToll-7 with spz-1, -2, -5 and -6 by Co-immunoprecipitation (CoIP) assay, and demonstrated that co-expression of DmToll-1 with and spz -1, -2, -5 could activate AMP gene promoters in S2 cells by dual luciferase assays. Likewise, co-expression of DmToll-7 with and spz -1, -2, -5 could activate AMP gene promoters in S2 cells measured by dual luciferase assays. In vivo assays showed that DmToll-1 mutant flies are highly susceptible to gram positive bacterial and fungal infection. However, DmToll-7 mutant male flies are more susceptible to *C.albicans* infection and to some extent to gram negative bacteria *P.aeruginosa* infections. This study may help better understand signaling pathways in insects, and the origin and evolution of animal innate immune signaling pathways. Since there are multiple Toll-spz pathways exists insects, but not all Toll-Spz pair play role in innate immunity like Toll-7-spz-6 pair, one future direction would be to investigate the function of such Toll-spz pair.

In *Drosophila*, Dorsal and DIF are activated by degradation of the inhibitor Cactus, while Relish is activated by cleavage of the C-terminal inhibitory ankyrin repeats. Dorsal, DIF and Relish can form homo- and heterodimers (Ganesan et al., 2011; Tanji et al., 2007; Tanji et al., 2010). In order to form DIF-Relish heterodimers, both the Toll and IMD pathways must be activated. Here, I identified short isoforms of Relish in *D. melanogaster* (Rel-S), which is a functionally active NF-κB factor. More importantly, I showed that Rel-S could form heterodimers with both Dorsal and DIF and regulate AMP genes. Other than this, I also studied mutational analysis of NF-kB binding sites on AMP gene promoters and found that certain bases are critical for AMP gene activity. This is a novel finding of NF-κB factors in the regulation of AMP gene expression. In other insects research also revealed the presence of Dorsal short isoforms. In future, it will be interesting to find other NF-kB transcription factor short isoforms which could play role in regulating AMP gene activity.

Lastly, NF-kB factors are not the only transcription factors that regulate the AMP genes. Factors like GATA also regulate the AMP gene activity. Research in *Drosophila* also showed the presence of other factors that could regulate AMP gene activity, namely forkhead factors (Fkh). Previous studies in *M.sexta* showed that both NF-κB and GATA-1 factors are required for activation of certain AMP gene promoters. Many AMP gene promoters contain extra regions for binding of additional transcription factors. Here, I found the presence of additional factor binding site in the AMP gene promoter in *M.sexta* named as Forkhead (Fkh). These Fkh binding sites are required for activation of AMP gene promoters by Fkh factors. This forkhead factor also binds with Relish and regulates AMP gene activity in *M. sexta*. In future, it will be worthwhile

finding other such Fkh factors in *M.sexta* and uncover their role in innate immunity. In *Drosophila* Fkh also controls Insulin signaling pathway. In *M.sexta* it will be interesting to test for Fkh role in insulin metabolism and other such functions.

References

Agrawal N., et al. 2013. Development associated profiling of chitinase and microRNA of Helicoverpa armigera identified chitinase repressive microRNA. Sci Rep 3: 2292.

Akira S.,Uematsu S., Takeuchi O. 2006. Pathogen recognition and innate immunity. Cell 124: 783-801.

Al Souhail Q., et al. 2016. Characterization and regulation of expression of an antifungal peptide from hemolymph of an insect, Manduca sexta. Dev Comp Immunol 61: 258-268.

Almeida P. F., Pokorny A. 2009. Mechanisms of antimicrobial, cytolytic, and cell-penetrating peptides: from kinetics to thermodynamics. Biochemistry 48: 8083-8093.

Anderson K. V.,Jurgens G., Nusslein-Volhard C. 1985. Establishment of dorsal-ventral polarity in the Drosophila embryo. Genetic studies on the role of the Toll gene product. Cell 42: 779-789.

Ao J. Q.,Ling E.,Rao X. J., Yu X. Q. 2008. A novel ML protein from Manduca sexta may function as a key accessory protein for lipopolysaccharide signaling. Mol Immunol 45: 2772-2781.

Arnot C. J.,Gay N. J., Gangloff M. 2010a. Molecular mechanism that induces activation of Spatzle, the ligand for the Drosophila Toll receptor. J Biol Chem 285: 19502-19509.

Arnot C. J.,Gay N. J., Gangloff M. 2010b. Molecular Mechanism That Induces Activation of Spätzle, the Ligand for the Drosophila Toll Receptor. J Biol Chem 285: 19502-19509.

Baeuerle P. A. 1991. The inducible transcription activator NF-kappa B: regulation by distinct protein subunits. Biochim Biophys Acta 1072: 63-80.

Baldwin A. S., Jr. 1996. The NF-kappa B and I kappa B proteins: new discoveries and insights. Annu Rev Immunol 14: 649-683.

Bangham J.,Jiggins F., Lemaitre B. 2006. Insect immunity: the post-genomic era. Immunity 25: 1-5.

Becker T., et al. 2010. FOXO-dependent regulation of innate immune homeostasis. Nature 463: 369-373.

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

Bettencourt R.,Tanji T.,Yagi Y., Ip Y. T. 2004. Toll and Toll-9 in Drosophila innate immune response. J Endotoxin Res 10: 261-268.

Bischoff V., et al. 2004. Function of the drosophila pattern-recognition receptor PGRP-SD in the detection of Gram-positive bacteria. Nat Immunol 5: 1175-1180.

Boman H. G., et al. 1989. Antibacterial and antimalarial properties of peptides that are cecropin-melittin hybrids. FEBS Lett 259: 103-106.

Brown K., et al. 1995. Control of I kappa B-alpha proteolysis by site-specific, signal-induced phosphorylation. Science 267: 1485-1488.

Buchon N., et al. 2009. A single modular serine protease integrates signals from pattern-recognition receptors upstream of the Drosophila Toll pathway. Proc Natl Acad Sci U S A 106: 12442-12447.

Bulet P., et al. 1991. Insect immunity. Isolation from a coleopteran insect of a novel inducible antibacterial peptide and of new members of the insect defensin family. J Biol Chem 266: 24520-24525.

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

Bulet P., et al. 1996. Enlarged scale chemical synthesis and range of activity of drosocin, an O-glycosylated antibacterial peptide of Drosophila. Eur J Biochem 238: 64-69.

Bulow M. H.,Aebersold R.,Pankratz M. J., Junger M. A. 2010. The Drosophila FoxA ortholog Fork head regulates growth and gene expression downstream of Target of rapamycin. PLoS One 5: e15171.

Busse M. S., et al. 2007. A kappaB sequence code for pathway-specific innate immune responses. Embo j 26: 3826-3835.

Cadigan K. M.,Grossniklaus U., Gehring W. J. 1994. Functional redundancy: the respective roles of the two sloppy paired genes in Drosophila segmentation. Proc Natl Acad Sci U S A 91: 6324-6328.

Cao J.,Chen Y.,Jin M., Ren Q. 2016. Enhanced antimicrobial peptide-induced activity in the mollusc Toll-2 family through evolution via tandem Toll/interleukin-1 receptor. R Soc Open Sci 3: 160123.

Carlsson A.,Nystrom T.,de Cock H., Bennich H. 1998. Attacin--an insect immune protein--binds LPS and triggers the specific inhibition of bacterial outer-membrane protein synthesis. Microbiology 144 (Pt 8): 2179-2188.

Carlsson P., Mahlapuu M. 2002. Forkhead transcription factors: key players in development and metabolism. Dev Biol 250: 1-23.

Chang C.-I., et al. 2005a. Structure of the ectodomain of Drosophila peptidoglycan-recognition protein LCa suggests a molecular mechanism for pattern recognition. Proceedings of the National Academy of Sciences of the United States of America 102: 10279-10284.

Chang C. I., et al. 2006. Structure of tracheal cytotoxin in complex with a heterodimeric pattern-recognition receptor. Science 311: 1761-1764.

Chang C. I., et al. 2005b. Structure of the ectodomain of Drosophila peptidoglycan-recognition protein LCa suggests a molecular mechanism for pattern recognition. Proc Natl Acad Sci U S A 102: 10279-10284.

Chasan R., Y. Jin, K. V. Anderson. 1992. Activation of the easter zymogen is regulated by five other genes to define dorsal-ventral polarity in the Drosophila embryo. Development: 115: 607–616.

Chen F. E.,Huang D. B.,Chen Y. Q., Ghosh G. 1998. Crystal structure of p50/p65 heterodimer of transcription factor NF-kappaB bound to DNA. Nature 391: 410-413.

Choe K. M.,Lee H., Anderson K. V. 2005. Drosophila peptidoglycan recognition protein LC (PGRP-LC) acts as a signal-transducing innate immune receptor. Proc Natl Acad Sci U S A 102: 1122-1126.

Choe K. M., et al. 2002. Requirement for a peptidoglycan recognition protein (PGRP) in Relish activation and antibacterial immune responses in Drosophila. Science 296: 359-362.

Christophides G. K., et al. 2002. Immunity-related genes and gene families in Anopheles gambiae. Science 298: 159-165.

Clark K. L.,Halay E. D.,Lai E., Burley S. K. 1993. Co-crystal structure of the HNF-3/fork head DNA-recognition motif resembles histone H5. Nature 364: 412-420.

Coffer P. J., Burgering B. M. 2004. Forkhead-box transcription factors and their role in the immune system. Nat Rev Immunol 4: 889-899.

Dathe M., Wieprecht T. 1999. Structural features of helical antimicrobial peptides: their potential to modulate activity on model membranes and biological cells. Biochim Biophys Acta 1462: 71-87.

De Gregorio E.,Spellman P. T.,Rubin G. M., Lemaitre B. 2001. Genome-wide analysis of the Drosophila immune response by using oligonucleotide microarrays. Proc Natl Acad Sci U S A 98: 12590-12595.

Diamond G.,Beckloff N.,Weinberg A., Kisich K. O. 2009. The Roles of Antimicrobial Peptides in Innate Host Defense. Curr Pharm Des 15: 2377-2392.

Dimarcq J. L., et al. 1990. Insect immunity: expression of the two major inducible antibacterial peptides, defensin and diptericin, in Phormia terranovae. Embo j 9: 2507-2515.

Dobrzanski P.,Ryseck R. P., Bravo R. 1993. Both N- and C-terminal domains of RelB are required for full transactivation: role of the N-terminal leucine zipper-like motif. Mol Cell Biol 13: 1572-1582.

Dunn P. E., Drake D. R. 1983. Fate of bacteria injected into naive and immunized larvae of the tobacco hornworm Manduca sexta. Journal of Invertebrate Pathology 41: 77-85.

Dushay M. S.,Asling B., Hultmark D. 1996a. Origins of immunity: Relish, a compound Rel-like gene in the antibacterial defense of Drosophila. Proc Natl Acad Sci U S A 93: 10343-10347.

Dushay M. S.,Asling B., Hultmark D. 1996b. Origins of immunity: Relish, a compound Rel-like gene in the antibacterial defense of Drosophila. Proceedings of the National Academy of Sciences of the United States of America 93: 10343-10347.

Engstrom Y., et al. 1993. kappa B-like motifs regulate the induction of immune genes in Drosophila. J Mol Biol 232: 327-333.

Evans J. D., et al. 2006. Immune pathways and defence mechanisms in honey bees Apis mellifera. Insect Mol Biol 15: 645-656.

Fehlbaum P., et al. 1994. Insect immunity. Septic injury of Drosophila induces the synthesis of a potent antifungal peptide with sequence homology to plant antifungal peptides. Journal of Biological Chemistry 269: 33159-33163.

Ferrandon D.,Imler J.-L.,Hetru C., Hoffmann J. A. 2007a. The Drosophila systemic immune response: sensing and signalling during bacterial and fungal infections. Nat Rev Immunol 7: 862-874.

Ferrandon D.,Imler J. L.,Hetru C., Hoffmann J. A. 2007b. The Drosophila systemic immune response: sensing and signalling during bacterial and fungal infections. Nat Rev Immunol 7: 862-874.

Ferrandon D.,Imler J. L., Hoffmann J. A. 2004. Sensing infection in Drosophila: Toll and beyond. Semin Immunol 16: 43-53.

Ganesan S.,Aggarwal K.,Paquette N., Silverman N. 2011. NF-kappaB/Rel proteins and the humoral immune responses of Drosophila melanogaster. Curr Top Microbiol Immunol 349: 25-60.

Gangloff M., A. Murali, J. Xiong, C. J. Arnot, A. N. Weber, A. M. Sandercock, C. V. Robinson, R. Sarisky, A. Holzenburg, C. Kao, N. J. Gay. 2008. Structural insight into

the mechanism of activation of the Toll receptor by the dimeric ligand Spätzle. J. Biol. Chem.: 283: 14629–14635.

Gangloff M., et al. 2008. Structural insight into the mechanism of activation of the Toll receptor by the dimeric ligand Spatzle. J Biol Chem 283: 14629-14635.

Gay N. J.,Gangloff M., Weber A. N. R. 2006. Toll-like receptors as molecular switches. Nat Rev Immunol 6: 693-698.

Gay N. J., Keith F. J. 1992. Regulation of translation and proteolysis during the development of embryonic dorso-ventral polarity in Drosophila. Homology of easter proteinase with Limulus proclotting enzyme and translational activation of Toll receptor synthesis. Biochim Biophys Acta 1132: 290-296.

Geisler R.,Bergmann A.,Hiromi Y., Nusslein-Volhard C. 1992. cactus, a gene involved in dorsoventral pattern formation of Drosophila, is related to the I kappa B gene family of vertebrates. Cell 71: 613-621.

Georgel P., et al. 2001. Drosophila immune deficiency (IMD) is a death domain protein that activates antibacterial defense and can promote apoptosis. Dev Cell 1: 503-514.

Ghosh S.,May M. J., Kopp E. B. 1998. NF-kappa B and Rel proteins: evolutionarily conserved mediators of immune responses. Annu Rev Immunol 16: 225-260.

Gottar M., et al. 2006. Dual detection of fungal infections in Drosophila through recognition of microbial structures and sensing of virulence factors. Cell 127: 1425-1437.

Gottar M., et al. 2002. The Drosophila immune response against Gram-negative bacteria is mediated by a peptidoglycan recognition protein. Nature 416: 640-644.

Graham McIlroy I. F., Jukka Aurikko, Jill S. Wentzell, Mei Ann Lim, Janine C. Fenton, Nicholas J. Gay, and Alicia Hidalgo. 2013. Toll-6 and Toll-7 function as neurotrophin receptors in the Drosophila central nervous system. Nat Neurosci.: 16(19): 1248–1256.

Grossniklaus U.,Pearson R. K., Gehring W. J. 1992. The Drosophila sloppy paired locus encodes two proteins involved in segmentation that show homology to mammalian transcription factors. Genes Dev 6: 1030-1051.

Habayeb M. S., et al. 2009. The Drosophila Nora virus is an enteric virus, transmitted via feces. J Invertebr Pathol 101: 29-33.

Hacker U., et al. 1995. The Drosophila fork head domain protein crocodile is required for the establishment of head structures. Embo j 14: 5306-5317.

Haeusler R. A.,Han S., Accili D. 2010. Hepatic FoxO1 ablation exacerbates lipid abnormalities during hyperglycemia. J Biol Chem 285: 26861-26868.

Haine E. R.,Moret Y.,Siva-Jothy M. T., Rolff J. 2008. Antimicrobial defense and persistent infection in insects. Science 322: 1257-1259.

Han J. H., et al. 2000. Gastrulation Defective is a serine protease involved in activating the receptor Toll to polarize the Drosophila embryo. Proc Natl Acad Sci U S A 97: 9093-9097.

Hancock R. E., Lehrer R. 1998. Cationic peptides: a new source of antibiotics. Trends Biotechnol 16: 82-88.

Hannenhalli S., Kaestner K. H. 2009. The evolution of Fox genes and their role in development and disease. Nat Rev Genet 10: 233-240.

Hansen I. A., et al. 2007. Forkhead transcription factors regulate mosquito reproduction. Insect Biochem Mol Biol 37: 985-997.

Hanson M. A.,Hamilton P. T., Perlman S. J. 2016. Immune genes and divergent antimicrobial peptides in flies of the subgenus Drosophila. BMC Evol Biol 16: 228.

Hedengren-Olcott M., et al. 2004. Differential activation of the NF-kappaB-like factors Relish and Dif in Drosophila melanogaster by fungi and Gram-positive bacteria. J Biol Chem 279: 21121-21127.

Hedengren M., et al. 1999. Relish, a central factor in the control of humoral but not cellular immunity in Drosophila. Mol Cell 4: 827-837.

Hetru C., Hoffmann J. A. 2009. NF-kappaB in the immune response of Drosophila. Cold Spring Harb Perspect Biol 1: a000232.

Hetru C., J. A. Hoffmann. 2009. NF-kappaB in the immune response of Drosophila. Cold Spring Harb. Perspect. Biol. : 1: a000232.

Hetru C.,Troxler L., Hoffmann J. A. 2003. Drosophila melanogaster Antimicrobial Defense. Journal of Infectious Diseases 187: S327-S334.

Hoffmann A.,Natoli G., Ghosh G. 2006. Transcriptional regulation via the NF-kappaB signaling module. Oncogene 25: 6706-6716.

Hoffmann J. A. 2003. The immune response of Drosophila. Nature 426: 33-38.

Hoffmann J. A., Hetru C. 1992. Insect defensins: inducible antibacterial peptides. Immunol Today 13: 411-415.

Honti V., et al. 2014. The cell-mediated immunity of Drosophila melanogaster: hemocyte lineages, immune compartments, microanatomy and regulation. Dev Comp Immunol 42: 47-56.

Horng T., Medzhitov R. 2001. Drosophila MyD88 is an adapter in the Toll signaling pathway. Proc Natl Acad Sci U S A 98: 12654-12658.

Hou F., et al. 2014. RNAi knock-down of shrimp Litopenaeus vannamei Toll gene and immune deficiency gene reveals their difference in regulating antimicrobial peptides transcription. Dev Comp Immunol 44: 255-260.

Huguet C.,Crepieux P., Laudet V. 1997. Rel/NF-kappa B transcription factors and I kappa B inhibitors: evolution from a unique common ancestor. Oncogene 15: 2965-2974.

Hui C. C.,Matsuno K., Suzuki Y. 1990. Fibroin gene promoter contains a cluster of homeodomain binding sites that interact with three silk gland factors. J Mol Biol 213: 651-670.

Hultmark D. 2003a. Drosophila immunity: paths and patterns. Curr Opin Immunol 15: 12-19.

Hultmark D. 2003b. Drosophila immunity: paths and patterns. Curr. Opin. Immunol. : 15: 12–19.

Hultmark D., Borge-Renberg K. 2007. Drosophila immunity: is antigen processing the first step? Curr Biol 17: R22-24.

Hwangbo D. S., et al. 2004. Drosophila dFOXO controls lifespan and regulates insulin signalling in brain and fat body. Nature 429: 562-566.

Imler J. L., Bulet P. 2005. Antimicrobial peptides in Drosophila: structures, activities and gene regulation. Chem Immunol Allergy 86: 1-21.

Imler J. L., Hoffmann J. A. 2001. Toll receptors in innate immunity. Trends Cell Biol 11: 304-311.

Imler J. L., et al. 2000. LPS-induced immune response in Drosophila. J Endotoxin Res 6: 459-462.

Ip Y. T., et al. 1993. Dif, a dorsal-related gene that mediates an immune response in Drosophila. Cell 75: 753-763.

Iwasaki A., Medzhitov R. 2004. Toll-like receptor control of the adaptive immune responses. Nat Immunol 5: 987-995.

Jackson B. C.,Carpenter C.,Nebert D. W., Vasiliou V. 2010. Update of human and mouse forkhead box (FOX) gene families. Hum Genomics 4: 345-352.

James Y. Ooi Y. Y., Xiaodi Hu, and Y. Tony Ip. 2002 The Drosophila Toll-9 activates a constitutive antimicrobial defense. EMBO Rep.: 3(1): 82–87.

Jang I. H., et al. 2006. A Spatzle-processing enzyme required for toll signaling activation in Drosophila innate immunity. Dev Cell 10: 45-55.

Janssens S.,Burns K.,Tschopp J., Beyaert R. 2002. Regulation of interleukin-1- and lipopolysaccharide-induced NF-kappaB activation by alternative splicing of MyD88. Curr Biol 12: 467-471.

Jiang H.,Vilcinskas A., Kanost M. R. 2010. Immunity in lepidopteran insects. Adv Exp Med Biol 708: 181-204.

Jonsson H., Peng S. L. 2005. Forkhead transcription factors in immunology. Cell Mol Life Sci 62: 397-409.

Jost P. J., Ruland J. 2007. Aberrant NF-κB signaling in lymphoma: mechanisms, consequences, and therapeutic implications. Blood 109: 2700-2707.

Jünger M. A., et al. 2003. The Drosophila Forkhead transcription factor FOXO mediates the reduction in cell number associated with reduced insulin signaling. J Biol 2: 20.

Kaestner K. H.,Knochel W., Martinez D. E. 2000. Unified nomenclature for the winged helix/forkhead transcription factors. Genes Dev 14: 142-146.

Kaneko T., et al. 2004. Monomeric and polymeric gram-negative peptidoglycan but not purified LPS stimulate the Drosophila IMD pathway. Immunity 20: 637-649.

Kaneko T., et al. 2006. PGRP-LC and PGRP-LE have essential yet distinct functions in the drosophila immune response to monomeric DAP-type peptidoglycan. Nat Immunol 7: 715-723.

Kappler C., et al. 1993. Insect immunity. Two 17 bp repeats nesting a kappa B-related sequence confer inducibility to the diptericin gene and bind a polypeptide in bacteria-challenged Drosophila. Embo j 12: 1561-1568.

Kaufmann E.,Hoch M., Jackle H. 1994. The interaction of DNA with the DNA-binding domain encoded by the Drosophila gene fork head. Eur J Biochem 223: 329-337.

Kaufmann E., Knochel W. 1996. Five years on the wings of fork head. Mech Dev 57: 3-20.

Khalil S.,Jacobson E.,Chambers M. C., Lazzaro B. P. 2015. Systemic Bacterial Infection and Immune Defense Phenotypes in Drosophila Melanogaster. J Vis Exp.

Khush R. S.,Leulier F., Lemaitre B. Drosophila immunity: two paths to NF-κB. Trends in Immunology 22: 260-264.

Kim Y. S., et al. 2000. Gram-negative bacteria-binding protein, a pattern recognition receptor for lipopolysaccharide and beta-1,3-glucan that mediates the signaling for the

induction of innate immune genes in Drosophila melanogaster cells. J Biol Chem 275: 32721-32727.

Kitamura T., et al. 2002. The forkhead transcription factor Foxo1 links insulin signaling to Pdx1 regulation of pancreatic beta cell growth. J Clin Invest 110: 1839-1847.

Kockum K., et al. 1984. Insect immunity. Isolation and sequence of two cDNA clones corresponding to acidic and basic attacins from Hyalophora cecropia. Embo j 3: 2071-2075.

Lai E., et al. 1990. HNF-3A, a hepatocyte-enriched transcription factor of novel structure is regulated transcriptionally. Genes Dev 4: 1427-1436.

Lam E. W. F.,Brosens J. J.,Gomes A. R., Koo C.-Y. 2013. Forkhead box proteins: tuning forks for transcriptional harmony. Nat Rev Cancer 13: 482-495.

Lamiable O., et al. 2016a. Analysis of the Contribution of Hemocytes and Autophagy to Drosophila Antiviral Immunity. J Virol 90: 5415-5426.

Lamiable O., et al. 2016b. Analysis of the Contribution of Hemocytes and Autophagy to Drosophila Antiviral Immunity. Journal of Virology 90: 5415-5426.

Lee W. J., et al. 1996. Purification and molecular cloning of an inducible gram-negative bacteria-binding protein from the silkworm, Bombyx mori. Proc Natl Acad Sci U S A 93: 7888-7893.

Lehmann O. J., et al. 2003a. Fox's in development and disease. Trends Genet 19: 339-344.

Lehmann O. J., et al. 2003b. Novel anterior segment phenotypes resulting from forkhead gene alterations: evidence for cross-species conservation of function. Invest Ophthalmol Vis Sci 44: 2627-2633.

Lemaitre B. 2004. The road to Toll. Nat Rev Immunol 4: 521-527.

Lemaitre B., Hoffmann J. 2007. The host defense of Drosophila melanogaster. Annu Rev Immunol 25: 697-743.

Lemaitre B., et al. 1995. A recessive mutation, immune deficiency (imd), defines two distinct control pathways in the Drosophila host defense. Proc Natl Acad Sci U S A 92: 9465-9469.

Lemaitre B., et al. 1996. The dorsoventral regulatory gene cassette spatzle/Toll/cactus controls the potent antifungal response in Drosophila adults. Cell 86: 973-983.

Leulier F., Lemaitre B. 2008. Toll-like receptors--taking an evolutionary approach. Nat Rev Genet 9: 165-178.

Leulier F., et al. 2000. The Drosophila caspase Dredd is required to resist gram-negative bacterial infection. EMBO Rep 1: 353-358.

Levashina E. A., et al. 1995. Metchnikowin, a novel immune-inducible proline-rich peptide from Drosophila with antibacterial and antifungal properties. Eur J Biochem 233: 694-700.

Levashina E. A.,Ohresser S.,Lemaitre B., Imler J. L. 1998. Two distinct pathways can control expression of the gene encoding the Drosophila antimicrobial peptide metchnikowin. J Mol Biol 278: 515-527.

Lewis M., et al. 2013. Cytokine Spatzle binds to the Drosophila immunoreceptor Toll with a neurotrophin-like specificity and couples receptor activation. Proc Natl Acad Sci U S A 110: 20461-20466.

Li S.,Weidenfeld J., Morrisey E. E. 2004. Transcriptional and DNA binding activity of the Foxp1/2/4 family is modulated by heterotypic and homotypic protein interactions. Mol Cell Biol 24: 809-822.

Lionakis M. S., Kontoyiannis D. P. 2012. Drosophila melanogaster as a model organism for invasive aspergillosis. Methods Mol Biol 845: 455-468.

Livak K. J., Schmittgen T. D. 2001. Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. Methods 25: 402-408.

Luo C.,Shen B.,Manley J. L., Zheng L. 2001. Tehao functions in the Toll pathway in Drosophila melanogaster: possible roles in development and innate immunity. Insect Mol Biol 10: 457-464.

Mach V., et al. 1995. Silk gland factor-1 involved in the regulation of Bombyx sericin-1 gene contains fork head motif. J Biol Chem 270: 9340-9346.

Maiese K.,Hou J.,Chong Z. Z., Shang Y. C. 2009. A fork in the path: Developing therapeutic inroads with FoxO proteins. Oxid Med Cell Longev 2: 119-129.

Manfruelli P., et al. 1999a. A mosaic analysis in Drosophila fat body cells of the control of antimicrobial peptide genes by the Rel proteins Dorsal and DIF. The EMBO Journal 18: 3380-3391.

Manfruelli P., et al. 1999b. A mosaic analysis in Drosophila fat body cells of the control of antimicrobial peptide genes by the Rel proteins Dorsal and DIF. Embo j 18: 3380-3391.

Manfruelli P., Reichhart, J. M., Steward, R., Hoffmann, J. A., & Lemaitre, B. 1999. A mosaic analysis in Drosophila fat body cells of the control of antimicrobial peptide genes by the Rel proteins Dorsal and DIF. . The EMBO Journal: 18(12), 3380–3391.

Margaret Nakamoto R. H. M., Jie Xu, Ari Yasunaga, Spencer S. Shelly, Beth Gold, Sara Cherry. 2012. Virus Recognition by Toll-7 Activates Antiviral Autophagy in Drosophila. Cell immunity: 36(34): 658–667.

Matova N., Anderson K. V. 2006. Rel/NF-kappaB double mutants reveal that cellular immunity is central to Drosophila host defense. Proc Natl Acad Sci U S A 103: 16424-16429.

Mazet F., et al. 2003. Phylogenetic relationships of the Fox (Forkhead) gene family in the Bilateria. Gene 316: 79-89.

McIlroy G., et al. 2013a. Toll-6 and Toll-7 function as neurotrophin receptors in the Drosophila melanogaster CNS. Nat Neurosci 16: 1248-1256.

McIlroy G., et al. 2013b. Toll-6 and Toll-7 function as neurotrophin receptors in the Drosophila melanogaster CNS. Nat Neurosci 16: 1248-1256.

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

Meister S., et al. 2005. Immune signaling pathways regulating bacterial and malaria parasite infection of the mosquito Anopheles gambiae. Proc Natl Acad Sci U S A 102: 11420-11425.

Mellroth P., et al. 2005. Ligand-induced dimerization of Drosophila peptidoglycan recognition proteins in vitro. Proc Natl Acad Sci U S A 102: 6455-6460.

Mellroth P.,Karlsson J., Steiner H. 2003. A scavenger function for a Drosophila peptidoglycan recognition protein. J Biol Chem 278: 7059-7064.

Meng X.,Khanuja B. S., Ip Y. T. 1999. Toll receptor-mediated Drosophila immune response requires Dif, an NF-kappaB factor. Genes Dev 13: 792-797.

Merkling S. H., van Rij R. P. 2013. Beyond RNAi: antiviral defense strategies in Drosophila and mosquito. J Insect Physiol 59: 159-170.

Michel T.,Reichhart J.-M.,Hoffmann J. A., Royet J. 2001a. Drosophila Toll is activated by Gram-positive bacteria through a circulating peptidoglycan recognition protein. Nature 414: 756-759.

Michel T.,Reichhart J. M.,Hoffmann J. A., Royet J. 2001b. Drosophila Toll is activated by Gram-positive bacteria through a circulating peptidoglycan recognition protein. Nature 414: 756-759.

Ming M.,Obata F.,Kuranaga E., Miura M. 2014. Persephone/Spatzle pathogen sensors mediate the activation of Toll receptor signaling in response to endogenous danger signals in apoptosis-deficient Drosophila. J Biol Chem 289: 7558-7568.

Miyamoto K. 2008. [FoxO3a is essential for the maintenance of hematopoietic stem cell pool]. Rinsho Ketsueki 49: 141-146.

Mohier E. 1993. [Determination of the dorso-ventral polarity of the Drosophila embryo]. Ann Genet 36: 5-15.

Morisato D., Anderson K. V. 1994. The spatzle gene encodes a component of the extracellular signaling pathway establishing the dorsal-ventral pattern of the Drosophila embryo. Cell 76: 677-688.

Morris O., et al. 2016a. Signal Integration by the IkappaB Protein Pickle Shapes Drosophila Innate Host Defense. Cell Host Microbe 20: 283-295.

Morris O., et al. 2016b. Signal Integration by the IκB Protein Pickle Shapes Drosophila Innate Host Defense. Cell Host Microbe 20: 283-295.

Moy R. H., et al. 2014. Antiviral autophagy restrictsRift Valley fever virus infection and is conserved from flies to mammals. Immunity 40: 51-65.

Muller U.,Vogel P.,Alber G., Schaub G. A. 2008. The innate immune system of mammals and insects. Contrib Microbiol 15: 21-44.

Myllymaki H., Ramet M. 2013. Transcription factor zfh1 downregulates Drosophila Imd pathway. Dev Comp Immunol 39: 188-197.

Myllymaki H.,Valanne S., Ramet M. 2014. The Drosophila imd signaling pathway. J Immunol 192: 3455-3462.

Nakae J., et al. 2002. Regulation of insulin action and pancreatic beta-cell function by mutated alleles of the gene encoding forkhead transcription factor Foxo1. Nat Genet 32: 245-253.

Nakamoto M., et al. 2012. Virus recognition by Toll-7 activates antiviral autophagy in Drosophila. Immunity 36: 658-667.

Narbonne-Reveau K.,Charroux B., Royet J. 2011. Lack of an Antibacterial Response Defect in Drosophila Toll-9 Mutant. PLOS ONE 6: e17470.

Neyen C.,Poidevin M.,Roussel A., Lemaitre B. 2012. Tissue- and ligand-specific sensing of gram-negative infection in drosophila by PGRP-LC isoforms and PGRP-LE. J Immunol 189: 1886-1897.

Nüsslein-Volhard C., E. Wieschaus. 1980. Mutations affecting segment number and polarity in Drosophila. Nature: 287: 795–801.

Oeckinghaus A., Ghosh S. 2009a. The NF-kappaB family of transcription factors and its regulation. Cold Spring Harb Perspect Biol 1: a000034.

Oeckinghaus A., Ghosh S. 2009b. The NF-κB Family of Transcription Factors and Its Regulation. Cold Spring Harb Perspect Biol 1.

Ooi J. Y.,Yagi Y.,Hu X., Ip Y. T. 2002. The Drosophila Toll-9 activates a constitutive antimicrobial defense. EMBO Rep 3: 82-87.

Oren Z., Shai Y. 1998. Mode of action of linear amphipathic alpha-helical antimicrobial peptides. Biopolymers 47: 451-463.

Paquette N., et al. 2010. Caspase-mediated cleavage, IAP binding, and ubiquitination: linking three mechanisms crucial for Drosophila NF-kappaB signaling. Mol Cell 37: 172-182.

Parker J. S.,Mizuguchi K., Gay N. J. 2001. A family of proteins related to Spatzle, the toll receptor ligand, are encoded in the Drosophila genome. Proteins 45: 71-80.

Parthier C., et al. 2014. Structure of the Toll-Spatzle complex, a molecular hub in Drosophila development and innate immunity. Proc Natl Acad Sci U S A 111: 6281-6286.

Pasare C., Medzhitov R. 2005. Toll-like receptors: linking innate and adaptive immunity. Adv Exp Med Biol 560: 11-18.

Pfaffl M. W. 2001. A new mathematical model for relative quantification in real-time RT-PCR. Nucleic Acids Res 29: e45.

Pieters R. J.,Arnusch C. J., Breukink E. 2009. Membrane permeabilization by multivalent anti-microbial peptides. Protein Pept Lett 16: 736-742.

Pili-Floury S., et al. 2004. In vivo RNA interference analysis reveals an unexpected role for GNBP1 in the defense against Gram-positive bacterial infection in Drosophila adults. J Biol Chem 279: 12848-12853.

Rao X. J.,Xu X. X., Yu X. Q. 2011. Manduca sexta moricin promoter elements can increase promoter activities of Drosophila melanogaster antimicrobial peptide genes. Insect Biochem Mol Biol 41: 982-992.

Rao X. J., et al. 2014a. Characterization of a novel Manduca sexta beta-1, 3-glucan recognition protein (betaGRP3) with multiple functions. Insect Biochem Mol Biol 52: 13-22.

Rao X. J., et al. 2014b. Characterization of a Novel Manduca sexta beta-1, 3-glucan recognition protein (βGRP3) with Multiple Functions. Insect Biochem Mol Biol 52: 13-22.

Ratnaparkhi G. S.,Duong H. A., Courey A. J. 2008. Dorsal interacting protein 3 potentiates activation by Drosophila Rel homology domain proteins. Dev Comp Immunol 32: 1290-1300.

Reichhart J. M., et al. 1993. Expression and nuclear translocation of the rel/NF-kappa B-related morphogen dorsal during the immune response of Drosophila. C R Acad Sci III 316: 1218-1224.

Rutschmann S., et al. 2000a. The Rel protein DIF mediates the antifungal but not the antibacterial host defense in Drosophila. Immunity 12: 569-580.

Rutschmann S., et al. 2000b. Role of Drosophila IKK in a Toll-independent antibacterial immune response. Nature Immunology 1: 342-347.

Sabin L. R.,Hanna S. L., Cherry S. 2010. Innate antiviral immunity in Drosophila. Curr Opin Immunol 22: 4-9.

Schneider D. S.,Jin Y.,Morisato D., Anderson K. V. 1994. A processed form of the Spatzle protein defines dorsal-ventral polarity in the Drosophila embryo. Development 120: 1243-1250.

Senger K., et al. 2004. Immunity regulatory DNAs share common organizational features in Drosophila. Molecular Cell 13: 19-32.

Servane Tauszig E. J., Jules A. Hoffmann, and Jean-Luc Imler. 2000. Toll-related receptors and the control of antimicrobial peptide expression in Drosophila. PNAS: 10520-10525.

Shi X. Z.,Zhong X., Yu X. Q. 2012. Drosophila melanogaster NPC2 proteins bind bacterial cell wall components and may function in immune signal pathways. Insect Biochem Mol Biol 42: 545-556.

Shin S. W.,Kokoza V.,Ahmed A., Raikhel A. S. 2002. Characterization of three alternatively spliced isoforms of the Rel/NF-kappa B transcription factor Relish from the mosquito Aedes aegypti. Proc Natl Acad Sci U S A 99: 9978-9983.

Silverman N., Maniatis T. 2001. NF-kappaB signaling pathways in mammalian and insect innate immunity. Genes Dev 15: 2321-2342.

Silverman N., et al. 2000a. A Drosophila IkappaB kinase complex required for Relish cleavage and antibacterial immunity. Genes Dev 14: 2461-2471.

Silverman N., et al. 2000b. A Drosophila IκB kinase complex required for Relish cleavage and antibacterial immunity. Genes Dev 14: 2461-2471.

Stein D.,Goltz J. S.,Jurcsak J., Stevens L. 1998. The Dorsal-related immunity factor (Dif) can define the dorsal-ventral axis of polarity in the Drosophila embryo. Development 125: 2159-2169.

Stenbak C. R., et al. 2004. Peptidoglycan molecular requirements allowing detection by the Drosophila immune deficiency pathway. J Immunol 173: 7339-7348.

Stoven S., et al. 2000. Activation of the Drosophila NF-kappaB factor Relish by rapid endoproteolytic cleavage. EMBO Rep 1: 347-352.

Stoven S., et al. 2003. Caspase-mediated processing of the Drosophila NF-kappaB factor Relish. Proc Natl Acad Sci U S A 100: 5991-5996.

Strand M. R. 2008. The insect cellular immune response. Insect Science 15: 1-14.

Strodicke M.,Karberg S., Korge G. 2000. Domina (Dom), a new Drosophila member of the FKH/WH gene family, affects morphogenesis and is a suppressor of position-effect variegation. Mech Dev 96: 67-78.

Sun H.,Bristow B. N.,Qu G., Wasserman S. A. 2002. A heterotrimeric death domain complex in Toll signaling. Proc Natl Acad Sci U S A 99: 12871-12876.

Sun S. C., Faye I. 1992. Cecropia immunoresponsive factor, an insect immunoresponsive factor with DNA-binding properties similar to nuclear-factor kappa B. Eur J Biochem 204: 885-892.

Sutcliffe B., et al. 2013. Neuron-type specific functions of DNT1, DNT2 and Spz at the Drosophila neuromuscular junction. PLoS One 8: e75902.

Taguchi T., J. L. Mitcham, S. K. Dower, J. E. Sims, J. R. Testa. 1996. Chromosomal localization of TIL, a gene encoding a protein related to the Drosophila transmembrane receptor Toll, to human chromosome 4p14. Genomics: 32: 486–488.

Takahashi Y., et al. 1997. A homozygous kinase-defective mutation in the insulin receptor gene in a patient with leprechaunism. Diabetologia 40.

Takeuchi O., Akira S. 2001. Toll-like receptors; their physiological role and signal transduction system. Int Immunopharmacol 1: 625-635.

Takeuchi O A. S. 2001. Toll-like receptors; their physiological role and signal transduction system. Int Immunopharmacol.: 1(4):625-635.

Tanaka H., et al. 2007. Identification and functional analysis of Relish homologs in the silkworm, Bombyx mori. Biochim Biophys Acta 1769: 559-568.

Tanji T.,Hu X.,Weber A. N., Ip Y. T. 2007a. Toll and IMD pathways synergistically activate an innate immune response in Drosophila melanogaster. Mol Cell Biol 27: 4578-4588.

Tanji T.,Hu X.,Weber A. N. R., Ip Y. T. 2007b. Toll and IMD pathways synergistically activate an innate immune response in Drosophila melanogaster. Molecular and Cellular Biology 27: 4578-4588.

Tanji T.,Hu X.,Weber A. N. R., Ip Y. T. 2007c. Toll and IMD Pathways Synergistically Activate an Innate Immune Response in Drosophila melanogaster. Mol Cell Biol 27: 4578-4588.

Tanji T.,Yun E. Y., Ip Y. T. 2010a. Heterodimers of NF-{kappa}B transcription factors DIF and Relish regulate antimicrobial peptide genes in Drosophila. Proceedings of the National Academy of Sciences of the United States of America 107: 14715-14720.

Tanji T.,Yun E. Y., Ip Y. T. 2010b. Heterodimers of NF-kappaB transcription factors DIF and Relish regulate antimicrobial peptide genes in Drosophila. Proc Natl Acad Sci U S A 107: 14715-14720.

Tanji T.,Yun E. Y., Ip Y. T. 2010c. Heterodimers of NF-κB transcription factors DIF and Relish regulate antimicrobial peptide genes in Drosophila. Proc Natl Acad Sci U S A 107: 14715-14720.

Tauszig S J. E., Hoffmann JA, Imler JL. 2000. Toll-related receptors and the control of antimicrobial peptide expression in Drosophila. Proc Natl Acad Sci USA: 97(19):10520-10525.

Tauszig S.,Jouanguy E.,Hoffmann J. A., Imler J. L. 2000. Toll-related receptors and the control of antimicrobial peptide expression in Drosophila. Proc Natl Acad Sci U S A 97: 10520-10525.

Towb P.,Bergmann A., Wasserman S. A. 2001. The protein kinase Pelle mediates feedback regulation in the Drosophila Toll signaling pathway. Development 128: 4729-4736.

Tuteja G., Kaestner K. H. 2007a. Forkhead transcription factors II. Cell 131: 192.

Tuteja G., Kaestner K. H. 2007b. SnapShot: forkhead transcription factors I. Cell 130: 1160.

Varma D., et al. 2014. Forkhead, a new cross regulator of metabolism and innate immunity downstream of TOR in Drosophila. J Insect Physiol 69: 80-88.

Wang G. 2015. Improved Methods for Classification, Prediction and Design of Antimicrobial Peptides. Methods Mol Biol 1268: 43-66.

Wasserman S. A. 1993. A conserved signal transduction pathway regulating the activity of the rel-like proteins dorsal and NF-kappa B. Mol Biol Cell 4: 767-771.

Webb A. E., Brunet A. 2014. FOXO transcription factors: key regulators of cellular quality control. Trends Biochem Sci 39: 159-169.

Weber A. N., et al. 2007. Role of the Spatzle Pro-domain in the generation of an active toll receptor ligand. J Biol Chem 282: 13522-13531.

Weber A. N., et al. 2003. Binding of the Drosophila cytokine Spatzle to Toll is direct and establishes signaling. Nat Immunol 4: 794-800.

Weigel D., Jackle H. 1990. The fork head domain: a novel DNA binding motif of eukaryotic transcription factors? Cell 63: 455-456.

Weigel D., et al. 1989. The homeotic gene fork head encodes a nuclear protein and is expressed in the terminal regions of the Drosophila embryo. Cell 57: 645-658.

Werner T., et al. 2003. Functional diversity of the Drosophila PGRP-LC gene cluster in the response to lipopolysaccharide and peptidoglycan. J Biol Chem 278: 26319-26322.

Werner T., et al. 2000. A family of peptidoglycan recognition proteins in the fruit fly Drosophila melanogaster. Proc Natl Acad Sci U S A 97: 13772-13777.

Wicker C., et al. 1990. Insect immunity. Characterization of a Drosophila cDNA encoding a novel member of the Diptericin family of immune peptides. Journal of Biological Chemistry 265: 22493-22498.

Wimley W. C., Hristova K. 2011. Antimicrobial Peptides: Successes, Challenges and Unanswered Questions. The Journal of Membrane Biology 239: 27-34.

Wu L. P., Anderson K. V. 1998. Regulated nuclear import of Rel proteins in the Drosophila immune response. Nature 392: 93-97.

Xia X., et al. 2015. Genome-wide characterization and expression profiling of immune genes in the diamondback moth, Plutella xylostella (L.). Sci Rep 5.

Xiao X., et al. 2014. Complement-related proteins control the flavivirus infection of Aedes aegypti by inducing antimicrobial peptides. PLoS Pathog 10: e1004027.

Xu P. X., et al. 1994. Promoter of the POU-M1/SGF-3 gene involved in the expression of Bombyx silk genes. J Biol Chem 269: 2733-2742.

Xu X. X.,Zhong X.,Yi H. Y., Yu X. Q. 2012. Manduca sexta Gloverin Binds Microbial Components and is Active against Bacteria and Fungi. Dev Comp Immunol 38: 275-284.

Yagi Y.,Nishida Y., Ip Y. T. 2010. Functional analysis of Toll-related genes in Drosophila. Dev Growth Differ 52: 771-783.

Yano T., et al. 2008. Autophagic control of listeria through intracellular innate immune recognition in drosophila. Nat Immunol 9: 908-916.

Yi H. Y.,Chowdhury M.,Huang Y. D., Yu X. Q. 2014. Insect antimicrobial peptides and their applications. Appl Microbiol Biotechnol 98: 5807-5822.

Yi H. Y., et al. 2013. Gloverins of the silkworm Bombyx mori: structural and binding properties and activities. Insect Biochem Mol Biol 43: 612-625.

Yvonne DeLotto R. D. 1998. Proteolytic processing of the Drosophila Spätzle protein by Easter generates a dimeric NGF-like molecule with ventralising activity. Mechanisms of Development: 72: 141–148.

Zaffran S.,Kuchler A.,Lee H. H., Frasch M. 2001. biniou (FoxF), a central component in a regulatory network controlling visceral mesoderm development and midgut morphogenesis in Drosophila. Genes Dev 15: 2900-2915.

Zambon R. A.,Nandakumar M.,Vakharia V. N., Wu L. P. 2005. The Toll pathway is important for an antiviral response in Drosophila. Proc Natl Acad Sci U S A 102: 7257-7262.

Zettervall C. J., et al. 2004. A directed screen for genes involved in Drosophila blood cell activation. Proc Natl Acad Sci U S A 101: 14192-14197.

Zhang L.,Wang Y. W., Lu Z. Q. 2015. Midgut immune responses induced by bacterial infection in the silkworm, Bombyx mori. J Zhejiang Univ Sci B 16: 875-882.

Zhong X.,Chowdhury M.,Li C. F., Yu X. Q. 2017. Transcription Factor Forkhead Regulates Expression of Antimicrobial Peptides in the Tobacco Hornworm, Manduca sexta. Sci Rep 7: 2688.

Zhong X., et al. 2016a. Co-expression of Dorsal and Rel2 Negatively Regulates Antimicrobial Peptide Expression in the Tobacco Hornworm Manduca sexta. Sci Rep 6.

Zhong X., et al. 2016b. Co-expression of Dorsal and Rel2 Negatively Regulates Antimicrobial Peptide Expression in the Tobacco Hornworm Manduca sexta. Sci Rep 6: 20654.

Zhong X., et al. 2012a. A Toll-Spatzle pathway in the tobacco hornworm, Manduca sexta. Insect Biochem Mol Biol 42: 514-524.

Zhong X., et al. 2012b. A Toll-Spätzle Pathway in the Tobacco Hornworm, Manduca sexta. Insect Biochem Mol Biol 42: 514-524.

Zhu B., et al. 2008. Drosophila Neurotrophins Reveal a Common Mechanism for Nervous System Formation. PLOS Biology 6: e284.

VITA

Munmun Chowdhury was born in Kolkata, India on August 12, 1980. She graduated from Rani Durgavati University in Jabalpur, MadhyaPradesh, India in 2001 with a Bachelor of Science in Biology. In 2006, she graduated from Rani Durgavati University in Jabalpur, MadhyaPradesh, India with a Master of Science in Zoology. In 2007, she received a Master of Philosophy in Fisheries Science from the Calcutta University, Kolkata, India. She received her master's degree in Cell, Molecular Biology and Biochemistry from University of Missouri-Kansas City in 2013. In 2012, she joined Dr. Xiao-Qiang Yu's lab to study insect immunology using the *Drosophila melanogaster*, and tobacco hornworm *Manduca sexta*.