REGULATION OF ANTIMICROBIAL PEPTIDE GENES IN THE
TOBACCO HORNWORM *MANDUCA SEXTA*

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

Xue Zhong

B. Sc., Xiamen University, 2006
M.Sc., Xiamen University, 2009
M.S., University of Missouri-Kansas City, 2010

Kansas City, Missouri
2014
REGULATION OF ANTIMICROBIAL PEPTIDE GENES IN THE
TOBACCO HORNWORM MANDUCA SEXTA

Xue Zhong, Candidate for the Doctor of Philosophy
University of Missouri-Kansas City, 2014

ABSTRACT

The innate immune system, also known as natural or non-specific immune system, is conserved from insects to humans. The insect innate immune system is composed of humoral and cellular components. Insect humoral responses mainly include synthesis of antimicrobial peptides (AMPs) and activation of the prophenoloxidase (PPO) system. Expression of AMPs is regulated by two signaling pathways, the Toll-SPätzle (Spz) and immune deficiency (IMD) pathways, in Drosophila melanogaster. My research has been focusing on pattern recognition receptors (PRRs) and regulation of AMP genes in the tobacco hornworm, Manduca sexta.

In chapter 2, I showed direct interaction between M. sexta MsToll and MsSpz-C108 by Co-immunoprecipitation (Co-IP), and demonstrated that co-expression of MsToll and MsSpz-C108 can activate AMP gene promoters in S2 cells by dual luciferase assays. My results confirm a Toll-Spz pathway in an insect other than Drosophila.

In chapter 3, a Toll-ML (MD2 (myeloid differentiation protein 2) - related lipid-recognition)-LPS signaling pathway was identified in M. sexta. The Co-IP assay showed that MsToll, MsML-1 and lipopolysaccharide (LPS) could form a receptor complex. More importantly, I showed that co-expression of MsToll and MsML-1 could up-regulate
AMP genes activated by LPS. My results for the first time showed that a Toll-ML-LPS signaling pathway is conserved from insects to humans.

In chapter 4, I identified Relish short isoforms and Dorsal in *M. sexta*. Dorsal and Relish belong to the nuclear factor-κB (NF-κB)/Rel family. I showed that *M. sexta* Dorsal and Rel2 (Relish) isoforms could regulate AMP genes differently, and Dorsal and Rel2 could form heterodimers, which negatively regulated AMP genes. This is a novel finding about NF-κB factors in regulation of AMP gene expression.

Characterization of a new member (βGRP3) of the β-1, 3-glucan recognition proteins in *M. sexta* was presented in chapter 5. I discovered novel functions of *M. sexta* βGRP3 in the βGRP family proteins: (1) calcium-dependent agglutination of bacterial cells, and (2) antibacterial (bacteriostatic) activity.

In Chapter 6, I characterized function of *M. sexta* gloverin. I showed that *M. sexta* gloverin can bind to different microbial cell wall components and has activity against different microorganisms. My results suggested that gloverin’s broad spectrum of activity may be related to its binding ability to microbial cell wall components.

In Chapter 7, I summarized my discoveries in insect innate immune signaling pathways and proposed future research directions.

My research has contributed to a better understanding of innate immunity in insects and vertebrates, regulation of gene expression by NF-κB factors, and evolution of the innate immune signaling pathways.
The faculty listed below, appointed by the Dean of the School of Graduate Studies, have examined a thesis titled “Regulation of Antimicrobial Peptide Genes in the Tobacco Hornworm *Manduca sexta*” presented by Xue Zhong, 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

Alexander Idnurm, Ph.D.

Cell Biology and Biophysics

Erika Geisbrecht, Ph.D.

Cell Biology and Biophysics

Leonard Dobens, Ph.D.

Molecular Biology and Biochemistry
## CONTENTS

ABSTRACT ................................................................................................................................. iii

ILLUSTRATIONS ....................................................................................................................... ix

TABLES ..................................................................................................................................... xii

ABBREVIATIONS ...................................................................................................................... xiii

ACKNOWLEDGEMENTS ........................................................................................................... xvii

Chapter

1. GENERAL INTRODUCTION OF INSECT INNATE IMMUNE SIGNALING PATHWAYS ................................................................. 1

   Beta-1,3-Glucan Recognition Proteins (BGRP)/Gram-Negative Binding Proteins (GNBPs) .......................................................... 2

   The *Drosophila* Toll Pathway ............................................................................................ 5

   Toll Receptors ..................................................................................................................... 9

   The *Drosophila* IMD Pathway ........................................................................................ 12

   The Mammalian Toll-like Receptors (TLRs) Pathways .............................................. 15

   NF-κB Transcription Factors ........................................................................................... 21

   Insect AMPs ....................................................................................................................... 25

   Questions and Hypotheses ............................................................................................... 29

2. A TOLL-SPÄTZLE PATHWAY IN THE TOBACCO HORNWORM, *MANDUCA SEXTA* ................................................................. 31

   Abstract ............................................................................................................................. 32

   Introduction ....................................................................................................................... 33

   Materials and Method .................................................................................................... 36

   Results .............................................................................................................................. 48

   Discussion ......................................................................................................................... 70

3. A TOLL-ML-LPS PATHWAY IN THE TOBACCO
7. RESEARCH SUMMARY AND FUTURE DIRECTION ........................................ 194

REFERENCES ........................................................................................................... 197

VITA ............................................................................................................................... 215
ILLUSTRATIONS

Figure                                           Page

1. GNBP1 and GNBP3 in D. melanogaster                    .................................................. 4
2. Drosophila Toll Pathway                             ................................................................. 8
3. Drosophila IMD Pathway                             ................................................................. 14
4. Mammalian TLRs Signaling Pathways                  ................................................................. 19
5. Mammalian members of the NF-κB, IκB, and IKK protein families ........................................... 23
6. The NF-κB and IκB proteins in Drosophila.          ................................................................. 24
7. Expression of recombinant M. sexta and D. melanogaster Toll and Spätzle proteins in S2 cells. .................................................................................................................. 50
8. Activation of recombinant M. sexta Spätzle by induced larval plasma. ........................................... 53
9. M. sexta Toll$^{\text{ecto}}$ (the ecto-domain) interacts with MsSpz-C108 but not with MsSpz... 56
10. D. melanogaster Toll$^{\text{ecto}}$ (the ecto-domain) interacts with DmSpz-C106 but not with DmSpz.................................................................................................................. 57
11. Over-expression of M. sexta Toll TIR domain (MsTIR) in S2 cells activates drosomycin promoter reporter .................................................................................................................. 60
12. Co-expression of MsToll with MsSpz-C108 but not MsSpz activates drosomycin promoter reporter in S2 cells .................................................................................................................. 61
13. Co-expression of MsToll with MsSpz-C108 but not MsSpz activates endogenous drosomycin gene in S2 cells .................................................................................................................. 63
14. Activation of AMP genes by MsSpz-C108 in M. sexta larvae ................................................................. 65
15. Activation of AMP genes in M. sexta larvae by MsSpz-C108 is blocked by pre-injection of antibody to MsToll .................................................................................................................. 69
16. Schematic diagram of the ChiToll. ................................................................................................. 83
17. Expression of $MsML-1$ (A), $MsToll^{ecto}$ (C), and $ChiToll$ (D) in S2 cells. ............... 84
18. Co-IP of $MsToll^{ecto}$ with $MsML-1$. ............................................................................. 87
19. LPS activates $drosomycin$ luciferase reporter in S2 cells co-expressing $ChiToll$ or $MsToll$ with $MsML-1$. ........................................................................................................... 89
20. Expression profile of $M. sexta$ Dorsal, Rel2A and Rel2B transcripts in different tissues by RT-PCR. ..................................................................................................................... 107
21. Western blot analysis of recombinant $MsDorsal$-RHD, $MsRel2A$, $MsRel2B$ and $MsRel2$-RHD in S2 and SF9 cells. ............................................................................................................. 109
22. $MsRel2$ and $MsDorsal$ activate expression of $M. sexta$ and $B. mori$ AMP gene promoters. .............................................................................................................................. 110
23. $Moricin$ promoter is regulated by $MsRel2$-RHD but not by $MsDorsal$-RHD. .......... 115
24. $Lysozyme$ promoter is regulated mainly by $MsDorsal$-RHD. ................................. 117
25. $M. sexta$ Rel2 interacts with $MsDorsal$. ........................................................................ 119
26. $MsDorsal$-$MsRel2$ heterodimer negatively regulates $M. sexta$ AMP expression. .... 121
27. Multiple sequence alignment of the active sites of glucanases and some BGRP/βGRP/GNBP proteins and phylogenetic tree of BGRP/βGRP/GNBP proteins. 141
28. Analysis of recombinant GFP and βGRP3 purified from $Drosophila$ S2 cells. ....... 144
29. Tissue distribution and expression of βGRP3 mRNA in response to immune challenge ......................................................................................................................... 145
30. Expression of βGRP3 protein in hemolymph in response to immune challenge 146
31. Binding of βGRP3 to microbial cell wall components ............................................. 149
32. Binding of βGRP3 to microorganisms ......................................................................... 150
33. Agglutination of microorganisms by βGRP3 .............................................................. 152
34. Antibacterial activity of βGRP3 against *B. cereus* ........................................ 154
35. Multiple sequence alignment of Gloverin proteins from some *lepidopteran* species.
........................................................................................................................................ 172
........................................................................................................................................ 173
37. Tissue distribution and expression of *Msglv* in *M. sexta* larvae after microbial infection.
........................................................................................................................................ 178
38. *Msglv* expression in *M. sexta* larvae is activated by *MsSpz-C108* and bacterial peptidoglycans.
........................................................................................................................................ 183
39. Activation of *Msglv* triggered by *MsSpz-C108* and bacterial peptidoglycans is blocked by antibody to *M. sexta* Toll. .............................................................. 184
40. Recombinant *MsGlv* binds to LPS, Gram-positive peptidoglycan (PG) and lipoteichoic acid (LTA), and laminarin. .............................................................. 188
41. Recombinant *MsGlv* binds to the O-specific antigen and out core carbohydrate moieties of LPS.............................................................. 189
42. Antimicrobial activities of recombinant *MsGlv* .................................................. 193
### TABLES

<table>
<thead>
<tr>
<th>Table</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. GRP/GNBP and Toll Receptors in <em>D. melanogaster</em></td>
<td>10</td>
</tr>
<tr>
<td>2. GRP/GNBP and Toll Receptors in <em>M. sexta</em></td>
<td>11</td>
</tr>
<tr>
<td>3. TLR-Targeted Therapeutics</td>
<td>20</td>
</tr>
<tr>
<td>4. Antimicrobial Peptides in <em>Drosophila</em></td>
<td>27</td>
</tr>
<tr>
<td>5. Antimicrobial Peptides in <em>M. sexta</em></td>
<td>28</td>
</tr>
<tr>
<td>6. PCR primers used in this study</td>
<td>46</td>
</tr>
<tr>
<td>7. PCR primers used in this study</td>
<td>103</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Description</td>
</tr>
<tr>
<td>--------------</td>
<td>-------------</td>
</tr>
<tr>
<td>AMP</td>
<td>Antimicrobial peptide</td>
</tr>
<tr>
<td>ANK</td>
<td>Ankyrin</td>
</tr>
<tr>
<td>AP</td>
<td>Alkaline Phosphatase</td>
</tr>
<tr>
<td>AP-1</td>
<td>Activator Protein 1</td>
</tr>
<tr>
<td>ATCC</td>
<td>American Type Culture Collection</td>
</tr>
<tr>
<td>BGRP</td>
<td>Beta-1,3-glucon recognition protein</td>
</tr>
<tr>
<td>bp</td>
<td>Base Pair</td>
</tr>
<tr>
<td>BPI</td>
<td>Bactericidal Permeability-Increasing domains</td>
</tr>
<tr>
<td>CD14</td>
<td>Cluster of Differentiation 14</td>
</tr>
<tr>
<td>cDNA</td>
<td>Complementary DNA</td>
</tr>
<tr>
<td>Co-IP</td>
<td>Co-immunoprecipitation</td>
</tr>
<tr>
<td>CREB</td>
<td>Cyclic AMP-responsive element-binding proteins</td>
</tr>
<tr>
<td>DAP</td>
<td>Diaminopimelic acid</td>
</tr>
<tr>
<td>DD</td>
<td>Death domain</td>
</tr>
<tr>
<td>ddH₂O</td>
<td>Distilled Deionized water</td>
</tr>
<tr>
<td>dH₂O</td>
<td>Deionized water</td>
</tr>
<tr>
<td>DIF</td>
<td>Dorsal-related immunity factor</td>
</tr>
<tr>
<td>DNA</td>
<td>Deoxyribonucleic acid</td>
</tr>
<tr>
<td>DREDD</td>
<td>Death related ced-3/Nedd2-like</td>
</tr>
<tr>
<td>dsRNA</td>
<td>Double-stranded RNA</td>
</tr>
<tr>
<td>FADD</td>
<td>FAS-associated death domain</td>
</tr>
<tr>
<td>GNBP</td>
<td>Gram-negative binding protein</td>
</tr>
<tr>
<td>Term</td>
<td>Definition</td>
</tr>
<tr>
<td>--------------</td>
<td>-------------------------------------------------------------------</td>
</tr>
<tr>
<td>Grass</td>
<td>Gram-positive-specific serine protease</td>
</tr>
<tr>
<td>IAP2</td>
<td>Inhibitor of apoptosis proteins 2 protein</td>
</tr>
<tr>
<td>IFNs</td>
<td>Type I interferons</td>
</tr>
<tr>
<td>IKK</td>
<td>IκB kinase</td>
</tr>
<tr>
<td>IL-1R</td>
<td>Interleukin-1 receptor</td>
</tr>
<tr>
<td>IMD</td>
<td>Immune deficiency</td>
</tr>
<tr>
<td>IPTG</td>
<td>Isopropyl β-D-thiogalactopyranoside</td>
</tr>
<tr>
<td>IRAK</td>
<td>IL-1R-associated kinase</td>
</tr>
<tr>
<td>IRD5</td>
<td>Immune-response deficient 5</td>
</tr>
<tr>
<td>IRF</td>
<td>Interferon regulatory factor</td>
</tr>
<tr>
<td>IκB</td>
<td>Inhibitor of κB</td>
</tr>
<tr>
<td>JNK</td>
<td>JUN N-terminal kinase</td>
</tr>
<tr>
<td>LBP</td>
<td>LPS Binding Protein</td>
</tr>
<tr>
<td>LPS</td>
<td>Lipopolysaccharide</td>
</tr>
<tr>
<td>LRR</td>
<td>Leucine-Rich Repeat</td>
</tr>
<tr>
<td>LTA</td>
<td>Lipoteichoic acid</td>
</tr>
<tr>
<td>LYS</td>
<td>Lysine</td>
</tr>
<tr>
<td>MAL</td>
<td>MYD88-adaptor-like protein</td>
</tr>
<tr>
<td>MAPKs</td>
<td>Mitogen-Activated Protein Kinases</td>
</tr>
<tr>
<td>MD2</td>
<td>Myeloid differentiation protein 2</td>
</tr>
<tr>
<td>ML</td>
<td>MD2-related lipid-recognition</td>
</tr>
<tr>
<td>ModSP</td>
<td>Modular Serine Protease</td>
</tr>
<tr>
<td>mRNA</td>
<td>Messenger Ribonucleic Acid</td>
</tr>
<tr>
<td>Term</td>
<td>Description</td>
</tr>
<tr>
<td>------</td>
<td>-------------</td>
</tr>
<tr>
<td>MYD88</td>
<td>Myeloid differentiation primary response gene (88)</td>
</tr>
<tr>
<td>NF-AT</td>
<td>Nuclear Factor of Activated T-cells</td>
</tr>
<tr>
<td>NF-κB</td>
<td>Nuclear factor kappa B</td>
</tr>
<tr>
<td>NOD</td>
<td>Nucleotide-binding oligomerization domain</td>
</tr>
<tr>
<td>ORF</td>
<td>Open Reading Frame</td>
</tr>
<tr>
<td>PAGE</td>
<td>Polyacrylamide Gel Electrophoresis</td>
</tr>
<tr>
<td>PAMP</td>
<td>Pathogen-associated molecular pattern</td>
</tr>
<tr>
<td>PAP</td>
<td>proPO activating proteinase</td>
</tr>
<tr>
<td>PCR</td>
<td>Polymerase chain reaction</td>
</tr>
<tr>
<td>PG</td>
<td>Peptidoglycan</td>
</tr>
<tr>
<td>PGRP</td>
<td>Peptidoglycan recognition protein</td>
</tr>
<tr>
<td>pI</td>
<td>Isoelectric point</td>
</tr>
<tr>
<td>PO</td>
<td>Phenoloxidase</td>
</tr>
<tr>
<td>proPO</td>
<td>Prophenoloxidase</td>
</tr>
<tr>
<td>PRR</td>
<td>Pattern recognition receptor</td>
</tr>
<tr>
<td>PSH</td>
<td>Protease Persephone</td>
</tr>
<tr>
<td>RACE</td>
<td>Rapid amplification of cDNA ends</td>
</tr>
<tr>
<td>RHD</td>
<td>Rel homolog domain</td>
</tr>
<tr>
<td>RIG-I</td>
<td>Retinoid acid-inducible gene I</td>
</tr>
<tr>
<td>RNA</td>
<td>Ribonucleic acid</td>
</tr>
<tr>
<td>S2</td>
<td>Schneider 2</td>
</tr>
<tr>
<td>SDS</td>
<td>Sodium dodecyl sulfate</td>
</tr>
<tr>
<td>SF9</td>
<td>Spodoptera frugiperda 9</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Description</td>
</tr>
<tr>
<td>--------------</td>
<td>-------------</td>
</tr>
<tr>
<td>siRNA</td>
<td>Small interfering RNA</td>
</tr>
<tr>
<td>SP</td>
<td>Serine protease</td>
</tr>
<tr>
<td>SPE</td>
<td>Spätzle processing enzyme</td>
</tr>
<tr>
<td>Spz</td>
<td>Spätzle</td>
</tr>
<tr>
<td>TAB</td>
<td>TAK1-binding protein</td>
</tr>
<tr>
<td>TAD</td>
<td>C-terminal transactivation domain</td>
</tr>
<tr>
<td>TAK1</td>
<td>TGFβ-activated kinase 1</td>
</tr>
<tr>
<td>TBS</td>
<td>Tris buffered saline</td>
</tr>
<tr>
<td>TBS-T</td>
<td>TBS with 0.05% Tween-20</td>
</tr>
<tr>
<td>TCT</td>
<td>Tracheal cytotoxin</td>
</tr>
<tr>
<td>TGF</td>
<td>Transforming growth factor</td>
</tr>
<tr>
<td>TIR</td>
<td>Toll/Interleukin-1 receptor</td>
</tr>
<tr>
<td>TLR</td>
<td>Toll-like receptor</td>
</tr>
<tr>
<td>TRAF</td>
<td>Tumor necrosis factor receptor-associated factor</td>
</tr>
<tr>
<td>TRAM</td>
<td>TRIF-related adaptor molecule</td>
</tr>
<tr>
<td>TRIF</td>
<td>TIR domain-containing adaptor protein inducing interferon-β</td>
</tr>
<tr>
<td>UBC5</td>
<td>Ubiquitin-conjugating enzyme 5</td>
</tr>
<tr>
<td>UEV1A</td>
<td>Ubiquitin-conjugating enzyme E2 variant 1</td>
</tr>
<tr>
<td>Unc13</td>
<td>Uncoordinated protein 13</td>
</tr>
</tbody>
</table>
ACKNOWLEDGEMENTS

I would like to thank my advisor and committee chair Dr. Xiao-Qiang Yu for his constant and patient 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. Erika Geisbrecht and Dr. Leonard Dobens for their helpful comments and suggestions while serving on my graduate advising committee.

I would like to thank all past and present members of the Yu lab for kindly sharing thoughts and ideas. I also want to thank Dr. Xiao-Xia Xu for cloning the *Manduca sexta* full length Toll receptor and gloverin, and Dr. Xiang-Jun Rao for cloning the *Manduca sexta* Relish, Dorsal, and βGRP3.
CHAPTER 1
GENERAL INTRODUCTION OF INSECT INNATE IMMUNE SIGNALING PATHWAYS

The innate and adaptive immune systems are two major branches of the defense system in multicellular organisms. The innate immune system, also known as natural or non-specific immune system, is the first line of defense against pathogens, such as bacteria and fungi. The innate immune system of insects is similar to that of vertebrates (Ferrandon et al., 2007; Lemaitre and Hoffmann, 2007; Muller et al., 2008). The generic recognition of pathogens by components of the innate immune system provides immediate defense against infection, which is conserved from insects to humans. On the other hand, adaptive immunity that is found in vertebrates but not insects develops over longer time periods, for example, the generation of antibody molecules may take up to one week. In innate immune responses, a group of germline-encoded pattern recognition receptors (PRRs) can recognize and bind to conserved pathogen-associated molecular patterns (PAMPs), such as lipopolysaccharides (LPS), peptidoglycans (PG), lipoteichoic acid (LTA), β-1,3-glucan, and mannan present on the invading microorganisms but not on the host cells. The receptors on the cell surface convey the information to the interior of the cell by series of adaptor proteins.

The innate immune system is composed of humoral and cellular components. Cellular responses mainly include blood cells (called hemocytes in insects) mediated immune responses such as nodule formation, phagocytosis and melanotic encapsulation, whereas synthesis of antimicrobial peptides (AMPs) and activation of the
prophenoloxidase (PPO) system are major components of humoral responses (Medzhitov and Janeway, 2000b). In this chapter, I focus on general introduction of insect humoral immunity. I will begin with recognition proteins in the host, which specifically recognize cell wall components of pathogens. Then, I will introduce two major pathways that have been well-studied in the *Drosophila melanogaster*: the Toll-Spätzle pathway and the immune deficiency (IMD) pathway. The two pathways regulate synthesis of most AMPs. I will also compare the difference between mammalian Toll-like receptor (TLR) pathway and *D. melanogaster* Toll pathway. Then I will discuss the family of nuclear factor kappa B (NF-κB) transcription factors that control the Toll and IMD pathways. Finally, I will present several AMPs as the targets of the Toll and IMD pathways.

**Beta-1,3-Glucan Recognition Proteins (BGRP)/Gram-Negative Binding Proteins (GNBPs)**

GNBPs are a class of immune recognition proteins, which share a conserved C-terminal β-glucanase-like domain (Werner et al., 2000). They were initially shown to recognize and bind to Gram-negative, but not Gram-positive, bacteria in *Bombyx mori* (Lee et al., 1996). These proteins are found only in invertebrate animals, and do not exist in mammals (Medzhitov and Janeway Jr, 2000). Members of the GNBPs were also discovered to recognize Gram-positive bacteria (Choe et al., 2002; Pili-Floury et al., 2004). GNBPs can bind to β-1, 3-glucans on fungi and therefore are also named β-1, 3-glucan recognition proteins (βGRPs or BGRPs) (Hughes, 2012; Jiang et al., 2004; Ma and Kanost, 2000; Wang et al., 2011). Based on their catalytic activity, GNBPs can be divided into two groups; one group possesses glucanase activity (Lee et al., 1996), while
another group has no catalytic activity due to lacking the critical cysteines required for the activity (Kim et al., 2000). In *Drosophila*, at least three GNBPs have been identified; however, only functions of *Dm*GNBP-1 and *Dm*GNBP-3 were characterized. Both contain an N-terminal GNBP homology domain and an evolutionarily conserved β-glucanase domain at the C-terminus (Hetru and Hoffmann, 2009) (Figure 1). *Dm*GNBP-1 exists as soluble and glycosylphosphatidylinositol (GIP)-anchored forms and can recognize LPS or β-1, 3-glucans (Kim et al., 2000). *Dm*GNBP-1 interacts with PGRP-SA to recognize the diaminopimelic acid (DAP)-type PG (Ferrandon et al., 2004). *Dm*GNBP-3 is crucial for the recognition of yeast (Buchon et al., 2009; Gottar et al., 2006). Both *Dm*GNBPs can initiate the Toll pathway (Ferrandon et al., 2004) (Table 1).

In *M. sexta*, three βGRPs and one GNBP have been identified. *Ms*βGRP-1 agglutinates yeast, Gram-negative and positive bacteria, and activates PPO, which triggers melanization in the presence of laminarin (Ma and Kanost, 2000). *Ms*βGRP-2 aggregates yeast, Gram-negative and Gram-positive bacteria, and stimulates PPO activation after binding to the laminarin (Jiang et al., 2004). *M. sexta* microbe binding protein (*Ms*MBP) alone enhances phenoloxidase (PO) activity at low concentration; however it dramatically inhibits PO activity when present at high concentration. PO activity is highly activated by *Ms*MBP via binding to a variety of cell wall components (Wang et al., 2011) (Table 2).
Figure 1. GNBP1 and GNBP3 in *D. melanogaster*

*Dm*GNBP1 and *Dm*GNBP3 contain an N-terminal GNBP homology domain and highly evolutional conserved β-glucanase domain at the C-terminus.

The *Drosophila* Toll Pathway

After PPRs detect and bind PAMPs to initiate immune responses, the signaling needs to be transmitted into cells to trigger the host defense mechanisms. There are two major pathways to integrate the extracellular signaling in *Drosophila melanogaster*, the Toll and IMD pathways. First, I introduce the *Drosophila* Toll pathway. Either in embryonic development or immune recognition, initiation of the *Drosophila* Toll pathway occurs through cleavage of the Toll receptor ligand Spätzle (or Spaetzle) (Spz) (Morisato and Anderson, 1994; Schneider et al., 1994). Spz, similar to other common secreted protein, is synthesized as an inactive precursor. Extracellular cleavage of the full-length Spz by a protease cascade into an active carboxyl-terminal fragment is required for activation of the Toll pathway. During early *Drosophila* embryogenesis, a Dorsal-Ventral (DV) signal initiates a serine protease (SP) cascade. In this cascade, Gastrulation Defective cleaves Snake, which activates the protease Easter (Chasan et al., 1992; Hong and Hashimoto, 1995). In another branch of the DV pattern, sulfotransferase Pipe also processes and activates Easter (Cho et al., 2010). The active form of Easter cleaves the full-length Spz (Munlix and Dunn, 1994). In pathogen recognition, the cell wall components β-glucan from fungi and Lysine (LYS)-type PG from Gram-positive bacteria initiate two protease cascades, which activate a Gram-positive-specific serine protease (Grass) (Royer and Dziarski, 2007). β-glucan and LYS-type PGs are recognized by GNBP3 and PGRP-SA, respectively (Kurokawa et al., 2011; Schleifer and Kandler, 1972). PGRP-SA physically interacts with GNBP1 to form a heterodimer (Gobert et al., 2003; Michel et al., 2001; Pili-Floury et al., 2004). New glycan reducing ends generated from the Lys-type PGs after proteolysis by active GNBP1 are captured by PGRP-SA
(Wang et al., 2006). LYS-type PG also can be recognized by PGRP-SD, which is a partially complementary pathway to the PGRP-SA-GNBP1 pathway (Bischoff et al., 2004). A modular serine protease (ModSP), which is highly conserved in insects, accepts the signals from GNBP3 and PGRP-SA (Buchon et al., 2009). ModSP converges the signal to Grass, which recruits another four serine proteases, named spirit, spheroid, and sphinx1/2 (El Chamy et al., 2008; Kambris et al., 2006). In immune response, there is a parallel proteolytic cascade mediated by the protease Persephone (PSH), which is activated by the secreted fungal virulence factor PR1 and Gram-positive bacterial virulence factors (El Chamy et al., 2008; Gottar et al., 2006).

In immune responses, all three proteolytic processes finally activate a Spz-processing enzyme (SPE), a serine protease that is responsible for Spz cleavage (Jang et al., 2006). Spz consists of a prodomain at the N-terminus and a carboxyl-terminal fragment (C-106), which includes 7 of the 9 Cysteine residues (DeLotto and DeLotto, 1998). The prodomain contains a conserved motif, which is an amphipathic helix. However, the C-terminal region is dominantly hydrophobic. In non-signaling conditions, the amphipathic domain masks the hydrophobic region, the determinant that is crucial for binding to the Toll receptor. The proteolysis by Easter or SPE leads to a conformational change, which releases and exposes the C-terminal 106 amino acids from the prodomain (Arnot et al., 2010). The active form of Spz dimerizes through three intrachain disulfide bridges. One Spz dimer binds to the N-terminus of one Toll receptor, triggering a conformational change that causes Toll receptor dimerization and the formation of an oligomer: 4Spz-2Toll. The complex initiates the intracellular signaling pathway (Gangloff et al., 2008). 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 and Medzhitov, 2001; Sun et al., 2002; Xiao et al., 1999). 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-signaling conditions, Cactus associates with the transcription factors of the Rel/NF-κB family Dorsal or a Dorsal-related immunity gene (Dif), retaining their cytoplasmic retention and inhibiting their nuclear localization (Wu and Anderson, 1998). During immune response, the active kinase Pelle phosphorylates and degrades Cactus (Towb et al., 2001). Dorsal or Dif is released from Cactus-Dorsal or Cactus-Dif complex, and translocates into the nucleus to activate the transcription of several target genes (Figure 2).
Figure 2. Drosophila Toll Pathway

Both early embryogenesis and immune recognition activate the Toll-Spz pathway via proteolytic processing. In the development and immune response, the protease Easter and SPE cleave full-length Spz, exposing the C-terminal Spz fragment, which is critical for the binding to Toll. One Spz dimer binds to one Toll receptor and triggers the conformation change leading to formation of a 4Spz:2Toll heteromultimer. The complex initiates the intracellular signal. Finally, kinase Pelle phosphorylates and degrades Cactus, which keeps Dif or Dorsal in the cytoplasm. Then Dif or Dorsal translocates to the nucleus and up-regulates target genes.

Toll Receptors

Nine Toll receptors and eleven Toll-like-Receptors (TLRs) have been identified in *D. melanogaster* and mammals, respectively. Both *Drosophila* Toll and mammalian TLRs contain a highly conserved extracellular domain called the Leucine-Rich Repeats (LRRs) and an intracellular motif named the Toll-IL-1-resistance (TIR) domain. In mammals, with the help of co-receptor proteins, TLRs directly detect, recognize, and bind to different PAMPs. These bindings trigger a conformation change, which induces the formation of TLR homo or heterodimers, to initiate the intracellular signal. In contrast, *Drosophila* Toll does not directly detect and bind to the PAMPs; however, similar to the TLR dimers, *Drosophila* Toll homodimers are crucial for 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). *Drosophila* Toll-6 and Toll-7 are expressed in central nervous system (CNS) and regulate the mature and formation of CNS (McIlroy et al., 2013). *Drosophila* Toll-7 is activated by virus infection (Nakamoto et al., 2012). One of *Drosophila* AMPs, *drosomycin* is up-regulated by Toll-9 in *D. melanogaster* Schneider 2 (S2) cell lines (Bettencourt et al., 2004b; 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 and Medzhitov, 2001). The functions of different *Drosophila* Toll are summarized in Table 1.

Previously, a Toll receptor in *M. sexta*, which can be induced by some of Gram-positive/negative bacteria or fungi, was reported, and the TIR domain of *MsToll* has high similarity to that of vertebrate TLR4 (Ao et al., 2008b) (Table 2).
Table 1 GRP/GNBP and Toll Receptors in *D. melanogaster*

<table>
<thead>
<tr>
<th>Gene</th>
<th>Type of Protein</th>
<th>Putative ligand</th>
<th>Evidence</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>GNBP1</td>
<td>Gram-negative binding protein</td>
<td>Potentially G(^+) bacterial determinants</td>
<td>Hydrolyzes G(^+) PGs. Acts in complex with PGRP-SA to activate the Toll pathway.</td>
<td>(Filipe et al., 2005; Gobert et al., 2003; Pili-Floury et al., 2004)</td>
</tr>
<tr>
<td>GNBP3</td>
<td>Gram-negative binding protein</td>
<td>Fungi</td>
<td>GNB3 recognizes β-1, 3-glucan and activates the Toll pathway.</td>
<td>(Buchon et al., 2009; Gottar et al., 2006)</td>
</tr>
<tr>
<td>18 wheeler, Toll-2</td>
<td>Toll-like Receptor</td>
<td>Unknown</td>
<td>Expression of <em>attacin</em> affected in mutant flies</td>
<td>(Eldon et al., 1994; Ligoxygakis et al., 2002; Williams et al., 1997)</td>
</tr>
<tr>
<td>Toll-6</td>
<td>Toll-like Receptor</td>
<td>Unknown</td>
<td>Toll-6 regulates the formation of <em>Drosophila</em> CNS</td>
<td>(McIlroy et al., 2013)</td>
</tr>
<tr>
<td>Toll-7</td>
<td>Toll-like Receptor</td>
<td>Unknown</td>
<td>Loss of Toll-7 leads to increased vesicular stomatitis virus (VSV) replication and mortality</td>
<td>(Nakamoto et al., 2012)</td>
</tr>
<tr>
<td>Toll-9</td>
<td>Toll-like Receptor</td>
<td>Unknown</td>
<td>Protein activates <em>drosomycin</em> in S2 cells through MyD88</td>
<td>(Bettencourt et al., 2004b; Ooi et al., 2002)</td>
</tr>
</tbody>
</table>

Table 2 GRP/GNBP and Toll Receptors in *M. sexta*

<table>
<thead>
<tr>
<th>Gene</th>
<th>Type of Protein</th>
<th>Putative ligand</th>
<th>Evidence</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>MsToll</td>
<td>Toll-like Receptor</td>
<td>Active Spätzle</td>
<td>Messenger ribonucleic acid (mRNA) was induced by yeast and Gram-positive bacteria</td>
<td>(Ao et al., 2008b)</td>
</tr>
<tr>
<td>MsPGRP</td>
<td>Peptidoglycan recognition protein</td>
<td></td>
<td>The mRNA level is rapidly increased at 2h post-injection and decreased after 24h injection</td>
<td>(Zhu et al., 2003)</td>
</tr>
<tr>
<td>MsβGRP-1</td>
<td>β-1, 3-glucan recognition proteins</td>
<td>Fungi, G⁻/G⁺ bacteria</td>
<td>MsβGRP-1 activates PPO in the presence of Laminarin</td>
<td>(Ma and Kanost, 2000)</td>
</tr>
<tr>
<td>MsβGRP-2</td>
<td>β-1, 3-glucan recognition proteins</td>
<td>Fungi, G⁻/G⁺ bacteria</td>
<td>MsβGRP-2 stimulates PO activity after binding to Laminarin</td>
<td>(Jiang et al., 2004)</td>
</tr>
<tr>
<td>MsMBP</td>
<td>Microbe binding protein</td>
<td>Fungi, G⁺/G⁻ bacteria</td>
<td>MsMBP enhances PO activity at low concentration and inhibit PO at high concentration</td>
<td>(Wang et al., 2011)</td>
</tr>
</tbody>
</table>
The *Drosophila* IMD Pathway

Gram-positive bacterial and fungal infections are mostly recognized by the *Drosophila* Toll pathway. In contrast, the IMD signaling pathway plays a key role in the recognition of Gram-negative bacteria. IMD signaling is triggered by DAP-type PG from Gram-negative and some Gram-positive bacteria. PGRP-LC, which encodes three alternative splice variants PGRP-LCa, PGRP-LCx and PGRP-LCy, is critical for the initiation of the IMD pathway (Werner et al., 2003). The PGRP-LCa and PGRP-LCx proteins contain a transmembrane domain and can recognize tracheal cytotoxin (TCT) in DAP-type PG (Chang et al., 2006; Chang et al., 2005; Kaneko et al., 2004; Kaneko and Silverman, 2005; Mellroth et al., 2005). The function of PGRP-LCy is partially redundant. DAP-type PG triggers the formation of the heterodimer of PGRP-LCa and PGRP-LCx or the homodimer of PGRP-LCx (Werner et al., 2000). PGRP-LE lacks the transmembrane domain. In the cell, PGRP-LE as the receptor 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).

Signal from extracellular region triggers the formation of PGRP-LC receptors hetero- or homo-dimers and the cytoplasmic domains of receptors 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). Cleaved IMD binds to the E3-ligase Inhibitor of apoptosis proteins 2 (IAP2), E2-ubiquitin-conjugating enzymes variant 1a (UEV1a), Bendless Uncoordinated protein 13 (Unc13), and Effete ubiquitin-conjugating enzyme 5 (Ubc5), which polyubiquitinate the cleaved IMD for degradation. The ubiquitin signaling activates the downstream Transforming growth factor-β activated kinase 1 (Tak1). In *Drosophila*, the N-terminal kinase (JNK) pathway is a branch of the IMD pathway, and Tak1 phosphorylates and activates transcription factor activator protein 1 (AP-1), which up-regulates the transcription of stress response or wound repair genes (Riesgo-Escovar et al., 1996; Sluss et al., 1996). In the IMD signaling pathway, Tak1 activates *Drosophila* inhibitor NF-κB proteins kinase (IKK) complex, which consists of two subunits, the kinase called Immune response deficient 5 (Ird5) and the regulatory subunit Kenney (Lu et al., 2001; Rutschmann et al., 2000; Silverman et al., 2003; Silverman et al., 2000; Vidal et al., 2001). Both IKK complex and DREDD are required for the cleavage of transcription factor Relish (Silverman et al., 2000). Relish contains an N-terminal Rel homology domain (RHD) and a C-terminal inhibitor NF-κB (IκB) part. In non-signaling condition, the IκB part inhibits RHD activity and nuclear localization. In immune challenge, Relish is phosphorylated and cleaved by kinase Ird5 and caspase DREDD. The IκB region is degraded and releases the NF-κB part, which translocates to the nucleus and activates the transcripts of AMPs (Stoven et al., 2000; Stoven et al., 2003) (Figure 3).
The IMD pathway is initiated by the binding of PGRP-LC dimer to the DAP-type PG. The dimer recruits IMD protein, which connects to the caspase DREDD through the adaptor protein FADD. DREDD cleaves IMD, which is polyubiquitinated and degraded. DREDD also cleaves Relish and releases the C-terminus of Relish. The N-terminal RHD region translocates into the nucleus and activates the transcription of downstream genes (Kleino, 2010).

(Kleino, A., The Imd Pathway-mediated Immune Response)
The Mammalian Toll-like Receptors (TLRs) Pathways

In vertebrate animals, TLRs, which are type I transmembrane receptors, serve in innate immune response to recognize a variety of PAMPs. The TLRs contain an evolutionarily conserved leucine-rich repeat (LRR) domain in the extracellular region and a TIR domain in the cells. Individual TLRs can distinguish different pathogens from self and directly bind to the conserved motifs found in distinct microorganisms, but not in hosts (Bischoff et al., 2004). TLR2, TLR1, TLR6, TLR5, and TLR11 localize to the plasma membrane. TLR4 localizes to both cell membrane and endosomes, while TLR3, TLR7, TLR8, and TLR13 localize to endosomes. The location of TLRs is dependent on the site of their ligands. TLRs localizing at the cell membrane detect specific cell wall components as ligands in blood; however, TLRs in the endosomes recognize microbial nucleic acid. The binding of ligands to the receptors triggers the formation of TLR dimers, which initiate the intracellular signaling pathway. TLR2 associates with TLR1 or TLR6 to form heterodimers that recognize a variety of microbial cell wall components, including LYS-type PG, lipoprotein/lipopeptides from several bacteria (Akira et al., 2001; Medzhitov, 2001). Diacylated and Triacylated lipopetides are recognized by TLR2-TLR6 and TLR2-TLR1, respectively (Takeuchi et al., 2001; Takeuchi et al., 2002). The proinflammatory signaling initiated by the bacterial flagellin is mediated by TLR5 (Hayashi et al., 2001). TLR11 plays an important role to detect uropathogenic bacteria (Zhang et al., 2004). In endosomes, homodimer TLR3 is involved in the recognition of double-stranded RNA (dsRNA) originating from viral replication within infected cells (Alexopoulou et al., 2001). The TLR7-TLR8 heterodimer recognizes many synthetic chemicals, like imiquimod, that is a potential antiviral chemical, and single-stranded viral
RNA (Diebold et al., 2004; Heil et al., 2004; Hemmi et al., 2002; Lund et al., 2004). Bacterial ribosomal RNA is recognized by the homodimer TLR13 (Hidmark et al., 2012; Li and Chen, 2012; Oldenburg et al., 2012). Homodimer TLR9 is crucial for the recognition of bacterial or viral nucleotides with rich hypomethylated CpG-deoxyribonucleic acid (DNA) motifs (Hemmi et al., 2000). LPS can be recognized by TLR4. The target PAMPs of all mammalian TLRs, specific immune cells expression, signal adaptors, and downstream effectors are summarized in Table 3.

Extracellular signals trigger the dimerization of receptors, which initiate the intracellular pathway. There are also two basic intracellular pathways, the MyD88-dependent and MyD88-independent pathways. TLR2, TLR1, TLR6, TLR5, TLR11, TLR7, TLR8, and TLR13 trigger a MyD88-dependent pathway in the cell. In contrast, TLR3 lead to a MyD88-independent pathway. Only TLR4 dimers can activate both pathways.

In the MyD88-dependent pathway, the dimers of TLRs recruit MyD88, an adaptor protein with the help of another adaptor protein MyD88-adaptor-like protein (MAL). TLRs interact with MyD88 via their intracellular TIR domain. MyD88 contains a death domain, which recruits the IL-1R-associated protein kinases (IRAKs) through DD domain interaction (Wesche et al., 1997). The interaction induces the recruitment of tumor necrosis factor receptor-associated factor 6 (TRAF6) and then triggers the formation of IRAKs/TRAF6 heterocomplex. IRAKs/TRAF6 heterodimer causes conformational changes and leaves the receptor dimers. The complex recruits kinase complex TAK1-binding proteins 2 (TAB2)/TAK1/TAK1-binding proteins 1 (TAB1). The TAB2/TAK1/TAB1/TRAF activates the mitogen-activated protein kinases (MAPKs),
which induce JNK and p38 pathways to up-regulate the proinflammatory cytokines (Flannery and Bowie, 2010; Wong et al., 1999). The complex also activates IKK, which phosphorylates and degrades IκBs. In nonsignaling conditions, IκBs sequester the transcription factor NF-κB in the cytoplasm. In immune recognition, IκBs are degraded and release NF-κB, which translocates into the nucleus and activates the transcription of effector genes (Silverman and Maniatis, 2001). In endosome signals, there is another pathway that can be activated. IRAKs/TRAF3 directly recruits IKK, which activates interferon-regulatory factor 7 (IRF7) translocating into the nucleus to induce the transcription of Type I interferons (IFNs) (O'Neill et al., 2013).

In the MyD88-independent pathway, also called the TRIF-dependent pathway, the dimer of TLRs recruits TIR domain-containing adaptor protein inducing interferon-β (TRIF) with the help of TRIF-related adaptor molecule (TRAM) via their TIR domain interaction (Fitzgerald et al., 2003; Oshiumi et al., 2003a; Yamamoto et al., 2003; Yamamoto et al., 2002). TRIF interacts with tumor necrosis factor receptor-associated factor 3 (TRAF3), which activates IKK. IKK leads to the translocation of transcription factor interferon-regulatory factor 3 (IRF3) into the nucleus and activates the expression of Type I IFNs (Oshiumi et al., 2003b).

Now I will interpret the TLR4-MD2-LPS pathway in more detail. In mammals, the TLR4-MD2-LPS pathway is elicited by a soluble acute-phase protein, LPS-binding protein (LBP), that rapidly responds to bacterial LPS. LBP contains two bactericidal permeability-increasing domains (BPI), which recognize and bind to LPS in the blood (Kirkland et al., 1990). LBP transfers LPS to a pattern recognition protein called cluster of differentiation 14 (CD14), which exists as glycosylphosphatidylinositol (GPI)-anchored
and soluble forms (Ferrero et al., 1990; Pugin et al., 1994). CD14 forms a dimer via its C-terminal domain, and accepts LPS at each N-terminal domain (Kim et al., 2005). Membrane-bound CD14 increases the local concentration of LPS and transfers LPS to a soluble myeloid differentiation factor 2 (MD2), which is a chaperone and co-receptor for TLR4 (Shimazu et al., 1999). The binding triggers a conformation change, which induces the formation of LPS-MD2-TLR4 complex (Schromm et al., 2001). The multimer triggers TLR4 dimerization, which initiates the intracellular signaling. TLR4 dimer leads to the translocation of transcription factors cyclic AMP-responsive element-binding proteins (CREB), AP1, and NF-κB to the nucleus via MyD88-dependent pathway and activates the transcription of target genes (Han et al., 1993; Kawai et al., 1999; Sen and Baltimore, 1986). Another transcription factor IRF3 can be activated by TLR4 dimer via MyD88-independent pathway (Hoebe et al., 2003). The active IRF3 translocates into the nucleus, then binds to the promoters of IFNβ genes. The general story of TLRs is summarized in Figure 4.
In mammalian systems, TLRs as PRRs recognize and directly bind to pathogens and initiate the intracellular signaling pathway (Bischoff et al., 2004). TLR2, TLR1, TLR6, TLR5, and TLR11 in the plasma membrane detect the cell wall components of microorganisms in the blood. TLR4 locating to both cell membrane and endosomes recognizes LPS from Gram-negative bacteria. Instead, TLR3, TLR7, TLR8, and TLR13 in endosomes bind to microbial nucleic acid. The binding of ligands to the receptors triggers the formation of TLR hetero- or homo-dimers. TLR2, TLR1, TLR6, TLR5, TLR11, TLR7, TLR8, and TLR13 activate a MyD88-dependent pathway in the cell. In contrast, TLR3 leads to a TRIF-dependent pathway. Only TLR4 initiates both pathways. In the two pathways, MyD88 or TRIF recruits a serial of adaptor proteins, kinases, or caspases. Activation of signaling pathways releases the transcription factors in the cytoplasm such as NF-κB, IRF7, IRF3, CREB, and AP1, which translocate into the nucleus and activate the transcription of target genes, such as AMPs, IFNβ gene, IFNs, Type I IFNs.

Table 3 TLR-Targeted Therapeutics

<table>
<thead>
<tr>
<th>TLR</th>
<th>Immune Cell Expression</th>
<th>PAMPs</th>
<th>Signal Adaptor</th>
<th>Production</th>
</tr>
</thead>
<tbody>
<tr>
<td>TLR1+ TLR2</td>
<td>Cell surface Mo, MΦ, DC, B</td>
<td>Triacylated lipoproteins (Pam3CSK4)</td>
<td>TIRAP, MyD88, Mal</td>
<td>IC</td>
</tr>
<tr>
<td></td>
<td></td>
<td>PG, LPS</td>
<td>TIRAP, MyD88, Mal</td>
<td>IC</td>
</tr>
<tr>
<td>TLR2+ TLR6</td>
<td>Cell surface Mo, MΦ, MC, B</td>
<td>Diacylated lipoproteins (FSL-1)</td>
<td>TIRAP, MyD88, Mal</td>
<td>IC</td>
</tr>
<tr>
<td>TLR3</td>
<td>Endosomes B, T, NK, DC</td>
<td>dsRNA (poly (I:C)) tRNA, siRNA</td>
<td>TRIF</td>
<td>IC, type1 IFN</td>
</tr>
<tr>
<td></td>
<td>Cell surface/endosomes Mo, MΦ, DC, MC, IE</td>
<td>LPS Paclitaxel</td>
<td>TRAM, TRIF TIRAP, MyD88 Mal</td>
<td>IC, type1 IFN</td>
</tr>
<tr>
<td>TLR4</td>
<td>Cell surface Mo, MΦ, DC, IE</td>
<td>Flagellin</td>
<td>MyD88</td>
<td>IC</td>
</tr>
<tr>
<td>TLR7</td>
<td>Endosomes Mo, MΦ, DC, B</td>
<td>ssRNA Imidazoquinolines (R848) Guanosine analogs (Loxoribine)</td>
<td>MyD88</td>
<td>IC, type1 IFN</td>
</tr>
<tr>
<td>TLR8</td>
<td>Endosomes Mo, MΦ, DC, MC</td>
<td>ssRNA, Imidazoquinolines (R848)</td>
<td>MyD88</td>
<td>IC, type1 IFN</td>
</tr>
<tr>
<td>TLR9</td>
<td>Endosomes Mo, MΦ, DC, B, T</td>
<td>CpG DNA CpG ODNs</td>
<td>MyD88</td>
<td>IC, type1 IFN</td>
</tr>
<tr>
<td>TLR10</td>
<td>Endosomes Mo, MΦ, DC</td>
<td>profilin-like proteins</td>
<td>MyD88</td>
<td>IC</td>
</tr>
</tbody>
</table>

NF-κB Transcription Factors

Nuclear factor kappa B (NF-κB), which consists of a family of transcription factors in insects and mammals, but not in bacteria, fungi, plants, or C. elegans, plays a crucial role in apoptosis, inflammation, immunity, cell proliferation, differentiation, and survival. In humans, there are five NF-κB members (Figure 5); and the Drosophila genome codes for three homologs (Figure 6). Based on the different regulatory mechanisms, secondary structure, and transactivation potential, 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 long RHD motif at the N-terminal region, which is critical for formation of homo- or hetero-dimers, 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 class I NF-κB factors is phosphorylation-induced proteolysis. In un-stimulated conditions, RHD is sequestered in the cell cytoplasm via interacting with ankyrin (ANK) repeats in IκB. In immune recognition, the IKK complex phosphorylates, ubiquitinates, and degrades IκB to release NF-κB. The active NF-κB translocates into the nucleus to activate the transcript of target genes (Brown et al., 1995). Drosophila Dorsal and Dif are 45% identical to the 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 an IκB conserved domain, ANK repeats, an N-terminal acidic domain
(Ac), and C-terminal proline, glutamic acid, serine, and threonine-rich segment (PEST) (Figure 6) (Oeckinghaus and Ghosh, 2009).

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 a 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 the immune challenge, p100 and p105 are proteolytic cleaved after phosphorylation and ubiquitination to remove ANK repeats, then the active p52 and p50 are released (Huguet et al., 1997; Hultmark, 2003).

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 region, which translocates to the nucleus and activates the transcript of target genes (Stoven et al., 2003).

Class IV NF-κB factors consists of a family of nuclear factors of activated T-cells (NF-AT) transcription factors, and induce the downstream genes in the adaptive immune system.
Figure 5. 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-kappaB essential modulator (NEMO)-binding domain; PEST, proline-, glutamic acid-, serine-, and threonine-rich region; TAD, transactivation domain; ZF, zinc finger

Figure 6. The 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.

Insect AMPs

Insects produce a battery of antimicrobial peptides (AMPs) and the synthesis of AMPs is controlled by the Toll and Imd pathways in *Drosophila*. If there is no microbial infection, the level of AMPs is non-detectable in the insect body. Once the hosts are infected by microorganisms, AMPs are rapidly synthesized within hours in the insect fat body, the functional analog of the mammalian liver. To defend the intruder from midgut or epidermis, AMPs are secreted and accumulated in the insect hemolymph. AMPs, which normally contain between 12 and 50 amino acids, are small cationic proteins. AMPs are synthesized as inactive precursors, which require proteolysis to form the active proteins/peptides. According to their secondary structures, AMPs can be classified into four groups: (i) alpha-helix proteins (Bulet et al., 1999; Dimarcq et al., 1998), (ii) Cysteine-rich AMPs (Barbault F, 2003), (iii) proline-rich peptides, (iv) glycine-rich peptides (Bulet and Stocklin, 2005; Otvos, 2000). AMPs include several positive charge residues, such as Arginine, Lysine, or Histidine and more than 50% hydrophobic residues. In solution, the AMPs are randomly folded. When AMPs detect and bind to bacteria, amphipathicity allows AMPs to partition into the bacterial membrane and form pores to change the permeability of the cell wall and lyze bacteria (Dathe and Wieprecht, 1999; Durr et al., 2006; Hancock and Lehrer, 1998; Oren and Shai, 1998; Papagianni, 2003; Pieters et al., 2009; Sitaram and Nagaraj, 2002; Wang, 2006).

In *Drosophila*, AMPs can be classified into seven families, including Attacin (Kockum et al., 1984), Diptericin (Dimarcq et al., 1988; Wicker et al., 1990), Cecropin (Boman et al., 1991), Drosocin (Bulet et al., 1993) that inhibits Gram-negative bacteria; Defensin (Hoffmann and Hetru, 1992) with activities against Gram-positive bacteria;
Drosomycin (Fehlbaum et al., 1994) and Metchnikowin (Levashina et al., 1995) with activities against fungi (Table 4).

In *M. sexta*, many AMPs were identified, including Lysozyme with bacteriolytic activity (Mulnix and Dunn, 1994; Russell and Dunn, 1991), Cecropin that can inhibit Gram-negative bacteria (Dickinson et al., 1988), Moricin with antimicrobial activity against Gram-negative and Gram-positive bacteria (Dai et al., 2008), Lebocin A, B, C and Gloverin, which can inhibit Gram-negative and Gram-positive bacteria, and fungi (Rao et al., 2012; Rayaprolu et al., 2010; Xu et al., 2012). Many AMPs are common in insects and have been identified in several insect orders; however, Moricin and Gloverin are only found in *Lepidoptera* so far (Table 5).
Table 4 Antimicrobial Peptides in *Drosophila*

<table>
<thead>
<tr>
<th>Peptides</th>
<th>Main Activity</th>
<th>Concentration</th>
<th>3-D structure</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Diptericin</td>
<td>Gram-negative bacteria</td>
<td>0.5 μM</td>
<td>nd</td>
<td>(Dimarcq et al., 1988; Wicker et al., 1990)</td>
</tr>
<tr>
<td>Attacin</td>
<td>Gram-negative bacteria</td>
<td>nd</td>
<td>nd</td>
<td>(Kockum et al., 1984)</td>
</tr>
<tr>
<td>Drosocin</td>
<td>Gram-negative bacteria</td>
<td>40 μM</td>
<td>(Bulet et al., 1993)</td>
<td></td>
</tr>
<tr>
<td>Cecropin</td>
<td>Gram-positive and Gram-negative bacteria, Fungi</td>
<td>20 μM</td>
<td>(Boman et al., 1991; Ekengren and Hultmark, 1999)</td>
<td></td>
</tr>
<tr>
<td>Defensin</td>
<td>Gram-positive bacteria</td>
<td>1 μM</td>
<td>(Hoffmann and Hetru, 1992; Imler and Bulet, 2005)</td>
<td></td>
</tr>
<tr>
<td>Drosomycin</td>
<td>Fungi</td>
<td>100 μM</td>
<td>(Fehlbaum et al., 1994; Landon et al., 1997)</td>
<td></td>
</tr>
<tr>
<td>Metchnikowin</td>
<td>Fungi</td>
<td>10 μM</td>
<td>nd</td>
<td>(Levashina et al., 1995)</td>
</tr>
</tbody>
</table>
Table 5 Antimicrobial Peptides in *M. sexta*

<table>
<thead>
<tr>
<th>Peptides</th>
<th>Main Activity</th>
<th>3-D structure</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lysozyme</td>
<td>Bacteriolytic</td>
<td>nd</td>
<td>(Mulnix and Dunn, 1994; Russell and Dunn, 1991)</td>
</tr>
<tr>
<td>Cecropin</td>
<td>Gram-negative bacteria</td>
<td>nd</td>
<td>(Dickinson et al., 1988)</td>
</tr>
<tr>
<td>Attacin</td>
<td>nd</td>
<td>nd</td>
<td></td>
</tr>
<tr>
<td>Moricin</td>
<td>Gram-positive and Gram-negative bacteria</td>
<td></td>
<td>(Dai et al., 2008)</td>
</tr>
<tr>
<td>Lebocin A, B, C</td>
<td>Gram-positive and Gram-negative bacteria, Fungi</td>
<td>nd</td>
<td>(Rao et al., 2012; Rayaprolu et al., 2010)</td>
</tr>
<tr>
<td>Gloverin</td>
<td>Gram-positive and Gram-negative bacteria, Fungi</td>
<td>nd</td>
<td>(Xu et al., 2012)</td>
</tr>
<tr>
<td>Defensin</td>
<td>nd</td>
<td>nd</td>
<td></td>
</tr>
</tbody>
</table>
Questions and Hypotheses

Based on the previous research, the best models to investigate the innate immune pathways are *Drosophila melanogaster* and *Manduca sexta* both in vivo and in vitro experiments. In my research, I identified several components of innate immunity in *M. sexta* and also demonstrated the synthesis of AMPs is mediated by highly conserved mechanism, Toll and Imd pathways in invertebrates.

As introduced above, the Toll pathway was initially identified for its role in the early dorsal – ventral (DV) patterning of the *Drosophila* embryo. The pathway has been well studied in *D. melanogaster*, but less characterized in other insect species. Although *Toll* and *Spz* genes have been identified in other insects, interactions between a Toll receptor and a Spz from the same insect species other than *D. melanogaster* have not been demonstrated. In chapter 2, I will propose that there is a Toll-Spz pathway in *M. sexta* and this pathway is conserved in insects.

In mammals, Toll-like receptor 4 (TLR4) is the receptor that recognizes LPS. MD2 is required for the recognition of LPS and binding to TLR4. In insects, the IMD and Toll signaling pathways have been identified; however receptors that can recognize LPS have not been discovered in invertebrates. In chapter 3, I hypothesize that insects have a similar Toll-ML-LPS pathway and the recognition of LPS is highly evolutionary conserved from insects to humans.

NF-κB transcription factors play essential roles in the regulation of innate immune responses. In *Drosophila*, Relish containing both a Rel-homology domain and an inhibitor κB (IκB)-like domain has well been studied. However, how the short isoforms of Relish containing only RHD or IκB-like domain regulate the gene expression is not
clear. In chapter 4, I investigated the functions of two short isoforms of Relish and Dorsal from *M. sexta* and the role of MsDorsal and MsRel2 heterodimer in the activation of AMP production.

Recognition of pathogens by insect pattern recognition receptors (PRRs) is critical to mount effective immune responses. Identifying new members of surveillance molecules and investigating functions of new members can advance our understanding on how the PRRs recognize pathogen-associated molecular patterns (PAMPs) and their crucial roles in immune responses. In chapter 5, I aim to characterize and functional analyze βGRP3, a new member of the *M. sexta* βGRP/BGRP family.

Insects can synthesize a variety of AMPs with activities against bacteria, fungi, viruses and some parasites. Some AMPs, such as Cecropin, Attacin and Defensin, are common and present in most insect species, while others like Moricin, Gloverin and Lebocin have been identified only in lepidopteran insects so far. Gloverin is a glycine-rich and heat stable antimicrobial protein. In chapter 6, I will investigate the expression profile, binding ability and antimicrobial activity of *M. sexta* Gloverin (MsGlv).
CHAPTER 2

A TOLL-SPÄTZLE PATHWAY IN THE TOBACCO HORNWORM, *MANDUCA SEXTA*

Any questions about whether a particular use is authorized and any requests for permission to publish, reproduce, distribute, display or make derivative works from any Elsevier published works, images or other material should be directed to the Elsevier Rights Department.

© 2012 Elsevier Ltd. All rights reserved.
Abstract

Insects synthesize a battery of antimicrobial peptides (AMPs) and expression of AMP genes is regulated by the Toll and Imd (immune deficiency) pathways in *Drosophila melanogaster*. The *Drosophila* Toll pathway is activated after Spätzle (Spz) is cleaved by Spätzle processing enzyme (SPE) to release the active C-terminal C106 domain (*DmSpz-C106*), which then binds to the Toll receptor to initiate the signaling pathway and regulate expression of AMP genes such as *drosomycin*. *Toll* and *Spz* genes have been identified in other insects, but interaction between Toll and Spz and direct evidence for a Toll-Spz pathway in other insect species have not been demonstrated. Our aim is to investigate a Toll-Spz pathway in *Manduca sexta*, and compare *M. sexta* and *D. melanogaster* Toll-Spz pathways. Co-immunoprecipitation (Co-IP) assays showed that *MsToll* (the ecto-domain of *M. sexta* Toll) could interact with *MsSpz-C108* (the active C-terminal C108 domain of *M. sexta* Spz), but not with full-length *MsSpz* and *DmToll* (the ecto-domain of *D. melanogaster* Toll) could interact with *DmSpz-C106* but not *DmSpz*, suggesting that the Toll receptor only binds to the active C-terminal domain of Spz. Co-expression of *MsToll-MsSpz-C108*, but not *MsToll-MsSpz*, could up-regulate expression of *drosomycin* gene in *Drosophila* S2 cells, indicating that *MsToll-MsSpz-C108* complex can activate the Toll signaling pathway. *In vivo* assays showed that activation of AMP genes, including *cecropin, attacin, moricin* and *lebocin*, in *M. sexta* larvae by purified recombinant *MsSpz-C108* could be blocked by pre-injection of antibody to *MsToll*, further confirming a Toll-Spz pathway in *M. sexta*, a lepidopteran insect.
Introduction

The innate and adaptive immune systems are two major branches of the defense system in multicellular organisms. Innate immunity is the first defensive line that controls initial steps of immune responses. It can also profoundly impact the establishment of adaptive immune responses (Medzhitov and Janeway, 2000a; Medzhitov and Janeway, 1997). In innate immune responses, a group of germline-encoded pattern recognition receptors (PRRs) can recognize and bind to conserved pathogen-associated molecular patterns (PAMPs) present on the invading microorganisms, such as bacteria and fungi, but not on the host cells (Janeway, 1989; Medzhitov and Janeway, 1997). The innate immune system is composed of humoral and cellular components. Cellular responses mainly include blood cells (hemocytes)-mediated immune responses such as nodule formation, phagocytosis and melanotic encapsulation, whereas synthesis of antimicrobial peptides (AMPs) and activation of the prophenoloxidase system are major components of humoral responses (Kanost et al., 2004; Rao et al., 2010). Expression of AMP genes in Drosophila melanogaster is regulated by the Toll and immune deficiency (Imd) pathways (Choe et al., 2002; De Gregorio et al., 2002; Lemaitre et al., 1995; Lemaitre et al., 1996; Ramet et al., 2002). The Drosophila Toll pathway is activated by Gram-positive bacteria and fungi, resulting in systemic production of AMPs (Aggarwal and Silverman, 2008; Hetru and Hoffmann, 2009). Moreover, the Toll signaling pathway and other pathways are involved in controlling hemocyte proliferation and density (Sorrentino et al., 2004; Zettervall et al., 2004), as well as melanization (Bettencourt et al., 2004a).

The Toll pathway was initially identified in early Drosophila embryonic development as, the dorsal–ventral (DV) patterning of the embryo, and the dorsal group
of genes includes \textit{Toll}, \textit{tube}, \textit{pelle}, \textit{cactus}, the NF-kB homolog \textit{dorsal}, and seven genes up-stream of the \textit{Toll} (Belvin and Anderson, 1996; Morisato and Anderson, 1995). Tolls and Toll-like receptors (TLRs) have been identified in many animal species, including mammals (Shinkai et al., 2006; Takeuchi and Akira, 2010), chicken (Fukui et al., 2001), fish (Tsujita et al., 2004), insects (Ao et al., 2008b; Christophides et al., 2002; Evans et al., 2006; Imamura and Yamakawa, 2002; Kanzok et al., 2004; Luna et al., 2002; Yamagata et al., 1994), shrimp (Yang et al., 2007), and sponge (Wiens et al., 2007).

However, mammalian TLRs function as pattern recognition receptors but do not have a role in development (Kimbrell and Beutler, 2001), whereas the \textit{Drosophila} Toll pathway is involved in both immunity (Lemaitre et al., 1996) and developmental processes (Belvin and Anderson, 1996; Halfon et al., 1995; Qiu et al., 1998).

Activation of the \textit{Drosophila} Toll pathway is preceded by activation of Spätzle (Spz), the Toll receptor ligand (Morisato and Anderson, 1994; Schneider et al., 1994). Under non-signaling conditions, a predominantly hydrophobic C-terminal domain of Spz is masked by a prodomain of Spz. Embryonic dorsal-ventral patterning, Gram-positive bacterial and fungal cell wall components, and virulence factors can activate Spz (Valanne et al., 2011). Spz is processed into its active C-terminal C-106 domain in a process that involves activation of a cascade of serine proteinases. Proteolysis of Spz causes a conformational change, which exposes determinants of C-106 domain that are critical for binding to the Toll receptor (Arnot et al., 2010). Two Spz-C106 dimers bind to two Toll receptors. Then, the binding triggers a conformational change in the Toll receptors to form stable dimers (Hu et al., 2004). The dimeric Toll complexes interact with an adaptor protein MyD88 via intracellular TIR (Toll-interleukin 1 resistance).
domains (Horng and Medzhitov, 2001; Sun et al., 2002; Tauszig-Delamasure et al., 2002). Tube and kinase Pelle are recruited by MyD88 to form a MyD88-Tube-Pelle heterotrimeric complex through their death domain (DD)-mediated interactions (Moncrieffe et al., 2008; Sun et al., 2002; Xiao et al., 1999). Intracellular signaling leads to phosphorylation and degradation of Cactus, which in turn releases Dorsal-related immunity factor (Dif) and/or Dorsal to translocate to the nucleus and activate transcription of AMP genes (Imler and Hoffmann, 2001; Wu and Anderson, 1998).

The Toll signaling pathway has been well studied in D. melanogaster, but less characterized in other insect species. Although Toll and Spz genes have been identified in other insects, including Anopheles gambiae (Christophides et al., 2002; Luna et al., 2002), Aedes aegypti (Kanzok et al., 2004), Apis mellifera (Evans et al., 2006), Bombyx mori (Imamura and Yamakawa, 2002; Wang et al., 2007), and Manduca sexta (An et al., 2010; Ao et al., 2008b), interaction between a Toll receptor and a Spz from the same insect species other than D. melanogaster has not been demonstrated. Previously a Toll receptor from M. sexta was discovered (Ao et al., 2008b), and Spz-1 gene has also been identified (An et al., 2010). M. sexta Spz-1A (MsSpz) was cleaved and activated by proteinase HP8 to release the active C-terminal domain MsSpz-C108 (An et al., 2010). Injection of MsSpz-C108 into M. sexta larvae up-regulates several AMP genes (An et al., 2010), suggesting that there is a Toll pathway in M. sexta. In this study, we showed direct interaction between M. sexta Toll (MsToll) and MsSpz-C108 and further confirmed a Toll-Spz pathway in M. sexta by both in vitro and in vivo assays. We established stable Drosophila S2 cell lines expressing M. sexta and D. melanogaster Tolls (MsToll and DmToll) and their ecto-domains (MsToll^{ecto} and DmToll^{ecto}), Spz proteins (MsSpz and
DmSpz) and their active C-terminal domains (MsSpz-C108 and DmSpz-C106). Co-immunoprecipitation (Co-IP) assays showed that MsToll^ecto and DmToll^ecto could interact with MsSpz-C108 and DmSpz-C106, but not MsSpz and DmSpz, respectively. Co-expression of MsToll-MsSpz-C108 and DmToll-DmSpz-C106 in S2 cells could up-regulate drosomycin but not diptericin gene. Activation of AMP genes, including cecropin-6, attacin-1, attacin-2, moricin and lebocin, by recombinant MsSpz-C108 in M. sexta larvae could be blocked by pre-injection of antibody to MsToll. Our results demonstrated a Toll-Spz pathway in M. sexta, a Lepidopteran insect.

Materials and Method

Insect rearing and cell line

M. sexta eggs were originally purchased from Carolina Biological Supplies (Burlington, NC, USA). Larvae were reared on an artificial diet at 25°C (Dunn, 1983), and the fifth instar larvae were used for the experiments. D. melanogaster Schneider S2 cells were purchased from the American Type Culture Collection (ATCC).

Construction of recombinant pMT/BiP/V5- His A expression vectors

Complementary DNA (cDNA) fragments encoding MsToll (residues 13-963), MsToll^ecto (residues 13-718), MsTIR (residues 766-963), MsSpz (residues 20-295), MsSpz-C108 (residues 188-295), DmToll (residues 28-1097), DmToll^ecto (residues 28-805), DmTIR (residues 858-1097), DmSpz (residues 26-244), and DmSpz-C106 (residues 139-244) were amplified by polymerase chain reaction (PCR) using forward and reverse
primers listed in Table 6. Forward primers for \textit{MsSpz}, \textit{MsSpz-C108}, \textit{DmSpz} and \textit{DmSpz-C106} contain codons for an in-frame Flag sequence and a \textit{KpnI} site, while reverse primers contain a stop codon followed by a \textit{PmeI} site. Forward primers for \textit{MsToll}, \textit{MsToll^{ecto}}, \textit{MsTIR}, \textit{DmToll}, \textit{DmToll^{ecto}} and \textit{DmTIR} contain a \textit{KpnI} site, while reverse primers contain an \textit{ApaI} site. 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). Plasmid DNAs in T-vectors were purified using PureYield™ Plasmid Miniprep System (A1222, Promega) according to the manufacturer’s instruction and digested with \textit{KpnI/PmeI} or \textit{KpnI/ApaI}, DNA fragments were recovered and inserted into \textit{KpnI/PmeI} or \textit{KpnI/ApaI} digested pMT/BiP/V5-His A vector (V413020, Invitrogen) using T4 DNA ligase (M0202L, New England Biolabs). Recombinant expression vectors 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 generate stable S2 cell lines.

\textbf{Cell culture and establishment of stable S2 cell lines}

\textit{D. melanogaster} S2 cells were maintained at 27°C in Insect Cell Culture Media (SH30610.02, Hyclone), supplemented with 10% heat-inactivated fetal bovine serum (catalog #10082063, Invitrogen) containing 1% penicillin-streptomycin solution (G6784, Sigma-Aldrich). For DNA transfection, cells were seeded overnight in serum-free
medium (SH30278.01, Hyclone). GenCarrier-1™ transfection reagent (#31-00110, Epoch Biolabs) was used for transient transfection based on the manufacturer’s instructions. Cells in culture dishes or plates were grown to 70% confluence prior to transfection. DES® – Inducible/Secreted Kit with pCoBlast (K5130-01, Invitrogen) was used to construct stable S2 cell lines. To select stable S2 cells expressing recombinant proteins, pCoBlast (Invitrogen) was cotransfected with recombinant pMT/BiP/V5-His A vectors. After 48h transfection, S2 cells were centrifuged and resuspended in complete growth medium containing 25 μg/ml Blasticidine S hydrochloride (No.15205, Sigma-Aldrich). Resistant colonies appeared 1 week later.

**Western blot analysis and immunoprecipitation (Co-IP) assay**

For Western blot analysis, copper sulfate (final concentration of 250 μM) was added to the stable S2 cell lines (2×10^6 cells/well) in 6-well plates, and protein expression was induced for 48 h. Cell culture medium (2 ml each) was collected, 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, 0.5 mM PMSF, protease inhibitor cocktail (P8340, Sigma-Aldrich)]. The cell homogenates were incubated on ice for 15 min and sonicated briefly several times, and then centrifuged at 15,000 g for 15 min at 4°C. The supernatants were collected as cell extracts for Western blot analysis. The 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% Sodium dodecyl sulfate (SDS) - Polyacrylamide Gel Electrophoresis (PAGE) and proteins were transferred to nitrocellulose membranes (162-0097, Bio-Rad). The membrane was blocked with 5% BSA in Tris-buffered saline (10 mM Tris-HCl, pH7.6,
150 mM NaCl) containing 0.05% Tween-20 (Tris buffered saline (TBS)-T) at room temperature for at least 3 h 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 2 h at room temperature. After washing four times with TBS-T (10 min each time), the signal was developed by using ECL Chemiluminescence Detection Kit (RPN2134, GE Healthcare) or 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, horseradish peroxidase-conjugate anti-mouse antibody (SC-2005, Santa Cruz Biotechnology, 1:10,000) was used as secondary antibody for chemiluminescence, and alkaline phosphatase-conjugate anti-mouse antibody (A4312, Sigma-Aldrich, 1:10,000) was used as secondary antibody for color development.

Immunoprecipitation (Co-IP) assay was performed by using 300 μl 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 pre-cleared for 30 min with 30 μl Protein G Sepharose (50% slurry, No.17-0618-01, GE Healthcare) in a total volume of 500 μl. After centrifugation, the supernatant was incubated with anti-Flag M2 or anti-V5 antibody (final concentration of 10 μg/ml) at 4°C for 10 h with gentle rocking. Then, 30 μl 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, re-suspended in 50 μl of 1×SDS
sample buffer, boiled at 95°C for 5 min, and used for subsequent immunoblotting analysis.

**Purification of recombinant MsSpz and MsSpz-C108**

To purify recombinant proteins, stable S2 cells expressing MsSpz or MsSpz-C108 in 75-cm² flasks were induced for protein expression after addition of copper sulfate (final concentration of 250 μM). Cell culture medium was collected continuously for 10 days starting at 24 h after protein expression by collecting culture medium every day and re-suspending the cells with fresh medium. To purify recombinant proteins, cell culture medium was combined, cell debris was removed by centrifugation at 1000 g for 10 min at room temperature, and cell-free medium was incubated overnight at 4°C with 500 μl of Anti-Flag M2 agarose beads (A2220, Sigma-Aldrich) equilibrated with initial buffer (50 mM Tris-HCl, 150 mM NaCl, pH7.4). Anti-Flag M2 agarose beads were then packed into a column (1.5 cm × 1 cm) and cell-free medium was reloaded to the column several times at a flow rate of 0.05 ml/min. Then, the column was washed with the initial buffer until A280 of the effluent was near zero. The bound proteins were sequentially eluted with one ml aliquots of the elution buffer (0.1 M glycine-HCl, pH 3.5, 1% Triton X-100) into vials containing 100 μl of 1 M Tris-base, pH 8.0. Fractions were analyzed by 12% or 15% SDS-PAGE. Fractions containing recombinant MsSpz or MsSpz-C108 were de-salted with D-salt™ Excellulose™ GF-5 desalting column (#1851850, Pierce) pre-equilibrated with deionized water (dH₂O), and fractions containing recombinant proteins were pooled and concentrated.
Cleavage of recombinant MsSpz by M. sexta larval plasma

To test whether MsSpz can be activated by M. sexta larval plasma, induced cell-free plasma was collected from immunized M. sexta larvae injected with a mixture of heat-killed yeast (Saccharomyces cerevisiae, $1.7 \times 10^7$ cells), dry Micrococcus luteus (33 μg) and heat-killed Escherichia coli XL-1 blue ($3.3 \times 10^7$ cells) at 24 h post-injection, and the control plasma was collected from naïve larvae. Purified MsSpz (50 μl of 100 ng/μl, 5 μg) was incubated with 10 μl induced or control plasma in a total of 75 μl at room temperature for 2 h. Then 25 μl 4×SDS loading buffer was added to the reaction mixture. The reaction mixture was heated to 95°C for 5 min and aliquots (10 μl, an equivalent of 1 μl original plasma and 0.5 μg recombinant protein) were analyzed by 15 % SDS-PAGE and immunoblotting using mouse monoclonal anti-Flag M2 antibody (F-1804, Sigma-Aldrich, 1:5000 dilution) or rabbit polyclonal anti-MsSpz-C108 antibody (a gift from Dr. Michael Kanost, Kansas State University) (1:1000 dilution) as primary antibody.

Dual-Luciferase Reporter Assay

For Dual-luciferase reporter assays, S2 cells were plated in 96-well culture plates ($10^4$ cells/well) overnight in serum-free medium. These S2 cells were then transiently co-transfected with recombinant pMT/BiP/V5-His A expression plasmid (0.3 μg), pGL3B, pGL3B-drosomycin, pGL3B-diptericin or pGL3B-attacin firefly luciferase reporter plasmid (0.15 μg) (Rao et al., 2011), and renilla luciferase reporter plasmid (0.015 μg) (as an internal standard) (pRL-TK, Promega). After overnight transfection, serum-free medium was replaced with complete growth medium containing 250 μM copper sulfate (final concentration) for protein expression, and firefly luciferase and renilla luciferase
activities were measured 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) was obtained as the ratio of firefly luciferase activity to renilla luciferase activity. RLA from S2 cells co-transfected with empty pMT/BiP/V5-His A and pGL3B (empty reporter vector) plasmids was 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.

Real-time PCR analysis

S2 cells were cultured in 6-well culture plates (10^5 cells/well) overnight in serum-free medium, and transiently transfected with recombinant pMT/BiP/V5-His A expression vectors (3 μg each vector). After overnight transfection, serum-free medium was replaced with complete growth medium containing 250 μM copper sulfate to induce expression of recombinant proteins. After protein expression for 48 h, total RNAs were extracted from these S2 cells using TRIzol® Reagent (T9424, Sigma-Aldrich) according to the manufacturer’s instructions. Residual genomic DNA was digested by RQ1 RNase-free DNase (M6101, Promega). cDNA was prepared from 1 μg total RNA in a 25 μl reaction using moloney murine leukemia virus (M-MLV) reverse transcriptase (M1701, Promega) with an anchor-oligo(dT)\textsubscript{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 6) were designed based
on the sequences of \textit{rp49}, \textit{drosomycin} and \textit{diptericin}. The real-time PCR was performed in 20 μl reactions containing 10 μl 2×SYBR® GreenER™ qPCR SuperMix Universal (No. 204141, Qiagen), 4 μl dH2O, 4 μl diluted cDNA template, and 1 μl forward and reverse primers (10 pmol each). Real-time PCR program was 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 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 was used as the calibrator. The expression levels of \textit{drosomycin} and \textit{diptericin} transcripts in other cDNA samples were calculated by the $2^{-\Delta\Delta CT}$ method, which stands for the n-fold difference in relative expression to the calibrator. All the data were presented as relative mRNA expression. These experiments were repeated at least three times.

**Injection of MsSpz and MsSpz-C108 into \textit{M. sexta} larvae**

Day 1 fifth instar \textit{M. sexta} naïve larvae were injected with water (as a negative control), purified recombinant MsSpz (3 μg /larva), or MsSpz-C108 (1 μg /larva). Twenty hours later, fat body and hemocyte samples were collected, total RNA was isolated with TRIzol® Reagent (Sigma-Aldrich), and cDNA was prepared with ImProm-II reverse transcriptase (Promega) as described above. Each cDNA sample (diluted 1:50) was used as template for real-time PCR analysis. \textit{M. sexta} ribosomal protein S3 (\textit{rpS3}) gene was used as an internal standard to normalize the amount of RNA template. AMP genes, including \textit{cecropin-6} (AY232302.1), \textit{attacin-1} (DQ072728.1), \textit{attacin-2} (AY232304.1),
lebocin-b and lebocin-c (GU563901.1 and GU563900.1), moricin (AY232301.1) and lysozyme (S71028.1) were detected with primer pairs (10 pmol each) listed in Table S1. cDNA sample from naïve larvae was used as the calibrator. The expression levels of AMP genes from other samples were calculated by the $2^{-\Delta\Delta CT}$ method. All the data were presented as relative mRNA expression. These experiments were repeated at least three times.

**Injection of antibody to MsToll into *M. sexta* larvae**

To test whether MsSpz-C108 binds to MsToll in *M. sexta* larvae to stimulate expression of AMP genes, day 1 fifth instar *M. sexta* naïve larvae were pre-injected with purified IgG to the ecto-domain of MsToll (Toll Ab, 5 μg/larva) or IgG from pre-immune rabbit serum (Control Ab, 5 μg/larva), and these larvae were then injected with water, purified recombinant MsSpz (3 μg/larva), MsSpz-C108 (1 μg/larva), TLR grade peptidoglycan from *Staphylococcus aureus* (PG-SA) or *Escherichia coli* (PG-K12) (Invivogen), or without second injection (control) at 1 h after pre-injection of antibody. Twenty hours later, fat body and hemocyte samples were collected for quantitative real-time PCR analysis. Total RNA and cDNA samples were prepared as described above. *M. sexta* ribosomal protein S3 (*rpS3*) gene was used as an internal standard to normalize the amount of RNA template. Expression of cecropin-6, attacin-1, lebocin-b/c, moricin and lysozyme genes were determined by real-time PCR as described above. These experiments were repeated at least three times.
Data analysis

One representative set of data was used to make figures using the Graphpad Prism software, and the significance of difference was determined by an unpaired t-test or by one way ANOVA followed by a Tukey’s multiple comparison test with the Graphpad InStat software (GraphPad, San Diego, CA).
Table 6 PCR primers used in this study

<table>
<thead>
<tr>
<th>Primers</th>
<th>Forward Primer (5' → 3')</th>
<th>Reverse Primer (5' → 3')</th>
</tr>
</thead>
<tbody>
<tr>
<td>For protein expression</td>
<td></td>
<td></td>
</tr>
<tr>
<td>DmTIR_N_KpnI</td>
<td>ATGGTACAAAGTTCG</td>
<td>DmTIR_C_Apal</td>
</tr>
<tr>
<td></td>
<td>ATGCCCTCATCTCG</td>
<td>ATGGGCCCTACGTCG</td>
</tr>
<tr>
<td></td>
<td>MsTIR_N_KpnI</td>
<td>MsTIR_C_Apal</td>
</tr>
<tr>
<td></td>
<td>ATGGTACACCGTGTCG</td>
<td>ATGGGCCCCTTTTGTAG</td>
</tr>
<tr>
<td></td>
<td>ACGCGTTTGTGTCTTTGCA</td>
<td>CAAGGACTCGCGCCC</td>
</tr>
<tr>
<td></td>
<td>DmToll_N_KpnI</td>
<td>DmToll_C_Apal</td>
</tr>
<tr>
<td></td>
<td>ATGGTACCAATGAGTC</td>
<td>ATGGGCCCTACGTCG</td>
</tr>
<tr>
<td></td>
<td>GACTAAAGGCC</td>
<td>CTCTGTTTGCG</td>
</tr>
<tr>
<td></td>
<td>MsToll_N_KpnI</td>
<td>MsToll_C_Apal</td>
</tr>
<tr>
<td></td>
<td>ATGGTACCAATGCAGA</td>
<td>ATGGGGCCCCTTTTGTAG</td>
</tr>
<tr>
<td></td>
<td>TGCTGAGCTGGCT</td>
<td>CAAGGACTCGCGC</td>
</tr>
<tr>
<td></td>
<td>MsToll_N_KpnI</td>
<td>MsToll_C_Apal</td>
</tr>
<tr>
<td></td>
<td>ATGGGCCCTTTGTAG</td>
<td>ATGGGGCCCCTTTTGTAG</td>
</tr>
<tr>
<td></td>
<td>CAAGGACTCGCGC</td>
<td>CAAGGACTCGCGC</td>
</tr>
<tr>
<td></td>
<td>MsSpz_N_KpnI</td>
<td>MsSpz_C_Pmel</td>
</tr>
<tr>
<td></td>
<td>ATGGTACCATGGACTACAAAGGACGAG</td>
<td>CGGTTTAAACTCATA</td>
</tr>
<tr>
<td></td>
<td>ACTGACAAGGCGGCCGCGC</td>
<td>TGATTGCTAACATTGG</td>
</tr>
<tr>
<td></td>
<td>ATACAAAATGAAAGACTGCTCAGTTCAGTGCACAC</td>
<td></td>
</tr>
<tr>
<td></td>
<td>MsSpz-C108_N_KpnI</td>
<td>MsSpz_C_Pmel</td>
</tr>
<tr>
<td></td>
<td>ATGGTACCATGGACTACAAAGGACGAG</td>
<td>CGGTTTAAACTCACC</td>
</tr>
<tr>
<td></td>
<td>ACTGACAAGGCGGCCGCGC</td>
<td>CAGTCTTCAACGC</td>
</tr>
<tr>
<td></td>
<td>ATACAAAATGAAAGACTGCTCAGTTCAGTGCACAC</td>
<td></td>
</tr>
<tr>
<td></td>
<td>DmSpz_N_KpnI</td>
<td>DmSpz_C_Pmel</td>
</tr>
<tr>
<td></td>
<td>ATGGTACCATGGACTACAAAGGACGAG</td>
<td>CGGTTTAAACTCACC</td>
</tr>
<tr>
<td></td>
<td>ACTGACAAGGCGGCCGCGC</td>
<td>CAGTCTTCAACGC</td>
</tr>
<tr>
<td></td>
<td>TATGATGACGCCCATGTG</td>
<td></td>
</tr>
<tr>
<td></td>
<td>MsSpz-C106_N_KpnI</td>
<td></td>
</tr>
<tr>
<td></td>
<td>ATGGTACCATGGACTACAAAGGACGAG</td>
<td></td>
</tr>
<tr>
<td></td>
<td>ACTGACAAGGCGGCCGCGC</td>
<td></td>
</tr>
<tr>
<td></td>
<td>ATACAAAATGAAAGACTGCTCAGTTCAGTGCACAC</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Dp_N_37</td>
<td>Dpt_C_306</td>
</tr>
<tr>
<td></td>
<td>GCAATCGCTTTCTACTTTG</td>
<td>TCTGTAGGTGTAGGT</td>
</tr>
<tr>
<td></td>
<td>TGG</td>
<td>GCT</td>
</tr>
<tr>
<td></td>
<td>rp49_N_1</td>
<td>rp49_C_317</td>
</tr>
<tr>
<td></td>
<td>ATGACCATCGCGCCAGCCAGCAT</td>
<td>TTCTTGGAGGAGAGC</td>
</tr>
<tr>
<td></td>
<td>CAT</td>
<td>CCGT</td>
</tr>
<tr>
<td></td>
<td>Cec6_N</td>
<td>Cec6_C</td>
</tr>
<tr>
<td></td>
<td>ATTCCTTCGTCTTCTGCATTG</td>
<td>ATGTAATACTGTAAT</td>
</tr>
<tr>
<td></td>
<td>GC</td>
<td>GACAC</td>
</tr>
<tr>
<td></td>
<td>Mor_N</td>
<td>Mor_C</td>
</tr>
<tr>
<td></td>
<td>ATAAGCAGGGCTGCGC</td>
<td>AGAAGATTCCGAAG</td>
</tr>
<tr>
<td></td>
<td>AAGG</td>
<td>GGAGAAC</td>
</tr>
<tr>
<td></td>
<td>Att1_N</td>
<td>Att1_C</td>
</tr>
<tr>
<td></td>
<td>CAATGGGACACGGCGC</td>
<td>ACGGCCCACCTTTGT</td>
</tr>
<tr>
<td></td>
<td>ACTCTTACC</td>
<td>CCTTGAAC</td>
</tr>
<tr>
<td></td>
<td>Leb_N</td>
<td>Leb_C</td>
</tr>
<tr>
<td></td>
<td>ACCTGCTAGTGAGCAGA</td>
<td>CGCAGATTATGAGTT</td>
</tr>
<tr>
<td></td>
<td>CGAG</td>
<td>ACGACGA</td>
</tr>
<tr>
<td></td>
<td>Lys_N</td>
<td>Lys_C</td>
</tr>
<tr>
<td></td>
<td>GTGTGCCCTGGGTAGAGA</td>
<td>ATGCGCTTGGTGATGT</td>
</tr>
<tr>
<td></td>
<td>ATG</td>
<td>CGTC</td>
</tr>
</tbody>
</table>

46
<table>
<thead>
<tr>
<th>Primers</th>
<th>Forward Primer (5’ → 3’)</th>
<th>Reverse Primer (5’ → 3’)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Glo_N</td>
<td>CCGCAATAACGCTCAGATA</td>
<td>Glo_C TGTGGAAAGAGACCTTGGAG</td>
</tr>
<tr>
<td>Att2_N</td>
<td>GCCAACGCTTTGCACAAGCA</td>
<td>Att2_C CTTGCCGCCCACAGAGTAGTCGTT</td>
</tr>
<tr>
<td>rpS3_N</td>
<td>GTTGCAGGGGTGGTGTGGTTCC</td>
<td>rpS3_C CCGTTCTTGCCCTGTGGTC</td>
</tr>
</tbody>
</table>
Results

Expression of recombinant *M. sexta* and *D. melanogaster* Toll and Spz proteins in S2 cells

The Toll-Spz signaling pathway has been well understood in *D. melanogaster*, but is not well characterized in other insect species. In *M. sexta*, *Toll* and *Spz-1* genes have been identified (An et al., 2010; Ao et al., 2008b). In order to investigate a Toll-Spz pathway in *M. sexta* and compare *M. sexta* and *D. melanogaster* Toll pathways in S2 cells, we established stable S2 cell lines expressing Toll receptors (*Ms*Toll and *Dm*Toll) and their TIR (*Ms*TIR and *Dm*TIR) and ecto-domains (*Ms*Toll\textsuperscript{ecto} and *Dm*Toll\textsuperscript{ecto}), as well as Spz proteins (*Ms*Spz and *Dm*Spz) and their active C-terminal domains (*Ms*Spz\textsuperscript{-C108} and *Dm*Spz\textsuperscript{-C106}). Immunoblotting results showed that recombinant *D. melanogaster* and *M. sexta* Spz proteins and their active C-terminal domains were detected in both cell culture media and cell lysates (Fig. 7A and B). For the active carboxyl-terminal domains of Spz, a single protein band was detected in the cells and the cell culture media (Fig. 7A). For the full-length Spz proteins, multiple protein bands were detected (Fig. 7B), suggesting differential post-translational modifications of Spz proteins, and different modified forms of Spz were present in the cell culture media and cells. For the Toll receptors, *M. sexta* Toll (*Ms*Toll), *Ms*Toll\textsuperscript{ecto} (the ecto-domain) and *Ms*TIR (Fig. 7C and E), as well as *D. melanogaster* Toll (*Dm*Toll) and *Dm*TIR (Fig. 7E and F) were detected only in S2 cells but not in cell culture media. However, *Dm*Toll\textsuperscript{ecto} (the ecto-domain of *Dm*Toll) was detected as multiple protein bands in S2 cells and a single protein band in the cell culture medium (Fig. 7D), also suggesting differential post-translational modifications of *Dm*Toll\textsuperscript{ecto}. One of the *Dm*Toll\textsuperscript{ecto} protein bands was just above the 80
kDa marker (Fig. 7D, lane 2), which may be a cleavage product because the calculated mass of $Dm$Toll$^{\text{ecto}}$ is 95.8 kDa. For $Dm$Toll, cleavage products with sizes slightly larger than $Dm$TIR were also detected (Fig. 7F, lane 4), suggesting that $Dm$Toll may be cleaved in the ecto-domain at a site close to the transmembrane domain.
Figure 7. Expression of recombinant *M. sexta* and *D. melanogaster* Toll and Spätzle proteins in S2 cells.

*M. sexta* and *D. melanogaster* Toll (MsToll and DmToll) and their TIR (MsTIR and DmTIR) and ecto-domains (MsToll\textsuperscript{ecto} and DmToll\textsuperscript{ecto}), Spätzle proteins (MsSpz and DmSpz) and their active C-terminal domains (MsC108 and DmC106) were expressed in *Drosophila* S2 cells, and recombinant proteins in cell culture media (10 μl each from 2 ml total) and cell lysates (10 μl each from 400 μl total) were identified by immunoblotting using anti-Flag M2 antibody or anti-V5 antibody as the primary antibody, alkaline phosphatase-conjugate anti-mouse antibody as the secondary antibody and alkaline phosphatase (AP) conjugate color development kit. A: MsSpz-C108 (lanes 1 and 2) and DmSpz-C106 (lanes 3 and 4); B: MsSpz (lanes 1 and 2) and DmSpz (lanes 3 and 4); C: MsToll\textsuperscript{ecto} (lanes 1 and 2) and MsToll (lanes 3 and 4); D: DmToll\textsuperscript{ecto}; E: MsTIR (lanes 1 and 2) and DmTIR (lanes 3 and 4); F: DmTIR (lanes 1 and 2) and DmToll (lanes 3 and 4). Lanes 1 and 3: culture media; lanes 2 and 4: cell lysates. The calculated masses of these recombinant proteins are: MsSpz-C108 (14.6 kDa), DmSpz-C106 (14.1 kDa), MsSpz (34.1 kDa), DmSpz (27.2 kDa), MsToll\textsuperscript{ecto} (81.9 kDa), MsToll (109.4 kDa), DmToll\textsuperscript{ecto} (95.8 kDa), DmToll (128.5 kDa), MsTIR (26.2 kDa), and DmTIR (30.8 kDa).
Recombinant *M. sexta* Spz is activated by induced larval plasma

To investigate activation of *M. sexta* Spz, recombinant *MsSpz* and *MsSpz*-C108 were purified from stable S2 cell lines by antibody affinity chromatography (Fig. 8A and B). Both recombinant *MsSpz* and *MsSpz*-C108 contain a Flag-tag at the amino-terminus, and they were recognized by anti-Flag antibody (Fig. 8B). To determine whether recombinant *MsSpz* can be activated by proteinases in *M. sexta* larval plasma, purified *MsSpz* was treated with induced *M. sexta* cell-free plasma at room temperature for 2 h, and the cleavage products were detected by monoclonal anti-Flag (Fig. 8C) or polyclonal anti-*MsSpz*-C108 (Fig. 8D) antibody. Purified *MsSpz* (calculated mass of 34.1 kDa) and *MsSpz*-C108 (calculated mass of 14.6 kDa) could be recognized by anti-Flag and anti-*MsSpz*-C108 antibodies (Fig. 8C and D, lanes 2 and 4, respectively). After treating with induced *M. sexta* larval plasma, the *MsSpz* band disappeared and a major band at ~20 kDa was recognized by anti-Flag antibody (Fig. 8C, lane 3, arrowhead), which corresponded to the N-terminal fragment of *MsSpz* since the Flag-tag was at the N-terminus, and a cleavage product at ~12 kDa was recognized by antibody to *MsSpz*-C108 (Fig. 8D, lane 3, arrowhead), which corresponded to the C-terminal *MsSpz*-C108. A control experiment using naïve plasma showed that very little *MsSpz* was cleaved. These results suggest that *MsSpz* can be activated by proteinases in the hemolymph of *M. sexta* larvae, and these proteinases may also be induced by microorganisms. The cleavage *MsSpz*-C108 was smaller than the recombinant *MsSpz*-C108 (Fig. 8D, comparing lanes 3 and 4) since recombinant protein contained a Flag-tag at the N-terminus. In the plasma sample alone, endogenous *MsSpz* or *MsSpz*-C108 was not detected by anti-*MsSpz*-C108 antibody (Fig. 8D, lane 1), probably due to low concentration of Spz protein in plasma.
(An et al., 2010). But a band at ~23 kDa in the plasma sample was recognized by anti-Flag antibody (Fig. 8C, lane 1, asterisk), and several bands were recognized by anti-\textit{MsSpz-C108} antibody (Fig. 8D, lane 1), indicating non-specific recognition of plasma proteins by antibodies.
Figure 8. Activation of recombinant *M. sexta* Spätzle by induced larval plasma.

Recombinant *M. sexta* Spätzle (MsSpz) and its active C-terminal domain (MsC108 for MsSpz-C108) were purified by antibody affinity chromatography and analyzed by SDS-PAGE (A) (1 μg MsSpz and 1.5 μg MsSpz-C108) and Western blotting (B) (0.5 μg each protein) using anti-Flag antibody. Purified recombinant MsSpz was activated by induced cell-free plasma collected at 24 h after larvae were injected with a mixture of *S. cerevisiae*, *M. luteus* and *E. coli*, and cleavage products were analyzed by Western blotting using anti-Flag (C) or anti-MsSpz-C108 (D) as primary antibody, alkaline phosphatase-conjugate anti-mouse antibody as the secondary antibody and alkaline phosphatase (AP) conjugate color development kit. Panels A and B: lanes 1 and 2 were MsSpz and MsC108, respectively. Panels C and D: lane 1, induced plasma (1 μl); lane 2, purified recombinant MsSpz (0.5 μg); lane 3, induced plasma (1 μl) + purified MsSpz (0.5 μg); lane 4, purified recombinant MsC108 (MsSpz-C108) (0.5 μg). Arrowhead indicates the N-terminal cleavage product of MsSpz (C, lane 3) or C-terminal cleavage product of MsSpz (D, lane 3), while arrow indicates purified recombinant MsC108 (MsSpz-C108) (C and D, lane 4).
MsToll<sup>ecto</sup> interacts with MsSpz-C108 but not with MsSpz

In *D. melanogaster*, after Spz is activated by proteolysis, the C-terminal active domain (*DmSpz-C106*) is released from prodomain and recognized by the Toll receptor to initiate intracellular signaling pathway (Arnot et al., 2010; Hu et al., 2004; Morisato and Anderson, 1994; Schneider et al., 1994; Valanne et al., 2011). To determine whether MsSpz or MsSpz-C108 can bind to MsToll receptor, we over-expressed the ecto-domain of MsToll (MsToll<sup>ecto</sup>) with a V5-His-tag at the C-terminus, and MsSpz and MsSpz-C108 with a Flag-tag at the N-terminus in S2 cells (Fig. 7). Co-immunoprecipitation (Co-IP) assays were performed by mixing individual cell lysates or using co-expression cell lysates. For MsSpz, both cell lysate and cell culture medium were used because different modified forms of MsSpz were present in culture medium and cells (Fig. 7B). Our results showed that when cell lysates containing MsSpz-C108 and MsToll<sup>ecto</sup> were mixed, Flag antibody could pull down both MsSpz-C108-Flag and MsToll<sup>ecto</sup>-V5 (Fig. 9A and B, lane 4), whereas V5 antibody could precipitate both MsToll<sup>ecto</sup>-V5 and MsSpz-C108-Flag (Fig. 9C and D, lane 4). Similar results were obtained when cells co-expressing MsToll<sup>ecto</sup>-V5 and MsSpz-C108-Flag were used for the Co-IP assay (Fig. 9A-D, lane 5). When cell lysates and culture media containing MsSpz and MsToll<sup>ecto</sup> were mixed, or cells co-expressing MsToll<sup>ecto</sup> and MsSpz were used for the Co-IP assays, Flag antibody only pulled down MsSpz-Flag but not MsToll<sup>ecto</sup>-V5 (Fig. 9E and F, lanes 5-8), while V5 antibody could precipitate only MsToll<sup>ecto</sup>-V5 but not MsSpz-Flag (Fig. 9G and H, lanes 5-8). These results suggest that MsToll<sup>ecto</sup> can interact with MsSpz-C108 but not with full-length MsSpz. We also performed Co-IP assays of *DmToll<sup>ecto</sup>* with *DmSpz-C106* or *DmSpz*, and the results showed that *DmToll<sup>ecto</sup>* could interact with *DmSpz-C106* but not
with *DmSpz* (Fig. 10). Together, these results indicate that the Toll receptor only binds to active form Spz.
Figure 9. *M. sexta* Toll<sub>ecto</sub> (the ecto-domain) interacts with MsSpz-C108 but not with MsSpz.

The ecto-domain of *M. sexta* Toll (MsToll<sub>ecto</sub>), Spätzle-1A (MsSpz) and the active C-terminal domain of MsSpz (MsSpz-C108) were expressed or co-expressed in *Drosophila* S2 cells for 48 h. Proteins in both culture media and cell lysates were used for co-immunoprecipitation (Co-IP) assays as described in the Materials and Methods. MsToll<sub>ecto</sub> contained a V5-tag, while MsSpz and MsSpz-C108 contained a Flag-tag. Monoclonal antibody (anti-FLAG or anti-V5) was added to combined cell lysates, combined culture media with cell lysates, co-expression culture media or cell lysates, and immunoprecipitated (IP) proteins or Co-IP proteins were detected by immunoblotting using anti-Flag or anti-V5 antibody as the primary antibody, horseradish peroxidase-conjugate anti-mouse antibody as the secondary antibody and ECL chemiluminescence detection kit (A-D), or alkaline phosphatase-conjugate anti-mouse antibody as the secondary antibody and alkaline phosphatase (AP) conjugate color development kit (E-F). Lanes 1-3 from A-D and lanes 1-4 from E-F were cell lysates or culture media alone (protein inputs). In the figure, M is for culture media, C for cell lysates, Co-M for culture media from co-expression, and Co-C for cell lysates from co-expression.
Figure 10. *D. melanogaster* Toll\(^{\text{ecto}}\) (the ecto-domain) interacts with *DmSpz*-C106 but not with *DmSpz*.

The ecto-domain of *D. melanogaster* Toll (*DmToll\(^{\text{ecto}}\)*), Spätzle-1 (*DmSpz*) and the active C-terminal fragment of *DmSpz* (*DmSpz*-C106) were expressed or co-expressed in *Drosophila* S2 cells for 48 h. Proteins in both cell culture media and cell lysates were used for co-immunoprecipitation (Co-IP) assays as described in the Materials and Methods and in Fig. 9 legend. *DmToll\(^{\text{ecto}}\)* contained a V5-tag, while *DmSpz* and *DmSpz*-C106 contained a Flag-tag. Monoclonal antibody (anti-FLAG or anti-V5) was added to combined cell lysates, combined culture media with cell lysates, co-expression culture media or cell lysates, and immunoprecipitated (IP) proteins or Co-IP proteins were detected by immunoblotting using anti-Flag or anti-V5 antibody as the primary antibody, alkaline phosphatase-conjugate anti-mouse antibody as the secondary antibody and alkaline phosphatase (AP) conjugate color development kit. Lanes 1-4 from A-D and lanes 1-5 from E-F were cell lysates or culture media alone (protein inputs). In the figure, M is for culture media, C for cell lysates, Co-M for culture media from co-expression, and Co-C for cell lysates from co-expression.
Co-expression of MsToll with MsSpz-C108, but not MsSpz, in S2 cells induces transcriptional of the drosomycin gene

In D. melanogaster, the Toll-Spz pathway activates NF-κB factors Dorsal and Dif to induce expression of the drosomycin gene, while the Imd pathway activates NF-κB factor Relish to induce diptericin gene expression (Tanji and Ip, 2005). To investigate whether MsToll-MsSpz-C108 can activate expression of drosomycin or diptericin genes, Drosophila S2 cell lines expressing or co-expressing MsToll and MsSpz-C108 were selected for experiments because no M. sexta cell line is available. We first over-expressed the TIR domains of MsToll and DmToll to determine whether MsToll can activate antimicrobial peptide (AMP) genes in S2 cells. Dual luciferase activity assay showed that over-expression of DmTIR and MsTIR could activate drosomycin promoter significantly, but did not activate diptericin or attacin promoter (Fig. 11), indicating that MsToll can activate drosomycin gene in S2 cells.

We next over-expressed individual D. melanogaster and M. sexta Toll (Toll and Toll^ecto) and Spz (Spz and the active C-terminal domain) proteins or co-expressed different combination of Toll and Spz proteins in S2 cells, and then determined activation of drosomycin or diptericin gene by dual luciferase reporter assays. As a positive control, co-expression of DmToll-DmSpz-C106, but not DmToll-DmSpz, significantly increased relative luciferase activity of drosomycin reporter (~40-fold), but did not activate diptericin reporter (Fig. 12). Similarly, co-expression of MsToll-MsSpz-C108, but not MsToll-MsSpz, also activated drosomycin reporter to a significantly higher level (~25-fold) than the control, although the level activated by MsToll-MsSpz-C108 was slightly lower than that activated by DmToll-DmSpz-C106. Over-expression of DmSpz-C106
alone also activated *drosomycin* reporter to a significantly higher level than the control. However, over-expression of Toll, Toll\textsuperscript{ecto} and Spz proteins or co-expression of other combinations of Toll and Spz, particularly *Dm*Toll-\textit{Ms}Spz-C108 and *Ms*Toll-*Dm*Spz-C106, did not activate *drosomycin* reporter (Fig. 12), suggesting that only a correct pair of Toll-Spz (*Dm*Toll-*Dm*Spz-C106 or *Ms*Toll-*Ms*Spz-C108) can activate the Toll signaling pathway.

To confirm the dual luciferase results, we also performed real-time PCR on RNA isolated from S2 cells over-expressing or co-expressing Toll and Spz proteins (Fig. 13). Over-expression of *Dm*Spz-C106 alone significantly increased *drosomycin* mRNA level (≈4-fold) compared to the control. Co-expression of *Ms*Toll-*Ms*Spz-C108 and *Dm*Toll-*Dm*Spz-C106 increased *drosomycin* transcript levels to significantly higher levels (14- and 18-fold, respectively) than over-expression of *Dm*Spz-C106 alone (Fig. 13A). However, over-expression or co-expression of these proteins did not significantly change the *diptericin* mRNA levels compared to the control (Fig. 13B). These results are consistent with those of dual luciferase assays (Fig. 12), and further confirm that *Ms*Toll-*Ms*Spz-C108 can activate the transcription of *drosomycin* but not *diptericin* gene in S2 cells.
Figure 11. Over-expression of *M. sexta* Toll TIR domain (*MsTIR*) in S2 cells activates *drosomycin* promoter reporter.

*Drosophila* S2 cells were transiently co-transfected with a recombinant expression plasmid (expressing *Ms*Toll TIR or *Dm*Toll TIR) and a promoter reporter (pGL3B, pGL3B-*drosomycin*, pGL3B-*dipterican*, or pGL3B-*attacinA*). These cells were induced for protein expression for 48 h, and relative luciferase activity (RLA) in these cells was measured by dual-luciferase assay. Activity in cells co-transfected with empty pMT/BiP/V5-His A and pGL3B plasmids was set as 1 (the calibrator). Bars represent the mean of three individual measurements ± SEM. Identical letters are not significant difference (p>0.05), while different letters indicate significant difference among groups (p<0.05) determined by one way ANOVA followed by a Tukey’s multiple comparison test.
Figure 12. Co-expression of MsToll with MsSpz-C108 but not MsSpz activates *drosomycin* promoter reporter in S2 cells.

S2 cells were transiently co-transfected with recombinant expression plasmid(s) (expressing MsToll, MsToll\textsuperscript{ecto}, DmToll, DmToll\textsuperscript{ecto}, MsSpz, MsC108 (MsSpz-C108), DmSpz or DmC106 (DmSpz-C106), or co-expressing any combinations of Toll and Spz) and a promoter reporter (pGL3B, pGL3B-*drosomycin* or pGL3B-*diptericin*). These cells were induced for protein expression for 48 h, and relative luciferase activity (RLA) in these cells was measured by dual-luciferase assay. Activity in cells co-transfected with empty pMT/BiP/V5-His A and pGL3B plasmids was set as 1 (the calibrator). Bars represent the mean of three individual measurements ± SEM. Identical letters are not significant difference (p>0.05), while different letters indicate significant difference among groups (p<0.05) determined by one way ANOVA followed by a Tukey’s multiple comparison test.
Figure 13. Co-expression of MsToll with MsSpz-C108 but not MsSpz activates endogenous drosomycin gene in S2 cells.

S2 cells were transiently transfected with recombinant expression plasmid(s) (expressing MsToll, MsToll$^{ecto}$, DmToll, DmToll$^{ecto}$, MsSpz, MsC108 (MsSpz-C108), DmSpz or DmC106 (DmSpz-C106), or co-expressing any combinations of Toll and Spz), and then induced for protein expression for 48 h. Total RNAs were prepared from these cells, and expression of endogenous drosomycin (A) and diptericin (B) genes was determined by real-time PCR. Ribosomal protein 49 (rps49) gene was used as an internal control. Bars represent the mean of three individual measurements ± SEM. Identical letters are not significant difference (p>0.05), while different letters indicate significant difference among groups (p<0.05) determined by one way ANOVA followed by a Tukey’s multiple comparison test.
The Toll-Spz pathway regulates expression of AMP genes in *M. sexta* larvae

Several antimicrobial peptide (AMP) genes, including *cecropin*-6, *attacin*-1, *attacin*-2, *moricin*, *gloverin*, *lebocin*-b and *lebocin*-c, and *lysozyme* have been identified in *M. sexta* (Kanost et al., 2004; Rao et al., 2012). It has been reported that injection of *MsSpz*-C108 but not *MsSpz* into *M. sexta* larvae can activate some AMP genes (An et al., 2010). To further confirm regulation of AMP genes in *M. sexta* larvae, purified recombinant *MsSpz*, *MsSpz*-C108, or water (the control) was injected into day 1 fifth instar *M. sexta* naïve larvae, and expression of AMP genes in hemocytes and fat body was determined. Quantitative real-time PCR results showed that water injection did not activate AMP genes in hemocytes and fat body compared to the naïve larvae, and injection of *MsSpz* activates AMP genes to low levels compared to the naïve larvae or the water injection control, probably due to activation of some *MsSpz* by hemolymph proteinases (Fig. 14). However, injection of *MsSpz*-C108 activated all AMP genes (except *lysozyme*) in hemocytes and fat body to significantly higher levels than the control (water-injection) and naïve larvae (Fig. 14). These results suggest that *M. sexta* AMP genes can be regulated by the Toll-Spz pathway. *Lysozyme* was activated by *MsSpz* and *MsSpz*-C108 to a similarly low level, which was still significantly higher than that of the naïve larvae or the water-injection control. *Lysozyme* mRNA level was significantly lower than that of any other AMP genes.
Figure 14. Activation of AMP genes by MsSpz-C108 in M. sexta larvae.

Day 1 fifth instar M. sexta naïve larvae were injected with purified recombinant MsSpz (3 μg/larva), MsSpz-C108 (1 μg/larva), or water (control), or left untreated (naïve), hemocytes and fat body were then collected at 20 h post-injection for preparation of total RNAs. Expression of AMP genes, including cecropin-6, attacin-1, attacin-2, lebocin-b/c (primers were for the common regions of the two mRNAs), moricin and lysozyme, in hemocytes (A) and fat body (B) was determined by real-time PCR. Ribosomal protein S3 (rpS3) gene was used as an internal control. The bars represent the mean of three individual measurements ± SEM. Relative expression of each AMP gene in naïve larvae was set as 1. Identical letters are not significant difference (p>0.05), while different letters indicate significant difference (p<0.05) among different treatments for each AMP gene determined by one way ANOVA followed by a Tukey's multiple comparison test.
Activation of AMP genes in *M. sexta* larvae by MsSpz-C108 is blocked by antibody to *MsToll*

In *Drosophila*, Gram-positive Lys-type peptidoglycan (PG) activates the Toll pathway, while Gram-negative meso-diaminopimelic acid (DAP)-type PG activates the Imd pathway (Gottar et al., 2002; Leulier et al., 2003; Michel et al., 2001). In *M. sexta*, both Lys-type and DAP-type PGs can activate expression of *M. sexta* AMP genes (Rao and Yu, 2010). To further confirm a Toll-Spz pathway in *M. sexta* and to test whether activation of AMP genes by Lys-type and DAP-type peptidoglycans in *M. sexta* is regulated by the Toll and/or Imd pathways, an antibody blocking assay was performed since attempts to silence *MsToll* gene by RNAi failed. *M. sexta* larvae were first injected with purified IgG to the ecto-domain of *MsToll* or control IgG from pre-bleed serum, then injected with water, recombinant *MsSpz* or *MsSpz*-C108, *S. aureus* PG (PG-SA, Lys-type PG), *E. coli* PG (PG-K12, DAP-type PG), or without a second injection (control), and induced expression of AMP genes in hemocytes and fat body was determined.

Real-time PCR results showed that in the control IgG pre-injected larvae, injection of water did not activate AMP genes, and injection of *MsSpz* could activate AMP genes to low levels in both hemocytes and fat body (Fig. 15), probably due to activation of some *MsSpz* in the hemolymph. But injection of *MsSpz*-C108 and PG-SA activated all the AMP genes (except lysozyme) to significantly higher levels in hemocytes and fat body compared to the injection of water (Fig. 15), and injection of PG-K12 also activated some AMP genes (e.g. moricin in hemocytes and fat body, and lebocin-b/c in fat body) to significantly higher levels than the water injection (Fig. 15D, H and I).
Among MsSpz-C108, PG-SA and PG-K12, MsSpz-C108 activated most AMP genes to significantly higher levels than PG-SA and PG-K12 did, and PG-SA activated most AMP genes to significantly higher levels than PG-K12 did, but PG-K12 activated moricin in hemocytes (Fig. 15D) and lebocin-b/c in fat body (Fig. 15H) to significantly higher levels than PG-SA did.

In the MsToll IgG pre-injected larvae, activation of AMP genes (except lysozyme) by MsSpz-C108, PG-SA and PG-K12 in hemocytes and fat body was all except for lebocin-b/c in hemocyte (Fig. 15C) significantly suppressed compared to those of the control IgG pre-injected larvae (Fig. 15). These results suggest that MsSpz-C108, PG-SA and PG-K12 may all activate AMP genes via the Toll-Spz pathway, and binding of MsToll IgG to MsToll blocks MsSpz-C108 from binding to MsToll and thus suppresses activation of the downstream AMP genes. Although overall activation level of lysozyme by MsSpz-C108, PG-SA and PG-K12 was significantly lower than any of the other AMP genes, PG-K12 activated lysozyme to significantly higher levels than MsSpz-C108 and PG-SA did, and pre-injection of MsToll antibody did not block PG-K12 activated expression of lysozyme in hemocytes and fat body (Fig. 15E and J). Pre-injection of MsToll antibody also stimulated activation of lebocin-b/c in hemocytes (Fig. 15C), which was different from suppression of lebocin-b/c in fat body (Fig. 15H). These results suggest that regulation of lysozyme in fat body and hemocytes and lebocin-b/c in hemocytes may not be regulated by the Toll-Spz pathway.
Figure 15. Activation of AMP genes in *M. sexta* larvae by *MsSpz*-C108 is blocked by pre-injection of antibody to *MsToll*.

Day 1 fifth instar *M. sexta* naïve larvae were pre-injected with purified IgG to the ecto-domain of *MsToll* (Toll Ab, 5 μg/larva) or IgG from pre-bleed serum (control Ab, 5 μg/larva). These larvae were then injected with purified recombinant *MsSpz* (3 μg/larva), *MsSpz*-C108 (1 μg/larva), PG-SA (1 μg/larva), PG-K12 (1 μg/larva), or water, or without second injection (control) at 1 h after pre-injection of antibody. Hemocytes and fat body were collected at 20 h after second injection for preparation of total RNAs. Expression of AMP genes (*cecropin*-6, *attacin*-1, *lebocin*-b/c and *moricin*) in hemocytes (A-E) and fat body (F-J) was determined by real-time PCR. Ribosomal protein S3 (*rpS3*) gene was used as an internal control. The bars represent the mean of three individual measurements ± SEM. Relative expression of each AMP gene after pre-injection of antibody but without second injection (control) was set as 1. Asterisks indicate significant difference (p<0.05) between Toll and Control antibody pre-injections for each AMP gene determined by an unpaired t-test.
Discussion

*MsToll* can interact with *MsSpz*-C108, but not with full-length *MsSpz*

Invertebrates, such as insects, mainly rely on innate immunity to fight against pathogens. Induced expression of antimicrobial peptide (AMP) genes is an important defense mechanism in insect innate immunity (Ashida M, 1998; Hancock and Scott, 2000; Imler and Bulet, 2005; Lemaitre and Hoffmann, 2007; Williams, 2007). AMP gene expression is regulated by signal transduction pathways, such as the Toll and Imd pathways in *D. melanogaster* (Choe et al., 2002; De Gregorio et al., 2002; Lemaitre et al., 1995; Lemaitre et al., 1996; Ramet et al., 2002). The *Drosophila* Toll-Spz signaling pathway controls dorsal-ventral patterning in the embryonic development and also retains a common function in stimulating expression of AMP genes. Although *Toll* and *Spz* genes have been identified in different insect species (An et al., 2010; Ao et al., 2008b; Christophides et al., 2002; Evans et al., 2006; Imamura and Yamakawa, 2002; Kanzok et al., 2004; Luna et al., 2002; Wang et al., 2007), the role of Toll-Spz pathway in regulating AMP gene expression in other insect species has not been well studied. Progress in understanding the Toll-Spz pathway that operates in the innate immune system requires more investigation of molecular and biochemical functions of Toll and Spz in diverse taxa. In this study, we have identified a Toll-Spz signaling pathway in a lepidopteran insect, *M. sexta*.

In *M. sexta*, Toll and Spz have been identified (An et al., 2010; Ao et al., 2008b), but interaction between *M. sexta* Toll and Spz and direct evidence for a Toll-Spz pathway have not been demonstrated. In *D. melanogaster*, it has been shown that proteolytic processing of *DmSpz* releases the active *DmSpz*-C106, which is required for interaction
with $Dm$Toll (Arnot et al., 2010). The binding of two active $Dm$Spz-C106 dimers to one $Dm$Toll receptor and the formation of Toll-Spz heterodimers are also essential (Hu et al., 2004). We showed by co-immunoprecipitation (Co-IP) assay that $Ms$Toll$^{\text{ecto}}$ could interact with $Ms$Spz-C108, but not with full-length $Ms$Spz (Fig. 9), a result consistent with that of $Dm$Toll$^{\text{ecto}}$ and $Dm$Spz-C106 (Fig. 10), suggesting that proteolytic activation of $Ms$Spz is required for interaction of active $Ms$Spz-C108 with $Ms$Toll. We also showed that $Ms$Spz could be processed to active $Ms$Spz-C108 by proteinases in the hemolymph of $M$. sexta larvae (Fig. 8C and D), and these hemolymph proteinases are also induced by microorganisms. One interesting result is that $Dm$Toll$^{\text{ecto}}$ was differentially post-translationally modified in S2 cells (Fig. 7D) and different modified forms of $Dm$Toll$^{\text{ecto}}$ could all interact with $Dm$Spz-C106 (Fig. 10), while only one form of $Ms$Toll$^{\text{ecto}}$ was detected (Fig. 7C).

**Co-expression of $Ms$Toll-$Ms$Spz-C108 activate drosomycin expression in S2 cells**

To test whether $Ms$Toll-$Ms$Spz-C108 complex can activate AMP genes, both in vitro experiments in Drosophila S2 cells (since no $M$. sexta cell line is available) and in vivo experiments in $M$. sexta larvae were performed. Over-expression of $Ms$TIR and $Dm$TIR in S2 cells both could activate drosomycin (a target gene of the Toll pathway) but not diptericin (a target gene of the Imd pathway) (Fig. 11), suggesting that $Ms$Toll can activate the Toll pathway in S2 cells. Co-expression of $Ms$Toll-$Ms$Spz-C108 and $Dm$Toll-$Dm$Spz-C106 could activate drosomycin to similarly high levels (Figs. 12 and 13), but did not activate diptericin, further confirming that $Ms$Toll-$Ms$Spz-C108 complex can activate the Toll signaling pathway in S2 cells. It has been shown that Bombyx mori Spz
can activate AMP genes in *M. sexta* larvae (Wang et al., 2007), but we showed that co-expression of MsToll-DmSpz-C106 or DmToll-MsSpz-C108 did not activate *drosomycin* (Figs. 12 and 13). MsSpz is 44% and 23% identical to *B. mori* and *D. melanogaster* Spz-1, respectively; together these results suggest that Toll and Spz binding may be specific and only the correct pair of Toll-Spz can activate the Toll pathway. *D. melanogaster* contains nine Toll and six Spz genes; it would be interesting to know whether other Toll-Spz pairs can trigger signaling pathways in *D. melanogaster*.

**Injection of MsSpz-C108 activated AMP genes in *M. sexta* larvae**

In *D. melanogaster*, expression of AMP genes is regulated by the Toll and immune deficiency (Imd) pathways (De Gregorio et al., 2002; Lemaitre et al., 1995; Lemaitre et al., 1996). The Toll pathway is activated by fungi and Lys-type peptidoglycan (PG) of Gram-positive bacteria via peptidoglycan recognition protein (PGRP)-SA, PGRP-SD and Gram-negative binding protein 1 (GNBP1) (Bischoff et al., 2004; Gobert et al., 2003; Michel et al., 2001), while the Imd pathway is activated by *meso*-diaminopimelic acid (DAP)-type PG of Gram-negative bacteria and some *Bacilli* species via PGRP-LC (Choe et al., 2002; Gottar et al., 2002; Kaneko et al., 2004). In *M. sexta*, several AMP genes, including *cecropin, attacin, moricin, lebocin* and *gloverin*, as well as *lysozyme* have been identified (Kanost et al., 2004), and they can be activated by different bacterial cell wall components (Rao and Yu, 2010). Among these *M. sexta* AMP genes, *cecropin* and *attacin* are common AMP genes found in most insect species, but *moricin, lebocin* and *gloverin* are only found in *lepidopteran* species (Axen et al., 1997; Chowdhury et al., 1995; Hara and Yamakawa, 1995; Kanost et al., 2004). It is not clear
whether *M. sexta* AMP genes, particularly *lepidoptera*-specific *moricin*, *lebocin* and *gloverin* genes, are regulated by the Toll and/or Imd pathways. Injection of *MsSpz*-C108 activated AMP genes (*cecropin*, *attacin*, *lebocin* and *moricin*) in *M. sexta* larvae to significantly higher levels than the control (water-injection) (Fig. 14), suggesting that these *M. sexta* AMP genes are regulated by the Toll-Spz pathway. *MsSpz*-C108 activated *lysozyme* mRNA expression to much lower level compared to other AMP genes (Fig. 14), but *Lysozyme* protein is highly induced by *MsSpz*-C108 in hemolymph (An et al., 2010).

**MsToll antibody blocks the activity of *MsSpz*-C108**

In *Drosophila*, though Lys-type and DAP-type PGs can activate the Toll and Imd pathways, respectively, PGRP-SD can bind to DAP-type PG and may be responsible for activation of the Toll pathway by Gram-negative bacteria (Leone et al., 2008), and *Anopheles gambiae* PGRP-LC is responsible for activation of AMP genes stimulated by *S. aureus* but not by *E. coli* (Meister et al., 2009). In *M. sexta*, it is not clear whether Lys-type and DAP-type PGs can activate the Toll-Spz and/or Imd pathways. We tried to silence *MsToll* by RNA interference (RNAi) in order to further confirm the Toll-Spz pathway and activation of AMP genes by *S. aureus* and *E. coli* PGs in *M. sexta* larvae, but all our attempts using siRNA and dsRNA failed. We then used antibody to the ecto-domain of *MsToll* to block *MsToll* in *M. sexta* larvae from binding to the injected *MsSpz*-C108. Our antibody blocking assay showed that activation of AMP genes in both hemocytes and fat body of *M. sexta* larvae by *MsSpz*-C108 and *S. aureus* PG (PG-SA, Lys-type PG) was significantly inhibited when larvae were pre-injected with antibody to *MsToll* but not the control antibody (Fig. 15). These results further confirm that *MsToll-
MsSpz-C108 can form a complex in *M. sexta* larvae to mediate the Toll-Spz signaling pathway and regulate AMP genes expression. *E. coli* PG (PG-K12, DAP-type PG) activated expression of some AMP genes (*moricin* in hemocytes and fat body, and *lebocin*-b/c in fat body) was also suppressed by pre-injection of antibody to *MsToll*. These results suggest that both the Lys-type PG-SA and DAP-type PG-K12 can activate the Toll-Spz pathway in *M. sexta*, but PG-K12 is a weaker elicitor than PG-SA in stimulation of the Toll-Spz pathway. Expression of *lebocin*-b/c in hemocytes was stimulated after *MsToll* Toll was blocked by antibody, suggesting that *lebocin*-b/c expression in hemocytes is not regulated by the Toll-Spz pathway. It is not clear why expression of *lebocin*-b/c in hemocytes and fat body is regulated differently. Expression of *gloverin* in hemocytes and fat body was also regulated in a similar pattern like *lebocin*-b/c (Xu et al., 2012). This may be related to expression pattern of *M. sexta* Spz, as it is expressed and induced in hemocytes but not induced in fat body (An et al., 2010). Activation of *lysozyme* by MsSpz-C108, PG-SA and PG-K12 was always lower than that of other *M. sexta* AMP genes, and the activation was not blocked by pre-injection of Toll antibody. In addition, PG-K12 is a stronger elicitor than MsSpz-C108 or PG-SA in activation of *lysozyme*. Thus, *lysozyme* is also not regulated by the Toll-Spz pathway. Expression of *lebocin*-b/c in hemocytes and *lysozyme* in both hemocytes and fat body may be regulated by other signaling pathways, such as the Imd pathway since Rel genes similar to *Drosophila* Relish have been identified in *M. sexta*.

In summary, we used a biochemical assay to show that *MsToll*\textsuperscript{ecto} and *DmToll*\textsuperscript{ecto} could interact with MsSpz-C108 and DmSpz-C106, respectively, but not with full-length Spz (Figs. 9 and 10), used *in vitro* assays to show that *MsToll-MsSpz-C108* and *DmToll-
DmSpz-C106 complexes could activate *drosomycin* but not *diptericin* gene in S2 cells (Figs. 12 and 13), used *in vivo* assays to show that activation of *M. sexta* AMP genes by *MsSpz*-C108 was significantly inhibited by pre-injection of antibody to *MsToll* (Fig. 15). Our results together demonstrated a Toll-Spz signaling pathway in a *lepidopteran* insect, *M. sexta*. This study may help better understand signaling pathways in *lepidopteran* insects, and the origin and evolution of animal innate immune signaling pathways.
CHAPTER 3
A TOLL-ML-LPS PATHWAY IN THE TOBACCO HORNWORM, MANDUCA SEXTA

Abstract

In mammals, Toll-like receptor 4 (TLR4) is the receptor that recognizes lipopolysaccharide (LPS). MD2 (Myeloid differentiation protein 2) is required for the recognition of LPS and binding to TLR4. In insects, there are two signaling pathways in the innate immunity system: the Imd (immune deficiency) and Toll pathways. However, receptors that can recognize LPS have not been identified in invertebrates. Our aims are to investigate whether insects have a Toll-ML-LPS pathway. A Toll receptor (MsToll) and an MD2-related lipid-recognition protein (MsML-1) have been identified in the tobacco hornworm, Manduca sexta. We constructed S2 cell lines expressing MsToll, a Chimeric Toll (ChiToll with the ecto-domain of MsToll, transmembrane and TIR domains of Drosophila Toll) and MsML-1. We showed that LPS could activate drosomycin gene expression in S2 cells only when both Toll and MsML-1 or ChiToll and MsML-1 were co-expressed, indicating that a Toll-ML-LPS pathway is present in insects.

Introduction

In mammals, Toll-like receptors (TLRs) can recognize a variety of pathogen-associated molecular patterns (PAMPs) to activate the NF-κB pathway. LPS released from Gram-negative bacteria is immediately captured by the LPS-binding protein (LBP), a specific lipid transfer protein, and binds to CD14 present on the surfaces of mononuclear phagocytes. CD14 is also a leucine-rich repeat protein and lacks a
transmembrane domain, and is, thus, incapable of transducing signals (An et al., 2010). Then LPS will be delivered to TLR4, the receptor that recognizes LPS. MD2 is required for the recognition of LPS and binding to TLR4. TLR4/MD2/LPS complex induces dimerization of TLR4 and subsequent activation of cytoplasmic signaling molecules (Poltorak et al., 1998). TLR4-mediated signaling through IκB and NF-κB leads to the robust induction of the genes involved in proinflammatory responses, which connects between innate and adaptive immunity. TLR4 resembles Toll receptor of insects, for example *Drosophila melanogaster*. Both receptors look alike in their extracellular domains as well as intracellular domains. Insects have multiple Toll and ML proteins resulting from gene duplication. However, there is no direct connection between Toll and ML. The *Drosophila* Toll-Spätzle pathway is similar to mammalian TLR-mediated signaling pathways, but also differs from TLR pathways in the recognition process (Valanne et al., 2011). LPS signaling pathway or receptors for LPS has not been identified in insects and other invertebrates. Thus, it is likely that there might be a Toll-ML-LPS signaling pathway in insects. A *M. sexta* Toll (*MsToll*) (Ao et al., 2008b) and a MD2-like protein (*MsML-1*) (Ao et al., 2008a) have been identified previously in our laboratory. Recombinant *MsML-1* protein specifically binds to LPS from several Gram-negative bacteria (Ao et al., 2008a). The TIR domain of *MsToll* also has high similarity to that of vertebrate TLR4. These results suggested that there might be a Toll-ML-LPS pathway in *M. sexta*. 
Materials and Method

*D. melanogaster* Schneider S2 cell line

*D. melanogaster* Schneider S2 cells were purchased from American Type Culture Collection (ATCC).

Construction of recombinant pMT/BiP/V5-His A expression vectors

cDNA fragments encoding *MsToll*, *MsToll*\(^{ecto}\), *ChiToll*, and *MsML-1* were amplified by PCR using forward and reverse primers. Forward primer for *MsML-1* contains codons for an in-frame Flag sequence and a *NotI* site, while reverse primer contains a stop codon followed by a *PmeI* site. Forward primers for *MsToll*, *MsToll*\(^{ecto}\), and *ChiToll* contain a *KpnI* site, while reverse primers contain an *ApaI* site. 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 45 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 digested with *NotI/PmeI* or *KpnI/ApaI*, DNA fragments were recovered and inserted into *NotI/PmeI* or *KpnI/ApaI* digested pMT/BiP/V5-His A vector (V413020, Invitrogen) using T4 DNA ligase (M0202L, New England Biolabs). Recombinant expression vectors 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 generate stable S2 cell lines.
**Cell culture and establishment of stable S2 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). For DNA transfection, cells were seeded overnight in serum-free medium (SH30278.01, HyClone). GenCarrier-1™ transfection reagent (#31-00110, Epoch Biolabs) was used for transient transfection based on the manufacturer’s instructions. Cells in culture dishes or plates were grown to 70% confluence prior to transfection. DES®–Inducible/Secreted Kit with pCoBlast (K5130-01, Invitrogen) was used to construct stable S2 cell lines.

**Western blot analysis and immunoprecipitation (Co-IP) assay**

For Western blot analysis, copper sulfate (final concentration of 250 μM) was added to the stable S2 cell lines (2×10^6 cells/well) in 6-well plates, and protein expression was induced for 48 h. Cell culture medium (2 ml each) was collected, 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, 0.5 mM PMSF, protease inhibitor cocktail (P8340, Sigma-Aldrich)]. After removing the cell debits, the supernatants were collected as cell extracts for Western blot analysis. The 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, and alkaline phosphatase-conjugate anti-mouse antibody (A4312, Sigma-Aldrich, 1:10,000) was used as secondary antibody for color development. The signal was developed by using the ECL Chemiluminescence Detection Kit (RPN2134, GE Healthcare) or alkaline phosphatase (AP)-conjugate color development Kit (#170-6432, Bio-Rad).

Immunoprecipitation (Co-IP) assay was performed by using 300 µl of cell extract, which is equivalent to approximately 10⁶ cells, or equivalent cell culture medium containing recombinant proteins. The cell extracts or cell culture media were pre-cleared for 30 min with 30 µl Protein G Sepharose (50% slurry, No.17-0618-01, GE Healthcare) in a total volume of 500 µl. After centrifugation, the supernatant was incubated with anti-Flag M2 or anti-V5 antibody (final concentration of 10 µg/ml) at 4°C for 10 h with gentle rocking. LPS or PG (peptidoglycan) (10 µg/ml final concentration) was also included in the cell lysates to test whether it can promote co-precipitation of MsToll<sup>ecto</sup> with MsML-1. Then, 30 µl 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, re-suspended in 50 µl of 1×SDS sample buffer, boiled at 95°C for 5 min, and used for subsequent immunoblotting analysis.

**Dual-Luciferase Reporter Assay**

Stable *Drosophila* S2 cell lines expressing *MsML-1, MsToll<sup>ecto</sup>* and *ChiToll* were established. *Drosomycin* (a readout antimicrobial peptide gene in the *Drosophila* Toll
pathway) promoter-luciferase reporter was constructed. Cell lines were transfected with the reporter alone or along with MsML-1 (for co-expression of MsToll\textsuperscript{ecto} or ChiToll with MsML-1). After protein expression for 48 h, cells were stimulated with LPS (10 μg/ml final concentration) for 7 h, and luciferase activity in cell lysates was determined using a dual luciferase assay kit. All results were expressed as the ratio of relative luciferase activity with and without LPS stimulation.

**Data analysis**

One representative set of data was used to make figures using the Graphpad Prism software, and the significance of difference was determined by an unpaired t-test or by one way ANOVA followed by a Tukey’s multiple comparison test with the Graphpad InStat software (GraphPad, San Diego, CA).

**Results**

**Expression of recombinant* M. sexta* Toll, ML, and Chimeric Toll proteins in S2 cells**

The TLR4-MD2-LPS signaling pathway is well understood in mammals; however the receptor recognizing LPS is not characterized in invertebrates. In *M. sexta*, *Toll* and *ML* genes have been identified (Ao et al., 2008a; Ao et al., 2008b). In order to investigate a Toll-ML-LPS pathway in *M. sexta*, we constructed a recombinant *ChiToll* receptor, with the ecto-domain of *MsToll*, transmembrane and TIR domains of *Drosophila* Toll 1 (Fig. 16). We established stable S2 cell lines expressing *MsToll* receptor and its ecto-domain *MsToll\textsuperscript{ecto}*, as well as *ChiToll* receptor and *MsML-1*. Immunoblotting results
showed that $MsToll^{ecto}$ (the ecto-domain) and $ChiToll$ (Fig. 17C and D) were detected only in S2 cells but not in cell culture media. For the $MsML-1$, a single protein band was detected in the cells and the cell culture media (Fig. 17A). But $MsML-1$, a typically secreted protein, could not be secreted into the medium when co-expressed with $MsToll^{ecto}$. Furthermore, the size of $MsML-1$ was smaller, suggesting differences in post-translational modification of $MsML-1$ when co-expressed with Toll receptor (Fig. 17A). Comparing to the native $MsML-1$, the size of recombinant $MsML-1$ co-expressed with $MsToll$ is similar as native protein (Fig. 17B). These results imply that there may be an interaction between $MsToll$ and $MsML-1$. 
Figure 16. Schematic diagram of the ChiToll.

The ecto-domain (LRRs) of *M. sexta* Toll is fused to the transmembrane region and intracellular TIR domain of *Drosophila* Toll-1.
Figure 17. Expression of MsML-1 (A), MsToll\textsuperscript{ecto} (C), and ChiToll (D) in S2 cells.

Proteins in cell culture media and cell lysates were detected by Western blotting using monoclonal antibody against FLAG-tag (MsML-1) or V5-tag (Toll), or polyclonal antibody against \textit{M. sexta} ML-1 (B). Panels A, C and D: lanes 1 (culture medium) and 2 (cell lysate) from individual expression; lanes 3 (culture medium) and 4 (cell lysate) from co-expression. Panel B: lane 1, native MsML-1 purified from \textit{M. sexta} hemolymph; lane 2, co-expressed MsML-1; lane 3: single-expressed MsML-1.
LPS enhances the interaction between MsToll\textsuperscript{ecto} and MsML-1

In mammals, membrane-bound CD14 accepts LPS captured by LBP and transfers LPS to MD2, which is a co-receptor for TLR4 (Shimazu et al., 1999). MD2 is required for the recognition of LPS and binding to TLR4. The binding of MD2 to LPS triggers a conformation change, which induces the formation of LPS-MD2-TLR4 complex (Schromm et al., 2001). The multimer triggers TLR4 dimerization to initiate the intracellular signaling (Poltorak et al., 1998). To determine whether MsML-1 can bind to MsToll receptor, we over-expressed the ecto-domain of MsToll (MsToll\textsuperscript{ecto}) with a V5-His-tag to the C-terminus, and MsML-1 with a Flag-tag to the N-terminus in S2 cells (Figure 17). Co-immunoprecipitation (Co-IP) assays were performed by mixing individual cell lysates or using co-expression cell lysates. For MsML-1, single transfection and co-expressing with MsToll\textsuperscript{ecto} were separated because the size of MsML-1 was different when co-expressed with Toll receptor. Our results showed that when cell lysates containing MsML-1 and MsToll\textsuperscript{ecto} were mixed, Flag antibody could pull down both MsML-1-Flag and MsToll\textsuperscript{ecto}-V5 (Fig.18 and B, lane 4-6), whereas V5 antibody could precipitate both MsToll\textsuperscript{ecto}-V5 and MsML-1-Flag (Fig. 18C and D, lane 4-6). Similar results were obtained when cells co-expressing MsToll\textsuperscript{ecto}-V5 and MsML-1-Flag were used for the Co-IP assay and showed that difference in the post-translational modification did not have an effect on the interaction between MsToll and MsML-1 (Fig. 18A-D, lane 5-8). These results suggest that MsToll\textsuperscript{ecto} can interact with MsML-1. In order to detect the role of LPS and PG in the interaction, LPS or PG (10 μg/ml final concentration) was included in the cell lysates. The results showed that LPS, but not PG,
could enhance the interaction between $M_s$Toll$^{\text{ecto}}$ and $M_s$ML-1 (Fig. 18B and D, lane 5-6 and 8-9).
Figure 18. Co-IP of MsToll\textsuperscript{ecto} with MsML-1.

Stable Drosophila S2 cell lines expressing the MsToll\textsuperscript{ecto} and MsML-1 were established, and cell lysates were prepared individually. MsML-1 was also co-expressed with MsToll\textsuperscript{ecto} using stable cell line expressing MsToll\textsuperscript{ecto} for the transfection, and cell lysates were also prepared. Cell lysates containing MsML-1 (FLAG-tagged) or MsToll\textsuperscript{ecto} (V5-tagged) and co-expressed cell lysates were used for Co-IP assays using monoclonal antibody (mAb) to FLAG- or V5-tag. LPS or PG (peptidoglycan) was also included in the cell lysates to test whether it can promote co-precipitation of MsToll\textsuperscript{ecto} with MsML-1. Lanes 1-3 contain only cell lysates without mAb. * indicates that LPS, but not PG, can enhance the interaction between MsToll\textsuperscript{ecto} and MsML-1.
Co-expression of MsToll or ChiToll with MsML-1 in S2 cells can activate *drosomycin* promoter stimulated by LPS

In mammals, the TLR4-MD2-LPS pathway mediates the activity of several NF-κB transcription factors, leading to robust induction of proinflammatory cytokines, interferon (IFN), or IFN-inducible genes. In *D. melanogaster*, the Toll-Spz pathway activates NF-κB factors Dorsal and Dif to induce expression of *drosomycin* gene, while the Imd pathway activates NF-κB factor Relish to induce *diptericin* gene expression (Tanji and Ip, 2005). To investigate whether co-expression of Toll with MsML-1 can activate expression of *drosomycin* or *diptericin* gene stimulated by LPS, *Drosophila* S2 cell line was applied because no *M. sexta* cell line is available. A recombinant ChiToll (the ecto-domain of MsToll combined with the transmembrane and TIR domains of *Drosophila* Toll 1) is designed for *Drosophila* S2 cell. Cell lines were transfected with the reporter alone or along with MsML-1 (for co-expression with MsToll<sup>ecto</sup>, MsToll or ChiToll). After protein expression for 48 h, cells were stimulated with LPS (10 μg/ml final concentration) for 7 h, and activation of *drosomycin* or *diptericin* gene promoter activity was determined by dual luciferase reporter assays. The results showed that LPS could significantly increase the relative luciferase activity of *drosomycin* reporter (∼4-fold), but not *diptericin* reporter, in S2 cells co-expressing ChiToll-MsML-1 (Fig. 19). Similarly, LPS also stimulated *drosomycin* reporter activity to a significantly higher level (∼2.5-fold) in cells co-expressing MsToll-MsML-1 compared to the control cells (Fig. 19). These results suggest that there might be a Toll-ML-LPS pathway in *M. sexta.*
Figure 19. LPS activates *drosomycin* luciferase reporter in S2 cells co-expressing *Chi*Toll or *Ms*Toll with *MsML-1*.

Stable *Drosophila* S2 cell lines expressing *MsML-1*, *MsToll<sup>ecto</sup>, *MsToll* and *Chi*Toll were established. *Drosomycin* and *Diptericin* promoter-luciferase reporters were constructed. Cell lines were transfected with the reporter alone or along with *MsML-1* (for co-expression with *MsToll<sup>ecto</sup>, *MsToll*, or *Chi*Toll). After protein expression for 48 h, cells were stimulated with LPS (10 μg/ml final concentration) for 7 h, and luciferase activity in cell lysates was determined using a dual luciferase assay kit. All results were expressed as the ratio of relative luciferase activity with and without LPS stimulation. Bars represent the mean of three individual measurements ± SEM. Identical letters are not significant difference (p>0.05), while different letters indicate significant difference among groups (p<0.05) determined by one way ANOVA followed by a Tukey’s multiple comparison test.
Summary

In mammals, Toll-like receptor 4 (TLR4) is the receptor that recognizes LPS. MD2 is required for the recognition and binding between TLR4 and LPS. However, in invertebrates, receptors that can recognize LPS and trigger signaling pathways have not been identified. We used co-immunoprecipitation experiments to show that MsToll ectodomain could directly interact with MsML-1, and LPS could enhance the interaction. In dual-luciferase reporter assay, we also confirmed that LPS could stimulate drosomycin promoter in S2 cells co-expressing MsToll and MsML-1 or ChiToll and MsML-1. Our results for the first time showed a conserved Toll-ML-LPS pathway from insects to mammals, and provided better insight into the evolution of the TLR4-MD2-LPS pathway.
CHAPTER 4
DORSAL-REL2 HETERODIMERS NEGATIVELY REGULATE ANTIMICROBIAL PEPTIDE EXPRESSION IN MANDUCA SEXTA

Abstract

NF-κB transcription factors play essential roles in regulation of innate immune responses. Here we report the identification and functional analysis of a Dorsal homologue (MsDorsal) and two short Relish isoforms (MsRel2A and MsRel2B) from the tobacco hornworm, Manduca sexta. The Rel homolog domain (RHD) of MsDorsal is similar to that of many Dorsal proteins in Arthropods. MsRel2A and MsRel2B have an identical RHD region and differ only in a few residues at the C-terminus. RT-PCR analysis showed that MsDorsal and MsRel2A transcripts were detected in all tissues tested, including fat body, hemocytes, midgut and testis; however, MsRel2B mRNA was detected only in the midgut and testis. Overexpression of the RHD domains from MsDorsal and MsRel2 in S2 cells could activate antimicrobial peptide (AMP) gene promoters from M. sexta, D. melanogaster and B. mori. MsDorsal-RHD specifically bound to an NF-κB site of M. sexta lysozyme promoter, while MsRel2-RHD bound to an NF-κB site of M. sexta moricin promoter. Interestingly, co-immunoprecipitation (Co-IP) analysis showed that MsDorsal-RHD could interact with MsRel2-RHD. More importantly, co-expression of MsDorsal-RHD with MsRel2-RHD abolished the activation of several M. sexta antimicrobial peptide gene promoters. These results suggest that the short MsRel2 isoforms may form heterodimers with MsDorsal to negatively regulate expression of antimicrobial peptides in M. sexta.
**Introduction**

Insects rely solely on the innate immune system to control or eliminate invading pathogens (Hoffmann, 2003; Lemaitre and Hoffmann, 2007). Pathogen-associated molecular patterns (PAMPs) from various pathogens are recognized by germ-line encoded pattern recognition receptors (PRRs) to trigger different cellular and humoral immune responses (Kumar et al., 2009; Mogensen, 2009). Synthesis of antimicrobial peptides (AMPs) is one of the major defense mechanisms in insects (Imler and Bulet, 2005; Yi et al., 2014; Yount et al., 2006) and expression of AMPs is regulated by the Toll and Imd (immune deficiency) pathways (De Gregorio et al., 2002; Tanji et al., 2007). The Toll pathway mediates the recognition of gram-positive bacteria and fungi (Valanne et al., 2011), while the Imd pathway is activated by most gram-negative bacteria (Georgel et al., 2001).

The nuclear factor κB (NF-κB) family of transcription factors plays essential roles in regulating the expression of immune-related genes (Ganesan et al., 2011; Li and Verma, 2002). These factors contain an N-terminal Rel homology domain (RHD) that interacts with DNA. In mammals, there are two classes of NF-κB factors. Class I factors include p105 and p100, which contain an N-terminal RHD and a C-terminal long inhibitory ankyrin repeats that must be cleaved off to release the N-terminal RHD for activation of gene expression. Class II factors include RelA (p65), RelB and c-Rel that contain an N-terminal RHD and a C-terminal transactivation domain (Silverman and Maniatis, 2001). NF-κB factors can form homo- and hetero-dimers in the nucleus, which bind to NF-κB elements in the promoter regions of many immune-related genes (Ganesan et al., 2011). In *Drosophila melanogaster*, three NF-κB factors, Dorsal, Dorsal-related
immunity factor (Dif) and Relish, have been identified (Hetru and Hoffmann, 2009). Dorsal and Dif belong to the Class II NF-κB factors and they are involved in the Toll pathway to regulate dorsal-ventral patterning in embryonic development (Anderson et al., 1985) or expression of AMPs such as drosomycin in Drosophila larvae and adults (Tian et al., 2008; Uttenweiler-Joseph et al., 1998; Whalen and Steward, 1993). Relish is a member of the Class I NF-κB factors and is cleaved to release the N-terminal fragment (Relish-N) containing RHD upon activation of the Imd pathway (Silverman et al., 2000; Stoven et al., 2000). Relish also regulates expression of AMPs including diptericin (Wicker et al., 1990). It has been suggested that Drosophila Dif and Relish may form heterodimers in in vivo to synergistically increase AMP production (Tanji et al., 2007; Tanji et al., 2010).

* M. sexta is a popular lepidopteran model organism to study insect immunity (Jiang et al., 2010). Various immune-related genes have been identified and characterized (Kanost et al., 2004), including AMP genes moricin, lebocin and gloverin (Dai et al., 2008; Rao et al., 2012; Xu et al., 2012); C-type lectins (immulectins) (Yu et al., 1999; Yu and Kanost, 2000; Yu et al., 2006; Yu et al., 2005); beta-1,3-glucan recognition proteins and so on (Jiang et al., 2004; Ma and Kanost, 2000; Rao et al., 2014; Wang et al., 2006). A Toll-Spätzle pathway has been confirmed in *M. sexta* and it plays an important role in immune responses (Zhong et al., 2012). Previously, we reported that *M. sexta moricin* is activated by both NF-κB and GATA factors (Rao et al., 2011).

NF-κB factors have been identified in the phylum of Arthropoda (Antonova et al., 2009; Fan et al., 2008; Huang et al., 2010; Huang et al., 2009; Li et al., 2010; Li et al., 2009; Montagnani et al., 2004; Shin et al., 2002; Shin et al., 2005; Tanaka et al., 2007;
Tanaka et al., 2005; Wang et al., 2013; Yu et al., 2013). In the mosquito *Aedes aegypti*,
two Dorsal isoforms, *Aa*REL1-A and *Aa*REL1-B, can cooperatively enhance activation
of immune genes (Shin et al., 2005), and three alternatively spliced isoforms of Relish are
generated from a single inducible *relish* gene. The 3.9 kb *relish* transcript contains both
the RHD domain and the inhibitor κB (IκB)-like domain, while the other two short
transcripts lack either the RHD or the IκB domain. The *Anopheles gambiae* REL2 gene
produces two spliced forms: a full-length (REL2-F) and a shorter (REL2-S) (Meister et
al., 2005). In the silkworm *Bombix mori*, two Rel (Dorsal) proteins, *Bm*RelA and
*Bm*RelB, activate antimicrobial peptide genes differentially (Tanaka et al., 2005) , and
two Relish homologs (*Bm*Relish1 and 2) have been identified (Tanaka et al., 2007).
*Bm*Relish2 lacks the acidic and hydrophobic amino acids (AHAA) rich regions and
ankyrin (ANK) repeats, and it serves as a dominant negative factor of *Bm*Relish1 active
form (Tanaka et al., 2007). In the pacific white shrimp *Litopenaeus vannamei*, *Lv*Relish
and its short isoform (*sLv*Relish) have been identified (Huang et al., 2009). However, no
NF-κB transcription factors and their functions have been reported in *M. sexta* so far.

Here we report cloning and functional studies of three NF-κB homologs from *M.
sexta*. Interestingly, we cloned two short isoforms of *M. sexta* Relish, named *Ms*Rel2A
and *Ms*Rel2B, which differed in tissue-specific expression patterns but both could
activate AMP gene promoters. More importantly, we confirmed formation of *Ms*Dorsal-
*Ms*Rel2 heterodimers, and showed for the first time that *Ms*Dorsal-*Ms*Rel2 heterodimers
could suppress expression of AMP gene promoters. Our results suggest that relish active
isoforms such as *Ms*Rel2 may activate AMP genes as homodimers, and may also
negatively regulate AMP gene expression as heterodimers with dorsal to prevent over-activation of AMPs.

**Materials and Method**

Insect rearing and *Drosophila S2* and *Spodoptera frugiperda* Sf9 cell lines

*M. sexta* eggs were originally purchased from Carolina Biological Supplies (Burlington, NC, USA). Larvae were reared on an artificial diet at 25°C (Dunn, 1983), and the fifth instar larvae were used for the experiments. *D. melanogaster* Schneider S2 cells were purchased from American Type Culture Collection (ATCC). *Spodoptera frugiperda* Sf9 cells were purchased from Invitrogen.

Cloning and sequence analysis of *M. sexta Dorsal* and *Rel2* cDNAs

After analyzing the *M. sexta* EST library (http://entoplp.okstate.edu/profiles/jiang.htm), two EST fragments were predicted to encode RHD-containing proteins (manduca.Contig2427 and manduca.Contig7025). We designed primers based on the EST sequence to clone the full-length cDNA. Briefly, total RNA was prepared from the fat body of day 3 naïve larvae using TRIZol® Reagent (T9424, Sigma–Aldrich). For reverse transcription, total RNA was treated with RQ1 RNase-free DNase I (Promega) at 37°C for 30 min to remove contaminated genomic DNA, and DNase was inactivated by heating to 65°C for 20 min. Reverse transcription was performed using oligo(dT) primer (Promega) and ImProm-II reverse transcriptase (Promega) following the manufacturer’s instructions. *M. sexta Rel2* and *Dorsal* cDNA
were cloned using forward and reverse primers listed in Table 7. 5’and 3’ RACE reactions were performed using smarter race kit (Clontech). The opening reading frame (ORF) was predicted from the nucleotide sequence using DNAMAN (Lynnon Corporation, Quebec, Canada). BLASTP (http://blast.ncbi.nlm.nih.gov/Blast.cgi) was used to search homologous RHD sequences. Figures were made from means of three independent biological replicates with the GraphPad Prism software (GraphPad, CA). Significance of difference was determined by one way ANOVA followed by a Tukey’s multiple comparison test using GraphPad Prism.

Construction of luciferase reporter plasmids

To construct different promoter mutants, site-directed mutagenesis was performed using the wild-type *M. sexta lysozyme* promoter (1203 bp) and *morincin* deletion promoter (240 bp) as templates (Rao et al., 2011). Primers with specific mutation sites were designed for each mutant and listed in Table 7. PCR program was 3 min at 95°C, 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 at 68°C for 30 min. The PCR products were recovered by agarose gel electrophoresis-Wizard® SV Gel and PCR Clean-Up System (A9285, Promega) and digested by *DpnI* and then transformed into competent *Escherichia 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 or Sf9 cell lines.
Tissue distribution and induced expression of *M. sexta* Rel2A, Rel2B, and Dorsal

To determine tissue distribution of *M. sexta* two Relish isoforms and Dorsal, 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). cDNAs were prepared as described previously (Zhong et al., 2012). Briefly, 1 μg total RNAs extracted with TRIzol® Reagent (T9424, Sigma-Aldrich) and digested by RQ1 RNase-free DNase I (Promega) to remove contaminated DNA were used for cDNA preparation. cDNA was prepared in a 25 μl reaction using moloney murine leukemia virus (M-MLV) reverse transcriptase (M1701, Promega) with an anchor-oligo(dT)₁₈ primer following the manufacturer’s instructions. To determine induced expression of *MsRel2A*, *MsRel2B*, and *MsDorsal*, day 2 fifth instar larvae were injected with distilled deionized (ddH₂O), heat-killed *E. coli* strain XL1-blue (5×10⁷ cells/larva), *S. aureus* (5×10⁷ cells/larva), or *C. neoformans* (10⁷ cells/larva). Twenty-four hours after injection, hemocytes, fat body and midgut were collected separately. Total RNA and cDNA were prepared as described above. The RT-PCR was performed in 20 μl reactions containing 2.2 μl Advantage® cDNA Polymerase Mix (No. 639105, Clontech), 15.2 μl dH₂O, 1 μl cDNA, 0.4 μl ultra-pure dNTP and 0.6 μl each reverse and forward diluted primer (10 pmol/μl). The ribosomal protein S3 (*rpS3*) gene was used as internal control, and the forward and reverse primers for *MsRel2A* and *MsRel2B* with 508 Base Pair (bp), *MsRel2B* with 468 bp, *MsDorsal* with 553bp, and *rpS3* with 317 bp genes listed in Table 7. RT-PCR program was 2 min at 95°C, followed by 28 cycles of 95°C for 30 s, 55°C for 30 s, 68°C
for 30 s, followed by a final extension at 68°C for 10 min. The products from three replicates of each sample were analyzed with 1% agarose gel electrophoresis and these experiments were repeated with 3 different biological samples.

**Expression and purification of *M. sexta* Dorsal in bacteria and preparation of polyclonal rabbit antiserum**

Recombinant *Ms*Dorsal was expressed in *E. coli* and purified for production of polyclonal antibody in rabbit. RT-PCR was performed to obtain cDNA sequence encoding *M. sexta* Dorsal-RHD domain (residues 92-263), using forward and reverse primers listed in Table 7. PCR fragment was purified by agarose gel electrophoresis, digested with *Nco*I and *Xho*I enzymes, ligated into the *Nco*I/*Xho*I digested expression vector pGEX-5X, and then transformed into competent *E. coli* BL21 (DE3) cells. Recombinant plasmids were prepared from positive clones and confirmed by restriction enzyme digestion and DNA sequencing. A single bacterial colony was inoculated into LB medium containing ampicillin (100 μg/ml) and grown at 37°C overnight. The overnight culture was diluted 1:100 in LB medium and incubated at 37°C to OD<sub>600</sub>=0.8 and then isopropyl-D-thiogalactoside (IPTG) was added (at 0.5 mM final concentration) to induce protein expression. After overnight incubation at 16°C, bacterial cells were harvested by centrifugation and lyzed with the lysis solution (50 mM Tris-HCl, pH 8.0, 50 mM NaCl, 0.5% Triton X-100), and recombinant protein was purified using Ni-NTA agarose beads (Qiagen) under native conditions following the manufacturer’s instructions. The purified recombinant GST-*Ms*Dorsal fusion protein was cleaved by thrombin, and the cleavage products were separated on 12% SDS-PAGE. The gel slice containing recombinant
MsDorsal was used as an antigen to produce rabbit polyclonal antiserum at Cocalico Biologicals, Inc (Pennsylvania, USA).

**Construction of recombinant pAC5.1/V5-His A and pIZ/V5-His expression vectors**

cDNA fragments encoding MsRel2A (residues 1-397), MsRel2B (residues 1-422), MsRel2-RHD (residues 58-227 of MsRel2B or Rel2A), and MsDorsal-RHD (residues 92-263), were amplified by PCR using forward and reverse primers listed in Table 7. For pAC5.1/V5-His A vector, forward primers of MsRel2A, Rel2B, and Dorsal contain 5’ non-coding region recognized by Drosophila ribosome, followed by a start codon and a KpnI site, while reverse primers contain an ApaI site. Forward primer of Rel RHD contains a start codon and an EcoRI site and reverse primer contains NotI site followed by an in-frame Flag sequence and a stop codon. For pIZ/V5-His vector, forward primers contain 5’ non-coding region suitable for Spodoptera frugiperda 9 (SF9) cell expression, followed by a start codon and a KpnI site, while reverse primer of Dorsal RHD contains an XbaI site and primer for Rel RHD contains an XbaI site followed by an in-frame Flag sequence and a stop codon. The primers listed in Table 7. 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 45 s to 1 min 30 s, 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 KpnI/ApaI or KpnI/XbaI digested pAC5.1/V5-His A or PIZ vector (V413020, Invitrogen) using T4 DNA ligase (M0202L, NEB). Recombinant expression vectors 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 protein expression in S2 or SF9 cell lines.

**Insect cell culture and transient transfection**

S2 cells or SF9 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). For DNA transfection, 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 based on the manufacturer’s instructions. After 7 h transfection, S2 or SF9 cells were centrifuged and resuspended in complete growth medium to induce protein expression for 48 h. The cell culture medium and cell lysate were analyzed by Western blot.

**Western blot analysis and immunoprecipitation (Co-IP) assay**

The Western blot analysis was performed as described previously (Zhong et al., 2012). Briefly, cell culture medium was collected, S2 or SF9 cells (2×10⁶ cells/well) were homogenized in lysis buffer [50 mM Tris-HCl, pH 7.4, 150 mM NaCl, 5 mM EDTA, 1% NP-40, 0.5 mM PMSF, protease inhibitor cocktail (P8340, Sigma-Aldrich)]. The cell debits were removed and the supernatants were collected as cell extracts for Western blot analysis. The cell culture media (10 μl each) and cell extracts (10 μl each, equivalent to ~5×10⁴ cells) were separated on 10% or 12% 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 and alkaline phosphatase-conjugate anti-mouse antibody (A4312, Sigma-Aldrich, 1:10,000) was used as secondary antibody for color development. The signal was developed by alkaline phosphatase (AP)-conjugate color development Kit (#170-6432, Bio-Rad).

For co-immunoprecipitation (Co-IP) assay, 300 μl of cell extract containing recombinant proteins, which is equivalent to approximately 5x10⁶ cells, was blocking for 30 min with 30 μl Protein G Sepharose (50% slurry, No.17-0618-01, GE Healthcare). The protein G Sepharose was removed by briefly centrifugation and the supernatant was incubated with anti-Flag M2 or anti-V5 antibody (final concentration of 10 μg/ml) at 4°C for 10 h with gentle rocking. Then, 30 μl 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 four times with lysis buffer and analyzed by subsequent immunoblotting (Zhong et al., 2012).

Dual-Luciferase Reporter Assay

For Dual-luciferase reporter assays, S2 or SF9 cells were seeded in 96-well culture plates (10⁴ cells/well) overnight in serum-free medium. These cells were then transiently co-transfected with recombinant pAC5.1V5-His A or pIZ/V5-His expression plasmid (0.3 μg), pGL3B (empty vector) or different pGL3B firefly luciferase reporter plasmids from the promoters of M. sexta or D. melanogaster antimicrobial peptide (AMP) genes, and several mutants of M. sexta lysozyme or moricin promoters (0.15 μg) (Rao et
al., 2011), and renilla luciferase reporter plasmid (0.015 μg) (as an internal standard) (pRL-TK, Promega). After overnight transfection, serum-free medium was replaced with complete growth medium for protein expression, and firefly luciferase and renilla luciferase activities were measured 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) was obtained as the ratio of firefly luciferase activity to renilla luciferase activity. RLA from S2 or SF9 cells co-transfected with empty pAC5.1/V5-His A or pIZ/V5-His and different reporter vectors was 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.
<table>
<thead>
<tr>
<th>Primers</th>
<th>Forward Primer (5’ $\rightarrow$ 3’)</th>
<th>Reverse Primer (5’ $\rightarrow$ 3’)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>For cloning</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rel2-F1</td>
<td>ATTTAAATATAGTTAGACTGAAGTTTTAGCGCTCA</td>
<td>Rel2-R1</td>
</tr>
<tr>
<td>Dorsal-F1</td>
<td>CAATCGCAGCCAATTCGCACAAGGTAAC</td>
<td>Dorsal-R1</td>
</tr>
<tr>
<td><strong>For RT-PCR</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rel2A-F</td>
<td>AGTGTGCGTCTACGAA</td>
<td>Rel2A,2B-R</td>
</tr>
<tr>
<td>Rel2B-F</td>
<td>GATCGCGGTTTCTCAGCACTG</td>
<td></td>
</tr>
<tr>
<td>Dorsal-F1</td>
<td>CGGGGTACCCAAAATGGCCCGTACGTGTCGTGTCAGCAG</td>
<td>Dorsal-R</td>
</tr>
<tr>
<td>RPS3-F</td>
<td>GTGTGGAGGTGGTGGT</td>
<td>RPS3-R</td>
</tr>
<tr>
<td><strong>For protein expression</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rel2A,2B-pAC5.1-KpnI-N</td>
<td>CGGGGTACC GGAATTTGCCACATTTTAGGTTGCTACCAATTTTAATAAACATCCACTGCGGTTAGTTGCCGCTTGAATTGCACAGGTCGTCGTGCAGCAG</td>
<td>Rel2A-pAC5.1-ApaI-C</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Rel2B-pAC5.1-ApaI-C</td>
</tr>
<tr>
<td>Dorsal-pAC5.1-KpnI-N</td>
<td>CGGGGTACCCAAAATGGCCCGTACGTGTCGTGTCAGCAG</td>
<td>Dorsal-pAC5.1-ApaI-C</td>
</tr>
<tr>
<td>Dorsal-pIZ-KpnI-N</td>
<td>CTGTTGACCCATTTAGGCAACTGCCGCGCGCGACGTGACGTGCTGTCGTG</td>
<td>Dorsal-pIZ-XbaI-C</td>
</tr>
<tr>
<td>Rel-pIZ-KpnI-N</td>
<td>CTTGTTGACCCATTTAGGCAACTGCCGCGCGCGACGTGACGTG</td>
<td>Rel-pIZ-XbaI-Flag-C</td>
</tr>
<tr>
<td>Rel-RHD-EcoRI-N</td>
<td>CCAGAATTCTAGCCCATCTACAGATCACA GA</td>
<td>Rel-RHD-NotI-Flag-C</td>
</tr>
</tbody>
</table>

Table 7 PCR primers used in this study
<table>
<thead>
<tr>
<th>Primers</th>
<th>Forward Primer (5’ → 3’)</th>
<th>Reverse Primer (5’ → 3’)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lysozyme&lt;sup&gt;κB5-rev&lt;/sup&gt;-N</td>
<td>CTAAACTAATATTGG GACTTTACTGCCAGT ATCAACCTGGATT</td>
<td>Lysozyme&lt;sup&gt;κB5-rev&lt;/sup&gt;-C ATACTGGCAGGTAAA GTCCCAATATTAGTT TAGGTACCTATCG</td>
</tr>
<tr>
<td>Moricin&lt;sup&gt;κB5-GATA&lt;/sup&gt;-rev-N</td>
<td>CTAAAATATTCCA GATAACGATAGGGAC TTTACTGCCAGTATC</td>
<td>Lysozyme&lt;sup&gt;κB5-GATA-rev&lt;/sup&gt;-C GTAAAGTCCCTATCG TTATCTGAATATTAG TTTAGGTACCTATCG</td>
</tr>
<tr>
<td>Moricin&lt;sup&gt;κB&lt;/sup&gt;-N</td>
<td>CTAAAATATTGTAAAGTCCCTGCCAGT ATCAACCTGGATT</td>
<td>Lysozyme&lt;sup&gt;κB&lt;/sup&gt;-C ATACTGGCAGGGACT TTACAATATTAGTT AGGTACCTATCG</td>
</tr>
<tr>
<td>Lysozyme&lt;sup&gt;κB5-GATA&lt;/sup&gt;-N</td>
<td>ATTGTAAGTCCCTAT CGTTATCTGTGCCAGT ATCAACCTGGAAAT</td>
<td>Lysozyme&lt;sup&gt;κB5-GATA-c&lt;/sup&gt; ATACTGGCAGAGGACA CACGATAGGGACTTTA CAATATTAGTT</td>
</tr>
<tr>
<td>Moricin&lt;sup&gt;κB-rev&lt;/sup&gt;-N</td>
<td>GCCTTTGTAAATTAG AAAGTCCCTATCGGT ATCTGAGAGTATAAAA</td>
<td>Moricin&lt;sup&gt;κB-rev&lt;/sup&gt;-C GATAACGATAGGGA CTTTCTAAATTTACA AGGCCCGGAATTAC</td>
</tr>
<tr>
<td>Moricin&lt;sup&gt;κB&lt;/sup&gt;-N</td>
<td>GCCTTTGTAAATTGG GACTTTTCTTGATCGTA ATCTGAGAGTATAAAA</td>
<td>Moricin&lt;sup&gt;κB-c&lt;/sup&gt; GATAACGATAAGAA AGTCCCAATTTACA AGGCCCGGAATTAC</td>
</tr>
<tr>
<td>Moricin&lt;sup&gt;κB-rev-ΔGATA&lt;/sup&gt;-N</td>
<td>GCCTTTGTAAATTAG AAAGTCCCTATGTAAT ATATTAGAGTATAAA</td>
<td>Moricin&lt;sup&gt;κB-rev-ΔGATA-c&lt;/sup&gt; TATAATTATAGGGAC TTTCATATTTACAAAGGCCCAGAAAATTAC</td>
</tr>
<tr>
<td>Moricin&lt;sup&gt;κB-ΔGATA&lt;/sup&gt;-N</td>
<td>GCCTTTGTAAATTAG GGGACCTTCTTATAATATATATTAGAGTATAAA</td>
<td>Moricin&lt;sup&gt;κB-ΔGATA-c&lt;/sup&gt; TATAATTATAGGAAGG GCCCAATTTACAAAGGCCCGGAATTAC</td>
</tr>
</tbody>
</table>
Results

Cloning and sequence analysis of *M. sexta Dorsal* and *Rel2*

Based on the partial sequences from the *M. sexta* EST database, we performed PCR amplification and RACE to obtain the full-length cDNA of two *M. sexta Relish* isoforms *Rel2A* (GenBank accession no. HM363513) and *Rel2B* (GenBank accession no. HM363514), and a *Dorsal* homologue (GenBank accession no. HM363515). *MsRel2A* cDNA is 1677 bp long with an open reading frame (ORF) of 1191 bp, which encodes a putative protein of 397 amino acids. *MsRel2B* cDNA is 2057 bp with an ORF of 1326 bp encoding a putative protein of 442 residues (Fig. 21). *MsRel2A* and *MsRel2B* have an identical RHD domain and only differ at the C-terminal regions. *MsRel2A* and *MsRel2B* share 91.7% identity, but *MsRel2B* is about 45 amino acids longer at the C-terminus. Sequence analysis showed that *MsDorsal*-RHD is most similar to RHDs of invertebrate species and it belongs to the class II NF-κB, while *MsRel2*-RHD is most similar to RHDs of Relish proteins, and thus *MsRel2A* and *MsRel2B* belong to the class I NF-κB.

The expression profile of *M. sexta Dorsal* and *Rel2*

Tissue distribution profile of *MsDorsal*, *MsRel2A* and *MsRel2B* in *M. sexta* naïve larvae was determined by RT-PCR. The results showed that *MsDorsal* mRNA was expressed at high levels in hemocytes and testis, at relatively high levels in fat body and midgut, and at a lower level in epidermis (Fig. 20A). *MsRel2A* transcript was expressed at high levels in all the tissues tested, but *MsRel2B* mRNA was only detected at high levels in the midgut and testis. These results suggest that *MsDorsal* and *MsRel2A* may be
ubiquitously expressed in tissues, while MsRel2B is more restricted to certain tissues (midgut and testis). To determine whether these NF-κB factors are induced by microbial infection, *M. sexta* larvae were injected with *S. aureus*, *E. coli* and *S. cerevisiae*, and MsDorsal and MsRel2 transcripts were measured by RT-PCR. Compared to the naïve or water injected larvae, the expression of MsDorsal, MsRel2A and MsRel2B mRNAs in the hemocytes, fat body and midgut didn’t change significantly after microbial injection (Fig. 20B and 20C), suggesting that expression of NF-κB factors in *M. sexta* larvae is not induced in response to microbial infection.
Figure 20. Expression profile of *M. sexta* Dorsal, Rel2A and Rel2B transcripts in different tissues by RT-PCR.

A, total RNA samples were prepared from different tissues of naïve larvae and cDNA was prepared by reverse transcription. RT-PCR was carried out using primers specific for MsDorsal, MsRel2A, MsRel2B and ribosomal protein S3. B and C, total RNA samples were prepared from different tissues of naïve larvae or larvae injected with water, *S. aureus*, *E. coli*, or *S. cerevisiae*. RT-PCR analysis was carried out as described above. The sizes of bands are: MsDorsal (553 bp), MsRel2A (508 bp), MsRel2B (468 bp), rps3 (317 bp). Hc, hemocyte; Fb, fat body; Mg, midgut; Epi, epidermis; Te, testis; M, DNA size ladders.
MsRel2 and MsDorsal activate promoter activity of M. sexta AMP genes.

D. melanogaster and M. sexta AMP genes can be significantly upregulated by various microbial components (Rao and Yu, 2010). Since NF-κB transcription factors play central roles in the activation of AMP expression, we carried out dual-luciferase reporter assays to determine if MsDorsal and MsRel2 could activate different AMP gene promoters (Rao et al., 2011). Recombinant MsDorsal-RHD (Dl-RHD) and MsRel2-RHD (Rel2-RHD) (only the RHD domains), as well as MsRel2A and MsRel2B (full length proteins) were successfully expressed in both Drosophila S2 cells and S. frugiperda Sf9 cells (Fig. 21) and could be detected in both the cytoplasm (Fig. 21, lanes 1 and 3) and the nucleus (Fig. 21, lanes 2 and 4). Over-expression of recombinant DI-RHD and Rel2-RHD in S2 cells could significantly activate promoter activity of several M. sexta AMP gene promoters and B. mori lebocin-4 promoter (Fig. 22A). Most AMP gene promoters, including M. sexta moricin, defensin-1 and two attacins and B. mori lebocin-4, were activated to significantly higher levels by Rel2-RHD than by DI-RHD, while M. sexta cecropin promoter was activated to an equally high level by DI-RHD and Rel2-RHD, but M. sexta lysozyme promoter was activated to a significantly higher level by Rel2-RHD than by DI-RHD (Fig. 22A). Similarly, over-expression of the full length Rel2A and Rel2B could also significantly stimulate the activity of AMP gene promoters, and Rel2A and Rel2B could activate the activity of most AMP gene promoters (except M. sexta attacin-2 and lysozyme) to similar high levels (Fig. 22B).
Figure 21. Western blot analysis of recombinant MsDorsal-RHD, MsRel2A, MsRel2B and MsRel2-RHD in S2 and SF9 cells.

Expression plasmids of V5-tagged MsDorsal-RHD, Flag-tagged MsRel2-RHD and V5-tagged MsRel2A and MsRel2B were used to transfect S2 or SF9 cells, respectively. Cytoplasmic or nuclear lysate was prepared separately for Western blot analysis using monoclonal V5 or Flag antibody. Odd lanes: cytoplasmic proteins; Even lanes: nuclear proteins; A and B: Lanes 1 & 2: S2 cells; lanes 3 & 4: SF9 cells; C: lanes 1 & 2, Rel2A; lanes 3 & 4, Rel2B in Sf2 cells.
Figure 22. MsRel2 and MsDorsal activate expression of *M. sexta* and *B. mori* AMP gene promoters.

A, Activation of AMP gene promoters by MsDorsal-RHD and MsRel2-RHD in S2 cells. B, Activation of AMP gene promoters by MsRel2A and MsRel2B in S2 cells. The S2 cells were transfected with 0.3 μg of NF-κB plasmid, 0.15 μg of reporter gene plasmid together with 0.015 μg pRL-TK *Renilla reniformis* luciferase vector as the internal control. 48 hours after transfection, cells were harvested and analyzed with Dual-Luciferase Reporter Assay System (E1980, Promega). Bars represent the mean of three individual measurements ± SEM. Asterisks indicate significant difference (p<0.05) determined by one-way ANOVA.
Activation of moricin and lysozyme promoters by MsDorsal and MsRel2

We have previously characterized a moricin promoter (1400 bp) and also cloned a lysozyme promoter (1203 bp) in M. sexta (Rao et al., 2011). Five NF-κB sites were predicted in the MsMoricin promoter, but only the proximal NF-κB5 (Mor-NF-κB5) was activated in Sf9 cells by peptidoglycan from E. coli (Rao et al., 2011). To test whether MsMoricin can be activated by MsDorsal and/or MsRel2 and whether the five predicted NF-κB sites are all functionally active, reporter luciferase assays were performed with MsMoricin promoter and its deletion and mutation promoters in Sf9 cells since MsMoricin promoter showed low activity in S2 cells but high activity in Sf9 cells (Rao et al., 2011). All the MsMoricin promoters showed almost no activity in Sf9 cells after overexpression of MsDorsal-RHD (Fig. 23A), indicating that MsMoricin is not activated by Dorsal. MsMoricin (1400 bp), MsMoricin-725 (725 bp) and MsMoricin-242 (242 bp) promoters were activated to similar high levels by MsRel2-RHD, but MsMoricin-40 (40 bp) promoter did not have any activity (Fig. 23A). Deletion of the predicted NF-κB1, NF-κB2, NF-κB3 or NF-κB4 site did not have an effect on the activity of MsMoricin promoter stimulated by MsRel2-RHD, but deletion or mutation of NF-κB5 significantly decreased the activity of MsMoricin promoter (Fig. 23B), indicating that MsMoricin is activated by Rel2 and only the NF-κB5 site is functionally active.

We showed that MsLysozyme promoter was activated by MsDorsal-RHD (Fig. 22A), and only one NF-κB site was predicted in the lysozyme promoter, which differs from moricin NF-κB5 site only at the two 3’-end nucleotides but the two NF-κB sites have opposite direction (Fig. 23C). To test whether the consensus sequence and direction of NF-κB site as well as other transcription factor binding sites are required for activation
of AMP promoters by Dorsal and Rel2, we made several mutations in the Mor-242 promoter (242 bp) by replacing NF-κB5 site with lysozyme NF-κB site (Lyz-κB, with an opposite direction to NF-κB5) or reversed lysozyme NF-κB site (Lyz-κB-Rev, with the same direction to NF-κB5) and with or without GATA-1 site (Fig. 23E), since GATA-1 is required for NF-κB5 to activate moricin promoter (Rao et al., 2011). It has also been reported that Drosophila Dif and Relish can form heterodimers to activate AMPs (Tanji et al., 2010). Thus, we also test whether co-expression of MsDorsal-RHD and MsRel2-RHD has an effect on the activity of the moricin promoters. Our results showed that among the five moricin promoters, only Mor-242 promoter (containing both NF-κB5 and GATA-1 sites) was activated by MsRel2-RHD, and only Mor\textsuperscript{Lyz-κB-Rev} promoter (containing both the reversed lysozyme NF-κB site that has the same direction to NF-κB5 and GATA-1) was activated by MsDorsal-RHD, but none of the five promoters was activated by co-expression of MsRel2-RHD and MsDorsal-RHD (Fig. 23D).

To determine activation of MsLysozyme promoter by Dorlal and Rel2, we constructed four deletion promoters and also four mutation promoters by replacing lysozyme NF-κB with moricin NF-κB5 (opposite direction to lysozyme NF-κB) or reversed NF-κB5 (same direction to lysozyme NF-κB) with or without GATA-1 site (Fig. 24C). Since lysozyme promoter showed similar high activities in both Drosophila S2 and S. frugiperda Sf9 cells (Rao et al., 2011), activation of lysozyme promoters was performed in S2 cells. The results showed that only the lysozyme promoter (1203 bp) but not the four deletion promoters was activated by MsDorsal-RHD, and all five lysozyme promoters showed low basal activities when MsRel2-RHD was overexpressed (Fig. 24A), indicating that the distal lysozyme NF-κB is functional active and it binds to Dorsal but
not Rel2. Among the lysozyme promoter and the four mutated promoters, only lysozyme promoter was activated by MsDorsal-RHD, and only $L_{yz}^{Mor-\kappa B5-GATA}$ promoter (containing moricin NF-κB5 and GATA-1 sites) was activated by MsRel2-RHD, and none of the five lysozyme promoters was activated by co-expression of MsDorsal-RHD and MsRel2-RHD (Fig. 24B).
Figure 23. *Moricin* promoter is regulated by *MsRel2-RHD* but not by *MsDorsal-RHD.*

A, the empty vector (PIZ) or expression vector (Dl-RHD or Rel2-RHD) was used to transfect SF9 cells along with different truncated *Moricin* promoter-driven luciferase reporters. The relative induction of *Moricin* promoter was determined by dual-luciferase assays. B, the empty vector (PIZ) or expression vector (Dl-RHD or Rel2-RHD) was used to transfect SF9 cells along with different NF-κB-deleted or the NF-κB-mutated reporter. The relative induction of each reporter was determined. C, DNA sequence of *Moricin* NF-κB5 and *Lysozyme* NF-κB binding site. D, the empty vector (PIZ), or expression vector (Dl-RHD, Rel2-RHD or Dl-RHD+Rel2-RHD) was used to transfect SF9 cells along with the *MsMoricin, Mor*<sup>Lyz-xB</sup>, *Mor*<sup>Lyz-xB-ΔGATA</sup>, *Mor*<sup>Lyz-xB-Rev</sup>, or *Mor*<sup>Lyz-xB-Rev-ΔGATA</sup> reporter. The relative induction of each reporter was determined. Bars represent the mean of three individual measurements ± SEM. Identical letters are not significant difference (p>0.05), while different letters indicate significant difference among groups (p<0.05) determined by one way ANOVA followed by a Tukey’s multiple comparison test. E, Schematic representation of different mutations in the *Mor-242* promoter.
Figure 24. *Lysozyme* promoter is regulated mainly by *Ms*Dorsal-RHD.

A, the empty vector (pAC5.1A) or the expression vector (pAC-Dorsal-RHD or pAC-Rel2-RHD) was used to transfect S2 cells along with the full-length *Ms*Lysozyme promoter (*Lyz*-1203) or the truncated forms (*Lyz*-888, *Lyz*-596, *Lyz*-345 and *Lyz*-67) and the pRL-TK internal control. The relative induction of *Lysozyme* promoter activity was determined by dual-luciferase assays. B, the relative induction of *Lyz*-1203, *Lyz*\(^{\text{Mor-κB5-Rev}}\), *Lyz*\(^{\text{Mor-κB5-GATA-Rev}}\), *Lyz*\(^{\text{Mor-κB5}}\), or *Lyz*\(^{\text{Mor-κB5-GATA}}\) by Dorsal-RHD, Rel2-RHD or Rel2+Dorsal was determined. Bars represent the mean of three individual measurements ± SEM. Identical letters are not significant difference (p>0.05), while different letters indicate significant difference among groups (p<0.05) determined by one way ANOVA followed by a Tukey’s multiple comparison test. C, Schematic representation of different mutations in the *Lyz*-1203 promoter.
Interaction of MsDorsal with MsRel2

It has been reported that *Drosophila* Dif and Relish can form heterodimers and expression of peptide linked Dif-Relish-N (the N-terminal domain of Relish) can activate AMP genes in the Toll and IMD pathways (Tanji et al., 2010). We showed above that co-expression of *Ms*Dorsal-RHD and *Ms*Rel2-RHD abolished activation of moricin promoter by *Ms*Rel2-RHD (Fig. 23C) and *lysozyme* promoter by *Ms*Dorsal-RHD (Fig. 24B). To test interaction between *Ms*Dorsal and *Ms*Rel2, we over-expressed *Ms*Rel2-RHD-Flag and *Ms*Dorsal-RHD-V5 in S2 cells and performed co-immunoprecipitation (Co-IP) experiments. Expression of *Ms*Rel2-RHD-Flag and *Ms*Dorsal-RHD-V5 in S2 cells was confirmed by Western blot analysis with monoclonal anti-Flag or anti-V5 antibody (Fig. 25A, B, D and E, lanes 2 and 3 for input), and expression of *Ms*Dorsal-RHD-V5 was also confirmed by polyclonal rabbit anti-Dorsal antibody (Fig. 25C, lane 3). Results from the Co-IP experiments showed that anti-Flag antibody could pull down *Ms*Rel2-RHD-flag (Fig. 25A, lane 4), and co-precipitated *Ms*Dorsal-RHD-V5, which was recognized by anti-V5 antibody (Fig. 25B, lane 4) and anti-Dorsal antibody (Fig. 25C, lane 4). Likewise, anti-V5 antibody could pull down *Ms*Dorsal-RHD-V5 (Fig. 25D, lane 4), and co-precipitated *Ms*DRel2-RHD-Flag, which was recognized by anti-Flag antibody (Fig. 25E, lane 4). These results suggest that *M. sexta* Dorsal and Rel2 may form heterodimers *in vivo*.

118
Figure 25. *M. sexta* Rel2 interacts with Dorsal.

*Ms*Rel2-RHD and *Ms*Dorsal-RHD were expressed in *Drosophila* S2 cells for 48 h. Proteins in cell lysates were used for co-immunoprecipitation (Co-IP) assays as described in the Materials and Methods. *Ms*Dorsal-RHD contained a V5-tag, while *Ms*Rel2-RHD contained a Flag-tag. Monoclonal antibody (anti-FLAG or anti-V5) was added to combined cell lysates and immunoprecipitated (IP) proteins or Co-IP proteins were detected by immunoblotting using anti-Flag (A and E), anti-V5 (B and D), anti-Dorsal antibody (C) as the primary antibody, alkaline phosphatase-conjugate anti-mouse or anti-rabbit antibody as the secondary antibody and alkaline phosphatase (AP) conjugate color development kit. Lanes 1-3 from A-E were cell lysates alone (protein inputs).
Dorsal-Rel2 heterodimers as negative regulators in AMP gene expression

We showed that co-expression of MsDorsal-RHD and MsRel2-RHD abolished activation of moricin promoter by MsRel2-RHD (Fig. 23C) and lysozyme promoter by MsDorsal-RHD (Fig. 24B), respectively, and MsDorsal-RHD and MsRel2-RHD could interact with each other to form heterodimers (Fig. 25). To test whether Dorsal-Rel2 heterodimers may also have an effect on activation of other AMP gene promoters, luciferase assays were performed in S2 cells for Drosophila AMP gene promoters and in Sf9 cells for M. sexta AMP gene promoters after overexpression of MsDorsal-RHD or MsRel2-RHD alone, or co-expression of MsDorsal-RHD and MsRel2-RHD. Overexpression of MsDorsal-RHD or MsRel2-RHD alone could activate all the Drosophila and M. sexta AMP gene promoters to certain levels (Fig. 26). Co-expression of MsDorsal-RHD and MsRel2-RHD in S2 cells stimulated Drosophila cecropin, diptericin and metchnikowin promoters to significantly higher levels than by MsDorsal-RHD or MsRel2-RHD alone, but inhibited activation of attacin (Fig. 26A). However, co-expression of MsDorsal-RHD and MsRel2-RHD in Sf9 cells abolished activation of all the AMP gene promoters tested, including M. sexta defensin-1, attacin-1, attacin-2, cecropin, and B. mori lebocin-4 (Fig. 26B). These results suggest that in M. sexta, Dorsal-Rel2 heterodimers may serves as negative regulators in activation of AMPs.
Figure 26. MsDorsal-MsRel2 heterodimer negatively regulates *M. sexta* AMP expression.

DI-RHD and Rel2-RHD were expressed separately or together in S2 (A) or SF9 (B) cells. The induction of *Drosophila* or *M. sexta* AMP gene promoters were measured, respectively. Drs: *Drosomycin*; Def: *Defensin*; Cec: *Cecropin*; Att: *Attacin*; Dpt: *Diptericin*; Dro: *Drosocin*; Met: *Metchnikowin*; Leb: *Lebocin*. Bars represent the mean of three individual measurements ± SEM. Identical letters are not significant difference (p>0.05), while different letters indicate significant difference among groups (p<0.05) determined by one way ANOVA followed by a Tukey’s multiple comparison test.
Discussion

In this study, we identified a Dorsal homolog (MsDorsal) and two short isoforms of Relish (MsRel2A and MsRel2B) in *M. sexta*, and investigated their roles in activation of AMP gene promoters. We found that MsDorsal and MsRel2A and 2B were functional active NF-κB factors that could activate *M. sexta* and *D. melanogaster* AMP gene promoters differently. We also showed that MsDorsal could interact with MsRel2 and may form Dorsal-Rel2 heterodimers, which served as negative regulators in activation of *M. sexta* AMP genes.

In *Drosophila*, Dorsal, Dif and Relish have been identified and characterized (Hetru and Hoffmann, 2009), and homo- and hetero-dimers among the three NF-κB factors have been observed in the flies (Ganesan et al., 2011). Peptide linked Dif-Relish-N (N-terminal fragment of Relish) heterodimers can activate AMPs regulated by both the Toll and IMD pathways (Tanji et al., 2010). Since the covalent linked Dif-Relish-N heterodimers may form dimers of heterodimers, in which Dif-Dif homodimers can be on one end and Relish-N-Relish-N homodimers can be on the other end, thus such dimers of heterodimers may still function as Dif-Dif and Relish-Relish homodimers. In addition, it is not clear how Dif-Relish heterodimers are form in vivo in flies as both the Toll and IMD pathways must be activated to release Dif and generate Relish-N for formation of heterodimers. Interestingly, alternative splice forms of Relish have been identified in the mosquito *Ae. Aegypti* (Shin et al., 2002; Shin et al., 2005), and two isoforms of Relish have also been identified in the silkworm *B. mori* (Tanaka et al., 2007). In vitro assays show that the short isoform of Relish containing the N-terminal fragment alone can serve as transcription activator in *Ae. aegypti*, but the *BmRelish2* isoform is a dominant
negative regulator of the active *BmRelish1* (Shin et al., 2002; Shin et al., 2005). The biological roles of short Relish isoforms in insects are not clear.

*M. sexta* Rel2A and Rel2B contain an identical RHD domain and only differ in the C-terminal residues. This might be a result of alternative splicing. Alternative splicing occurs frequently in insects. In mosquito *Aedes aegypti*, there are three spliced forms of Relish homolog (Shin et al., 2002). The 3.9 kb transcript contains both the RHD domain and the inhibitor kappa B (I kappa B)-like domain, while the other two shorter transcripts lack either the RHD or the I kappa B domain. The *Anopheles gambiae* REL2 gene produces two spliced forms: a full-length REL2-F and a shorter REL2-S (Meister et al., 2005). In *B. mori*, two transcripts *BmRelA* and *BmRelB* differ only in the 5’ region, *BmRelB* could activate the *Attacin* gene strongly, whereas *BmRelA* activated *Lebocin 4* gene strongly (Tanaka et al., 2007). Both *M. sexta* Rel2A and Rel2B were functionally active NF-κB factors with different tissue distribution, Rel2B was more restricted to the midgut and testis.

We showed that *MsDorsal* and *MsRel2* could form heterodimer to negatively regulate AMP expression. In *Drosophila melanogaster*, Dif and Relish can form heterodimers to activate AMP production (Tanji et al., 2010). Dimerization of NF-κB factors has been reported in both vertebrates and invertebrates. p65 and p50 can form heterodimers to bind NF-κB sites (Urban et al., 1991). p65/c-Rel dimer is related to neuro protection through activation of metabotropic glutamate receptors type 5 (Sarnico et al., 2008). In *Drosophila*, homodimers and heterodimers can form among Dorsal, Dif and Relish with different degrees of efficiency (Tanji et al., 2010). Therefore, further study is
needed to reveal NF-κB sites recognized by homo- and hetero-dimers of NF-κB factors in insects.
CHAPTER 5
CHARACTERIZATION OF A NOVEL MANDUCA SEXTA BETA-1, 3-GLUCAN RECOGNITION PROTEIN-3 (BGRP3) WITH MULTIPLE FUNCTIONS

Any questions about whether a particular use is authorized and any requests for permission to publish, reproduce, distribute, display or make derivative works from any Elsevier published works, images or other material should be directed to the Elsevier Rights Department.

© 2014 Elsevier Ltd. All rights reserved.
Abstract

Recognition of pathogens by insect pattern recognition receptors is critical to mount effective immune responses. In this study, we reported a new member (βGRP3) of the β-1, 3-glucan recognition protein (βGRP) family from the tobacco hornworm Manduca sexta. Unlike other members of the M. sexta βGRP family proteins, which contain an N-terminal small glucan binding domain and a C-terminal large glucanase-like domain, βGRP3 is 40-45 residues shorter at the N-terminus and lacks the small glucan binding domain. The glucanase-like domain of βGRP3 is most similar to that of M. sexta microbe binding protein (MBP) with 78% identity. βGRP3 transcript was mainly expressed in the fat body, and both its mRNA and protein levels were not induced by microorganisms in larvae. Recombinant βGRP3 purified from Drosophila S2 cells could bind to several Gram-negative and Gram-positive bacteria and yeast, as well as to laminarin (β-1, 3-glucan), mannan, lipopolysaccharide (LPS), lipoteichoic acid (LTA), and meso-diaminopimelic acid (DAP)-type peptidoglycan (PG), but did not bind to Lysine-type PG. Binding of βGRP3 to laminarin could be competed well by free laminarin, mannan, LPS and LTA, but almost not competed by free PGs. Recombinant βGRP3 could agglutinate Bacillus cereus and Escherichia coli in a calcium-dependent manner and showed antibacterial (bacteriostatic) activity against B. cereus, novel functions that have not been reported for the βGRP family proteins before. M. sexta βGRP3 may serve as an immune surveillance receptor with multiple functions.
**Introduction**

The innate immune system is the first line of defense against pathogenic invaders in animals (Mogensen, 2009; Sukhithasri et al., 2013) and relies on surveillance molecules, named pattern recognition receptors (PRRs), to recognize pathogen-associated molecular patterns (PAMPs) present in pathogens but not in the hosts (Charroux et al., 2009; Charroux and Royet, 2010; Kanost et al., 2004; Lemaitre and Hoffmann, 2007). PRRs include C-type lectins, β-1,3-glucan recognition/binding proteins (βGRPs/BGRPs) and Gram-negative bacteria binding proteins (GNBPs), peptidoglycan recognition proteins (PGRPs), Toll-like receptors (TLRs), retinoic acid-inducible gene 1 (RIG-I)-like receptors, nucleotide-binding oligomerization domain (NOD) receptors, and Dectin receptors (Pal and Wu, 2009; Takeuchi and Akira, 2010). Examples of PAMPs include lipopolysaccharide (LPS), lipoteichoic acid (LTA) and peptidoglycans (PG) from bacteria and β-1,3-glucan from fungi (Rao and Yu, 2010). Upon binding to PAMPs, PRRs can stimulate humoral and cellular immune responses, such as phagocytosis, nodule formation, encapsulation and melanization, synthesis of antimicrobial peptides and activation of the prophenoloxidase (proPO) system (Jiang et al., 2010; Kanost et al., 2004; Thompson et al., 2011). In insect hemolymph, activation of serine proteinase cascade leads to the proteolytic activation of proPO to active phenoloxidase (PO) (Gupta et al., 2005; Jiang et al., 2010). PRRs such as C-type lectins, βGRPs/BGRPs and PGRPs can stimulate proPO activation in hemolymph when binding to various PAMPs (Jiang et al., 2010; Lee et al., 2004; Ma and Kanost, 2000; Rao et al., 2010; Wang et al., 2011).

Gram-negative bacteria binding protein (GNBP) was first characterized as a 50-kDa hemolymph protein from the silkworm *Bombyx mori* that can bind to *Escherichia*
coli (Lee et al., 1996). GNBPs actually belong to the β-1, 3-glucan recognition protein (βGRP/BGRP) family, which is one of the major pattern recognition receptors that can bind to β-1, 3-glucans on bacteria and fungi (Hughes, 2012; Jiang et al., 2004; Lee et al., 1996; Ma and Kanost, 2000; Wang et al., 2011). Members of the βGRP family proteins contain a small (~100 residues) N-terminal glucan binding domain and a large (~350 residues) C-terminal glucanase-like domain that lacks key residues in the active sites for glucanase activity (Ma and Kanost, 2000; Ochiai and Ashida, 2000). The small glucan binding domains of βGRPs bind to β-1, 3-glucan with a mechanism different from that of glucanase-like domains (Dai et al., 2013; Kanagawa et al., 2011; Mishima et al., 2009; Takahasi et al., 2009). βGRPs have been identified in invertebrates, including insects and crustaceans, and they can bind to microbial cell wall components, leading to activation of proPO (Cerenius et al., 1994; Duvic and Soderhall, 1990; Vargas-Albores et al., 1996; Vargas-Albores et al., 1997; Zheng and Xia, 2012). In the tobacco hornworm Manduca sexta, βGRP1 and βGRP2 have been identified and they can greatly stimulate proPO activation after binding to laminarin (β-1, 3-glucan) (Jiang et al., 2004; Ma and Kanost, 2000). M. sexta microbe binding protein (MBP), a β-1, 3-glucanase related protein, binds to bacteria and fungi, and MBP alone weakly stimulates proPO activation, but can significantly activate proPO when combined with different microbial elicitors (Wang et al., 2011). An inducible GNBP was purified from the silkworm, Bombyx mori (Hughes, 2012; Lee et al., 1996), and silkworm βGRP can bind to β-1, 3-glucan to initiate activation of the proPO cascade (Ochiai and Ashida, 2000). Drosophila DGNBP-1 can bind to LPS and β-1, 3-glucan and enhance immune gene expression induced by LPS and
β-1, 3-glucan (Kim et al., 2000). *Anopheles gambiae* GNBPs are involved in anti-*Plasmodium* responses (Warr et al., 2008).

In this paper, we reported the characterization and functional analysis of βGRP3, a new member of the *M. sexta* βGRP family. We investigated tissue distribution of βGRP3 transcript and induced expression of βGRP3 mRNA in fat body, hemocytes and midgut as well as βGRP3 protein in hemolymph by immune challenge. We also expressed and purified recombinant βGRP3 from *Drosophila* S2 cells and studied binding of βGRP3 to microorganisms and to various microbial cell wall components, including laminarin (β-1, 3-glucan), mannan, LPS, LTA, meso-diaminopimelic acid (DAP)-type and Lysine-type PGs. Interestingly, we found that βGRP3, which is 40-45 residues shorter at the N-terminus and lacks the small glucan binding domain, possessed novel properties with calcium-dependent agglutinating activity against *Bacillus cereus* and *E. coli* and strong antibacterial (bacteriostatic) activity against *B. cereus*.

**Materials and Methods**

**Insect rearing and *Drosophila* S2 cell line**

*M. sexta* eggs were purchased from Carolina Biological Supply (Burlington, NC). Larvae were reared on artificial diet at 25°C (Dunn, 1983), and the fifth instar larvae were used for the experiments. *D. melanogaster* Schneider S2 cells were purchased from American Type Culture Collection (ATCC).
cDNA cloning of $\beta$GRP3 and sequence analysis

An *M. sexta* EST sequence was predicted to encode a partial BGRP-like protein (Accession number: GR922389.1). We then designed primers based on the EST sequence to clone the full-length cDNA. Briefly, total RNA was prepared from the fat body of day 3 naïve larvae using TRIzol® Reagent (T9424, Sigma–Aldrich). For reverse transcription, total RNA was treated with RQ1 RNase-free DNase I (Promega) at 37°C for 30 min to remove contaminated genomic DNA, and DNase was inactivated by heating to 65°C for 20 min. Reverse transcription was performed using oligo(dT) primer (Promega) and ImProm-II reverse transcriptase (Promega) following the manufacturer’s instructions. *M. sexta* $\beta$GRP3 full-length cDNA was cloned using the forward primer $\beta$GRP3-F1: 5’-ACG ACT CGA TCA CAA GCA AC-3’ and the reverse primer $\beta$GRP3-R1: 5’-CAG AAC TTG AGC ATG GCT TT-3’. 5’and 3’ RACE reactions were performed using smarter race kit (Clontech). The ORF of $\beta$GRP3 was predicted from the nucleotide sequence using DNAMAN (Lynnon Corporation, Quebec, Canada). BLASTP (http://www.ncbi.nlm.nih.gov/) was used to search homologous $\beta$GRP, BGRP or GNBP protein sequences. Multiple sequence alignment was performed by ClustalW2 (http://www.ebi.ac.uk/Tools/msa/clustalw2/) with protein sequences retrieved from the NCBI database using default settings, and the phylogenetic tree was generated from the conserved regions of 18 proteins (not including the N-terminal regions) by Neighbor-Joining method with bootstrap of 1,000 replications using MEGA5 (Tamura et al., 2011). Figures were made from one representative set of data with the GraphPad Prism software (GraphPad, CA). Significance of difference was determined by one way ANOVA followed by a Tukey’s multiple comparison test using the same software (GraphPad, CA).
Expression and purification of recombinant βGRP3 and GFP

Recombinant βGRP3 was expressed in *E. coli* and purified for production of polyclonal antibody in rabbit. The cDNA sequence encoding mature βGRP3 (residues 24-441) was obtained by PCR using primers βGRP3-F2 (5’-CAT GCC ATG GTT TAT CGG TCC CGT TCC ACG TCTTTG-3’) and βGRP3-R2 (5’-CCC AAG CTT TTA CAG CGC GAC TAC TTT GAC ATAGTC-3’). The PCR fragment was purified by agarose gel electrophoresis, digested with *Nco*I and *Hind*III enzymes, ligated into the *Nco*I/*Hind*III-digested expression vector H6pQE-60 (Lee et al., 1994) and then transformed into competent *E. coli* XL1-Blue cells. Recombinant plasmids were prepared from positive clones and confirmed by restriction enzyme digestion and DNA sequencing. Single bacterial colony on petri dish plates containing recombinant plasmid DNA was inoculated into LB medium containing ampicillin (100 μg/ml) and incubated at 37°C overnight. The overnight culture was diluted 1:100 in LB medium and incubated at 37°C to OD600=0.8 and then isopropyl-D-thiogalactopyranoside (IPTG) (0.5 mM final concentration) was added to induce protein expression for another 6 h at 37°C. Bacterial cells were harvested by centrifugation and lyzed with the lysis buffer (50 mM Tris-HCl, pH 8.0, 50 mM NaCl, 0.5% Triton X-100, 2 mg/ml lysozyme), and recombinant protein was purified using Ni-NTA agarose beads (Qiagen) following the manufacturer’s instructions. The purified protein was further separated on 12% SDS-PAGE and the gel slice containing recombinant βGRP3 was used as an antigen to produce rabbit polyclonal antiserum at Cocalico Biologicals, Inc (Pennsylvania, USA). Recombinant βGRP3 was also expressed in *Drosophila* S2 cells and purified. The cDNA sequence encoding mature βGRP3 (residues 24-441) was amplified by PCR using primers βGRP3-F3 (5’-GGA
AGA TCT TAT CGG TCC CGT TCC ACG TCT TTG A-3') and βGRP3-R3 (5'-CCG CTC GAG CAG CGC GAC TAC TTT GAC ATA GTC-3'). The PCR product was recovered by agarose gel electrophoresis-Wizard® SV Gel and PCR Clean-Up System (A9285, Promega), subcloned into T-Easy vector (A1360, Promega), and recombinant plasmid was isolated. After digested with Bg/II/XhoI, cDNA fragment was recovered and inserted into Bgl/II/XhoI digested expression vector pMT/BiP/V5-His A (V413020, Invitrogen) using T4 DNA ligase (M0202L, NEB). Green fluorescent protein (GFP) in the expression vector pMT/Bip/V5-His/GFP (V413020, Invitrogen) was used as a control protein. Recombinant expression vectors were purified using PureYield™ Plasmid Miniprep System (A1222, Promega) according to the manufacturer’s instruction and the insert sequences were confirmed by DNA sequencing. These plasmids were used to transfect S2 cells and generate stable S2 cell lines following the manufacturer’s instructions. To purify recombinant βGRP3 and GFP, copper sulfate (final concentration of 250 μM) was added to stable S2 cells to induce expression of βGRP3 or GFP. Cell culture medium was collected starting at 24 h after protein expression for several days by collecting culture medium very day and re-suspending the cells with fresh medium, and recombinant βGRP3 and GFP were purified by affinity chromatography using anti-V5 agarose beads (A7345, Sigma–Aldrich) as described previously (Xu et al., 2012; Zhong et al., 2012). Fractions were analyzed by 10% SDS-PAGE, and those containing recombinant βGRP3 or GFP were combined and desalted using D-salt™ Excellulose™ GF-5 desalting column (#1851850, Pierce) pre-equilibrated with dH₂O. Recombinant βGRP3 or GFP was eluted with water, and fractions containing βGRP3 or GFP were pooled and concentrated for the following assays.
Tissue distribution and expression of βGRP3 mRNA in response to immune challenge

To determine tissue distribution of βGRP3 mRNA, day 2 fifth instar M. sexta naïve larvae were used in the following experiments. Larvae were injected with 50 μl of heated-killed E.coli DH5α, Serratia marcescens, Staphylococcus aureus, Bacillus subtilis (5 × 10⁷ cells/larva) or yeast (Saccharomyces cerevisiae) (10⁷ cells/larva), or ddH₂O (as a control), hemocytes, fat body, midgut, epidermis and testis were collected at 24 h post-injection 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% polyvinyl pyrrolidone, 1.7 mM PIPES). Total RNAs were isolated from these tissues using TRIzol® Reagent, and cDNAs were prepared from total RNAs (1 μg from each tissue) in a 25 μl reaction as described above. Real-time PCR was performed with SYBR Premix (Takara) on a 7500 system (Applied Biosystems) using primers βGRP3-F4 (5’-AAG AAG CGA CGG CTC TGC TTG G-3’) and βGRP3-R4 (5’-TGGAC GCC AGG AGT TGG TGC-3’) for βGRP3 gene, and RPS3-F (5’-GTT GCG AGG TGGTGG TTT C-3’) and RPS3-R (5’-CCG TTC TTG CCC TGT TGG TC-3’) for the control ribosomal protein S3 (rpS3) gene, and the reaction conditions were: 95°C for 10 min, followed by 40 cycles of 95°C for 15 s and 60°C for 1 min. Then dissociation curve analysis was performed. Data from three replicates of each sample were analyzed with SDS software (ABI) using a comparative method (2^-ΔΔCt) with the rpS3 gene as an internal control gene. Expression in naïve or water-injected larvae was set to 1.
Time-course induction of βGRP3 protein in hemolymph

To determine time-course induction of βGRP3 protein in hemolymph by Western blot analysis, day 2 fifth instar naïve larvae were injected with *E. coli*, *Micrococcus luteus* or *S. cerevisiae* as described above. Hemolymph (~50 μl) was collected every 2 h after injection (at least 4 larvae in each group). Hemolymph samples were centrifuged at 3000 g for 10 min and cell-free plasma samples were collected. Equal volumes of plasma samples (from at least 4 larvae) were mixed and 0.5 μl mixed plasma samples were analyzed by 10% SDS-PAGE. Western blot was performed using rabbit anti-βGRP3 antiserum (1:2000 dilution) and goat anti-rabbit IgG conjugated to alkaline phosphatase (1:10,000 dilution) (Sigma-Aldrich). To determine the concentration of βGRP3 protein in hemolymph, cell-free plasma samples from fifth instar naïve or *E. coli*-injected larvae were collected and pooled (from at least 30 larvae in each group). The amounts of βGRP3 in plasma samples were estimated by comparing the intensity of endogenous βGRP3 in the plasma samples with that of recombinant βGRP3 purified from *Drosophila* S2 cells by Western blot analysis using myImage AnalysisV1.1 (Thermo Scientific).

Binding of βGRP3 to microbial cell wall components

To test binding of βGRP3 to microbial cell wall components, plate ELISA assays were performed. Ultrapure TLR grade lipopolysaccharide (LPS) and peptidoglycan (PG) from *E. coli* K12 (LPS-K12 and PG-K12), lipoteichoic acid (LTA) and PG from *B. subtilis* (LTA-BS and PG-BS) and *S. aureus* (LTA-SA and PG-SA) were from Invivogen. Mannan and laminarin (β-1,3-glucan) were from Sigma-Aldrich. Briefly, wells of a flat bottom 96-well plate (Costar, Fisher) were coated with different microbial components (2
μg/well) as described previously (Yu and Kanost, 2000; Yu et al., 2005). The plates were placed overnight at room temperature until the water evaporated completely, heated to 60°C for 30 min, and then blocked with 200 μl/well of 1 mg/ml BSA in Tris buffer (TB) (50 mM Tris-HCl, 50 mM NaCl, pH 8.0) for 2 h at 37°C. The plates were rinsed and increasing concentrations of purified recombinant βGRP3 or GFP (a control protein) in TB containing 0.1 mg/ml BSA were added to the coated plates (50 μl/well). For competitive binding assay, recombinant βGRP3 (diluted to 2 μg/ml) was pre-incubated with increasing concentrations of free microbial components for 1 h at room temperature as described previously (Yu et al., 2005) and then added to the laminarin-coated plates. Binding of recombinant proteins to microbial components was for 3 h at room temperature. Then, the plates were incubated with monoclonal anti-polyhistidine antibody (Sigma-Aldrich) (1:2,000 in TB containing 0.1 mg/ml BSA) (100 μl/well) overnight at 4°C, followed by incubation with goat anti-mouse IgG conjugated to alkaline phosphatase (Sigma–Aldrich; 1:3,000 in TB containing 0.1 mg/ml BSA) (100 μl/well) for 2 h at 37°C. The plates were rinsed, p-nitro-phenyl phosphate (1 mg/ml in 10 mM diethanolamine, 0.5 mM MgCl₂) was added (50 μl/well), and absorbance at 405 nm of each well was determined every minute for 30 min using a microtiter plate reader (BioTek Instrument, Inc.). The data were analyzed by the GraphPad Prism software.

**Binding of βGRP3 to microorganisms**

Binding of recombinant βGRP3 to microorganisms was performed using a published method (Du et al., 2009) with slight modifications. Briefly, Gram-negative *E. coli*, *Pseudomonas aeruginosa*, *Serratia marcescens*, Gram-positive *S. aureus*, *B. subtilis*,

135
M. luteus and Bacillus cereus, as well as S. cerevisiae were cultured overnight. The microorganisms were pelleted, washed twice with Tris-buffered saline (TBS) (25 mM Tris-HCl, 2.7 mM KCl, 137 mM NaCl, pH 7.5), then thoroughly resuspended in TBS containing 2 mM CaCl$_2$ and 1 mM MgCl$_2$. Purified recombinant βGRP3 or GFP (as a control) (each at 2 μg in 50 μl) were added to 50 μl bacteria (4x10$^8$ cells/ml) or yeast (4x10$^7$ cells/ml) with rotation for 1 h at room temperature. The microorganisms were pelleted and the supernatants were collected as unbound proteins. These microorganisms were washed four times with TBS, subjected to elution with 7% SDS (100 μl) for 10 min, washed in 1 ml TBS four times, and finally lyzed in lysis buffer. Samples of the unbound proteins, TBS wash, 7% SDS elution and microbial lysates (to detect tightly bound proteins) were subjected to immunoblotting analysis using monoclonal anti-V5 antibody (V8012, Sigma-Aldrich).

Agglutination of microorganisms by βGRP3

E. coli, S. marcescens, P. aeruginosa, S. aureus, B. subtilis, B. cereus, M. luteus, and S. cerevisiae were used for the agglutination assays. Aliquots (3.5 μl) of bacteria (3x10$^9$ cells/ml) and yeast (2.5x10$^8$ cells/ml) suspensions were mixed with purified recombinant βGRP3 or GFP protein (final concentration of 8 μg/ml) in a total of 25 μl TBS containing 2 mM CaCl$_2$ and 1 mM MgCl$_2$ or 5 mM EDTA. The mixtures were incubated at room temperature for 1 h and then observed by microscopy as described previously (Yu et al., 1999; Yu et al., 2006).
Antimicrobial activity assay

The antimicrobial inhibition zone assay was performed according to a published protocol (Hultmark, 1998). A single colony of *E. coli*, *S. marcescens*, *B. cereus*, *B. subtilis*, or *S. cerevisiae* was grown overnight in LB broth (for bacteria) or YPD medium (for *S. cerevisiae*) at 37°C. Overnight bacterial cultures were then sub-cultured in LB medium and *S. cerevisiae* culture was sub-cultured in YPD medium until mid-log phase (OD600=~0.6). Fifteen micro liters (15 μl) of log-phase cultures were diluted into 20 ml LB or YPD medium containing 1.5% agarose (kept in a 45°C heating block), mixed well and spread immediately on a petri dish. After the medium was solidified, 5-mm wells were punched in the agarose. Buffer, ampicillin, or purified βGRP3 was added into each well. The plates were incubated at 37°C overnight, and the diameters of inhibition zones were measured and recorded. *B. cereus* growth curves were also determined in the presence of recombinant βGRP3 and various microbial components as described previously (Rao et al., 2012). Briefly, overnight *B. cereus* culture was sub-cultured in LB medium until mid-log phase, centrifuged at 1000 g for 10 min at 4°C, washed once in 10 mM Tris–HCl, pH 7.5, 0.1 mM EDTA, and diluted to OD600 =10^{-3} with LB medium. The diluted *B. cereus* cultures (75 μl/well) was mixed with purified βGRP3 alone (final concentration of 2 μg/ml), βGRP3 mixed with different microbial components (final concentration of 20 μg/ml), or water (Control) in 96-well plates, and the total volume was adjusted to 100 μl with LB medium. *B. cereus* cultures were incubated at 37°C with 220 rpm shaking. OD600 was measured every hour by PowerWave XS plate reader (Bio-Tek, USA). Bacterial growth curves were generated using the Graphpad Prism software (GraphPad, CA).
Results

cDNA cloning and sequence analysis of *M. sexta* βGRP3

An *M. sexta* EST sequence was found to encode a partial BGRP-like protein (Accession number: GR922389.1), which is different from *M. sexta* βGRP1 and βGRP2 proteins reported previously (Jiang et al., 2004; Ma and Kanost, 2000). To clone the new βGRP/BGRP homologue, we designed gene-specific primers based on the EST sequence. The full-length cDNA sequence of *M. sexta* βGRP3 was obtained by 5’ and 3’ RACE. The complete ORF of βGRP3 is 1323 nucleotides long encoding a protein of 441 amino acids with a putative signal peptide of 23 residues. The calculated molecular mass of mature βGRP3 is 48.9 kDa with an isoelectric point (pI) of 8.33. There are four potential N-linked glycosylation sites in βGRP3 (Asn46, Asn79, Asn124 and Asn177), and the glycosyl hydrolase like domain is located close to the C-terminus of βGRP3. βGRP3 is ~40-45 residues shorter at the N-terminus compared to *M. sexta* βGRP1 (487 residues), βGRP2 (482 residues), and microbe binding protein (MBP) (482 residues), and it lacks the small glucan binding domain. The N-terminal region (~57 residues, not including the 23-residues putative signal peptide) of βGRP3 shows no similarity to any sequence in the NCBI database (data not shown).

BLASTP search showed that the glucanase-like domain of *M. sexta* βGRP3 protein is most similar to *M. sexta* MBP, a β-1, 3-glucanase related protein (Wang et al., 2011), *Hyphantria cunea* β-1, 3-glucan binding protein (BGRP) (Shin et al., 1998), and *Bombyx mori* BGRP-2 (Lee et al., 1996) with 63% identity, followed by *Helicoverpa armigera* BGRP-3 (59% identity) (Pauchet et al., 2009), *Ostrinia furnacalis* BGRP (60%
identity), *Papilio xuthus* Gram-negative bacteria binding protein-3 (GNBP-3) (58% identity). *M. sexta* βGRP3 also shows similarity to *M. sexta* βGRP1 (36% identity) (Ma and Kanost, 2000), βGRP2 (35% identity) (Jiang et al., 2004) and β-1,3-glucanase (36% identity), but lacks the four key residues (W, E, D, E) required for glucanase activity (Fig. 27A) (Ma and Kanost, 2000). Therefore, βGRP3 may serve as a pattern recognition receptor in the immune system of *M. sexta*.

The BGRP family proteins have been identified in different invertebrate species. The phylogenetic tree constructed based on the glucanase-like domains of 17 BGRP/βGRP related proteins and *M. sexta* beta-1, 3-glucanase showed that *M. sexta* βGRP3 and microbe binding protein (*Ms*-MBP) are clustered in the same clade, βGRP1 and βGRP2 are in a different clade, while *M. sexta* β-1,3-glucanase is in a distant clade (Fig. 27B). These results suggest that the common ancestor of *M. sexta* βGRP family proteins diverged into three clades with possible functional variances.
Figure 27. Multiple sequence alignment of the active sites of glucanases and some BGRP/βGRP/GNBP proteins and phylogenetic tree of BGRP/βGRP/GNBP proteins.

Glucanase-related proteins and BGRP/βGRP/GNBP proteins were retrieved from the NCBI database, multiple sequence alignment was performed by ClustalW2, and only residues around the active site of glucanase were shown (A). The Glycine residue conserved in all sequences and the two highly conserved hydrophobic residues are shaded. Residues corresponding to the active site (W, E, D and E) of the *Bacillus macerans* glucanase are shown in bold. A phylogenetic tree of 17 BGRP/βGRP/GNBP proteins and *M. sexta* beta-1, 3-glucanase was constructed by MEGA5 using the Neighbor-Joining method with bootstraps of 1,000. *Ms*-βGRP1, *M. sexta* beta-1, 3-glucan recognition protein 1 (AF177982); *Ms*-βGRP2, *M. sexta* beta-1, 3-glucan recognition protein 2 (Q8ISB6); *Ms*-βGRP3, *M. sexta* beta-1, 3-glucan recognition protein 3 (ADK39022); *Ms*-MBP, *M. sexta* microbe binding protein (GNBP-like) (ADT82662); *Ms*-Glucanase, *M. sexta* beta-1, 3-glucanase (AEV66276); *Af*-BGRP, *Apis florea* beta-1, 3-glucan-binding protein (XP_003694585); *Hc*-BGRP, *Hyphantria cunea* beta-1, 3-glucan-binding protein (O96363.1); *Bm*-GNBP, *Bombyx mori* beta-1, 3-glucan recognition protein 2 precursor (NP_001037450); *Ha*-BGRP3, *Helicoverpa armigera* beta-1,3-glucan recognition protein 3 (ACI32828.1); *Px*-GNBP3, *Papilio xuthus* Gram-negative bacteria binding protein 3 (BAM19646.1); *Dp*-BGRP, *Danaus plexippus* beta-1,3-glucan-binding protein (EHJ65186.1); *Pr*-BGRP3, *Pieris rapae* beta-1, 3-glucan recognition protein 3 (ACI32823.1); *Sl*-BGRP2, *Spodoptera littoralis* beta-1, 3-glucan recognition protein 2 (ACI32819.1); *Cq*-GRP, *Culex quinquefasciatus* beta-1, 3-glucan recognition protein (XP_001850624.1); *Dw*-GK12242, *Drosophila willistoni* GK12242 (XP_002068068.1); *Gm*-GNBP, *Glossina morsitans* Gram-negative binding protein 1-like protein (ABC25063.1); *Aa*-GRP, *Aedes aegypti* AAEL007626-PA (EAT40654.2); *Ad*-GRP, *Anopheles darling* hypothetical protein (EFR19153.1).
Expression profiles of βGRP3 mRNA and protein

We expressed *M. sexta* βGRP3 in *E. coli* and purified the recombinant protein for production of a rabbit polyclonal antibody. We also expressed and purified V5-tagged βGRP3 and GFP from *Drosophila* S2 cells for functional study. Recombinant βGRP3 purified from S2 cells migrated close to 58 kDa, which is larger than the calculated mass of 48.9 kDa, probably due to the addition of V5-His tag and post-translational modifications (Fig. 28). Recombinant βGRP3 could be recognized by both monoclonal anti-V5 and polyclonal anti-βGRP3 antibodies (Fig. 28B and C).

The mRNA level of βGRP3 in different tissues of *M. sexta* naïve larvae was determined by real-time PCR, and the results showed that βGRP3 transcript was mainly expressed in fat body at a significantly higher level than in hemocytes, midgut and other tissues (Fig. 29A). Expression of βGRP3 mRNA in fat body, hemocytes and midgut was not significantly up-regulated in response to microbial challenges (Fig. 29B-D), which is similar to that of *M. sexta* βGRP1 (Ma and Kanost, 2000). We also determined βGRP3 protein concentration in hemolymph by Western blot. The results showed that βGRP3 protein concentration in the hemolymph of larvae injected with *E. coli, M. luteus* or *S. cerevisiae* did not change significantly from 0 to 12 h post-injection (Fig. 30A-C). To estimate the concentration of endogenous βGRP3 protein in hemolymph, Western blot was performed with increasing volumes of hemolymph from naïve and *E. coli*-injected larvae along with increasing amounts of purified recombinant βGRP3: endogenous βGRP3 protein was estimated at ~ 260 μg/ml in the hemolymph of naïve larvae and ~255 μg/ml in the hemolymph of *E. coli*-injected larvae (Fig. 30D and E). These results
suggest that both βGRP3 transcript and protein levels were not induced by immune challenges.
V5-tagged recombinant GFP and βGRP3 proteins were purified from *Drosophila S2* cells and analyzed by SDS-PAGE (1.5 μg each protein) (A) and Western blot (0.2 μg each protein) using monoclonal anti-V5 (B) or polyclonal anti-βGRP3 (C) antibody. (A-C): GFP (lane 1) and βGRP3 (lane 2).
(A) Tissue distribution of $\beta$GRP3 mRNA in hemocytes, fat body, midgut, epidermis and testis of the 5th instar *M. sexta* naïve larvae was determined by real-time PCR and normalized to *rpS3* gene. (B-D): Expression of $\beta$GRP3 mRNA in fat body (B), hemocytes (C), and midgut (D) after immune challenge Expression of $\beta$GRP3 mRNA in hemocytes, fat body and midgut of the 5th instar *M. sexta* naïve larvae or larvae injected with ddH$_2$O, *E. coli*, *S. marcescens*, *S. aureus*, *B. subtilis*, or *S. cerevisiae* at 24 h post-injection was determined by real-time PCR. The bars represent the mean of three individual measurements ±SEM. Relative expression of $\beta$GRP3 mRNA in naïve larval hemocytes (A) or in tissues of naïve larvae (B-D) was set as 1. Comparing expression of $\beta$GRP3 mRNA in different tissues (A) or after microbial injections, identical letters among tissues or treatments indicate not significant difference (p>0.05), while different letters indicate significant difference (p<0.05) determined by one way ANOVA followed by a Tukey’s multiple comparison test.
Figure 30. Expression of βGRP3 protein in hemolymph in response to immune challenge

The 5th instar *M. sexta* naïve larvae were injected with *E. coli* (A), *M. luteus* (B), or *S. cerevisiae* (C), and hemolymph samples were collected at 0, 2, 4, 6, 8, 10 and 12 h post-injection as described in the Materials and Methods. Cell-free mixed plasma samples (0.5 μl each) were subjected to Western blot analysis using rabbit anti-βGRP3 antibody. (D) and (E): Concentration of endogenous βGRP3 in hemolymph Increasing amounts of purified recombinant βGRP3 (lanes 1-4) or increasing volumes of mixed plasma samples (lanes 5-8) from naïve larvae (D) or *E. coli*-injected larvae (E) were analyzed by Western blot using rabbit anti-βGRP3 antibody. Lanes 1-4: 10, 25, 50 and 75 ng purified recombinant βGRP3, respectively; lanes 5-8: 0.1, 0.2, 0.3 and 0.4 μl cell-free mixed plasma samples, respectively.
Binding of βGRP3 to microbial cell wall components and microorganisms

To determine whether recombinant βGRP3 can bind to different microbial cell wall components, plate ELISA assay was performed. Compared to the control GFP protein, βGRP3 could bind to LPS-K12 and PG-K12 from *E. coli* K12, LTA-BS and PG-BS from *B. subtilis*, mannan and laminarin from fungi, and LTA-SA from *S. aureus* (Fig. 31), but did not bind to PG-SA from *S. aureus* (Fig. 31D). To confirm binding of βGRP3 to microbial components is specific, a competitive binding assay was performed. Binding of βGRP3 to laminarin was well competed by free laminarin, mannan, LPS-K12, LTA-SA and LTA-BS, but only slightly competed by PG-K12 and PG-BS (Fig. 31I), indicating that βGRP3 has broad binding spectra to microbial cell wall components with different affinities.

To further confirm broad binding spectra of βGRP3, a direct binding of βGRP3 to different microorganisms was performed. We first tested binding of GFP and βGRP3 to *E. coli* and found that GFP did not bind to *E. coli* and remained in the unbound fraction (Fig. 32A), but βGRP3 bound tightly to *E. coli* since only a small fraction of βGRP3 was eluted by 7% SDS and more βGRP3 remained in the *E. coli* lysates (Fig. 32D). We then tested binding of GFP and βGRP3 to several Gram-negative and Gram-positive bacteria as well as yeast. GFP did not bind to any of the microorganisms tested (Fig. 32B and C), but βGRP3 bound to all eight microorganisms tested (Fig. 32E and F). More βGRP3 protein bound to *E. coli, P. aeruginosa, S. aureus* and *B. subtilis* than to the other four microorganisms (Fig. 32E and F). Binding of βGRP3 to *S. marcescens, S. aureus, B. subtilis, M. luteus*, and *S. cerevisiae* seems to be stronger since little or no βGRP3 was eluted by 7% SDS and more βGRP3 remained in microbial lysates (comparing Fig. 32E
and F). These results suggest that βGRP3 is able to bind to multiple types of PAMPs on various microorganisms.
Figure 31. Binding of βGRP3 to microbial cell wall components

Increasing concentrations of recombinant βGRP3 or GFP purified from *Drosophila* S2 cells were added to 96-well plates coated with LPS-K12 (A), PG-K12 (B), LTA-SA (C), PG-SA (D), LTA-BS (E), PG-BS (F), mannan (G) or laminarin (H), and binding of recombinant proteins to microbial components was determined by plate ELISA using anti-V5 antibody. Recombinant βGRP3 (2 μg/ml) was pre-incubated with increasing amounts of free microbial components and then added to the laminarin-coated 96-well plates (I), and binding of βGRP3 to laminarin was determined by plate ELISA assay. The figures showed total binding of recombinant proteins to microbial components. Each point represents the mean of three individual measurements ± SEM, and the lines in (A-H) represent nonlinear regression calculation of one-site binding curves for βGRP3 (solid line) and GFP (dotted line).
Figure 32. Binding of βGRP3 to microorganisms

Microorganisms were incubated with recombinant GFP (A-C) or βGRP3 (D-F) at room temperature for 1 h, washed with TBS, and bound proteins were eluted with 7% SDS. The remaining microorganisms were lyzed in lysis buffer. Recombinant proteins in each fraction were detected by Western blot using anti-V5 antibody. (A) and (D): Binding of recombinant GFP (A) and βGRP3 (D) to E. coli. Lane 1: purified recombinant proteins (0.2 μg each); lane 2: unbound proteins; lane 3: the fourth TBS wash; lane 4: 7% SDS elution; lane 5: E. coli lysates. In lanes 2-5, equivalent volumes to lane 1 were loaded in each lane. (B, C, E and F): Binding of GFP and βGRP3 to various microorganisms. Only SDS elution fractions (B and E) and microbial lysates after SDS eluting (C and F) were analyzed by Western blot using anti-V5 antibody. Ec: E. coli, Pa: P. aeruginosa, Sm: S. marcescens, Sa: S. aureus, Bs: B. subtilis, Ml: M. luteus, Bc: B. cereus, Sc: S. cerevisiae.
Agglutination of *E. coli* and *B. cereus* by βGRP3

To test whether βGRP3 also has agglutinating activity against microorganisms, purified recombinant βGRP3 or GFP was incubated with eight microorganisms, and agglutination of microbial cells was observed under microscope. GFP did not agglutinate any of the eight microorganisms tested (Fig. 33A-C and data not shown), but βGRP3 could agglutinate *E. coli* and *B. cereus* (Fig. 33D and E), but not *S. cerevisiae* (Fig. 33F) and the other five bacteria (data not shown), and agglutination of *E. coli* and *B. cereus* was calcium-dependent (Fig. 33G and H). In addition, *E. coli* cells formed aggregates in the presence of βGRP3 (Fig. 33D), while *B. cereus* cells were cross-linked by βGRP3 to form large networks (Fig. 33E).
Figure 33. Agglutination of microorganisms by βGRP3

Purified recombinant GFP or βGRP3 was incubated with *E. coli*, *B. cereus* or *S. cerevisiae* in TBS containing 2 mM CaCl$_2$ and 1 mM MgCl$_2$ (A-F) or 5 mM EDTA (G and H) at room temperature for 60 min, samples of microbial cells were examined by microscopy. Bar = 20 μm.
Antibacterial activity of βGRP3 against *B. cereus*

Since βGRP3 could bind to microorganisms (Fig. 32) and agglutinate *E. coli* and *B. cereus* (Fig. 33), it may have antimicrobial activity and an inhibition zone assay was performed. Ampicillin and buffer were used as positive and negative controls. Ampicillin showed high activity against *E. coli* and relatively high activity against *B. subtilis*, but recombinant βGRP3 did not show activity against *E. coli* or *B. subtilis* (Fig. 34A and B). Interestingly, ampicillin was not active against *B. cereus*, but βGRP3 showed strong activity against *B. cereus* in a dose-dependent manner (Fig. 34C). Both ampicillin and βGRP3 did not have activity against *S. cerevisiae* (Fig. 34D). To test whether binding of microbial components to βGRP3 may inhibit its activity against *B. cereus*, bacterial growth curves were determined in presence of βGRP3 alone or βGRP3 mixed with microbial components. The results showed that βGRP3 alone almost completely inhibited *B. cereus* growth (OD600 = 0.104 compared to OD600 = 1.013 in the control group (water) at 20 h after incubation). LPS-K12, LTA-BS, LTA-SA, laminarin, mannann and PG-BS could all completely block the antibacterial activity of βGRP3 against *B. cereus* (OD600 =0.998 – 1.087 compared to OD600 = 1.013 in the control group) (Fig. 34E). PG-K12 also significantly blocked the activity of βGRP3 against *B. cereus* growth (OD600 = 0.805 vs. OD600 =0.104) (p<0.001), but *B. cereus* growth was still significantly inhibited compared to the control group (OD600 = 0.805 vs. OD600 = 1.013) (p<0.001) (Fig. 34E).
Figure 34. Antibacterial activity of βGRP3 against *B. cereus*

Antibacterial activity of recombinant βGRP3 was determined by inhibition zone assay (A-D) and broth micro-dilution assay (E) as described in the Materials and Methods. Buffer and different amounts of ampicillin (Amp) were used as controls in the inhibition zone assay, the plates were incubated at 37°C overnight, and the diameters of inhibition zones were measured and recorded (A-D). Diluted *B. cereus* culture was incubated with recombinant βGRP3 alone (2 μg/ml) or βGRP3 (2 μg/ml) mixed with different free microbial components (20 μg/ml each) in 96-well plates at 37°C, and OD600 was recorded every hour up to 22 h. The points represent the mean of four individual measurements ± SEM. Asterisks indicate significant difference in OD600 values at 20 h among the control, βGRP3, and βGRP3+PG-K12 groups (p<0.001) determined by one way ANOVA followed by a Tukey’s multiple comparison test.
Discussion

*M. sexta* βGRP3, a novel β-1, 3-glucan recognition proteins

Invertebrates lack the adaptive immune system and rely on pattern recognition receptors (PRRs) as surveillance molecules to detect microorganisms and mount immune responses. These surveillance molecules circulate in the animal body to recognize pathogen-associated molecular patterns (PAMPs) (Mogensen, 2009; Pal and Wu, 2009; Sukhithasri et al., 2013). β-1, 3-glucan recognition proteins (BGRPs/βGRPs) are a family of proteins containing a glycosyl hydrolase like domain, but lack key residues for glucanase activity (Fig. 27A). In this study, we have cloned a new member (βGRP3) of the *M. sexta* βGRP family, and investigated functions of βGRP3 in innate immunity.

*M. sexta* βGRP1 and βGRP2 transcripts are mainly expressed in fat body. βGRP1 mRNA is expressed in both feeding and wandering stages of larvae, while βGRP2 transcript is only expressed in the wandering stage larvae (Jiang et al., 2004; Ma and Kanost, 2000). We found that βGRP3 mRNA was predominantly expressed in fat body of naïve larvae (Fig. 29A). Expression of βGRP3 transcript in fat body, hemocytes and midgut as well as βGRP3 protein in hemolymph was not induced by different immune challenges (Figs. 29B-D and 30), which is similar to that of *M. sexta* βGRP1 but different from that of βGRP2 (Jiang et al., 2004; Ma and Kanost, 2000). *M. sexta* MBP mRNA is mainly expressed in fat body and is significantly induced by immune challenge, and MBP protein concentration in hemolymph also increases after immune stimulations (Wang et al., 2011). These results indicate that the βGRP family proteins are differentially regulated by development and immune challenge.
**M. sexta βGRP3 had a broad binding ability**

Recombinant βGRP3 had a broad binding spectrum to microbial components (Fig. 31) and microorganisms (Fig. 32). Broad binding spectra of insect PRRs, including BGRPs/βGRPs, GNBP, PGRPs and C-type lectins, have been reported. For example, two beetle 1,3-β-D-glucan-binding proteins can recognize β-1, 3-D-glucan and peptidoglycan (Lee et al., 2004). *M. sexta* C-type lectins can bind to several different microbial components, including LPS, LTA and PG (Yu et al., 1999; Yu and Kanost, 2000; Yu et al., 2006; Yu et al., 2005). *Drosophila* GNBP-1 can bind to LPS and β-1, 3-glucan (Kim et al., 2000). *M. sexta* βGRP1 can bind to *E. coli*, *M. luteus* and *S. cerevisiae* (Ma and Kanost, 2000), βGRP2 can bind to laminarin and LTA but not LPS (Jiang et al., 2004), and MBP can bind to LPS, LTA and DAP-type PG (Wang et al., 2011). Thus, the BGRP/βGRP family proteins may have broad binding spectra for various pathogens but differ in binding affinities. Less *M. sexta* βGRP3 bound to DAP-type PGs than to LPS, LTA and laminarin (Fig. 31), which is consistent with the results of competitive binding assay that binding of βGRP3 to laminarin was only slightly competed by PGs but well competed by LPS, LTA and laminarin (Fig. 31).

**M. sexta** βGRP3 agglutinates *E. coli* and *B. cereus*, and inhibits the growth of *B. cereus*

The βGRP family proteins are mainly involved in activation of the proPO system (Fabrick et al., 2003; Jiang et al., 2004; Ma and Kanost, 2000; Ochiai and Ashida, 2000), and some βGRPs can also aggregate microorganisms (Fabrick et al., 2003; Jiang et al., 2004; Beschin et al., 1998). Like most βGRP proteins, *M. sexta* βGRP3 could activate proPO in plasma (Zhong X and Yu X, unpublished results), more importantly, it possess
some unique properties that have not been reported for the BGRP/βGRP family proteins: 1) calcium-dependent agglutination of *E. coli* and *B. cereus* (Fig. 33), which is similar to that of C-type lectins (Yu et al., 1999; Yu and Kanost, 2000; Yu et al., 2006; Yu et al., 2005), and 2) direct antibacterial activity against *B. cereus* that is resistant to ampicillin (Fig. 34C). Compared to *M. sexta* βGRP1, βGRP2 and MBP, βGRP3 is ~40-45 residues shorter at the N-terminus and lacks the small glucan binding domain. BlastP search showed that the shorter N-terminal domain (~57 residues) of *M. sexta* βGRP3 does not have a hit in the NCBI database (data not shown), indicating that the N-terminus of βGRP3 is not a truncated form of the small glucan binding domain but a new sequence that may have unknown functions. We speculate that the shorter N-terminal domain of βGRP3 may be involved in protein-protein interactions for formation of dimers or oligomers, which are required for agglutination of bacterial cells, since βGRP3 only contains a glucanase domain. Such a protein-protein interaction may also be calcium-dependent. The coelomic cytolytic factor 1 (CCF-1) of earthworm *Eisenia fetida*, which is a member of the βGRP family, also lacks the small N-terminal glucan binding domain, but it can still agglutinate Gram-negative bacteria, indicating that the N-terminal glucan binding domain is not absolutely required for agglutination (Beschin et al., 1998). Antibacterial activity of βGRP3 is probably due to bacteriostatic effect but not bactericidal effect as the activity of βGRP3 was completely blocked by laminarin, mannan, LPS, LTA and PG-BS, and significantly blocked by PG-K12 (Fig. 34E). Future work is to understand how βGRP3 agglutinates bacterial cells and exerts antibacterial activity.
CHAPTER 6

MANDUCA SEXTA GLOVERIN BINDS MICROBIAL COMPONENTS AND IS ACTIVE AGAINST BACTERIA AND FUNGI

Any questions about whether a particular use is authorized and any requests for permission to publish, reproduce, distribute, display or make derivative works from any Elsevier published works, images or other material should be directed to the Elsevier Rights Department.

© 2012 Elsevier Ltd. All rights reserved.
Abstract

Gloverin is a glycine-rich and heat stable antimicrobial protein with activity mainly against *Escherichia coli*. However, *Spodoptera exigua* Gloverin is active against a Gram-positive bacterium but inactive against *E. coli*. In this study, we investigated expression profile, binding ability and antimicrobial activity of *Manduca sexta* Gloverin (*MsGlv*). *MgGlv* transcript was detected in several tissues of naïve larvae with higher levels in the midgut and testis. Expression of *MgGlv* mRNA in larvae was up-regulated by the active form Spätzle-C108, *E. coli* and *Staphylococcus aureus* peptidoglycans (PGs), and the activation was blocked by pre-injection of antibody to *M. sexta* Toll receptor, suggesting that *MgGlv* expression is regulated by the Toll-Spätzle pathway. Recombinant *MsGlv* bound to lipopolysaccharide (LPS) to the O-specific antigen and outer core carbohydrate moieties, Gram-positive lipoteichoic acid (LTA) and PG, and laminarin, but did not bind to *E. coli* PG or mannan. *MsGlv* was active against *Bacillus cereus*, *Saccharomyces cerevisiae* and *Cryptococcus neoformans*, but had low activity against *E. coli* and *S. aureus*. Our results suggest that Gloverins may have various activities against different microorganisms.

Introduction

Insects rely on the innate immune system to fight against microbial infections. The insect innate immune system shares similarities to the innate immune system of vertebrates, and is also composed of humoral and cellular responses (Ferrandon et al., 2007; Lemaitre and Hoffmann, 2007; Muller et al., 2008). Insect cellular immune responses include hemocyte (blood cell) mediated nodule formation, phagocytosis and
encapsulation, while synthesis of antimicrobial peptides/proteins (AMPs) and activation of proteinase cascades are major components of humoral immune responses (Bulet and Stocklin, 2005; Bulet et al., 2004; Chae et al., 2012; Ferrandon et al., 2007; Jiang et al., 2010; Lemaitre and Hoffmann, 2007). Insects can synthesize a variety of AMPs with activities against bacteria, fungi, viruses and some parasites (Bulet and Stocklin, 2005; Bulet et al., 2004). Some AMPs, such as Cecropin, Attacin and Defensin, are common and present in most insect species, while some other AMPs like Moricin, Gloverin and Lebocin have been identified only in lepidopteran insects so far.

Gloverin was first isolated from the hemolymph of immunized Hyalophora gloveri pupae (Axen et al., 1997). H. gloveri Gloverin (HgGlv) is a heat-stable, glycine-rich and basic antibacterial protein (~14kDa) that is active against Escherichia coli (Axen et al., 1997). Homologous Gloverin proteins or cDNAs have also been isolated in other lepidopteran species, including Helicoverpa armigera (Mackintosh et al., 1998), Trichoplusia ni (Lundstrom et al., 2002; Seitz et al., 2003), Galleria mellonella (Seitz et al., 2003), Bombyx mori (Cheng et al., 2006; Kaneko et al., 2007; Kawaoka et al., 2008; Mrinal and Nagaraju, 2008), Diatraea saccharialis (Silva et al., 2010), Plutella xylostella (Etebari et al., 2011), Spodoptera exigua (Hwang and Kim, 2011), and Manduca sexta (Abdel-latif and Hilker, 2008; Zhu et al., 2003). H. armigera Gloverin is active against Gram-negative bacteria but inactive against Gram-positive bacteria and the fungus Candida albicans (Mackintosh et al., 1998). Recombinant T. ni pro-Gloverin (containing the pro-segment) is active against E. coli, and its activity is comparable to that of mature HgGlv protein (Lundstrom et al., 2002). B. mori has four Gloverin genes, and all four recombinant Gloverins are active against E. coli (Kawaoka et al., 2008; Mrinal and
Nagaraju, 2008). Gloverins have been reported to be active almost exclusively against Gram-negative bacteria, however, recombinant *S. exigua* Gloverin (SeGlv) is active against a Gram-positive bacterium (*Flavobacterium* sp.) but inactive against *E. coli*, and knockdown expression of Seglv by RNA interference (RNAi) increases susceptibility of *S. exigua* larvae to Gram-positive *Bacillus thuringiensis* infection (Hwang and Kim, 2011). A recent report also shows that *T.ni* Gloverins have anti-viral activity (Moreno-Habel et al., 2012).

*HgGlv* has a random-coil conformation in aqueous solution, but can convert to a more α-helical structure in a hydrophobic membrane-like environment (Axen et al., 1997). *HgGlv* can inhibit synthesis of *E. coli* outer membrane proteins to increase the permeability of bacterial outer membrane (Axen et al., 1997). Pre-incubation of Rd mutant LPS with *HgGlv* can inhibit its activity against *E. coli*, suggesting that Gloverin may interact with the lipid A moiety of LPS, since lipid A is negatively charged, which may interact with basic Gloverin through electrostatic interaction (Axen et al., 1997). But the isoelectric point (pI) of Gloverins from different insect species varies from slightly acidic to neutral (pI 5.5-7 for four *B. mori* Gloverins), basic (pI ~8.3) to highly basic (pI>9.3). In addition, direct binding of Gloverin to lipid A or LPS has not been demonstrated, and it is not known whether Gloverin can also bind to other microbial components such as bacterial lipoteichoic acid (LTA) and peptidoglycan (PG), or fungal β-1, 3-glucan and mannan. In this study, we investigate expression profile of *M. sexta gloverin* (Msglv), binding of recombinant MsGlv to microbial components and antimicrobial activity of MsGlv against Gram negative and Gram-positive bacteria as well as fungi.
Materials and Methods

Insect rearing and *Drosophila* S2 cell line

*M. sexta* eggs were purchased from Carolina Biological Supplies (Burlington, NC, USA). Larvae were reared on an artificial diet at 25°C (Dunn, 1983), and the fifth instar larvae were used for the experiments. *D. melanogaster* Schneider S2 cells were purchased from American Type Culture Collection (ATCC).

Microorganisms and microbial components

*Escherichia coli* strain XL1-blue was purchased from Strata gene (CA, USA). *Serratia marcescens* and *Bacillus thuringiensis* were from American Type Culture Collection (ATCC). *Staphylococcus aureus* and *Bacillus cereus* were kindly provided by Professor Brian Geisbrecht, *Saccharomyces cerevisiae* (strain BY4741) and *Cryptococcus neoformans* (KN99alpha) were provided by Professor Alexander Idnurm, *Bacillus subtilis* was provided by Professor Michael O’Connor, School of Biological Sciences at University of Missouri-Kansas City.

Smooth lipopolysaccharide (LPS) from *Salmonella enterica*, *S. marcescens*, *E. coli* 055:B5, *E. coli* 026:B6 and *E. coli* 0111:B4, rough mutants of LPS from *E. coli* EH100 (Ra mutant), *E. coli* J5 (Re mutant), *E. coli* F583 (Rd mutant) and *S. enterica* serotype minnesota Re 595 (Re mutant), lipid A monophosphoryl from *E. coli* F583 (Rd mutant) and lipid A diphosphoryl from *E. coli* F583 (Rd mutant), laminarin, mannan, and zymosan were from Sigma-Aldrich (MO, USA), TLR grade LPS and PG from *E. coli* K12 (LPS-K12 and PG-K12), peptidoglycan (PG) and lipoteichoic acid (LTA) from *B.
*subtilis* (LTA-BS and PG-BS) and *S. aureus* (LTA-SA and PG-SA) were from Invivogen (CA, USA).

**Sequence analysis**

Sequence similarity search was carried out using blast biological software ([http://www.ncbi.nlm.nih.gov/blastp](http://www.ncbi.nlm.nih.gov/blastp)). Multiple sequence alignments were performed using ClustalW ([http://www.ch.embnet.org/software/ClustalW.html](http://www.ch.embnet.org/software/ClustalW.html)). A phylogenetic tree of the mature gloverin proteins from some insect species was constructed by MEGA 5.05 software (Tamura et al., 2011). Signal peptide sequences were predicted with SignalP 3.0 Server ([http://www.cbs.dtu.dk/services/SignalP/](http://www.cbs.dtu.dk/services/SignalP/)) (Bendtsen et al., 2004). Figures were made with the GraphPad Prism software (GraphPad, San Diego, CA) with one representative set of data. Significance of difference was determined by an unpaired t-test or by one way ANOVA followed by a Tukey’s multiple comparison test using the same software (GraphPad, San Diego, CA).

**Tissue distribution and induced expression of *M. sexta gloverin***

To determine tissue distribution of *M. sexta gloverin* (*Msglv*), 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 from these tissues were extracted with TRIzol® Reagent (T9424, Sigma-Aldrich). cDNA was prepared from 1 μg total RNA in a 25 μl reaction using moloney murine leukemia virus (M-MLV) reverse
transcriptase (M1701, Promega) with an anchor-oligo(dT)$_{18}$ primer following the manufacturer’s instructions. To determine induced expression of *Msglv*, day 2 fifth instar larvae were injected with ddH$_2$O, heat-killed *E. coli* strain XL1-blue (5×10$^7$ cells/larva), *S. marcescens* (5×10$^7$ cells/larva), *S. aureus* (5×10$^7$ cells/larva), *B. subtilis* (5×10$^7$ cells/larva) or *C. neoformans* (10$^7$ cells/larva). Twenty-four hours after injection, hemocytes, fat body and midgut were collected separately. Total RNA and cDNA were prepared as described above. The real-time PCR was performed in 20 μl reactions containing 10 μl 2×SYBR® GreenER™ qPCR SuperMix Universal (No. 204141, Qiagen), 4 μl dH$_2$O, 4 μl diluted (1:50) cDNA, and 1 μl each reverse and forward diluted primer (10 pmol/μl). For *Msglv* gene, primers MsGlv-F (5’-CCC GCA ATA CGC TCA GAT A-3’) and MsGlv-R (5’-TGC TGG AAG AGA CCT TGG A-3’) were used, for the control ribosomal protein S3 (rps3) gene, primers RPS3-F (5’-GTT GCG AGG TGG TGG TTT C-3’) and RPS3-R (5’-CCG TTC TTG CCC TGT TGG TC-3’) were used. Real-time PCR program was 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$^{-ΔΔCT}$) and these experiments were repeated with 3 different biological samples.

**Activation of *M. sexta* gloverin by Spätzle and peptidoglycans**

Recombinant full-length *M. sexta*Spätzle-1A (*MsSpz*) and the C-terminal active domain of Spätzle-1A (*MsSpz-C108*) were expressed in *Drosophila* S2 cells and purified by affinity chromatography as described previously (Zhong et al., 2012). Day 1 fifth instar *M. sexta* larvae were injected with *MsSpz* (3 μg/larva), *MsSpz-C108* (1 μg/larva), *S.
*aureus* PG-SA or *E. coli* PG-K12 (1 μg/larva) (Zhong et al., 2012). Twenty hours later, fat body, hemocytes and midgut samples were collected for preparation of total RNA and cDNA, and expression of *Msglv* in these tissues was determined by quantitative RT-PCR as described above.

For antibody blocking assay, day 1 fifth instar *M. sexta* naïve larvae were pre-injected with purified IgG to the ecto-domain of *M. sexta* Toll (Toll Ab, 5 μg/larva) or IgG from pre-bleed rabbit serum (Control Ab, 5 μg/larva). One hour later, these larvae were injected with water, *Ms*Spz (3 μg/larva), *Ms*Spz-C108 (1 μg/larva), *S. aureus* PG-SA (1 μg/larva), *E. coli* PG-K12 (1 μg/larva), or without second injection (control) as described previously (Zhong et al., 2012). Twenty hours later, fat body, hemocyte and midgut samples were collected for quantitative RT-PCR analysis. *M. sexta* ribosomal protein S3 (*rpS3*) gene was used as an internal standard to normalize the amount of RNA template. The expression levels of *Msglv* were calculated by the $2^{-\Delta\Delta CT}$ methods as described above.

**Expression and purification of *M. sexta* Gloverin in bacteria and preparation of polyclonal rabbit antiserum**

RT-PCR was performed to obtain a cDNA encoding *M. sexta* pro-Gloverin (pro-*MsGlv*, residues 23-177) using primers pro-*MsGlv*-F (5’-GGA CCA TGG CCC AAT ACG CTC AGA T-3’) and pro-*MsGlv*-R (5’-CCA CTC GAG CCA TCT ATG CTG GAA GAG ACC-3’). The PCR fragment was purified by agarose gel electrophoresis, digested with *NcoI* and *XhoI* enzymes, ligated into the *NcoI /XhoI*-digested expression vector pET-32a (+) (Novagen), and then transformed into competent *E. coli* BL21 (DE3)
cells. A single positive bacterial colony, which was confirmed by restriction enzyme digestion and sequencing, was inoculated into LB medium containing Ampicillin (100 μg/mL) and grown overnight. The overnight culture was diluted 1:100 in LB medium and grown at 37°C to OD$_{600}$=0.8 and then isopropyl-D-thiogalactopyranoside (IPTG) (0.5 mM) was added to induce protein expression. After 6 h incubation at 28°C, bacterial cells were harvested by centrifugation and lysed with the lysis solution (50 mM Tris-HCl, pH 8.0, 50 mM NaCl, 0.5% Triton X-100, 2 mg/mL lysozyme). Protein purification was performed with a His-Bind® Buffer Kit (Novagen). The purified protein was further separated on 15% SDS-PAGE and the gel slice containing recombinant pro-MsGlv was used as an antigen to produce rabbit polyclonal antiserum at Cocalico Biologicals, Inc (Pennsylvania, USA).

Establishment of stable S2 cell lines expressing *M. sexta* mature Gloverin (MsGlv) and green fluorescent protein (GFP)

The cDNA encoding mature MsGlv (residues 46-177) was amplified by PCR using primers MsGlvM-F (5'-GCG AGA TCT GAC GTG ACC TGG GAC AAG CAA G-3') and MsGlvMF-R (5'-GGA CTC GAG CCA TCT ATG CTG GAA GAG ACC T-3'). The forward primer contained a BglII site at the 5’ end and the reverse primer contained an XhoI site at the 3’ end. PCR product was recovered by agarose gel electrophoresis-Wizard® SV Gel and PCR Clean-Up System (A9285, Promega), subcloned into T-Easy vector (A1360, Promega), and recombinant plasmid was isolated. After digested with BglII/XhoI, cDNA fragment was recovered and inserted into the same restriction sites of the expression vector pMT/BiP/V5-His (V413020, Invitrogen) using
T4 DNA ligase (M0202L, NEB). Green Fluorescent Protein (GFP) in the expression vector pMT/Bip/V5-His/GFP (V413020, Invitrogen) was used as a control. Recombinant expression vectors were purified using PureYield™ Plasmid MiniprepSystem (A1222, Promega) according to the manufacturer’s instruction and used to generate stable S2 cell lines after the insert sequences were confirmed by DNA sequencing. *Drosophila* 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) and 1% penicillin-streptomycin solution (G6784, Sigma-Aldrich). For transfection assay, S2 cells (in 6-well plates) were seeded overnight in serum-free medium (SH30278.01, Hyclone) and GenCarrier-1™ transfection reagent (#31-00110, Epoch Biolabs) was used for transient transfection based on the manufacturer’s instructions. Culture dishes or plates were prepared to 70% confluence prior to transfection. DES®–Inducible/Secreted Kit with pCoBlast (K5130-01, Invitrogen) was used to construct stable S2 cell lines as described previously (Zhong et al., 2012). For Western blot analysis, copper sulfate (250 μM) was added to stable S2 cells in a 6-well plate to induce protein expression for 48 h as described previously (Zhong et al., 2012).

**Purification of recombinant MsGlv and GFP from S2 cells**

To purify recombinant MsGlv and GFP, copper sulfate (final concentration of 250 μM) was added to stable S2 cells expressing MsGlv or GFP in 75-cm² flasks to induce protein expression. Cell culture medium was collected starting at 24 h after protein expression for 10 days by collecting culture medium every day and re-suspending the cells with fresh medium. Cell culture medium was combined, cell debris was removed by centrifugation, and cell-free medium was incubated overnight at 4°C with 500 μl of Anti-
V5-agarose Affinity Gel (A7345, Sigma-Aldrich) equilibrated with initial buffer (50 mM Tris-HCl, 150 mM NaCl, pH 7.4). Purification of MsGlv and GFP by Anti-V5-agarose Affinity Gel was performed essentially the same as described previously for purification of *M. sexta* Spätzle proteins (Zhong et al., 2012). Recombinant MsGlv and GFP were sequentially eluted from the column with 1 ml aliquots of the elution buffer (0.1 M glycine-HCl, pH 3.5, 1% Triton X-100) into vials containing 100 μl of 1M Tris-base, pH 8.0. Fractions were analyzed by 15% SDS-PAGE, and those containing recombinant MsGlv or GFP were combined and desalted using D-salt™ Excellulose™ GF-5 desalting column (#1851850, Pierce) pre-equilibrated with dH2O. Recombinant MsGlv or GFP was eluted with water, and fractions containing MsGlv or GFP were pooled and concentrated for the following binding and activity assays.

**Binding of MsGlv to microbial components**

To test binding of MsGlv to microbial cell wall components, plate ELISA assays were performed using different microbial components (LPS, LTA, PG, lipid A, laminarin, zymosan and mannan). Briefly, wells of a flat bottom 96-well plate (Costar, Fisher) were coated with different microbial components (2 μg/well) as described previously (Yu and Kanost, 2000; Yu et al., 2005). The plates were placed overnight at room temperature until the water evaporated completely, heated to 60°C for 30 min, and then blocked with 200 μl/well of 1 mg/ml BSA in Tris buffer (TB) (50 mM Tris-HCl, 50 mM NaCl, pH 8.0) for 2 h at 37°C. Then, plates were rinsed four times with 200 μl/well of TB, and increasing concentrations or 120 nM of purified MsGlv or GFP (a control protein) diluted in TB containing 0.1 mg/ml BSA were added to the coated plates (50 μl/well), and
binding was allowed to occur for 3 h at room temperature. The plates were rinsed four times with 200 μl/well of TB, and monoclonal anti-polyhistidine antibody (Sigma-Aldrich, USA) (1:2,000 in TB containing 0.1 mg/ml BSA) was added (100 μl/well) and incubated overnight at 4°C. The plates were rinsed four times with TB (200 μl/well), and alkaline phosphatase-conjugated goat anti-mouse-IgG (Sigma-Aldrich, USA) (1:3,000 in TB containing 0.1 mg/ml BSA) was added (100 μl/well) and incubated for 2 h at 37°C. The plates were rinsed, p-nitro-phenyl phosphate (1 mg/ml in 10 mM diethanolamine, 0.5 mM MgCl2) was added (50 μl/well), and absorbance at 405 nm of each well was determined every minute for 30 min using a microtiter plate reader (Bio-Tek Instrument, Inc.).

**Antimicrobial activity assays**

Antimicrobial activity of purified MsGlv was tested against six bacterial strains (*B. cereus*, *E. coli* DH5α, *S. marcescens*, *B. subtilis*, *B. thuringiensis* and *S. aureus*) and two fungal strains *S. cerevisiae* and *C. neoformans*. A broth micro dilution assay was used to generate growth curves as described previously (Rao et al., 2012). Briefly, overnight bacterial cultures were sub-cultured in LB medium and fungal cultures were sub-cultured in YPD medium (1% yeast extract, 2% peptone, 2% dextrose) until mid-log phase. The bacterial and fungal cultures were centrifuged at 1,000 g for 10 min at 4°C and washed once with 10 mM Tris-HCl, pH7.5, 0.1 mM EDTA. The bacterial and fungal cells were diluted to OD600=10⁻⁵ in LB and YPD media, respectively, and the diluted cell cultures (75 μl) were mixed with purified MsGlv (25 μl of ~85 μg/ml) (final concentration of 21 μg/ml or 1.5 μM) or water (Control) in 96-well plates. Bacteria were cultured at 37°C
with 220 rpm shaking, while fungi were cultured at 30°C with 220 rpm shaking. OD600 was measured every hour by Powerwave XS plate reader (BioTek, VT, USA). Bacterial growth curves were generated using the Graphpad Prism version 4.0 for Windows (GraphPad Software, CA, USA).

**Results**

**Sequence analysis of *M. sexta* Gloverin**

*M. sexta* Gloverin (MsGlv) (Genbank accession number: CAL25129) is 177-residues long, with a predicted 19-residue signal peptide, 26-residue pro-segment, and 132-residue mature protein (Fig. 35). Mature MsGlv contains 28 glycines (21.2%) with a theoretical mass of 13986 Da and pI of 9.35. MsGlv is most similar to Gloverins from *Antheraea mylitta* (74.2% identity in the mature proteins), *Antheraea pernyi* (73.5%), *B. mori* (gloverin-4, 71.2%), *H. gloveri* (69.7%), and also has 48-69% identities to other mature Gloverins (Fig. 35).

Among the four *B. mori* gloverin genes, Bmglv1 is the ancestral gene, whereas Bmglv2-4 are derived from duplication (Mrinal and Nagaraju, 2008). Phylogenetic analysis of mature Gloverin protein sequences showed that *A. mylitta*, *A. pernyi* and *H. gloveri* Gloverins clustered in one group, BmGlv2-4 clustered in one group, and *H. virescens*, *S. exigua* and *T. ni* (Glv-1) Gloverins clustered in one group (Fig. 36A). These Gloverins along with PxGlv and TnGlv2 may come from the same ancestral gene. However, MsGlv and BmGlv1 did not cluster with any of the three groups, thus MsGlv may be an ancestral gene in *M. sexta*.
Expression and purification of recombinant *MsGlv*

Pro-*MsGlv* (residues 23–177) was expressed in bacteria, purified by nickel affinity column (Fig. 36B, lane 1), and used as an antigen to generate polyclonal antibody in a rabbit. Mature *MsGlv* (residues 46–177) was expressed in *Drosophila* S2 cells and purified by affinity chromatography (Fig. 36B, lane 2) for binding and activity assays (see below). Polyclonal anti-pro-*MsGlv* antibody could recognize recombinant *MsGlv* (Fig. 36C, lane 1) and natural *MsGlv* in the cell-free hemolymph of larvae immunized with *E. coli* (Fig. 36C, lane 3). Western blot analysis showed that *MsGlv* protein was not detected in the cell-free hemolymph of naïve larvae (Fig. 36D, lane 1), but was present at relatively high levels in the hemolymph of larvae immunized with *S. aureus*, *E. coli* and *S. cerevisiae* (Fig. 36D, lanes 2–4).
Figure 35. Multiple sequence alignment of Gloverin proteins from some lepidopteran species.

Gloverin protein sequences from *M. sexta* (CAL25129), *Antheraea mylitta* (ABG72699), *Antheraea pernyi* (ACB45565), *B. mori* (NP_001036930, NP_001037683, NP_001093312 and NP_001037684), *Heliothis virescens* (ACR78446), *Plutella xylostella* (ACM69342), *Spodoptera exigua* (ADL27731), *Trichoplusia ni* (ABV68856 and AAG44367), and *Hyalophora gloveri* (Axen et al., 1997) were aligned by ClustalW and residues conserved in all 13 proteins are indicated by asterisks. Predicted signal peptide sequences are underlined, and identity between *M. sexta* Gloverin and a Gloverin from other species is indicated in the parenthesis. The box indicates a conserved endopeptidase cleavage site and the arrow indicates the beginning of mature Gloverins.
Figure 36. Phylogenetic analysis of mature Gloverin sequences and expression of *M. sexta* Gloverin.

Thirteen mature gloverin protein sequences from some *lepidopteran* species from Fig. 35 were used to construct the NJ tree by the MEGA5.05 software. Numbers above the nodes indicate percent support from bootstrap analysis. (A). Recombinant *M. sexta* pro-Gloverin (pro-*MsGlv*) and mature Gloverin (*MsGlv*) were purified from bacteria and *Drosophila* S2 cells, respectively, and analyzed by SDS–PAGE (B). Lane 1: pro-*MsGlv* (2 μg) and lane 2: *MsGlv* (0.5 μg). Western blot analysis of recombinant and hemolymph *MsGlv* (C). Purified recombinant *MsGlv* (lane 1, 0.2 μg), cell-free hemolymph (1 μl each) from naïve larvae (lane 2) and larvae immunized with *E. coli* (lane 3) were analyzed by Western blot using polyclonal rabbit antiserum against pro-*MsGlv*. Induced expression of *MsGlv* in hemolymph by Western blot (D). Cell-free hemolymph (1 μl each) from naïve larvae (lane 1) and larvae immunized with *S. aureus* (lane 2), *E. coli* (lane 3) and *S. cerevisiae* (lane 4) were analyzed by Western blot using polyclonal rabbit antiserum against pro-*MsGlv*. 
Tissue distribution and induced expression of *M. sexta gloverin*

Real-time PCR analysis showed that *Msglv* mRNA was detected in the epidermis, hemocytes, fat body, midgut and testis of *M. sexta* naïve larvae, with significantly higher expression levels in the midgut and testis (Fig. 37A). Expression of *Msglv* transcript was induced to a significantly higher level in the fat body by Gram-positive *S. aureus* than the other microorganisms tested (Gram-negative *E. coli* and *S. marcescens*, Gram-positive *B. subtilis*, and the fungus *C. neoformans*) (Fig. 37B, upper panel), and was induced to significantly higher levels in the hemocytes and midgut by Gram-negative *E. coli* than other microorganisms (Fig. 37C and D, upper panels), but *Msglv* mRNA was not induced by water injection (Fig. 37B–D, upper panels). Activation of *Msglv* gene by different microorganisms showed tissue-specific pattern, which is consistent with tissue-specific activation of some *Drosophila* AMP genes (Imler and Bulet, 2005).

*MsGlv* protein was almost not detected in fat body, hemocytes and midgut of naïve larvae (Fig. 37B–D, low panels, lanes 1) and was present at very low levels in these tissues when larvae were injected with water (Fig. 37B–D, low panels, lanes 2). *MsGlv* protein in fat body, hemocytes and midgut was induced by Gram-negative and Gram-positive bacteria as well as *C. neoformans* (Fig. 37B–D, low panels, lanes 3–7), a result consistent with the induction of *Msglv* mRNA by microorganisms in these tissues (Fig. 37B–D, upper panels). *MsGlv* protein in hemolymph was also induced after larvae were injected with *S. aureus*, *E. coli* and *S. cerevisiae* (Fig. 36D). Together, these results suggest that *M. sexta gloverin* gene expression can be induced by different microorganisms at both the transcriptional and protein levels.
In *B. mori*, *Bmglv1* gene is expressed in larval but not adult gonads, while *Bmglv* 2–4 genes are expressed in adult but not larval gonads (Mrinal and Nagaraju, 2008), suggesting that *Bmglv* genes may be developmentally regulated. Knockdown expression of *Bmglv*2 gene by RNAi in *B. mori* embryos reduces hatching (Mrinal and Nagaraju, 2008), and RNAi of *Seglv* gene in *S. exigua* larvae reduces pupation and prolongs larval period (Hwang and Kim, 2011). In the *M. sexta* eggs parasitized by *Trichogramma evanescens*, *Msglv* gene expression is suppressed at 2 days after parasitization, while expression of *like-moricin* (*l-mor*), *leureptin* and *attacin-2* genes do not change significantly (Abdel-latief and Hilker, 2008). We showed that *Msglv* gene was expressed at a higher level in the testis of *M. sexta* naïve larvae (Fig. 37A). Altogether, these results suggest that Gloverin may play a role in development in addition to being an AMP.
Figure 37. Tissue distribution and expression of *Msglv* in *M. sexta* larvae after microbial infection.

Day 2 fifth instar *M. sexta* naïve larvae were dissected, epidermis, hemocytes, fat body, midgut and testis were collected for preparation of total RNAs. Expression of *Msglv* mRNA in these tissues was determined by quantitative real-time PCR (A). Total RNAs were also prepared from fat body, hemocytes and midgut of fifth instar larvae immunized with different microorganisms at 24 h post-injection, and expression of *Msglv* mRNA was also determined by quantitative real-time PCR (B–D, upper panels). *M. sexta* ribosomal protein S3 (*rpS3*) gene was used as an internal control. The bars represent the mean of three individual measurements ± SEM. Relative expression of *Msglv* mRNA in epidermis (A), fat body (B), hemocytes (C) or midgut (D) of naïve larvae was set as 1. Comparing expression of *Msglv* mRNA in different tissues (A) or after microbial injections, identical letters among tissues or treatments indicate not significant difference (p > 0.05), while different letters indicate significant difference (p < 0.05) determined by one way ANOVA followed by a Tukey’s multiple comparison test. Protein extracts from fat body, hemocytes and midgut (B-D, lower panels) of naïve larvae (lane 1) and larvae injected with water (lane 2), *E. coli* (lane 3), *S. marcescens* (lane 4), *S. aureus* (lane 5), *B. subtilis* (lane 6) and *C. neoformans* (lane 7) (60 μg total protein per lane), and cell-free hemolymph from larvae immunized with *S. aureus* (lane 8, 1 μg per lane) were also analyzed by SDS–PAGE, and *MsGlv* was detected by Western blot using polyclonal rabbit antiserum against pro-*MsGlv*. 
Expression of *M. sexta gloverin* is regulated by the Toll pathway

In *Drosophila*, expression of AMP genes is regulated by the Toll and Imd pathways. *Drosophila* Toll pathway is activated by Gram-positive Lys-type peptidoglycan (PG), while the Imd pathway is activated by Gram-negative meso-diaminopimelic acid (DAP)-type PG (Ganesan et al., 2011; Lemaitre and Hoffmann, 2007; Lemaitre et al., 1995; Lemaitre et al., 1996; Valanne et al., 2011). In *M. sexta*, both Lys-type and DAP-type PGs can activate expression of *M. sexta* AMP genes (Rao and Yu, 2010). To determine whether expression of *Msglv* is regulated by the Toll and/or Imd pathways in *M. sexta*, naïve larvae were injected with purified recombinant active MsSpz-C108 (a Toll pathway ligand) (Zhong et al., 2012), *S. aureus* and *E. coli* peptidoglycans (PGs), and expression of *Msglv* mRNA was determined. Real-time PCR results showed that expression of *Msglv* mRNA was up-regulated to significantly higher levels in the fat body (Fig. 38A) and hemocytes (Fig. 38B) by MsSpz-C108, *S. aureus* PG-SA (Lys-type) and *E. coli* PG-K12 (DAP-type) compared to the water-injection control. *Msglv* mRNA was also induced by MsSpz-C108 and PG-K12 in the midgut, but was not induced by PG-SA (Fig. 38C). MsSpz also activated *Msglv* in the fat body to a significantly higher level compared to water-injection (Fig. 38A), probably due to activation of MsSpz by hemolymph proteinases. In the fat body and hemocytes, MsSpz-C108 activated *Msglv* expression to a significantly higher level than PG-SA and PG-K12 did, and PG-SA and PG-K12 activated *Msglv* expression to a similarly high level (Fig. 38A and B). However, in the midgut PG-K12 activated *Msglv* expression to a significantly higher level than MsSpz-C108 and PG-SA did (Fig. 38C). These results suggest that expression of *Msglv* in the fat body and hemocytes is mainly regulated by the
Toll pathway, and that PG-K12 (DAP-type) activated \textit{Msglv} expression is regulated by the Toll-Spätzle pathways.

To confirm that expression of \textit{Msglv} is regulated by the Toll pathway and to determine whether PG-K12 activates the Toll or Imd pathway, an antibody blocking assay was performed. \textit{M. sexta} naive larvae were pre-injected with purified IgG from rabbit antiserum against the ecto-domain of \textit{M. sexta} Toll (\textit{MsToll}) or pre-bleed rabbit serum (Zhong et al., 2012), and then injected with \textit{MsSpz}, \textit{MsSpz-C108}, PG-SA or PG-K12. Expression of \textit{Msglv} in fat body, hemocytes and midgut was determined by real-time PCR (Fig. 39). Our results showed that in the control IgG pre-injected larvae, expression of \textit{Msglv} in the fat body, hemocyte and midgut was up-regulated by \textit{MsSpz-C108} (Fig. 39), and PG-SA activated \textit{Msglv} in the fat body and hemocytes (Fig. 39A and B), while PG-K12 activated \textit{Msglv} in the midgut (Fig. 39C). However, in the \textit{MsToll} IgG pre-injected larvae, \textit{MsSpz-C108}-activated expression of \textit{Msglv} in the fat body, hemocytes and midgut, PG-SA-activated \textit{Msglv} expression in the fat body and hemocytes, and PG-K12 activated expression of \textit{Msglv} in the fat body and midgut were all significantly suppressed (Fig. 39). In contrast, PG-K12-activated \textit{Msglv} expression in the hemocytes was stimulated after pre-injection of \textit{MsToll} antibody (Fig. 39B). These results suggest that pre-injection of larvae with IgG to \textit{MsToll} blocks the Toll receptor from binding to \textit{MsSpz-C108}, resulting in blocking the Toll pathway to activate \textit{Msglv} gene expression.

We have previously shown that a Toll–Spätzle pathway regulates expression of \textit{cecropin}, \textit{attacin}, \textit{moricin} and \textit{lebocin} genes in \textit{M. sexta} (Zhong et al., 2012). Thus, our results suggest that systematic expression of AMP genes in fat body, hemocytes and
midgut of *M. sexta* larvae are mainly regulated by the Toll pathway. Interestingly, the Toll receptor is involved in regulation of *Msglv* gene in midgut (Figs. 38 and 39C), *moricin* gene in hemocytes and *lebocin*b/c genes in fat body (Zhong et al., 2012) of *M. sexta* larvae activated by PG-K12 (DAP-type PG), a result differing from *D. melanogaster* in that DAP-type PG activates the Imd pathway (Leulier et al., 2003). In addition, *MsSpz*-C108 was more potent than PG-SA and PG-K12 in activation of *Msglv* gene in fat body and hemocytes, but PG-K12 was more potent than *MsSpz*-C108 and PG-SA in activation of *Msglv* gene in midgut (Fig. 38), suggesting that even though all three ligands can activate the Toll pathway, the recognition process at the cell surface may differ, and tissue-specific co-activators/receptors may also be involved. Activation of *Msglv* gene in hemocytes by PG-K12 was not blocked but stimulated after pre-injection of *MsToll* antibody (Fig. 39B), which is similar to activation of *lebocin*b/c in hemocytes by PG-K12 (Zhong et al., 2012). These results suggest that PG-K12-activated expression of *Msglv* and *lebocin*b/c genes in hemocytes may not be Toll-dependent and may be regulated by other pathways such as the Imd pathway. Thus, our results also suggest tissue-specific regulation of insect AMP genes.

The promoter regions of insect AMP genes contain NF-κB binding sites for Rel/NF-κB factors such as Dorsal, Dif and Relish. In *Drosophila*, Dorsal and Dif regulate the Toll pathway, while Relish regulates the Imd pathway. Therefore, whether an AMP gene is regulated by the Toll, Imd or both pathways depends upon the NF-κB binding sites in the promoter regions. For example, *D. melanogaster drosomycin* gene is synergistically regulated by the Toll and Imd pathways, since *drosomycin* gene promoter contains NF-κB binding sites for both Dif/Dorsal and Relish (Ganesan et al., 2011; Tanji
et al., 2007; Tanji et al., 2010). On the other hand, *diptericin* gene promoter contains NF-κB binding sites only for Relish and it is predominantly regulated by the Imd pathway (Ganesan et al., 2011). Some species/tissue-specific activators or suppressors may also contribute to regulation of AMP genes. *M. sexta moricin* gene promoter has high activity in Sf9 cells but has low or no activity in S2 cells, because the promoter region may contain a binding site for species-specific activator(s) (Rao et al., 2011).
Figure 38. *Msglv* expression in *M. sexta* larvae is activated by MsSpz-C108 and bacterial peptidoglycans.

Day 1 fifth instar *M. sexta* naïve larvae were injected with purified recombinant MsSpz (3 μg/larva), MsSpz-C108 (1 μg/larva), *S. aureus* PG (PG-SA) (1 μg/larva), *E. coli* PG (PG-K12) (1 μg/larva), or water (control), or left untreated (naïve), fat body, hemocytes and midgut were then collected at 20 h post-injection for preparation of total RNAs. Expression of *Msglv* mRNA was determined by real-time PCR. Ribosomal protein S3 (rpS3) gene was used as an internal control. The bars represent the mean of three individual measurements ± SEM. Relative expression of *Msglv* in naïve larvae was set as 1. Identical letters are not significant difference (p>0.05), while different letters indicate significant difference (p<0.05) among different treatments determined by one way ANOVA followed by a Tukey’s multiple comparison test.
Figure 39. Activation of Msglv triggered by MsSpz-C108 and bacterial peptidoglycans is blocked by antibody to M. sexta Toll.

Day 1 fifth instar M. sexta naïve larvae were pre-injected with purified IgG to the ecto-domain of M. sexta Toll (Toll Ab, 5 μg/larva) or IgG from pre-bleed rabbit serum (Control Ab, 5 μg/larva). One hour later, these larvae were injected with purified recombinant MsSpz (3 μg/larva), MsSpz-C108 (1 μg/larva), S. aureus PG (PG-SA) (1 μg/larva), E. coli PG (PG-K12) (1 μg/larva), or water (control), or without second injection (control), fat body, hemocytes and midgut were then collected at 20h after second injection for preparation of total RNAs. Expression of Msglv mRNA was determined by real-time PCR. Ribosomal protein S3 (rpS3) gene was used as an internal control. The bars represent the mean of three individual measurements ± SEM. Relative expression of Msglv mRNA after pre-injection of antibody but without second injection (control) was set as 1. Asterisks indicate significant difference (p<0.05) between Toll and Control antibody pre-injections for Msglv determined by an unpaired t-test.
MsGlv binds to microbial components and to the O-specific antigen and outer core moieties of LPS

It was suggested that Gloverin can interact with LPS to inhibit synthesis of E. coli outer membrane proteins (Axen et al., 1997). To demonstrate direct binding of MsGlv to LPS and other microbial cell wall components, a plate ELISA assay was performed using recombinant MsGlv purified from Drosophila S2 cells. The results showed that MsGlv bound to E. coli LPS, Gram-positive (S. aureus and B. subtilis) lipoteichoic acid (LTA-SA and LTA-BS) and peptidoglycan (PG-SA and PG-BS), and fungal laminarin (β-1, 3-glucan), but did not bind to E. coli PG (PG-K12) or fungal mannan (Fig. 40). More MsGlv protein bound to LPS (Fig. 40A) compared to other microbial components, and more MsGlv protein bound to S. aureus LTA-SA and PG-SA (Fig. 40C and D) than to B. subtilis LTA-BS and PG-BS (Fig. 40E and F). Our results confirm direct binding of Gloverin to LPS and showed broad binding of MsGlv to other microbial components.

LPS is composed of three moieties: the O-specific antigen, the core (outer and inner core) carbohydrate, and the lipid A (Raetz, 1990; Yu and Kanost, 2002). It was suggested that HgGlv may bind to the lipid A moiety because pre-incubation of HgGlv with Rd-LPS inhibits its activity against E. coli (Axen et al., 1997). To test binding of MsGlv to different moieties of LPS, several smooth LPS containing different O-specific antigens, different rough mutants of LPS, and lipid A were used in the binding assay. Plate ELISA results showed that various amounts of MsGlv bound to all five smooth LPS and Ra-LPS, but almost no binding of MsGlv to Rc-, Rd-, Re-LPS or lipid A was observed (Fig. 41). These results suggest that MsGlv can bind to the O-specific antigen and the outer core moieties of LPS.
HgGlv can interact with Rd-LPS and Re-LPS, but may not bind to Ra-LPS, since *E. coli* K12 D21 (with Ra-LPS) is less sensitive to HgGlv than *E. coli* K12 D21f2 (with Re-LPS) (Axen et al., 1997). Thus, MsGlv bound to both the O-specific antigen and the outer core carbohydrate of LPS (Fig. 41), while HgGlv may bind to the lipid A moiety of LPS. MsGlv has a pI of 9.35, while HgGlv has a pI of 8.23. The difference in the pI values of MsGlv and HgGlv may not completely account for the difference in binding to LPS. HgGlv has a random conformation in solution, but has a more defined structure in a membrane-like environment (Axen et al., 1997). Therefore, it is possible that differences in the defined structures of Gloverins after contacting microbial components or microorganisms account for the differences in binding to microbial components or activities against different microorganisms. A 3-dimentional structure of Gloverin may provide insight into the relationship between microbial binding and antimicrobial activity of Gloverin. So far, no structure of glycine-rich antimicrobial proteins is available.
Figure 40. Recombinant MsGlv binds to LPS, Gram-positive peptidoglycan (PG) and lipoteichoic acid (LTA), and laminarin.

Wells of 96-well fat-bottom microtiter plates were coated with TLR grade *E. coli* LPS-K12 (A) and PG-K12 (B), *B. subtilis* LTA-BS (E) and PG-BS (F), *S. aureus* LTA-SA (C) and PG-SA (D), laminarin (G) and mannan (H). Recombinant MsGlv and GFP purified from *Drosophila* S2 cells were diluted to different concentrations and added to the ligand-coated plates, and the binding assay was performed as described in the Materials and Methods. Each point represents the mean of four individual measurements ± SEM, and the lines represent nonlinear regression calculation of one-site binding curve.
Figure 41. Recombinant MsGlv binds to the O-specific antigen and out core carbohydrate moieties of LPS.

Wells of 96-well fat-bottom microtiter plates were coated with smooth LPS from several bacteria, Ra-, Rc-, Rd- and Re-LPS, monophosphoryl and diphosphoryl lipid A. Recombinant MsGlv and GFP purified from Drosophila S2 cells were diluted to 120 nM and added to the ligand-coated plates, and the binding assay was performed as described in the Materials and Methods. The figure showed specific binding of recombinant MsGlv to LPS and lipid A after subtracting the total binding of the control GFP from the total binding of MsGlv. Each bar represent the mean of three individual measurements ± SEM. Identical letters are not significant difference (p>0.05), while different letters indicate significant difference (p<0.05) among different ligands determined by one way ANOVA followed by a Tukey’s multiple comparison test.
MsGlv is active against bacteria and fungi

Gloverins are active almost exclusively against *E. coli* (Axen et al., 1997; Kawaoka et al., 2008; Lundstrom et al., 2002; Mackintosh et al., 1998). However, SeGlv is active against a Gram-positive bacterium (*Flavobacterium* sp.) but is inactive against *E. coli* (Hwang and Kim, 2011). Our binding assays showed that MsGlv could bind to Gram-negative LPS, Gram-positive LTA and PG, and fungal laminarin (Figs. 40 and 41), suggesting that Gloverin may be active against different microorganisms. To determine the activity of recombinant MsGlv against microorganisms, microbial growth curves were performed. Our results showed that MsGlv at low concentration (~1.5 μM or ~20 μg/mL) could inhibit the growth of Gram-positive *B. cereus* (Fig. 42C), and fungi *S. cerevisiae* (Fig. 42G) and *C. neoformans* (Fig. 42H) even after long incubation time (over 20h for bacteria and 30h for fungi). MsGlv also had activity against Gram-negative *E. coli* (Fig. 42A) and *S. marcescens* (Fig. 42B), and Gram-positive *B. subtilis* (Fig. 42D), *B. thuringiensis* (Fig. 42E) and *S. aureus* (Fig. 42E) when the incubation time was short (less than 10h, indicated by asterisks), but could not completely inhibit the growth of these microorganisms when incubation time was longer (over 15 h). In addition, when the initial CFU of *S. cerevisiae* or *C. neoformans* was increased, MsGlv could not completely inhibit their growth and the growth curves had a pattern similar to that of *E. coli* or *S. aureus* (data not shown). These results suggest that MsGlv is active against Gram-negative and Gram-positive bacteria as well as fungi to some extent depending on the microorganism.

MsGlv could bind to Gram-positive LTA and PG, as well as fungal laminarin (Fig. 41), which may account for its activity against *B. cereus*, *S. cerevisiae* and *C. neoformans*. 
MsGlv also bound to LPS (Figs. 40A and 41), but it was almost inactive against *E. coli* (Fig. 42A), while most Gloverins from other insect species are active against *E. coli* but inactive against Gram-positive bacteria (Axen et al., 1997; Kawaoka et al., 2008; Lundstrom et al., 2002; Mackintosh et al., 1998). We think that in vitro assay to test whether a Gloverin is active against Gram-negative, Gram-positive bacteria, and/or fungi may depend upon microbial strains, initial number (CFU) of microorganisms, incubation time, and the concentration of Gloverin. For example, *B. mori* Gloverins (*BmGlvs*) are active against *E. coli* in phosphate buffer within 6 h of incubation, but the activity is weak (Kawaoka et al., 2008). We observed that when the initial CFU of *S. cerevisiae* and *C. neoformans* was increased, MsGlv (at ~1.5 μM) was almost inactive against the two fungi with a growth curve similar to that of *E. coli* or *S. aureus* (data not shown). Thus, Gloverin may play a role in defense against initial infection in insects when the number of microbes is low.
Figure 42. Antimicrobial activities of recombinant MsGlv.

Mid-log phase bacteria (Gram-negative *E. coli* and *S. marcescens*, Gram-positive *B. cereus*, *B. subtilis*, *B. thuringiensis* and *S. aureus*) and fungi (*S. cerevisiae* and *C. neoformans*) were diluted to OD$_{600}$=$10^{-5}$ and incubated with recombinant MsGlv purified from S2 cells (final concentration of ~21 μg/ml or 1.5 μM) or water (Control) in 96-well plates with 220 rpm shaking at 37°C (for bacteria) or 30°C (for fungi). OD$_{600}$ was recorded every hour up to 20 h (for bacteria) or 30 h (for fungi) after incubation. The points represent the mean of four individual measurements ± SEM. The asterisks in the boxes indicate significant difference (p<0.05) between MsGlv and the control at the indicated time points determined by an unpaired t-test.
The innate immune system is conserved from insects to humans. In insects, the Toll and IMD signaling pathways, PGRPs, and Dorsal, Dif and Relish NF-κB factors in regulation of immune genes have been well studied in the model organism *D. melanogaster*. However, signaling pathways and NF-κB factors in regulation of gene expression in other insect species have not been well characterized. Even in *Drosophila*, it is also debatable whether LPS can activate AMP gene expression. In some insect species including *M. sexta*, it has been reported that LPS can activate innate immune responses (Kawabata S, 2009; Rao and Yu, 2010; Zeidler et al., 2004); however, whether there is an LPS-triggered signaling pathway in insects or in invertebrates remains elusive.

I have confirmed a Toll-Spz pathway in *M. sexta*, a *Lepidopteran* insect. We showed direct interaction between MsToll and MsSpz-C108 by Co-immunoprecipitation (CoIP), and demonstrated that co-expression of MsToll and MsSpz-C108 could activate AMP gene promoters in S2 cells by dual luciferase assays. *In vivo* assays showed that MsToll antibody could block stimulation of AMP genes by MsSpz-C108 in *M. sexta* larvae. This study may help better understand signaling pathways in *lepidopteran* insects, and the origin and evolution of animal innate immune signaling pathways. Since there are multiple Toll receptors and Spz proteins in insects, for example, there are nine Toll receptors and six Spz proteins in *Drosophila*, one future direction is to investigate whether there are multiple Toll-Spz pathways and their roles in innate immune response.
Mammalian TLR4-MD2-LPS signaling has been well studied. However, it is not known whether a similar LPS signaling pathway exists in invertebrates. I showed by in vitro assays that MsToll, MsML-1 and LPS could form a receptor complex. More importantly, I showed that co-expression of chimeric Toll (ChiToll by combining the ecto-domain of MsToll with the transmembrane and TIR domains of Drosophila Toll) with MsML-1 could up-regulate AMP genes activated by LPS. Our results for the first time showed that a Toll-ML-LPS signaling pathway is conserved from insects to humans. Drosophila has eight MD2-like proteins, and it is not clear which Toll receptor and/or MD2-like protein are involved in LPS signaling. Also, it is interesting to see whether MsToll-MsML-1 could replace TLR4-MD2 to sense LPS and activate NF-κB factors in mammalian cells. In human HEK293T cell line, NF-κB factors cannot be activated without co-expression of both human TLR4 and MD2. Thus, it is possible to test whether chimeric human TLR4 (ChiTLR4, combining the ecto-domain of MsToll with the transmembrane and TIR domains of human TLR4) can form a complex with MsML-1 to bind LPS and activate the NF-κB pathway in HEK293T cells.

In Drosophila, Dorsal and Dif are activated by degradation of the inhibitor Cactus, while Relish is activated by cleavage off the C-terminal ankyrin repeats. Dorsal, Dif and Relish can form homo- and hetero-dimers (Ganesan et al., 2011; Tanji et al., 2007; Tanji et al., 2010). In order to form Dif-Relish heter-dimers, both the Toll and IMD pathways must be activated. I identified two short isoforms of Relish in M. sexta (Rel2A and Rel2B), which were functionally active NF-κB factors. More importantly, I showed that Dorsal and Rel2 could form heterodimers, which could negatively regulate AMP genes. This is a novel finding about NF-κB factors in regulation of gene expression. Both
MsDorsal and the short isoforms of MsRel2 may form homodimers to activate AMP gene expression. If the Toll pathway is over-activated, Dorsal-Rel2 heterodimers can be formed, thus Rel2 short isoforms not only can serve as activators but also may serve as dampen factors to prevent over-activation of the Toll pathway. It would be interesting to see whether Relish short isoforms are also present in Drosophila or not, whether they play a role similar to M. sexta Rel2 isoforms.

β-1, 3-glucan recognition proteins (βGRPs) have been identified in several insect species and their main role is to stimulate PPO activation. I identified a novel member of the β-1, 3-glucan recognition proteins (βGRP3) in M. sexta. I found βGRP3 possesses novel properties, it has calcium-dependent agglutinating activity for B. cereus and E. coli, and bactericidal activity against B. cereus. These functions are new to the β-1, 3-glucan recognition proteins, and it is necessary to investigate the mechanisms underlying agglutination of bacteria and bactericidal activity, which have been observed for C-type lectins (immulectins) and AMPs in M. sexta and other insects.
REFERENCES


larvae of the dipteran Phormia terranovae and complete amino-acid sequence of the predominant member, diptericin A. Eur J Biochem 171, 17-22.


Kleino, A., 2010. The Imd Pathway-mediated Immune Response in Drosophila. the University of Tampere, Finland, p. 132.


and ubiquitination: linking three mechanisms crucial for Drosophila NF-kappaB signaling. Mol Cell 37, 172-182.
nitric oxide synthase (NOS) and induce defense genes. Proc Natl Acad Sci U S A 101, 15811-15816.
VITA

Xue Zhong was born in Chongqing, China on November 18, 1983. She graduated from Xiamen University in Xiamen, Fujian, China in 2006 with a Bachelor of Science in Biology. In 2009, she graduated from Xiamen University in Xiamen, Fujian, China with a Master of Science in Biology. In 2010, she received a Master of Science in Cell Biology and Biophysics from the University of Missouri-Kansas City. In 2009, she joined Dr. Xiao-Qiang Yu’s lab studying insect immunology using the tobacco hornworm Manduca sexta.