UC San Diego

UC San Diego Electronic Theses and Dissertations

Title

Transcriptional regulation of antimicrobial peptide induction by NF-[kappa]B family members during the Drosophila melanogaster immune response

Permalink

https://escholarship.org/uc/item/00g0061d

Author

Busse, Matthew Schmidt

Publication Date

2007

Peer reviewed|Thesis/dissertation

eScholarship.org Powered by the California Digital Library
University of California

UNIVERSITY OF CALIFORNIA, SAN DIEGO

Transcriptional Regulation of Antimicrobial Peptide Induction by NF-kappaB Family Members During the *Drosophila Melanogaster* Immune Response

A dissertation submitted in partial satisfaction of the requirements for the degree

Doctor of Philosophy

in

Biology

by

Matthew Schmidt Busse

Committee in charge:

Professor Steven P. Briggs, Chair
Professor Ethan Bier
Professor Christopher K. Glass
Professor Stephen M. Hedrick
Professor Steven A. Wasserman

2007

The Dissertation of Matthew Schmidt Busse is approved, and it is acceptable in quality and form for publication on microfilm:

_____

_____

_____

_____

_____

Chair

University of California, San Diego
2007

EPIGRAPH

A process cannot be understood by stopping it.
Understanding must move with the flow of the
process, must join it and flow with it.

-The First Law of Mentat
From *Dune*, by Frank Herbert

TABLE OF CONTENTS

Signature Page.....iii

Epigraph.....iv

Table of Contents.....v

List of Figures.....viii

List of Tables.....x

Acknowledgements.....xi

Vita.....xii

Abstract.....xiii

I. Introduction

NF-κB.....1

Drosophila Immunity

Overview.....5

Background.....9

Summary.....31

II. Results

Establishment of A Cell Culture Based Luciferase Assay

A Transfected Attacin A Construct Faithfully Reports

Toll and Imd Signaling.....32

Promoter Proximal κB Sites Govern AttA Induction

by the Toll and Imd Pathways.....35

Bioinformatic Analysis

Toll and Imd Responsive Loci Differ in κB Site Structure and Sequence……..37

Toll and Imd Responsive Loci Differ in κB Site Number.....40

A κB Sequence Code Governs Drosophila Innate Immunity

Confirmation of Bioinformatic Analysis in Cell Culture.....41

The κB Sequence Code Functions by Specifying NF-κB Protein Binding…….....43

Determinants of Dif and Relish Binding…….....45

A Single GGGGA Half-site is Sufficient for Relish Binding.....49

Analysis of Additional Reporters…….51

The Drosomycin Reporter Contains Two Functional κB Motifs….52

The IM1 Reporter Contains One Functional κB Motif……..55

The Attacin D Reporter Contains Three Functional κB Motifs……56

Imd Induction of Pathway Specific κB Motifs in Additional Promoter Contexts…….58

Toll Induction of Pathway Specific κB Motifs in Additional Promoter Contexts…….59

Multiple κB Motifs Function Cooperativly to Regulate Transcription……60

Activation of Both Signaling Pathways Results in a Synergistic Response…….61

Two Pathway Cooperativity Requires Two κB motifs and

Two Transcription Factors……..….62

Dorsal Contributes to the Regulation of AttA and Drs, But Not IM1…..….64

III. Discussion

Overview……..67

Nature of the Sequence Code…….68

The Role of Cooperativity in Immune Responsive Genes…….…72

Cooperativity Between Two Pathways……74

Transcriptional Regulation of Immune Responsive Genes……76

Dorsal…….……..77

IV. Experimental Procedures

Reagents and Site-Directed Mutagenesis…….……78

Cell Culture and Transient Transfections…….……..…78

Bioinformatic Analysis of *Drosophila* Innate Immune Response Loci.....79

Protein Purification.....80

Gel-Shift Assays..……...81

Works Cited…….82

LIST OF FIGURES

Figure 1: Activating Signals and Targets of Mammalian NF-κB Proteins……2

Figure 2: A Simplified Paradigm of Antimicrobial Peptide Induction……8

Figure 3: Recognition of Microbes Through Pathogen Recognition Receptors...11

Figure 4: Components of the Toll and Imd Pathways……..14

Figure 5: Signaling Pathway Activation by Pathogens……16

Figure 6: Pathway Induction of Antimicrobial Peptide Genes……21

Figure 7: Gram(-) induction of Antimicrobial Peptide Genes……..22

Figure 8: Gram(+) bacteria induction of Antimicrobial Peptide Genes……..24

Figure 9: Induction of Antimicrobial Peptide Genes by Fungal Injection…..……25

Figure 10: An AttA Reporter System Recapitulates Endogenous
Innate Immune Responses……..…….34

Figure 11: κB Sites Determine Pathway Specific Transcriptional Responses….36

Figure 12: Imd and Toll responsiveness of consensus motifs
identified by MEME……..……..41

Figure 13: Dif and Relish Exhibit Selective Binding to Synthetic
and Endogenous κB sites……..…...……44

Figure 14: Spacing is one feature that determines specificity…….……46

Figure 15: Spacing determines specificity in three G motifs…….…….47

Figure 16: Gel shift assays with Gst-Relish (R) and Gst-Dif
reveals that a single GGGGA half-site is often sufficient
for Relish binding……..……..48

Figure 17: Specific changes in the 3’ half-site disrupt binding……..…..49

Figure 18: Toll and Imd Inductions of Additional Luciferase Reporters……50

Figure 19: κB Motifs in the Drosomycin reporter……52

Figure 20: The κB motif in the IM1 reporter……53

Figure 21: κB Motifs in the Attacin D reporter……54

Figure 22:Imd inductions of synthetic κB motifs in additional promoter contexts.....56

Figure 23:Toll inductions of synthetic κB motifs in additional promoter contexts.....57

Figure 24: Imd and Toll inductions of multiple synthetic motifs in the AttA reporter……58

Figure 25: Activation of both the Imd and Toll pathways induces a synergistic response……59

Figure 26: Two κB motifs are required for a synergistic response to both pathways……60

Figure 27: A Synergistic response to both pathways requires two transcription factors……61

Figure 28: The role of Dorsal in the Toll response of Drosomycin and IM1…...62

Figure 29: Summary of mammalian NF-κB binding to different κB motifs……67

LIST OF TABLES

Table 1: Function of Antimicrobial Peptide Families……..……..……..7

Table 2: Direct activation of signaling pathways differentially induces AMPs…..18

Table 3: Results of MEME motif analysis……..……..……..38

Table 4: Innate immune motifs……..……..…….40

Table 5: Mode of binding to different κB motifs..……..…….68

ACKNOWLEDGEMENTS

First, I wish to thank my mentor, Steven Wasserman. His guidance and influence set the tone for an ideal work environment. I also thank all members of the Wasserman Lab, past and present, for fostering a supportive, collaborative, and entertaining lab environment. In particular, I thank Hauiyu Sun and Par Towb for their aid in teaching me how to perform many of the experiments here. Also, Tammie Stone and Nick Reeves in the Posakony lab helped immeasurably with protein purification and gel-shifts. Special thanks go to Jim Katrivesis, who actually performed many of the gel-shift assays presented below, and to Chris Arnold, for many of the RNAi experiments. I also thank Steven Ho, Chris Glass, and Micheal David for extensive use of their luminometers.

The pGex-Dif expression plasmid was a generous gift from Tony Ip, and the Drosomycin reporter was a gift from Jean-Luc Imler.

Portions of Chapter Two are taken from Busse, MS, Towb, PT, Arnold, CP, Katrivesis, J, and Wasserman, SA. “A κB Sequence Code for Pathway-Specific Innate Immune Responses” Which was under revision at Molecular Cell at the time of submission of this dissertation. The dissertation author was the primary researcher and author of this paper.

VITA

2000 Bachelor of Arts, Wabash College

2000-2001 Research Technician, The Scripps Research Institute

2007 Doctor of Philosophy, University of California, San Diego

PUBLICATIONS

"A κB Sequence Code for Pathway-Specific Innate Immune Responses"

Under Revision

ABSTRACT OF THE DISSERTATION

Transcriptional Regulation of Antimicrobial Peptide Induction by NF-kappaB Family Members During the *Drosophila Melanogaster* Immune Response

by

Matthew Schmidt Busse

Doctor of Philosophy in Biology

University of California, San Diego, 2007

Professor Steven P. Briggs, Chair

The innate immune system of *Drosophila melanogaster* is a well-studied model of evolutionarily conserved innate immunity. The humoral branch of *Drosophila* immunity is regulated by the Toll and Imd pathways, which signal through the NF-κB family members Dif and Relish, respectively. Here is presented evidence that κB motifs, the target DNA sequences of NF-κB factors, direct pathway-specific gene expression of innate immune loci by determining the binding of Dif and Relish to the promoters of target genes. Using a combination

of immune gene luciferase reporters, bioinformatic analyses, and *in vitro* binding studies, the regulatory code through which κB motifs direct this pathway specific expression was deciphered. The binding specificity of κB motifs is determined by the number of guanine bases in the 5' G cluster, and the number of A or T bases between the 5' G cluster and the 3' C cluster. Examination of four immune responsive genes led to the discovery that cooperative interaction of multiple κB motifs is a crucial determinant of pathway responsiveness. Multiple κB motifs responsive to the same pathway cooperatively increase the response to that pathway, and multiple motifs responsive to different pathways allows a synergistic response to simultaneous activation of both pathways. Together, these studies elucidate the mechanism by which κB motifs direct binding by particular *Drosophila* NF-κB family members and thereby regulate the induction of immune responsive genes.

I. Introduction

NF-κB

The family of transcription factors known as Nuclear Factor-κB (NF-κB) has come to light as one of the cardinal regulators of the immune system, involved in both activation of the immune response and the development of the immune system [1, 2]. The name is derived from the initial discovery of NF-κB as a DNA-binding activity directed at the immunoglobulin κ light chain gene in activated B cells [3]. Since then, extensive research has uncovered the role of NF-κB in the regulation of innate immune response factors, including cytokines, chemokines and antimicrobial peptides, and adaptive immune response effectors, including MHC proteins and co-stimulatory molecules. A developmental role in apoptosis and cell proliferation, including biogenesis of lymphoid organs and immune cell differentiation [2, 4, 5], has also been found for NF-κB.

Research into the mechanism through which the five mammalian NF-κB family members regulate their target genes has revealed a deliciously complex regulatory scheme which generates specific responses at multiple levels [2, 6, 7]. The first level of regulation varies with the type of stimuli. In the absence of stimulation, NF-κB factors are kept inactive in the cytoplasm by inhibitor proteins belonging to the family of Inhibitor of κB (IκB). Activating signals lead to phosphorylation, ubiquitination, and proteosomal degradation of specific IκB proteins, allowing particular NF-κB proteins to initiate transcription of target

genes. Inflammation primarily activates the classical p50 and p65 family members, which often function as a heterodimer. cRel is active during hematopoiesis, and p52 and RelB respond primarily to the non-inflammatory signals of lymphoid organogenesis [2, 8] (Figure 1). NF-κB factors work as dimers, and twelve different combinations of homo- and heterodimers have been observed.

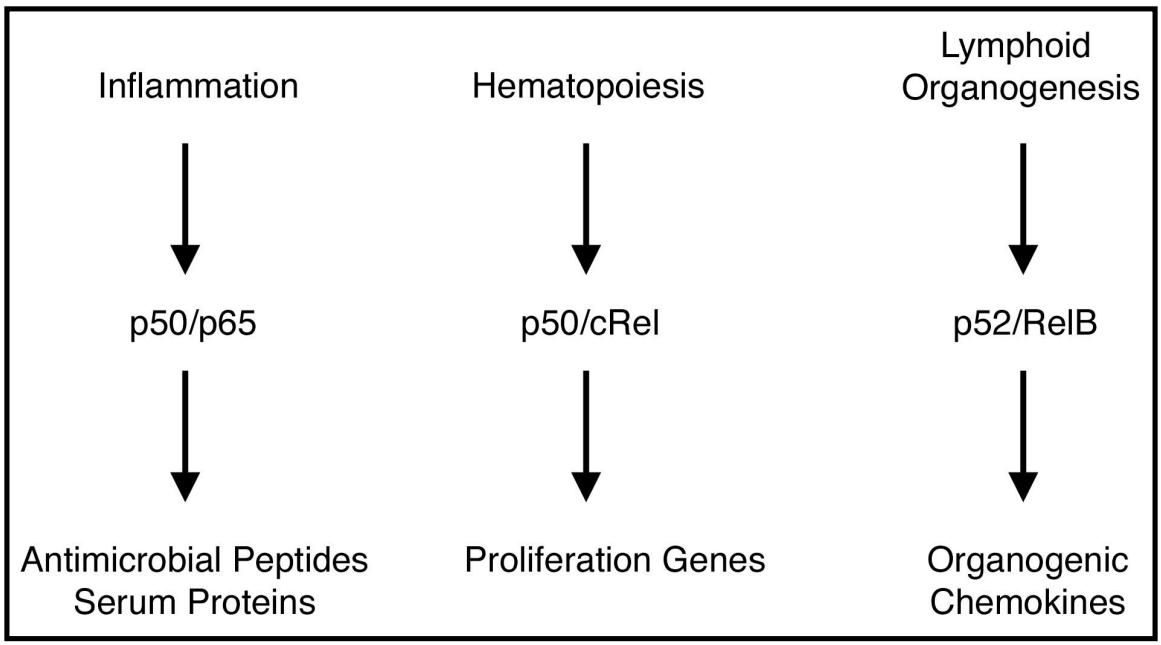

Figure 1: Activating Signals and Targets of Mammalian NF-κB Proteins

On top of the induction of specific NF-κB members lies a layer of temporal control, generated by waves of IκB synthesis [9] and IκB Kinase activation [10], as well as regulated access to chromatin [6, 11]. An additional layer of regulation is generated by the interaction of NF-κB members with other transcription factors [11, 12].

All NF-κB family members bind to *cis*-regulatory DNA sequences known as κB motifs, and the most obvious level for gene-specific activation by different NF-κB dimers is via recognition of different κB motifs in the enhancers of target genes. However, specific gene expression is generated without clear preferences of NF-κB dimers for specific κB motifs [2, 7]. While it is clear that correct gene expression requires specific dimer pairs, this requirement does not correlate with specific κB motifs [13]. Furthermore, several kB motifs are bound indiscriminately by several different NF-κB dimers [14]. In at least one documented case, the κB motif sequence determines specific gene expression not by dictating which NF-κB factors are bound, but by determining the capability of bound NF-κB factors to interact with specific co-factors [12].

The intrinsic affinity of NF-κB proteins for κB motifs does come into play in some cases [15]. For example, selection studies have revealed that p65 homodimers, p50 homodimers, and p50:p65 heterodimers each prefer different target sequences[16]. In addition, chemokine genes involved in the organization of lymphoid organs are specifically activated by RelB:p52 dimers via κB motifs which are selectively bound by RelB:p52 dimers and not p65:p50 [17]. Another example of regulation through differential affinity is seen in the *Il12b* gene, whose regulation depends the affinity of c-Rel for a broader range of κB motifs than p65 [18]. Furthermore, while different κB motifs are often bound with only minor differences in affinity by specific dimers, these minor differences can be biologically relevant [19].

The consensus sequence for κB motifs is conventionally described as GGGRNNYYCC (R is any purine, N is any nucleotide, and Y is any pyrimidine) [1]. Although this consensus defines a broad range of sequences, there exist some sites which deviate even from this loose consensus and are still high affinity sites [6]. The major cause of the broad range of binding exhibited by NF-κB is the overall flexibility of these proteins: contact with DNA is made through five separate flexible loops, the two domains of each subunit are connected by a flexible linker, and the two subunits dimerize through flexible contacts [20-23]. The result of this remarkable flexibility is the ability of NF-κB proteins to adopt different conformations, and bind with similar affinity to different κB motifs. These altered conformations allow the same dimer to use entirely different amino acids to form different base-specific hydrogen bonds when bound to different motifs [24].

As a consequence of the plasticity in binding conformations, clear concrete rules governing the binding of NF-κB dimers to specific κB motifs remain elusive [2, 7]. Indeed, efforts to define separate consensus motifs for specific dimers are confounded by internucleotide interactions in which the same single nucleotide change can have wildly different effects on binding affinity depending on the rest of the sequence. For example, changing GGGATA<u>C</u>CCC to GGGATA<u>T</u>CCC has no effect on p50 homodimer affinity, but changing GGGGCT<u>C</u>CCC to GGGGCT<u>T</u>CCC markedly decreases binding affinity [19].

In light of this intricate complexity, investigation of a model organism, such as *Drosophila melanogaster*, can yield insights into the basic principles governing transcriptional regulation by NF-κB.

Drosophila **Immunity**

Overview

Although *Drosophila,* like all other invertebrates, lack the somatic DNA arrangements and high-specificity T- and B-cell receptors that characterize an adaptive immune system, their innate immune systems provide sufficient protection to survive and proliferate successfully in microbe-laden environments. Innate immune systems are evolutionarily ancient, being found in plants, animals and insects. Only recently has the importance of mammalian innate immunity been fully appreciated for its role in initiating the adaptive immune response, as well as providing protection during the week or so required for the adaptive system to mount a response.

The *Drosophila* immune response can roughly be divided into two branches: a cellular response, including phagocytosis and a phenol oxidase cascade leading to melanization and encapsulation, and a humoral response, composed primarily of the rapid synthesis of a cocktail of antimicrobial peptides and complement like proteins in the fat body and at the site of infection (for reviews of *Drosophila* immunity, see [25-30]).

Antimicrobial peptides (AMPs) are a critical immune component in virtually all multicellular organisms [31], and the *Drosophila* immune repertoire consists of at least 34 AMPs, divided into eight families. All AMPs are small (4-30 kDa) and cationic [32]. In response to infection, they are induced to micromolar concentrations in the hemolymph (the *Drosophila* blood equivalent) within a matter of hours [28].

Antimicrobial peptides target conserved features of microbes that are absent in multicellular organisms. One such feature is the concentration of negatively charged lipids on the outer layer of the cell membrane. The membranes of plant and animal cells contain no net charge; as a result, the cationic immune peptides specifically interact with microbial membranes [31]. Since AMPs target these conserved features, microbes are virtually unable to develop resistance to AMPs, which show little evidence of selective pressure to evolve among insects [33]. Some microbes achieve resistance by evolving secreted proteases that inactivate peptides.

While all microbes share some features absent in multicellular organisms, they are unique in other features. For example, Gram(-) bacteria contain an extra LPS membrane surrounding the cell membrane that is absent in Gram(+) bacteria. As a result, different peptides are necessary to combat different microbes. Table 1 summarizes known specificity of identified peptides.

Table 1: Function of Antimicrobial Peptide Families

Peptide family	Microbes inhibited by family
Attacin	Gram(-) bacteria
Cecropin	Gram(-) and Gram(+) bacteria
Defensin	Gram(+) bacteria
Diptericin	Gram(-) bacteria
Drosocin	Gram(-) bacteria
Drosomycin	Fungi
Metchnikowin	Gram(+) bacteria and Fungi

AMPs function by adopting an amphipathic structure, inserting into the microbial membrane, and causing death through a variety of mechanisms, such as depolarization of membrane potentials, creation of physical holes, activation of hydrolases, or disruption of lipid distribution [31].

Several reasons could explain why multiple AMP genes are induced in response to one class of pathogen. First, since different AMPs act though slightly different mechanisms, some peptides potentiate the activity of each other in a cooperative manner [34]. Second, inducing multiple peptides increases the number of proteases a microbe would need to evolve in order to gain resistance.

Interestingly, immune peptides with different activity are differentially induced in response to different types of infection [35], raising the question: How does *Drosophila* discriminate between different types of microbes and mount an appropriate immune response? During the past decade, much time and effort has been devoted to understanding how *Drosophila* recognizes distinct types of infection, and then transduces that recognition into distinct transcriptional profiles of antimicrobial peptides.

From the research in this field, a paradigm has emerged which labels fungi and Gram(+) bacteria as inducers of the Toll pathway [36], which acts through the NF-κB member Dif to activate anti-fungal and anti-Gram(+) bacterial peptides [37-39]; and Gram(-) bacteria as inducers of the Imd pathway [40, 41], which acts through the NF-κB member Relish to activate peptides effective against Gram(-) bacteria [42-44] (Figure 2).

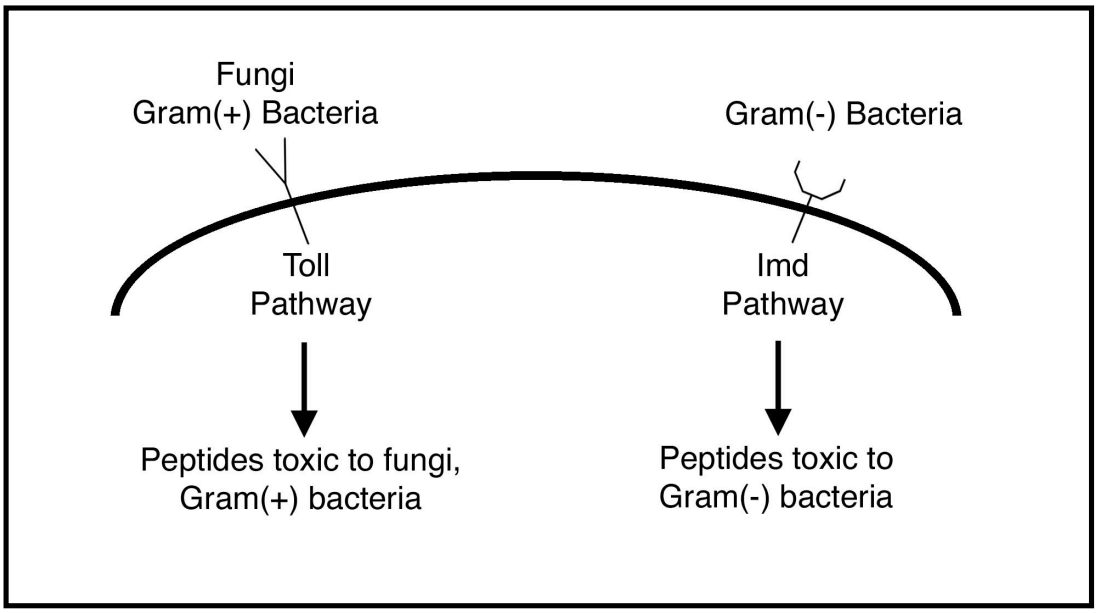

Figure 2: A Simplified Paradigm of Antimicrobial Peptide Induction

While this paradigm has dominated research in the field, data in several publications, including the original publications upon which this paradigm is based, clearly indicate that the division into fungi and Gram(+) bacteria versus Gram(-) bacteria is an oversimplification [35, 45, 46]. The reason this paradigm has been maintained is that no other model can cleanly and simply relate pathogen recognition, signal transduction, transcription factor activation, and

peptide induction. A careful review of the literature reveals that some aspects of the regulation of antimicrobial peptides are clearly established, while some aspects are not so clear.

Background

-Pattern recognition

The innate immune systems of vertebrates and invertebrates recognize invading pathogens through pathogen recognition receptors, which recognize features of microbes that are essential and invariant among microbes, but absent from eukaryotic cells. Peptidoglycan of bacteria, mannans of yeast, and double-stranded RNA of viruses are all examples of pathogen-associated molecular patterns (PAMPs) that are essential to the survival of broad classes of microbes. By targeting these PAMPs, the innate immune system ensures recognition of infectious non-self in contrast to noninfectious self [47-49].

Recent studies have demonstrated that the initial recognition of infection in *Drosophila* occurs through Peptidoglycan Recognition Proteins (PGRPs) [50], and Gram-Negative Binding Proteins (GNBPs) [51]. A major breakthrough in this field was the discovery that bacterial recognition is based on the type of peptidoglycan present in the cell wall, not on Gram type. Even though the first paper elucidating the differential activation of antimicrobial peptides demonstrated that some of the Gram(+) bacteria tested activated the Imd pathway and others activated the Toll pathway [35], subsequent studies used a

species of Gram(+) bacteria, *M. luteus*, that primarily turned on the Toll pathway, leading to the perception that all Gram(+) bacteria only activate the Toll pathway.

Subsequent research demonstrated that Gram(+) bacteria containing peptidoglycan with a lysine residue in the third position of the stem peptide (Lys-PGN) activate the Toll pathway through PGRP-SA [52, 53] and GNBP-1 [54, 55] (Gram-Negative Binding Protein-1 recognizes Gram(+) bacteria *in vivo*, but was named for an *in vitro* binding to LPS). These receptors work together to recognize Lys-PGN: GNBP-1 hydrolyzes Lys-PGN polymers into dimers and tetramers, which are then recognized by PGRP-SA **[56]**. Gram(+) bacteria containing Lys-PGN in their cell membrane include *M. luteus*, *E. faecalis*, and *S. aureus*.

All Gram(-) and some Gram(+) bacteria, including the *Bacillus* genus, contain peptidoglycan with diaminopilemic acid instead of lysine at the third position of the stem peptide. This DAP-PGN is recognized by PGRP-LC [57-59] and PGRP-LE [46, 60]. PGRP-LC contains a transmembrane domain, and functions as a transmembrane receptor. PGRP-LE exhibits extremely complex behavior, involving a short version of the protein which acts extracellularly in cooperation with PGRP-LC [61], and a longer version which acts intracellularly [46]. A third, unidentified receptor may be involved as well [46]. Furthermore, there is evidence of additional discrimination between different types of DAP-PGN. Gram(+) *Bacillus* and Gram(-) bacteria differ in the amount of meso- versus amidated DAP contained in peptidoglycan, leading to different levels of immune

activation [62]. Further investigation will likely reveal more mechanisms of discrimination between pathogens.

Less is known about the recognition of fungi. A serine protease, Persephone, is known to be involved [63], and unpublished data suggests the involvement of GNBP-3, which recognizes b-1,3-glucan [64].

The discovery and characterization of PGRPs and GNBPs have significantly clarified and refined the paradigm (Figure 3), yet questions remain. Thirteen members of the PGRP family have been identified in *Drosophila* [50], and the role of other family members is under investigation. Two of these family members, PGRP-SC1 and PGRP-SC2 have catalytic activity and appear to down-regulate the immune response, perhaps to reduce immune responses to non-pathogenic, ingested peptidoglycan [65].

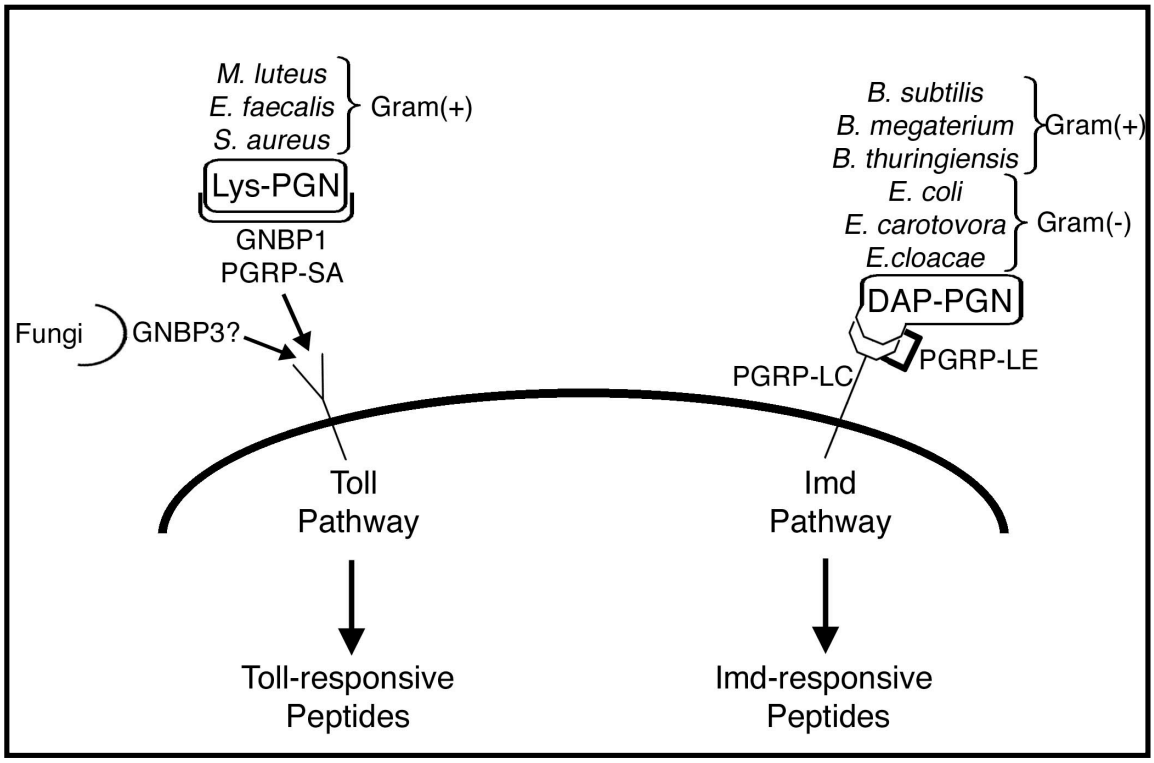

Figure 3: Recognition of Microbes Through Pathogen Recognition Receptors

-Signal Transduction Pathways

The Toll and Imd pathways are clearly the two most important signaling pathways controlling antimicrobial peptide genes, as mutation of both pathways render flies incapable of inducing any AMP genes [36, 66]. What is unclear is the exact role each plays in responding to different species of pathogens and inducing appropriate immune peptides.

The Toll pathway was originally discovered for its role in dorsal-ventral patterning during *Drosophila* development. The entire signaling cascade was then shown to play a role in the immune response [36], including the extracellular ligand, Spaetzle, which activates the transmembrane receptor, Toll, causing the formation of a complex involving the adaptor proteins Tube and MyD88 with the

kinase Pelle [67]. During development, Toll signaling leads to the activation of the transcription factor Dorsal (Dl), a member of the NF-kB family, through degradation of its inhibitor, Cactus. During immune signaling, it appears that Dif (Dorsal-related immunity factor) is the major effector in adults [37, 38]. However, in larvae, Dif and Dl appear to play redundant roles [39]. A more thorough discussion of Dif and Dl will be presented later.

The Imd pathway, which was discovered through screens for immune deficiencies [40], consists of a death domain containing adapter protein [41], an IκB kinase complex [68] (composed of IKKβ [43] and IKKγ [69]), the caspase Dredd [70], and Relish [42], another NF-κB family member. Unlike Dif and Dl, which are held inactive in the cytoplasm by the inhibitor Cactus, Relish is synthesized as a compound molecule, containing both a Rel Homology domain and an ankyrin-rich inhibitor domain very similar to cactus. This inhibitory domain is cleaved off in response to Imd signaling, allowing Relish translocation into the nucleus [71].

Less clear is the role of another protein, dTak1 [44], and another signaling module, JNK [72] in Imd signaling. Initial evidence placed the JNK pathway downstream of the Imd receptor, branching away from the IKK/Relish pathway at dTAK1 [73]. The JNK pathway and IKK/Relish pathway are both involved in regulating immune responsive genes, and appear to negatively regulate each other [73, 74]. However, two recent reports suggest that Imd responsive antimicrobial peptides actually require components of the JNK pathway for full

induction [75, 76]. Further experiments are required to understand these complex interactions.

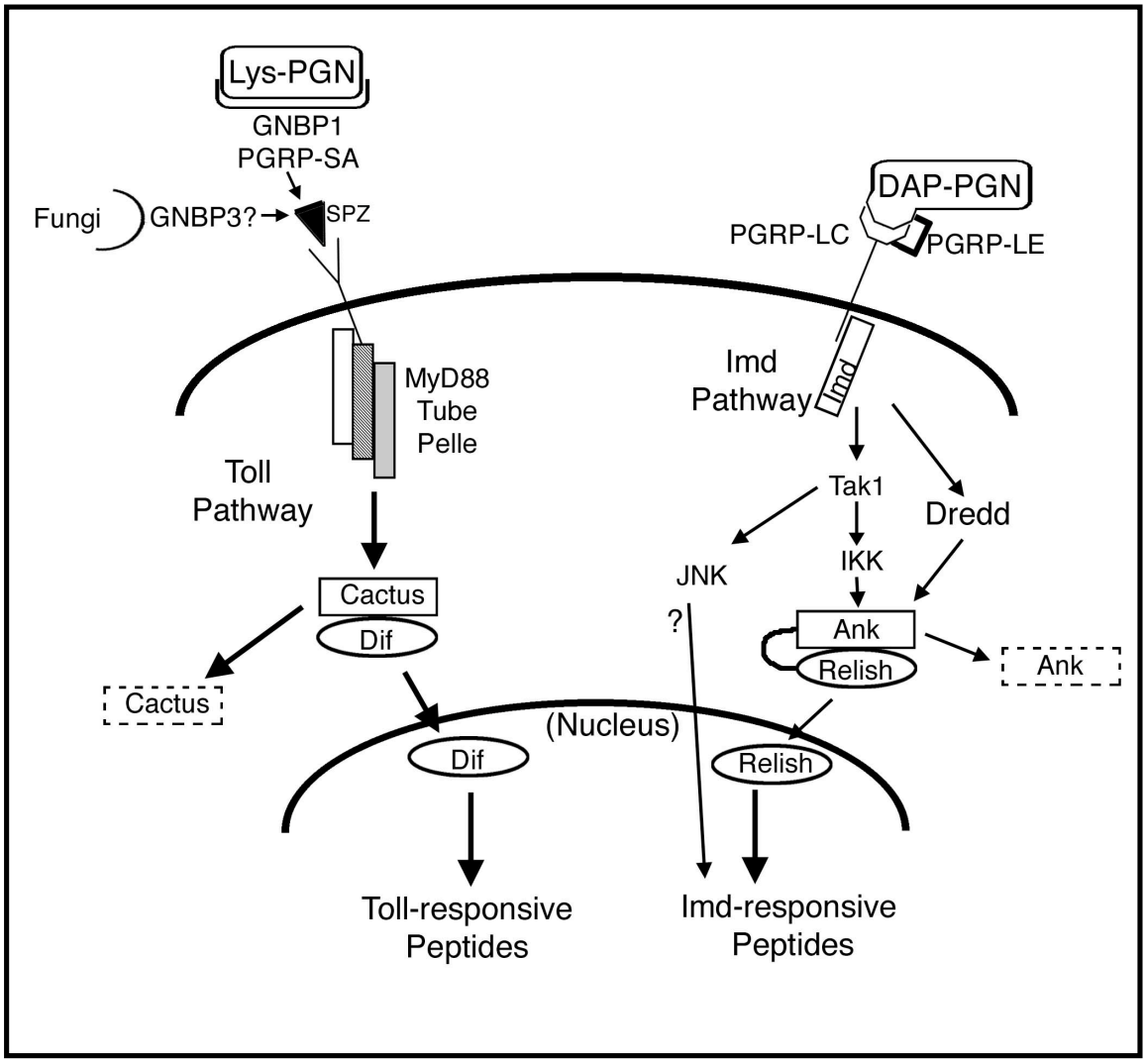

Figure 4: Components of the Toll and Imd Pathways

-Requirements of signaling pathways for immune responses

Survival experiments have been extensively used to determine which pathways are necessary to mount a successful immune response to different

species of pathogen, and provide the most compelling argument for a major role of the Toll pathway in mediating the antifungal response and for the Imd pathway in mediating the response to Gram(-) bacteria.

The Toll pathway is clearly the major regulator of the antifungal response, since *Spaetzle, Toll*, and *Dif* loss-of-function mutants exhibit a drastic drop in survival (down to 0-20% of wild type) after infection with several types of fungi [36, 39, 44, 52, 58, 66]. The precise role of the Imd pathway in antifungal defense is unclear, since some studies find no effect of *Imd, IKK*γ*,* or *Relish* mutants on fungal survival [36, 52, 58], while some reports do find a slight decrease in fungal survival [42, 44, 66]. Several reports find that flies with mutations in both the Toll and Imd pathways show a greater decrease in survival than Toll mutants alone after infection by the fungi *A. fumigatus*, but not *B. bassiana* [36, 44, 66], indicating that the Imd pathway is involved in antifungal defense, and that different species evoke slightly different responses.

The Imd pathway is clearly the major regulator of the response to Gram(-) bacteria, since *Imd, IKK*γ, and *Relish* mutants are unable to survive infection with *E. coli, E. carotovora,* or *E. cloacae* [36, 42, 44, 52, 58, 66]. Whether or not the Toll pathway is involved is less clear, since several studies have found no effect on survival in *Spaetzle*, *Toll*, and *Dif* mutants [36, 52, 66], while other studies have seen a slight decrease in survival [44, 61]. All studies which have examined Toll/Imd pathway double mutants have found a greater effect of the double mutants than single mutants in the Imd pathway [36, 44, 66], suggesting that the Toll pathway does in fact contribute to the defense against Gram(-) bacteria.

A careful examination of the literature suggests that both Imd and Toll are required to resist infection by most species of Gram(+) bacteria. After infection by *M. luteus*, a Lys-PGN (Toll inducing) Gram(+) bacterium, flies with a mutation in either the Toll pathway or Imd pathway exhibit an 80% survival rate. Only when both pathways are mutated is a drastic decrease in survival rate seen [44, 66]. Thus, either pathway appears sufficient to defend against this pathogen. In contrast, defense against *S. faecalis* and *E. faecalis*, both Lys-PGN Gram(+) bacteria, appears to rely heavily on the Toll pathway, with a minor contribution by the Imd pathway. *Spaetzle* and *Dif* mutants show a drastic decrease in survival, whereas *IKK*β, *Relish*, and *PGRP-LC* show weak to modest decreases [44, 52, 58, 66]; double mutants are worse off than Toll single mutants [66]. Curiously, *B. megaterium*, a DAP-PGN (Imd activating) Gram(+) bacterium, rapidly kills flies harboring mutations in Spaetzle and PGRP-SA (the receptor shown to respond to Lys-PGN) [52], and mutations in PGRP-LC, -LE, and Imd only reduce survival to 50-60% [61]. Clearly, the response to Gram(+) bacteria involves complex interactions of both pathways in a species-specific manner.

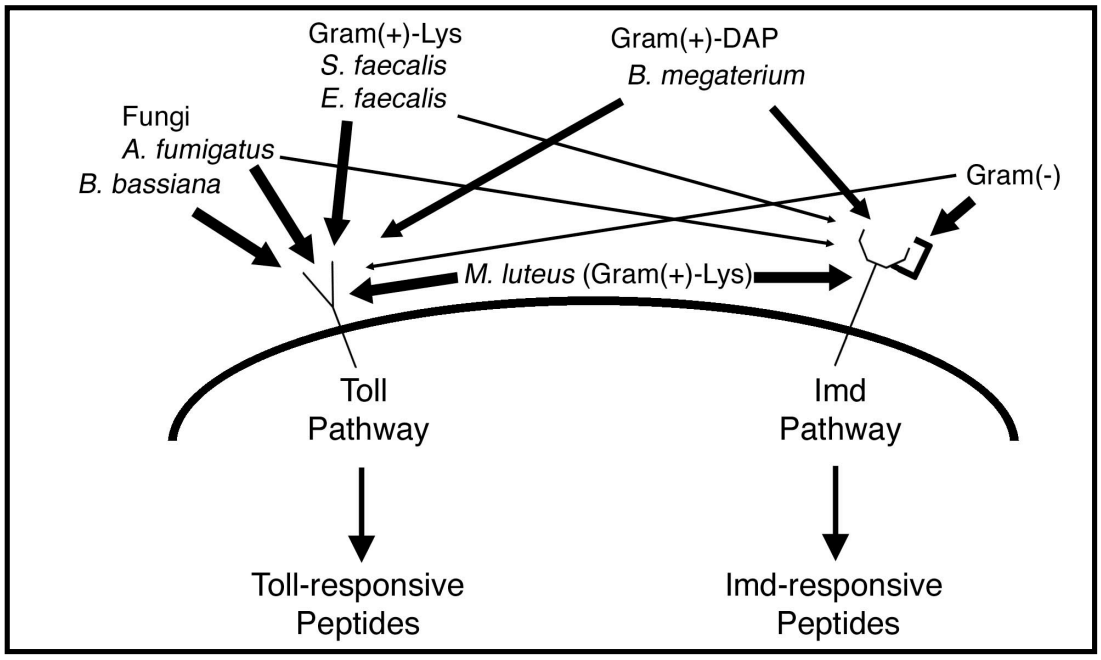

Figure 5: Signaling Pathway Activation by Pathogens

-Antimicrobial Peptide Gene Induction

The ultimate question of antimicrobial peptide induction is: which peptides respond to which pathogens? This question is actually two questions: what signaling pathways are activated by each pathogen, and which peptides respond to the signaling pathways? As described above, it appears very few pathogens activate only one pathway, so one way to simplify this question is to bypass the pathogens and directly activate the signaling pathways.

A constitutively active mutant of Toll, Toll^{10b}, strongly induces Drosomycin (Drs), and moderately induces Metchnikowin (Mtk). Several genes, including Diptericin (Dpt), Attacin A (AttA), Cecropin A (CecA), Drosocin (Dro), and Defensin (Def), respond very weakly to Toll^{10b} [36, 45, 77, 78]. However, one

report employing microarrays found that Dpt, Met, AttA, and Dro responded as strongly to Toll^{10b} as Drs [66]. Taken together, these data suggest that all antimicrobial peptides are capable of responding to the Toll pathway, albeit at low levels, and Drs is the only AMP of this set that is consistently strongly responsive to Toll.

No constitutively active mutant is available for the Imd receptor. However, over-expression of Imd causes a strong induction of Diptericin, Drosocin, and CecropinA. Drosomycin levels, however, are barely above background [41]. Over-expression of PGRP-LE (a receptor upstream of Imd) causes a strong activation of Diptericin, a weak activation of Drosomycin, and an even weaker activation of AttacinA [60]. These studies directly activating signaling pathways suggest that Drs is largely Toll responsive, and most other peptides are largely Imd responsive.

Table 2: Direct activation of signaling pathways differentially induces AMPs

	Imd	PGRP-LE	Toll^{10b}
Diptericin	+++	+++	+
Drosocin	+++		+
CecropinA	+++		+
AttacinA		+	+
Drosomycin	+	+	+++

Note: blank cells indicate lack of experimental evidence

Another way to bypass induction by pathogens is to over-express the transcription factors, Dif, Dorsal and Relish. However, this experiment yields confusing results. Over-expression of Dif, which is downstream of Toll, strongly

activates CecropinA and Diptericin, two peptides which are not very responsive to the Toll pathway. Drosomycin, the only peptide strongly induced by the Toll pathway, is only highly induced by a mixture of Relish and Dif. Two other peptides, Defensin and AttacinA, are induced most strongly by an mixture of Dl and Relish [79]. These data might be artifacts of expressing transcription factors at high concentrations, or might belie additional complexity during the activation of immune peptides by the signaling pathways.

When real pathogens are employed in conjunction with pathway mutations, a more complex picture emerges. Most infections are performed by puncturing the abdomen of flies with a microbe-coated needle, or by injecting a small volume of a concentrated solution of microbes. A major factor complicating all of these studies is the fact that injuring the flies with a sterile needle evokes a response, resulting in low levels of induction of all antimicrobial peptides [36, 45]. In order to properly control these experiments, all inductions need to be compared to the induction caused by sterile wounding, which some studies neglect to do, requiring extra caution during interpretation.

After close examination of the available studies, it appears the antimicrobial peptide genes lie on a spectrum of Toll versus Imd responsiveness. Diptericin lies far on the Imd side, Drs lies far on the Toll side, and most other peptide genes, while close to the Imd side, fall somewhere in between. Metchnikowin lies in the middle.

When infected with a mixture of *E. coli* and *M. luteus*, which should strongly turn on both the Toll and Imd pathways, Diptericin induction is

consistently abolished in *Imd*, *IKKβ,* or *Relish* mutants, and unaffected in *Toll* or *Dif* mutants [36, 40, 43-45, 77, 78, 80]. In contrast, Drosomycin induction is always strongly decreased (but not abolished) in *Toll* or *Dif* mutants [36, 40, 43-45, 77, 78, 80]. The effect of Imd pathway mutants on Drs varies. Sometimes, no effect is seen [40, 44, 80], sometimes a decrease of 20-50% is seen [36, 42, 66, 77, 78]. These experiments suggest that Dpt is almost purely Imd responsive. Drs is largely Toll responsive, but has a larger response to the Imd pathway than Dpt has to the Toll pathway.

The limited data available for another antimicrobial peptide gene, Drosocin, indicate it behaves similar to Dpt [36, 40, 43]

Cecropin A requires the activation both pathways for induction, as mutation of either pathway alone decreases its induction to 0-20% in response to several pathogens [36, 42-44, 61, 77, 78], except in one microarray study, which only found an effect when both pathways were mutated [66].

Similar to CecropinA, full induction of Attacin A and Defensin requires both the Toll and Imd pathways. However, 30-80% of AttA and Def induction persists in Toll pathway mutants, and 10-30% remains in Imd pathway mutants [36, 39, 43, 44, 78]. Thus, a single pathway is able to induce these genes, in contrast to Cecropin, which is only induced upon activation of both pathways. This result for AttA is somewhat surprising, since AttA is not strongly induced by over-expression of PGRP-LE, or by Toll^{10b}. Perhaps PGRP-LE alone is not able to fully activate the Imd pathway, since it normally acts in cooperation of PGRP-LC.

The induction of Metchnikowin presents a puzzle. As stated above, it is the only gene besides Drosomycin that responds significantly to the Toll pathway. However, when Mtk is induced by *E. coli* and *M. luteus,* a Toll or Dif mutation has no effect [39, 77], but *Toll* and *Spaetzle* mutations do reduce the response to *E. coli, M. luteus,* or the fungus *A. fumigatus* when infected separately [66]. Mutations in Imd components always cause a strong decrease in Mtk induction, irrespective of the pathogen [42, 43, 66, 77, 81]. The puzzle is that even though Mtk responds to the Toll pathway, only mutations in the Imd pathway consistently reduce its induction. Curiously, a mutation was isolated, *ird27*, which only reduced the induction of Met, leaving all other peptides unaffected, indicating that Metchnikowin is regulated through at least one unique component.

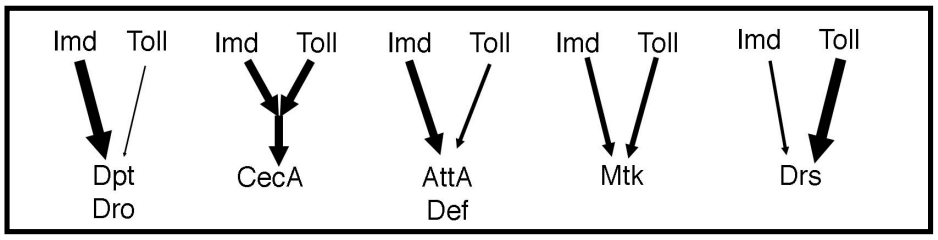

Figure 6: Pathway Induction of Antimicrobial Peptide Genes

-Pathogen-specific induction suggests additional discrimination

The majority of experiments outlined above have defined the framework of the Imd and Toll pathways by studying the immune response to a small group of pathogens. A few studies have begun to examine more species of pathogens.

The available data suggests that within a broad class of pathogen, particular species elicit slightly different immune responses.

E. coli is the species of Gram(-) bacteria most often studied. As mentioned above, survival experiments suggest that the Imd pathway provides the major defense against *E. coli*, with a minor contribution provided by the Toll pathway. This result is confirmed by looking at the induction of Diptericin and Drosomycin by *E. coli*. Dpt is strongly induced by *E. coli* [35] through DAP-PGN [61]; this induction is abolished by mutants in the Imd pathway [43, 45, 61, 66, 81], and Toll pathway mutants have no effect [39, 61]. Drosomycin is reproducibly induced by *E. coli* to levels slightly greater than sterile injury [35, 39, 45]. This induction could result from a weak response to the strong activation of the Imd pathway, or a strong response to the weak activation of the Toll pathway. The available data suggest both are happening [39, 61, 66, 81], although two reports saw no effect of Imd pathway mutations [43, 45]. The reason for this discrepancy is likely the difficulties associated with studying a very weak signal.

A small number of studies have examined other species of Gram(-) bacteria. At least one other species of Gram(-) bacteria, *P. aeruginosa,* elicits an immune response very similar to *E. coli*, activating only the Imd pathway [44, 45]. However, two other species appear to weakly activate the Toll pathway as well. In *Relish* mutants, which leave the Toll pathway as the sole activator, Drosomycin and Attacin A exhibit a modest response to *E. carotovora* and *E. cloacae* [45]. These data suggest that the *Drosophila* immune system recognizes additional distinctions between particular species of Gram(-) bacteria.

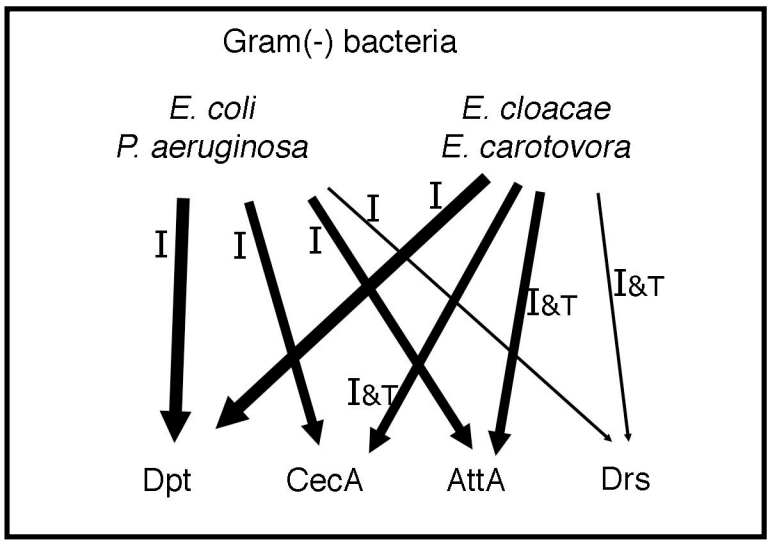

Figure 7: Gram(-) induction of Antimicrobial Peptide Genes. T indicates induction is through the Toll pathway, I indicates induction through the Imd pathway

Several studies have already established that the *Drosophila* immune systems discriminates between different species of Gram(+) bacteria based on the type of peptidoglycan present in the cell membrane. *M. luteus* is the most often studied Gram(+) bacteria. Survival studies suggest that either the Imd or Toll pathway alone is able to defend against *M. luteus* (which has Lys-PGN that is known to activate the Toll pathway). *M. luteus* infection strongly induces Drosomycin [35, 39, 45, 66], but there are conflicting reports concerning the induction of other peptide genes. In two studies, Diptericin (the purely Imd responsive peptide) is not induced above background by *M. luteus* [35, 39], but in two other studies, it is strongly induced [45, 66]. Similar conflicting data is seen for other more Imd responsive peptides, such as Cecropin A and Attacin A, sometimes they are induced by *M. luteus* [45], and sometimes not [35]. The reason for this discrepancy is unclear. *S. aureus*, another Lys-PGN bacterium,

appears to induce antimicrobial peptide genes in a manner similar to *M. luteus* [45].

The DAP-PGN (Imd activating) Gram(+) bacteria, such as the *Bacillus* genus, strongly induce Diptericin and other Imd-responsive peptides, and weakly or moderately induce Drosomycin [35, 45]. This induction is absent in *Relish* mutants [45], indicating that induction of Dpt, AttA, CecA, and Drs by DAP-PGN occurs mainly through the Imd pathway. Toll activation of these bacteria has not been extensively studied. The discrimination of Gram(+) bacteria based on the Lys versus DAP peptidoglycan is sufficient to explain the available data.

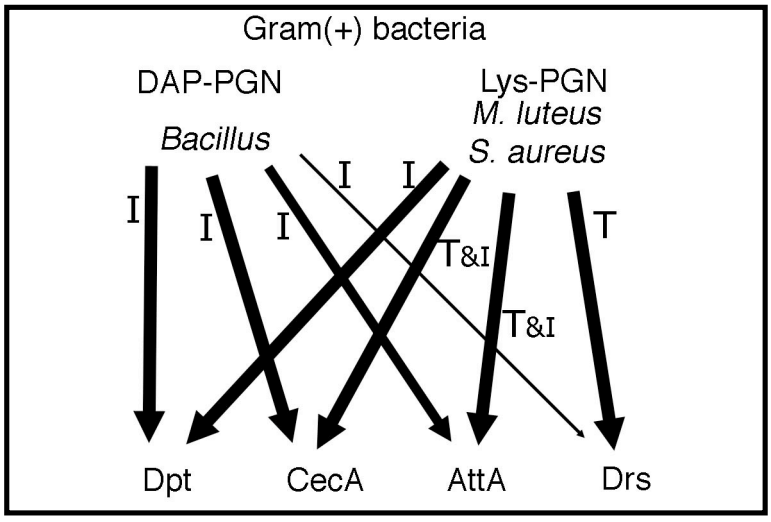

Figure 8: Gram(+) bacteria induction of Antimicrobial Peptide Genes

The immune response to fungi appears to change dramatically depending whether the infection occurs through a septic injury, or through an injury-free infection, i.e., coating flies with fungal spores. Drosomycin is always strongly induced by *B. bassiana* in a Toll-dependent manner [35, 36, 39, 44, 45], and Metchnikowin is moderately induced through Toll [35]. The other more Imd-

responsive peptides, such as Diptericin, Attacin A, and Cecropin, are only induced when the fungi are injected [35, 45], not when the infection occurs naturally [35, 39, 44]. Comparing different species, natural infection by *B. bassiana, M. anisopliae,* and *P. fumoroseus* strongly induces Drosomycin, but fails to induce Diptericin and Cecropin A [35]. Metchnikowin is moderately induced by two of those species, *B. bassiana*, and *M. anisopliae*. *A. fumigatus*, however, only weakly induces Drosomycin through natural infection [35].

When these fungi are injected into adult flies, a much different response is seen. *B. bassiana*, *M. anisopliae*, *G. candidum*, and *S. cerevisiae* all strongly induce the more Imd-responsive peptides, Diptericin, AttacinA, and CecropinA, as well as Drosomycin. The response of Diptericin to two of these fungi, *M. anisopliae* and *S. cerevisiae*, is abolished in *Dif* or *Relish* single mutants, suggesting that Diptericin induction by these fungi requires both the Toll and Imd pathways. This is surprising, since most data suggests that Dpt barely responds to the Toll pathway. The response to the other two fungi is abolished in *Relish*, not *Dif,* mutants, indicating that those two species only activate Dpt through the Imd pathway. Curiously, the response of all four peptides to one species, *G. candidum* is entirely Dif-independent, suggesting it only activates the Imd pathway. In light of these studies, it is very likely that fungi contain additional species-specific molecular patterns which are detected by the *Drosophila* immune response [45].

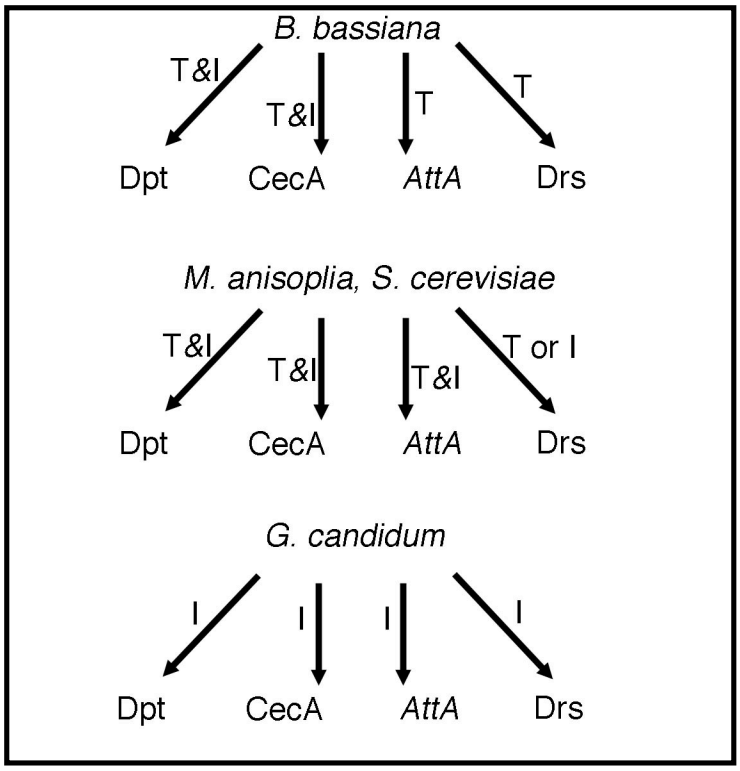

Figure 9: Induction of Antimicrobial Peptide Genes by Fungal Injection.

-Locations of the immune response

The difference in the immune response between a natural fungal infection and injection of fungi could arise from another level of complexity in the *Drosophila* immune response: the location of the stimulus and response. Most of the studies detailed above address the systemic immune response, in which antimicrobial peptides are synthesized in the fat body and secreted into the hemolymph. However, recent studies have revealed complex mechanisms of tissue specific expression. In addition to the systemic antimicrobial peptide induction, several tissues display both constitutive and inducible AMP expression.

The systemic response, originating in the fat body, requires Serpent (dGATAb), a member of the GATA family of transcription factors. Serpent works in conjunction with NF-kB members to regulate fat body specific expression [82, 83].

The local induced response is independent of the Toll pathway, even for Drosomycin, and instead utilizes only Imd pathway [84, 85]. Expression in the midgut has been show to require another GATA family member, dGATAe [86]. The three other GATA family members exhibit different patterns of expression, and are likely involved in tissue-specific regulation as well [86]. However, GATA factors may only be active in larvae, not adults [82], suggesting that other mechanisms may be at work in adults.

The constitutive expression of immune peptides in the absence of any infection results from complex, tissue and gene-specific regulation. For example, Drosomycin is always present in high levels in the salivary glands and female reproductive organs; Cecropin A, Defensin, and Drosocin are also present in reproductive organs; and Metchnikowin is constitutively expressed in the anteno-maxillary organ [87]. Curiously, this expression of Drosomycin is independent of both the Toll and Imd pathways [84, 87]. The constitutive expression of Drosomycin and Cecropin A depends instead on Caudal, a homeotic transcription factor. Caudal is only involved in the local constitutive expression in certain tissues; the systemic induced response, as well as the local induced response, are independent of Caudal [88].

Careful experimentation will be necessary to fully understand the mechanisms underlying tissue specific regulation of antimicrobial peptide induction.

-Regulatory Motifs

One of the clues which originally led to the discovery of role of the Toll pathway in the immune response was the presence of NF-κB binding sites in the promoters of immune responsive genes. The importance of these κB motifs for immune responsive loci was initially demonstrated in Cecropin A [89] and Diptericin [90]. Several other potential regulatory motifs, including the Interleukin 6 Response Element, the Interferon Consensus Response Element, and the Hepatic Nuclear Factor 5 Response Element, and GAAANN motifs were also identified [91, 92], suggesting that multi-protein complexes are likely involved in the regulation of antimicrobial peptides. However, only κB motifs by themselves are sufficient to mediate an immune response, and other regulatory elements likely modulate the strength of the response [93], or the location of the response. For example, GATA motifs are necessary, but not sufficient, for tissue-specific induction in the fatbody and midgut of larvae [82, 83, 86], and many immune peptides contain κB and GATA motifs arranged in similar orientation and spacing [83]. The R1 element, a putative regulatory motif found in the promoter of Cecropin family and Defensin genes, was also shown to be necessary for the induction Cecropin in adults and larvae [94].

Of particular importance are the observations that full induction of immune peptide genes often requires multiple κB motifs. The Diptericin promoter contains two identical κB motifs. Mutation of either site abolished Imd induction of reporter constructs [90, 93], suggesting that the response to the Imd pathway is achieved through cooperative interactions between multiple sites.

Since κB motifs are the only motifs sufficient for immune responsiveness, the NF-κB transcription factors, Dif and Relish, are obviously of paramount importance for the immune response. Since Dif is downstream of the Toll pathway, and Relish is downstream of the Imd pathway, yet Toll and Imd induce different subset of immune peptides, a logical conclusion is that Dif mediates Toll responsiveness through Dif-specific κB motifs, and Relish mediates Imd responsiveness through Relish-specific κB motifs. However, data supporting that conclusion is lacking.

Extracts from immune induced cells contain a binding activity capable of binding κB motifs from CecropinA [89] and Diptericin [90], which provided some of the initial proof that NF-κB proteins were involved in the *Drosophila* immune response. This binding activity is disrupted by α-Dif sera [37, 95], and purified Dif is able to bind to the κB motif in CecA [37]. There is clearly some specificity in binding to different κB motifs, the motif from CecA and the motif from DptA are bound with different intensity by extracts of immune activated cells [96]. Interestingly, the binding activity to the CecA κB sequence can be competed by a mammalian κB motif [89], and purified human p50 can bind to the Dpt κB motif

[90], demonstrating the similarity between mammalian and *Drosophila* NF-κB proteins.

A SELEX selection assay, a high-throughput assay designed to identify the range of motifs bound by DNA-binding proteins, demonstrated that Dif binds to a highly degenerate site, displaying little specificity, and Relish and Dorsal bind to very similar motifs. Relish/Dif heterodimers displayed a binding preference different from each homodimer [83]. As is the case in mammalian studies, high-throughput assays are unable to identify consensus motifs for individual family members. Given the flexibility of NF-κB members and the ability to bind different sequences, this result is not surprising. The question remains, how do the Toll and Imd pathways direct the expression of specific subsets of antimicrobial peptides?

-The Role of Dif and Dl in the Immune Response

The exact role of Dorsal in the immune response is also unclear. Dif and Dorsal are closely related, sharing 48% amino acid residue identity [37], and Dif can largely substitute for Dl during DV patterning, showing only slight defects in the expression of some target genes [97]. Dif and Dl are located very close to each other, approximately 7kb apart from each other on chromosome 2 [78], and this proximity has made isolating mutations in Dif difficult. As mentioned above, *Dorsal* mutants do not show any defects in AMP induction [36, 96]. However, when *Dif*, *Dorsal* double mutants are examined, Dif or Dl can rescue AMP induction in larvae [78], but only Dif can rescue induction in adults [38].

The binding specificity of the two proteins also shows overlap. Dif can bind to the κB motif in the Zen promoter, a known Dorsal target gene, with an affinity similar to Dl [37]. However Dif, but not Dl, binds to the κB motif from CecA [37, 96]. Dif can also bind to one or two copies of the DptA κB motif, but Dl can only bind to two copies [96]. Furthermore, the binding activity of extracts from immune induced cells for the CecA motif is only disrupted by α-Dif sera, not α-Dl sera [37, 95].

When the proteins are over-expressed in cell culture, Dl is a stronger inducer than Dif of a DptA construct, but Dif activates a CecA construct more strongly than Dl [95, 96]. Dif and Dl can heterodimerize *in vitro* [96], and combining Dl with Dif reduces Dif's ability to activate the CecA reporter [95], suggesting that Dif/Dl heterodimers are less active than Dif homodimers.

The available data do not clearly define the roles of Dif and Dorsal. Some of their functions appear to overlap, and some do not. It is fairly certain that Dif is the major regulator of the immune response. However, based on the data at hand, a distinct role for Dorsal in the immune response cannot be ruled out.

Summary

A solid groundwork has been laid for understanding the differential activation of immune genes by microbial pathogens. The components of the Toll and Imd signaling pathways, the major regulators of the humoral immune response, are firmly established. Understanding of the exact microbial features

that activate these pathways through pathogen recognition receptors is advancing rapidly. A thorough understanding of the *cis*-regulatory code governing the expression of targets of the Toll and Imd pathways is the missing piece necessary to fully understand the differential regulation of immune loci.

In this dissertation is presented the first evidence that κB motifs are able to direct specific transcriptional responses to the Toll and Imd pathways in *Drosophila*. Furthermore, this specificity is determined by the specific binding of Dif and Relish to κB motifs in the promoters of target genes. A bioinformatic approach was used to identify the features of κB motifs unique to Toll and Imd responsive genes. Experimentally, these features are shown to determine Toll and Imd responsiveness in a cell-culture model, and determine Dif and Relish binding through *in vitro* gel-shift assays.

An examination of several immune responsive promoters, revealed that Toll responsiveness of κB motifs is largely independent of surrounding promoter context, however Imd responsiveness requires a permissive context. Multiple κB sites interact in a cooperative manner, resulting in a cooperative (non-additive) increase in expression. This Cooperativity can occur in two forms: first, if multiple κB motifs responsive to the same pathway are present in a promoter, a cooperative response is seen upon activation of a single pathway. If at least one κB motif responsive to each pathway is present, then simultaneous activation of both pathways results in a cooperative response. Furthermore, the role that cooperative interactions between multiple motifs play in the regulation of immune loci is examined.

Finally, a brief investigation into the role of Dorsal in the immune response in the cell-culture system suggests that Dorsal has a distinct role in regulating immune responsive genes that is non-redundant with Dif.

II. Results

Establishment of A Cell Culture Based Luciferase Assay

A Transfected Attacin A Construct Faithfully Reports Toll and Imd Signaling

To explore how pathway responsiveness is encoded at the DNA level, a molecular genetic analysis of a single locus was used as the foundation for a global bioinformatic approach. Attacin A (AttA), an antimicrobial peptide gene, is responsive to both Toll and Imd signaling [35, 45, 66, 98]. In cultured *Drosophila* Schneider (S2*) cells, which express all three fly NF-κB proteins and mediate robust Toll and Imd responses [79, 99], the transcriptional response of an AttA reporter construct was assayed in response to innate immune signaling (Figure 10A).

To stimulate the Toll pathway, epidermal growth factor (EGF) was applied to cells expressing EGFR-Toll, a chimera fusing the extracellular and transmembrane domains of the human EGF receptor to the intracellular domain of Toll [67]. To specifically induce the Imd pathway, a peptidoglycan (PG) preparation from *B. subtilis* was used [53, 62]. The AttA reporter construct exhibited a robust response to either inducer, with 5-11 fold activation upon EGF treatment and a 12-28 fold increase upon exposure to PG.

To further validate the cultured cell system, RNA interference (RNAi) was used to inactivate components of the Toll or Imd signaling cascades. RNAi

against genes in the Toll pathway – *MyD88*, *tube*, or *Pelle* – specifically blocked induction by EGF, but not PG (Figure 10B). Similarly, inactivation of Imd pathway components – *Imd*, *Tak1*, or *key* (IKKγ) – eliminated induction by PG, but not EGF. AttA induction by either innate immune pathway was strictly dependent on endogenous NF-κB factors (Figure 10C). For the Toll response, Dif and Dorsal had overlapping function, with either factor alone being sufficient for some signaling. For the Imd response, Relish alone was necessary and sufficient. The S2* cell system thus effectively recapitulates endogenous regulation of the AttA gene.

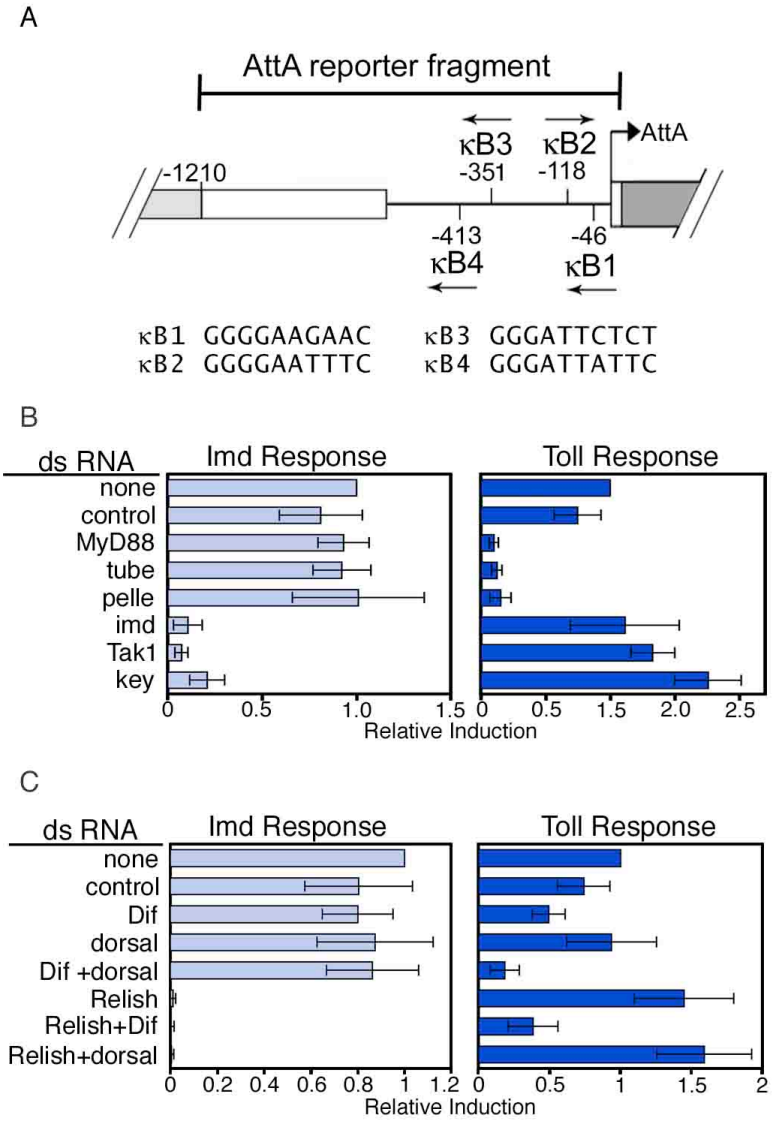

Figure 10: An AttA Reporter System Recapitulates Endogenous Innate Immune Responses. (A). AttA Reporter Construct. The AttA genomic fragment extending from the 3' UTR of the Drosocin gene (white box) to the AttA transcriptional start site was fused to the luciferase gene (bar). The locations of four potential κB motifs [83, 92], numbered from proximal to distal relative to the start site, are indicated. Arrows designate the orientation of each κB motif, with the forward sequence shown below. (B). RNAi against Toll and Imd pathway components, followed by EGF or *B. subtilis* PG stimulation. EGFR-Toll cells were treated with the indicated dsRNA and transfected with the AttA reporter. For comparison, cells were incubated without dsRNA (none) or treated with dsRNA for *Easter* (control), which acts upstream of Toll in embryonic patterning. Values were normalized to the induction measured without dsRNA and are each the average of at least four independent experiments. Capped lines indicate standard deviation. (C). Role of fly NF-kB proteins in AttA regulation. EGFR-Toll cells were treated singly or in combination with dsRNA for *Dif*, *dorsal*, or *Relish*; transfected with the AttA reporter; and subjected to Toll stimulation (EGF) or Imd stimulation (PG). Controls and analysis were as in (B).

Promoter Proximal κB Sites Govern AttA Induction by the Toll and Imd Pathways

The transcriptional start site for Attacin A lies approximately 650 bp downstream of the 3' end of the neighboring transcription unit (Figure 10A). Within the intergenic region lie four potential κB sites [83, 92]. As is characteristic of κB motifs, the 5' half-site in each case contains either GGG or GGGG.

Using site-directed mutagenesis to inactivate or reposition the κB-related motifs, motif function in directing expression of the AttA reporter gene was assayed. Inactivation of pairs of sites revealed that only the κB motifs at positions -46 and –118 were necessary for activation by either Imd or Toll (Figure 11A). Furthermore, each of these two sites, κB1 and κB2, preferentially mediates signaling by one innate immune pathway (Figure 11B). A construct containing one copy of κB1 responded more strongly to Imd than to Toll. Introducing a second copy of κB1 caused a greater-than-additive increase in Imd induction, indicating that multiple sites can act together in a cooperative manner, as seen previously in the Diptericin promoter [90, 93]. Likewise, a single copy of κB2 directed a stronger response to Toll than to Imd, and two copies of κB2 yielded a cooperative increase in Toll induction.

These data provide the first evidence that a sequence of a single κB site can determine the signaling pathway to which a *Drosophila* innate immune gene responds.

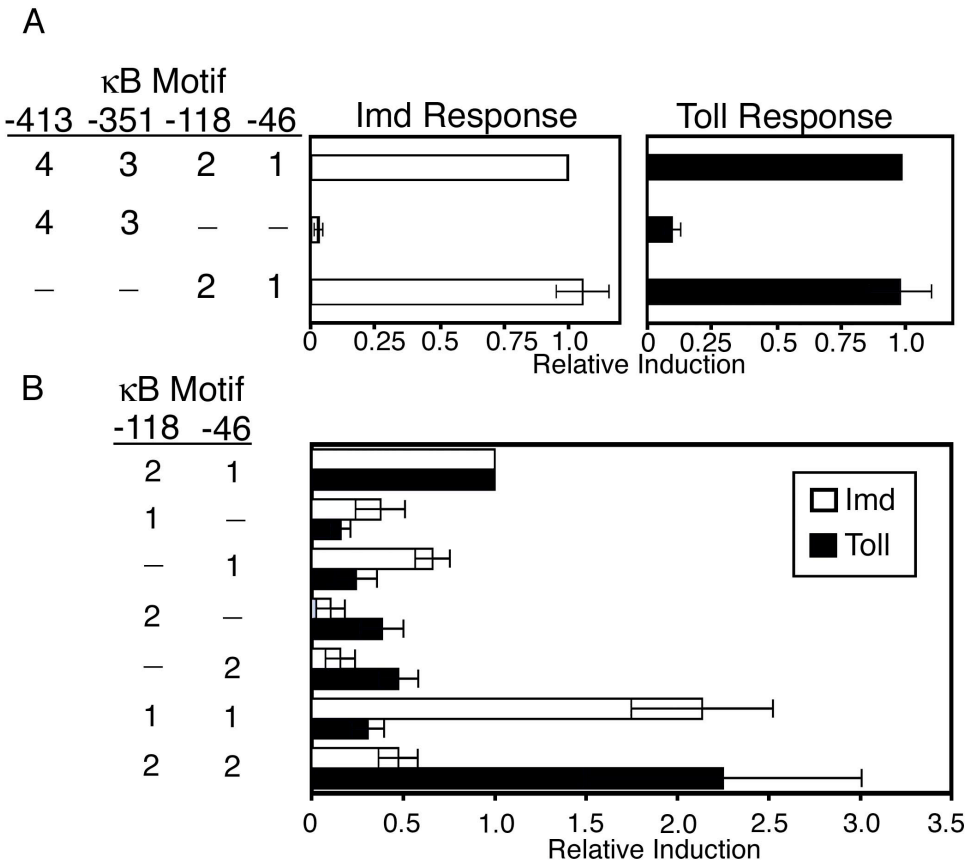

Figure 11: κB Sites Determine Pathway Specific Transcriptional Responses. (A) Proximal κB sites govern AttA induction. Mutational inactivation of κB motifs was achieved by converting both the second and third G residues in the 5' core element to C residues. AttA reporter constructs in which pairs of κB motifs were inactivated (-) were transfected into EGFR-Toll cells. Following Toll or Imd pathway stimulation, the induction of the reporter construct was measured and normalized to the wild-type level. Data are the average of at least six independent transfections. (B) Effects of κB motif sequence, context, and number on Toll and Imd induction of AttA. Site-directed mutagenesis was used to inactivate (-) or replace κB sites. All values are normalized to the wild-type construct (top row).

Bioinformatic Analysis

Toll and Imd Responsive Loci Differ in κB Site Structure and Sequence

To establish whether there is a general κB sequence code for innate immunity, a bioinformatics analysis was conducted using published microarray

data sets for innate immune responses in wild-type and mutant strains of *Drosophila* [66, 100]. To identify loci responsive to Toll but not Imd, parameters were set to screen for strong induction by fungal infection and by a constitutively active Toll receptor, as well as a dependence on a functional Toll pathway for induction by bacterial infection. Similarly, genes specifically responsive to Imd were identified by screening for genes that responded robustly to bacterial infection only in the presence of a functional Imd pathway, but were not appreciably induced by fungal infection. Using quantitative expressions of these criteria to define screening algorithms (see Experimental Procedures), 16 Toll responsive genes and 11 Imd responsive loci were identified. Good agreement was found between these two gene sets and a classification based on clustering of temporal expression patterns [80], eleven of the Toll loci and nine of the Imd loci being common to both analyses.

To narrow the focus of the analysis, the assumption was made that the κB sites relevant to innate immune induction would typically lie within 200 bp upstream of the transcriptional start site, as for AttA and as reported for rapid response genes from both flies and mammals [89, 90, 101, 102]. Extracting the corresponding genomic sequence for each locus (see Supplementary Figure 1), the MEME motif discovery program [103] was used to search for overrepresented sequence motifs.

The results of the bioinformatic analysis were striking. For both the Toll and Imd responsive gene sets, MEME analysis identified a κB-type sequence as

the highest scoring motif (Table 4). The Toll and Imd gene sets were decidedly different, however, with regard to κB site composition and number.

Table 3: Results of MEME Motif Analysis.

Innate Immune Genes			Properties of κB Motifs			Consensus κB Motifs		
Gene Set	# Genes in Set	# κB Sites in Set	5' $(G)_4A$ Half-Site	$(A/T)_{2-3}$ Central Region	$(A/T)_{4-5}$ Central Region	Consensus κB Motif From MEME Analysis	Mean Mismatch To Toll Set	Mean Mismatch To Imd Set
Toll	16	17	0%	6%	94%	G G G A A A C C C T/G T	1.1 ± 0.6	3.0 ± 1.0
Imd	11	21	62%	52%	14%	G G G G A T T T T T C C C	3.7 ± 0.8	1.4 ± 1.0

Percentages indicate the fraction of putative κB sites in each gene set containing the indicated sequence element. Central elements are defined as the sequence beginning with the first A after two or more consecutive G residues and ending with the last A or T before the first C residue. Mismatches between each consensus and the members of each gene set are reported as the average plus or minus the standard deviation.

Whereas nearly two-thirds (62%) of κB motifs in the Imd set contained a GGGGA 5' half-site, such a half-site was absent from the κB motifs in the Toll set. The Toll κB motifs instead typically had either GGGA (12 examples) or GGAA (three examples) as the 5' half-site (Table 5). Moreover, the difference in half-site sequence was not an accident of motif definition during MEME analysis. A scan of the entire 3.2 kb sequence space comprising the upstream regions for the Toll gene set detected only a single example of GGGGA or its reverse complement, TCCCC; a parallel scan of the smaller (2.2 kb) sequence space for the Imd upstream regions detected 15 such instances.

Divergence between the potential κB sites in the Toll and Imd gene sets extended throughout the motifs. Representative Toll-responsive κB motifs had

four or five bases between the G cluster and the first C residue, e.g., GGGAAAACCC. Conversely, those Imd elements containing strings of G's and C's typically had a two or three base separation, e.g., GGGGATTCCT. Statistically, these differences were marked. Overall, a 4-5 bp (A,T)-rich region separated G's and C's in 94% of the Toll motifs, but only 14% of the Imd motifs. Similarly, we found a 2-3 bp (A,T)-rich region separating G's and C's in 52% of the Imd motifs, but only one of the Toll motifs.

In approximately half of the Imd motifs, the 3' half-site diverged significantly from a canonical κB motif. In place of two or more C residues, the 3' half-site of these motifs consisted largely or entirely of a string of T residues, e.g., GGGGATTTTT. Studies in mammalian systems have demonstrated that there are motifs of this type, i.e., having only a single cognate κB half-site, that can nevertheless bind specifically to Rel proteins in vitro and exhibit cis-regulatory activity in vivo [see, e.g., 101].

Toll and Imd Responsive Loci Differ in κB Site Number

The Toll and Imd responsive gene sets differed not only in κB site sequence, but also κB site number. Of the fourteen Toll genes for which κB sites were detected, twelve had only a single presumptive κB site and none had more than two sites. In contrast, eight of the nine Imd genes with predicted κB sites had two or more sites.

Table 4: Innate immune motifs

Toll Responsive Loci				Imd Responsive Loci			
#	Name		κB Sequence	#	Name		κB Sequence
1	Nec		GGGAAAAGCCC	1	CecA1	a	GGGATTTCA
2	Ser7		GGGAAAACCC			b	GGGGATTTTT
3	CG4757		GGGTAAAGCCC	2	CecA2	a	GGGGATTTTT
4	CG5778		GGGAATTGCCC			b	GGGGATTTTT
5	CG5791		CGGAAAACCCC			c	GGGATTTTT
6	CG11841		GGGAAAACCT	3	CecC	a	GGGGATCCAC
7	GNBP		GGGAAATTCAC			b	GGGGAACTTTT
8	IM23	a	GGAAAACCCC	4	Def	a	GGGAGTCCC
		b	GGGATTCGCC			b	GGGGTTTTTA*
9	CG15067		none	5	CG2056		none
10	CG16705		GGGTTTGCCC	6	PGRP-SD	a	GGGGATTTT
11	CG16772		GGGAAATTCC			b	GGGGAGTGCCC
12	CG18067		GGGAATTGCC	7	AttD	a	GGGAGTCCC
13	IM2	a	GGGAAAACAC			b	GGGGAAATCAC
		b	GGAAATTCC	9	PGRP-SB1		GGGGTATTTA*
14	IM1		GGGAAAACAC	10	DptB	a	GGGATTCCTT
15	IM10	a	GGGAAATTCT			b	GGGGAATCTCA
		b	GGGAATTTCT			c	GGGATTCCCA
16	CG30080		none			d	GGGGATCCAC
				10	Dpt	a	GGGGATTCCT
						b	GGGGATTCCT
						c	GGGGATCTTT
				11	TotM		none

Motifs identified by MEME analysis for Toll and Imd responsive loci. The designations a, b, and c refer to multiple motifs identified for a single locus; none indicates the absence of a predicted motif for a given locus. Asterisks indicate sequences that gave negative results in binding studies (see text).

A κB Sequence Code Governs *Drosophila* Innate Immunity

Confirmation of Bioinformatic Analysis in Cell Culture

To determine whether the observed differences between the κB sites in the Toll and Imd gene sets correspond to a *cis*-regulatory code, the cultured cell system was again employed. In the AttA reporter construct, the -118 site was

inactivated, and synthetic versions of the κB consensus motifs were introduced at the -46 position. Responses to innate immune signaling were then assayed.

The consensus κB motifs defined by the MEME analysis directed the expected pathway specific transcriptional responses in the context of the AttA reporter (Figure 12). Upon induction of Toll signaling, the Toll consensus sequence, GGGAAAACCC, directed robust reporter gene expression. Indeed, this site responded at a level an order of magnitude greater that that observed with either one of the two endogenous sites in AttA. In contrast, the two Imd consensus sites – GGGGATCCCC and GGGGATTTTT – were only induced approximately ten percent as much as the Toll consensus upon Toll induction; but mediated a significant response to Imd. In this case the response was at a level comparable to that detected with a wild-type AttA site. These data support the conclusion that the consensus sites defined as overrepresented motifs in the two genomic sequence sets mediate pathway specific innate immune responses in cells.

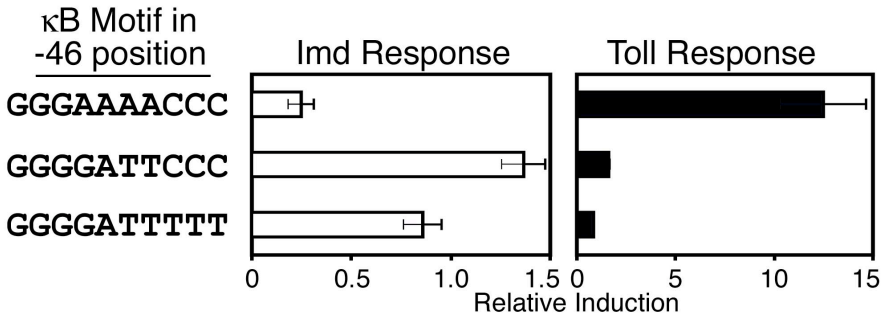

Figure 12: Imd and Toll responsiveness of consensus motifs identified by MEME. The ¯118 site was inactivated in the AttA reporter and the ¯46 site was replaced with synthetic sites. Induction for the Toll consensus motif (top row) was normalized to that for κB2 alone; induction by either of the two Imd consensus motifs (2nd and 3rd rows) was normalized to that for κB1 alone.

The κB Sequence Code Functions by Specifying NF-κB Protein Binding

There are at least two distinct mechanisms by which a κB sequence code could dictate specific transcriptional responses. First, NF-κB proteins activated by different pathways could vary significantly in their affinity for particular κB motifs. Available data implicate this mechanism in the IKKβ-independent NF-κB response in humans [17]. Second, a single κB sequence could bind a range of NF-κB proteins indiscriminately, but motif sequence could determine which co-activators interact with bound NF-κB proteins. Such a model can explain the regulation of several mammalian genes by pairs of κB motifs [12].

In exploring which mechanism is operative in flies, a gel shift assay was used to determine whether or not NF-κB proteins differ in their target site binding preference. Relish and Dif were expressed from pGEX vectors in bacteria, purified by affinity chromatography, and used in binding assays with labeled oligonucleotides.

The gel shift assays demonstrated that the Relish and Dif GST fusion proteins bind synthetic κB motifs in vitro with a specificity identical to that observed in vivo for signaling by Imd and Toll, respectively (Figure 13A). Using GST-Relish, a strong gel shift signal was observed for the Imd consensus motifs GGGGATTCCC and GGGGATTTTT, but not to the Toll consensus sequence, GGGAAAACCC. Results with GST-Dif were the inverse: strong binding to the Toll consensus site, but not the Imd consensus motifs.

The correlation between interaction seen in the gel shift assay and pathway specificity in vivo held true not only for synthetic κB sites, but also for κB sites found upstream of innate immune loci (Figure 13B). For example, GST-Relish gave a strong gel shift signal with the κB site from an Imd-specific locus, Diptericin, but had no detectable interaction with the κB site from a Toll-specific gene, IM1. Furthermore, GST-Relish appeared to interact to a much greater extent with the -46 motif from AttA than with the -118 motif, consistent with the specificity observed in cells. GST-Dif exhibited complementary specificity, binding to the greatest extent with the IM1 motif and the -118 motif from AttA. When the Dpt and IM1 motifs were introduced into the AttA promoter, their pathway responsiveness correlated with transcription factor binding. The Dpt κB responded in a manner almost identical to the AttA-46 motif, and the IM1 motif responded very similarly to the Toll consensus motif (data not shown).

Given the consistent relationship between pathway responsiveness and Rel protein binding specificity, gel shift analysis was extended to all 38 of the sites identified by the MEME program. Of the sixteen κB sites identified in the Toll gene set, fifteen had significant binding to GST-Dif and none bound to GST-Relish. The results with the Imd motifs were likewise very clear. Of the 18 potential κB sites in the Imd gene set, fifteen bound well to Relish and none bound to Dif.

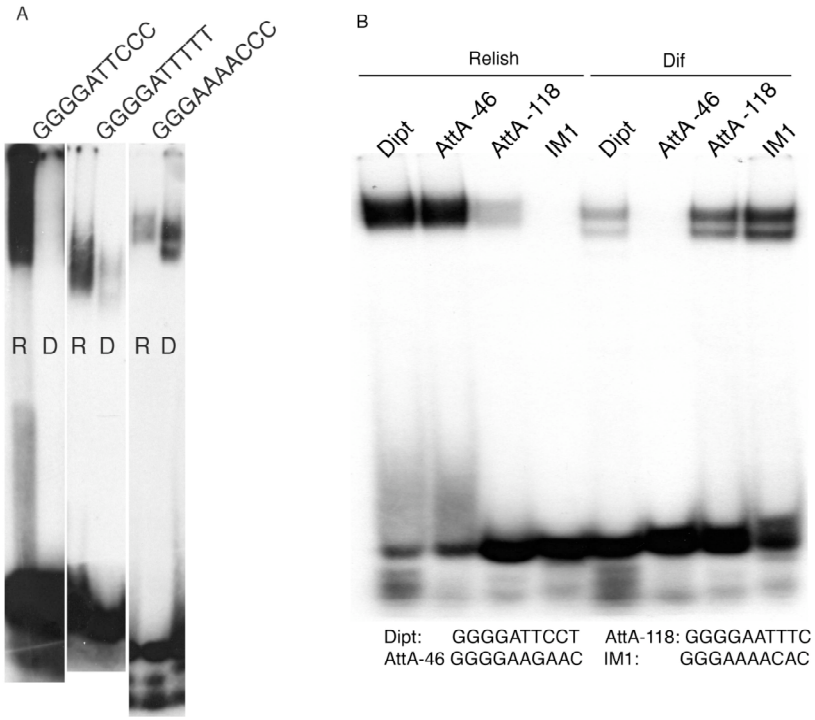

Figure 13: Dif and Relish Exhibit Selective Binding to Synthetic and Endogenous κB sites. GST-tagged Dif or Relish was incubated with labeled oligonucleotides and the resulting nucleoprotein complexes resolved by native polyacrylamide electrophoresis. (A) Consensus κB sites. R indicates a lane with Relish, D indicates a lane with Dif. (B) Endogenous κB sites.

Determinants of Dif and Relish Binding

After establishing that Dif and Relish binding determines Toll and Imd responsiveness, respectfully, the next step was to explore further the features of κB motifs that determine binding and pathway responsiveness. For example, it is not clear why the Dipt (GGGGATTCCT) and AttA-46 (GGGGAAGAAC) motifs were bound by Relish, whereas the AttA-118 motif (GGGGAATTTC) was bound by Dif. All three motifs contain a GGGGA sequence, which was a feature common to the motifs in Imd responsive genes, yet the GGGGAATTTC motif is preferentially bound by Dif and is Toll responsive. In preparation for this exploration, it is first necessary to establish conventions to facilitate discussion

(Figure 14). This first string of guanine nucleotides will be referred to as the 5' G cluster, the last string of cytosine nucleotides will be referred to as the 3' C cluster, even if this cluster sometimes contains bases other than cytosine. The region in between the last G of the 5' cluster and the first C of the 3' cluster will be referred to as the A/T central region. Note that these labels are not to be confused with half-sites, which are an alternate way to describe κB motifs.

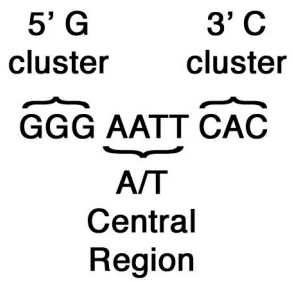

Figure 14: Naming conventions used to describe features of κB motifs.

The MEME analysis suggested that Imd responsive motifs often contain a short, 2-3 bp A/T central region, in contrast to Toll responsive motifs, which often contain a 4-5 bp A/T region. Experiments were performed to test whether spacing between G and C clusters could explain why the sequence GGGGAATTTC was preferentially bound by Dif and was Toll responsive.

The starting point was a Relish/Imd specific sequence, GGGGATCCCC. Bases were added in between the G and C clusters to lengthen the A/T central region (Figure 15). When the A/T central region reached a length of four nucleotides, the motif became dual responsive: bound by both Dif and Relish, and induced by both the Toll and Imd pathways. Adding one additional base caused a dramatic drop in Relish binding and Imd induction, creating a Dif

specific motif. These data demonstrate that the spacing between the 5' G cluster and the 3' C cluster plays a crucial rule in determining transcription factor binding and pathway responsiveness.

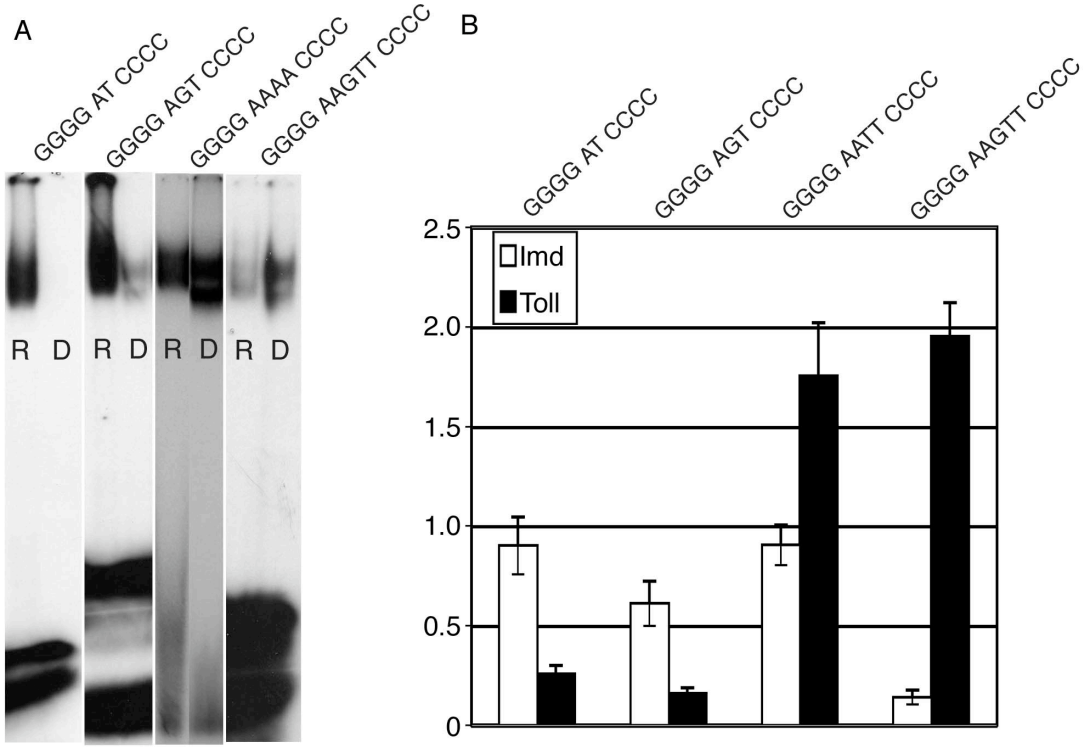

Figure 15: Spacing is one feature that determines specificity. (A) Gel shift assays with Gst-Relish (R) and Gst-Dif (D). (B) Luciferase assays of κB motifs in the –46 position of the AttA reporter. Imd and Toll inductions were normalized to the wild-type AttA inductions.

Next, the examination turned to the binding and pathway responsiveness of κB motifs with three-G clusters (Figure 16). Similar to four-G motifs, Dif binding and Toll induction were only seen when the A/T central region was at least four nucleotides in length. However, the motifs with three G's yielded a much stronger

response than motifs with four G's to the Toll pathway (compare 14B to 15B), indicating that a three-G motif provides a stronger response to Toll than a four-G motif.

Curiously, the only three-G motif that was bound strongly by Relish was GGG AATT CCC. This motif also responded to the Imd pathway as strongly as the wild-type AttA reporter. It seems Relish can only bind to a three G motif if it is a perfect palindrome, since the motif GGG AAAA CCC is not bound by Relish (Compare Figure 16 and 13). Also, any deviation from GGGAATTCCC resulted in a loss of Relish binding (data not shown).

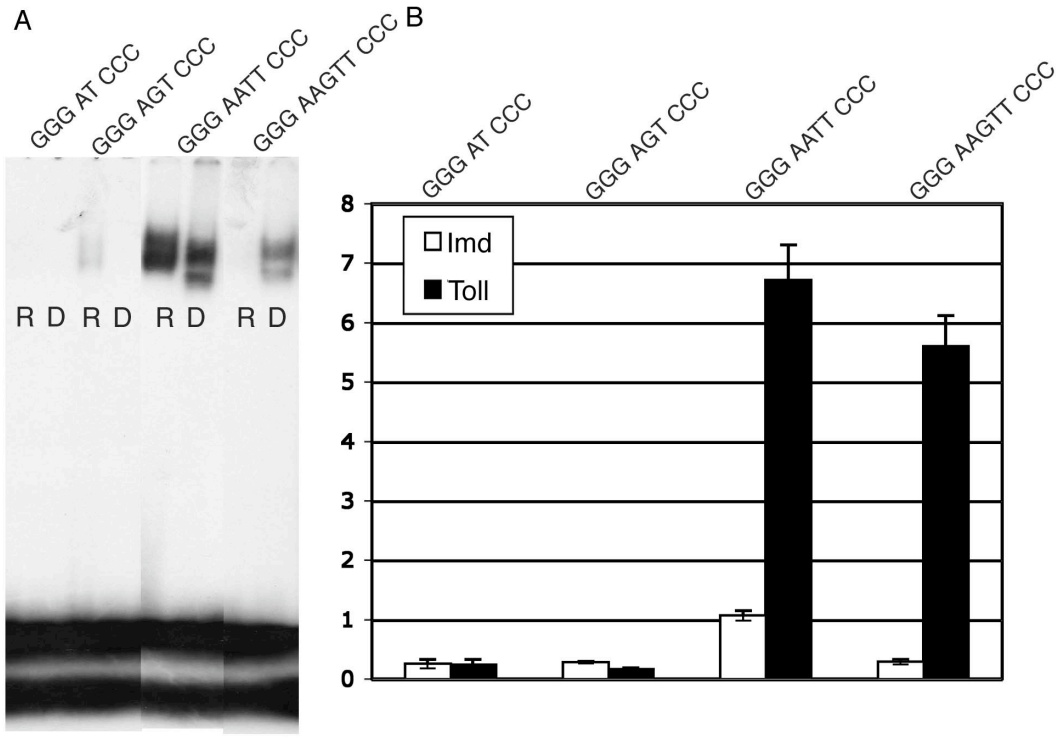

Figure 16: Spacing determines specificity in three G motifs. (A) Gel shifts with Gst-Relish (R) and Gst-Dif (D). (B) Luciferase assays of κB motifs in the –46 position of the AttA reporter. Imd and Toll inductions were normalized to the wild-type AttA inductions.

A Single GGGGA Half-site is Sufficient for Relish Binding

In addition to palindromic κB motifs, many Imd responsive genes contain motifs lacking a 3' C cluster. As previously demonstrated, this second consensus motif was bound by Relish and mediated Imd responsiveness, warranting a more in-depth analysis of this alternative consensus. Motifs with one GGGGA half site can exhibit wide variation in the 3' half of the motif and still be bound by Relish (Figure 17A). However, certain sequences were not bound by Relish (Figure 17B). No pattern common to the bound and unbound motifs was readily apparent.

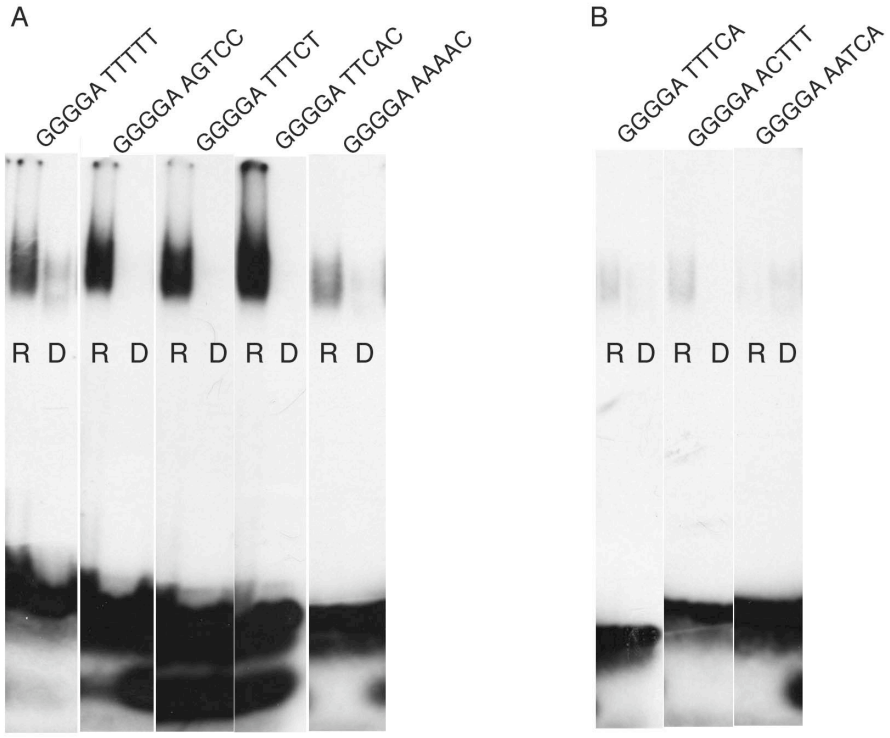

Figure 17: Gel shift assays with Gst-Relish (R) and Gst-Dif reveals that a single GGGGA half-site is often sufficient for Relish binding. (A) Motifs that bind Relish. (B) Motifs that do not bind Relish.

To elucidate further the role of a second half-site, investigation proceeded into how many changes were necessary to disrupt binding to a palindromic 3' motif. Deviations from the palindromic sequence were introduced into the second half-site, and the effect on binding was observed (Figure 18). The presence of a G in place of the 2nd or 3rd C of the 3' half-site disrupted binding, but a G in the 1st or 4th position had little effect. The presence of two consecutive G's, regardless of position, completely disrupted binding. This disruption could result from either the absence of a necessary C, or the presence of an inhibitory G.

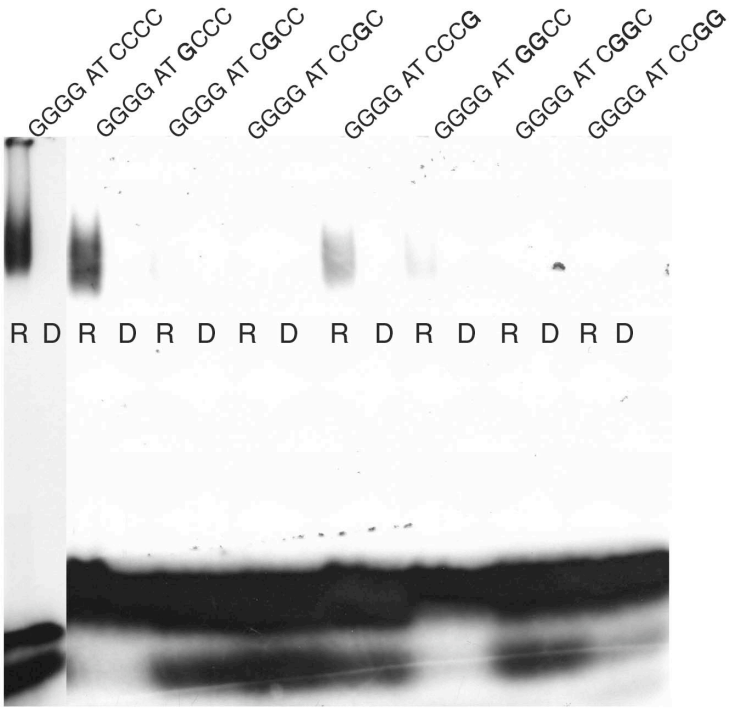

Figure 18: Specific changes in the 3' half-site disrupt binding. Gel shift assays with Gst-Relish (R) or Gst-Dif (D) and oligonucleotides contains variation in the 3' half site of a palindromic motif.

Analysis of Additional Reporters

To determine if the context of the promoter surrounding κB motifs contributes to pathway responsiveness, additional reporters were subjected to the same analysis as AttA. Since AttA is dual responsive, reporters that were Toll specific or Imd specific were sought. Two Toll specific reporters were characterized: Drosomycin, an antimicrobial peptide gene; and IM1, a gene of unknown function that is strongly induced by the Toll pathway (Figure 19). Completely Imd-specific genes were hard to find; all reporters that were strongly Imd responsive maintained a low level of Toll Induction. AttacinD, another antimicrobial peptide of the Attacin family, exhibits a strong Imd response and weak Toll was response (Figure 19), and was selected to represent Imd-responsive genes. It is somewhat surprising that Drs does not respond to *B. subtilis* PGN, since data in the literature suggest that Drs is weakly activated by other DAP-PGN bacteria. However, other reports have found that Drs responds weakly to injection of whole *B. subtilis* [61], but not to injection of peptidoglycan purified from *B. subtilis* [62]. A possible explanation is that S2* cells are lacking some factor necessary for Drs alone to respond to the Imd pathway. This hypothesis is supported by the observation that in a different cell line, S2-C cells, Drs exhibits a weak, but reproducible induction by *B. subtilis* (CP Arnold, personal communication).

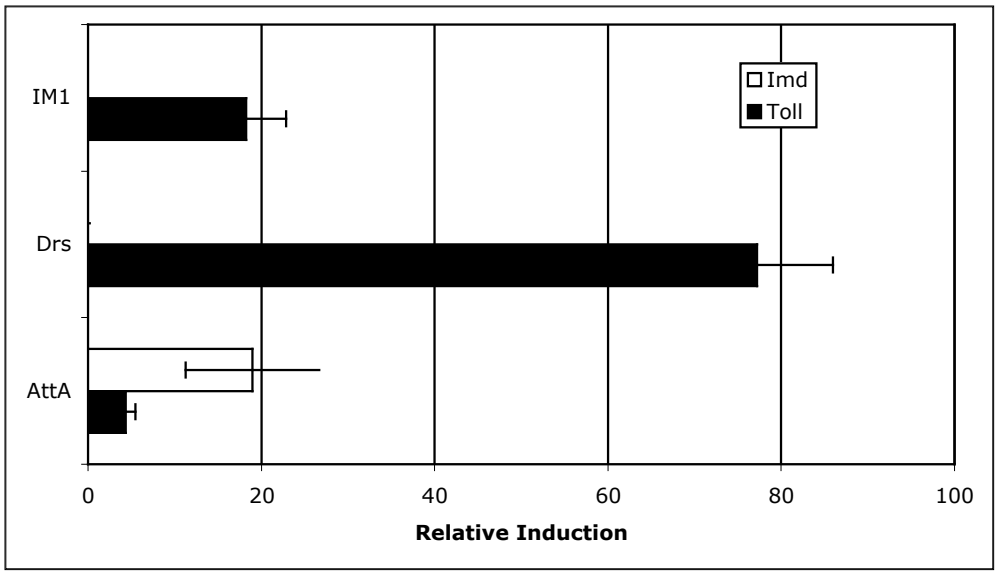

Figure 19: Toll and Imd Inductions of Additional Luciferase Reporters. Induction of each reporter is presented relative to the uninduced level of that reporter.

Analysis of κB Motifs in the Drosomycin Reporter

The 2400 bp region upstream of the Drosomycin gene contains five possible κB motifs. Since this Drosomycin reporter does not respond to the Imd pathway, only data concerning the Toll response is reported. Mutational analysis revealed that two of the κB motifs contribute to the Toll induction of this reporter (Figure 20). Although mutation of the motifs at –9 and –38 yields a slight decrease in Toll induction (Figure 20B), neither is bound by Dif or Relish (Figure 20C), suggesting they do not contribute significantly to regulation. Curiously, the motif at –139 is bound weakly by Relish, yet yields a significant decrease in Toll induction when mutated. The –303 motif is the critical motif, since mutating it eliminates the Toll response, and it is strongly bound by Dif.

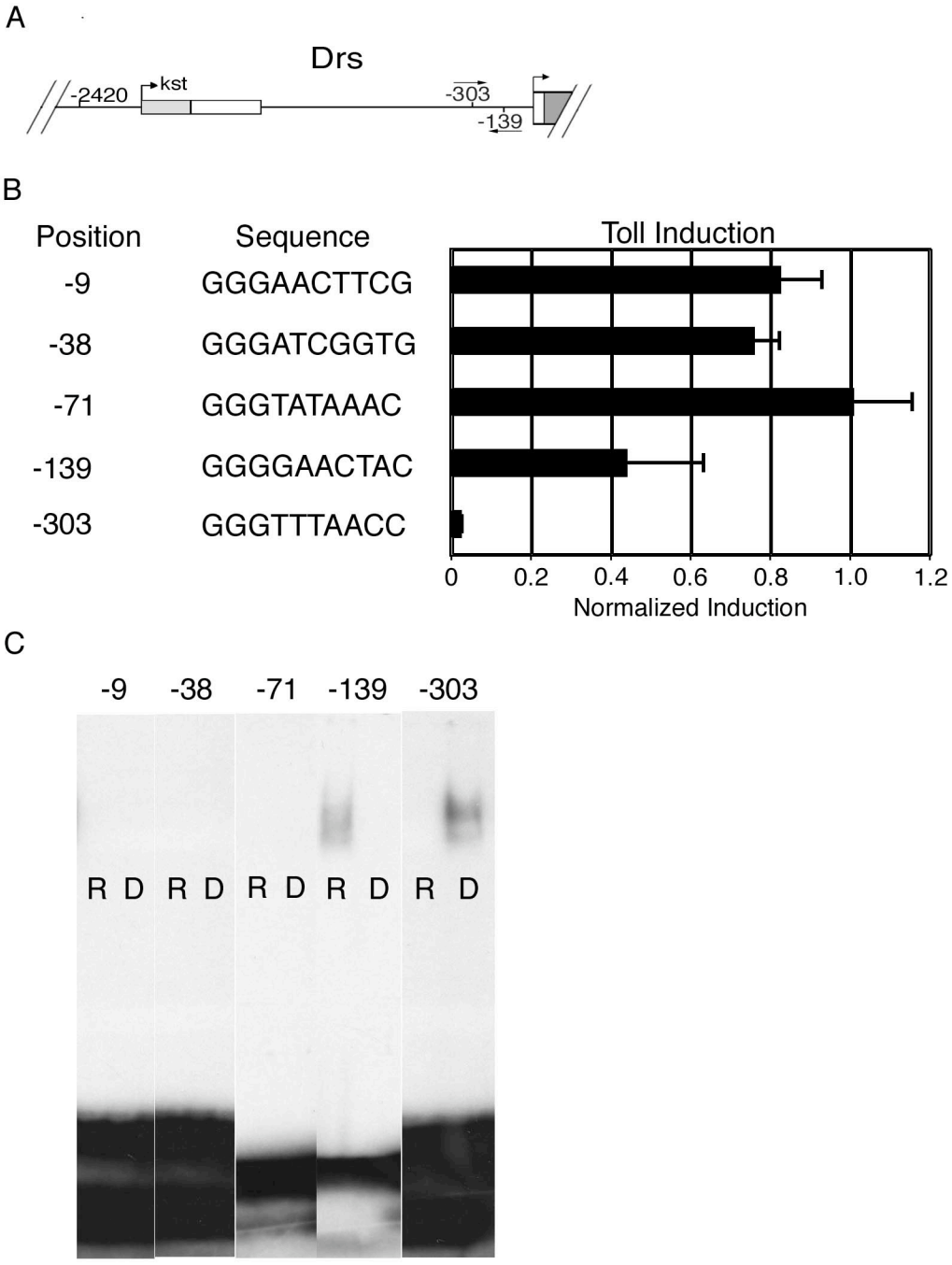

Figure 20: κB motifs in the Drosomycin reporter. (A) Diagram of the Drs reporter, showing the position of the functional motifs. The reporter contained the ORF of another gene, *kst* oriented in the same direction as Drs. (B) Luciferase assay of potential κB motifs in the Drs promoter. Inductions were normalized to the wild-type Drs Induction. (C) Gel shift assays of potential κB motifs with Gst-Relish (R) and Gst-Dif (D).

Analysis of κB Motifs in the IM1 Reporter

The 2000 bp region upstream of the IM1 gene contains only one potential κB motif, at –86. This motif is responsible for the Toll induction of this reporter and is bound only by Dif (Figure 21).

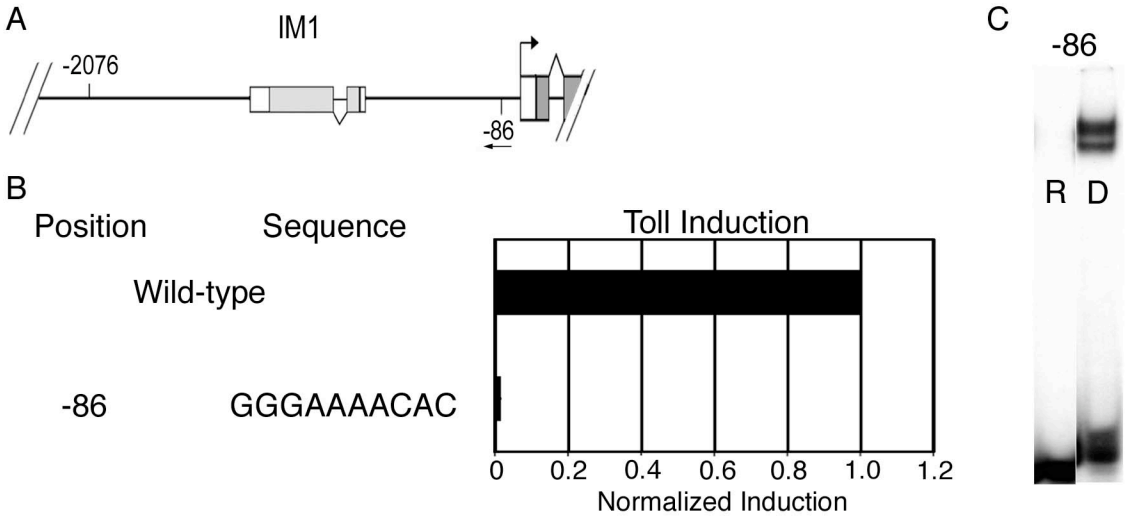

Figure 21: The κB motif in the IM1 reporter. (A) Diagram of the IM1 reporter, showing the position of the functional motif. (B) Luciferase assay of potential κB motifs in the IM1 promoter. Inductions were normalized to the wild-type IM1 Induction. (C) Gel shift assay of the κB motif with Gst-Relish (R) and Gst-Dif (D).

Analysis of κB Motifs in the Attacin D Reporter

The 2000 bp region upstream of the AttD gene contains four potential κB motifs, three of which are functional (Figure 22). Mutation of any of the three motifs reduces both the Toll and Imd inductions, although the –89 motif makes a stronger contribution to the Toll response, and the motifs at –155 and –189 make stronger contributions to the Imd response. The –89 motif is modestly bound by

Relish and weakly by Dif, whereas the –155 and –189 motifs are both bound strongly by Relish.

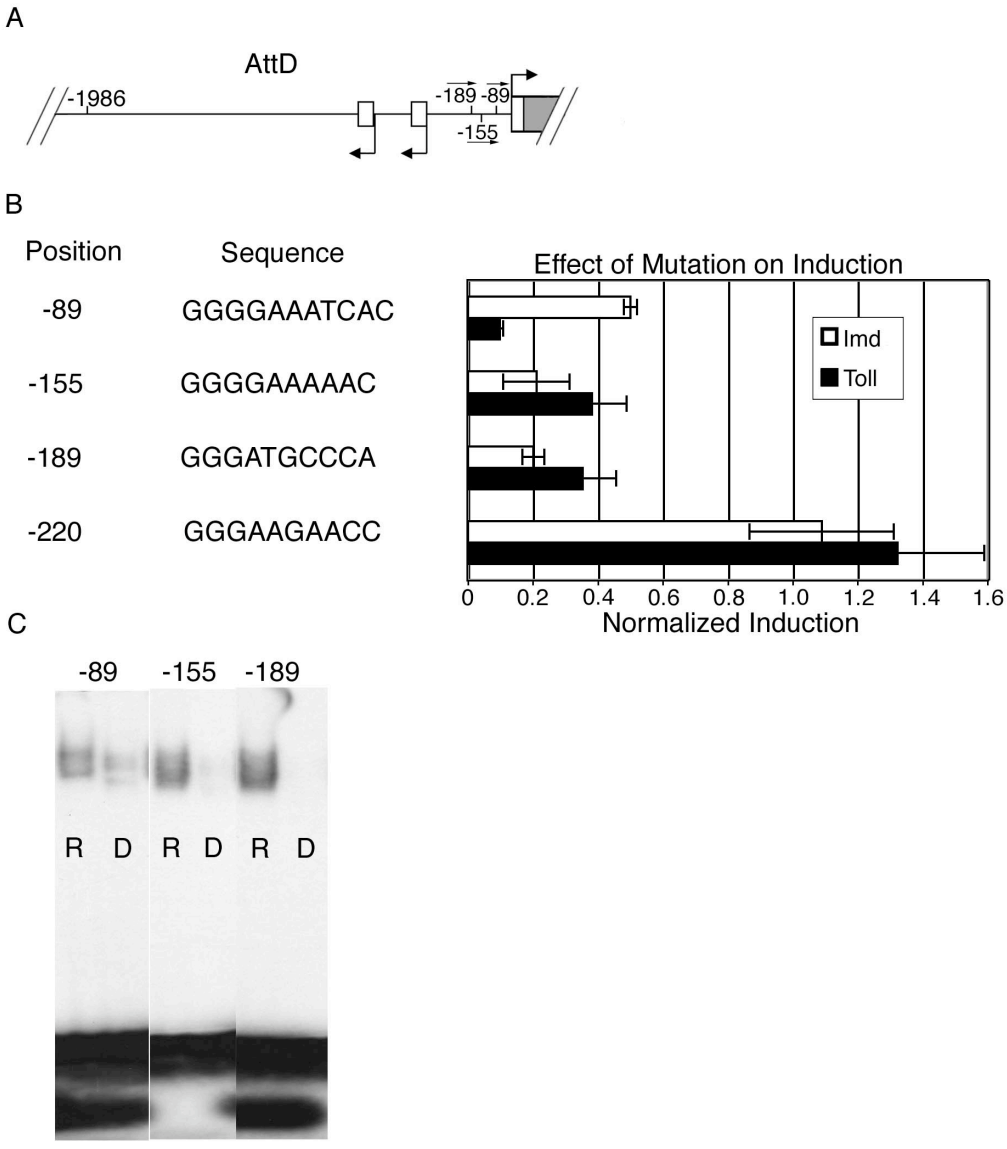

Figure 22: κB motifs in the Attacin D reporter. Two tRNA sequences were located just upstream of the AttD start. (A) Diagram of the AttD reporter, showing the position of the functional motifs. (B) Luciferase assay of potential κB motifs in the AttD promoter. Inductions were normalized to the wild-type AttD Induction. (C) Gel shift assays of functional κB motifs with Gst-Relish (R) and Gst-Dif (D).

Imd Induction of Pathway-Specific, Synthetic κB Motifs in Additional Promoter Contexts

To determine if promoter context affects the pathway responsiveness encoded in κB motifs, synthetic motifs were introduced to the Drs, IM1, and AttD reporters. As shown previously in the AttA context (Figure 15,16), a Relish specific motif, GGGGATCCCC, specifically responded to the Imd pathway; a Dif specific motif, GGGAAGTTCCC, responded specifically to the Toll pathway; and a motif bound by both Relish and Dif, GGGAATTCCC, responded to both the Imd and Toll pathways. This correlation was tested in the context of other immune responsive promoters by replacing endogenous κB motifs with synthetic κB motifs, and assaying pathway responsiveness.

In the additional contexts, the correlation between transcription factor binding and pathway response was maintained, although there were effects specific to each promoter (Figure 23). The two motifs bound by Relish and responsive to the Imd pathway in the AttA context, GGGAATTCCC and GGGGATCCCC, were also Imd-responsive in the Drosomycin and Attacin D contexts. The strength of the response relative to the wild-type induction varied between the two reporters, probably as a result of the starting strength of the wild-type induction. In contrast, The Dif/Toll specific motif never yielded a significant Imd response.

Interestingly, when the two Rel/Imd motifs were placed into the IM1 context, neither was capable of generating an Imd response. This result suggests

there is some difference in the IM1 promoter that renders it unresponsive to the Imd pathway.

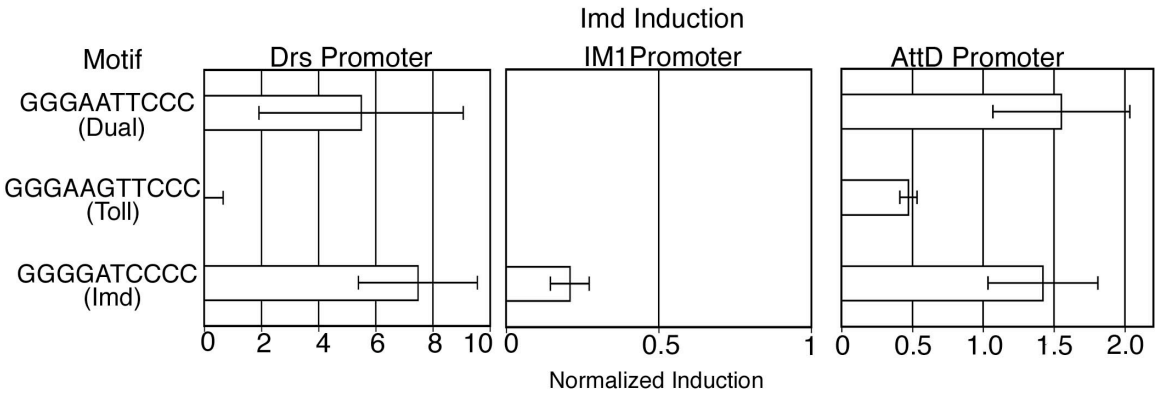

Figure 23:Imd inductions of synthetic κB motifs in additional promoter contexts. In each promoter, one κB motif was replaced with the indicated synthetic motif. In Drosomycin, the –303 site was inactivated, and the –139 site was replaced; in IM1, the –86 site was replaced; and in AttD, the –155 site was replaced. Inductions are presented normalized to the wild-type response of each reporter.

Toll Induction of Pathway-Specific, Synthetic κB Motifs in Additional Promoter Contexts

Similar to the correlation between Relish binding and Imd induction, the correlation between Dif binding and Toll induction also held true in all three additional promoter contexts (Figure 24). The only difference between the three motifs was the strength of the Toll response relative to the wild-type response. Similar to the Imd response, this difference likely results from differences in the wild-type induction of these promoters (Figure 20), and from the number of κB motifs in the promoter. This concept will be discussed in more detail later.

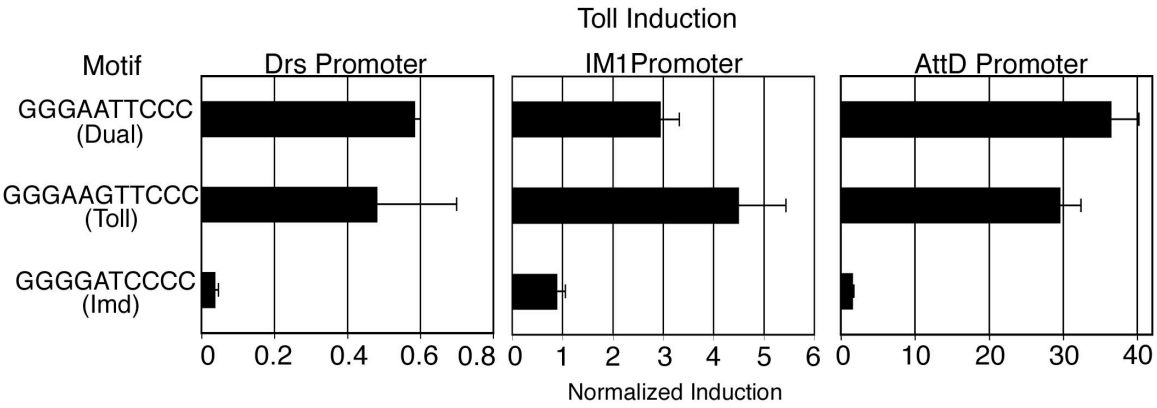

Figure 24: Toll inductions of synthetic κB motifs in additional promoter contexts. In each promoter, one κB motif was replaced with the indicated synthetic motif as in Figure 22. Inductions are presented normalized to the wild-type response of each reporter.

Multiple κB Motifs Function Cooperatively to Regulate Transcription

Several pieces of data suggest that multiple κB sites can act cooperatively in a non-additive fashion. First, three of the reporters, AttA, AttD, and Drs, contain multiple κB sites. Mutation of each site singly often causes a significant drop in induction (see Figures 12, 21, 23). For example, in the AttA context, mutation of the –118 motif (leaving the –46 motif as the sole κB motif regulating induction) results in a response to the Imd pathway that is 60% of wild-type. Similarly, when the –118 motif is the only κB motif present, the Imd induction is 10% of wild-type. If multiple motifs interacted in an additive manner, than the inductions resulting from a single motif should add up to 100%. However, 60% plus 10% yields 70%, indicating there is a non-additive effect of multiple motifs. Second, two copies of either the –46 or –118 motifs of AttA yield inductions by the Imd or Toll pathways that are multiplicatively larger than a single copy (Figure 12). This interaction of multiple sites was explored further using synthetic motifs

in the AttA reporter. In contrast to the endogenous –46 motif, two copies of the Relish/Imd motif, GGGGATCCCC, yield an Imd response which is additive, not cooperative (Figure 25). For both the endogenous and synthetic motifs, two copies elevate the Imd response to a level roughly twice that of the wild-type AttA reporter, raising the possibility that this is the maximum Imd induction possible for this reporter.

The Dif/Toll motif, GGGAAGTTCCC, yields an enormous Toll response when present in two copies, reaching a level approximately six times as great as a singly copy of the motif (Figure 25). Even when present as a single copy, the Toll response is five times greater than the wild-type AttA response. These data suggest that the Toll pathway possesses a higher capacity for Cooperativity than the Imd pathway, which will be discussed in more detail later.

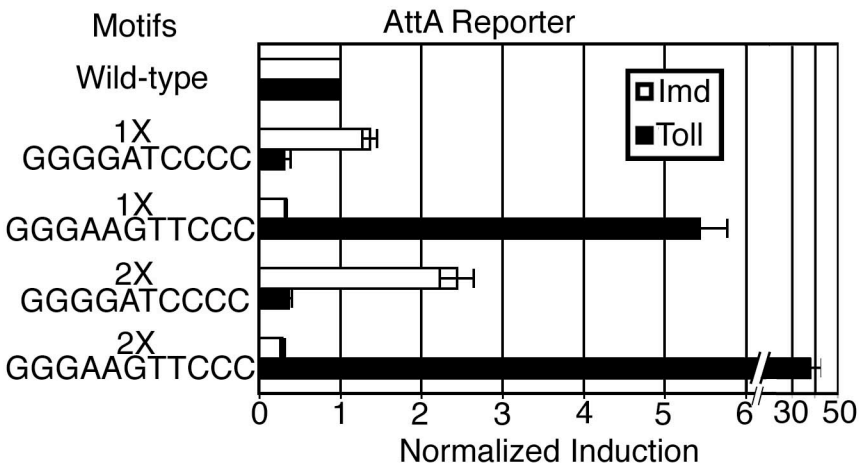

Figure 25: Imd and Toll inductions of multiple synthetic motifs in the AttA reporter. In constructs with one substitution, the –46 motif was replaced. In constructs with two motifs, both the –46 and –118 motifs were replaced. Inductions were normalized to the wild-type AttA induction.

Examination of a Two-Pathway Cooperative Response

A similar Cooperativity was seem upon simultaneous stimulation of both the Imd and Toll for reporters containing multiple κB motifs - Drs, AttD, and AttA - but not for IM1, which only contains a singly κB motif (Figure 26).

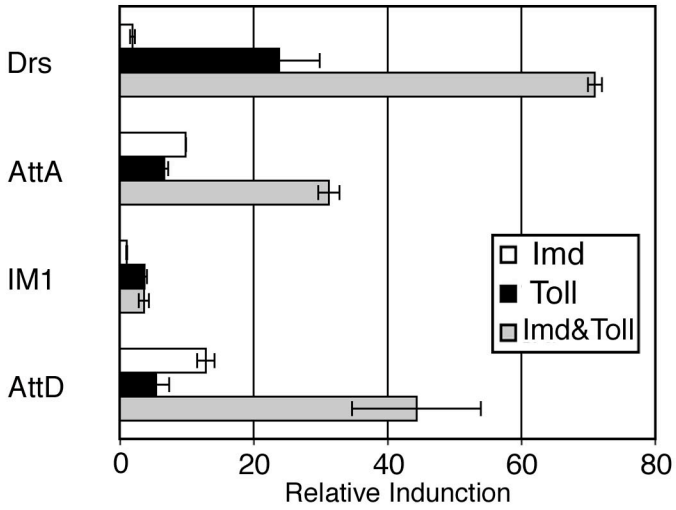

Figure 26: Activation of both the Imd and Toll pathways induces a cooperative response. Inductions are relative to the uninduced level of each reporter.

Two hypotheses could explain the Cooperativity between two pathways. First, Cooperativity could arise from the interaction of different NF-κB homodimers bound to two different κB motifs, similar to the Cooperativity seen when two Toll or Imd specific motifs are present in a reporter. Second, two-pathway Cooperativity could result from the formation of Dif/Relish heterodimers that bind to a single site and boost induction. To test these hypotheses, the effect of mutating a single site in the AttA and Drs promoters was observed. Both of these promoters contain one motif bound by Relish and one motif bound by Dif. Mutation of either site eliminated the cooperative response to both pathways,

indicating that two motifs bound by two different homodimers are required for a cooperative response (Figure 27).

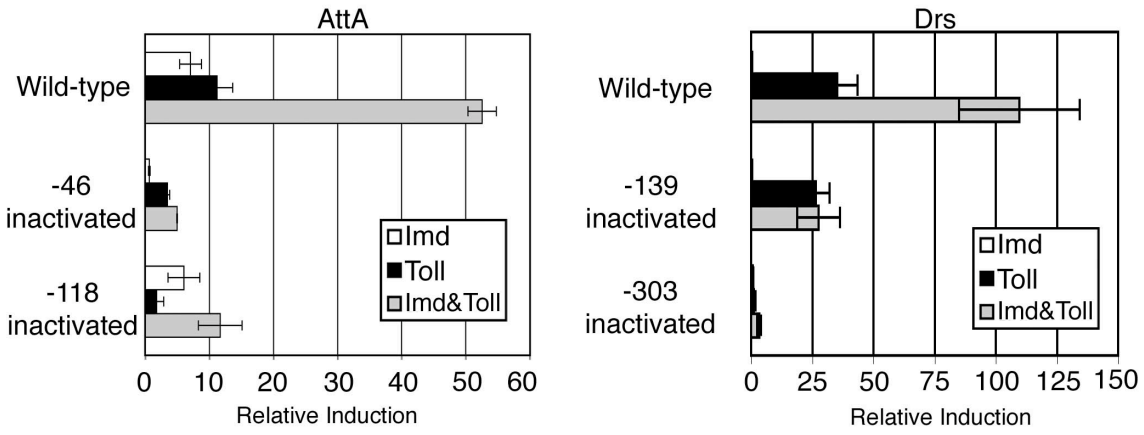

Figure 27: Two κB motifs are required for a cooperative response to both pathways. Inductions are relative to the uninduced level of each construct.

Since two κB motifs were required for Cooperativity, the next test was to determine if two different NF-κB proteins were also required for Cooperativity between pathways. Using RNAi to knock-down two of the three NF-κB factors at a time revealed that a single factor was capable of responding to a single pathway, but was unable to produce a cooperative response to both pathways (Figure 28).

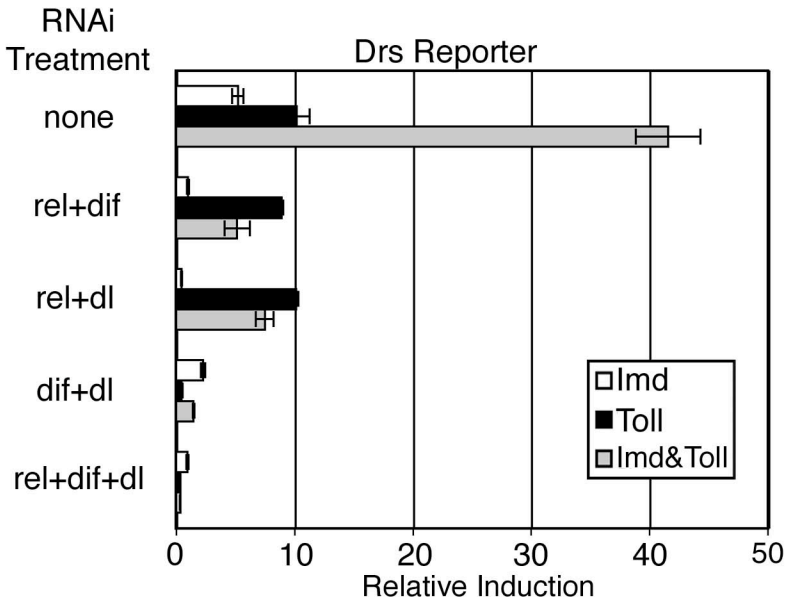

Figure 28: A Cooperative response to both pathways requires two transcription factors. RNAi was used to knock-down two of the three NF-κB factors, and the effect upon pathway induction was observed. Inductions are relative to the uninduced level of each treatment.

Taken together, these results suggest that the cooperative response to simultaneous activation of both signaling pathways results from cooperative interactions between Dif homodimers bound to one motif and Relish homodimers bound to another motif.

Dorsal Contributes to the Regulation of AttA and Drs, But Not IM1

As mentioned in the introduction, Dorsal appears to play a role in regulation of the immune response that is redundant with Dif in some settings. RNAi was used to investigate the involvement of Dorsal in regulating the AttA, Drs, and IM1 luciferase reporters. As seen previously in Figure 10, RNAi against

Dl does not reduce the Toll response of AttA significantly. However, RNAi against Dif alone only reduces Toll induction to 50% of wild-type, and RNAi against both Dif and Dl reduces the Toll response to 20% of wild-type, suggesting that Dl is involved in the Toll response.

Upon examination of other immune genes, Dorsal was found to play a similar role in the regulation of Drosomycin, but not in the response of IM1 to the Toll pathway (Figure 29). For IM1, RNAi against Dif alone reduces the Toll response to 20% of wild-type, and RNAi against Dif and Dl shows no additional decrease. A possible explanation is that AttA and Drs contain κB motifs that can be bound by Dif and Dl, but the sole κB motif in IM1 can only be bound by Dif.

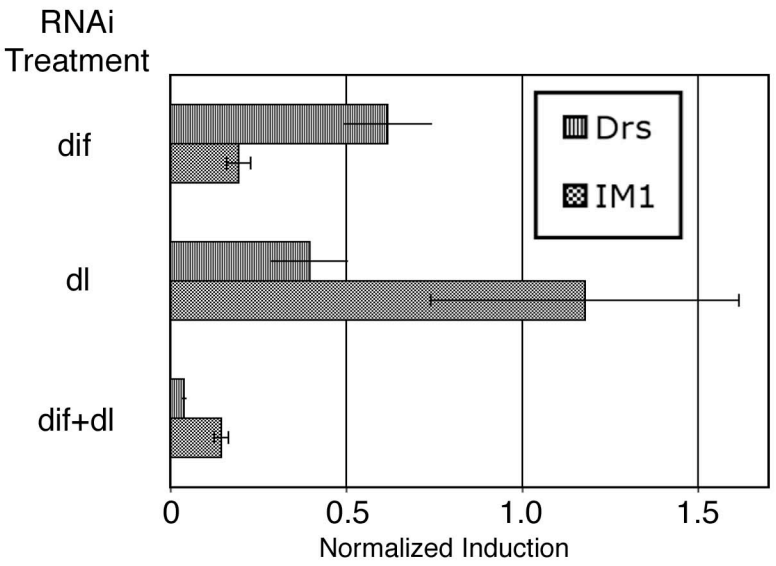

Figure 29: The role of Dorsal in the Toll response of Drosomycin and IM1. Inductions were normalized to the wild-type induction of each reporter.

Acknowledgement

Portions of Chapter Two are taken from “A κB Sequence Code for Pathway-Specific Innate Immune Responses,” Busse, MS, Towb, PT, Arnold, CP, Katrivesis, J, and Wasserman, SA., which was under revision at Molecular Cell at the time of submission of the dissertation. The dissertation author was the primary researcher and author of this paper.

III. Discussion

Overview

The experiments described above represent the first evidence that *Drosophila* κB motifs are intrinsically different in their responsiveness to the Toll and Imd pathways. Previous work had demonstrated the crucial importance of κB motifs in immune induction, but had not demonstrated that different κB sequences actually direct immune loci to respond preferentially to the Toll or Imd pathways. A likely reason previous studies were unable to determine a sequence code is ability of NF-κB proteins to adopt different conformations to bind disparate κB motifs. The resulting rules governing pathway responsiveness are difficult to dissect using common consensus motif techniques based on single base pair mismatches to a motif of fixed length and structure. In addition, many immune loci contain both pathway specific and dual responsive κB motifs, further complicating efforts to classify motifs based on pathway responsiveness.

Subsequent experiments carried out in our lab demonstrated that κB motifs direct specific responses to the Toll and Imd pathways by dictating whether Dif or Relish binds to the motif. This result was unexpected, since studies in a mammalian system had shown that κB motifs were bound by the same NF-κB factors after different pathways were activated, and directed specific pathway responses by dictating which cofactors were able to bind to the NF-κB

factors [12]. A plausible explanation is that the mammalian response to TNFα and LPS is regulated differently than the *Drosophila* response to the Toll and Imd pathways. However, there are some clear parallels in NF-κB binding preferences between mammals and *Drosophila*. p50/p50 dimers preferentially bind motifs very similar to those preferred by Relish, e.g. GGGGATTCCC and GGGGATCCCC, which both contain a GGGGA half site and a central region of three or fewer base pairs. Similarly, p65/p65 homodimers select motifs very similar to those preferred by Dif, such as GGGAATTCCC, which contains a GGGAA half site and four base pair central region [16]. However, in vivo these two mammalian proteins almost always act as heterodimers of p50/p65, especially during regulation of the inflammatory response. This difference illustrates a major difference between the mammalian and *Drosophila* systems, and provides a likely explanation of the lack of an apparent sequence code in mammals. Regulation of the inflammatory response is only one role performed by NF-κB proteins. The regulation of hematopoiesis and lymphoid development might be influenced by the sequence code established here.

Nature of the Sequence Code

X-ray crystallography studies of mammalian p50 and p65 homodimers bound to different κB motifs assisted greatly in understanding the rules governing the specific binding of Relish and Dif. As demonstrated in Figure 15, the distance between the 5' G cluster and the 3' C cluster is a critical determinant of

specificity: Dif can bind when the spacing is at least four bp, and Relish can bind if the spacing is four bp or less. Structural studies provide a plausible explanation for this specificity. For p50 (related to Relish) and p65 (related to Dif), both subunits of the dimer make critical base specific contacts through a pair of arginine residues, which contact two guanine bases at position +3 and +4 (-3 and –4 for subunit 2) (Figure 30 A). These arginine pairs must be separated by at least five bp. The critical difference between p50 and p65 is that p50 requires a G at the +2 and –2 position, whereas p65 requires an A at the +2 and –2 positions. In other words, the critical arginine of p65 binds to the G closest to the A/T central region, while arginine of p50 binds to the G one base away from the A/T region. As a result, p50 requires three bases between the G and C clusters for optimal, symmetric binding, and p65 requires five bases in between the G and C clusters.

However, as stated above, these are flexible proteins, especially with regard to where the second subunit contacts DNA. p65 can bind to a motif with only four bases separating the G and C clusters by shifting the contact bases of the second subunit. Instead of binding to the first and second G closest to the A/T region, it binds non-symmetrically to the second and third G's away from the A/T region (Figure 30B). In order to accommodate this altered binding, the first subunit loses contact between another arginine and the +2 A/T pair, and makes an additional contact between a serine residue and the G^{+5} base. The second subunit also loses a contact between an arginine residue and the –2 position of the motif. Presumably, the second subunit of Dif shifts in a similar manner when

binding to a GGGAATTCCC motif, and Relish makes a similar shift when binding to a GGGGATCCCC motif.

A structure of p50 bound to a GGGAATTCCC motif reveals that p50 is able to accommodate a space of six A/T bases between the arginine contacts (Figure 30 C). This binding is accomplished through several additional base-specific contacts from each subunit covering both half-sites of the motif, explaining why Relish can only bind to a GGGAATTCCC half-site if it is a perfect palindrome. Presumably, Relish cannot bind to other three-G motifs because it is unable to make sufficient contacts. It is unlikely that Dif homodimers can accommodate a similar stretch, since Dif did not bind to a GGGAAATTTCCC (data not shown).

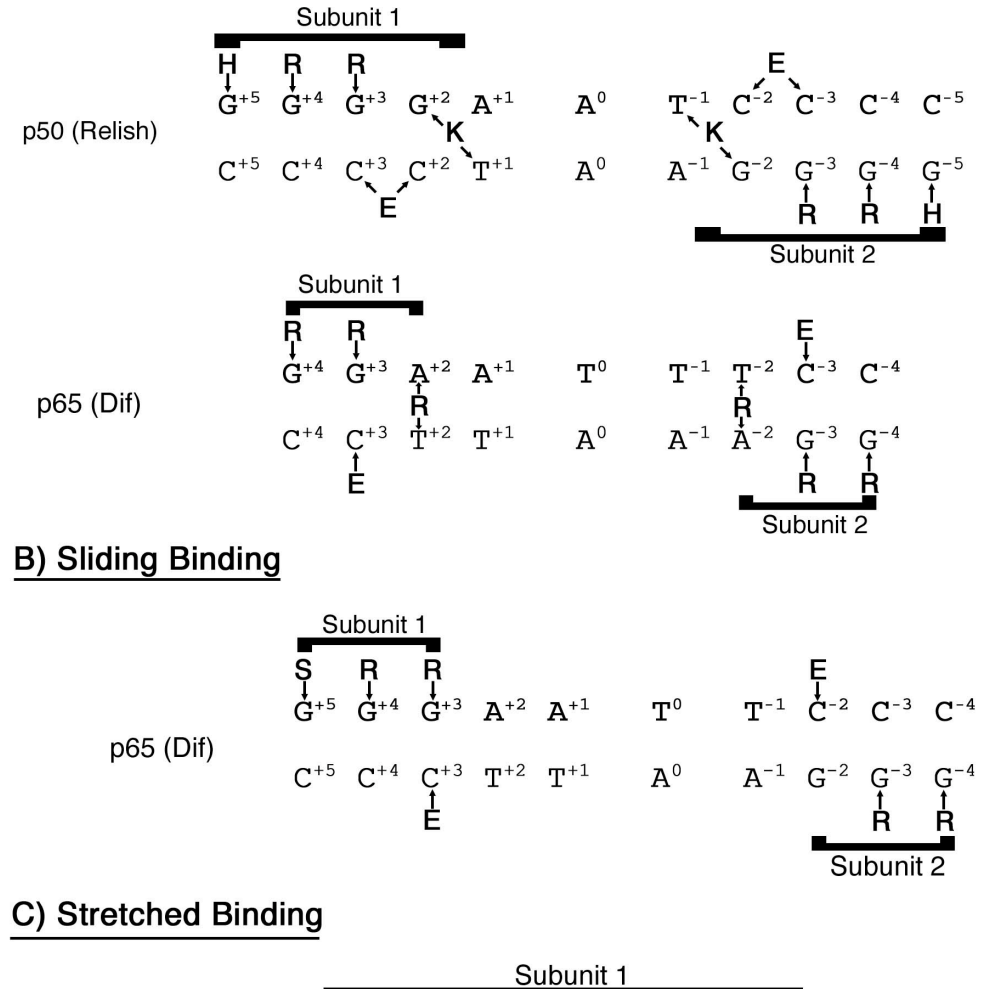

Figure 30: Summary of mammalian NF-κB binding to different κB motifs from published results [20, 24]. The *Drosophila* NF-κB member most closely related to the mammalian family member is indicated in parentheses.

The results of Figure 14 and 15 are correlated to Relish and Dif binding in Table 5.

Table 5: Mode of binding of different κB motifs

	Mode of Binding	
Motif	Rel	Dif
GGGGATCCCC	slide	none
GGGGAGTCCCC	symmetric	none
GGGGAATTCCCC	stretch	slide
GGGGAAGTTCCCC	none	symmetric
GGGATCCC	none	none
GGGAGTCCC	none	none
GGGAATTCCC	stretch	slide
GGGAAGTTCCC	none	symmetric

The reason Relish is able to bind to motifs containing only one half-site is less clear. A study of p65 (which is similar to Dif, not Relish), found that p65 is able to bind to a GGGAAATTTC motif by making base specific contacts with one half-site only. When bound to this motif, the second subunit pulls back from the major groove of DNA and only makes non-specific contacts to the phosphate backbone. No structure of p50 bound to a similar motif is available, so it is unknown if p50 can bind in this manner. A mechanism similar to that seen for p65 could explain why Relish is able to bind to motifs such as GGGGATTTTT. Our data suggest that Dif is unable to bind in this manner. Instead, Dif appears to require a high affinity second half-site in order to bind efficiently.

The apparently sporadic disruption of Relish binding caused by variations in the 3' half-site is puzzling. One possible explanation is that a single GGGGA site is sufficient for binding as long as the 3' half of the motif does not contain a sequence which attracts the second subunit of the transcription factor, allowing it

to swing out of the major groove and bind to the phosphate backbone through non-base-specific contacts. However, a “pseudo-half-site” (a half-site which varies from the perfect sequence, but is still recognizable as a κB half-site) in the 3’ end of the motif might disrupt binding if the second subunit binds to it weakly, preventing the second subunit from swinging out of the major groove and binding non-specifically to the phosphate backbone. Even though some base specific binding occurs, if the sequence is not a high-affinity site, the second subunit will periodically fall off the DNA. Thus, all contact with DNA is lost, including the non-specific contact with the phosphate backbone that would be maintained if the second half-site did not attract the second subunit at all, allowing it to swing out of the major groove.

Exactly what constitutes a “pseudo-half-site” is still unknown. Changing a single T to a C in any position of the GGGGATTTTT motif is not sufficient to disrupt binding (data not shown), suggesting that more than one base is necessary to attract the second subunit into the major groove. There is likely a balance that must be achieved. The second half-site must either be of sufficient affinity that the second subunit is stably bound, or of low enough affinity that the second subunit does not make any base-specific contacts.

A perfect second half-site might have a similar effect if it is spaced too far away from the first half-site, simultaneously attracting both subunits, but not allowing either to maintain stable contact. In effect, each subunit might periodically pull the other subunit off its half-site, preventing a stable transcription complex from forming.

The Role of Cooperativity in Immune Responsive Genes

The concept of Cooperativity - non-additive increases in the induction resulting from multiple κB motifs - is clearly illustrated in Figure 11. In the AttA promoter, one copy of the –46 motif (κB 1) yields an Imd induction 40-60% of the wild-type AttA response, depending on where in the promoter it is positioned. Two copies of this motif give an Imd induction greater than 200% of wild-type, a four-fold increase over a single motif. Similar results are seen with Toll induction of the –118 motif (κB 2). This cooperative increase is likely a result of cooperative interactions between transcription factors bound to multiple motifs.

Several results suggest there is some intrinsic difference in the strength of Imd and Toll inductions. The first such result was presented in Figure 12. In the context of the AttA promoter, the Toll response to the Toll consensus κB motif was more than ten-fold higher than the wild-type AttA response, while the Imd response to the Imd consensus motifs was at best equal to the wild-type response. Similarly, in Figure 24, a single Toll-specific motif, GGGAAGTTCCC gives a Toll response five times greater than the wild-type AttA response, and two copies of that motif yield and induction 40 times as great as wild-type, an eight-fold increase over the response of a single site. An Imd specific motif, GGGGATCCCC is only marginally stronger than the wild-type AttA response.

The reason for this difference in the strength of Imd and Toll response is not clear. It could be an artifact of the cultured cell system, or the way in which

Toll signaling is activated using the chimeric EGF receptor. However, the fact that almost all of the genes in the Toll set only contain a single site suggests that this difference is physiologically relevant. It seems that the Imd response requires the cooperative activity of multiple Imd-responsive motifs to reach the same level of induction as a single Toll motif.

Cooperativity Between Two Pathways

In addition to Cooperativity between two κB motifs responsive to the same pathway, Cooperativity also occurs between two κB motifs responsive to different pathways if both pathways are stimulated at the same time. Our results suggest that this two-pathway Cooperativity is a result of Relish binding to one site and cooperatively interacting with Dif bound to another site. Since the IM1 promoter only contains a single, Toll-responsive motif, it is not cooperatively induced by both pathways. The other three reporters examined, AttA, AttD, and Drs, all contain multiple sites, and respond cooperatively to the activation of both pathways. It is somewhat surprising that Drosomycin responds cooperatively to both pathways, since it does not respond to the Imd pathway alone. Since the –139 motif in Drs is only bound weakly by Relish, perhaps not enough Relish is able to bind when the Imd pathway alone is stimulated, but stimulation of both pathways allows Relish to bind to the promoter in a cooperative manner with Dif bound to the –303 site.

This two-pathway Cooperativity could represent the fly's response to a massive infection by multiple pathogens. If multiple pathogens were infecting the fly and turning on both the Imd and Toll pathways, the fly's only chance of survival at that point might be a massive induction of all antimicrobial peptides.

Transcriptional Regulation of Immune Responsive Genes

In light of the experiments presented above, we can now begin to address the question of how the Toll and Imd pathways direct expression of specific subsets of antimicrobial peptide genes. The synergistic interaction of multiple motifs is a key feature of the regulation of these genes, especially with regard to Imd induction.

One striking finding is the inability of the IM1 reporter to respond to the Imd pathway, even when it contains a strongly Imd responsive κB motif. One difference between IM1 and the other three AMPs studied is the presence in IM1 of only one κB motif. However, in other reporters, when only a single κB motif is present, an Imd response is still seen. Furthermore, addition of a second, Imd responsive κB motif to IM1 does not render the reporter Imd responsive. It appears that there is something intrinsically different about the IM1 context that prevents it from responding to the Imd pathway. Perhaps the IM1 promoter lacks a regulatory site for a co-factor required for Imd responsiveness, or contains an element that represses the Imd response. Given the recent reports of the

requirement for JNK signaling for the induction of some Imd responsive genes, AP-1 sites are interesting candidates for required co-factor binding sites.

We have a good understanding of the regulation of the other three AMP genes studied. In previous studies and our hands, Drosomycin is strongly Toll responsive, and minimally Imd responsive. This pattern of induction results from the presence of a high-affinity Dif site at –303 coupled with a low-affinity Relish site at -139. Since Drs displays an induction much greater than IM1, it is likely that the site at –139 contributes cooperatively to this high Toll induction. There are two possible explanations for this contribution, which are not mutually exclusive. First, Relish is known to be present at low levels in the nucleus in the absence of any stimulus, and perhaps this low level of Relish is able to synergize with Dif bound to the –303 site. Alternatively, the –139 site might have an affinity for Dif that is too low to detect through a gel shift, but sufficient to attract very low levels of Dif to the site. Given the enormous capacity for Cooperativity in the Toll pathway, a low level of Dif on the –139 site might be enough to boost the Toll response. Another possible reason for the presence of this Relish site might involve localized immune inductions. Since local immune responses in the epithelia are regulated by the Imd pathway only [84, 85], this –139 site might be required for the local induction of Drosomycin.

Attacin A is a dual-responsive gene, although it responds more strongly to the Imd pathway than the Toll pathway. This induction pattern results from the presence of one high affinity Relish site at –46, and a motif at –118 with a moderate affinity for Dif and low affinity for Relish. The high affinity and low

affinity Relish sites combine to produce a strong Imd response, and the moderate Dif site provides the moderate Toll response.

Attacin D exhibits the strongest Imd response of the reporters examined, due to the presence of two high affinity and one low affinity Relish sites at –155, -189, and –89, respectively. These three sites act cooperatively to produce a high Imd response. The weak Toll induction of AttD results from the low affinity Dif site at –89.

Dorsal

Dorsal is often ignored during studies of the immune response, perhaps rightly so, since available data suggests that in adults, Dorsal is unable to mediate immune responses. The data presented here indicate a distinct role for Dorsal in the immune response.

The regulation of two loci, AttA and Drs, exhibits a role for Dorsal which is not entirely redundant with Dif. Although RNAi against Dorsal does not significantly reduce the Toll response of these genes, knocking down both Dif and Dorsal has a greater effect than RNAi against Dif alone. Reciprocal to the results of replacing Dorsal with Dif during development, Dorsal can partially substitute for the function of Dif during immune signaling in S2 cells. Another gene, IM1, appears to be independent of Dorsal. There are at least two possible explanations for this distinct behavior of IM1. First, IM1 might exhibit less regulation by Dorsal because its κB motif is not bound by Dorsal. It would be

interesting to see if any of the κB motifs from the AttA and Drs promoter are bound more strongly by Dorsal than the motif in IM1. Second, the lack of Dorsal involvement in the regulation of IM1 might be related to the inability of Relish to induce IM1. Perhaps the IM1 promoter is a minimal promoter, lacking binding sites for other cofactors present in AttA, Drs and other genes. It is possible that Relish and Dorsal require the recruitment of large regulatory complexes in order to promote transcription, and the IM1 promote does not support the construction of these complexes.

The results presented here concerning Dorsal could be an artifact of the S2 cells, which have embryonic characteristics. Given the involvement of Dorsal in immune regulation of larvae, but not adults, it is not clear how directly these results in S2 cells will translate to adult flies. However, several existing studies demonstrate that Dif and Dorsal have distinct binding and activating capabilities, and it is possible that Dorsal plays a distinct role in regulating immune genes. Available data suggest that even if Dorsal does play a role which is distinct from Dif, it is probably a minor role, meaning the time and resources necessary to fully elucidate this role is unlikely to result in a large pay-off.

Conclusion

The findings described in this document can help clarify the field of *Drosophila* immunity. Since the expression of antimicrobial peptides is often used as a read-out for Toll and Imd signaling, it is useful to understand how that

expression is regulated. Studies explaining the activation of the Toll and Imd pathways by different types of peptidoglycan refined the usefulness of peptidoglycan as a research tool. Similarly, understanding exactly how the κB motifs in AMPs direct Imd and Toll responsiveness will refine the usefulness of AMPs as reporters of the Imd and Toll pathways. For example, Drosomycin is commonly employed as a Toll-specific reporter, despite conflicting reports of weak Imd responsiveness. The presence of a weak Imd site in the Drosomycin reporter suggests it is not an ideal reporter to measure Toll-specific activation. A better choice is the IM1 gene, which is truly Toll-specific.

As described in the introduction, there is much variability and even conflicting data concerning pathway activation and gene induction. The two-pathway Cooperativity described here might explain some of this confusion, in three different ways. First, few studies are carried out in truly axenic conditions. While the flies may often appear healthy, a minor infection might activate either the Toll or Imd pathway, such that when an infection is purposely introduced, an unintended cooperative response occurs, leading to different levels of induction in different individuals. Second, injection is often used to introduce microbe infections, and it is known that the injury itself elicits some sort of response. Exactly which pathways are activated is unknown, however both Imd and Toll-responsive genes are induced, meaning both pathways might be activated. Injection of a microbe might again activate an unintended synergist response involving both the Toll and Imd pathways when the intention was to activate a single pathway. In particular, this could explain the huge difference in response

between natural infection of fungi and injection of fungi [45]. Third, a common method of immune challenge is injection of a mixture of *E. coli* (Imd activating) and *M. luteus* (Toll activating). This technique is used to study the Toll response and the Imd response at the same time. However, the data presented here suggest that instead of studying the Toll response and the Imd response at the same time, what is actually being studied is the cooperative response. As a result, genes which are weakly Toll responsive will show a greater reduction in the response to this mixed bacterial infection when Toll pathway components are mutated than they would show if only the Toll pathway were activated, and vice versa.

IV. Experimental Procedures

Reagents and Site-Directed Mutagenesis

Murine EGF was purchased from Calbiochem and the *B. subtilis* peptidoglycan (PG) preparation was purchased from Sigma. All luciferase reporter constructs are based on the pBL3-Basic vector (Promega). PCR amplification of the AttA promoter utilized a 5' primer that introduced a KpnI site and a 3' primer that introduced an NheI site.

Mutagenic oligonucleotides and the Expand High-Fidelity PCR System (Roche) were used to introduce site-directed alterations into the AttA reporter construct. The construct length was kept invariant and all motifs were introduced such that the position of the last G in the 5' core element was kept constant.

Cell Culture and Transient Transfections

S2*/EGFR-Toll cells were maintained and transfected as described previously [104]. Six-well plates were seeded with 3 ml of 10^6 cells/ml. After six hours, transfections were performed as described for the *Drosophila* Expression System (Invitrogen), with 100 ng each of reporter plasmid and pAc-LacZ. After 16-20 hours, each transfection was split into thirds in fresh media and transferred to 12-well plates. One third (1 ml) was treated with 10 μg/ml PGN, one third was treated with 0.4 μg/ml EGF, and one third was left untreated. After 4 hours, cells were washed with 500 μl PBS and then lysed in 100μl Luciferase Reporter Buffer

(Promega). Fifteen μl of lysate was assayed for luciferase activity (Luciferase Assay System, Promega), and fifteen μl was assayed for β-galactosidase (β-gal) activity (Galacto-Light Plus, Tropix). Luciferase activity was normalized against β-gal activity to control for transfection efficiency. The percentage increase in expression upon induction was calculated by comparing treated and untreated samples from the same transfection. RNA interference experiments were carried out as previously described [104].

Bioinformatic Analysis of *Drosophila* Innate Immune Response Loci

Using published microarray data [66], we defined a set of loci that were strongly induced in adult flies by mixed [Gram(+) and Gram(-)] bacterial infection or by fungal infection. We then selected loci for which induction was substantially diminished by a mutation in *spätzle* (*spz*), blocking Toll signaling, or a mutation in *relish* (*rel*), disrupting Imd signaling. The specific criteria for assigning loci to each class are presented in Supplemental Table 2.

For each locus, we used available cDNA sequence data, as well as sequence profiles for TATA box and initiator elements [105], to define the core promoter region. For eleven genes this resulted in assignment of a transcriptional start site more than 10 bp away from the annotated gene start (http://flybase.bio.indiana.edu/); the sequences for the 5' ends of all genes are presented in Supplemental Figure S1. For MEME analysis [103], we extracted

160 bp of genomic sequence extending from -190 to -40 relative to the calculated transcriptional start site of each gene.

Protein Purification

The pGex-DifNX plasmid [37] was a gift from Tony Ip. A pGex-Rel construct was generated by inserting bp 1-1635 of Relish (amplified from cDNA) into the BamHI and EcoRI sites of the poly-linker site of pGex. This fragment corresponds to amino acids 1-545, ending at the natural proteolytic cleavage site of Relish [70]. Plasmids were transformed into BL21, and purified as described previously [106], with the following modifications. A 500 ml culture was grown at 30°C to OD550 ~ 0.5, then induced for one hour with 1.6 mM IPTG. Cells were pelleted by centrifugation at 6,000 g for 15 minutes at 4°C. The pellet was resuspended in 10ml cold lysis buffer [25 mM HEPES (pH 7.5), 20 mM KCl, 2.5 mM EDTA, 1% Triton-X 100, 1 mM DTT, 0.5 mM PMSF, 1X EDTA-free protease inhibitor cocktail (Roche)]. Cells were lysed by performing three freeze-thaw cycles, followed by three cycles of 10 second sonication and 30 second incubation on ice. To 5 ml of bacterial cell suspension, 500 μl of 5M NaCl was added, and the suspension was incubated on ice for 15 minutes. Cellular debris was pelleted by centrifugation at 12,000 g for 15 minutes at 4°C. Next, 750 μL of a 50% slurry of cleaned glutathione-Sepharose 4B (Amersham Biosciences) in PBS was added to the lysate and rocked gently at 4°C for 1 hour. Sepharose beads were pelleted by centrifugation at 500 g for 5 minutes at 4°C, then washed

three times with 10 ml cold lysis buffer and once with 10 ml cold 50 mM Tris (pH 8.0). Three elutions were performed with 500 μl of 50 mM Tris (pH 8.0) plus 10 mM glutathione and gentle rocking at 4°C for 1 hour.

Gel-Shift Assays

Electrophoretic mobility shift assays were performed as previously described [107] with the following modifications. Reactions were carried out in a volume of 10 μl containing 0.1 μg/μl poly(dI-dC), 10 mM HEPES, 50 mM NaCl, 1 mg/ml BSA, 3 mM $MgCl_2$, 6 mM βME, 10 mM EDTA, and 10% (v/v) glycerol. End-labeled probe (0.8 pmol) was combined with protein, and incubated for 10 minutes at room temperature. Protein was used at a concentration of 1.5 μg per reaction in experiments with synthetic binding sites and 7.5 μg protein per reaction with endogenous binding sites. Binding complexes were resolved by electrophoresis in a 4% polyacrylamide gel in TBE buffer for 140 min at 7.5 V/cm. All oligonucleotides were of 24 bp in length, centered on the κB motif. Synthetic κB motifs were substituted for the -46 motif in Attacin A such that the position of the last G in the 5' core element was kept constant. The κB motifs from Attacin A (-46 and -118), Diptericin (-145), and IM1 (-86) were each analyzed in their endogenous context.

Works Cited

1. Ghosh, S., M.J. May, and E.B. Kopp, *NF-kappa B and Rel proteins: evolutionarily conserved mediators of immune responses.* Annu Rev Immunol, 1998. **16**: p. 225-60.

2. Hoffmann, A. and D. Baltimore, *Circuitry of nuclear factor kappaB signaling.* Immunol Rev, 2006. **210**: p. 171-86.

3. Sen, R. and D. Baltimore, *Multiple nuclear factors interact with the immunoglobulin enhancer sequences.* Cell, 1986. **46**(5): p. 705-16.

4. Ghosh, S. and M. Karin, *Missing pieces in the NF-kappaB puzzle.* Cell, 2002. **109 Suppl**: p. S81-96.

5. Piva, R., G. Belardo, and M.G. Santoro, *NF-kappaB: a stress-regulated switch for cell survival.* Antioxid Redox Signal, 2006. **8**(3-4): p. 478-86.

6. Natoli, G., et al., *Interactions of NF-kappaB with chromatin: the art of being at the right place at the right time.* Nat Immunol, 2005. **6**(5): p. 439-45.

7. Natoli, G. and F. De Santa, *Shaping alternative NF-kappaB-dependent gene expression programs: new clues to specificity.* Cell Death Differ, 2006. **13**(5): p. 693-6.

8. Pomerantz, J.L. and D. Baltimore, *Two pathways to NF-kappaB.* Mol Cell, 2002. **10**(4): p. 693-5.

9. Hoffmann, A., et al., *The IkappaB-NF-kappaB signaling module: temporal control and selective gene activation.* Science, 2002. **298**(5596): p. 1241-5.

10. Werner, S.L., D. Barken, and A. Hoffmann, *Stimulus specificity of gene expression programs determined by temporal control of IKK activity.* Science, 2005. **309**(5742): p. 1857-61.

11. Natoli, G., *Tuning up inflammation: how DNA sequence and chromatin organization control the induction of inflammatory genes by NF-kappaB.* FEBS Lett, 2006. **580**(12): p. 2843-9.

12. Leung, T.H., A. Hoffmann, and D. Baltimore, *One nucleotide in a kappaB site can determine cofactor specificity for NF-kappaB dimers.* Cell, 2004. **118**(4): p. 453-64.
13. Hoffmann, A., T.H. Leung, and D. Baltimore, *Genetic analysis of NF-kappaB/Rel transcription factors defines functional specificities.* Embo J, 2003. **22**(20): p. 5530-9.

14. Saccani, S., S. Pantano, and G. Natoli, *Modulation of NF-kappaB activity by exchange of dimers.* Mol Cell, 2003. **11**(6): p. 1563-74.

15. Saccani, S., S. Pantano, and G. Natoli, *Two waves of nuclear factor kappaB recruitment to target promoters.* J Exp Med, 2001. **193**(12): p. 1351-9.

16. Kunsch, C., S.M. Ruben, and C.A. Rosen, *Selection of optimal kappa B/Rel DNA-binding motifs: interaction of both subunits of NF-kappa B with DNA is required for transcriptional activation.* Mol Cell Biol, 1992. **12**(10): p. 4412-21.

17. Bonizzi, G., et al., *Activation of IKKalpha target genes depends on recognition of specific kappaB binding sites by RelB:p52 dimers.* Embo J, 2004. **23**(21): p. 4202-10.

18. Sanjabi, S., et al., *A c-Rel subdomain responsible for enhanced DNA-binding affinity and selective gene activation.* Genes Dev, 2005. **19**(18): p. 2138-51.

19. Udalova, I.A., et al., *Quantitative prediction of NF-kappa B DNA-protein interactions.* Proc Natl Acad Sci U S A, 2002. **99**(12): p. 8167-72.

20. Muller, C.W. and S.C. Harrison, *The structure of the NF-kappa B p50:DNA-complex: a starting point for analyzing the Rel family.* FEBS Lett, 1995. **369**(1): p. 113-7.

21. Chen, Y.Q., S. Ghosh, and G. Ghosh, *A novel DNA recognition mode by the NF-kappa B p65 homodimer.* Nat Struct Biol, 1998. **5**(1): p. 67-73.

22. Ghosh, G., et al., *Structure of NF-kappa B p50 homodimer bound to a kappa B site.* Nature, 1995. **373**(6512): p. 303-10.

23. Muller, C.W., et al., *Structure of the NF-kappa B p50 homodimer bound to DNA.* Nature, 1995. **373**(6512): p. 311-7.

24. Chen, Y.Q., et al., *NF-kappaB p65 (RelA) homodimer uses distinct mechanisms to recognize DNA targets.* Structure, 2000. **8**(4): p. 419-28.

25. Engstrom, Y., O. Loseva, and U. Theopold, *Proteomics of the Drosophila immune response.* Trends Biotechnol, 2004. **22**(11): p. 600-5.

26. Hultmark, D., *Insect immunology. Ancient relationships.* Nature, 1994. **367**(6459): p. 116-117.

27. Hoffmann, J.A., *The immune response of Drosophila.* Nature, 2003. **426**(6962): p. 33-8.

28. Hultmark, D., *Drosophila immunity: paths and patterns.* Curr Opin Immunol, 2003. **15**(1): p. 12-9.

29. Kim, T. and Y.J. Kim, *Overview of innate immunity in Drosophila.* J Biochem Mol Biol, 2005. **38**(2): p. 121-7.

30. Royet, J., J.M. Reichhart, and J.A. Hoffmann, *Sensing and signaling during infection in Drosophila.* Curr Opin Immunol, 2005. **17**(1): p. 11-7.

31. Zasloff, M., *Antimicrobial peptides of multicellular organisms.* Nature, 2002. **415**(6870): p. 389-95.

32. Bulet, P., et al., *Antimicrobial peptides in insects; structure and function.* Dev Comp Immunol, 1999. **23**(4-5): p. 329-44.

33. Jiggins, F.M. and K.W. Kim, *The evolution of antifungal peptides in Drosophila.* Genetics, 2005. **171**(4): p. 1847-59.

34. Rabel, D., et al., *Primary structure and in vitro antibacterial properties of the Drosophila melanogaster attacin C Pro-domain.* J Biol Chem, 2004. **279**(15): p. 14853-9.

35. Lemaitre, B., J.M. Reichhart, and J.A. Hoffmann, *Drosophila host defense: differential induction of antimicrobial peptide genes after infection by various classes of microorganisms.* Proc Natl Acad Sci U S A, 1997. **94**(26): p. 14614-9.

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

37. Ip, Y.T., et al., *Dif, a dorsal-related gene that mediates an immune response in Drosophila.* Cell, 1993. **75**(4): p. 753-763.

38. Meng, X., B.S. Khanuja, and Y.T. Ip, *Toll receptor-mediated Drosophila immune response requires Dif, an NF-kappaB factor.* Genes Dev, 1999. **13**(7): p. 792-7.

39. Rutschmann, S., et al., *The Rel protein DIF mediates the antifungal but not the antibacterial host defense in Drosophila.* Immunity, 2000. **12**(5): p. 569-80.

40. Lemaitre, B., et al., *A recessive mutation, immune deficiency (imd), defines two distinct control pathways in the Drosophila host defense.* Proceedings of the National Academy of Sciences of the United States of America, 1995. **92**(21): p. 9465-9.

41. Georgel, P., et al., *Drosophila immune deficiency (IMD) is a death domain protein that activates antibacterial defense and can promote apoptosis.* Dev Cell, 2001. **1**(4): p. 503-14.

42. Hedengren, M., et al., *Relish, a central factor in the control of humoral but not cellular immunity in Drosophila.* Molecular Cell, 1999. **4**(5): p. 827-37.

43. Lu, Y., L.P. Wu, and K.V. Anderson, *The antibacterial arm of the drosophila innate immune response requires an IkappaB kinase.* Genes Dev, 2001. **15**(1): p. 104-10.

44. Vidal, S., et al., *Mutations in the Drosophila dTAK1 gene reveal a conserved function for MAPKKKs in the control of rel/NF-kappaB-dependent innate immune responses.* Genes Dev, 2001. **15**(15): p. 1900-12.

45. Hedengren-Olcott, M., et al., *Differential activation of the NF-kappaB-like factors relish and Dif in drosophila melanogaster by fungi and gram-positive bacteria.* J Biol Chem, 2004. **In press.**

46. Kaneko, T., et al., *PGRP-LC and PGRP-LE have essential yet distinct functions in the drosophila immune response to monomeric DAP-type peptidoglycan.* Nat Immunol, 2006. **7**(7): p. 715-23.

47. Janeway, C.A., Jr., *The immune system evolved to discriminate infectious nonself from noninfectious self.* Immunology Today, 1992. **13**(1): p. 11-6.

48. Medzhitov, R. and C.A. Janeway, Jr., *Innate immunity: the virtues of a nonclonal system of recognition.* Cell, 1997. **91**(3): p. 295-8.

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

50. Werner, T., et al., *A family of peptidoglycan recognition proteins in the fruit fly Drosophila melanogaster.* Proc Natl Acad Sci U S A, 2000. **97**(25): p. 13772-7.

51. Kim, Y.S., et al., *Gram-negative bacteria-binding protein, a pattern recognition receptor for lipopolysaccharide and beta-1,3-glucan that mediates the signaling for the induction of innate immune genes in Drosophila melanogaster cells.* J Biol Chem, 2000. **275**(42): p. 32721-7.

52. Michel, T., et al., *Drosophila Toll is activated by Gram-positive bacteria through a circulating peptidoglycan recognition protein.* Nature, 2001. **414**(6865): p. 756-9.

53. Leulier, F., et al., *The Drosophila immune system detects bacteria through specific peptidoglycan recognition.* Nat Immunol, 2003. **4**(5): p. 478-84.

54. Gobert, V., et al., *Dual activation of the Drosophila toll pathway by two pattern recognition receptors.* Science, 2003. **302**(5653): p. 2126-30.

55. Pili-Floury, S., et al., *In vivo RNA interference analysis reveals an unexpected role for GNBP1 in the defense against Gram-positive bacterial infection in Drosophila adults.* J Biol Chem, 2004. **279**(13): p. 12848-53.

56. Filipe, S.R., A. Tomasz, and P. Ligoxygakis, *Requirements of peptidoglycan structure that allow detection by the Drosophila Toll pathway.* EMBO Rep, 2005. **6**(4): p. 327-33.

57. Choe, K.M., et al., *Requirement for a peptidoglycan recognition protein (PGRP) in Relish activation and antibacterial immune responses in Drosophila.* Science, 2002. **296**(5566): p. 359-62.

58. Gottar, M., et al., *The Drosophila immune response against Gram-negative bacteria is mediated by a peptidoglycan recognition protein.* Nature, 2002. **416**(6881): p. 640-4.

59. Ramet, M., et al., *Functional genomic analysis of phagocytosis and identification of a Drosophila receptor for E. coli.* Nature, 2002. **416**(6881): p. 644-8.

60. Takehana, A., et al., *Overexpression of a pattern-recognition receptor, peptidoglycan-recognition protein-LE, activates imd/relish-mediated antibacterial defense and the prophenoloxidase cascade in Drosophila larvae.* Proc Natl Acad Sci U S A, 2002. **99**(21): p. 13705-10.

61. Takehana, A., et al., *Peptidoglycan recognition protein (PGRP)-LE and PGRP-LC act synergistically in Drosophila immunity.* Embo J, 2004. **23**(23): p. 4690-700.

62. Stenbak, C.R., et al., *Peptidoglycan molecular requirements allowing detection by the Drosophila immune deficiency pathway.* J Immunol, 2004. **173**(12): p. 7339-48.

63. Ligoxygakis, P., et al., *Activation of Drosophila Toll during fungal infection by a blood serine protease.* Science, 2002. **297**(5578): p. 114-6.

64. Bangham, J., F. Jiggins, and B. Lemaitre, *Insect immunity: the post-genomic era.* Immunity, 2006. **25**(1): p. 1-5.

65. Bischoff, V., et al., *Downregulation of the Drosophila Immune Response by Peptidoglycan-Recognition Proteins SC1 and SC2.* PLoS Pathog, 2006. **2**(2): p. e14.

66. De Gregorio, E., et al., *The Toll and Imd pathways are the major regulators of the immune response in Drosophila.* Embo J, 2002. **21**(11): p. 2568-79.

67. Sun, H., et al., *Regulated assembly of the Toll signaling complex drives Drosophila dorsoventral patterning.* Embo J, 2004. **23**(1): p. 100-10.

68. Silverman, N., et al., *A Drosophila IkappaB kinase complex required for Relish cleavage and antibacterial immunity.* Genes Dev, 2000. **14**(19): p. 2461-71.

69. Rutschmann, S., et al., *Role of Drosophila IKK gamma in a toll-independent antibacterial immune response.* Nat Immunol, 2000. **1**(4): p. 342-7.

70. Stoven, S., et al., *Caspase-mediated processing of the Drosophila NF-kappaB factor Relish.* Proc Natl Acad Sci U S A, 2003. **100**(10): p. 5991-6.

71. Stoven, S., et al., *Activation of the Drosophila NF-kappaB factor Relish by rapid endoproteolytic cleavage.* EMBO Rep, 2000. **1**(4): p. 347-52.

72. Silverman, N., et al., *Immune activation of NF-kappaB and JNK requires Drosophila TAK1.* J Biol Chem, 2003. **278**(49): p. 48928-34.

73. Park, J.M., et al., *Targeting of TAK1 by the NF-kappa B protein Relish regulates the JNK-mediated immune response in Drosophila.* Genes Dev, 2004. **18**(5): p. 584-94.

74. Kim, T., et al., *Downregulation of lipopolysaccharide response in Drosophila by negative crosstalk between the AP1 and NF-kappaB signaling modules.* Nat Immunol, 2005. **6**(2): p. 211-8.

75. Kallio, J., et al., *Functional analysis of immune response genes in Drosophila identifies JNK pathway as a regulator of antimicrobial peptide gene expression in S2 cells.* Microbes Infect, 2005. **7**(5-6): p. 811-9.

76. Delaney, J.R., et al., *Cooperative control of Drosophila immune responses by the JNK and NF-kappaB signaling pathways.* Embo J, 2006. **25**(13): p. 3068-77.

77. Levashina, E.A., et al., *Two distinct pathways can control expression of the gene encoding the Drosophila antimicrobial peptide metchnikowin.* J Mol Biol, 1998. **278**(3): p. 515-27.

78. Manfruelli, P., et al., *A mosaic analysis in Drosophila fat body cells of the control of antimicrobial peptide genes by the Rel proteins Dorsal and DIF.* Embo Journal, 1999. **18**(12): p. 3380-91.

79. Han, Z.S. and Y.T. Ip, *Interaction and specificity of Rel-related proteins in regulating Drosophila immunity gene expression.* Journal of Biological Chemistry, 1999. **274**(30): p. 21355-61.

80. Boutros, M., H. Agaisse, and N. Perrimon, *Sequential activation of signaling pathways during innate immune responses in Drosophila.* Dev Cell, 2002. **3**(5): p. 711-22.

81. Wu, L.P., et al., *Drosophila immunity: genes on the third chromosome required for the response to bacterial infection.* Genetics, 2001. **159**(1): p. 189-99.

82. Petersen, U.M., et al., *Serpent regulates Drosophila immunity genes in the larval fat body through an essential GATA motif.* Embo J, 1999. **18**(14): p. 4013-22.

83. Senger, K., et al., *Immunity regulatory DNAs share common organizational features in Drosophila.* Mol Cell, 2004. **13**(1): p. 19-32.

84. Ferrandon, D., et al., *A drosomycin-GFP reporter transgene reveals a local immune response in Drosophila that is not dependent on the Toll pathway.* Embo J, 1998. **17**(5): p. 1217-27.

85. Tzou, P., et al., *Tissue-specific inducible expression of antimicrobial peptide genes in Drosophila surface epithelia.* Immunity, 2000. **13**(5): p. 737-48.

86. Senger, K., K. Harris, and M. Levine, *GATA factors participate in tissue-specific immune responses in Drosophila larvae.* Proc Natl Acad Sci U S A, 2006. **103**(43): p. 15957-62.

87. Tzou, P., J.M. Reichhart, and B. Lemaitre, *Constitutive expression of a single antimicrobial peptide can restore wild-type resistance to infection in immunodeficient Drosophila mutants.* Proc Natl Acad Sci U S A, 2002. **99**(4): p. 2152-7.

88. Ryu, J.H., et al., *The homeobox gene Caudal regulates constitutive local expression of antimicrobial peptide genes in Drosophila epithelia.* Mol Cell Biol, 2004. **24**(1): p. 172-85.

89. Engstrom, Y., et al., *kB-like motifs regulate the induction of immune genes in Drosophila.* J Mol Biol, 1993. **232**(2): p. 327-333.

90. Kappler, C., et al., *Insect immunity. Two 17 bp repeats nesting a kappa B-related sequence confer inducibility to the diptericin gene and bind a polypeptide in bacteria-challenged Drosophila.* Embo J, 1993. **12**(4): p. 1561-8.

91. Georgel, P., et al., *Insect immunity: the diptericin promoter contains multiple functional regulatory sequences homologous to mammalian acute-phase response elements.* Biochem Biophys Res Commun, 1993. **197**(2): p. 508-17.

92. Dushay, M.S., et al., *Two attacin antibacterial genes of Drosophila melanogaster.* Gene, 2000. **246**(1-2): p. 49-57.

93. Meister, M., et al., *Insect immunity. A transgenic analysis in Drosophila defines several functional domains in the diptericin promoter.* Embo J, 1994. **13**(24): p. 5958-66.

94. Uvell, H. and Y. Engstrom, *Functional characterization of a novel promoter element required for an innate immune response in Drosophila.* Mol Cell Biol, 2003. **23**(22): p. 8272-81.

95. Petersen, U.M., et al., *The dorsal-related immunity factor, Dif, is a sequence-specific trans-activator of Drosophila Cecropin gene expression.* Embo J, 1995. **14**(13): p. 3146-58.

96. Gross, I., et al., *Drosophila immunity: a comparative analysis of the Rel proteins dorsal and Dif in the induction of the genes encoding diptericin and cecropin.* Nucleic Acids Res, 1996. **24**(7): p. 1238-45.

97. Stein, D., et al., *The Dorsal-related immunity factor (Dif) can define the dorsal-ventral axis of polarity in the Drosophila embryo.* Development, 1998. **125**(11): p. 2159-69.

98. Asling, B., M.S. Dushay, and D. Hultmark, *Identification of early genes in the Drosophila immune response by PCR-based differential display: the Attacin A gene and the evolution of attacin-like proteins.* Insect Biochem Mol Biol, 1995. **25**(4): p. 511-8.

99. Samakovlis, C., et al., *The immune response in Drosophila: pattern of cecropin expression and biological activity.* Embo J, 1990. **9**(9): p. 2969-76.

100. De Gregorio, E., et al., *Genome-wide analysis of the Drosophila immune response by using oligonucleotide microarrays.* Proc Natl Acad Sci U S A, 2001. **98**(22): p. 12590-5.

101. Whitley, M.Z., et al., *A striking similarity in the organization of the E-selectin and beta interferon gene promoters.* Mol Cell Biol, 1994. **14**(10): p. 6464-75.

102. Thanos, D. and T. Maniatis, *Identification of the rel family members required for virus induction of the human beta interferon gene.* Mol Cell Biol, 1995. **15**(1): p. 152-64.

103. Bailey, T.L. and C. Elkan, *Fitting a mixture model by expectation maximization to discover motifs in biopolymers.* Proc Int Conf Intell Syst Mol Biol, 1994. **2**: p. 28-36.

104. Sun, H., et al., *A heterotrimeric death domain complex in Toll signaling.* Proc Natl Acad Sci USA, 2002. **99**(20): p. 12871-12876.

105. Ohler, U., et al., *Computational analysis of core promoters in the Drosophila genome.* Genome Biol, 2002. **3**(12): p. 087.

106. Ip, Y.T., et al., *The dorsal gradient morphogen regulates stripes of rhomboid expression in the presumptive neuroectoderm of the Drosophila embryo.* Genes & Development, 1992. **6**(9): p. 1728-39.

107. Ip, Y.T., et al., *The dorsal morphogen is a sequence-specific DNA-binding protein that interacts with a long-range repression element in Drosophila.* Cell, 1991. **64**(2): p. 439-46.