

# A $\kappa$ B sequence code for pathway-specific innate immune responses

Matthew S Busse, Christopher P Arnold, Par Towb, James Katrivesis and Steven A Wasserman\*

Section of Cell and Developmental Biology, University of California at San Diego, La Jolla, CA, USA

**The Toll and Imd pathways induce humoral innate immune responses in *Drosophila* by activating NF- $\kappa$ B proteins that bind  $\kappa$ B target sites. Here, we delineate a  $\kappa$ B site sequence code that directs pathway-specific expression of innate immune loci. Using bioinformatic analysis of expression and sequence data, we identify shared properties of Imd- and Toll-specific response elements. Employing synthetic  $\kappa$ B sites in luciferase reporter and *in vitro* binding assays, we demonstrate that the length of the (G)<sub>n</sub> element in the 5' half-site and of the central (A,T)-rich region combine to specify responsiveness to one or both pathways. We also show that multiple sites function to enhance the response to either or both pathways. Together, these studies elucidate the mechanism by which  $\kappa$ B motifs direct binding by particular *Drosophila* NF- $\kappa$ B family members and thereby induce specialized innate immune repertoires.**

*The EMBO Journal* (2007) 26, 3826–3835. doi:10.1038/sj.emboj.7601798; Published online 26 July 2007

**Subject Categories:** chromatin & transcription; immunology  
**Keywords:** Dif; Imd; motif; Relish; Toll

## Introduction

In vertebrates and invertebrates alike, innate immune systems mediate pathogen recognition, direct expression of defensive molecules, and trigger secondary responses. The signaling pathways of innate immunity are broadly conserved, utilizing as primary effectors members of the NF- $\kappa$ B family of transcription factors, for example, vertebrate p50 and p65 (Ghosh *et al.*, 1998; Karin and Ben-Neriah, 2000; Brennan and Anderson, 2004; Xiao and Ghosh, 2005). In the absence of a stimulus, NF- $\kappa$ B proteins are retained in the cytoplasm through direct interaction with I $\kappa$ B-family inhibitors. Signals from receptors in the Toll, Interleukin-1, and TNF receptor families trigger proteolytic events that allow NF- $\kappa$ B factors to enter the nucleus and direct expression of antimicrobial agents and secondary signals.

NF- $\kappa$ B proteins function as sequence-specific DNA-binding proteins. Their target sequences, termed  $\kappa$ B sites for the

context in which they were first described, have a consensus in mammals of GGG G/A NN T/C T/C CC (Baeuerle, 1991). NF- $\kappa$ B proteins bind as dimers to  $\kappa$ B sites, with each half-site containing a cluster of guanines (5' end) or cytosines (3' end).

In *Drosophila melanogaster*, three NF- $\kappa$ B family members govern antimicrobial responses under the control of the Toll and Imd signaling cascades. Two of the NF- $\kappa$ B proteins, Dorsal and DIF, each interact with Cactus, an I $\kappa$ B-related inhibitor. Degradation of Cactus in response to signaling by the Toll receptor releases Dorsal or DIF for nuclear translocation and interaction with target gene regulatory sites. The third family member in flies is Relish. Like mammalian p105, Relish is a composite protein, containing both an NF- $\kappa$ B-like domain and an I $\kappa$ B-like inhibitory domain (Dushay *et al.*, 1996; Stoven *et al.*, 2000). Activation of the Imd pathway triggers endoproteolytic cleavage of Relish. Relish processing, carried out at least in part by the caspase Dredd, releases an active NF- $\kappa$ B-like fragment (Stoven *et al.*, 2003).

Acting through Relish, Dif, and Dorsal, flies respond to infection by expressing a plethora of defense molecules, including antimicrobial peptide genes and pathogen recognition proteins. For the most part, the Toll pathway mediates the *Drosophila* response to Gram (+) bacteria and fungi, whereas the Imd pathway activates defenses against Gram (–) bacteria. Flies mutant for both pathways fail to induce any of the known antimicrobial peptides (Tzou *et al.*, 2002). To effect responses specific to pathogen class, the two pathways each direct expression of a set of immune response loci; some loci are pathway-specific, whereas others can be induced by either the Toll or Imd pathway (Lemaitre *et al.*, 1997; De Gregorio *et al.*, 2001, 2002; Irving *et al.*, 2001; Hedengren-Olcott *et al.*, 2004).

It has been thought that variation in the affinity of distinct NF- $\kappa$ B dimers for particular  $\kappa$ B sequences might underlie differential transcription of innate immune loci in insects and mammals. Consistent with this hypothesis, studies *in vitro* have revealed small but significant differences in binding specificity among vertebrate NF- $\kappa$ B proteins (Kunsch *et al.*, 1992). Furthermore, the IKK $\beta$ -independent pathway for NF- $\kappa$ B activation in mammals acts via target sites that are distinct from canonical  $\kappa$ B sites (Bonizzi *et al.*, 2004). However, studies of the classical pathway in mammals have suggested that  $\kappa$ B site sequence does not direct binding by particular NF- $\kappa$ B proteins (Leung *et al.*, 2004). Similarly, in *Drosophila*, NF- $\kappa$ B protein-binding specificity *in vitro* has not for the most part been found to correlate with *in vivo* requirements for gene activation (Gross *et al.*, 1996; Han and Ip, 1999; Senger *et al.*, 2004). Indeed, there is evidence in the mammalian system that variations in  $\kappa$ B sequence determine which combination of NF- $\kappa$ B protein and coactivator can productively interact at a particular site (Leung *et al.*, 2004).

We report here a molecular genetic, bioinformatic, and biochemical dissection of the  $\kappa$ B code for innate immunity in *Drosophila*. These experiments reveal an unanticipated interplay between  $\kappa$ B motif elements in determining the specificity

\*Corresponding author. Section of Cell and Developmental Biology, University of California, San Diego, Bonner Hall Rm 4402, MC 0349, 9500 Gilman Drive, La Jolla, CA 92093-0349, USA.  
Tel.: +1 858 822 2408; Fax: +1 858 822 3201;  
E-mail: stevenw@ucsd.edu

Received: 14 May 2007; accepted: 22 June 2007; published online: 26 July 2007

of response and delineate the logic linking signaling pathways, transcription factors, binding sites, and effector loci.

## Results

### A transfected attacin A construct faithfully reports Toll and Imd signaling

To explore how pathway responsiveness is encoded at the DNA level, we used a molecular genetic analysis of a single locus as the foundation for a global bioinformatic approach. Attacin A (AttA), an antimicrobial peptide gene, is responsive to both Toll and Imd signaling (Asling *et al*, 1995; Lemaitre *et al*, 1997; De Gregorio *et al*, 2002; Hedengren-Olcott *et al*, 2004). Using cultured *Drosophila* Schneider (S2\*) cells, which express all three fly NF- $\kappa$ B proteins and mediate robust Toll and Imd responses (Samakovlis *et al*, 1990; Han and Ip, 1999), we assayed transcription of an AttA reporter construct (Figure 1A) upon innate immune signaling.

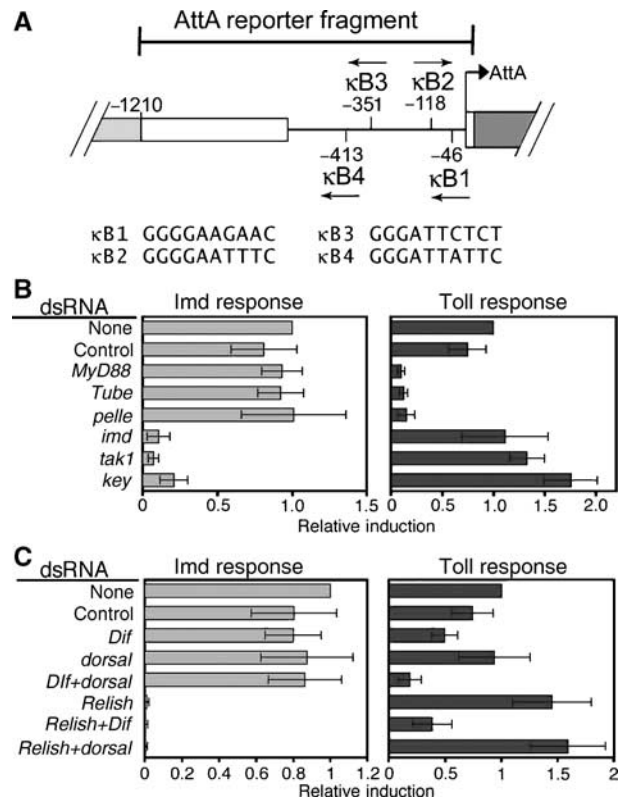
To stimulate the Toll pathway, we applied epidermal growth factor (EGF) to cells expressing EGFR-Toll, a chimera fusing the extracellular and transmembrane domains of the human EGF receptor to the intracellular domain of Toll (Sun *et al*, 2004). To specifically induce the Imd pathway, we used a peptidoglycan (PG) preparation from *Bacillus subtilis* (Leulier *et al*, 2003; Stenbak *et al*, 2004). The AttA reporter construct exhibited a robust response to either inducer, with 5- to 11-fold activation on EGF treatment and a 12- to 28-fold increase on exposure to PG.

To further validate the cultured cell system, we used RNA interference (RNAi) to inactivate innate immune response loci in 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 1B). Similarly, inactivation of Imd pathway components—*imd*, *Tak1*, or *key* (IKK $\gamma$ )—eliminated induction by PG, but not EGF. AttA induction by either innate immune pathway was strictly dependent on endogenous NF- $\kappa$ B factors (Figure 1C). 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.

### Promoter proximal $\kappa$ B sites govern AttA induction by the Toll and Imd pathways

The transcriptional start site for AttA lies approximately 650 bp downstream of the 3' end of the neighboring transcription unit (Figure 1A). Within the intergenic region lie four potential  $\kappa$ B sites (Dushay *et al*, 2000; Senger *et al*, 2004). As is characteristic of prototypical  $\kappa$ B motifs, the 5' half-site in each case contains either GGG or GGGG.

Using site-directed mutagenesis to inactivate or reposition the  $\kappa$ B-related motifs, we assayed motif function in directing expression of the AttA reporter gene. Inactivation of pairs of sites revealed that only the  $\kappa$ B motifs at positions -46 and -118 were necessary for activation by either Imd or Toll (Figure 2A). Furthermore, each of these two sites,  $\kappa$ B1 and  $\kappa$ B2, preferentially mediates signaling by one innate immune pathway (Figure 2B). Constructs containing one or more copies of  $\kappa$ B1, but not  $\kappa$ B2, responded more strongly to Imd than to Toll. Likewise, the presence of  $\kappa$ B2, but not  $\kappa$ B1, directed a stronger response to Toll than to Imd.

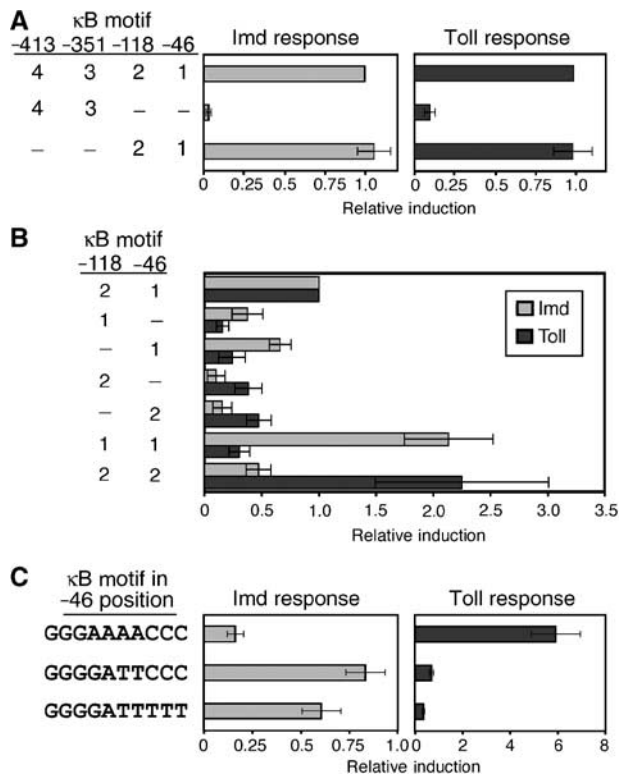


**Figure 1** An AttA reporter system recapitulates endogenous innate immune responses. (A) AttA reporter construct. The AttA genomic fragment extending from the 3' UTR of the drosocin gene (white box) to the AttA translational start site was fused to the luciferase gene (bar). The locations of four potential  $\kappa$ B motifs (Dushay *et al*, 2000; Senger *et al*, 2004), numbered from proximal to distal relative to the transcriptional start site, are indicated. Arrows designate the orientation of each  $\kappa$ B 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- $\kappa$ B 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). A full-color version of this figure is available at the EMBO Journal online.

The sequence of a single  $\kappa$ B site can thus determine the signaling pathway to which a *Drosophila* innate immune gene is preferentially responsive.

### Bioinformatic analysis reveals that Toll- and Imd-responsive loci differ in $\kappa$ B site structure and sequence

To establish whether there is a general  $\kappa$ B sequence code for innate immunity, we carried out a bioinformatics analysis based on published microarray data sets for innate immune responses in wild-type and mutant strains of *Drosophila* (De Gregorio *et al*, 2001, 2002). To identify loci responsive to Toll but not Imd, we screened 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, we identified



**Figure 2**  $\kappa$ B sites determine pathway-specific transcriptional responses. (A) Proximal  $\kappa$ B sites govern AttA induction. Mutational inactivation of  $\kappa$ B 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  $\kappa$ B 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  $\kappa$ B motif sequence, context, and number on Toll and Imd induction of AttA. Site-directed mutagenesis was used to inactivate (-) or replace  $\kappa$ B sites. All values are normalized to the wild-type construct (top row). (C) 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 a synthetic site. Normalization was as in (B). A full-color version of this figure is available at the EMBO Journal online.

genes specifically responsive to Imd as those 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 Materials and methods), we identified 16 Toll-responsive genes and 11 Imd-responsive loci. We noted good agreement of these two gene sets with a classification based on clustering of temporal expression patterns (Boutros *et al*, 2002), eleven of the Toll loci and nine of the Imd loci being common to both analyses.

To narrow the focus of our analysis, we postulated that the  $\kappa$ B 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 (Engstrom *et al*, 1993; Kappler *et al*, 1993; Whitley *et al*, 1994; Thanos and Maniatis, 1995). Extracting the corresponding genomic sequence for each locus (see Supplementary Figure 1), we used the MEME motif discovery program (Bailey and Elkan, 1994) 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  $\kappa$ B-type sequence as the highest scoring motif (Table I). The Toll and Imd gene sets were decidedly different, however, with regard to  $\kappa$ B site composition and number.

Whereas nearly two-thirds (62%) of  $\kappa$ B motifs in the Imd set had a GGGGA 5' half-site, such a half-site was absent from the  $\kappa$ B motifs in the Toll set. The Toll  $\kappa$ B motifs instead typically had either GGGGA (12 examples) or GGAA (3 examples) as the 5' half-site (Table II). 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  $\kappa$ B sites in the Toll and Imd gene sets extended throughout the motifs. Representative Toll-responsive  $\kappa$ B motifs had four or five bases between the G cluster and the first C residue, for example, GGGAAAACCC. Conversely, those Imd elements containing strings of G's and C's typically had a two or three base separation, for example, 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  $\kappa$ B 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, for example, GGGGATTTTT. Studies in mammalian systems have demonstrated that there are motifs of this type, that is, having only a single cognate  $\kappa$ B half-site, that can nevertheless bind specifically to Rel proteins *in vitro* and exhibit *cis*-regulatory activity *in vivo* (see, e.g., Whitley *et al*, 1994).

The Toll- and Imd-responsive gene sets differed not only in  $\kappa$ B site sequence, but also  $\kappa$ B site number. Of the 14 Toll genes for which  $\kappa$ B sites were detected, 12 had only a single presumptive  $\kappa$ B site and none had more than two sites. In contrast, eight of the nine Imd genes with predicted  $\kappa$ B sites had two or more sites.

### A $\kappa$ B sequence code governs fly innate immunity

To determine whether the observed differences between the  $\kappa$ B sites in the Toll and Imd gene sets correspond to a *cis*-regulatory code, we returned to the cultured cell system. We inactivated the -118 site in the AttA reporter construct, introduced synthetic versions of the  $\kappa$ B consensus motifs at the -46 position, and assayed responses to innate immune signaling.

The consensus  $\kappa$ B motifs defined by the MEME analysis mediated pathway-specific transcriptional responses in the context of the AttA reporter (Figure 2C). In response to Toll signaling, the Toll consensus sequence, GGGAAAACCC, directed reporter gene expression many times that seen for the wild-type gene. In contrast, the response to Imd signaling was well below that of the wild type. The two Imd consensus sites—GGGGATCCCC and GGGGATTTTT—also discriminated between pathways (Figure 2C). On Imd induction,

**Table I** Pathway-specific properties of innate immune  $\kappa$ B motifs

Innate immune genes			Properties of $\kappa$ B motifs			Consensus $\kappa$ B motifs						
Gene set	No. of genes in set	No. of $\kappa$ B sites in set	5' (G) <sub>4</sub> A half-site (%)	(A/T) <sub>2-3</sub> central region (%)	(A/T) <sub>4-5</sub> central region (%)	Consensus $\kappa$ B motif from MEME analysis					Mean mismatch to Toll set	Mean mismatch to Imd set
Toll	16	17	0	6	94	<u>G G G A A</u> A	A	C	C	C	1.1±0.6	3.0±1.0
							T/GT					
Imd	11	21	62	52	14	<u>G G G G A</u> T	T	T	T	T	3.7±0.8	1.4±1.0
							C	C	C			

Percentages indicate the fraction of putative  $\kappa$ B 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.

both sites provided a significant increase in reporter gene expression, as substantial as that seen for any single site introduced into AttA (Figure 2B and unpublished results). On Toll induction, the increase in reporter expression with either Imd consensus site was an order of magnitude less than with the Toll consensus site.

#### The $\kappa$ B sequence code functions by specifying NF- $\kappa$ B protein binding

There are at least two distinct mechanisms by which a  $\kappa$ B sequence code could dictate specific transcriptional responses. First, NF- $\kappa$ B proteins activated by different pathways could vary significantly in their affinity for particular  $\kappa$ B motifs. Available data implicate this mechanism in the IKK $\beta$ -independent NF- $\kappa$ B response in humans (Bonizzi *et al*, 2004). Second, a single  $\kappa$ B sequence could bind a range of NF- $\kappa$ B proteins indiscriminately, but motif sequence could determine which coactivators interact with bound NF- $\kappa$ B proteins. Such a model can explain the regulation of several mammalian genes by pairs of  $\kappa$ B motifs (Leung *et al*, 2004).

In exploring which mechanism is operative in flies, we used a gel shift assay to determine whether or not NF- $\kappa$ B proteins differ in their target site-binding preference. We expressed Relish and DIF from pGEX vectors in bacteria, purified the GST fusion proteins by affinity chromatography, and carried out binding assays with labeled oligonucleotides.

The gel shift assays demonstrated that the Relish and DIF GST fusion proteins bind synthetic  $\kappa$ B motifs *in vitro* with a specificity identical to that observed *in vivo* for signaling by Imd and Toll, respectively (Figure 3A). Using GST-Relish, we observed a strong gel shift signal for the Imd-specific sites GGGGATCCCC and GGGGATTCCC. Relish also bound well to the noncanonical Imd site, 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  $\kappa$ B sites, but also for  $\kappa$ B sites found upstream of innate immune loci (Figure 3B). For example, GST-Relish gave a strong gel shift signal with the  $\kappa$ B site from an Imd-specific locus, dipteracin, but had no detectable interaction with the  $\kappa$ B 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.

Given the consistent relationship between pathway responsiveness and Rel protein-binding specificity, we extended our gel shift analysis to all 38 of the sites identified by the MEME program. Thirty-four of the 38 sites bound to either DIF or Relish (Table II). Of the 16  $\kappa$ B sites identified in the Toll gene set, 15 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  $\kappa$ B sites in the Imd gene set, 15 bound well to Relish and none bound to DIF.

#### DIF and Relish tolerate $\kappa$ B site variation

Together, the bioinformatic, cell culture, and gel shift analyses revealed that the differential response to the Toll and Imd pathways of *Drosophila* was governed by a  $\kappa$ B sequence code that directs binding of DIF and Relish. We noted, however, variation in both the sequence and length of the active  $\kappa$ B sites bound by DIF and Relish. Work in mammalian systems had demonstrated that NF- $\kappa$ B proteins interact with a variety of related target sites via alternative patterns of hydrogen bonding DNA binding mediated by flexible polypeptide loops (Chen *et al*, 2000). To determine the extent to which DIF and Relish accommodate  $\kappa$ B sequence variation, we extended our DNA binding and cell culture studies of synthetic  $\kappa$ B motifs.

Illustrative examples from our parallel studies of  $\kappa$ B site activity *in vivo* and *in vitro* are presented in Figure 3A and Table III. The preferred half-site for Toll responsiveness and DIF binding was 5'GGGAA, with an additional G at the 5' end being well tolerated. The preferred 5' half-site for Imd responsiveness and Relish binding was GGGGA, with a GGGAA tolerated in palindromic sites. Both pathways tolerate some variation in motif length (Table III, compare rows 1 and 2 and rows 3 and 4). For example, the sites that bind Dif and are most responsive to Toll can have either one or no central residue separating half-sites based on GGGAA.

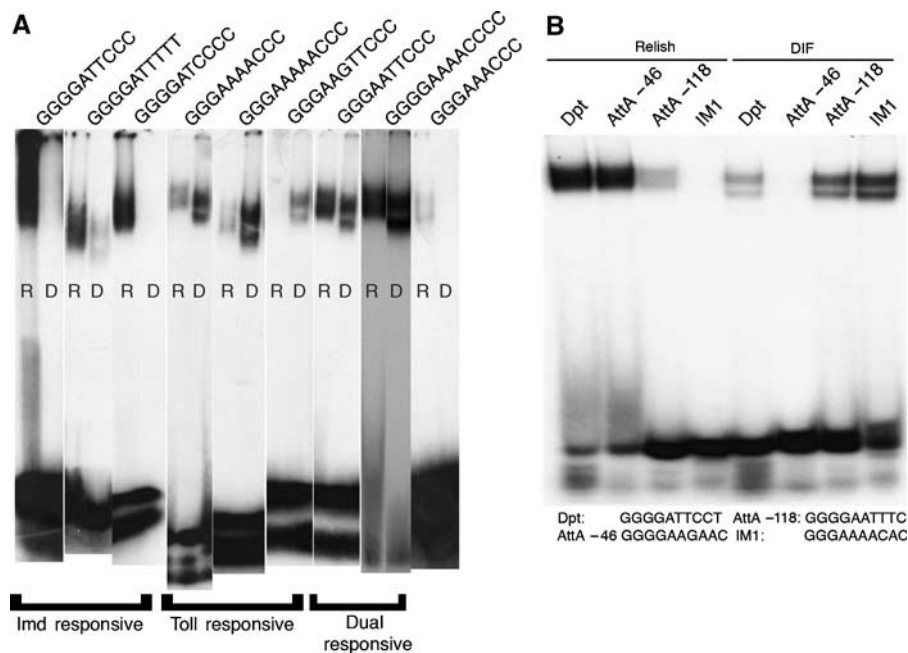
#### The $\kappa$ B sequence code can specify dual responsiveness

Expression studies have shown that many innate immune loci are dual-responsive, that is, activated by either Toll or Imd signaling. The cell culture studies revealed that dual responsiveness can reside in a single  $\kappa$ B site. As shown by the examples in Table III (rows 6 and 7), some  $\kappa$ B sequences bind

**Table II** Innate immune  $\kappa$ B motifs

Toll-responsive loci				Imd-responsive loci			
No.	Name	Position	$\kappa$ B sequence	No.	Name	Position	$\kappa$ B sequence
1	Nec	-168	GGGAAAAGCCC	1	CecA1	-80	GGGGATTTTT
2	Ser7	-144	GGGAAAACCC			-159	GGGATTCA
3	CG4757	-60	GGGTAAAGCCC	2	CecA2	-74	GGGGATTTTT
4	CG5778	-168	GGGAATTGCC			-141	GGGGATTTTT
5	CG5791	-160	CGGAAAACCCC			-175	GGGATTTTT
6	CG11841	-72	GGGAAAACCT	3	CecC	-65	GGGGAACTTTT
7	CG13422	-89	GGGAAATTCAC			-100	GGGGATCCAC
8	IM23	-77	GGGATTCGCC*	4	Def	-65	GGGATCCC
		-124	GGAAAACCCC			-129	GGGGTTTTTA*
9	CG15067		None	5	CG2056		None
10	CG16705	-133	GGGTTTGCCC	6	PGRP-SD	-111	GGGGATTTT
11	CG16772	-156	GGGAAATTC			-147	GGGGAGTGCCC
12	CG18067	-161	GGGAATGCC	7	AttD	-65	GGGAAAATCAC
13	IM2	-116	GGGAAAACAC			-165	GGGATCCC
		-170	GGAAATTC	9	PGRP-SB1	-67	GGGGTATTTA*
14	IM1	-86	GGGAAAACAC	10	DptB	-68	GGGGAATCTC*
15	IM10	-83	GGGAAATCT			-81	GGGATCCCA
		-131	GGGAATTTCT			-119	GGGGATCCAC
16	CG30080		None			-184	GGGATTCCTT
				11	Dpt	-49	GGGGATTCCT
						-145	GGGGATTCCT
						-176	GGGGATCTTT
				12	TotM		None

Motifs identified by MEME analysis for Toll- and Imd-responsive loci. Position is indicated relative to the transcriptional start site; none indicates the absence of a predicted motif for a given locus. Asterisks indicate sequences that gave negative results in  $\kappa$ B protein binding studies (see text).



**Figure 3** Dif and Relish exhibit selective binding to synthetic and endogenous  $\kappa$ B sites. GST-tagged Relish (R) or DIF (D) was incubated with labeled oligonucleotides and the resulting nucleoprotein complexes resolved by native polyacrylamide electrophoresis. (A) Synthetic  $\kappa$ B sites. (B) Endogenous  $\kappa$ B sites.

strongly to Relish and DIF. Such sites were not identified in our bioinformatic study, as expected for an analysis restricted to loci responsive only to a single pathway.

We envision two mechanisms for dual responsiveness. One would be regulation by a single site that responds to either pathway. Alternatively, a dual-responsive gene could have multiple *cis*-regulatory sites, with at least one responsive to each pathway. Regulation of metchnikowin (Mtk), a

locus that has a robust response to both Toll and Imd (De Gregorio *et al*, 2001), appears to reflect an amalgamation of the two strategies. The Mtk gene contains three potential  $\kappa$ B sites upstream of the transcriptional start. As predicted from bioinformatic analysis and confirmed by gel shift studies (Supplementary Figure 2), one, GGGAAGTCCCC, binds to both DIF and Relish, whereas two others, GGGGACTTTTT and GGGGAACCC, bind only to Relish.

**Table III** Assays with synthetic sites delineate the  $\kappa$ B sequence code

	Response	Site	Type	Relish binding	AttA Imd response	DIF binding	AttA Toll response
1	Imd	GGGGATTCCC	433	+++	100%	–	12% $\pm$ 2
2	Imd	GGGGATCCCC	424	+++	108% $\pm$ 17	–	4% $\pm$ 1
3	Toll	GGGAAAACCC	343	+/-	22% $\pm$ 4	+++	100%
4	Toll	GGGAAGTCCC	353	+/-	37% $\pm$ 4	+++	95% $\pm$ 9
6	Dual	GGGAATTCCC	343	+++	121% $\pm$ 9	+++	113% $\pm$ 10
7	Dual	GGGAAAACCCC	444	+++	79% $\pm$ 11	+++	83% $\pm$ 13
8	None	GGGAAACCC	333	+/-	11% $\pm$ 3	–	4% $\pm$ 1

Binding was measured by gel shift analysis, as shown in Figure 3. Transcriptional responses were measured as in Figure 2C and normalized to the value obtained for either the consensus Imd site (shown in row 1) or Toll site (shown in row 3). The three digits in the site type designation refer to the 5' string of G's, the central (A,T)-rich region, and the 3' string of C's.

### $\kappa$ B-site sequence and number regulate the immune response

Turning to the questions as to how  $\kappa$ B site number influences transcriptional responses, we investigated two additional innate immune loci, IM1 and AttD (Figure 4A). Using luciferase reporter genes and site-directed mutagenesis, we demonstrated that the Toll-specific IM1 gene has a single functional  $\kappa$ B site, whereas the Imd-specific AttD gene contains three  $\kappa$ B sites that are each required for a wild-type response (Supplementary Figure 3). We generated a pair of additional constructs for each gene, replacing one endogenous site with a synthetic site specific for Relish or DIF binding.

Within the divergent contexts of the IM1 and AttD loci, the Relish- and Dif-specific  $\kappa$ B sites had behaviors comparable to those observed in the AttA reporter. In particular, when we compared expression induced by each pathway, the ratio of the Imd response to the Toll response was in each case greater for the Relish site than for the Dif site. The  $\kappa$ B sequence code thus functions predictably in the context of loci varying in pathway responsiveness and  $\kappa$ B site number.

In the course of our investigation of both the AttA and AttD genes, we observed a greater than additive effect of inducing the Toll and Imd pathways simultaneously (Figure 4B). For each gene, the fold-increase in reporter gene expression on activating both Toll and Imd was significantly greater than the sum of the increases seen upon activating the two pathways individually. Furthermore, concomitant Toll and Imd activation had an apparently multiplicative effect not only in these loci, but also in drosomycin, which in S2 cells have a significant response only to Toll (Figure 4C).

To explore the origin of these effects in drosomycin expression, we mapped and mutated the functional  $\kappa$ B sites (Supplementary Figure 3). There were two such sites, varying significantly in their behavior. The site at  $-303$  bound DIF *in vitro*; inactivation of this site eliminated responsiveness to Toll. The site at  $-139$  bound to Relish; inactivation of this site left the response to Toll signaling largely intact, but abrogated the synergistic effect of activating both pathways. We interpret these results to indicate that Relish bound at  $-139$  is in itself insufficient to direct significant gene expression, but that the presence of Relish at the  $-139$  site enhances the transcriptional response directed by Dif bound at the  $-303$  site. Consistent with this hypothesis, we observed only additive effects for the pathways with a drosomycin construct having a single, albeit dual-responsive, site (Figure 4C).

Taken together, the studies of AttA, AttD, and Drs reveal that pairs of  $\kappa$ B sites can act together to enhance the strength of a transcriptional response. Furthermore, the greater than

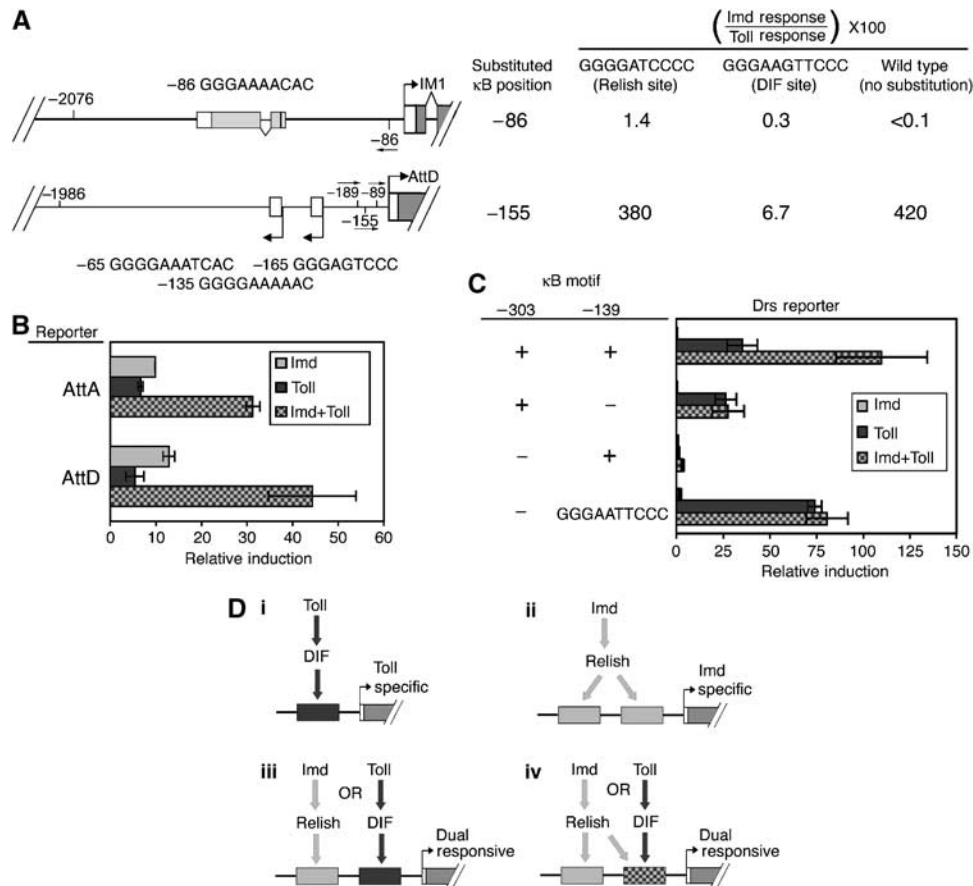
additive effect results from pairs of sites responsive to the same pathway (Figure 2B) or to different pathways (Figure 4B and C). The most parsimonious explanation for these observations is an interaction between  $\kappa$ B protein homodimers bound to distinct  $\kappa$ B sites. Based on similar findings, Ip and co-workers suggest the formation of active heterodimers when both Dif and Relish sites are present (Tanji *et al*, 2007), an explanation we cannot rule out.

### Discussion

Although innate immune loci vary with regard to their responsiveness to the Toll and Imd pathways, sorting out the regulation of individual loci has been difficult. Experiments in the past typically involved flies injected with a mixture of Gram (+) and Gram (–) bacteria, leading to the activation of at least three defensive systems—the Toll, Imd, and wound-response pathways (Lemaitre *et al*, 1996; Rutschmann *et al*, 2000; Khush *et al*, 2001; De Gregorio *et al*, 2002). To avoid this complication, we have employed a cell culture system in which we can reproducibly and rapidly activate individual response pathways. We have shown here that such an approach, specifically the use of transiently transfected reporter constructs and endogenously expressed Rel family members, recapitulates the response specificity seen in whole flies.

Taken together, the data from our molecular genetic, bioinformatic, and biochemical studies lead us to a general model for innate immune gene regulation in *Drosophila* (Figure 4D). The critical components of this model are as follows: first, pathway responsiveness is dictated by a small number of  $\kappa$ B sites located 5' to binding sites for the basal transcriptional machinery; second, the sequence of each site dictates specific responses through differences in binding affinity for the Imd effector Relish and the Toll effector DIF (and, by extrapolation, Dorsal); third, genes that respond to both Toll and Imd possess either a combination of pathway-specific sites or at least one site capable of binding both to DIF and to Relish. Such a scheme allows the fine-tuning of response strength by the influence of selective pressures on site sequence and number. Thus, for example, a gene that responds detectably only when both pathways are active might have a pair of weak binding sites, one for DIF and one for Relish.

Our results, as well as previously published studies with cecropin and dipterin (Engstrom *et al*, 1993; Kappler *et al*, 1993), indicate that  $\kappa$ B sites within several hundred base pairs of the transcriptional start site control induction of



**Figure 4** (A) The IM1 and AttD reporters. The 2 kb upstream of IM1 contains the transcript of IM23, oriented in the opposite direction. The 2 kb upstream of AttD contains two tRNA genes oriented in the opposite direction. Functional  $\kappa$ B sites were identified through mutational analysis as described in Figure 2A. Imd and Toll responses were calculated as described in Materials and methods, then presented as a ratio of Imd:Toll response  $\times 100$ . Luciferase assays were performed as described in Figure 2. Values indicated are for ratios measured with synthetic binding site (left, middle) or endogenous sites (right). (B) Simultaneous stimulation of Imd and Toll pathways results in greater than additive activation of AttA and AttD reporter constructs. Toll and Imd responses were calculated as described in Materials and methods. To directly compare Toll, Imd, and simultaneous induction, responses were not normalized to wild type. (C) Multiplicative effects require two functional  $\kappa$ B sites. The mutated  $\kappa$ B sequences are indicated in lower case letters. An asterisk marks the substitution at -139 with the GGGAAATCCC synthetic motif. (D) A mechanistic model for transcriptional regulation of innate immunity by the Toll and Imd pathways. (i, ii) Pathway-specific transcriptional responses. (iii, iv) Alternative strategies used for dual responsiveness. A full-color version of this figure is available at the EMBO Journal online.

*Drosophila* innate immune loci. Such findings are reminiscent of earlier work on the  $\beta$ -interferon and IFN- $\beta$  promoters (Whitley *et al*, 1994; Thanos and Maniatis, 1995). Those studies defined the critical function of  $\kappa$ B motifs in regulating human acute phase genes and demonstrated that induction is mediated by a combination of binding factors with promoter-proximal targets. A compact arrangement of *cis*-regulatory sequences thus appears to be a conserved property of rapid response loci.

By systematic site-directed mutagenesis and bioinformatic analysis, we have demonstrated that  $\kappa$ B motifs differ from one another in their intrinsic ability to support signaling by the Toll or Imd pathway. That is,  $\kappa$ B motifs are not only necessary for induction, but they are also sufficient to determine the specificity of induction. Whereas such a result is not unanticipated, evidence for this mechanism has been lacking, particularly on a global scale. Our results suggest a reason for the paucity of supporting data. The underlying scheme, while simple, involves variation in motif length as well as sequence. Such a code is to some extent transparent to an analysis based on varying sequence at particular positions

within a motif of fixed length and structure. Moreover, the  $\kappa$ B sites of endogenous innate immune loci vary substantially in sequence, masking the general characteristics of active sites that emerged from the bioinformatic analysis. Lastly, the presence of both pathway-specific and dual-responsive motifs in known innate immune gene sets has likely further confounded previous attempts to discern the transcriptional response code.

#### Nature of the sequence code

Our findings demonstrate a remarkable agreement among the sequence rules derived from our bioinformatic studies, the binding specificity we detect *in vitro*, and the pathway responsiveness observed in cells. Moreover, the rules make sense. For example, it is at first puzzling that 5'GGGGA half-sites are absent from the Toll-responsive gene set, as the presence of this half-site does not preclude a motif from supporting Toll-directed transcription. Note, however, that a 5'GGGGA renders a site Imd-responsive, whether or not it is Toll-responsive. The key is therefore to recognize that a

Toll-specific motif must be responsive to the Toll pathway and unresponsive to the Imd pathway.

Although the sequence rules have allowed us to identify previously unannotated innate immune loci (SH Sze, PA Pevzner, and SA Wasserman, unpublished results), we have not identified  $\kappa$ B sites upstream of every Toll- or Imd-responsive gene. There are two likely explanations. First, as the structure and extent of the spliced transcript is not well established for some loci, the sequence examined may not in fact be immediately upstream of the promoter. Second,  $\kappa$ B sites further removed from the promoter could contribute to gene expression in some cases. Indeed, the  $-303$  Toll-responsive site in drosomycin (see Figure 4C) lies outside of the boundaries within which we carried out our analysis. Furthermore, innate immune response loci are often tandemly arrayed or nested (e.g., CG15067), raising the possibility of control elements shared among genes.

Although the mean number of  $\kappa$ B sites differs between Imd- and Toll-responsive genes, a single  $\kappa$ B site is sufficient to confer pathway specificity. In Imd-specific genes, additional sites apparently function to increase the strength of the transcriptional response. It has been shown previously that the two  $\kappa$ B sites upstream of dipterin, an Imd-specific gene, act cooperatively to direct gene expression (Kappler *et al*, 1993). Some Toll loci may behave similarly. For example, the IM10 gene has two sites that bind DIF relatively weakly, but is strongly responsive to Toll activation. It may be, therefore that innate immune responses frequently rely on the cooperative effects of multiple sites for robust gene expression.

The scheme presented in Figure 4D invokes DIF and Relish homodimers, not heterodimers. Consistent with this idea, published studies indicate that only 1–2% of the Relish protein in *Drosophila* cells is in a complex with either DIF or Dorsal (Han and Ip, 1999). Although the formation of functional heterodimers might underlie cooperative effects observed upon co-transfection of Relish and Dif constructs, cooperativity might also reflect the interaction of dissimilar homodimers bound to pairs of specific  $\kappa$ B sites. Supershift experiments demonstrating the presence of both Relish and DIF at a given promoter (Han and Ip, 1999) are similarly open to alternative explanations. We cannot at present say, therefore, whether or not NF- $\kappa$ B protein heterodimers have a role in fly innate immunity.

### A $\kappa$ B code in mammals?

With regard to potential conservation of the sequence code, the closest mammalian counterparts to Relish and Dif are p50 and p65, respectively (Silverman and Maniatis, 2001; Friedman and Hughes, 2002). Data on p50 and p65 from biochemical and biophysical analyses (Kunsch *et al*, 1992; Ghosh *et al*, 1995; Muller *et al*, 1995; Chen *et al*, 1998) reveal striking parallels with the fly system. The most frequently selected targets for p50/p50 dimers are GGGGATTCCC and GGGGATCCCC. Both sites contain a GGGGA half-site and a central region of three or fewer base pairs, just as observed for Relish binding in Imd pathway targets. Similarly, p65/p65 dimers select motifs typified by GGGAATTCCC, which contains the GGGA half-site and four-residue long central region typical of motifs responsive to Toll and Dif. However, the fact that *in vivo* these two mammalian proteins exist almost exclusively as p50/p65 heterodimers argues against conservation of the sequence code. The types of

synergistic effects we see in flies may nevertheless contribute to the observed role of pairs of  $\kappa$ B sites in regulating some mammalian innate immune loci (Leung *et al*, 2004).

### Transcriptional control of innate immunity

Whereas our data suggest that pathway specificity resides exclusively in  $\kappa$ B sites, a range of studies has identified other sequence elements as bona fide or potential regulators of innate immune gene transcription. The best characterized of these at the functional level are the GATA and R1 elements (Kadalayil *et al*, 1997; Uvell and Engstrom, 2003; Senger *et al*, 2004). Such sequences, or others identified on the basis of similarity to mammalian regulatory elements (Georgel *et al*, 1993; Levashina *et al*, 1998), might contribute to discrimination between signaling pathways by providing contextual features required for NF- $\kappa$ B binding and activity. It appears more likely, however, that the major function of these additional motifs is to modulate the level or location of the response, as indicated by *in vivo* experiments with GATA elements (Senger *et al*, 2004). As the promoter-proximal region of AttA contains a well-conserved GATA element, as well as an R1 element, it should be possible to begin to address these questions by the types of experimental approaches presented here.

## Materials and methods

### Reagents and site-directed mutagenesis

Murine EGF was purchased from Calbiochem and the *B. subtilis* 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 (Sun *et al*, 2002). Six-well plates were seeded with 3 ml of  $10^6$  cells/ml. After 6 h, transfections were performed as described for the *Drosophila* Expression System (Invitrogen), with 100 ng each of reporter plasmid and pAc-LacZ. After 16–20 h, each transfection was split into thirds in fresh media and transferred to 12-well plates. One-third (1 ml) was treated with 10  $\mu$ g/ml PGN, one-third was treated with 0.4  $\mu$ g/ml EGF, and one-third was left untreated. After 4 h, cells were washed with 500  $\mu$ l PBS and then lysed in 100  $\mu$ l Luciferase Reporter Buffer (Promega). A 15  $\mu$ l portion of lysate was assayed for luciferase activity (Luciferase Assay System, Promega), and 15  $\mu$ l was assayed for  $\beta$ -galactosidase ( $\beta$ -gal) activity (Galacto-Light Plus, Tropix). Luciferase activity was normalized against  $\beta$ -gal activity to control for transfection efficiency. The increase in expression on induction was calculated by subtracting the uninduced value from the induced value. Dividing this number by the uninduced value yielded the percentage increase in expression. As indicated in the figure legends, the percentage increase was in many cases normalized to a value of one for the response of a wild-type construct to a particular pathway. RNAi experiments were carried out as described previously (Sun *et al*, 2002).

### Bioinformatic analysis of *Drosophila* innate immune response loci

Using published microarray data (De Gregorio *et al*, 2002), 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*, blocking Toll signaling, or a mutation in *relish*, disrupting Imd signaling. The specific



criteria for assigning loci to each class are presented in Supplementary Table 2.

For each locus, we used available cDNA sequence data, as well as sequence profiles for TATA box and initiator elements (Ohler *et al*, 2002), to define the core promoter region. For 11 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 Supplementary Figure S1. For MEME analysis (Bailey and Elkan, 1994), 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 (Ip *et al*, 1993) was a gift from Tony Ip. A pGex-Rel construct was generated by inserting bp 1-1635 of Relish (amplified from cDNA) into the *Bam*HI and *Eco*RI 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 (Stoven *et al*, 2003). Plasmids were transformed into BL21 and expressed proteins purified as described previously (Ip *et al*, 1992), with the following modifications. A 500 ml culture was grown at 30°C to OD<sub>550</sub> ~0.5, then induced for 1 h with 1.6 mM IPTG. Cells were pelleted by centrifugation at 6000 g for 15 min at 4°C. The pellet was resuspended in 10 ml 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, 1 × EDTA-free protease inhibitor cocktail (Roche)). Cells were lysed by performing three freeze-thaw cycles, followed by three cycles of 10 s sonication and 30 s incubation on ice. To 5 ml of bacterial cell suspension, 500  $\mu$ l of 5 M NaCl was added, and the suspension was incubated on ice for 15 min. Cellular debris was pelleted by centrifugation at 12 000 g for 15 min at 4°C. Next, 750  $\mu$ l 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 h. Sepharose beads were pelleted by centrifugation at 500 g for 5 min 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  $\mu$ l of 50 mM Tris (pH 8.0) plus 10 mM glutathione and gentle rocking at 4°C for 1 h.

### Gel-shift assays

Electrophoretic mobility shift assays were performed as described previously (Ip *et al*, 1991) with the following modifications. Reactions were carried out in a volume of 10  $\mu$ l containing 0.1  $\mu$ g/ $\mu$ l poly(dI-dC), 10 mM HEPES, 50 mM NaCl, 1 mg/ml BSA, 3 mM MgCl<sub>2</sub>, 6 mM  $\beta$ ME, 10 mM EDTA, and 10% (v/v) glycerol. End-labeled probe (0.8 pmol) was combined with protein, and incubated for 10 min at room temperature. Protein was used at a concentration of 1.5  $\mu$ g per reaction in experiments with synthetic binding sites and 7.5  $\mu$ g 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  $\kappa$ B motif. Synthetic  $\kappa$ B motifs were substituted for the -46 motif in AttA such that the position of the last G in the 5' core element was kept constant. The  $\kappa$ B motifs from AttA (-46 and -118), dipterocin (-145), and IM1 (-86) were each analyzed in their endogenous context.

### Supplementary data

Supplementary data are available at *The EMBO Journal* Online (<http://www.embojournal.org>).

## Acknowledgements

We thank S Ho and C Glass for use of luminometers in their laboratories, K Pogliano for supplying an initial PG preparation, the laboratory of J Posakony for assistance with gel shift experiments, and Y Ip for both plasmid DNA and informative conversations. E Bier, G Ghosh, A Hoffmann, A Letsou, W McGinnis, and L Zipursky provided helpful comments on a draft manuscript. This work was supported by NIH Grant 5R01-GM50545 to SAW. MSB was supported by funds from the NIH Training Program in Basic Clinical Genetics, 2T32GM008666-06. CPA was supported in part by Cal (IT)<sup>2</sup>.

## References

- Asling B, Dushay MS, Hultmark D (1995) 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* **25**: 511-518
- Baeuerle PA (1991) The inducible transcription activator NF-kappa B: regulation by distinct protein subunits. *Biochim Biophys Acta* **1072**: 63-80
- Bailey TL, Elkan C (1994) Fitting a mixture model by expectation maximization to discover motifs in biopolymers. *Proc Int Conf Intell Syst Mol Biol* **2**: 28-36
- Bonizzi G, Bebiën M, Otero DC, Johnson-Vroom KE, Cao Y, Vu D, Jegga AG, Aronow BJ, Ghosh G, Rickert RC, Karin M (2004) Activation of IKKalpha target genes depends on recognition of specific kappaB binding sites by RelB:p52 dimers. *EMBO J* **23**: 4202-4210
- Boutros M, Agaisse H, Perrimon N (2002) Sequential activation of signaling pathways during innate immune responses in *Drosophila*. *Dev Cell* **3**: 711-722
- Brennan CA, Anderson KV (2004) *Drosophila*: the genetics of innate immune recognition and response. *Annu Rev Immunol* **22**: 457-483
- Chen FE, Huang DB, Chen YQ, Ghosh G (1998) Crystal structure of p50/p65 heterodimer of transcription factor NF-kappaB bound to DNA. *Nature* **391**: 410-413
- Chen YQ, Sengchanthalangsy LL, Hackett A, Ghosh G (2000) NF-kappaB p65 (RelA) homodimer uses distinct mechanisms to recognize DNA targets. *Structure* **8**: 419-428
- De Gregorio E, Spellman PT, Rubin GM, Lemaitre B (2001) Genome-wide analysis of the *Drosophila* immune response by using oligonucleotide microarrays. *Proc Natl Acad Sci USA* **98**: 12590-12595
- De Gregorio E, Spellman PT, Tzou P, Rubin GM, Lemaitre B (2002) The Toll and Imd pathways are the major regulators of the immune response in *Drosophila*. *EMBO J* **21**: 2568-2579
- Dushay MS, Asling B, Hultmark D (1996) Origins of immunity: Relish, a compound Rel-like gene in the antibacterial defense of *Drosophila*. *Proc Natl Acad Sci USA* **93**: 10343-10347
- Dushay MS, Roethele JB, Chaverri JM, Dulek DE, Syed SK, Kitami T, Eldon ED (2000) Two attacin antibacterial genes of *Drosophila melanogaster*. *Gene* **246**: 49-57
- Engstrom Y, Kadalayil L, Sun SC, Samakovlis C, Hultmark D, Faye I (1993)  $\kappa$ B-like motifs regulate the induction of immune genes in *Drosophila*. *J Mol Biol* **232**: 327-333
- Friedman R, Hughes AL (2002) Molecular evolution of the NF-kappaB signaling system. *Immunogenetics* **53**: 964-974
- Georgel P, Meister M, Kappler C, Lemaitre B, Reichhart JM, Hoffmann JA (1993) Insect immunity: the dipterocin promoter contains multiple functional regulatory sequences homologous to mammalian acute-phase response elements. *Biochem Biophys Res Commun* **197**: 508-517
- Ghosh G, van Duyn G, Ghosh S, Sigler PB (1995) Structure of NF-kappa B p50 homodimer bound to a kappa B site. *Nature* **373**: 303-310
- Ghosh S, May MJ, Kopp EB (1998) NF-kappa B and Rel proteins: evolutionarily conserved mediators of immune responses. *Annu Rev Immunol* **16**: 225-260
- Gross I, Georgel P, Kappler C, Reichhart JM, Hoffmann JA (1996) *Drosophila* immunity: a comparative analysis of the Rel proteins dorsal and Dif in the induction of the genes encoding dipterocin and cecropin. *Nucleic Acids Res* **24**: 1238-1245
- Han ZS, Ip YT (1999) Interaction and specificity of Rel-related proteins in regulating *Drosophila* immunity gene expression. *J Biol Chem* **274**: 21355-21361
- Hedengren-Olcott M, Olcott MC, Mooney DT, Ekengren S, Geller BL, Taylor BJ (2004) Differential activation of the NF-kappaB-like factors relish and Dif in *Drosophila melanogaster* by fungi and gram-positive bacteria. *J Biol Chem* **279**: 21121-21127

- Ip YT, Kraut R, Levine M, Rushlow CA (1991) The dorsal morphogen is a sequence-specific DNA-binding protein that interacts with a long-range repression element in *Drosophila*. *Cell* **64**: 439–446
- Ip YT, Park RE, Kosman D, Bier E, Levine M (1992) The dorsal gradient morphogen regulates stripes of rhomboid expression in the presumptive neuroectoderm of the *Drosophila* embryo. *Genes Dev* **6**: 1728–1739
- Ip YT, Reach M, Engstrom Y, Kadalayil L, Cai H, Gonzalez-Crespo S, Tatei K, Levine M (1993) Dif, a dorsal-related gene that mediates an immune response in *Drosophila*. *Cell* **75**: 753–763
- Irving P, Troxler L, Heuer TS, Belvin M, Kopczynski C, Reichhart JM, Hoffmann JA, Hetru C (2001) A genome-wide analysis of immune responses in *Drosophila*. *Proc Natl Acad Sci USA* **98**: 15119–15124
- Kadalayil L, Petersen UM, Engstrom Y (1997) Adjacent GATA and kappa B-like motifs regulate the expression of a *Drosophila* immune gene. *Nucleic Acids Res* **25**: 1233–1239
- Kappler C, Meister M, Lagueux M, Gateff E, Hoffmann JA, Reichhart JM (1993) Insect immunity. Two 17 bp repeats nesting a kappa B-related sequence confer inducibility to the dipterin gene and bind a polypeptide in bacteria-challenged *Drosophila*. *EMBO J* **12**: 1561–1568
- Karin M, Ben-Neriah Y (2000) Phosphorylation meets ubiquitination: the control of NF-kappaB activity. *Annu Rev Immunol* **18**: 621–663
- Khush RS, Leulier F, Lemaitre B (2001) *Drosophila* immunity: two paths to NF-kappaB. *Trends Immunol* **22**: 260–264
- Kunsch C, Ruben SM, Rosen CA (1992) 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* **12**: 4412–4421
- Lemaitre B, Emmanuelle N, Michaut L, Reichhart J-M, Hoffmann JA (1996) The dorsoventral regulatory gene cassette spatzle/Toll/cactus controls the potent antifungal response in *Drosophila* adults. *Cell* **86**: 973–983
- Lemaitre B, Reichhart JM, Hoffmann JA (1997) *Drosophila* host defense: differential induction of antimicrobial peptide genes after infection by various classes of microorganisms. *Proc Natl Acad Sci USA* **94**: 14614–14619
- Leulier F, Parquet C, Pili-Floury S, Ryu JH, Caroff M, Lee WJ, Mengin-Lecreux D, Lemaitre B (2003) The *Drosophila* immune system detects bacteria through specific peptidoglycan recognition. *Nat Immunol* **4**: 478–484
- Leung TH, Hoffmann A, Baltimore D (2004) One nucleotide in a kappaB site can determine cofactor specificity for NF-kappaB dimers. *Cell* **118**: 453–464
- Levashina EA, Ohresser S, Lemaitre B, Imler JL (1998) Two distinct pathways can control expression of the gene encoding the *Drosophila* antimicrobial peptide metchnikowin. *J Mol Biol* **278**: 515–527
- Muller CW, Rey FA, Sodeoka M, Verdine GL, Harrison SC (1995) Structure of the NF-kappa B p50 homodimer bound to DNA. *Nature* **373**: 311–317
- Ohler U, Liao GC, Niemann H, Rubin GM (2002) Computational analysis of core promoters in the *Drosophila* genome. *Genome Biol* **3**: 087
- Rutschmann S, Jung AC, Hetru C, Reichhart JM, Hoffmann JA, Ferrandon D (2000) The Rel protein DIF mediates the antifungal but not the antibacterial host defense in *Drosophila*. *Immunity* **12**: 569–580
- Samakovlis C, Kimbrell DA, Kylsten P, Engstrom A, Hultmark D (1990) The immune response in *Drosophila*: pattern of cecropin expression and biological activity. *EMBO J* **9**: 2969–2976
- Senger K, Armstrong GW, Rowell WJ, Kwan JM, Markstein M, Levine M (2004) Immunity regulatory DNAs share common organizational features in *Drosophila*. *Mol Cell* **13**: 19–32
- Silverman N, Maniatis T (2001) NF-kappaB signaling pathways in mammalian and insect innate immunity. *Genes Dev* **15**: 2321–2342
- Stenbak CR, Ryu JH, Leulier F, Pili-Floury S, Parquet C, Herve M, Chaput C, Boneca IG, Lee WJ, Lemaitre B, Mengin-Lecreux D (2004) Peptidoglycan molecular requirements allowing detection by the *Drosophila* immune deficiency pathway. *J Immunol* **173**: 7339–7348
- Stoven S, Ando I, Kadalayil L, Engstrom Y, Hultmark D (2000) Activation of the *Drosophila* NF-kappaB factor Relish by rapid endoproteolytic cleavage. *EMBO Rep* **1**: 347–352
- Stoven S, Silverman N, Junell A, Hedengren-Olcott M, Erturk D, Engstrom Y, Maniatis T, Hultmark D (2003) Caspase-mediated processing of the *Drosophila* NF-kappaB factor Relish. *Proc Natl Acad Sci USA* **100**: 5991–5996
- Sun H, Bristow BN, Qu G, Wasserman SA (2002) A heterotrimeric death domain complex in Toll signaling. *Proc Natl Acad Sci USA* **99**: 12871–12876
- Sun H, Towb P, Chiem DN, Foster BA, Wasserman SA (2004) Regulated assembly of the Toll signaling complex drives *Drosophila* dorsoventral patterning. *EMBO J* **23**: 100–110
- Tanji T, Hu X, Weber ANR, Ip YT (2007) Toll and IMD pathways synergistically activate an innate immune response in *Drosophila melanogaster*. *Mol Cell Biol* **27**: 4578–4588
- Thanos D, Maniatis T (1995) Identification of the rel family members required for virus induction of the human beta interferon gene. *Mol Cell Biol* **15**: 152–164
- Tzou P, Reichhart JM, Lemaitre B (2002) Constitutive expression of a single antimicrobial peptide can restore wild-type resistance to infection in immunodeficient *Drosophila* mutants. *Proc Natl Acad Sci USA* **99**: 2152–2157
- Uvell H, Engstrom Y (2003) Functional characterization of a novel promoter element required for an innate immune response in *Drosophila*. *Mol Cell Biol* **23**: 8272–8281
- Whitley MZ, Thanos D, Read MA, Maniatis T, Collins T (1994) A striking similarity in the organization of the E-selectin and beta interferon gene promoters. *Mol Cell Biol* **14**: 6464–6475
- Xiao C, Ghosh S (2005) NF-kappaB, an evolutionarily conserved mediator of immune and inflammatory responses. *Adv Exp Med Biol* **560**: 41–45