THE JOURNAL OF BIOLOGICAL CHEMISTRY
Vol. 278, No. 49, Issue of December 5, pp. 49428–49437, 2003

*Printed in U.S.A.*

# Transcriptional Regulation of Limulus Factor C

## REPRESSION OF AN NFκB MOTIF MODULATES ITS RESPONSIVENESS TO BACTERIAL LIPOPOLYSACCHARIDE*

Received for publication, June 23, 2003, and in revised form, August 22, 2003
Published, JBC Papers in Press, August 29, 2003, DOI 10.1074/jbc.M306641200

**Lihui Wang‡, Bow Ho§, and Jeak Ling Ding‡¶**

*From the Departments of ‡Biological Sciences and §Microbiology, National University of Singapore, 14 Science Drive 4, Singapore 117543*

**Serine proteases play fundamental roles in invertebrate development, hemostasis, and innate immunity. This is exemplified by the limulus Factor C, which is a serine protease that binds a pathogen-associated molecule, lipopolysaccharide (LPS) to trigger a blood coagulation cascade. As a central molecule in the limulus innate immunity and hemostasis, Factor C gene expression has been detected in two major immune defense tissues, the amebocytes and hepatopancreas. Infection of the limulus with live Gram-negative bacteria induces a 2–3-fold increase in mRNA transcripts in both tissues. However, *in vitro* studies in *Drosophila* cell lines using Factor C promoter-reporter chimera DNA constructs, and site-directed mutagenesis of the promoter demonstrated that a proximal κB binding site, aided by an adjacent dorsal-like binding motif responds dramatically to LPS and dorsal transcription factor overexpression. Electrophoretic mobility shift assay further confirmed a strong interaction of the limulus κB motif with Rel proteins. However, deletion constructs of the Factor C promoter harboring different numbers of dorsal-like binding sites upstream of the κB motif as well as the electrophoretic mobility shift assay of these motifs with Rel proteins strongly suggest that the up-regulation of Factor C gene expression is attenuated during microbial challenge. The repression of the dramatic activation of this pathogen-responsive gene by LPS is probably effected via competition between the dorsal-like motifs over the proximal LPS-responsive κB unit, or through inhibition from the upstream repressive element(s), which accounts for the gene expression pattern observed *in vivo*. Our findings demonstrate that blood coagulation and innate immune response are integrated at the transcriptional level in this ancient organism, and that this LPS-responsive serine protease is controlled by an evolutionarily conserved NFκB pathway.**

Serine protease cascades are indispensable in various fundamental biological processes in both invertebrates and vertebrates (1). They play major roles in signal transduction in development, immunity, and hemostasis. In *Drosophila*, a well established serine protease cascade is involved in generating a signal that promotes ventral and lateral development of the embryo (2). The proteolytically processed product of this cascade, Spätzle, is thought to be recognized by the Toll receptor, which finally results in the translocation of dorsal transcription factor from the cytoplasm into the nucleus (2, 3). The Dorsal concentration gradient regulated by this pathway controls the dorsoventral polarity in the early *Drosophila* embryo. Dorsal belongs to the Rel protein family, which is commonly referred to as NFκB protein in mammals. In *Drosophila*, in addition to Dorsal, two other Rel proteins have been identified: Dif (4) and Relish (5). The Rel protein family in particular controls transcription of many immune responsive genes (6, 7). Therefore, it is not surprising that the Spätzle/Toll/Dorsal pathway was also found to be involved in the expression of antimicrobial peptides during fungal and Gram-positive bacterial infection in *Drosophila* larvae (8). Nevertheless, in the adult fly, Dif substitutes the role of Dorsal to control the antifungal and anti-Gram-negative bacteria responses. In contrast, Relish is the Rel protein required in the lipopolysaccharide (LPS)[1]-mounted immune response against Gram-negative bacteria. Thus, the dorsoventral determination pathway and the antifungal and Gram-positive bacterial immune responses converge on the proteolytic processing of Spätzle. However, another uncharacterized serine protease cascade(s) distinguishable from the GD/SNK/Ea cascade has been suggested to cleave Spätzle into its active form (9).

Fig. 1 illustrates an interesting parallel in the two serine protease cascades: the *Drosophila* dorsoventral polarity determination and the limulus LPS-sensitive blood coagulation (9, 10). In the limulus blood coagulation, the serine protease, Factor C is activated by LPS (also known as endotoxin, a component of the outer cell wall of Gram-negative bacteria) and subsequently undergoes autocatalytic cleavage to expose its catalytic activity. Consequently, the LPS-activated Factor C initiates proteolytic processing to convert coagulogen to coagulin, which polymerizes to form an insoluble clot at the site of infection/injury (11, 12). Coagulogen is so far the only reported structural homologue to Spätzle in invertebrates (13, 14). The upstream serine proteases in both cascades show a significant sequence homology between Factor B and proclotting enzyme *versus* Snake and Easter (10). Remarkably, Persephone, a newly discovered serine protease involved in signal transduc-

* This work was supported by the Agency for Science Technology and Research, A*STAR, Grant LS/99/004 (to J. L. D. and B. H.). The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

*The nucleotide sequence(s) reported in this paper has been submitted to the GenBank™/EBI Data Bank with accession number(s) AF517565 and S77063.*

¶ To whom correspondence should be addressed: Dept. of Biological Sciences, National University of Singapore, 14 Science Drive 4, Singapore 117543. Tel.: 65-68742776; Fax: 65-67792486; E-mail: dbsdjl@nus.edu.sg.

[1] The abbreviations used are: LPS, lipopolysaccharide; CAT, chloramphenicol acetyltransferase; CrFC, *Carcinoscopius rotundicauda* Factor C; EMSA, electrophoretic mobility shift assay; mbn-2, *Drosophila melanogaster* malignant blood leukemia 2; S2, *Drosophila melanogaster* Schneider line-2; TF, tissue factor; TK, thymidylate kinase; GST, glutathione *S*-transferase; PIPES, 1,4-piperazinediethanesulfonic acid; IL, interleukin; TLR, Toll-like receptor.

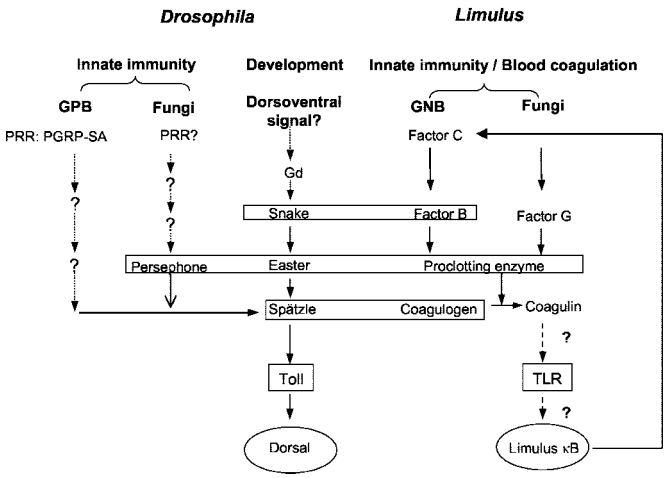


FIG. 1. **Serine protease cascades in development, innate immunity, and blood coagulation.** On the *left* are the serine protease cascades in dorsoventral determination, immune response against Gram-positive bacteria (*GPB*) and fungal infection in *Drosophila*. *Dotted arrows* with "?" indicate unidentified components in the cascades. On the *right* are the serine protease cascades in limulus blood coagulation and innate immunity, which are activated by Gram-negative bacteria (*GNB*) and fungi, respectively. Factor G is the upstream serine protease in the alternate blood coagulation pathway that is triggered by β1,3-glucan. *Discontinuous arrows* annotate the putative LPS-mediated signaling pathway as proposed under "Discussion." Homologues in all the cascades are *boxed*. *Gd*, gastrulation defective.

tion of anti-Gram-positive bacteria and anti-fungal immune responses, is deduced to contain structural domains homologous to those in Easter, Snake (of the dorsoventral determination serine protease cascade), and proclotting enzymes (of the limulus blood coagulation) (15). Although serine proteases involved in blood coagulation and innate immunity have not been well characterized in *Drosophila*, its genome, however, contains a high number of genes encoding putative proteases, including 199 members that appear to code for trypsin-like serine proteases and many others that code for proteins with fibrinogen- and complement-like domains (16).

In mammalian hemostasis, the tissue factor (TF)-dependent pathway is the primary initiator of blood coagulation. Upon injury, Factor VII complexes with the exposed mature TF on the cell surface to activate Factor X to Factor Xa. Factor Xa converts prothrombin to thrombin, which eventually leads to fibrin clot formation (17–19). During inflammation, blood coagulation can also be modulated by LPS and various cytokines through regulating the expression of coagulation related factors. These two fundamental biological processes, *viz.* inflammation and blood coagulation, are integrated into an essential part of the host immune response in higher vertebrates. The simultaneous activation of the innate immune response and the coagulation system after injury is a phylogenetically ancient, adaptive response that can be traced back to the early stages of evolution in eukaryotes as exemplified by the limulus LPS initiated blood coagulation cascade (20, 21). Therefore, it is tantalizing to speculate that a single ancestral developmental/immunity cascade has given rise to the developmental process, blood coagulation, and innate immunity in the animal kingdom (9, 15, 22). In this respect, the limulus blood coagulation cascade contributes an invaluable system for studying the evolution of serine protease cascades in hemostasis and innate immune response.

In innate immune response against microbial infection, pattern recognition receptors recognize and bind bacterial cell wall components and metabolic products like LPS and β-1,3-glucan (23) to transduce the signal to a downstream serine protease cascade or other signaling pathway. Peptidoglycan recognition proteins constituting a large protein family have been characterized as pattern recognition receptors that sense microbial infection in both vertebrates and invertebrates (24, 25). Gram-positive bacterial infection in *Drosophila* causes peptidoglycan recognition protein SA to recognize and activate Toll for antimicrobial peptide synthesis (26) (Fig. 1). To date, the limulus Factor C is the only known pattern recognition receptor that has a unique mosaic structure combining LPS recognition and serine protease activity (27, 28). Unlike Factor C, other pattern recognition receptors including peptidoglycan recognition proteins show no known serine protease activity and do not share any common structure, and indirectly activate serine proteases involved in the immune response.

The functional and evolutionary connection of blood coagulation with innate immunity makes it an interesting subject to investigate the expression pattern of Factor C during microbial infection. Once activated by LPS, Factor C must be regulated spatially and temporally to limit and localize the immune response in the hemolymph. Currently, it is reported that the regulation of the limulus coagulation cascade appears to be limited to three types of serpin-related serine protease inhibitors. To date, research on the regulation of serine proteases involved in blood coagulation or developmental processes in other organisms has focused on the feedback regulation from within the cascade or the protein-protein interaction from protease inhibitors under normal physiological conditions. Although the limulus blood coagulation constitutes a primitive serine protease cascade, the knowledge we gained previously from the study on the protein structure and function has provided insights into the roles of serine proteases in fundamental biological processes (11, 29, 30). By examining how the Factor C gene is regulated under bacterial infection at the transcriptional level, we sought to provide in-depth knowledge on the regulatory mechanisms controlling other serine proteases in immunologically challenging conditions. We report here the characterization of the *Carcinoscorpius rotundicauda* Factor C (*CrFC*) 5′-flanking region, *cis*-acting sites in the promoter region, and cognate transcription factors involved in its regulation. We demonstrate the transcriptional regulation of this unique serine protease during microbial infection and propose an *in vivo* model involving the interplay and competition among multiple *cis*-acting sites and/or probable upstream repressive elements that oppose a proximal κB binding element.

## EXPERIMENTAL PROCEDURES

*Materials*—Horseshoe crabs (*C. rotundicauda*) were collected from Kranji estuary, Singapore. Gram-negative bacteria, *Pseudomonas aeruginosa*, ATCC 27853, were cultured in tryptic soya broth (Difco, Detroit, MI). LPS from *Escherichia coli* O55:B5 was a product of Sigma. Human recombinant NFκB protein, p50, was from Promega. The mbn-2 cell line was a gift from Professor Y. Engstrom, Stockholm University, Sweden. Glutathione-Sepharose™ 4B was from Amersham Biosciences. Unless otherwise stated, all enzymes were from PerkinElmer Life Sciences, and molecular biology grade chemicals were from Sigma.

*Infection of Limulus—P. aeruginosa* was cultured overnight in TSB at 37 °C. Bacteria was pelleted at 5,000 × *g* for 10 min at 4 °C, washed in saline, and resuspended to the original culture volume in saline. Serial dilution and colony count were performed to determine colony forming units. Each time point constituted three experimental animals. The horseshoe crabs were injected with $2.5 \times 10^6$ colony forming units/kg of body weight. At various time points, the viable bacterial count in the hemolymph was determined by plating 100 μl of cell-free hemolymph or dilutions of it on TSA and cetrimide agar. Colonies were counted after overnight incubation at 37 °C.

*Northern Analysis*—The tissue distribution of *CrFC* gene expression was studied by Northern analysis (31). Hemolymph obtained by cardiac puncture (29) were centrifuged at 100 × *g* for 5 min at 4 °C to obtain amebocytes. Hepatopancreas, heart, intestine, and muscle were excised and immediately frozen in liquid nitrogen prior to homogenization just before use. Total RNA was isolated with TRIZOL™ RNA extraction reagent (Invitrogen). Twenty μg of total RNA from each tissue was

electrophoresed in a formaldehyde-treated 1% agarose gel and transferred overnight to a nylon membrane in 20× SSC solution (3 M NaCl and 0.3 M sodium citrate). The membrane was hybridized with a 2.1-kb fragment of the *CrFC* cDNA (GenBank™ accession number S77063) probe labeled by random priming with Rediprimer™ II (Amersham Biosciences). The membrane was then washed sequentially with 1× SSC, 0.1% SDS; 0.5× SSC, 0.1% SDS; and 0.1× SSC, 0.1% SDS for 15 min each at 37 °C. Finally, the membrane was exposed overnight to x-ray film (Kodak) at −80 °C. For loading control, an 800-bp limulus actin-3 gene (32) was used as probe. The existence of limulus Dorsal homologues was also probed by using a 2-kb fragment of *Drosophila* dorsal cDNA (33).

*Real Time PCR*—Quantitative kinetic analysis of *CrFC* gene expression in response to *Pseudomonas* injection was conducted with the ABI PRISM™ 7000 system (Applied Biosystems). Total RNA from amebocytes and hepatopancreas were extracted at 0, 3, 6, 9, 12, and 24 h after injection. mRNAs were extracted using the Oligotex-direct mRNA purification kit (QIAgen). Aliquots of 200 ng of mRNA were used for every 20 μl of cDNA synthesis reaction, and 2 μl of the template was used in a 25-μl PCR reaction. The cDNA template was synthesized using Taqman™ reverse transcription reagents (Applied Biosystems). PCR was performed using SYBR™ Green PCR master mixture (Applied Biosystems). The primer concentrations were consistently at 600 nM. The results were analyzed by a comparative quantitative method using actin-3 as the endogenous control.

*Isolation of the CrFC Promoter*—The genomic DNA was extracted from amebocytes and various genomic subclones were obtained by a PCR-based method of genome walking with the Universal Genome Walker system (Clontech). Briefly, various pools of adaptor-ligated genomic DNA fragments were generated. These genome walker libraries were used for PCR amplification, employing an outer adaptor primer (AP1) and a primary gene-specific primer (GSP1). A secondary nested PCR amplification was carried out with AP1 and another gene-specific primer (GSP2). The PCR-amplified genomic fragments were blunt end-ligated into the pT7Blue-3 Perfectly Blunt™ vector system (Novagen). Thus, the Factor C promoter and the 5′ region of its genomic sequence were obtained from the genomic library using gene-specific primer: 5′-GACACCAAAAACGACGCTAAGACCATAC-3, which contains the Factor C exon 1 sequence (GenBank™ accession number S77063). A fragment of 1.5 kb containing the leader signal and exon 1 was isolated and subcloned into pT7B3 vector. The 5′-flanking region in the resulting pT7B3pro1.5 clone was sequenced (GenBank™ accession number AF517565).

*Mapping Transcription Initiation Site by Primer Extension Assay*—The primer extension assay was performed according to Mason *et al.* (34) with various modifications. A probe derived entirely from within exon 1, 5′-GACACCAAAAACGACGCTAAGACCATAC-3′, was labeled with [$\gamma$-$^{32}$P]ATP by T4 polynucleotide kinase. Hybridization was carried out with 2.5 μg of mRNA and 2.5 fmol of the radiolabeled probe (specific activity of $10^6$ cpm/pmol) in hybridization buffer (400 mM NaCl and 10 mM PIPES, pH 6.4) for 5–6 h at 55 °C. The primer extension reaction was performed using the annealed primer/mRNA mixture with the Superscript™ II RNase H$^-$ reverse transcriptase (Invitrogen). The primer extension product was resolved in 6% denaturing polyacrylamide gel, alongside the A, T, G, and C sequencing ladder generated by the TaqTrack manual sequencing system (Promega).

*Construction of CrFC Promoter-Reporter Plasmids*—To define essential *cis*-elements that transcriptionally control the *CrFC* gene expression, a nest of primers was designed to obtain sequential deletion constructs with the end fixed at the +1 site. The CAT reporter gene was used in the transfection assay for analyzing the promoter strength. pTKCAT vector was mutated to pMutTKCAT, in which two endogenous dorsal binding sites in the original vector were deleted using QuikChange™ site-directed mutagenesis kit (Stratagene). Thus, using pT7B3pro1.5 as the template and a series of primers, different deletion fragments of the *CrFC* promoter were subcloned into compatible sites of the pMutTKCAT vector. In all the constructs, the *CrFC* promoter sequence (up to +43) was fused in-frame with the CAT reporter gene. The deletion constructs tested in the transient transfection assay extended to positions −1252, −1017, −682, −395, and −186, relative to the transcription start site (+1), and were termed pFL (full-length), p1017, p682, p395, and p186, respectively. All the plasmids used in the transfection assay were prepared by alkaline lysis and further purified through CsCl/ethidium bromide density gradient ultracentrifugation.

*Cell Culture and Transfection Assay*—*Drosophila* S2 cells were used in transient transfection assays to define the regulatory sequence in the *CrFC* promoter region. Mbn-2 cells, which are responsive to LPS, were used to study the effect of LPS on the *CrFC* promoter. The S2 cells were maintained in DES expression medium (Invitrogen), supplemented with 10% fetal bovine serum, 50 IU/ml penicillin, and 50 μg/ml streptomycin (Invitrogen). Mbn-2 cells were cultured in Schneider's *Drosophila* medium (Sigma), supplemented with 10% fetal bovine serum, 100 IU/ml penicillin, and 100 μg/ml streptomycin. Cells were seeded 24 h before transfection at a density of 1.5–2 × $10^6$ cells/3.5-cm culture well. Transfections were performed using calcium phosphate/DNA co-precipitation (Invitrogen) with 5 μg of reporter DNA, 1 μg of pACH110 (containing the β-galactosidase gene fused with the *Drosophila* actin promoter) as internal control (33), and 13 μg of pBluescript SKII plasmid to adjust the total amount of DNA in each transfection to 19 μg. In cotransfection assays, 1 μg of *Drosophila* pPACfl(dorsal) (33) expression vector was used in each transfection with 12 μg of the pBluescript SKII. After 24 h, the culture medium was renewed and 48 h later, cells were harvested. CAT and β-galactosidase activities in the cell extracts were measured using the CAT and β-galactosidase enzyme-linked immunosorbent assay kits, respectively (Roche Diagnostics). Expression of CAT reporter gene was divided by β-galactosidase expression in each transfection experiment to normalize the transfection efficiency difference. To investigate the effect of LPS on the *CrFC* gene promoter, LPS from *E. coli* O55:B5 (Sigma) was added at 10 μg/ml into each well of transfected mbn-2 cells and incubated for 8 h before harvesting.

*Mutation of CrFC-κB and Its Adjacent GAAAA Motif*—To further study the putative *cis*-elements in the *CrFC* promoter, the CrFC-κB motif at −133 bp and its adjacent GAAAA motif at −146 bp were mutated. By designing three sets of primers, the CrFC-κB motif was mutated by deleting its canonical GGGAAA sequence. The adjacent GAAAA motif was mutated by removing the AAAA site. The two sites were also combinatorially mutated using primers complementary to the template containing mutation of the CrFC-κB site. Primers used for this purpose were: 5′-GCACGAGAAAAAAGCCTCCATTAGAAATCAGCGCG-3′ (sense) and 5′-CGCGCTGATTTCTAATGGAGGCTTTTTTCTCGTGC-3 (antisense) for mutation of the CrFC-κB; 5′-CTGTCATGCACGAGGCCGGGAAATCC-3′ (sense) and 5′-GGATTTCCCGGCCTCGTGCATGACAG-3 (antisense) for mutation of the adjacent GAAAA motif; 5′-GTCTGTCATGCACGAGGCCTCCATTAGAAATCAGCG-3′(sense) and 5′-CGCTGATTTCTAATGGAGGCCTCGTGCATGACAGAC-3′ (antisense) for double mutation of the two sites.

*Recombinant Expression and Partial Purification of Dorsal Protein*—Rel and DNA binding domains of the dorsal transcription factor were amplified by PCR from the pPACfl(dorsal) expression vector. SmaI and NotI enzyme sites were introduced during PCR cycles and the fragments were subcloned into pGEX-4T-1 (Amersham Biosciences). The primers were: 5′-GCTGCGGCCGCGAGGGTGTGAGCTCAGG-3′ (antisense) for the Rel domain amplification and 5′-GCTGCGGCCGCCTACACAGCCGGCAGAT-3′ (antisense) for the DNA binding domain. For PCR cycles, both antisense primers were paired with the same sense primer: 5′-GATCCCGGGGAACCAGAACAATGGAGCCG-3′. Briefly, the recombinant protein was expressed in *E. coli* BL21. After 3 h induction with 0.1 mM isopropyl-1-thio-β-D-galactopyranoside at 30 °C, 500 ml of culture was pelleted and resuspended in 40 ml of phosphate-buffered saline. Cell lysate was prepared by using a French press (Basic Z 0.75KW Benchtop Cell Disruptor, UK) operated at 30 k.p.s.i., and immediately centrifuged at 15,000 × *g* at 4 °C for 30 min. The supernatants containing either recombinant DNA binding domain or Rel domain of dorsal protein were separately applied at a flow rate of 0.5 ml/min through the Sepharose™ 4B GST affinity column of 1-ml bed volume. The column was washed with 10 ml of phosphate-buffered saline at a flow rate of 1.5 ml/min, followed by elution of the bound protein with 5 ml of elution buffer (50 mM Tris-HCl, 10 mM reduced glutathione, pH 8.0) at the same flow rate. The recombinant proteins were confirmed by Western analysis with anti-GST antibody (Amersham Biosciences).

*Electrophoretic Mobility Shift Assay (EMSA) with Dorsal and Human p50*—In each binding reaction with human recombinant NFκB p50 (Promega), 1 × $10^5$ cpm/pmol probe was incubated with 20 ng of p50 protein in 30 μl of binding buffer at 25 °C for 30 min before electrophoresis on a native 6% PAGE gel (with acrylamide and bisacrylamide ratio of 79:1). The binding buffer for p50 contained 10 mM HEPES, 50 mM KCl, 1 mM EDTA, 2.5 mM dithiothreitol, 10% glycerol, and 0.05% Nonidet P-40. Binding assay with the recombinant *Drosophila* dorsal protein was performed according to Ip *et al.* (35), with some modifications. Recombinant Rel or DNA binding domain of dorsal protein at 50–100 ng each was used in each binding reaction, which was incubated at 25 °C for 30 min before PAGE gel analysis. The binding buffer for recombinant Dorsal contained 10 mM HEPES, pH 7.9, 50 mM NaCl, 3 mM $MgCl_2$, 6 mM β-mercaptoethanol, 10 mM EDTA, and 10% glycerol. In competition assay, 10× and 100× of cold probes were used in each

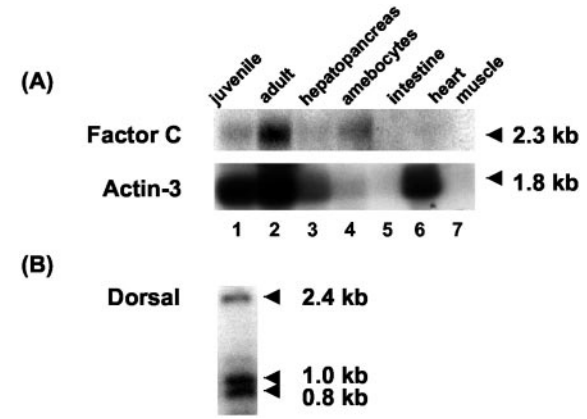


FIG. 2. **Tissue-specific expression of *CrFC* and elucidation of dorsal transcription factor homologues in the limulus.** *A*, Northern analysis of *CrFC* with total RNA from various tissues. Lanes: *1*, all tissues excluding amebocytes of a juvenile animal; *2*, all tissues excluding amebocytes of an adult animal; *3*, hepatopancreas; *4*, amebocytes; *5*, intestine; *6*, heart; and *7*, muscle. Limulus actin-3 in the *lower panel* was probed to normalize loading variations in the amount of RNA sample in each lane. *B*, Northern hybridization with a 2.0-kb *Drosophila* Dorsal cDNA probe. Three major bands of putative limulus dorsal transcript are evident, with molecular sizes of 2.4, 1.0, and 0.8 kb.

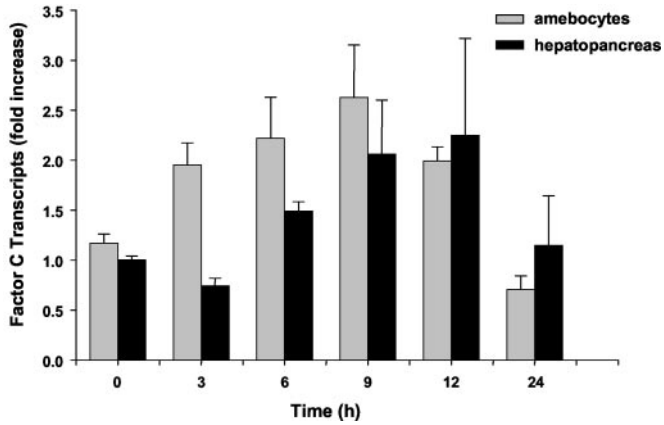


FIG. 3. **Determination of *P. aeruginosa*-induced *CrFC* expression by real time PCR.** The kinetic profile of *CrFC* expression in amebocytes and hepatopancreas were analyzed over 24 h of infection and expressed as -fold increase in transcript at each time point compared with 0 h (naive uninfected) using primers: 5′-TCCACAGAACCATATGATGG-3′ (sense) and CCACGGTTACCCCTCCCATTAGCTAGA-3′(antisense). All values were normalized against limulus actin-3 control, where the following primers were used: 5′-TGATGCCCCTCGTGC-3′(sense) and 5′-ATACCACGTTTGCTTTGGG-3′(antisense). Data are presented as mean ± S.D. of three independent experiments.

reaction. Nuclear extract from the limulus amebocytes was prepared according to Dignam *et al.* (36). EMSA with the amebocyte nuclear extract was performed under the same conditions as in the binding assay with the recombinant Dorsal.

## RESULTS

*Tissue Specificity of CrFC Gene Expression and Bacteria-induced Expression—CrFC* expression is specifically localized to the amebocytes and hepatopancreas. Differential expression levels of the gene were observed in these two tissues. Amebocytes express more abundantly than hepatopancreas (Fig. 2*A*). Although the biological function of hepatopancreas Factor C remains unknown, it is highly likely that this tissue is also involved in the innate immune response of limulus. As a pivotal molecule in limulus immune response, it is important to determine the expression pattern of *CrFC* gene under Gram-negative bacterial challenge. Thus, real time PCR indicates that within 9–12 h of *P. aeruginosa* infection, expression of *CrFC* was increased by 2–3-fold in both amebocytes and hepatopancreas (Fig. 3).

*Functional Characterization of CrFC 5′-Flanking Region—* To fully understand the transcriptional regulation of the *CrFC* gene, it is pertinent to characterize its upstream regulatory sequence. Primer extension assay using mRNA from the limulus amebocytes mapped the transcription start site to a single "T" (shown as the antisense strand in Fig. 4), which is located 43 bp upstream of the ATG start codon, and represented by the complementary "A" on the promoter sequence in sense strand (Fig. 5). A 1.5-kb fragment of DNA containing the *CrFC* promoter and exon 1 was obtained by genome walking. A noncanonical TATA sequence, TATTAATA, is located 23 bp upstream of the transcription start site. No consensus CAAT box was found in the *CrFC* promoter region (Fig. 5). Other mammalian coagulation proteins: Factor IX (37), Factor X (38, 39), Factor XII (40), Factor VII (41, 42), and prothrombin (43) have also been shown to lack typical TATA and CAAT box in their 5′-flanking region.

Sequential deletion constructs were generated to define the regulatory elements that direct the expression of the *CrFC* gene. Because of the lack of homologous cell lines from limulus, but its evolutionary closeness to insects, *Drosophila* S2 were chosen as host cells in which the *CrFC* gene promoter was characterized. The S2 cell line was derived from a primary

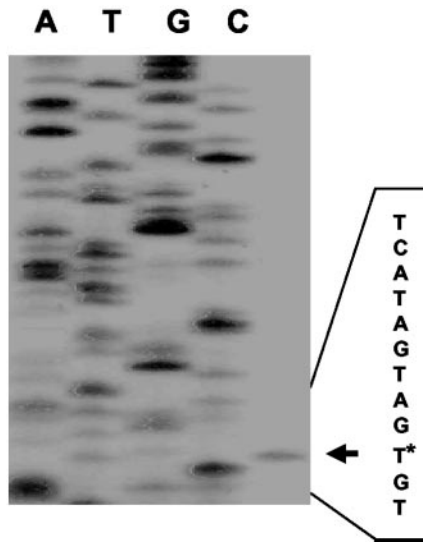


FIG. 4. **Mapping of the transcription start site of the *CrFC* gene.** The chemical sequencing ladders are ATCG. The *arrow* indicates the primer extension product, which shows the transcription start site at T in the antisense strand.

culture of late stage (20–24 h old) *Drosophila melanogaster* embryo. The close similarity between limulus amebocytes and the *Drosophila* S2 cells make the latter an ideal *in vitro* system for examining *CrFC* transcriptional regulation. Transient transfection was carried out to delineate the basal promoter region in S2 cells.

Deletion of promoter sequences from −1252 down to −1017 did not significantly alter CAT reporter gene expression, but a longer deletion to −682 caused a 12-fold increase in CAT expression. Further truncation of the promoter down to position −395 retained a higher basal promoter activity with a 4-fold increase in CAT expression. But a longer deletion to −186 restored the promoter activity to an 8-fold increase in CAT expression (Fig. 6*A*). Therefore, promoter sequence from +1 to −682 confers the highest basal promoter activity. However, sequence from +1 to −186 is sufficient to maintain a high expression level. Further analysis of the whole 1.26-kb *CrFC* promoter with TRANSFAC 6.0 (BIOBASE, Germany) revealed

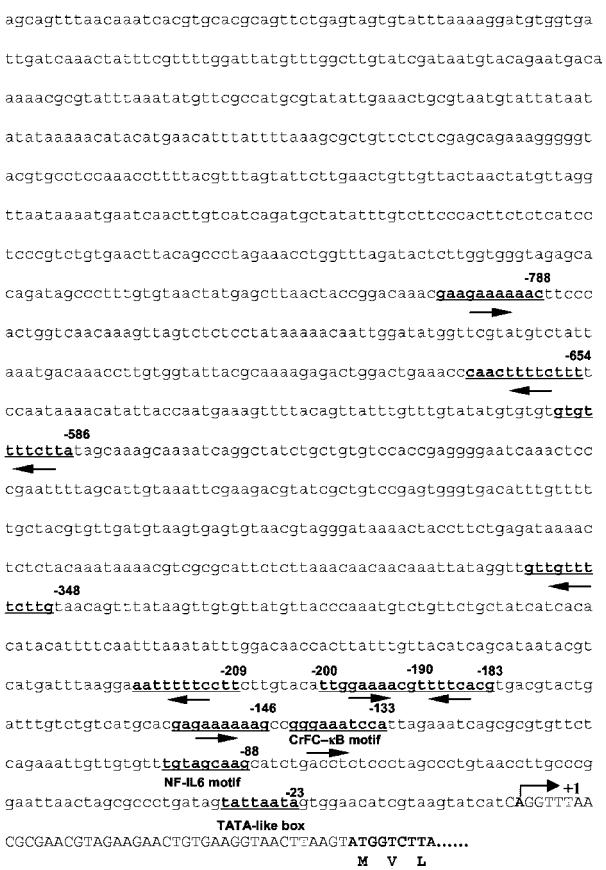


FIG. 5. ***CrFC* promoter sequence with illustrations of immunity related *cis*-acting sites.** *Numbers in bold* indicate the relative positions of the dorsal-like binding sites with respect to the transcription start site (+1), which is an A in the sense strain of the promoter sequence. The *CrFC-κB* and *NF-IL6* motifs are *labeled*. All the putative binding sites are *bold* and *underlined*. The orientation of each dorsal-like site in the promoter region is indicated by the *direction of arrows*. The sequence shown here is in the sense strand.

seven putative dorsal-like binding sites within p682 and another located at −788. All 8 dorsal-like motifs contain a central GAAAA sequence. However, the 5′ and 3′ ends of the central motif show significant variation within the *CrFC* dorsal-like motifs and with *Drosophila* counterparts (33). The distribution of the 8 dorsal-like binding motifs is illustrated in Fig. 5. Strikingly, within the region from +1 to −186, a *cis*-acting site found at −133 (5′-GGGAAATCCA-3′) resembles the classic NFκB binding sequence: 5′-GGGRNNYYCC-3′ (6). Henceforth this motif is referred to as CrFC-κB. Intriguingly, a dorsal-like motif bearing the sequence, GAGAAAAAAG, is only 2 bp away from the CrFC-κB motif. In addition to these potential *cis*-acting sites, an NF-IL6 binding site is present at position −88 bp, with a conserved sequence similar to the mammalian counterparts: 5′-T(T/G)NNGNAA(T/G)-3′ (44). Both the NFκB and NF-IL6 are important transcription factors involved in immune responsive gene regulation.

Is the CrFC-κB site the positive element that determines the high promoter activity observed in short constructs: p682, p395, and p186? Mutagenesis study on the CrFC-κB motif was performed to further investigate the role of CrFC-κB in *CrFC* transcriptional regulation. The dorsal-like binding motif at −146, which is 2 bp away from the CrFC-κB site, was also mutated to examine its effects on the CrFC-κB, because it has been reported that a GAAANN sequence adjacent to the κB-like motif modulates the κB transcription factor-controlled expression of antimicrobial peptides in *Drosophila* (45). We observed that mutation of the CrFC-κB motif alone completely abolished the higher basal promoter activity observed in constructs p682, p395, and p186, whereas no significant change in reporter gene expression was observed in longer constructs, pFL and p1017. Mutation of CrFC-κB completely abolished the promoter activity, whereas mutation of the adjacent dorsal-like binding site only reduced the reporter gene expression. The combinatorial mutation of these two sites is comparable with the single mutation of the CrFC-κB binding motif (Fig. 6*A*).

*LPS Induces Factor C Promoter Activity*—In light of a 2–3-fold up-regulation of *CrFC* gene expression during infection by *P. aeruginosa* (Fig. 3), we sought to examine the influence of LPS over the promoter. *Drosophila* mbn-2 cells derived from *Drosophila* blood leukemia were chosen for this study because of its responsiveness to LPS (46). After LPS stimulation, the pFL construct with the full-length promoter sequence and the p1017 construct lacking the sequence upstream of −1017, showed a 2-fold increase in reporter gene expression, a level that matches the *in vivo* situation. Exposure of transfected mbn-2 cells to 10 μg/ml LPS increased the expression of p682, p395, and p186 by 10-, 4-, and 16-fold, respectively. However, transfectants with constructs harboring the mutated CrFC-κB motif remained unresponsive to LPS. Deletion of the GAAAA sequence also caused a significant down-regulation of the promoter activity after LPS treatment (Fig. 6*B*).

*Dorsal Transcription Factor Strongly Activates the CrFC-κB*—The LPS-induced up-regulation of *CrFC* promoter activity in mbn-2 transfectants prompted us to investigate the potential *trans*-acting factors in *CrFC* gene transcription. The effect of overexpression of *Drosophila* Dorsal on the promoter was conducted in S2 cells. It was observed that co-expression of Dorsal increased the transcriptional activities of pFL, p1017, p682, and p395 by 11-, 20-, 65-, and 22-fold, respectively. Unexpectedly, the reporter gene in the shortest construct, p186, was enhanced by 400-fold. As anticipated, in all the constructs the dramatic up-regulation of CAT reporter gene expression exerted by dorsal transcription factor overexpression was completely eliminated by mutation of the CrFC-κB motif alone. Deletion of the GAAAA sequence in the adjacent dorsal-like motif alone did not totally abolish but significantly diminished the reporter gene up-regulation (Fig. 6*A*).

*Differential Binding Characteristics of Dorsal-like Motifs*—EMSA was performed to study the binding characteristics of the CrFC-κB and other dorsal-like motifs with Rel proteins. A synthetic 22-mer oligonucleotide containing the CrFC-κB motif exhibited sequence-specific binding activity with p50, the human NFκB protein. Compared with oligonucleotide probes containing consensus mammalian NFκB binding sequences, CrFC-κB probes were less efficient in binding p50 (Fig. 7*A*). To further delineate the interaction of all the dorsal-like binding sites and the CrFC-κB motif with dorsal transcription factors, recombinant DNA binding and Rel domains were expressed and partially purified from *E. coli*. Table I lists the CrFC-κB and dorsal-like motifs in the *CrFC* promoter used in EMSA. ComκB probes contain the CrFC-κB and its adjacent dorsal-like motifs. Prox3Dor probes were synthesized with three dorsal-like motifs: two overlapping sites spanning from −183 to −200 and another at −209. Dor348, Dor586, Dor654, and Dor788 probes each contain a dorsal site at −348, −586, −654, and −788, respectively. ComκB and Prox3Dor oligonucleotides showed specific binding to both the Rel and DNA binding domains of Dorsal. However, only ComκB probe, which contains the CrFC-κB motif, can bind to human p50 protein (Fig. 7, *B* and *C*). Dor788 showed specific recognition and binding of dorsal protein. However, no interaction was observed between Dor788 and p50 (Fig. 7*D*). On the other hand, Dor348, Dor586, and Dor654, did not exhibit binding activity to either Dorsal or

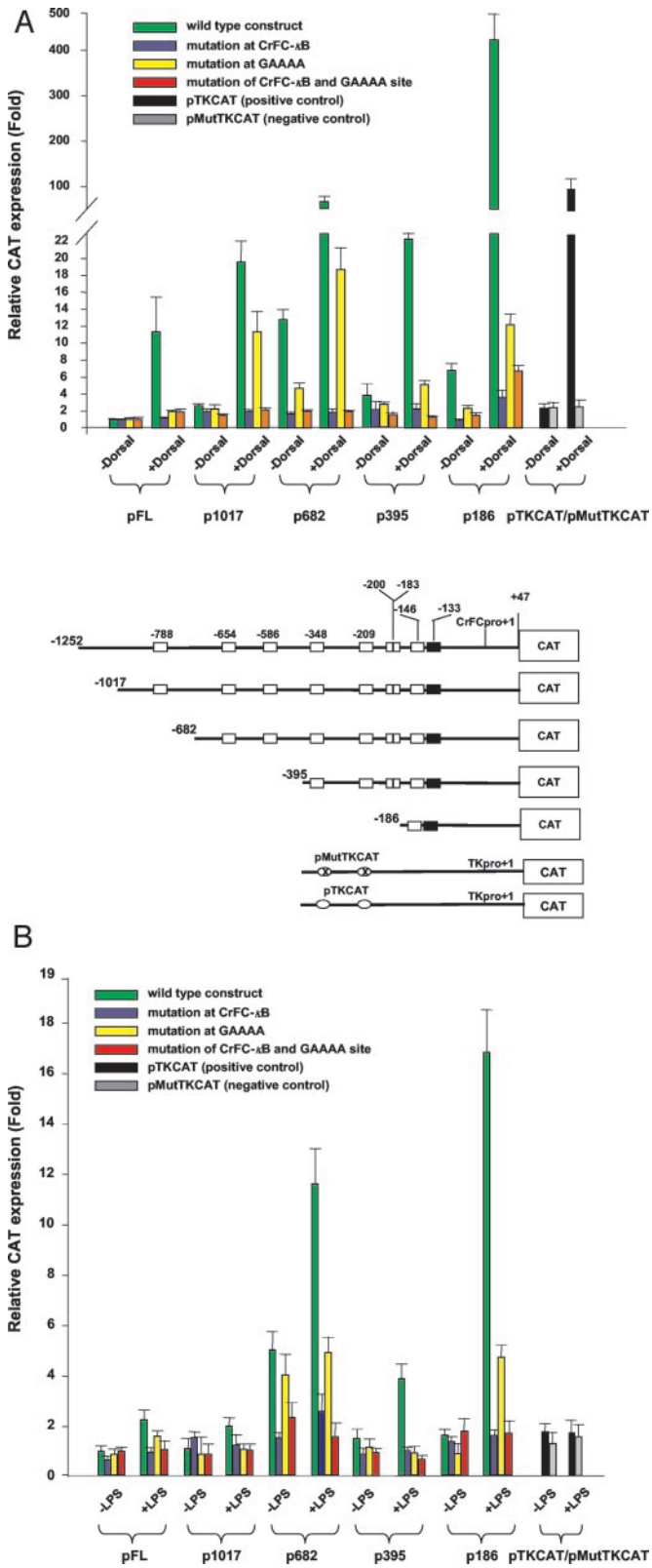


FIG. 6. **Effects of LPS and Dorsal overexpression on *CrFC* promoter activity before and after mutagenesis of CrFC-κB and its adjacent dorsal-like binding motif.** *A*, overexpression of Dorsal in S2 cell-mediated *CrFC* promoter activity. The increase in the relative CAT activity driven by the different deletion constructs of the *CrFC* promoter is compared with the wild type pFL (full-length) construct. The pTKCAT and pMutTKCAT are used as positive and negative controls, respectively. Schematic illustration of various *CrFC* promoter and CAT fusion constructs is shown on the *lower panel*. The *numbers to the left* of each construct represent the position relative to the transcription start site (+1). Dorsal-like binding motifs are represented by *open squares*. The CrFC-κB motif is shown as a *closed square*. The two overlapping dorsal motifs are shown as *connected open squares*. The position of each *cis*-acting site is indicated in the construct. *B*, the effects of LPS on the *CrFC* gene promoter expression studied in mbn-2 cells. After 36 h post-transfection, the cells were treated for 8 h with 10 μg/ml LPS from *E. coli* strain O55:B5. Results are shown as -fold increase of the CAT reporter gene expression after normalization against the wild type full-length (*pFL*) construct without LPS treatment. Data are presented as mean ± S.D. of three independent experiments. Primers for the construction of plasmids used in transfection assay are forward primers: 5′-CAAGAAGCTTCACGTGACGTACTGATTTG for p186; 5′-CAAGAAGCTTGTCGCGCATTCTCTTAAAC-3′ for p395; 5′-CAAGAAGCTTGAGACTGGACTGAAACCCA-3′ for p682; 5′-CAAGAAGCTTATGTTCGCCATGCGTATAT-3′ for p1017; 5′-CAAGAAGCTTTAACATATCACGTGCACGC-3′ for pFL with the same reverse primer: 5′-GGCAGATCTACTTAAGTTACCTTCACAGTTCTT-3′ containing the +1 site of the *CrFC* gene.

p50 (Fig. 7*E*). The binding characteristics of all the probes with Rel proteins are summarized in Table I.

The existence of limulus Dorsal/NFκB protein was confirmed by both EMSA and Northern analysis. CrFC-κB probe formed a prominent complex with nuclear extract from amebocytes (Fig. 7*F*). In Northern analysis, possible homologues of dorsal transcripts were also evidently cross-hybridized with *Drosophila* dorsal cDNA (Fig. 2*B*).

## DISCUSSION

*Transcriptional Regulation of CrFC under Gram-negative Bacterial Infection*—The horseshoe crab, often referred to as a "living fossil," has survived unchanged for ~400 million years of evolution. Besides the unique evolutionary role this species plays in the animal kingdom, its potent innate immune response has been a subject of intense research. In particular, significant progress has been made on the biochemical isolation and characterization of various immunity-related molecules from this organism (12, 47). The well characterized blood coagulation cascade in limulus, which combines the function of hemostasis with immune response against Gram-negative bacteria, plays a unique role in the development of serine protease cascades in both blood coagulation and innate immune defense from invertebrates to vertebrates.

Factor C, the first zymogen of this cascade to recognize and bind LPS, has long been thought to be expressed exclusively in amebocytes and only released by degranulation into the hemolymph upon Gram-negative bacterial infection. However, by Northern analysis, we discovered the expression of *CrFC* in hepatopancreas, which strongly suggests a potential circulating form of Factor C that may function like a coagulant in the blood vessels of higher vertebrates. It is highly likely that this extracellular Factor C recognizes LPS to initiate the degranulation of amebocytes to release more intracellular Factor C for hemostasis and incapacitation of the invading pathogen.

The blood clotting process has apparently co-evolved with the innate immune response because of the early differentiation and divergence in metazoans. In mammals, cytokines not only activate the inflammatory response but also the clotting cascade, which is evidenced in the cytokine-controlled TF expression. Besides cytokines like tumor necrosis factor α and interleukin 1β (IL-1β), LPS has been discovered to up-regulate TF transcription through the NFκB pathway in monocytes via the κB binding site characterized in the proximal promoter region of the TF gene (48, 49). Although cytokines are also reported to be involved in serine protease regulation of blood coagulants like Factor VII (50, 51) and Factor XII (52), there are no reports on their transcriptional regulation during inflammation induced by microbial infection. Instead, research on the transcriptional regulation of mammalian blood coagulation factors has focused mainly on the tissue-specific expression pattern in

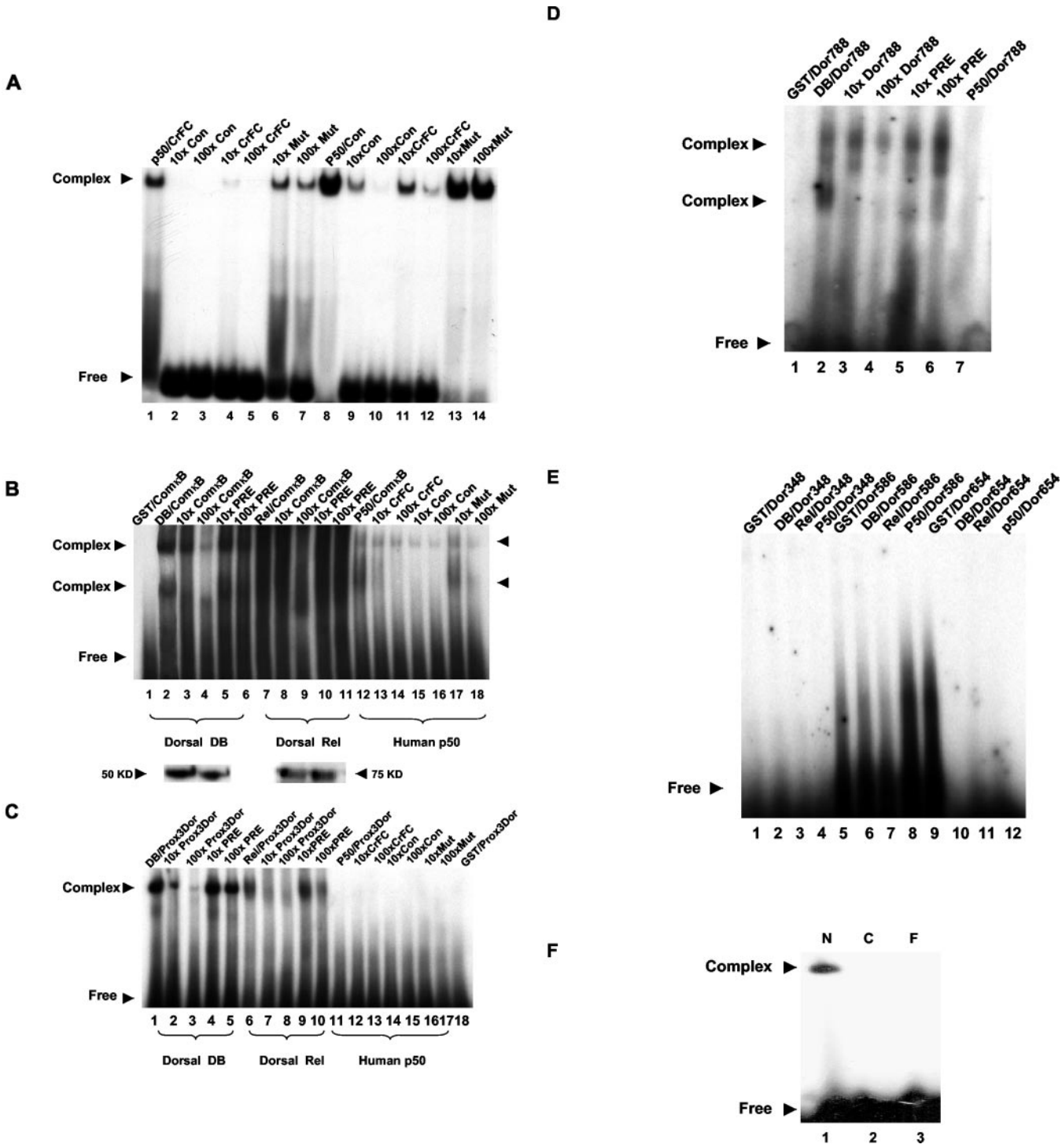


FIG. 7. **EMSA of CrFC-κB motif and dorsal-like binding sites with Rel proteins.** *A*, competition assay of CrFC-κB probes (oligonucleotides containing the CrFC-κB motif) with human p50 NFκB protein. Sequences of all probes used in this assay are listed in Table I. *Con*, *CrFC*, and *Mut*, respectively, represent probes with the consensus mammalian NFκB motif (5′-AGTTGAGGGGACTTTCCCAGGC-3′), CrFC-κB motif, and the mutated CrFC-κB motif (5′-AAAAGCCattAAATCCATTAGA-3′; where the 5′ end of the binding motif, GGG, was mutated to att as indicated by *lowercase* and *underlined*). Excess 10× and 100× of the cold probes (unlabeled) were applied to compete against DNA and p50 complex as indicated. *B*, competition assay of the ComκB probe (oligonucleotide containing complete sequences of CrFC-κB and its adjacent dorsal-like motif) with human p50, DNA binding domain (*DB*), and Rel domains of recombinant *Drosophila* Dorsal. Binding reaction with GST protein without the dorsal expression represents the negative control (*lane 1*). Two specific DNA-protein complexes were formed with both the DNA binding and Rel domains (*lanes 2* and *7*). Excess 10× and 100× unlabeled ComκB probe were used to compete against the DNA-protein complex of ComκB with DNA binding domain and with Rel (*lanes 3*, *4*, and *8*, *9*). Unlabeled PRE probes (containing progesterone responsive elements: 5′-CTAGACTGCTTTGTTTGCT-GAGAACATGGTGTTCTGCACTTAAACACACT-3′) at 10× and 100× excess were applied as nonspecific competitor to the DNA-protein complex (*lanes 5*, *6* and *10*, *11*). In *lanes 13–18*, 10× and 100× CrFC, Con, and Mut probes were added to compete against the complex formed between p50 and ComκB (*lane 12*, p50/ComκB). Of the two DNA-protein complexes, only the *lower complex* shows binding specificity with p50 in the competition assay. Western analysis of DNA binding and Rel domains of dorsal protein after GST affinity column chromatography is shown in the *lower panel*. *C*, competition assay of Prox3Dor probe (oligonucleotides containing three dorsal-like motifs proximal to the CrFC-κB motif), 10× and 100× excess of unlabeled Prox3Dor and PRE were applied as competitors. The binding assay of GST protein alone is the negative control in *lane 18*. *D*, competition assay of Dor788 (Dorsal-like motif at −788 of the *CrFC* promoter) with the DNA binding domain of Dorsal. *Lane 1* is the GST negative control. From *lanes 3* to *6*, 10× and 100× excess of cold probes: unlabeled Prodor788 and PRE probes were used to compete against the

TABLE I
*Probes used in EMSA*

| Oligonucleotide probes | DNA sequence | *cis*-Elements studied | Binding characteristics | |
|---|---|---|---|---|
| | | | p50 | Dorsal |
| CrFC-κB | 5′-AAAAGCCGGGAAATCCATTAGA-3′ | (−143) 5′-GGGAAATCCA-3′ (−133) | + | + |
| ComκB | 5′-CATGCACGAGAAAAAAGCCGGGAAATCCATTAGA-3′ | (−143) 5′-GGGAAATCCA-3′ (−133) | + | + |
| | | (−156) 5′-CGA**GAAAA**AAG-3′ (−146) | | |
| | | (−200) 5′-TTG**GAAAA**CGT-3′ (−190) | | |
| Prox3Dor | 5′-GAAATTTTTCCTTCTTGTACATTGGAAAACGTTTT CACGTGACGTACTGATTTGTCTGTCATGCA-3′ | (−183) 5′-CGT**GAAAA**CGT-3′ (−193)[a] | | |
| | | (−209) 5′-AAG**GAAAA**ATT-3′ (−219)[a] | − | + |
| Dor348 | 5′-GTTGTTGTTTTCTTGTAACAG-3′ | (−348) 5′-CAA**GAAAA**CAA-3′ (−359)[a] | − | − |
| | | (−586) 5′-TAA**GAAAA**CAC-3′ (−596)[a] | | |
| Dor586 | 5′-GTGTGTGTTTTCTTATAGCA-3′ | | − | − |
| Dor654 | 5′-AACCCAACTTTTCTTTCCAA-3′ | (−654) 5′-AAA**GAAAA**GTT-3′ (−664)[a] | − | − |
| Dor788 | 5′-CAAACGAAGAAAAAACTTCC-3 | (−798) 5′-GAA**GAAAA**AAC-3′ (−788) | − | + |

[a] Sequences shown are in the antisense strand. The central sequence of dorsal-binding motifs is shown in bold.

hepatocytes, and liver-specific transcription factors have been implicated in the transcriptional regulation of Factor VII (42), Factor X (53), and prothrombin (43) under normal physiological circumstances. Although both NFκB and NF-IL6 binding motifs were found in the 5′-flanking region of Factor VII, no further research was performed to elucidate the roles of these immune responsive *cis*-acting sites under inflammatory conditions. Being an invertebrate free of adaptive immunity, plus the convenience of microbial challenge, the horseshoe crab is a model of choice for exploring the transcriptional regulation of serine proteases involved in innate immunity and blood coagulation.

First, we examined the expression profile of CrFC after *P. aeruginosa* infection. Instead of a prompt and dramatic up-regulation of expression as observed for antimicrobial peptides in *Drosophila*, expression of *CrFC* in both the amebocytes and hepatopancreas was only moderately increased by 2–3-fold. Analysis of the *CrFC* gene promoter reveals a non-canonical but functional TATA box in *Drosophila* cell lines. The lack of typical TATA and absence of CAAT box is a common characteristic shared by other blood coagulants as described earlier. When devoid of further upstream sequence, the *CrFC* promoter between +1 to −682 (which contains seven dorsal-like binding sites) is capable of directing the highest basal promoter activity in S2 cells. Strikingly, we observe an NFκB motif (CrFC-κB), located proximal to the TATA-like box, which is the recognition and binding site for NFκB proteins. More interestingly, an NF-IL6 binding site, which is another important immune and inflammation responsive transcription factor binding site, was found 40 bp downstream of the CrFC-κB motif.

Only deletion constructs that are devoid of sequences upstream of −682 and containing the intact CrFC-κB motif exhibited significant response to LPS, indicating that the CrFC-κB motif is the essential positive *cis*-acting site during the LPS challenge. The corresponding *trans*-activating factors were further investigated by overexpression of dorsal transcription factor. The dramatic increase (400-fold) observed in p186 promoter substantially indicates a strong interaction and transactivation of the dorsal transcription factor with the CrFC-κB motif in the promoter devoid of sequences upstream of −186. The mutation of the CrFC-κB site completely eliminated the reporter gene up-regulation after LPS treatment and overexpression of Dorsal in all the promoter constructs (Fig. 6, *A* and *B*), thus further corroborating the importance of this CrFC-κB site in regulating the *CrFC* gene in innate immune response. It also further proves that a potential limulus Dorsal/Rel transcription factor has a strong interaction with its cognate CrFC-κB sequence, which occurs in the absence of its upstream sequences. Furthermore, attenuation of the *CrFC* promoter activity because of mutation of the GAAAA motif adjacent to the CrFC-κB indicates the cooperation between these two motifs.

In EMSA using human NFκB protein (p50) and recombinant *Drosophila* Dorsal, the oligonucleotide probe containing both the CrFC-κB and its adjacent dorsal-like motif interacted more prominently with Dorsal than with human p50, suggesting that the transcription factor recognizing the CrFC-κB motif is more closely related to *Drosophila* Dorsal. The Prox3Dor probe contains 3 proximal dorsal-like binding motifs from −183 to −219: two overlapping sites at position −183 and one site at −209. Prox3Dor was also able to bind the recombinant dorsal transcription factor but not the human p50, whereas Dor348 with the dorsal-like motif at position −348 showed no binding to any Rel proteins. Removal of −395 to −186 (containing cis-acting sites in Prox3Dor and Dor348 probes) resulted in p186, which exhibited the highest promoter activity in response to LPS and during Dorsal overexpression. This strongly suggests that dorsal-like sites in the region from −183 to −219 (in Prox3Dor) have a repressive effect on the CrFC-κB site. However, neither Dor586 nor Dor654 was able to bind the Rel proteins. Thus, the reduction in promoter activity in p395 (devoid of −682 to −395) cannot be attributed to these two sites. Therefore, other positive elements residing in this region (from −682 to −395), together with the CrFC-κB, must be responsible for the highest basal promoter activity and the significant CAT gene up-regulation during LPS treatment and Dorsal overexpression in p682. Dor788, the only dorsal-like binding motif found upstream of −682, showed specific binding with recombinant dorsal-binding protein. This motif might also play a role in the repressive effect because pFL and p1017 (both

Dor788-DNA binding domain complex (see *lane 2*). *Lane 7* shows the binding reaction between Dor788 and p50. No DNA-protein complex was observed. *E*, EMSA of Dor348, Dor586, and Dor654 (Dorsal-like motifs at −348, 586 and −654, respectively, of the promoter) with Dorsal DNA binding domain, Rel, and p50. *Lanes 1–4* are Dor348 with GST negative control, DNA binding domain, Rel, and p50. *Lanes 5–8* and *9–12* show binding reactions of Dor586 and Dor654, respectively. No specific DNA-protein complexes were observed with these probes. *F*, CrFC-κB probe forms DNA-protein complex with the nuclear extracts from limulus amebocytes. *N*, *C*, and *F* stand for the nuclear extract, cytoplasmic proteins, and free probes as indicated from *lanes 1–3*, respectively. Note: primers designed for subcloning of Rel and DNA binding domains of Dorsal were: 5′-GCTGCGGCCGCGAGGGTGTGAGCTCAGG-3′ (antisense) for the Rel domain amplification; 5′-GCTGCGGCCGCCTACACAGCCGGCAGAT-3′ (antisense) for the DNA binding domain. For PCR cycles, both the antisense primers were paired with the same sense primer: 5′-GATCCCGG-GGAACCAGAACAATGGAGCCG-3′.

containing Dor788) showed no significant change in promoter activity during LPS stimulation and in the presence of Dorsal overexpression.

Therefore, we conclude: (i) that the proximal CrFC-κB motif in cooperation with the adjacent dorsal-like binding motif, constitutes the positive LPS-response unit, which is facilitated by the Rel transcription factor to become strongly responsive to Gram-negative bacterial challenge; (ii) that the dorsal-like binding motifs within −183 to −219 and at −788 may be involved in a repressive effect, acting competitively against the proximal LPS-response unit to counterbalance the LPS-induced overexpression of the *CrFC* gene during Gram-negative infection, which would otherwise be lethal to the organism; and (iii) the existence of potential repressive elements in the promoter region further upstream of −682 (in constructs pFL and p1017), which antagonizes the strong transactivation potential of the limulus Rel transcription factor directed on the positive element, CrFC-κB. The models for the transcriptional regulation of *CrFC* are illustrated in Fig. 8. Although the specific repressive/negative element(s) needs to be further verified, the modulation of the Rel/NFκB transcription factor by various repressors and corepressors has been well documented in both invertebrates and vertebrates (35, 54–57). Whereas the minimal promoter activity is commonly reported to be confined within a short sequence proximal to the transcription start site in many serine protease blood coagulants (39, 41, 42, 53), we report here, the repressive effect of the upstream sequence on the promoter that regulates the expression of a serine protease that is involved in both blood coagulation and innate immune response.

Accordingly, it is logical to envisage that during Gram-negative bacterial invasion or during overexpression of Dorsal, as soon as the pre-existing pool of Factor C serine protease zymogen is activated by LPS, its action must be tightly regulated to control and prevent the spread of hemostasis because of LPS challenge. At the transcriptional level, this tight control of Factor C expression is achieved as described above. These findings have led to our search for a limulus κB transcription factor. The prominent complex of the CrFC-κB oligonucleotide with proteins from the nuclear extract of the limulus amebocytes has formed the basis for a related study in our lab, which has led to the isolation and cloning of this novel Rel transcription factor from the limulus.[2]

*Cryptic Role of Limulus LPS-responsive Serine Protease Cascade?*—Multiple evidence points to other potential role(s) that may be played by the limulus LPS-responsive serine protease cascade. Recently, the cDNA of a Toll-like receptor (TLR) in the Japanese horseshoe crabs (*Tachypleus tridentatus*) has been cloned from a hemocyte cDNA library (58) and a *C. rotundicauda* TLR cDNA clone has also been isolated from a hepatopancreas cDNA library.[2] Although the functional homology of the limulus TLRs with the *Drosophila* Toll receptor remains to be determined, the structural homology between coagulogen and Spätzle has raised speculations that the processed coagulogen might be the candidate ligand for limulus TLR. In our recent study on the differential display of cDNA clones on *P. aeruginosa* infection, coagulogen cDNA was highly inducible and abundantly expressed upon infection.[2] Furthermore, sequence homology of other limulus coagulation factors with the *Drosophila* dorsoventral serine protease cascade has been consistently documented. In addition, the newly discovered serine protease, Persephone, in *Drosophila* antifungal response also shares sequence similarity with serine proteases from both cascades. Our finding that the Dorsal/NFκB transcription factors regulate Factor C expression parallels the report on a negative feedback regulation from Dorsal on Snake and Easter, which are upstream components of the serine protease cascade that controls the activity of Dorsal (59). Moreover, the unique domain structures of CrFC (30, 60) and the circulating form of Factor C also suggest that this protein must have other unidentified roles. Besides blood coagulation and removal of endotoxin, CrFC may be involved in cell signaling and protein-

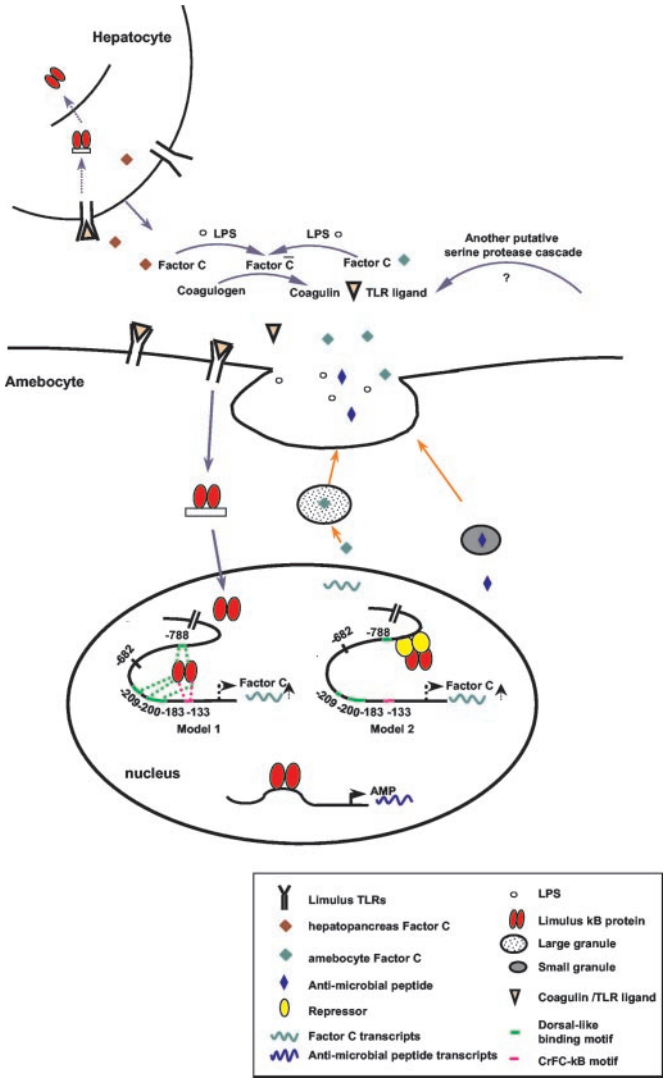


FIG. 8. **Proposed model for the signaling pathway in the limulus innate immune response and transcriptional regulation of Factor C by LPS.** Circulating Factor C derived from hepatocytes binds Gram-negative bacteria or LPS and triggers a further exocytosis of cellular Factor C from the large granule of amebocytes. The LPS-activated Factor C initiates the serine protease cascade to proteolytically process coagulin (or unidentified TLR ligand cleaved either by the blood coagulation cascade or a distinguishable serine protease cascade), which binds to the limulus Toll-like receptor on the amebocyte and/or hepatocyte membranes and elicits a signaling pathway in the immune response (similar to mammals and *Drosophila*), resulting in the up-regulation of Factor C gene through the limulus κB transcription factor. However, any potential dramatic up-regulation of Factor C expression in amebocytes is limited by *models 1* and *2*. In *model 1*, the dorsal-like binding motifs (two overlapping sites between −183 to −200, and the ones at −209 and −788) interplay and compete against the CrFC-κB site (−133) for the limulus κB transcription factor. In *model 2*, putative repressive element(s) with their cognate repressors upstream of −682 directly inhibits the interaction of the limulus κB transcription factor at the CrFC-κB site. We postulate that the same signaling pathway also regulates the expression of immune-related molecules, for instance, antimicrobial peptides as in *Drosophila*. The *symbols* used in this figure are keyed in the *lower right box*.

[2] X. W. Wang, L. H. Wang, B. Ho, and J. L. Ding, unpublished data.

protein interaction. Is the LPS-responsive blood coagulation cascade reutilized in innate immune response? Or does a separate serine protease cascade perform the "signaling" role as in *Drosophila*? We propose that the limulus blood coagulation or another uncharacterized serine protease cascade transduces an extracellular signal into the amebocytes and/or hepatopancreas during the innate immune response initiated by LPS. The bacterial infection signaled by the serine protease cascades ultimately results in the activation of the Dorsal/NF$\kappa$B transcription factors as illustrated in Fig. 8. Similar to its counterparts in *Drosophil*a, the putative limulus Dorsal/NF$\kappa$B transcription factors also transactivate other antimicrobial genes to defend against microbial infection.

In conclusion, the limulus hemolymph coagulation cascade is probably an ancient innate immune process with a combination of functions: recognition of microbial pathogens, blood coagulation to regulate hemostasis, and possibly, signal transduction for subsequent immune responses as illustrated in Fig. 8. Our finding on the transcriptional regulation of CrFC fortifies the idea that blood clotting and innate immune responses are closely linked processes pervasive from arthropods to humans. It opens an important avenue to the ultimate understanding of the molecular basis on the sophisticated and orchestrated regulation of serine proteases by the blood clotting and immune defense networks in both invertebrates and vertebrates.

*Acknowledgments*—We thank Professor Y. Engstrom for kindly providing the mbn-2 cells, Professor J.-M. Reichhart (Institute de Biologie Moléculaire et Cellulaire, France) for the pPACfl(dorsal) and pACH110 vectors, and Professor P. Matsudaira (Whitehead Institute, MIT) for limulus actin-3 probe. The initial genome walker libraries were obtained by C. Y. Ng. We thank B. Y. Teo for excellent technical advice.

REFERENCES

1. Rawlings, N. D., and Barrett, A. J. (1993) *Biochem. J.* **290,** 205–218
2. Morisato, D., and Anderson, K. V. (1995) *Annu. Rev. Genet.* **29,** 371–399
3. Belvin, M., and Anderson, K. V. (1996) *Annu. Rev. Cell. Dev. Biol.* **12,** 393–416
4. Petersen, U-M., Bjorklund, G., and Engstrom, Y. (1995) *EMBO J.* **14,** 3146–3158
5. Dushay, M., Asling, B., and Hultmark, D. (1996) *Proc. Natl. Acad. Sci. U. S. A.* **93,** 10343–10347
6. Baldwin, A. S., Jr. (1996) *Annu. Rev. Immunol.* **16,** 649–681
7. Ghosh, S., May, M. J., and Kopp, E. B. (1998) *Annu. Rev. Immunol.* **16,** 225–260
8. Lemaitre, B., Nicolas, E., Michaut, L., Reichhart, J-M., and Hoffmann, J. A. (1996) *Cell* **86,** 973–983
9. Hoffmann, J. A., Kafatos, F. C., Janeway, C. A., Jr., and Ezekowitz, R. A. B. (1999) *Science* **284,** 1313–1318
10. Gay, N. J., and Keith, J. F. (1992) *Biochim. Biophys. Acta* **1132,** 290–296
11. Ding, J. L., Navas, M. A., III, and Ho, B. (1993) *Biochim. Biophys. Acta* **1202,** 149–156
12. Muta, T., and Iwanaga, S. (1996) *Curr. Opin. Immunol.* **8,** 41–47
13. Bergner, A., Oganessyan, V., Muta, T., Iwanaga, S., Typke, D., Huber, R., and Bode, W. (1996) *EMBO J.* **15,** 6789–6797
14. Mizuguchi, K., Parker, J. S., Blundell, T. L., and Gay, N. J. (1998) *Trends Biochem. Sci.* **23,** 239–242
15. Ligoxygakis, P., Pelte, N., Hoffmann, J. A., and Reichhart, J-M. (2002) *Science* **297,** 114–116
16. Khush, R. S., and Lemaitre, B. (2000) *Trends Genet.* **16,** 442–449
17. Davie, E. W., Fujikawa, K., and Kisiel, W. (1991) *Biochemistry* **30,** 10363–10370
18. Hoffman, M., Monroe, D. M., and Roberts, H. R. (1996) *Haemostasis* **26,** Suppl. 1, 12–16
19. Hoffman, M., and Monroe, D. M. (2001) *Thromb. Haemostasis* **85,** 958–965
20. Cicala, C., and Cirino, G. (1998) *Life Sci.* **62,** 1817–1824
21. Opal, S. M. (2000) *Crit. Care Med.* **28,** S77–S80
22. Krem, M. M., and Cera, E. D. (2002) *Trends Biochem. Sci.* **27,** 67–74
23. Janeway, C. A., Jr. (1989) *Cold Spring Harbor Symp. Quant. Biol.* **54,** 1–13
24. Liu, C., Xu, Z., Gupta, D., and Dziarski, R. (2001) *J. Biol. Chem.* **276,** 34686–34694
25. Leulier, F., Parquet, C., Pili-Floury, S., Ryu, J. H., Caroff, M., Lee, W. J., Mengin-Lecreulx, D., and Lemaitre, B. (2003) *Nat. Immunol.* **4,** 478–484
26. Michel, T., Reichhart, J-M., Hoffmann, J. A., and Royet, J. (2001) *Nature* **414,** 756–759
27. Söderhäll, K., and Cerenius, L. (1998) *Curr. Opin. Immunol.* **10,** 23–28
28. Tan, N. S., Ho, B., and Ding, J. L. (2000) *FASEB J.* **14,** 859–870
29. Ding, J. L., Navas, M. A., III, and Ho, B. (1995) *Mol. Marine Biol. Biotechnol.* **4,** 90–103
30. Wang, J., Tan, N. S., Ho, B., and Ding, J. L. (2002) *J. Biol. Chem.* **277,** 36363–36372
31. Sambrook, J., Fritsch, E. F., and Maniatis, T. (1988) *Molecular Cloning: A Laboratory Manual*, 2nd Ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY
32. Way, M., Sanders, M., Chafel, Garcia, C., Sakai J., and Matsudaira, P. (1995) *J. Cell Biol.* **128,** 51–60
33. Thisse, C., Perrin-Schmitt, F., Stoetzel, F., and Thisse, B. (1991) *Cell* **65,** 1191–1201
34. Mason, P. J., Enver, T., Wilkinson, D., and Williams, J. G. (1993) in *Gene Transcription: A Practical Approach* (Rickwood, D., and Hames, B. D., eds) pp. 47–53, Oxford University Press, Oxford
35. Ip, Y. T., Kraut, R., Levine, M., and Rushlow, C. A. (1991) *Cell* **64,** 439–446
36. Dignam, J. D., Lebowitz, R. M., and Roeder, R. G. (1983) *Nucleic Acids Res.* **11,** 1475–1489
37. Kurachi, S., Furukawa, M., Salier, J. P., Wu, C. T., Wilson, E. J., French, F. S., and Kurachi, K. (1994) *Biochemistry* **33,** 1580–1591
38. Miao, C. H., Leytus, S. P., Chung, D. W., and Davie, E. W. (1992) *J. Biol. Chem.* **267,** 7395–7401
39. Huang, M. N., Hung, H. L., and Stanfield-Oakley, S. A. (1992) *J. Biol. Chem.* **267,** 15440–15446
40. Citarella, F., Misiti, S., Felici, A., Aiuti, A., La Porta, C., and Fantoni, A. (1993) *Biochim. Biophys. Acta* **172,** 197–199
41. Erdmann, D., and Heim, J. (1995) *J. Biol. Chem.* **270,** 22988–22996
42. Pollak, E. S., Hung, H. L., Godin, W., Overton, G. C., and High, K. A. (1996) *J. Biol. Chem.* **271,** 1738–1747
43. Chow, B. K., Ting, V., Tufaro, F., and MacGillivray, R. T. (1991) *J. Biol. Chem.* **266,** 18927–18933
44. Akira, S., and Kishimoto, T. (1997) *Adv. Immunol.* **65,** 1–46
45. Georgel, P., Kappler, C., Langley, E., Gross, I., Nicolas, E., Reichart, J-M., and Hoffmann, J. A. (1995) *Nucleic Acids Res.* **23,** 1140–1145
46. Gateff, E., Gissman, L. V., and Shrestha, R. (1980) in *Invertebrate Systems in Vitro* (Kurstak, E., Mara, K., and Dubenderfer, A., eds) pp. 517–533, Elsevier Science B.V., Amsterdam
47. Iwanaga, S. (2002) *Curr. Opin. Immunol.* **14,** 87–95
48. Mackman, N. (1995) *FASEB J.* **9,** 883–889
49. Mackman, N. (1996) *Haemostasis* **26,** Suppl. 1, 17–19
50. Moon, D. K., and Geczy, C. L. (1988) *J. Immunol.* **141,** 1536–1542
51. Hassan, J. H., Chelucci, C., Peschle, C., and Sorrentino, V. (1992) *Thromb. Haemostasis* **67,** 78–83
52. Citarella, F., Felici, A., Brouwer, M., Wagstaff, J., Fantoni, A., and Hack, C. E. (1997) *Blood* **90,** 1501–1507
53. Wilberding, J. A., and Castellino, F. J. (2000) *Thromb. Haemostasis* **84,** 1031–1038
54. Lee, S. K., Kim, J. H., Lee, Y. C., Cheong, J., and Lee, J. W. (2000) *J. Biol. Chem.* **275,** 12470–12474
55. Nissen, R. M., and Yamamoto, K. R. (2000) *Genes Dev.* **14,** 2314–2329
56. Hader, T., Wainwright, D., Shandala, T., Saint, R., Taubert, H., Bronner, G., and Jackle, H. (2000) *Curr. Biol.* **10,** 51–54
57. Tetsuka, T., Uranishi, H., Imai, H., Ono, T., Sonta, S., Takahashi, N., Asamitsu, K., and Okamoto, T. (2000) *J. Biol. Chem.* **275,** 4383–4390
58. Inamori, K., Koori, K., Mishima, C., Muta, T., and Kawabata, S. (2000) *J. Endotoxin Res.* **6,** 397–399
59. Misra, S., Hecht, P., Maeda, R., and Anderson, K. V. (1998) *Development* **125,** 1261–1267
60. Muta, T., Miyata, T., Misumi, Y., Tokunaga, F., Nakamura, T., Toh, Y., Ikehara, Y., and Iwanaga, S. (1991) *J. Biol. Chem.* **266,** 6554–6561