

Microbial Pattern Recognition in the *Euprymna scolopes* - *Vibrio fischeri* Symbiosis

By

Benjamin C. Krasity

A dissertation submitted in partial fulfillment

of the requirements for the degree of

Doctor of Philosophy

(Microbiology)

at the

University of Wisconsin-Madison

2015

Date of final oral examination: 5/5/2015

Approved by the following members of the Final Oral Committee:

Margaret McFall-Ngai, Professor, Medical Microbiology and Immunology

Joseph Dillard, Associate Professor, Medical Microbiology and Immunology

Anna Huttenlocher Professor, Medical Microbiology and Immunology and Pediatrics

Shigeki Miyamoto, Professor, Oncology

Edward Ruby, Professor, Medical Microbiology and Immunology

MICROBIAL PATTERN RECOGNITION IN THE *EUPRYMNA SCOLOPES* - *VIBRIO FISCHERI* SYMBIOSIS

Benjamin C. Krasity

Under the Supervision of:

Margaret J. Mcfall-Ngai

In the mutualistic relationship of the Hawaiian bobtail squid *Euprymna scolopes* and the bioluminescent marine bacterium *Vibrio fischeri*, numerous host developmental events are triggered by *V. fischeri*'s microbe-associated molecular pattern molecules, or MAMPs, such as lipopolysaccharide (LPS) and peptidoglycan. The mechanisms by which these effects are exerted are largely uncharacterized. Here are described new insights into biology of host and microbe, concerning MAMPs and their detection by the host. In mammals, lipopolysaccharide binding proteins (LBPs) are soluble proteins which bind LPS, transferring it to membrane-bound Toll-like receptor-4 by way of CD14 and MD-2. LBPs have previously been scarcely characterized in animal/bacterial mutualism. Structure and function of an *E. scolopes* LBP family protein, EsLBP1, were examined to provide evidence for binding symbiont LPS. Like human LBP, EsLBP1 bound *Francisella tularensis* LPS only poorly; however, failure to bind lipooligosaccharide-CD14 complexes, combined with a lack of reports of CD14 in *E. scolopes*, suggests that different intermediates in the transfer of LPS exist in *E. scolopes* and humans. *eslbp1* expression correlated with the onset of MAMP-driven developmental changes, induced by peptidoglycan products, showing evidence of synergy in MAMP signaling in *E. scolopes*. The structure of the *V. fischeri* LPS was further characterized, showing a short, unusual O-antigen of yersiniose, 8-epi-legionaminic acid, and *N*-acetylfucosamine attached to the core oligosaccharide. An O-antigen ligase mutant, *waaL*, lacking the O-antigen, was markedly

outcompeted by wild-type *V. fischeri* in colonization of *E. scolopes*. This phenotype was most obviously attributed to reduced motility, but a role for the O-antigen in conversation with the host immune system must also be considered. Potential cell surface receptors of MAMPs, Toll-like receptors, were also investigated in *E. scolopes*. Five previously unknown Toll-like receptors were sequenced and placed in the phylogenetic context of other lophotrochozoan Toll and Toll-like receptors, suggesting an ancient split between the *E. scolopes* proteins that may relate to functional differences. Toll-like receptor expression varied between tissues, and in one case was regulated by the symbiotic colonization of the light organ. Combined, these data provide new evidence of the immune system of *E. scolopes* adjusting to accommodate *V. fischeri*.

Acknowledgements

This work, itself an examination of certain deeply, perhaps murkily in-between regions of host and bacterial biology, rests on more pragmatic sorts of in-betweenness, as it would never have been possible if there were not those who supported and believed in me as a would-be physician and scientist, and if such people did not exist on both sides of the Mississippi, in Madison and in Coralville. The list in Madison includes my mentor, Margaret McFall-Ngai; the former and current Medical Scientist Training Program directors, Deane Mosher and Anna Huttenlocher; and the remaining members of my committee, Joe Dillard, Shigeki Miyamoto and Ned Ruby. Beyond these, I would like to thank the undergraduates who worked with me: Kirsten Guckes, Julian Cagnazzo and Jedediah Seltzer. I will cease counting in Madison there, as I must stop at some point, but I must also mention those who made me feel welcome in Iowa. In addition to Jerrold Weiss and Athmane Teghanemt, special gratitude is owed to the late, dearly missed Theresa Gioannini of the University of Iowa.

TABLE OF CONTENTS

Abstract	i
Acknowledgements	iii
Contents	iv
CHAPTER 1: Introduction and Outline	1
Preface	2
Part I: <i>Euprymna Scolopes</i> , <i>Vibrio Fischeri</i> , and Microbe-Associated Molecular Patterns	3
Part II: LBP/BPI Proteins and Their Relatives: Conservation over Evolution and Roles in Mutualism	7
Part III: Toll and Toll-Like Receptors across Evolution and in Mutualism	21
References	28
CHAPTER 2: Structural and Functional Features of a Developmentally Regulated Lipopolysaccharide-Binding Protein	39
Preface	40
Abstract	41
Importance	41
Introduction	42
Materials and Methods	45
Results	51
Discussion	64
Acknowledgements	68
References	69

CHAPTER 3: O-antigen and Core Carbohydrate of <i>Vibrio fischeri</i>	76
Lipopolysaccharide: Composition and Analysis of Their Role in <i>Euprymna scolopes</i> Light Organ Colonization	
Preface	77
Abstract	78
Introduction	78
Materials and Methods	80
Results	86
Discussion	123
Acknowledgements	130
References	130
CHAPTER 4: Novel Toll-like Receptors in the <i>E. scolopes</i> Light Organ	137
Preface	138
Abstract	139
Introduction	139
Materials and Methods	141
Results	147
Discussion	158
References	162
CHAPTER 5: Conclusions and Future Directions	166
Preface	167
References	171
APPENDIX A: NF-κB Pathway Activation in <i>E. scolopes</i>	173

Preface	174
Introduction	175
Materials and Methods	178
Results	184
Discussion	193
References	199
APPENDIX B: Genome Walking in <i>E. scolopes</i>	203
Preface	204
Introduction	205
Materials and Methods	206
Results	208
Discussion	212
References	214
APPENDIX C: <i>Euprymna scolopes</i> Acyloxyacyl Hydrolase	216
Preface	217
Introduction	218
Materials and Methods	218
Results and Discussion	219
References	224

CHAPTER 1

Introduction and Outline

PREFACE

Sections I and III, “*Euprymna scolopes*, *Vibrio fischeri*, and Microbe-Associated Molecular Patterns” and “Toll and Toll-like Receptors across Evolution and in Mutualism” are entirely the work of BCK.

Section II is published as: Krasity BC, Troll JV, Weiss JW, and McFall-Ngai MMN. “LBP/BPI proteins and their relatives: Conservation over evolution and roles in mutualism.” *Biochem Soc Trans.* 2011 Aug;39(4):1039-44, as a contribution to "Proteins with a BPI/LBP/Plunc-like domain: Revisiting the old and characterizing the new", 5-7 January 2011, New Business School, University of Nottingham, UK.

The complete EsLBP3 sequence, based on a portion obtained by JVT, was obtained by BCK. Sequences for EsLBP1 and EsLBP2, as well as a portion of EsLBP3, were obtained by JVT. BCK performed other analysis and experiments and wrote the paper. JVT, JPW and MMN provided ideas and assisted in revision of the paper. The figure numbering has been changed from the published version.

I: *EUPRYMNA SCOLOPES*, *VIBRIO FISCHERI*, AND MICROBE-ASSOCIATED MOLECULAR PATTERNS

The symbiosis between the Hawaiian bobtail squid *Euprymna scolopes* and the bioluminescent bacterium *Vibrio fischeri* is established as a model for beneficial host-microbial interactions (reviewed in (1, 2)). In this relationship, the host is provided with a form of camouflage called counterillumination from the light of the bacteria (3), which are able to grow to high density on nutrients available in the crypt spaces of the host light organ (4). The symbiosis is horizontally transmitted, with each new generation of squid acquiring its symbionts from the environment; the majority of the bacteria are expelled from the light organ each day. There is evidence of coevolution of host and symbiont (5).

It has long been known that the *E. scolopes* light organ, which must fulfill the functions of both initial recruitment of *V. fischeri* from the environment and management of the mature symbiosis, undergoes morphological changes in response to colonization (6). Investigation of the bacterial factors responsible for these changes has focused on microbe-associated molecular pattern (MAMP) molecules (also called PAMPs, for “pathogen-associated”, with unnecessary specificity), conserved components of broad classes of microorganism that provoke a response from the host innate immune system upon recognition by pattern recognition receptors (PRRs) (7, 8). Many MAMPs are surface components of microbes, including peptidoglycan, lipopolysaccharide (LPS), and flagellin, though others are internal materials in the cell the host does not itself produce, such as unmethylated CpG DNA (9). Of these, it is LPS and peptidoglycan, in particular, which have been found to be important in the *Euprymna-Vibrio* symbiosis. These molecules, between them, recapitulate the prominent apoptosis and regression that occur in the ciliated fields of the light organ (important in the initial recruitment of bacteria)

in the first days of development. LPS causes chromatin condensation and early stage apoptosis in the first 12-24 hours post-inoculation, but regression will not proceed unless peptidoglycan products are also applied (8, 10).

Other host phenotypes also depend on LPS and peptidoglycan. Peptidoglycan treatment results in secretion of mucus (11) and migration of hemocytes to the light organ epithelium (12). LPS promotes regrowth of the microvilli of epithelial lining of the crypt spaces (13), and the two MAMPs synergize to cause the turndown in host nitric oxide synthase production associated with colonization (14). Furthermore, a beneficial effect of these MAMPs on the host is not unique to *E. scolopes*. The LPS of beneficial bacteria is important for gut homeostasis and resistance to injury in mice (15), and renewal of gut epithelium in *Drosophila* depends on pathways responsive to the peptidoglycan of the microbiota (16, 17).

As LPS and peptidoglycan are MAMPs, and thus broadly conserved (though not identical) across categories of microbe, it is unsurprising that the efficacy of these effector molecules is often not limited to *V. fischeri*. *V. fischeri* LPS, added pharmacologically to seawater, is actually less potent at inducing apoptosis than the LPS of other species (10); why other, copious marine bacteria should not induce this phenotype is thought to relate to the location of antigen presentation, and is a subject considered in Chapter 2 of this thesis. The induction of mucus shedding is not specific to *V. fischeri* peptidoglycan; indeed, this phenotype can be induced by intact, Gram-positive bacteria that do not colonize the light organ (11).

Peptidoglycan is a polymer including of a sugar backbone, with alternating *N*-acetyl glucosamine and *N*-acetylmuramic acid residues, linked to the next row of sugars by short peptide chains. Peptidoglycan structure varies fundamentally between most Gram-positive and Gram-negative bacteria. The third amino acid position in Gram-negative organisms is generally

occupied by meso-diaminopimelic acid (Dap-type peptidoglycan), whereas Gram-positive organisms typically feature L-lysine instead (Lys-type peptidoglycan), with notable exceptions including *Bacillus* spp. and other Gram-positive rods with Dap-type peptidoglycan (18, 19). Because of the polymeric nature of peptidoglycan, a single organism can present a variety of products to the environment and a host, depending on where the bonds between sugar and amino acid residues are broken. *Vibrio fischeri* exports a disaccharide tetrapeptide Dap-type species identical to the tracheal cytotoxin (TCT) of *Bordetella pertussis* causative of whooping cough (8) (reviewed in (20)). Despite considerable turnover of the cell wall during growth (21), most Gram-negative bacteria do not export TCT, possessing enzymes capable of promoting the reuptake and recycling of the compound (22). *V. fischeri* is on a short list of TCT-exporters (23) including *Neisseria gonorrhoeae* (24) and *Bordetella pertussis* (25).

The response to peptidoglycan, and TCT in particular, in *E. scolopes* is believed to be mediated by peptidoglycan recognition proteins. At least five of these exist (26, 27), with different observed or predicted locations, including secreted, transmembrane or GPI-anchored, and nuclear. The two best studied are EsPGRP1 and EsPGRP2. EsPGRP1 has been detected in the nuclei of epithelial cells – a very unusual finding – and is degraded or dispersed over the course of apoptosis (28). EsPGRP2 is exported into the light organ crypt spaces during colonization and has amidase activity rendering it capable of degrading TCT and thus reducing the immunological potency of material exported by *V. fischeri* (29). It is suspected that one or more of these PGRPs signals through the NF- κ B system (27, 28), though this has not been directly demonstrated. *Drosophila* PGRPs can activate NF- κ B transcription factors through either the Toll or IMD pathways, but this type of signaling has not been observed in mammals (reviewed in (30)).

This dissertation will focus more heavily on LPS in the symbiosis than peptidoglycan. As a Gram-negative bacterium, *V. fischeri* contains LPS in its outer membrane. The host is certainly exposed to plentiful *V. fischeri* LPS, because in addition to whatever material the host senses on intact *V. fischeri* or dead cells, *V. fischeri* produces outer membrane vesicles (OMVs) (31). The rotation of the membrane-sheathed flagellum also promotes the release of LPS in vesicles, which may be different than cell-derived OMVs (32). While LPS, unlike peptidoglycan, does not form polymers, it is a large molecule with distinct components whose makeup may vary considerably between bacterial species, or even within a species. There are three major components of LPS: the hydrophobic lipid A, which attaches the molecule to the membrane with its acyl groups, the core oligosaccharide, and the outermost O-antigen, a sugar polymer (reviewed in (33)). Some organisms, including *Neisseria*, have smaller molecules consisting of lipid A and a core oligosaccharide, but no O-antigen, that are called lipooligosaccharide (34). Lipopolysaccharide molecules lacking the O-antigen, especially in a species that often does exhibit an O-antigen, may also be referred to as rough or semi-rough (33).

Lipid A is required for viability in almost all Gram-negative bacteria (35). Lipid A (“endotoxin”) of an invasive organism is well recognized by the host when disphosphorylated and hexaacetylated, but this configuration affords mucosal organisms some resistance against secreted antimicrobial peptides. It may thus be an optimal configuration for a peaceful relationship between many mucosal organisms and their hosts (36). *V. fischeri*, however, does not have a typical lipid A structure. Both mono- and diphosphorylated lipid A molecules have been observed, with between four and eight acyl chains. Molecules with five or more acyl chains were observed to contain a unique phosphoglycerol moiety as well (37). Exactly which of this family of molecules observed *in vitro* is most important in the symbiosis is unclear, though some clues

exist. Acylation mutants in the *htrB1*, *htrB2* and *msbB* genes in *V. fischeri* have been characterized. Only *htrB1*, which is located on the smaller, more variable *Vibrio* chromosome, had a symbiosis defect, albeit a small one (38). Sites have been proposed for these modifications, with HtrB1 and HtrB2 likely sharing an acylation site; two different groups were observed at this site, but it seems to always be acylated in wild-type *V. fischeri* (37). The potential genes responsible for tetra- to octaacylation variability have not been characterized. *Pseudomonas aeruginosa* varies its acylation, shifting from a penta- to hexa- or heptaacylated lipid A in the host environment (39, 40). It is possible that *V. fischeri*'s multitude of LPS structures are adaptations to the different environments of its life cycle.

The expression of an alkaline phosphatase in the interior regions of the symbiotic light organ indicates that the host promotes *V. fischeri* LPS processing into a less-than-fully phosphorylated state. It is possible that this modification renders *V. fischeri* LPS less reactive, and a parallel may be drawn to the TCT-degrading EsPGRP2 (29). The previously uncharacterized core oligosaccharide and O-antigen are the focus of Chapter 3 of this dissertation.

II. LBP/BPI PROTEINS AND THEIR RELATIVES: CONSERVATION OVER EVOLUTION AND ROLES IN MUTUALISM

Abstract: LBP [LPS (lipopolysaccharide)-binding protein] and BPI (bactericidal/permeability-increasing protein) are components of the immune system that have been principally studied in mammals for their involvement in defense against bacterial pathogens. These proteins share a basic architecture and residues involved in LPS binding. Putative orthologues, *i.e.*, proteins encoded by similar genes that diverged from a common

ancestor, have been found in a number of non-mammalian vertebrate species and several non-vertebrates. Similar to other aspects of immunity, such as the activity of Toll-like receptors and NOD (nucleotide-binding oligomerization domain) proteins, analysis of the conservation of LBPs and BPIs in the invertebrates promises to provide insight into features essential to the form and function of these molecules. This review considers state-of-the-art knowledge in the diversity of the LBP/BPI proteins across the eukaryotes and also considers their role in mutualistic symbioses. Recent studies of the LBPs and BPIs in an invertebrate model of beneficial associations, the Hawaiian bobtail squid *Euprymna scolopes*' alliance with the marine luminous bacterium *Vibrio fischeri*, are discussed as an example of the use of non-vertebrate models for the study of LBPs and BPIs.

Introduction

LBP [LPS (lipopolysaccharide)-binding protein] and BPI (bactericidal/permeability-increasing protein) are closely related proteins involved in innate immunity. LBP, which is produced largely by hepatocytes, is secreted into the bloodstream, where it binds LPS and catalyzes the extraction and transfer of individual LPS molecules to CD14, forming a monomeric LPS-CD14 complex that is a key intermediate in delivery of LPS to MD-2/TLR4 (Toll-like receptor 4) and TLR4-dependent cell activation. BPI, which is produced by neutrophils, has higher affinity for LPS and bacteria, is bactericidal and represses inflammation by preventing LBP from delivering LPS to CD14 (41). Much of our knowledge of the function of LBP and BPI relates to their roles in host response to acute pathogenesis involving Gram-negative bacteria or LPS in the bloodstream. LBP is also believed to play a role in the handling of LPS that has been absorbed across the intestinal barrier (a more common process with high-fat diets), helping to

shuttle LPS to lipoproteins and chylomicrons, modulating monocyte/macrophage activation and pro-inflammatory cytokine secretion (42).

Members of this protein family and their relatives also function in other sites that interact with bacteria, most notably along the mucosal surfaces. For example, BPI is not only found in neutrophils and their secretions; it is also produced in human and mouse epithelia, including intestinal epithelia. As in the blood, it has bactericidal effects and blocks endotoxin signaling (43, 44). Pro-inflammatory mediators have little effect on its expression, but an anti-inflammatory eicosanoid of the lipoxin pathway has been shown to increase epithelial BPI expression (44). Because diverse assemblages of bacteria promote health of the mucosa, it is likely that the activities of the LBPs, BPIs and their relatives at these sites function not only in defense but also to modulate responses to the essential microbial partners. Although these immune proteins have not been studied in this capacity, their principal ligand, LPS, has. Specifically, host recognition of the normal microbiota in the gut is important for resistance to epithelial injury and critical for gut homeostasis. Oral administration of LPS can mimic the effects of intact bacteria (15). The circumstantial evidence taken together suggests that members of LPB/BPI protein family participate in control of the normal microbiota.

LBP and BPI have been most extensively studied in mammals, but myriad examples of related proteins occur throughout the animal kingdom and even in other eukaryotes. Several invertebrate species that have LBP and BPI orthologues provide simpler, more tractable models of bacterial interaction with animal epithelia. Unlike the mammalian mucosal surfaces, which typically harbour hundreds to thousands of bacterial phylotypes, these sites in the invertebrates often support partnerships of low diversity, with single to a few microbial phylotypes (45).

Several experimentally tractable invertebrate systems are being exploited as models of animal-bacterial interactions (46) and promise to shed light on the role of LBP and BPI in both pathogenic and non-pathogenic associations. In addition to offering relative simplicity, the study of invertebrate systems reveals features that are evolutionarily conserved across the animal kingdom and, as such, can provide insight into the essential functional features of proteins, such as LBP and BPI. Analogous contributions to the study of toll-like receptors (TLRs) in humans were the result of the discovery and characterization of these proteins in the fruit fly *Drosophila melanogaster*. Further, similarities between regulation of the transcription pathway in *Drosophila* development and NF- κ B activation in the mammalian immune system prompted investigations that revealed the immune role of Toll in *Drosophila* and spurred advances in the study of mammalian immune systems (47).

This review will first discuss the relevant features of mammalian LBP and BPI to set the stage for comparisons across the animal kingdom. Then, we will introduce examples of similar proteins in the non-mammalian vertebrates, the invertebrates, and in other eukaryotes. Finally, we will discuss applications of several of these as models, with special emphasis on the squid-vibrio system, for research on the basic nature of the biochemistry and physiology of members of the LBP and BPI protein family.

Characteristics of LBP and BPI in mammals

LBP and BPI have a characteristic, conserved two-domain “boomerang” structure, with an N-terminal domain and a C-terminal domain that share little sequence identity, but are very similar in overall architecture. The N-terminal domain carries out binding of LPS, but the precise LPS-binding site is still a matter of some conjecture. LBP and BPI are part of a wider family of lipid-binding proteins that includes members whose functions are not directly related to bacterial

pathogenesis, such as CETP (cholesterol ester transfer protein) and PLTP (phospholipid transfer protein). While these proteins are not the focus of this discussion, it is worth mentioning that they share this basic architecture (48). Certain PLUNCs (palate, lung, and nasal epithelium clones) have only a single BPI domain (49).

More basic residues occur in the N-terminus of human BPI than in LBP, a pattern common to the other mammalian LBP/BPI proteins; this feature is believed to promote improved binding to bacterial membranes (41, 50). However, several basic residues are conserved between LBP and BPI of humans, cattle, mice, rats, and rabbits. These occur at the positions corresponding to human LBP residues R42, R48, K92, K95, and K99 (50). R94 has also been implicated as important for LPS binding (51), though it is not universally conserved. Based on the crystal structure of BPI (52, 53), these residues are near each other, close to the tip of the protein's N-terminal domain (50). Four of these residues (all but K95) form a structural motif for binding LPS that is also present in proteins such as *Escherichia coli* FhuA, lactoferrin, lysozyme, and LALF (*Limulus* anti-bacterial and anti-LPS factor) (54), which are proteins that also interact with the surfaces of Gram-negative bacteria (55-57).

Data from mutagenesis experiments generally support the importance of the conserved residues. The replacement of three positively charged human LBP residues, K92, R94 and K95, with alanine dramatically reduces the protein's LPS binding capability (51, 58). The positive charge corresponding to LBP R94 is common, but not universally conserved among mammalian LBP/BPI proteins: in the aligned sequence of human BPI, a glutamine residue fills this position instead (50). These residues are not necessarily present in more distant members of this protein family, such as human CETP and PLTP, even where the basic architecture of BPI is believed to be retained (48).

Human BPI contains two apolar lipid-binding pockets. In the crystal structure, these contain phosphatidylcholine, but their role in interactions with LPS is unclear (53, 59). Unlike CD14 and MD-2, complexes of BPI (or LBP) with individual LPS monomers have not been described. Existing evidence strongly suggests that the primary interaction of BPI and LBP with LPS is with interfaces containing large numbers of LPS molecules packed closely together, such as aggregates of purified LPS, outer membranes of Gram-negative bacteria or shed outer membrane vesicles (41). These interactions appear to be driven by electrostatic interactions between multiple anionic groups clustered within the inner core/lipid A region of LPS and clusters of cationic residues concentrated at the tip of the N-terminal domain. The higher concentration of net basicity of BPI against LBP in this region correlates with the higher affinity of BPI (compared with LBP) for these LPS-rich interfaces (41, 50-53).

Analysis of amino acid sequences of other proteins in the LBP/BPI family suggests that they contain these pockets as well (50). The LBP/BPI family has a conserved disulfide bond, which is present between residues C132 and C175 of the N-terminal domain in human BPI (50). This disulfide bond and associated residues are also present in related proteins, such as CETP and PLTP (52, 60).

Beyond mammalian LBP/BPI

In recent years, the study of LBP and BPI has expanded beyond mammals. This family of proteins appears to be ancient. Proteins with BPI-like domains occur even outside the metazoans, such as in *Monosiga brevicollis*, which belongs to a group of marine choanoflagellate protists considered ancestral to the metazoans (61, 62). Phylostratigraphic analysis, a method in which all available sequence information for the biological world is considered in the construction of phylogenetic relationships, suggests that the CETP family emerged before the last common

ancestor of today's eukaryotes, although this analysis does not make mention of the point at which LBP and BPI emerged (63).

Apart from mammals, members of the LBP/BPI family have been reported in other vertebrates, including fish such as the rainbow trout *Onchorhynchus mykiss* (64) and Atlantic cod *Gadus morhua* (60), and birds, such as the chicken, *Gallus gallus*, which has a BPI but not LBP (65, 66). They also exist in various invertebrates. *Caenorhabditis elegans* has multiple proteins with high sequence similarity to LBP (65), and the freshwater snail *Biomphalaria glabrata* has several variants as well (67, 68). These few examples serve to demonstrate that LBPs and BPIs are widespread among the non-mammalian animal groups.

A few themes emerge upon analysis of the structure of these non-mammalian proteins. The BPI disulfide bond is ubiquitous, being present in BPI family members as distant as the tunicate *Ciona intestinalis* (60) and, a simple ClustalW alignment suggests, as far away as the protist *M. brevicollis*. Many non-mammalian LBP/BPI proteins are predicted to have the basic “boomerang” two-domain fold of human LBP and BPI (65). It is generally believed that the ancestor of LBP and BPI was a single-domain protein whose gene was duplicated (65), though the two-domain structure is common. Cases of LBP/BPI proteins with one domain have been reported, such as one of three BPI-family proteins in the sponge *Amphimedon queenslandica* (69). However, in this case it is believed to have arisen from a neighboring, two-domain LBP/BPI gene.

Some animals have abandoned LBP/BPI altogether. For example, *D. melanogaster* does not have an LBP/BPI and uses PGRPs (peptidoglycan recognition proteins) to detect Gram-negative bacteria (62). *D. melanogaster* also has a GGBP (Gram-negative bacterial-binding protein) capable of binding lipopolysaccharide and β -1,3-glucan (70). GGBP and the related

LGBP (lipopolysaccharide- and β -1,3-glucan-binding protein) are present in numerous arthropods and molluscs, e.g., the mosquito *Anopheles gambiae* (71) and freshwater crayfish *Pacifastacus leniusculus* (72). They are similar to CD14 and resemble defective β -1,3 or β -1,3-1,4 glucanases, possessing a functioning β -1,3 glucan binding site, but missing, in at least some cases, two glutamate residues believed to be important for catalytic activity (71, 73).

In animals that do have an LBP/BPI-like protein, there may be divergence from characteristics conserved in these proteins in mammals. Some features of mammalian LBP and BPI are better conserved than others. The Atlantic cod *G. morhua* has positively charged residues corresponding to positions 42, 48, 92, and 99 of human BPI or LBP, but not position 95; *i.e.*, it retains most sites believed to be important to LBP/BPI function (60). More distantly, a protein of this family in the oyster *Crassostrea gigas* lacks basic residues corresponding to R42 and R48 in human LBP, although it does share basic residues by alignment with LBP and BPI in other positions, including the positions corresponding to K92, K95 and K99 (74). In contrast, a BPI-like protein obtained from haemocytes of the snail *B. glabrata* (67, 68), when subjected to a ClustalW alignment by the authors of this review, did not obviously retain any of the positive charges corresponding to human LBP residues R42, R48, K92, K95, and K99, although in several cases positively charged residues were close and may be functionally relevant in LPS binding. These invertebrate proteins also have small insertions or deletions relative to human LBP and BPI, and their structures are not available, and so it is difficult to know exactly how these differences affect binding of LPS and other factors. Both of the invertebrate proteins mentioned in this paragraph appear to retain the conserved disulfide bond. Nevertheless, this variation across the animals does raise questions about how these proteins function, and what elements are required for LBP or BPI activity.

Because of differences from canonical residues in the non-mammalian proteins, and because mammalian LBP and BPI are believed to have arisen from a gene duplication after the radiation of the mammals (64), one might plausibly question whether LBP and BPI are meaningful as distinct categories outside of mammals. Some insight may be obtained from considering other aspects of the proteins' chemistry and activity. Human BPI has a higher pI (roughly 9.4) than human LBP (roughly 6.3) (75, 76), and it is generally assumed that invertebrate proteins will have a similar pattern. The most effective approach to determine whether isoelectric point correlates with antibacterial activity typical of BPI-like proteins, and thereby to distinguish candidate BPI from an LBP-like protein, is experimental analysis. The protein of the cod *G. morhua* mentioned above, for example, is up-regulated in the blood and peritoneum of the fish after intraperitoneal injection of bacteria. Its expression pattern is more similar to mammalian BPI than LBP, though to the best of our knowledge it has not been determined whether the protein has bactericidal activity like BPI (77); the predicted pI of this protein based on information provided by Solstad *et al.* and computed using ExPASy ProtParam (76), is roughly 10, supporting its identity as a BPI. The *C. gigas* protein binds LPS, is bactericidal and has a predicted isoelectric point of 9.3, making it more functionally similar to human BPI than LBP. This protein is constitutively expressed in epithelial tissues and is also up-regulated in haemocytes in response to bacterial challenge from non-pathogenic marine organisms (74).

We are not aware of a comparably well characterized non-mammalian protein that functions more similarly to human LBP than BPI, although there are related invertebrate proteins with a charged residue profile more similar to human LBP than human BPI, such as those in the sponge *A. queenslandica* (69). It is possible that comparing the LPS-binding functionality of

LBP/BPI variants, between and within species, will contribute to improved knowledge of the molecular activities of these proteins.

Toward models of LBP/BPI function in mutualism and development

Several of the animals mentioned above offer the opportunity to study LBP and BPI function in experimentally tractable systems. For example, researchers have begun using molluscs to study LBP and BPI in development. In the snail *B. glabrata*, egg mass fluid contains significant quantities of LBP/BPI, suggesting its utilization in parental immune protection of offspring (68). Notably, this protein's sequence differed from a previously characterized LBP/BPI from haemocytes of the same species, suggesting that different isoforms serve alternative functions in *B. glabrata* (67, 68). In the oyster *C. gigas*, larvae develop as free-swimming forms in the plankton, exposed to $\sim 10^6$ bacteria/ml of seawater (78). BPI transcript was detected throughout larval development of this species, increasing markedly around the time of the differentiation of epithelia. Experiments in which these larvae were challenged with both Gram-positive and Gram-negative non-pathogenic bacteria, including two *Vibrio* species and *Micrococcus luteus*, demonstrated that transcription of this BPI increases in the larval stage in a bacterial-dose dependent manner (79). These observations suggest a role for the BPI in immune defense during the developmental process, a time when the larvae are highly vulnerable to bacterial settlement.

Another system ripe for study of LBP and BPI is the Hawaiian bobtail squid *Euprymna scolopes* and its partner, the bioluminescent bacterium *Vibrio fischeri*. These organisms form a binary (one animal, one bacterial species), easily evaluated and easily manipulated symbiosis, in which bacteria colonize the squid's light organ and produce light. As such, these organisms

provide an excellent model system for a number of research questions, among them the role of LPS in chronic, beneficial colonization of an epithelial surface by Gram-negative bacteria (80).

During the process of bacterial colonization of the light organ, the MAMPs (microbe-associated molecular patterns) of *V. fischeri* play a central role in symbiont-host communication. Most notably, these molecules induce the transformation from a host-organ morphology that promotes symbiont colonization to one that facilitates the organ's mature function in bioluminescence production. The most conspicuous feature of this process is the loss of a superficial ciliated field that facilitates harvesting of the symbionts from the seawater. This developmental program involves a series of cellular and biochemical events, including haemocyte trafficking, apoptosis and attenuation of the levels of nitric oxide synthase and its product, nitric oxide. *V. fischeri* MAMPs work in synergy to trigger all of these events (8, 14). Specifically, lipid A induces early-stage apoptosis of *E. scolopes* epithelial cells, as the light organ adapts to the presence of the symbiont (10), and LPS, working in concert with the PGN (peptidoglycan) monomer, triggers the completion of the apoptotic process. This pattern of LPS-induced apoptosis is not unique to the *Euprymna/Vibrio* system, as the ability of LPS to induce host cell apoptosis has been investigated in other species (81, 82), generally with a focus on pathogenesis rather than mutualistic association.

Many of the details of this morphogenetic process in the squid-*Vibrio* symbiosis, however, are still incompletely understood. Among these, given the apparent importance of LPS in the developmental process, are how *E. scolopes* detects the presence of bacterial LPS in the light organ and how LPS functions in synergy with the PGN monomer. As it turns out, *E. scolopes* possesses at least three light-organ proteins in the LBP/BPI family, which have recently been sequenced (Fig. 1-1). The expression of one of these proteins, EsLBP1, has been evaluated

Protein	Length (AAs)	Mass (kDa)	pI	Predicted Domain Structure	Predicted Cellular Localization
EsLBP1	487	54.4	6.86		Cytosolic and/or nuclear
EsLBP2	477	53.8	9.63		Secreted
EsLBP3	531	60	9.41		Secreted

FIG 1-1 Characteristics of the predicted *E. scolopes* LBP proteins. The derived amino acid sequence of each of the EsLBP cDNAs was analyzed for biochemical parameters and protein family domains using ExPASy ProtParam (76) and the SMART algorithm (83). AAs, amino acids. NCBI accession numbers for sequences: JF514880, JF514881, JF514882.

in the context of establishment of mutualism. Eighteen hours into colonization, at which point apoptosis and epithelial regression are well underway, *eslbp1* mRNA is up-regulated roughly 9-fold over comparable, symbiont-free control animals. EsLBP1 protein itself has been detected in the animal's bacteria-containing crypt spaces at the same time point (Fig. 1-2) (84).

The patterns of occurrence of three distinct EsLBPs in the light organ may enable the tissues to use LPS to signal the variety of processes in which it is implicated. This type of strategy has already been described in this system. Specifically, previous studies of the symbiosis have characterized some of the EsPGRPs in the light organ; genes encoding four members of this family are expressed in the hatchling organ. In-depth analyses of two of the proteins, EsPGRP1 and 2, have demonstrated that a mechanism by which the animal can respond to PGN over the trajectory of early development is to deploy the isoforms at different times and in different locations in the organ (28, 29). The biochemical properties of both the EsLBPs and the EsPGRPs, along with their presence and relative importance in haemocytes, the light organ epithelium and extracellular crypts, are at present being evaluated. The results of such studies promise to shed light on how animal epithelia interact with the LPS and PGN as individual MAMPs, as well as how these MAMPs synergize to trigger host responses.

Conclusion

In light of the fact that all animal body plans arose in the Cambrian, some 520-540 million years ago (85), in the context of the bacteria-rich environment of the oceans, it is not surprising that the animals have developed mechanisms to respond to the surface molecules of members of the bacterioplankton. These responses may involve exploiting bacterial molecules for defensive purposes or as signals for normal development and homeostasis. Analyses of the current genomic databases are revealing that members of the LBP/BPI family of proteins

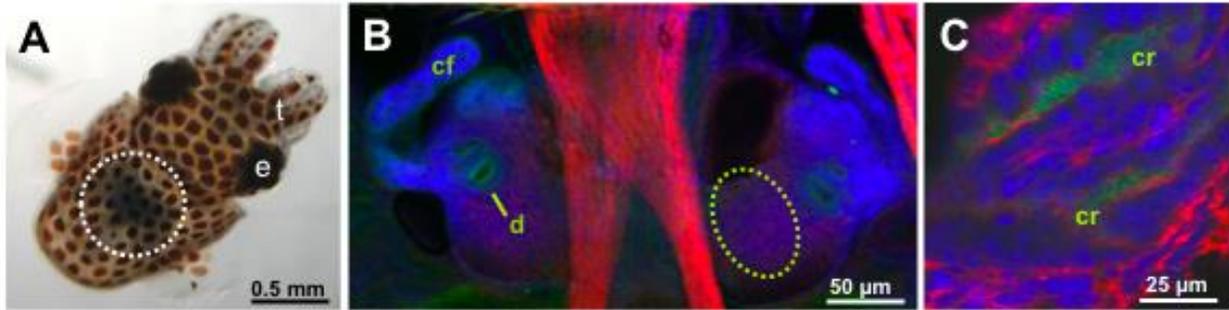


FIG 1-2 **Immunocytochemical localization of EsLBP1 in the light organ of juvenile *E. scolopes*.** (A) The position of the organ in the whole, live juvenile squid. The organ can be seen through the translucent dorsal surface of the animal as a dark region in the center of the body (white dashed circle). (B) Confocal microscopy image of the ventral surface of the juvenile light organ. Each lateral lobe of the organ bears a complex ciliated field (cf), which promotes harvesting of *V. fischeri* from the environment, and ducts (d), passageways through which the symbiont cells enter host tissues. Anti-EsLBP1 antibodies (green) label regions along the apical surfaces of the epithelia of the cells of the ciliated fields and the pores. The dashed oval region is where deeper images were taken in (C). (C) Confocal image of the deep crypt (cr) region of the light organ. Anti-EsLBP1 antibodies (green) label the deep crypt spaces where the symbiont cells reside following colonization of the organ. In (B, C), nuclei, blue (TOTO-3); actin, red (rhodamine phalloidin). e, eye; t, tentacles. For confocal microscopy methods, see (84).

represent an ancient means of recognizing MAMPs. Proteins similar to mammalian LBP and BPI exist in diverse organisms, with some residues strictly conserved and others forming common motifs. Opportunities exist to examine these proteins and determine core features of LBP and BPI. Additionally, these discoveries are occurring as we become increasingly aware of the full spectrum of symbiosis, from pathogenesis to mutualism.

Acknowledgements

We wish to thank Natacha Kremer, Mallory Agard and Ned Ruby for assistance and advice. This work was supported by grants from the National Institutes of Health (R01 AI50661) and National Science Foundation (IOS 0817232).

III. TOLL AND TOLL-LIKE RECEPTORS ACROSS EVOLUTION AND IN MUTUALISM

Toll and Toll-like Receptors across Evolution and in Mutualism

Toll-like receptors (TLRs) are an evolutionarily ancient family of proteins present in most animal phyla. Originally discovered (as Toll) serving to guide *Drosophila* morphogenesis (86), they have since become well known for serving crucial roles in vertebrate innate immune signaling (9), while evidence for a developmental role in numerous organisms is also expanding, as discussed below. Toll-like receptors are commonly referred to as PRRs, a straightforward characterization in vertebrates. The role of Tolls in *Drosophila* immunity is believed to be several steps removed from initial recognition of the MAMP (87), however, and beyond these two branches of the metazoans, there are few well documented instances of TLRs directly binding MAMPs. Even within *Drosophila* and vertebrates, important questions regarding TLR function remain. Here, an overview of this proteins over a broad phylogenic range is presented,

beginning with general questions of structure and function, considering the function of Toll-like receptors in less-studied groups such as mollusks and Lophotrochozoa more generally, and concluding with thoughts on the role of Toll-like receptors in beneficial animal-bacterial associations.

Toll-like receptors have well conserved structural features, including an extracellular region featuring leucine-rich repeats (LRRs), a transmembrane domain, and an intracellular Toll/Interleukin-1 Receptor (TIR) domain responsible for conveying signal to downstream, intracellular elements. Generally, ligand-bound TLRs dimerize and interact in the cytoplasm with another TIR domain, such as that of Myd88 (88). Several of these TIR-containing adapters are involved in NF- κ B signaling (TLRs activate NF- κ B transcription in both vertebrates and in *Drosophila*), and it is common, but not quite universal, for TLRs and NF- κ B proteins to occur together (89). Neither LRRs nor TIR domains are unique to TLRs; indeed, R proteins in plants contain both of them, though they are cytoplasmically located and otherwise different from TLRs (90). TLRs are apparently specific to eumetazoans (including some of the Radiata), though very similar proteins have been detected in Porifera; interestingly, a sponge TLR-like protein provides some of the limited evidence for direct recognition of bacterial ligands outside of vertebrate TLRs, though it lacks clearly identifiable LRRs (91).

LRRs, as suggested by their name, occur multiple times in a given protein, with each motif consisting of 20-30 amino acids and including both a variable region and a conserved region with a characteristic pattern of hydrophobic (not necessarily leucine) and cysteine, serine, threonine or asparagine residues. The LRR-containing extracellular portions of TLRs bend into an arc or horseshoe shape (92). Leucine-rich repeats in other proteins are very often involved in protein/protein interactions (92, 93); protein/protein interactions involving the LRRs are crucial

for the function of *Drosophila* Toll, but less prominent in the MAMP-binding action of vertebrate TLRs (89).

Vertebrate Toll-like receptors fall into six major families with different subcellular localization and ligand specificities (ligands may be endogenous or microbe-derived): TLR1, 3, 4, 7, and 11. There is further branching within some of these families; for example, TLR6 and 10 cluster very tightly with TLR1, while TLR2 and 14 are in the same family but somewhat more divergent. There is, additionally, a handful of outlier TLRs. Humans have a total of 10 TLRs, and other vertebrates have similar numbers (94). One characteristic feature of vertebrate Toll-like receptors is the presence of a single cysteine cluster at the C-terminal end of the leucine-rich repeats (CF) motif. Invertebrate TLRs may share this organization, as in the case of *Drosophila* Toll 9, but commonly have multiple cysteine clusters, both CF and NF motifs (95). Vertebrate TLRs have an important role as PRRs, with the preferred ligand varying by TLR family. An example is TLR4; TLR4-knockout mice are hyporesponsive to LPS (96), and polymorphisms in TLR4 can affect the outcome of Gram-negative infection (97). Vertebrate TLRs can directly bind MAMP ligands, but may require the involvement of other protein factors for performing this role, as is true in the case of TLR4, which requires extracellular MD-2 to bind LPS (98).

It is possible that in some cases, vertebrate TLRs convey signal indicating presence of a microbe without actually binding a MAMP, but instead respond to a host factor released in response to infection. Human nerve growth factor β (NGF β) and zebrafish tropomyosin-related kinase receptor A (TRKA) are cysteine knot proteins similar to *Drosophila* Spätzle (discussed below). NGF β is released in response to *Staphylococcus aureus* exposure, and mutations are associated with susceptibility to infection by that organism (99). No specific TLR was identified as responding to NGF β , but the effect of zebrafish TRKA was dependent on the TLR adapters

Myd88 and TRIF. A vertebrate TLR binding NGF β or TRKA as *Drosophila* Toll binds Spätzle, and responding similarly, would suggest commonalities of function in ancestral TLRs.

A role for vertebrate TLRs in development is less well attested than an immune role (89), but evidence is beginning to take shape that such a function does exist, at least in some limited contexts. Multiple TLRs are expressed in the brain, including the developing brain (100), with different studies reporting dynamic expression for TLR7, TLR8, and TLR9 over development (101, 102) and implicating TLR3 and TLR8 in axon development (102, 103). Deletion of TLR2 in mice induces brain morphological and behavioral abnormalities reminiscent of schizophrenia (104). TLR9 deletion also affects sensory and motor function as well (105).

In other deuterostomes apart from the vertebrates, information on TLRs is more limited. These TLRs do not appear to fall neatly into the families of vertebrate TLRs (94, 106, 107). However, it is notable that two TLRs investigated in the tunicate *Ciona intestinalis* (tunicates being perhaps the closest relatives of the vertebrates) have been implicated in directly binding MAMP ligands in the manner of vertebrate TLRs, as measured by NF- κ B-driven transcription in a reporter cell line transfected with ciTLRs and treated with MAMPs. Their ligand affinity may be much less specialized than vertebrate TLRs, as both were reported to interact with Poly I:C (dsRNA analogue), zymosan (fungal cell wall component), flagellin, and heat-killed *Legionella pneumophila* (108). Elsewhere in the deuterostomes, massive expansions of TLRs have occurred, with 72 TLRs detected in *Branchiostoma floridae* (109) and 222 in *Strongylocentrotus purpuratus* (110), many of them extremely similar to each other and implicated more in immunity than in embryogenesis by expression pattern (111). Non-vertebrate deuterostome TLRs may have one or more cysteine clusters associated with their LRRs.

The field of Toll-like receptor studies began with observations in Ecdysozoa, more specifically, of the involvement of *Drosophila* Toll in dorsoventral patterning in development (86). *Drosophila* Toll was found to induce this phenotype through response to a cleaved, activated form of the endogenous ligand Spätzle (112, 113). Spätzle also allows Toll to influence the *Drosophila* immune response; activated Toll in the fat body leads to expression of antimicrobial peptides, through a different NF- κ B family transcription factor, Dif (87). All but one of the nine *Drosophila* Tolls have multiple cysteine clusters in their LRRs (95). Tolls and Spätzle are present in other arthropods as well, and are up-regulated in response to bacterial challenge in the shrimp *Litopenaeus vannamei* (114, 115). Elsewhere in Ecdysozoa, the nematode *Caenorhabditis elegans* has one TLR, TOL-1 (116). *C. elegans* is an exception to the general rule that TLRs and NF- κ B co-occur (62). TOL-1 is generally regarded as primarily development-related, as its deletion is lethal to *C. elegans* embryos (116). It has been implicated in antimicrobial peptide release and defense against *Salmonella enterica* and *Serratia marcescens* (117, 118), but its function in host defense seems narrower than that of TIR-1, the only obvious potential adapter for TOL-1 (119).

In Lophotrochozoa, TLRs have been studied comparatively little. Platyhelminthes, notably, have not been found to contain TLRs (106). The NCBI Protein database lists TLR sequences (some predicted) for the mollusk genera *Aplysia*, *Azumapekten*, *Biomphalaria*, *Crassostrea*, *Cyclina*, *Euprymna*, *Haliotis*, *Lottia* and *Mytilus*, as well as the annelid genera *Capitella*, *Eisenia*, *Helobdella*, and *Hirudo*. There are additional reports of TLRs in lophotrochozoans outside of the mollusks and annelids (120).

Probably the best-characterized lophotrochozoan TLRs are four in *Crassostrea gigas* (121). These four CgTLRs are apparently present at both the plasma membrane and endosomes

and are implicated in hemocyte degranulation. They are capable of activating an NF- κ B reporter when transfected into a reporter cell line, but the reporter showed no further upregulation upon treatment with various MAMPs, so their ligands are unknown. These proteins are a small fraction of the total in *Crassostrea gigas*, which was recently reported to possess a remarkable 83 TLRs, most, but not all, of which have multiple cysteine clusters. Many of these are likely the product of recent duplications (122). *Capitella capitata* is an example of another lophotrochozoan with a remarkable number of TLR homologs, at 105 (123).

To the best of our knowledge, clear evidence of neither lophotrochozoan TLR-MAMP interaction nor a lophotrochozoan equivalent of Spätzle has been reported. Despite a paucity of information on ligand specificity, there is evidence that lophotrochozoan TLRs are involved in immunity. As mentioned above, *Crassostrea gigas* TLRs appear to be involved in the NF- κ B system (121), and a *Hirudo medicinalis* TLR also appears to be involved in cytokine regulation (124). TLR expression has been found to be regulated by MAMPs or pathogen challenge in numerous lophotrochozoans, including *Azumapecten farreri*, *Crassostrea gigas*, *Hirudo medicinalis* and *Mya arenaria* (121, 122, 124-127).

In summary, Tolls and Toll-like receptors occur throughout most phyla of the eumetazoans and are commonly associated with NF- κ B machinery. Originally, ecdysozoan TLR activity was associated with developmental phenotypes, responding to the endogenous ligand Spätzle; vertebrate TLRs were associated with direct association with MAMPs and resultant immune activation. Since then, more similarities in function have been realized: both immune and developmental functions are under investigation in diverse phyla.

Like LPS-binding proteins, vertebrate Toll-like receptors have roles in homeostasis and the response to beneficial bacteria, in addition to their well characterized role in defense against

pathogens (and their role in development in response to endogenous ligands, as noted above). Bacteria in the mammalian gut are recognized by Toll-like receptors, and the effect of LPS in protecting the mouse gut from injury is believed to depend on Toll-like receptors (15). There are several lines of evidence for this supposition, including the severe vulnerability of Myd88 (TLR adapter)-deficient mice to colonic injury and the possibility of rescuing the effect of depleted microbiota with LPS alone in wild-type, but not TLR4-deficient, mice. As signaling through some Toll-like receptors can have an ultimately pro- or anti-inflammatory impact, depending on the cytokines eventually produced, (128, 129), bacteria (whether beneficial or pathogenic) may exploit TLR signaling to reduce a host response to their presence. Anti-inflammatory TLR2 signaling can be seen in response to the pathogen *Staphylococcus aureus* (130) or the generally non-pathogenic *Bacteroides fragilis*. A specific molecule, polysaccharide A, allows *B. fragilis* to signal through TLR2 to induce tolerance in mucosal T cells and colonize normally (128).

Region-specific expression of TLRs may help control bacterial populations in different niches, across regions of the gastrointestinal tract (131) or within a single villus (132). Thus, the benefits host and microbe receive from colonization may be impacted by the receipt of both MAMP and species-specific signaling molecules by TLRs, themselves varying in expression by both location and developmental stage (133, 134) in host tissue. Of this complex interplay between host and symbiont, it is fair to say, “good fences make good neighbors” (135). What form host TLR- symbiont MAMPs interactions may take in invertebrates is still an open question. Chapter 4 of this dissertation focuses on new insights in Toll-like receptors in *E. scolopes*.

REFERENCES

1. **McFall-Ngai MJ.** 2014. The importance of microbes in animal development: lessons from the squid-vibrio symbiosis. *Annu Rev Microbiol* **68**:177-194.
2. **McFall-Ngai M, Heath-Heckman EA, Gillette AA, Peyer SM, Harvie EA.** 2012. The secret languages of coevolved symbioses: insights from the *Euprymna scolopes-Vibrio fischeri* symbiosis. *Semin Immunol* **24**:3-8.
3. **Jones B, Nishiguchi M.** 2004. Counterillumination in the Hawaiian bobtail squid, *Euprymna scolopes* Berry (Mollusca : Cephalopoda). *Marine Biology* **144**:1151-1155.
4. **Visick KL, McFall-Ngai MJ.** 2000. An exclusive contract: specificity in the *Vibrio fischeri-Euprymna scolopes* partnership. *J Bacteriol* **182**:1779-1787.
5. **Nishiguchi MK, Ruby EG, McFall-Ngai MJ.** 1998. Competitive dominance among strains of luminous bacteria provides an unusual form of evidence for parallel evolution in Sepiolid squid-vibrio symbioses. *Appl Environ Microbiol* **64**:3209-3213.
6. **McFall-Ngai MJ, Ruby EG.** 1991. Symbiont recognition and subsequent morphogenesis as early events in an animal-bacterial mutualism. *Science* **254**:1491-1494.
7. **Medzhitov R, Janeway CA.** 1997. Innate immunity: the virtues of a nonclonal system of recognition. *Cell* **91**:295-298.
8. **Koropatnick TA, Engle JT, Apicella MA, Stabb EV, Goldman WE, McFall-Ngai MJ.** 2004. Microbial factor-mediated development in a host-bacterial mutualism. *Science* **306**:1186-1188.
9. **Medzhitov R.** 2001. Toll-like receptors and innate immunity. *Nat Rev Immunol* **1**:135-145.
10. **Foster JS, Apicella MA, McFall-Ngai MJ.** 2000. *Vibrio fischeri* lipopolysaccharide induces developmental apoptosis, but not complete morphogenesis, of the *Euprymna scolopes* symbiotic light organ. *Dev Biol* **226**:242-254.
11. **Nyholm SV, Deplancke B, Gaskins HR, Apicella MA, McFall-Ngai MJ.** 2002. Roles of *Vibrio fischeri* and nonsymbiotic bacteria in the dynamics of mucus secretion during symbiont colonization of the *Euprymna scolopes* light organ. *Appl Environ Microbiol* **68**:5113-5122.
12. **Koropatnick TA, Kimbell JR, McFall-Ngai MJ.** 2007. Responses of host hemocytes during the initiation of the squid-Vibrio symbiosis. *Biol Bull* **212**:29-39.
13. **Heath-Heckman E.** 2014. Day/Night Cycles in the *Euprymna scolopes – Vibrio fischeri* Symbiosis. Ph.D. University of Wisconsin, Madison, WI.
14. **Altura MA, Stabb E, Goldman W, Apicella M, McFall-Ngai MJ.** 2011. Attenuation of host NO production by MAMPs potentiates development of the host in the squid-vibrio symbiosis. *Cell Microbiol* **13**:527-537.

15. **Rakoff-Nahoum S, Paglino J, Eslami-Varzaneh F, Edberg S, Medzhitov R.** 2004. Recognition of commensal microflora by toll-like receptors is required for intestinal homeostasis. *Cell* **118**:229-241.
16. **Buchon N, Broderick NA, Chakrabarti S, Lemaitre B.** 2009. Invasive and indigenous microbiota impact intestinal stem cell activity through multiple pathways in *Drosophila*. *Genes Dev* **23**:2333-2344.
17. **Zaidman-Rémy A, Hervé M, Poidevin M, Pili-Floury S, Kim MS, Blanot D, Oh BH, Ueda R, Mengin-Lecreulx D, Lemaitre B.** 2006. The *Drosophila* amidase PGRP-LB modulates the immune response to bacterial infection. *Immunity* **24**:463-473.
18. **Swaminathan CP, Brown PH, Roychowdhury A, Wang Q, Guan R, Silverman N, Goldman WE, Boons GJ, Mariuzza RA.** 2006. Dual strategies for peptidoglycan discrimination by peptidoglycan recognition proteins (PGRPs). *Proc Natl Acad Sci U S A* **103**:684-689.
19. **Schleifer KH, Kandler O.** 1972. Peptidoglycan types of bacterial cell-walls and their taxonomic implications. *Bacteriological Reviews* **36**:407-477.
20. **Cloud-Hansen KA, Peterson SB, Stabb EV, Goldman WE, McFall-Ngai MJ, Handelsman J.** 2006. Breaching the great wall: peptidoglycan and microbial interactions. *Nat Rev Microbiol* **4**:710-716.
21. **Goodell EW.** 1985. Recycling of murein by *Escherichia coli*. *J Bacteriol* **163**:305-310.
22. **Jacobs C, Huang LJ, Bartowsky E, Normark S, Park JT.** 1994. Bacterial cell wall recycling provides cytosolic muropeptides as effectors for beta-lactamase induction. *EMBO J* **13**:4684-4694.
23. **Adin DM, Engle JT, Goldman WE, McFall-Ngai MJ, Stabb EV.** 2009. Mutations in *ampG* and lytic transglycosylase genes affect the net release of peptidoglycan monomers from *Vibrio fischeri*. *J Bacteriol* **191**:2012-2022.
24. **Sinha RK, Rosenthal RS.** 1980. Release of soluble peptidoglycan from growing gonococci: demonstration of anhydro-muramyl-containing fragments. *Infect Immun* **29**:914-925.
25. **Rosenthal RS, Nogami W, Cookson BT, Goldman WE, Folkening WJ.** 1987. Major fragment of soluble peptidoglycan released from growing *Bordetella pertussis* is tracheal cytotoxin. *Infect Immun* **55**:2117-2120.
26. **Salazar KA, Joffe NR, Dinguirard N, Houde P, Castillo MG.** 2015. Transcriptome analysis of the white body of the squid *Euprymna tasmanica* with emphasis on immune and hematopoietic gene discovery. *PLoS One* **10**:e0119949.
27. **Goodson MS, Kojadinovic M, Troll JV, Scheetz TE, Casavant TL, Soares MB, McFall-Ngai MJ.** 2005. Identifying components of the NF-kappaB pathway in the beneficial *Euprymna scolopes-Vibrio fischeri* light organ symbiosis. *Appl Environ Microbiol* **71**:6934-6946.

28. **Troll JV, Adin DM, Wier AM, Paquette N, Silverman N, Goldman WE, Stadermann FJ, Stabb EV, McFall-Ngai MJ.** 2009. Peptidoglycan induces loss of a nuclear peptidoglycan recognition protein during host tissue development in a beneficial animal-bacterial symbiosis. *Cell Microbiol* **11**:1114-1127.
29. **Troll JV, Bent EH, Pacquette N, Wier AM, Goldman WE, Silverman N, McFall-Ngai MJ.** 2010. Taming the symbiont for coexistence: a host PGRP neutralizes a bacterial symbiont toxin. *Environ Microbiol* **12**:2190-2203.
30. **Royet J, Gupta D, Dziarski R.** 2011. Peptidoglycan recognition proteins: modulators of the microbiome and inflammation. *Nat Rev Immunol* **11**:837-851.
31. **Shibata S, Visick KL.** 2012. Sensor kinase RscS induces the production of antigenically distinct outer membrane vesicles that depend on the symbiosis polysaccharide locus in *Vibrio fischeri*. *J Bacteriol* **194**:185-194.
32. **Brennan CA, Hunt JR, Kremer N, Krasity BC, Apicella MA, McFall-Ngai MJ, Ruby EG.** 2014. A model symbiosis reveals a role for sheathed-flagellum rotation in the release of immunogenic lipopolysaccharide. *Elife* **3**:e01579.
33. **Anwar MA, Choi S.** 2014. Gram-negative marine bacteria: structural features of lipopolysaccharides and their relevance for economically important diseases. *Mar Drugs* **12**:2485-2514.
34. **Preston A, Mandrell RE, Gibson BW, Apicella MA.** 1996. The lipooligosaccharides of pathogenic gram-negative bacteria. *Crit Rev Microbiol* **22**:139-180.
35. **Steeghs L, de Cock H, Evers E, Zomer B, Tommassen J, van der Ley P.** 2001. Outer membrane composition of a lipopolysaccharide-deficient *Neisseria meningitidis* mutant. *EMBO J* **20**:6937-6945.
36. **Munford RS, Varley AW.** 2006. Shield as signal: lipopolysaccharides and the evolution of immunity to gram-negative bacteria. *PLoS Pathog* **2**:e67.
37. **Phillips NJ, Adin DM, Stabb EV, McFall-Ngai MJ, Apicella MA, Gibson BW.** 2011. The lipid A from *Vibrio fischeri* lipopolysaccharide: a unique structure bearing a phosphoglycerol moiety. *J Biol Chem* **286**:21203-21219.
38. **Adin DM, Phillips NJ, Gibson BW, Apicella MA, Ruby EG, McFall-Ngai MJ, Hall DB, Stabb EV.** 2008. Characterization of htrB and msbB mutants of the light organ symbiont *Vibrio fischeri*. *Appl Environ Microbiol* **74**:633-644.
39. **Ernst RK, Yi EC, Guo L, Lim KB, Burns JL, Hackett M, Miller SI.** 1999. Specific lipopolysaccharide found in cystic fibrosis airway *Pseudomonas aeruginosa*. *Science* **286**:1561-1565.
40. **Moskowitz SM, Ernst RK.** 2010. The role of *Pseudomonas* lipopolysaccharide in cystic fibrosis airway infection. *Subcell Biochem* **53**:241-253.
41. **Weiss J.** 2003. Bactericidal/permeability-increasing protein (BPI) and lipopolysaccharide-binding protein (LBP): structure, function and regulation in host defence against Gram-negative bacteria. *Biochem Soc Trans* **31**:785-790.

42. **Manco M, Putignani L, Bottazzo GF.** 2010. Gut microbiota, lipopolysaccharides, and innate immunity in the pathogenesis of obesity and cardiovascular risk. *Endocr Rev* **31**:817-844.
43. **Canny G, Levy O, Furuta GT, Narravula-Alipati S, Sisson RB, Serhan CN, Colgan SP.** 2002. Lipid mediator-induced expression of bactericidal/ permeability-increasing protein (BPI) in human mucosal epithelia. *Proc Natl Acad Sci U S A* **99**:3902-3907.
44. **Canny G, Cario E, Lennartsson A, Gullberg U, Brennan C, Levy O, Colgan SP.** 2006. Functional and biochemical characterization of epithelial bactericidal/permeability-increasing protein. *Am J Physiol Gastrointest Liver Physiol* **290**:G557-567.
45. **Dethlefsen L, McFall-Ngai M, Relman DA.** 2007. An ecological and evolutionary perspective on human-microbe mutualism and disease. *Nature* **449**:811-818.
46. **Rosenstiel P, Philipp EE, Schreiber S, Bosch TC.** 2009. Evolution and function of innate immune receptors--insights from marine invertebrates. *J Innate Immun* **1**:291-300.
47. **Imler JL, Hoffmann JA.** 2001. Toll receptors in innate immunity. *Trends Cell Biol* **11**:304-311.
48. **Qiu X, Mistry A, Ammirati MJ, Chrnyk BA, Clark RW, Cong Y, Culp JS, Danley DE, Freeman TB, Geoghegan KF, Griffor MC, Hawrylik SJ, Hayward CM, Hensley P, Hoth LR, Karam GA, Lira ME, Lloyd DB, McGrath KM, Stutzman-Engwall KJ, Subashi AK, Subashi TA, Thompson JF, Wang IK, Zhao H, Seddon AP.** 2007. Crystal structure of cholesteryl ester transfer protein reveals a long tunnel and four bound lipid molecules. *Nat Struct Mol Biol* **14**:106-113.
49. **Bingle CD, LeClair EE, Havard S, Bingle L, Gillingham P, Craven CJ.** 2004. Phylogenetic and evolutionary analysis of the PLUNC gene family. *Protein Sci* **13**:422-430.
50. **Beamer LJ, Carroll SF, Eisenberg D.** 1998. The BPI/LBP family of proteins: a structural analysis of conserved regions. *Protein Sci* **7**:906-914.
51. **Lamping N, Hoess A, Yu B, Park TC, Kirschning CJ, Pfeil D, Reuter D, Wright SD, Herrmann F, Schumann RR.** 1996. Effects of site-directed mutagenesis of basic residues (Arg 94, Lys 95, Lys 99) of lipopolysaccharide (LPS)-binding protein on binding and transfer of LPS and subsequent immune cell activation. *J Immunol* **157**:4648-4656.
52. **Beamer LJ, Carroll SF, Eisenberg D.** 1997. Crystal structure of human BPI and two bound phospholipids at 2.4 angstrom resolution. *Science* **276**:1861-1864.
53. **Beamer LJ, Carroll SF, Eisenberg D.** 1999. The three-dimensional structure of human bactericidal/permeability-increasing protein: implications for understanding protein-lipopolysaccharide interactions. *Biochem Pharmacol* **57**:225-229.
54. **Ferguson AD, Welte W, Hofmann E, Lindner B, Holst O, Coulton JW, Diederichs K.** 2000. A conserved structural motif for lipopolysaccharide recognition by procaryotic and eucaryotic proteins. *Structure* **8**:585-592.

55. **Ellison RT, Giehl TJ.** 1991. Killing of gram-negative bacteria by lactoferrin and lysozyme. *J Clin Invest* **88**:1080-1091.
56. **Killmann H, Benz R, Braun V.** 1996. Properties of the FhuA channel in the *Escherichia coli* outer membrane after deletion of FhuA portions within and outside the predicted gating loop. *J Bacteriol* **178**:6913-6920.
57. **Weiss CA, Wasiluk KR, Kellogg TA, Dunn DL.** 2000. Bactericidal and endotoxin neutralizing activity of a peptide derived from *Limulus* antilipopolysaccharide factor. *Surgery* **128**:339-344.
58. **Reyes O, Vallespi MG, Garay HE, Cruz LJ, González LJ, Chinaea G, Buurman W, Araña MJ.** 2002. Identification of single amino acid residues essential for the binding of lipopolysaccharide (LPS) to LPS binding protein (LBP) residues 86-99 by using an Ala-scanning library. *J Pept Sci* **8**:144-150.
59. **Bruce C, Beamer LJ, Tall AR.** 1998. The implications of the structure of the bactericidal/permeability-increasing protein on the lipid-transfer function of the cholesteryl ester transfer protein. *Curr Opin Struct Biol* **8**:426-434.
60. **Stenvik J, Solstad T, Strand C, Leiros I, Jørgensen T T.** 2004. Cloning and analyses of a BPI/LBP cDNA of the Atlantic cod (*Gadus morhua* L.). *Dev Comp Immunol* **28**:307-323.
61. **King N, Westbrook MJ, Young SL, Kuo A, Abedin M, Chapman J, Fairclough S, Hellsten U, Isogai Y, Letunic I, Marr M, Pincus D, Putnam N, Rokas A, Wright KJ, Zuzow R, Dirks W, Good M, Goodstein D, Lemons D, Li W, Lyons JB, Morris A, Nichols S, Richter DJ, Salamov A, Sequencing JG, Bork P, Lim WA, Manning G, Miller WT, McGinnis W, Shapiro H, Tjian R, Grigoriev IV, Rokhsar D.** 2008. The genome of the choanoflagellate *Monosiga brevicollis* and the origin of metazoans. *Nature* **451**:783-788.
62. **Irazaqui JE, Urbach JM, Ausubel FM.** 2010. Evolution of host innate defence: insights from *Caenorhabditis elegans* and primitive invertebrates. *Nat Rev Immunol* **10**:47-58.
63. **Domazet-Lošo T, Brajković J, Tautz D.** 2007. A phylostratigraphy approach to uncover the genomic history of major adaptations in metazoan lineages. *Trends Genet* **23**:533-539.
64. **Inagawa H, Honda T, Kohchi C, Nishizawa T, Yoshiura Y, Nakanishi T, Yokomizo Y, Soma G.** 2002. Cloning and characterization of the homolog of mammalian lipopolysaccharide-binding protein and bactericidal permeability-increasing protein in rainbow trout *Oncorhynchus mykiss*. *J Immunol* **168**:5638-5644.
65. **Beamer LJ, Fischer D, Eisenberg D.** 1998. Detecting distant relatives of mammalian LPS-binding and lipid transport proteins. *Protein Sci* **7**:1643-1646.
66. **Chiang SC, Veldhuizen EJ, Barnes FA, Craven CJ, Haagsman HP, Bingle CD.** 2011. Identification and characterisation of the BPI/LBP/PLUNC-like gene repertoire in chickens reveals the absence of a LBP gene. *Dev Comp Immunol* **35**:285-295.

67. **Mitta G, Galinier R, Tisseyre P, Allienne JF, Girerd-Chambaz Y, Guillou F, Bouchut A, Coustau C.** 2005. Gene discovery and expression analysis of immune-relevant genes from *Biomphalaria glabrata* hemocytes. *Dev Comp Immunol* **29**:393-407.
68. **Hathaway JJ, Adema CM, Stout BA, Mobarak CD, Loker ES.** 2010. Identification of protein components of egg masses indicates parental investment in immunoprotection of offspring by *Biomphalaria glabrata* (gastropoda, mollusca). *Dev Comp Immunol* **34**:425-435.
69. **Gauthier ME, Du Pasquier L, Degnan BM.** 2010. The genome of the sponge *Amphimedon queenslandica* provides new perspectives into the origin of Toll-like and interleukin 1 receptor pathways. *Evol Dev* **12**:519-533.
70. **Kim YS, Ryu JH, Han SJ, Choi KH, Nam KB, Jang IH, Lemaitre B, Brey PT, Lee WJ.** 2000. Gram-negative bacteria-binding protein, a pattern recognition receptor for lipopolysaccharide and beta-1,3-glucan that mediates the signaling for the induction of innate immune genes in *Drosophila melanogaster* cells. *J Biol Chem* **275**:32721-32727.
71. **Warr E, Das S, Dong Y, Dimopoulos G.** 2008. The Gram-negative bacteria-binding protein gene family: its role in the innate immune system of *Anopheles gambiae* and in anti-*Plasmodium* defence. *Insect Mol Biol* **17**:39-51.
72. **Lee SY, Wang R, Söderhäll K.** 2000. A lipopolysaccharide- and beta-1,3-glucan-binding protein from hemocytes of the freshwater crayfish *Pacifastacus leniusculus*. Purification, characterization, and cDNA cloning. *J Biol Chem* **275**:1337-1343.
73. **Lee WJ, Lee JD, Kravchenko VV, Ulevitch RJ, Brey PT.** 1996. Purification and molecular cloning of an inducible gram-negative bacteria-binding protein from the silkworm, *Bombyx mori*. *Proc Natl Acad Sci U S A* **93**:7888-7893.
74. **Gonzalez M, Gueguen Y, Destoumieux-Garzón D, Romestand B, Fievet J, Pugnère M, Roquet F, Escoubas JM, Vandenbulcke F, Levy O, Sauné L, Bulet P, Bachère E.** 2007. Evidence of a bactericidal permeability increasing protein in an invertebrate, the *Crassostrea gigas* Cg-BPI. *Proc Natl Acad Sci U S A* **104**:17759-17764.
75. **Eilers B, Mayer-Scholl A, Walker T, Tang C, Weinrauch Y, Zychlinsky A.** 2010. Neutrophil antimicrobial proteins enhance *Shigella flexneri* adhesion and invasion. *Cell Microbiol* **12**:1134-1143.
76. **Gasteiger E, Gattiker A, Hoogland C, Ivanyi I, Appel RD, Bairoch A.** 2003. ExPASy: The proteomics server for in-depth protein knowledge and analysis. *Nucleic Acids Res* **31**:3784-3788.
77. **Solstad T, Stenvik J, Jørgensen T.** 2007. mRNA expression patterns of the BPI/LBP molecule in the Atlantic cod (*Gadus morhua* L.). *Fish Shellfish Immunol* **23**:260-271.
78. **Glavin DP, Cleaves HJ, Schubert M, Aubrey A, Bada JL.** 2004. New method for estimating bacterial cell abundances in natural samples by use of sublimation. *Appl Environ Microbiol* **70**:5923-5928.

79. **Tirapé A, Bacque C, Brizard R, Vandenbulcke F, Boulo V.** 2007. Expression of immune-related genes in the oyster *Crassostrea gigas* during ontogenesis. *Dev Comp Immunol* **31**:859-873.
80. **McFall-Ngai M.** 2008. Host-microbe symbiosis: the squid-Vibrio association--a naturally occurring, experimental model of animal/bacterial partnerships. *Adv Exp Med Biol* **635**:102-112.
81. **Norimatsu M, Ono T, Aoki A, Ohishi K, Takahashi T, Watanabe G, Taya K, Sasamoto S, Tamura Y.** 1995. Lipopolysaccharide-induced apoptosis in swine lymphocytes in vivo. *Infect Immun* **63**:1122-1126.
82. **Norimatsu M, Ono T, Aoki A, Ohishi K, Tamura Y.** 1995. In-vivo induction of apoptosis in murine lymphocytes by bacterial lipopolysaccharides. *J Med Microbiol* **43**:251-257.
83. **Letunic I, Doerks T, Bork P.** 2012. SMART 7: recent updates to the protein domain annotation resource. *Nucleic Acids Res* **40**:D302-305.
84. **Chun CK, Troll JV, Koroleva I, Brown B, Manzella L, Snir E, Almabrazi H, Scheetz TE, Bonaldo MeF, Casavant TL, Soares MB, Ruby EG, McFall-Ngai MJ.** 2008. Effects of colonization, luminescence, and autoinducer on host transcription during development of the squid-vibrio association. *Proc Natl Acad Sci U S A* **105**:11323-11328.
85. **Budd GE.** 2008. The earliest fossil record of the animals and its significance. *Philos Trans R Soc Lond B Biol Sci* **363**:1425-1434.
86. **Anderson KV, Jürgens G, Nüsslein-Volhard C.** 1985. Establishment of dorsal-ventral polarity in the *Drosophila* embryo: genetic studies on the role of the Toll gene product. *Cell* **42**:779-789.
87. **Manfrulli P, Reichhart JM, Steward R, Hoffmann JA, Lemaitre B.** 1999. A mosaic analysis in *Drosophila* fat body cells of the control of antimicrobial peptide genes by the Rel proteins Dorsal and DIF. *EMBO J* **18**:3380-3391.
88. **Means TK, Golenbock DT, Fenton MJ.** 2000. Structure and function of Toll-like receptor proteins. *Life Sci* **68**:241-258.
89. **Leulier F, Lemaitre B.** 2008. Toll-like receptors--taking an evolutionary approach. *Nat Rev Genet* **9**:165-178.
90. **DeYoung BJ, Innes RW.** 2006. Plant NBS-LRR proteins in pathogen sensing and host defense. *Nat Immunol* **7**:1243-1249.
91. **Wiens M, Korzhev M, Perovic-Ottstadt S, Luthringer B, Brandt D, Klein S, Müller WE.** 2007. Toll-like receptors are part of the innate immune defense system of sponges (demospongiae: Porifera). *Mol Biol Evol* **24**:792-804.
92. **Enkhbayar P, Kamiya M, Osaki M, Matsumoto T, Matsushima N.** 2004. Structural principles of leucine-rich repeat (LRR) proteins. *Proteins* **54**:394-403.

93. **Kobe B, Kajava AV.** 2001. The leucine-rich repeat as a protein recognition motif. *Curr Opin Struct Biol* **11**:725-732.
94. **Roach JC, Glusman G, Rowen L, Kaur A, Purcell MK, Smith KD, Hood LE, Aderem A.** 2005. The evolution of vertebrate Toll-like receptors. *Proc Natl Acad Sci U S A* **102**:9577-9582.
95. **Imler JL, Zheng L.** 2004. Biology of Toll receptors: lessons from insects and mammals. *J Leukoc Biol* **75**:18-26.
96. **Hoshino K, Takeuchi O, Kawai T, Sanjo H, Ogawa T, Takeda Y, Takeda K, Akira S.** 1999. Cutting edge: Toll-like receptor 4 (TLR4)-deficient mice are hyporesponsive to lipopolysaccharide: evidence for TLR4 as the Lps gene product. *J Immunol* **162**:3749-3752.
97. **Lorenz E, Mira JP, Frees KL, Schwartz DA.** 2002. Relevance of mutations in the TLR4 receptor in patients with gram-negative septic shock. *Arch Intern Med* **162**:1028-1032.
98. **Nagai Y, Akashi S, Nagafuku M, Ogata M, Iwakura Y, Akira S, Kitamura T, Kosugi A, Kimoto M, Miyake K.** 2002. Essential role of MD-2 in LPS responsiveness and TLR4 distribution. *Nat Immunol* **3**:667-672.
99. **Hepburn L, Prajsnar TK, Klapholz C, Moreno P, Loynes CA, Ogryzko NV, Brown K, Schiebler M, Hegyi K, Antrobus R, Hammond KL, Connolly J, Ochoa B, Bryant C, Otto M, Surewaard B, Seneviratne SL, Grogono DM, Cachat J, Ny T, Kaser A, Török ME, Peacock SJ, Holden M, Blundell T, Wang L, Ligoxygakis P, Minichiello L, Woods CG, Foster SJ, Renshaw SA, Floto RA.** 2014. Innate immunity. A Spaetzle-like role for nerve growth factor β in vertebrate immunity to *Staphylococcus aureus*. *Science* **346**:641-646.
100. **Kielian T.** 2009. Overview of toll-like receptors in the CNS. *Curr Top Microbiol Immunol* **336**:1-14.
101. **Kaul D, Habel P, Derkow K, Krüger C, Franzoni E, Wulczyn FG, Bereswill S, Nitsch R, Schott E, Veh R, Naumann T, Lehnardt S.** 2012. Expression of Toll-like receptors in the developing brain. *PLoS One* **7**:e37767.
102. **Ma Y, Li J, Chiu I, Wang Y, Sloane JA, Lü J, Kosaras B, Sidman RL, Volpe JJ, Vartanian T.** 2006. Toll-like receptor 8 functions as a negative regulator of neurite outgrowth and inducer of neuronal apoptosis. *J Cell Biol* **175**:209-215.
103. **Cameron JS, Alexopoulou L, Sloane JA, DiBernardo AB, Ma Y, Kosaras B, Flavell R, Strittmatter SM, Volpe J, Sidman R, Vartanian T.** 2007. Toll-like receptor 3 is a potent negative regulator of axonal growth in mammals. *J Neurosci* **27**:13033-13041.
104. **Park SJ, Lee JY, Kim SJ, Choi SY, Yune TY, Ryu JH.** 2015. Toll-like receptor-2 deficiency induces schizophrenia-like behaviors in mice. *Sci Rep* **5**:8502.
105. **Khariv V, Pang K, Servatius RJ, David BT, Goodus MT, Beck KD, Heary RF, Elkabes S.** 2013. Toll-like receptor 9 deficiency impacts sensory and motor behaviors. *Brain Behav Immun* **32**:164-172.

106. **Coscia M, Giacomelli S, Oreste U.** 2011. Toll-like receptors: an overview from invertebrates to vertebrates. *Isj-Invertebrate Survival Journal* **8**:210-226.
107. **Satake H, Sekiguchi T.** 2012. Toll-like receptors of deuterostome invertebrates. *Front Immunol* **3**:34.
108. **Sasaki N, Ogasawara M, Sekiguchi T, Kusumoto S, Satake H.** 2009. Toll-like receptors of the ascidian *Ciona intestinalis*: prototypes with hybrid functionalities of vertebrate Toll-like receptors. *J Biol Chem* **284**:27336-27343.
109. **Huang S, Yuan S, Guo L, Yu Y, Li J, Wu T, Liu T, Yang M, Wu K, Liu H, Ge J, Huang H, Dong M, Yu C, Chen S, Xu A.** 2008. Genomic analysis of the immune gene repertoire of amphioxus reveals extraordinary innate complexity and diversity. *Genome Res* **18**:1112-1126.
110. **Rast JP, Smith LC, Loza-Coll M, Hibino T, Litman GW.** 2006. Genomic insights into the immune system of the sea urchin. *Science* **314**:952-956.
111. **Hibino T, Loza-Coll M, Messier C, Majeske AJ, Cohen AH, Terwilliger DP, Buckley KM, Brockton V, Nair SV, Berney K, Fugmann SD, Anderson MK, Pancer Z, Cameron RA, Smith LC, Rast JP.** 2006. The immune gene repertoire encoded in the purple sea urchin genome. *Dev Biol* **300**:349-365.
112. **Schneider DS, Jin Y, Morisato D, Anderson KV.** 1994. A processed form of the Spätzle protein defines dorsal-ventral polarity in the *Drosophila* embryo. *Development* **120**:1243-1250.
113. **Stein D, Roth S, Vogelsang E, Nüsslein-Volhard C.** 1991. The polarity of the dorsoventral axis in the *Drosophila* embryo is defined by an extracellular signal. *Cell* **65**:725-735.
114. **Wang PH, Liang JP, Gu ZH, Wan DH, Weng SP, Yu XQ, He JG.** 2012. Molecular cloning, characterization and expression analysis of two novel Toll (LvToll2 and LvToll3) and three putative Spätzle-like Toll ligands (LvSpz1-3) from *Litopenaeus vannamei*. *Dev Comp Immunol* **36**:359-371.
115. **Yang LS, Yin ZX, Liao JX, Huang XD, Guo CJ, Weng SP, Chan SM, Yu XQ, He JG.** 2007. A Toll receptor in shrimp. *Mol Immunol* **44**:1999-2008.
116. **Pujol N, Link EM, Liu LX, Kurz CL, Alloing G, Tan MW, Ray KP, Solari R, Johnson CD, Ewbank JJ.** 2001. A reverse genetic analysis of components of the Toll signaling pathway in *Caenorhabditis elegans*. *Curr Biol* **11**:809-821.
117. **Tenor JL, Aballay A.** 2008. A conserved Toll-like receptor is required for *Caenorhabditis elegans* innate immunity. *EMBO Rep* **9**:103-109.
118. **Pradel E, Zhang Y, Pujol N, Matsuyama T, Bargmann CI, Ewbank JJ.** 2007. Detection and avoidance of a natural product from the pathogenic bacterium *Serratia marcescens* by *Caenorhabditis elegans*. *Proc Natl Acad Sci U S A* **104**:2295-2300.

119. **Couillault C, Pujol N, Reboul J, Sabatier L, Guichou JF, Kohara Y, Ewbank JJ.** 2004. TLR-independent control of innate immunity in *Caenorhabditis elegans* by the TIR domain adaptor protein TIR-1, an ortholog of human SARM. *Nat Immunol* **5**:488-494.
120. **Halanych KM, Kocot KM.** 2014. Repurposed transcriptomic data facilitate discovery of innate immunity toll-like receptor (TLR) Genes across Lophotrochozoa. *Biol Bull* **227**:201-209.
121. **Zhang Y, He X, Yu F, Xiang Z, Li J, Thorpe KL, Yu Z.** 2013. Characteristic and functional analysis of toll-like receptors (TLRs) in the lophotrochozoan, *Crassostrea gigas*, reveals ancient origin of TLR-mediated innate immunity. *PLoS One* **8**:e76464.
122. **Zhang L, Li L, Guo X, Litman GW, Dishaw LJ, Zhang G.** 2015. Massive expansion and functional divergence of innate immune genes in a protostome. *Sci Rep* **5**:8693.
123. **Davidson CR, Best NM, Francis JW, Cooper EL, Wood TC.** 2008. Toll-like receptor genes (TLRs) from *Capitella capitata* and *Helobdella robusta* (Annelida). *Dev Comp Immunol* **32**:608-612.
124. **Schikorski D, Cuvillier-Hot V, Boidin-Wichlacz C, Slomianny C, Salzet M, Tasiemski A.** 2009. Deciphering the immune function and regulation by a TLR of the cytokine EMAPII in the lesioned central nervous system using a leech model. *J Immunol* **183**:7119-7128.
125. **Zhang L, Li L, Zhang G.** 2011. A *Crassostrea gigas* Toll-like receptor and comparative analysis of TLR pathway in invertebrates. *Fish Shellfish Immunol* **30**:653-660.
126. **Wang M, Yang J, Zhou Z, Qiu L, Wang L, Zhang H, Gao Y, Wang X, Zhang L, Zhao J, Song L.** 2011. A primitive Toll-like receptor signaling pathway in mollusk Zhikong scallop *Chlamys farreri*. *Dev Comp Immunol* **35**:511-520.
127. **Mateo DR, Greenwood SJ, Araya MT, Berthe FC, Johnson GR, Siah A.** 2010. Differential gene expression of gamma-actin, Toll-like receptor 2 (TLR-2) and interleukin-1 receptor-associated kinase 4 (IRAK-4) in *Mya arenaria* haemocytes induced by in vivo infections with two *Vibrio splendidus* strains. *Dev Comp Immunol* **34**:710-714.
128. **Round JL, Lee SM, Li J, Tran G, Jabri B, Chatila TA, Mazmanian SK.** 2011. The Toll-like receptor 2 pathway establishes colonization by a commensal of the human microbiota. *Science* **332**:974-977.
129. **Siegemund S, Sauer K.** 2012. Balancing pro- and anti-inflammatory TLR4 signaling. *Nat Immunol* **13**:1031-1033.
130. **Peres AG, Stegen C, Li J, Xu AQ, Levast B, Surette MG, Cousineau B, Desrosiers M, Madrenas J.** 2015. Uncoupling of Pro- and Anti-Inflammatory Properties of *Staphylococcus aureus*. *Infect Immun* **83**:1587-1597.
131. **Mowat AM, Agace WW.** 2014. Regional specialization within the intestinal immune system. *Nat Rev Immunol* **14**:667-685.

132. **Furrie E, Macfarlane S, Thomson G, Macfarlane GT, Group MGB, Bank TTT.** 2005. Toll-like receptors-2, -3 and -4 expression patterns on human colon and their regulation by mucosal-associated bacteria. *Immunology* **115**:565-574.
133. **Kolluri G, Ramamurthy N, Churchil RR, Dhinakar Raj G, Kannaki TR.** 2014. Influence of age, sex and rearing systems on Toll-like receptor 7 (TLR7) expression pattern in gut, lung and lymphoid tissues of indigenous ducks. *Br Poult Sci* **55**:59-67.
134. **Uddin MJ, Kaewmala K, Tesfaye D, Tholen E, Looft C, Hoelker M, Schellander K, Cinar MU.** 2013. Expression patterns of porcine Toll-like receptors family set of genes (TLR1-10) in gut-associated lymphoid tissues alter with age. *Res Vet Sci* **95**:92-102.
135. **Moens E, Veldhoen M.** 2012. Epithelial barrier biology: good fences make good neighbours. *Immunology* **135**:1-8.

CHAPTER 2

Structural and Functional Features of a Developmentally Regulated Lipopolysaccharide-Binding Protein

PREFACE

A version of this chapter is in review for publication in MBio, submitted 04/10/2015, with the authors Benjamin C. Krasity, Joshua V. Troll, Erik M. Lehnert, Kathleen T. Hackett, Joseph P. Dillard, Michael A. Apicella, William E. Goldman, Jerrold P. Weiss, Margaret J. McFall-Ngai, and the title: “Structural and functional features of a developmentally regulated lipopolysaccharide-binding protein.” All supplemental material has been included here and the figure and table numbering altered.

JVT conducted immunocytochemistry and EML conducted hybridization chain reaction fluorescent *in situ* hybridization. Other experiments were conducted by BCK, with input from KTH, JPD and JPW. The manuscript was primarily written by BCK, with significant revisions from JPW and MMN and contributions from other authors.

ABSTRACT Lipopolysaccharide (LPS) binding proteins, (LBPs), occur mainly in extracellular fluids of mammals and promote LPS delivery to specific host cell receptors. The function of LBP has been studied principally in the context of host defense; the possible role of LBPs in non-pathogenic host-microbe interactions has not been well characterized. Using the *Euprymna scolopes-Vibrio fischeri* model, we analyzed the structure and function of an LBP family protein, EsLBP1, and provide evidence for its role in triggering a symbiont-induced host developmental program. Previous studies showed that, during initial host colonization, the LPS of *V. fischeri* synergizes with peptidoglycan (PGN) monomer to induce morphogenesis of epithelial tissues of the host animal. Computationally modeled EsLBP1 shares certain but not all structural features of mammalian LBPs thought important for LPS binding. Similar to human LBP, recombinant EsLBP1 expressed in insect cells bound *V. fischeri* LPS and *Neisseria meningitidis* lipooligosaccharide (LOS) with nM or greater affinity, but only weakly bound *Francisella tularensis* LPS and did not bind PGN monomer. Unlike human LBP, EsLBP1 did not bind *N. meningitidis* LOS:CD14 complexes. The *eslbp1* transcript was up-regulated ~22-fold by *V. fischeri* at 24 h post-inoculation. Surprisingly, this upregulation was not induced by exposure to LPS, but rather to the PGN monomer alone. Hybridization chain reaction fluorescent *in situ* hybridization (HCR-FISH) and immunocytochemistry (ICC) localized *eslbp1* transcript and protein to crypt epithelia, where *V. fischeri* induces morphogenesis. The data presented here provide a window into the evolution of LBPs and the scope of their roles in animal symbioses.

IMPORTANCE The mechanism by which *V. fischeri* lipopolysaccharide (LPS) is conveyed to host cells, and the scope of its action as a signal molecule, remain unknown. Mammalian LPS-binding protein (LBP) is implicated in these activities, and has been a reliable biomarker for

chronic bacterial colonization. Although specific amino acids putatively critical for LPS binding in mammalian LBP are substituted in EsLBP1, high-affinity ligand binding is retained. These data suggest that overall charge of the N-terminal tip region, conserved in the modeled tertiary structure, is essential. Expression of *eslbp1* is induced by peptidoglycan derivatives, not LPS, possibly contributing to the ability of these two bacterial signals to act synergistically in induction of epithelial morphogenesis in the host. The widespread occurrence of EsLBP1 in tissues other than the symbiotic organ suggests it was recruited from a general defensive role to one that mediates specific interactions with its symbiont.

INTRODUCTION

Over the last decade, several studies have demonstrated that mutualistic and pathogenic associations share a common molecular language (for reviews see (1, 2)). Key elements of this dialogue are microbe-associated molecular patterns (MAMPs) of microbial cell surfaces, including lipopolysaccharide (LPS) and peptidoglycan (PGN) derivatives, and their cognate host sensors and receptors. The model mutualistic association between the Hawaiian bobtail squid *Euprymna scolopes* and the bioluminescent marine bacterium *Vibrio fischeri* has provided an experimental system for the study of these interactions. In the squid-vibrio association, symbiont lipid A, the lipid component of LPS, and the PGN monomer TCT (tracheal cytotoxin) direct various symbiont-induced developmental programs of the ‘light organ’, the set of tissues that harbors the symbionts and modifies their luminescence for its use in host behaviors. MAMPs-induced morphogenetic programs include the apoptotic loss of the ciliated epithelium that potentiates initial colonization, recruitment of macrophage-like blood cells into the organ, and the transformation of the biochemical environment where the symbionts take up residence in

host tissues (for reviews see (3, 4)). Subsequent studies in mammals have demonstrated that MAMPs also drive development, such as immune system maturation, in complex host/microbial associations including in the gut (2). This study uses the squid-vibrio model to provide experimental evidence for the role of lipopolysaccharide binding protein (LBP), a member of the LBP/BPI (bactericidal/permeability-increasing protein) protein family, as a key player in host responses to symbiont LPS.

As MAMPs recognition proteins, mammalian LBP and BPI are closely related, structurally similar proteins that serve complementary roles in innate immunity. LBP relays LPS to host cells whereas BPI acts as a bactericidal/LPS-neutralizing effector molecule to help clear infection and resolve infection-induced inflammation (5, 6). LBP is present in plasma and tissue fluids, and recent reports have determined that LBP and BPI are expressed in the gut and other epithelia, where they likely mediate responses to the interfacing microbiota (7, 8). During an inflammatory response, the LBP levels present constitutively are sufficient to catalyze extraction and delivery of individual LPS molecules to CD14 and MD-2/TLR4 to induce potent pro-inflammatory responses (9-12). LBP levels can increase up to 100-fold during acute-phase reactions following the initial inflammatory responses, promoting, along with BPI, “silent” clearance of LPS and eventual resolution of LPS-triggered inflammation (10, 13, 14). Unlike BPI, LBP binding to Gram-negative bacteria (e.g., *Escherichia coli*) does not produce lethal or sublethal alterations of the bacteria (15).

Despite differences in activity between mammalian LBP and BPI, the structural organization conferring function is generally similar. Thus, the N-terminal domain of each protein binds LPS, whether presented in the form of intact Gram-negative bacteria (15, 16), shed outer membrane blebs, or aggregates of purified LPS (17-19). The C-terminal domain of both

LBP and BPI is responsible for delivery of the bound LPS-containing material to host cells (17, 19, 20). Conserved, positively charged residues near the N-terminal tip of both mammalian LBP and BPI are likely important in initial electrostatic interactions with typically polyanionic LPS (21-24). Net cationicity is greater in BPI than in LBP in both the charged N-terminal tip and across the molecule; the isoelectric point (pI) of LBP is near neutral, whereas the BPI pI is ~10. This difference in charge likely accounts for the higher LPS and bacterial binding affinity of BPI and the distinct effects that LBP and BPI produce on LPS-rich interfaces, including bacterial outer membranes (20, 25). BPI causes bacterial sublethal and lethal injury, whereas LBP promotes extraction of individual LPS molecules by CD14 at substoichiometric concentrations (11, 18). Although the gene duplication giving rise to mammalian LBP and BPI is believed to have occurred after the radiation of the mammals (26), related proteins have been reported in several invertebrate groups (27, 28). While a given invertebrate species may have multiple LBP/BPI gene isoforms, an LBP/BPI dichotomy has not been well characterized outside of the mammals. Sequence analysis across the animal kingdom shows that these proteins are relatively quickly evolving, with ~21% identity between EsLBP1 and mammalian LBP/BPI proteins indicating ~1% change in the amino acid sequence every 7 million years. A comparative analysis of structure and function across these deep divergences, over greater than 500 million years, offers the opportunity to define biochemical features that are essential for function of these molecules, as well as to explore how evolutionary tinkering can give rise to diversity of function.

Full length transcripts of four members of the LBP/BPI protein family have been identified in *E. scolopes* (27, 29). One of these proteins, EsLBP1 (previously called EsLBP (30)), is of particular interest as a candidate for responding to symbiont LPS during development. The pI of EsLBP1 is near neutral, suggesting it functions similarly to mammalian LBPs, *i.e.*,

presenting LPS as a signal molecule. In addition, the gene encoding EsLBP1 increases in expression at 18 h following the onset of host-symbiont interaction, a time when morphogenesis is being signaled by MAMPs (30). In the present study, to provide insight into the possible role of an LBP in inducing an animal developmental program, we sought to characterize the biochemical properties of EsLBP1, comparing them to those of mammalian LBP, and to examine the timing and location of EsLBP1 gene expression and protein production through the early trajectory of host development.

MATERIALS AND METHODS

Alignment of EsLBP1 with mammalian LBP and BPI. Human BPI (NCBI accession number CAD99178.1), human LBP (AAB31143.1) and EsLBP1 (JF514880.1) sequences were aligned by clustalW2 (31, 32). LBP/BPI family N-terminal domains were defined by SMART (33). A 3D model of EsLBP1 was formed with SWISS-MODEL (34-37); the model with the best QMEAN4 score (38) was chosen. Images were generated with pv Java Script Protein Viewer (DOI10.5281/zenodo.12620).

Expression of EsLBP1 in insect cells. The complete open reading frame of EsLBP1 was amplified from *E. scolopes* cDNA using primers LBP1pBAC3F: 5'-ATACACCATGGTAATGTCTTGCCCCACTCAA-3' and LBP1pBAC3R: 5'-TATCACTCGAGAATAGATGTAATTGCCAAGTC-3'. Two SNPs resulting in amino acid substitutions relative to the published EsLBP1 sequence, T253S and A218T, were consistently noted in this cDNA preparation and were included in the expressed recombinant EsLBP1. The PCR product was digested with EcoRI and XhoI and ligated into the plasmid pBAC-3 (EMD Millipore, Billerica, MA), adding a leader peptide and His₆-tag, and sequenced. The construct

was transfected into and expressed from Sf9 insect cells by Kinnakeet Biotechnology (Midlothian, VA). Conditioned medium from these cells was blotted with anti-tetra-His antibodies (Qiagen), by which means the concentration of EsLBP1 was estimated at 2 μ M in the undiluted medium. Control medium including His-tagged *E. coli* β -glucuronidase was prepared from cells transfected with the BacMagic3 transfection control plasmid (EMD Millipore).

Preparation of MAMPs. [3 H]LOS (5000 cpm/ng LOS) and [14 C]LOS (6 cpm/ng LOS) were extracted and purified from metabolically labeled *Neisseria meningitidis* and used as aggregates of purified LOS as described previously (39). The [14 C]LOS was used as unlabeled LOS in competition experiments with [3 H]LOS (Fig. 2-2B). Metabolic labeling with radiolabeled acetate yielded equivalent radiolabeling of all LOS molecules (40). Unlabeled lipopolysaccharide and lipid A from wild-type *V. fischeri*, *F. tularensis* and *N. meningitidis* were prepared by the water-phenol method (41); diphosphoryl *E. coli* lipid A was purchased from Sigma Aldrich (St. Louis, MO). LPS and LPS derivative stock solutions were sonicated before use as previously (42); preparations used in animal experiments were initially solubilized at 1 mg/mL in 10 mM PIPES buffer pH 6.3 before dilution. Unlabeled TCT was purified from *Bordetella pertussis* (43); endotoxin contamination was undetectable as assayed with Pyrochrome chromogenic reagent (Associates of Cape Cod, East Falmouth, MA). To prepare [3 H]TCT, peptidoglycan of *N. gonorrhoeae* strain KH619 (MS11 *ldcA*) was metabolically labeled using [6 - 3 H]-glucosamine (44) and sacculi were purified as described (45). [3 H]TCT was produced by digestion of sacculi with gonococcal LtgD, and TCT was purified by reversed-phase HPLC (45).

EsLBP1 - endotoxin, TCT and LOS:CD14 complex binding assays. Conditioned medium containing EsLBP1 was used in parallel with equal volumes of control (β -

glucuronidase-containing) conditioned medium (Fig. 2-2A, 2-2B) or diluted with control conditioned medium (Fig. 2-2D) to permit testing of a range of EsLBP1 concentrations. Medium containing EsLBP1 or control medium was mixed with a [³H]-radiolabeled *N. meningitidis* MAMP, either 5000 cpm of LOS (Fig. 2-2A) or LOS complexed with human CD14 (12) (Fig. 2-2C), or 1000 cpm TCT (Fig. 2-2D), along with 0.1% human serum albumin in Dulbecco's phosphate-buffered saline (PBSA) in 0.2 mL reactions and incubated 30 min at 27°C. 10 µL Ni²⁺-agarose resin was added, the reactions were brought to 0.5 mL total with PBSA and the reactions were incubated on a rotating wheel for an additional 30 min. The beads were pelleted by gentle centrifugation for 1 min and washed twice with PBSA for 5 min each. Recovery of radiolabeled material was measured in each of the recovered fractions by liquid scintillation spectroscopy. Co-capture of radiolabeled material was calculated as percent of total recovered radioactive material present in beads. For competition experiments (Fig. 2-2B), unlabeled LOS/LPS was pre-incubated with 0.33 nM EsLBP1 for 30 minutes at 27°C before addition of ³H-LOS and incubation and co-capture as described above.

Assay for bactericidal activity of proteins. Cultures of PL2 *E. coli* were grown in tryptic soy broth (TSB) at 37°C to mid-log phase, then diluted 100-fold in TSB with no recombinant protein or 30 nM of one of the following: hBPI-21 (N-terminal fragment of human BPI), hLBP (human LBP) (both from Xoma, Berkeley, CA) or EsLBP1. Cultures were incubated a further 60 min at 37°C. Aliquots of each culture were diluted 1:6250 and 3 x 10µL were plated on tryptic soy agar, incubated overnight at 37°C; then CFUs were counted.

General procedures for animal experiments. Adult *E. scolopes* animals were collected from the sand flats of Oahu, Hawaii and transported and maintained as described in previous publications (46); experiments used newly hatched juveniles in artificial seawater collected from

the table on which they hatched. Symbiotic animals were exposed to ~5000 CFU/mL of *V. fischeri* ES114 (47); aposymbiotic animals were not. Symbiosis was verified with luminescence using a TD-20/20 luminometer (Turner Designs, Sunnydale, CA). In MAMP-treatment experiments, lipid A was used at 10 ng/mL and TCT at 1 μ M.

Quantitative reverse-transcriptase PCR. Quantitative reverse-transcriptase PCR.

We stabilized light organ tissues, extracted RNA, prepared cDNA and conducted qRT-PCR experiments in accordance with MIQE guidelines (48), as described previously (49), except qRT-PCR used gene-specific primers given in Table 2-1 and the protocol: 3 min at 94°C, 40×[15 s at 94°C, 20s at 59°C, 20 s at 68°C]. We used the comparative $\Delta\Delta C_q$ method to determine expression levels (50). *eslbp1* levels were normalized to the mean levels of control transcripts for the 40S ribosomal subunit and serine hydroxymethyltransferase (HMT).

Hybridization chain reaction fluorescent *in situ* hybridization (HCR-FISH). HCR-FISH to visualize transcripts of *E. scolopes* and *V. fischeri* genes was performed according to established protocols (51). All animals examined were collected at 24 h after exposure to ES114 *V. fischeri*. *E. scolopes* transcripts probed were those of *eslbp1* and *hsp90*, a counterstain for *E. scolopes* tissue. 16S ribosomal subunit transcript was probed to counterstain *V. fischeri*. Probe sequences are given in Table 2-2.

Immunocytochemistry. Immunocytochemistry for EsLBP1 in 18h post-inoculation hatchlings was performed with anti-EsLBP1 chicken antibody and viewed with a LSM510 laser-scanning confocal microscope (Zeiss, Thornwood, NY) as described previously (30), except the 1° antibody incubation was reduced to 7 days. Fluorophores included goat-anti-chicken-FITC antibody (for EsLBP1), rhodamine-phalloidin (actin cytoskeleton) and TOTO-3 (DNA).

Table 2-1: Primers used for qRT-PCR experiments

Gene	Primer name	Primer sequence
EsLBP1	LBP1qRTF4	TGGTTATCAGCTTTCAGGCCACCT
EsLBP1	LBP1qRTR4	AGACCTCCGTTTGTTCGCCATAGTT
Serine HMT	SerHMTqF	GTCCTGGTGACAAGAGTGCAATGA
Serine HMT	SerHMTqR	TTCCAGCAGAAAGGCACGATAGGT
40S ribosomal subunit	40SF2	AATCTCGGCGTCCTTGAGAA
40S ribosomal subunit	40SR2	GCATCAATTGCACGACGAGT

Table 2-2: Probe sequences for HCR-FISH experiments

Probe	Amplifier (52)/ Fluorophore	Probe sequence
<i>E. scolopes</i> EsLBP1 Probe 1	B2/Alexa 546	CCTCGTAAATCCTCATCAATCATCCAGTAAACCGCCAAAAACACGTATACCTGGATTACAG GGCTACTGTTCCCTTTGAGTGGGGCAAGACAAAAAAGCTCAGTCCATCCTCGTAAATCCTCAT CAATCATC
<i>E. scolopes</i> EsLBP1 Probe 2	B2/Alexa 546	CCTCGTAAATCCTCATCAATCATCCAGTAAACCGCCAAAAACAATGTGCTGCATTCAAATT CAGACCTCCGTTTGTGCCATAGTTGGGGAAAAAAGCTCAGTCCATCCTCGTAAATCCTCAT CAATCATC
<i>E. scolopes</i> EsLBP1 Probe 3	B2/Alexa 546	CCTCGTAAATCCTCATCAATCATCCAGTAAACCGCCAAAAATAACCGTGGCTGATCGAGCGA GTCACCTGGATAATTGGCAAGTTTTGTAGCAAAAAAGCTCAGTCCATCCTCGTAAATCCTCAT CAATCATC
<i>E. scolopes</i> EsLBP1 Probe 4	B2/Alexa 546	CCTCGTAAATCCTCATCAATCATCCAGTAAACCGCCAAAAACGGAAATTTACCACTGACTGC CTTACACTGGCAACTTGTTCCTTGTGAAAAAAGCTCAGTCCATCCTCGTAAATCCTCATC AATCATC
<i>E. scolopes</i> EsLBP1 Probe 5	B2/Alexa 546	CCTCGTAAATCCTCATCAATCATCCAGTAAACCGCCAAAAATGAATTGGAGAGGCACGTGTT GGAAGGCACCACAAGAGGAAATTGTGACAAAAAAGCTCAGTCCATCCTCGTAAATCCTCA TCAATCATC
<i>E. scolopes</i> HSP90 Probe 1	B5/Alexa 628	CTCACTCCCAATCTCTATCTACCCTACAAATCCAATAAAAAATCCAGATGTTTTCTGGCTGCCA TATATCCCATCGTTGAGGTGTCTCGCAGATTTTCACTTCATATCACTCACTCCCAATCTCTAT CTACCC
<i>E. scolopes</i> HSP90 Probe 2	B5/Alexa 628	CTCACTCCCAATCTCTATCTACCCTACAAATCCAATAAAAAATATCAGCATCCACTTTCTCCTT CAGGGATTTGATGATAGGGTGGTCGGGGATTTTCACTTCATATCACTCACTCCCAATCTCTAT CTACCC
<i>E. scolopes</i> HSP90 Probe 3	B5/Alexa 628	CTCACTCCCAATCTCTATCTACCCTACAAATCCAATAAAAAATCATCATCTTCCAAAGGTGGG AGTTCAGCATCAACAGAGTCTCCTGTCCATTTTCACTTCATATCACTCACTCCCAATCTCTA TCTACCC
<i>E. scolopes</i> HSP90 Probe 4	B5/Alexa 628	CTCACTCCCAATCTCTATCTACCCTACAAATCCAATAAAAAAGCTGGATCAGTTTTTGGAACT ACACCAATCCGCAATGATACGCAGCGCATTTTCACTTCATATCACTCACTCCCAATCTCTAT CTACCC
<i>E. scolopes</i> HSP90 Probe 5	B5/Alexa 628	CTCACTCCCAATCTCTATCTACCCTACAAATCCAATAAAAAATTTTTTTTCCGTTAGCAGTCAC CTCTTTCCAGACCAAGCTCTAGCCGCGCATTTTCACTTCATATCACTCACTCCCAATCTCTAT CTACCC
<i>V. fischeri</i> 16S subunit Probe 1	B3/Alexa 488	GTCCCTGCCTCTATATCTCCACTCAACTTTAACCCGTACAAAGACATCGTTTACGGCGTGGAC TACCAGGGTATCTAATCCTGTTTGTCTCTAAAAAAGTCTAATCCGTCCCTGCCTCTATATCT CCACTC
<i>V. fischeri</i> 16S subunit Probe 2	B3/Alexa 488	GTCCCTGCCTCTATATCTCCACTCAACTTTAACCCGTACAATACTTAAACGCGTTAGCTCCGAA AGCCACTCCTCAAGGGAACAACCTCCAATAAAAAAAGTCTAATCCGTCCCTGCCTCTATATCT TCCACTC
<i>V. fischeri</i> 16S subunit Probe 3	B3/Alexa 488	GTCCCTGCCTCTATATCTCCACTCAACTTTAACCCGTACAATGTGCGGGCCCCCGTCAATTCA TTTGAGTTTTAATCTTGGCACCCTACTCTAAAAAAGTCTAATCCGTCCCTGCCTCTATATCT CCACTC
<i>V. fischeri</i> 16S subunit Probe 4	B3/Alexa 488	GTCCCTGCCTCTATATCTCCACTCAACTTTAACCCGTACAAGTAGGTAAGGTTCTTCGCGTTG CATCGAATTAACCACATGCTCCACCCTAAAAAAGTCTAATCCGTCCCTGCCTCTATATCT TCCACTC
<i>V. fischeri</i> 16S subunit Probe 5	B3/Alexa 488	GTCCCTGCCTCTATATCTCCACTCAACTTTAACCCGTACAACAGCACCTGTCTCAGAGTTCCC GAAGGCACTAAGCTATCTTAGCGAATTTAAAAAAGTCTAATCCGTCCCTGCCTCTATATCT TCCACTC
<i>V. fischeri</i> 16S subunit Probe 6	B3/Alexa 488	GTCCCTGCCTCTATATCTCCACTCAACTTTAACCCGTACAAATTACGTGCTGGCAACAAGG ATAAGGGTTGCGCTCGTTGCGGGACTTAATAAAAAAAGTCTAATCCGTCCCTGCCTCTATAT CTCCACTC
<i>V. fischeri</i> 16S subunit Probe 7	B3/Alexa 488	GTCCCTGCCTCTATATCTCCACTCAACTTTAACCCGTACAAATTGTAGCAGTGTGTAGCCCT ACTCGTAAGGGCCATGATGACTTGACGTTAAAAAAGTCTAATCCGTCCCTGCCTCTATATCT TCCACTC
<i>V. fischeri</i> 16S subunit Probe 8	B3/Alexa 488	GTCCCTGCCTCTATATCTCCACTCAACTTTAACCCGTACAAACTTCATGGAGTCGAGTTGCAG ACTCCAATCCGGACTACGACGCATTTTTAAAAAAGTCTAATCCGTCCCTGCCTCTATATCT CCACTC
<i>V. fischeri</i> 16S subunit Probe 9	B3/Alexa 488	GTCCCTGCCTCTATATCTCCACTCAACTTTAACCCGTACAAGTGTGTACAAGGCCGGGAAC GTATTCACCGTAGCATCTGATCTACGATTAAAAAAAGTCTAATCCGTCCCTGCCTCTATATCT TCCACTC
<i>V. fischeri</i> 16S subunit Probe 10	B3/Alexa 488	GTCCCTGCCTCTATATCTCCACTCAACTTTAACCCGTACAATFCCCCTAGGGCTACCTTGTTA CGACTTCACCCAGTCATGAACCACAAATAAAAAAAGTCTAATCCGTCCCTGCCTCTATATCT TCCACTC

Statistics. For experiments with quantitative comparisons, with the sole exception of comparison of CFU levels (Fig. 2-3), data were log-transformed to provide for normality prior to statistical analysis. The highest [*eslbp1*] level from each treatment was removed as an outlier in the timecourse experiment (Fig. 2-4A). Comparisons between treatments were made with ANOVA (Repeated Measures ANOVA for CFUs), followed by *post hoc* pairwise comparisons with Tukey multiple comparisons of means.

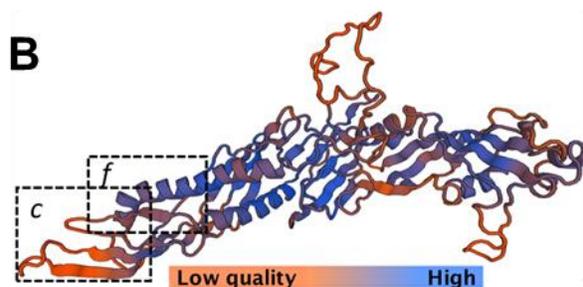
RESULTS

Structural comparison of LPS-binding domain of mammalian LBP and BPI to corresponding region of EsLBP1. We focused on the LPS-binding N-terminal domain of human LBP/BPI (hLBP/BPI) proteins, for comparison with EsLBP1 by primary sequence alignment (Fig. 2-1A). Overall conservation with human LBP/BPI proteins is limited; EsLBP1 is 23.6% identical to hBPI and 25.7% identical to hLBP within this region. In contrast to the extensive differences in overall primary structure, the polycationic region of mammalian BPI and LBP most strongly implicated in LPS binding (residues 86-102) (21-23) also represents the most cationic local region of EsLBP1 (Fig. 2-1A). The sequence alignment of human and squid proteins predicts the disulfide bond of the mammalian LBP/BPI family within the N-terminal domain (53) is conserved in EsLBP1 (Fig. 2-1A). The tertiary structure of EsLBP1, as derived by SWISS-MODEL, is most similar to the 2.40 Å structure of human BPI (SMTL ID 1BP1.1.A) (21). The QMEAN4 score, a protein-model quality measure based on scoring functions, such as torsion and solvation, related to model geometry (38) is -5.63 (Fig. 2-1B). Thus, the overall model quality is low; *i.e.*, confidence that the three-dimensional structure of BPI can predict that of EsLBP1 is limited. However, it should be noted that model quality varies greatly by region

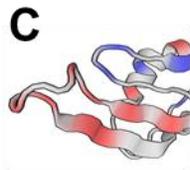
A

hBPI	--RISQKGLDYASQQGTAALQKRLKRLKIPDYSDSRKIKH	45
hLBP	--RITDRGLQYAAQGLLALQSPLLRITLDFFTGDRIPH	45
EsLBP1	RVQTTKKGLDYANGLALNRMVCLSHYQFDLMSD-----	50
hBPI	LGRGHYSFYSMIRRFQLPSSQISMVFNVGLRFSISNANI	85
hLBP	VGRGRYDFEISLNIHSCGLLHLSALRFVPGQLSLSSIS	85
EsLBP1	---GDFSLFGITISRVSHPSPGQVLLIPGQGLRWSLSINGY	87
hBPI	RISG--KWAQKRFLEMSGNFLLSIPGMSISADLLGNSP	123
hLBP	RVQG--RWRVRRSFFRLQGSFVSVKGISISVNLGSGP-	122
EsLBP1	QLSGHURYKTRVSRIRVRRKTVNIQMHINQIALDLIFTPTM	127
hBPI	-TSGKPTITCSCSSHINSVHVISKSRVGVLIQLFHKKI	162
hLBP	-SSGRPTVTASSCSSIADVFDMS-GDLGWLLNLFHNQI	160
EsLBP1	ATNGGLNLNAAHCTS-----NVDVSRVSSGNLIRKMFNDIF	162
hBPI	ESALNRMNSQVCEAVTNSVSSKLPYFQTLFVMTIDSV	202
hLBP	ESRFQVLESPICEMIQHSVSSDLQPYLQTLFVTTIDSF	200
EsLBP1	NSKINKTIGKKICDSALRIINNDFATRLANYPVTRISISIG	202
hBPI	AGINYGLVAPPATTAETLVQMKGFYSDFNH	233
hLBP	ADIDYSLVDAFATAQMLVVMFKGDFPHDNH	231
EsLBP1	YTRDYGLVSNPVFTSAYMQTNHRKAEVFNDRK	233

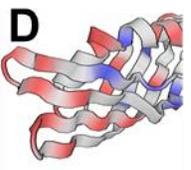
B



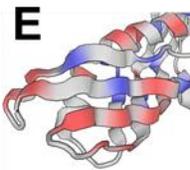
C



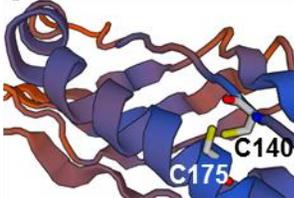
D



E



F



G

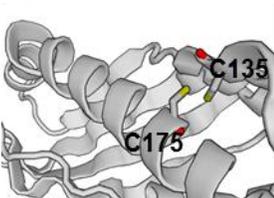


FIG 2-1 Characteristics of the EsLBP1 protein. (A) Amino acid sequence alignment of N terminal LBP/BPI domains of human BPI (hBPI), human LBP (hLBP) and EsLBP1. Cysteine residues of the conserved disulfide bond (53), yellow; negatively charged residues (Glu, Asp), blue; positively charged residues (Arg, Lys, His), red. Numbering for human LBP and BPI is based on (53). (B) SWISS-MODEL-predicted 3D-structure of EsLBP1 based on human BPI template (SMTL ID 1BP1.1.A). Color gradient indicates quality of structure determined by local QMEAN4 score. Dashed boxes, regions depicted in C and F. (C) N-terminal “tip” of EsLBP1 model from B. Charged amino acids: red, positive - Arg, Lys, His; blue, negative - Glu, Asp). (D) Corresponding N-terminal region in human BPI structure, marked for charge as in C. (E) Corresponding N-terminal region in mouse LBP (4m4d.1) marked for charge as in C. (F) Region of predicted disulfide bond in EsLBP1 structure, with the position of side chains of residues C140 and C175 indicated. Colors of backbone illustration depict QMEAN4 score as in B. (G) Region of disulfide bond in human BPI structure, with the positions of side chains of residues C135 and C175 indicated.

(Fig. 2-1B), conforming well within the extended protein regions that define the unique boomerang-like configuration of mammalian LBP and BPI (21, 54), but deviating greatly in regions within both the N- and C-terminal domains that impart interactive properties specific to LBP or BPI. For example, the highest net concentration of positive charge present in mammalian LBP and BPI, at the tip region of the N-terminal domain (Fig. 2-1D and E), is also manifest in EsLBP1 (Fig. 2-1C), but in an area of high divergence from the template. The residues predicted by ClustalW2 alignment to be involved in the conserved disulfide bond (Fig. 2-1A) are in an area of relatively good fit to the BPI template, and in close proximity to each other, as in human BPI (Fig. 2-1F-G). The choice of template structure, human BPI or mouse LBP, for the EsLBP1 model did not affect these general characteristics.

LOS/LPS-binding properties of EsLBP1. The structural comparisons described above are consistent with LPS-binding properties of EsLBP1. To test this hypothesis more directly, we assayed binding of metabolically labeled [³H] meningococcal LOS to recombinant his-tagged EsLBP1. Binding was measured by quantifying co-capture of the radiolabeled LOS to nickel beads to which His-tagged proteins bind. Co-capture of [³H]LOS was dependent on the dose of conditioned medium containing EsLBP1. Control conditioned medium lacking EsLBP1 produced significantly lower co-capture of [³H]LOS, without dose-dependency. High levels of co-capture occurred following incubations of 1 nM LOS with ~0.1-1 nM EsLBP1, indicating very high affinity binding of meningococcal LOS to the squid protein (Fig. 2-2A). Pre-incubation of EsLBP1 with comparable amounts of unlabeled *N. meningitidis* or *V. fischeri* LOS/LPS (up to 100-fold excess) prior to addition and incubation with [³H]LOS reduced co-capture of [³H]LOS of radiolabeled substrate. *Francisella tularensis* LPS, which binds human

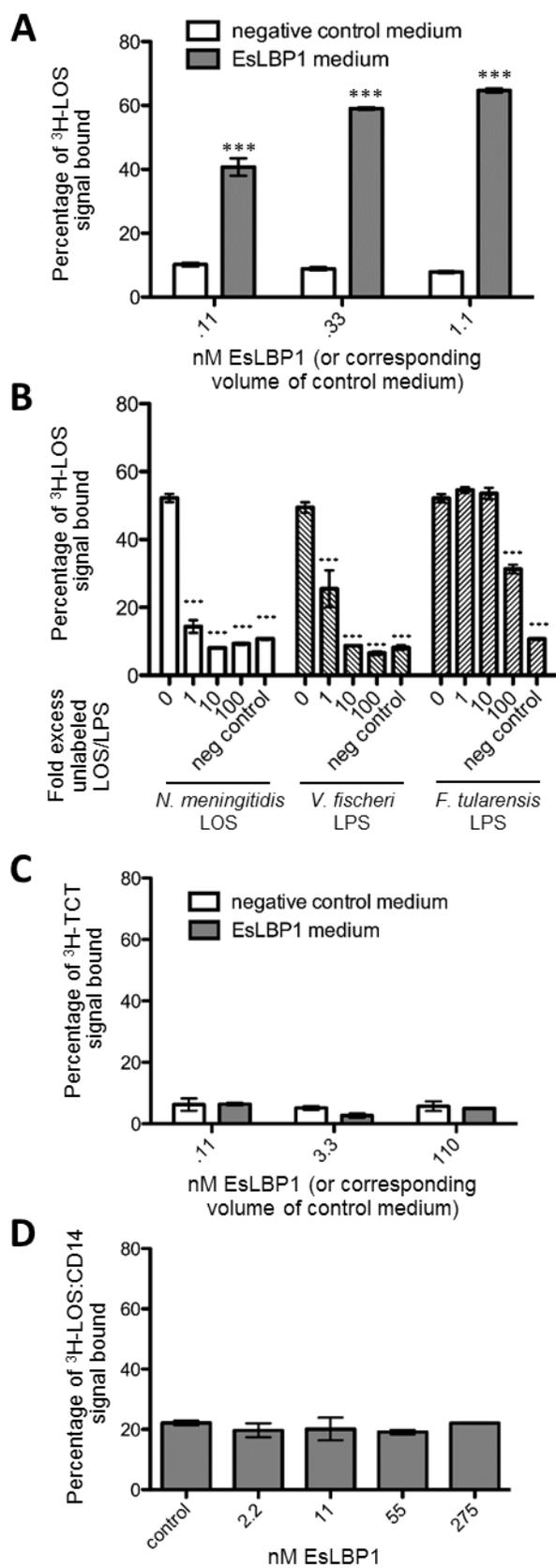


FIG 2-2 Binding of EsLBP1-His₆ to MAMP substrates. (A) Capture assay with EsLBP1-His₆ medium or control medium, incubated with 5000 cpm [³H]-*Neisseria* LOS, pelleted to nickel beads. (B) Competition experiment for 0.3 nM EsLBP1-His₆ or control medium, 5000 cpm [³H]-*Neisseria* LOS and 0-100 fold excess by weight of unlabeled *N. meningitidis* LOS, *V. fischeri* LPS or *F. tularensis* LPS. (C) As A, but measuring capture of 1000 cpm [³H]-labeled TCT. (D) As A, but measuring capture of 5000 cpm [³H]LOS:CD14 aggregates. EsLBP1-His₆ was diluted in D with control medium such that all treatments have a constant total volume of conditioned medium. Error bars indicate SEM of three technical replicates. Statistical comparisons by ANOVA and pairwise tests compare capture by EsLBP1-His₆ medium to corresponding volume of control medium (A or C) or the sole control reaction (D); in B, comparisons are to the capture assay with LBP1-His₆ medium and no competing unlabeled endotoxin (column 0). P values: *** < .001.

LBP poorly (55), only caused this decrease in capture when added in 100-fold excess of the [³H]LOS (Fig. 2-2B).

As mammalian LBP has a role in binding pneumococcal cell wall fragments (56), we sought to determine whether EsLBP1 binds the peptidoglycan monomer, TCT. Additionally, human LBP reacts with monomeric endotoxin:CD14 complexes to form supra-molecular complexes containing LBP, endotoxin, and CD14 (J. Weiss unpublished data). We adapted the LOS/LPS capture assay described above to measure binding to [³H]TCT and to [³H]LOS:human CD14 complexes, but in neither case was EsLBP1-dependent co-capture observed (Fig. 2-2C-D). Thus, we found no evidence for EsLBP1 binding of TCT or LOS:CD14 complexes.

EsLBP1 does not kill *Escherichia coli*. The pI of EsLBP1 suggests that it more likely functions like LBP and not BPI. To test this hypothesis more directly, we assayed the effect of EsLBP1 and, for comparison, recombinant human LBP and BPI-21 (the LPS-binding and bactericidal N-terminal M_r 21,000 fragment of human BPI (57)) on the viability of *E. coli*. As shown before (15, 16), BPI, but not LBP, produced killing of *E. coli*, as manifested by reduced colony-forming units (CFU) (Fig. 2-3). At the same protein concentration tested, EsLBP1 had no effect on bacterial viability, resembling mammalian LBP.

***eslbp1* transcript expression is induced by symbiosis and TCT, but not by the binding partner, LPS.** Previous microarray results (30) showed that EsLBP1 is up-regulated at 18 h following colonization of *E. scolopes* by *V. fischeri*, *i.e.*, following first full colonization of host crypts. Here we used qRT-PCR with gene-specific primer sets (Table 2-1) to establish the time course of *eslbp1* expression over the trajectory of early development. We also investigated the role of symbiont MAMPs in the induction of *eslbp1*. A reproducible, significant ~3-fold difference in *eslbp1* expression between apo- and symbiotic light organs was first observed at 12

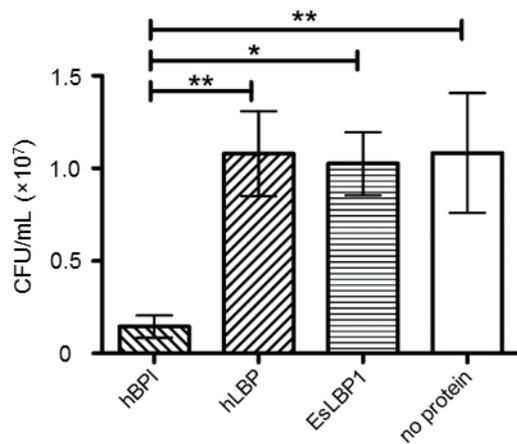


FIG 2-3 Effect of EsLBP1 on viability of *E. coli*. Shown are CFUs of PL2 *E. coli* after treatment of log-phase culture for 1h with 30 nM hBPI-21 (N-terminal fragment of human BPI), hLBP (human LBP) or EsLBP1, or with no protein. Error bars are SEM of 5 matched biological replicates. P values of comparisons by ANOVA and pairwise tests: ** < .01 < * < .05.

h post-inoculation; upregulation in sym- over aposymbiotic light organs at 24 h varied, but was typically greater than 20-fold (Fig. 2-4A). Lipid A, the LPS derivative involved in symbiont-induced development, from either *V. fischeri* or *E. coli* was used at 10 ng/mL, a level that optimally induces host cell phenotypes, but did not induce changes in *eslbp1* expression. However, TCT added at 1 μ M was shown to increase *eslbp1* ~11-fold over background; the TCT effect was not further amplified by lipid A addition (Fig. 2-4B-C).

***eslbp1* transcript localizes to light organ tissues interfacing with *V. fischeri*.** Using hybridization chain reaction fluorescent *in situ* hybridization (HCR-FISH), we located expression of *eslbp1* within the light organ to determine whether the gene is expressed in proximity to *V. fischeri*, and to ascertain the steps of the establishment of symbiosis with which EsLBP1 may play a role. We evaluated the location of transcript in 24 h post-inoculation light organs using gene-specific probes for *eslbp1*, for *hsp90* (as a counterstain for *E. scolopes* tissue), and, to label the symbionts, for *V. fischeri* 16S ribosomal subunit (Table 2-2). *eslbp1* transcript signal was elevated in symbiotic light organs, relative to those of non-symbiotic animals, throughout epithelial surfaces closely associated with *V. fischeri*, including the pores, ducts, and crypts of the organ (Fig. 2-5).

EsLBP1 protein is present at epithelial surfaces that directly associate with bacteria and with the environment. We compared the data on the location of *eslbp1* transcript with the location of protein. Using chicken anti-EsLBP1 antibody, we confirmed and expanded upon previously reported EsLBP1 localization at 18 h post-inoculation in symbiotic animals (30). EsLBP1 was highly abundant throughout the light organ (Fig. 2-6B). The antibody cross-reactivity was high in anterior appendage epithelial cells, as well as the apical surfaces of pore and duct cells. EsLBP1 protein was also abundant in the extracellular crypt spaces of symbiotic

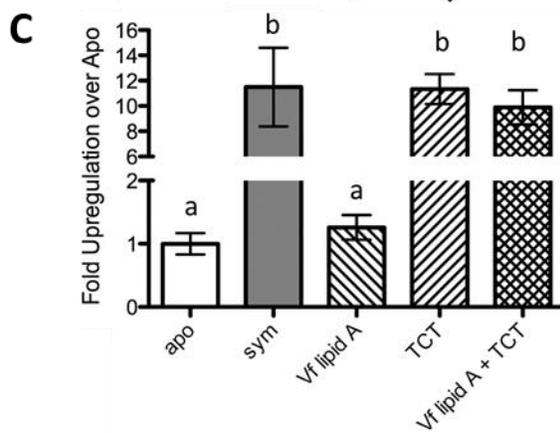
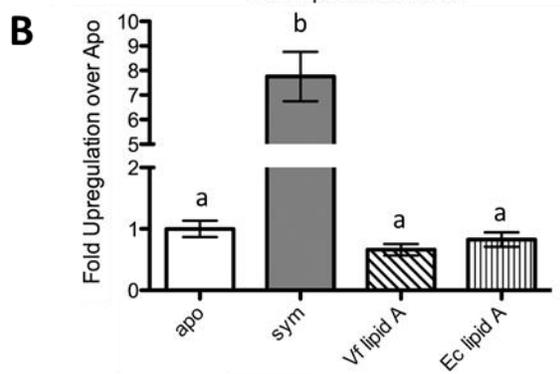
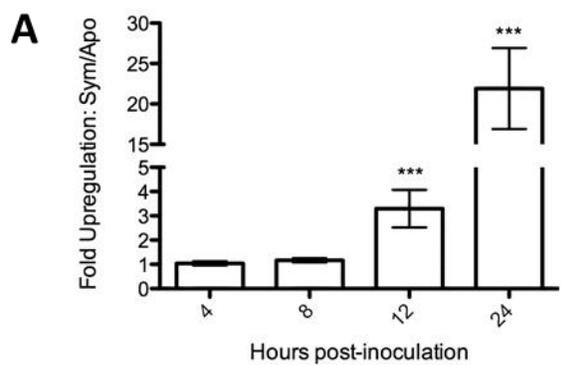


FIG 2-4 *eslbp1* transcript induction by *V. fischeri* colonization or MAMPs. (A) *eslbp1* normalized transcript levels in *E. scolopes* light organs at 4, 8, 12, and 24 h post-inoculation are represented as a ratio of transcript levels in symbiotic light organs to those in control, aposymbiotic light organs. (B) Light organ *eslbp1* normalized transcript levels are shown for 24 h apo- and symbiotic animals as above, along with animals treated with lipid A (10 ng/mL) from *V. fischeri* or *E. coli* for 24 h, but no bacteria. (C) Light organ *eslbp1* transcript levels are shown for 24h apo- and symbiotic animals as above, along with animals treated with lipid A (10 ng/mL) from *V. fischeri* and/or TCT treatment (1 μ M) for 24h. In A, *** indicates $p < .001$ for comparison of symbiotic to corresponding aposymbiotic treatment. In B-D, small letters indicate statistically indistinguishable groups of treatments, determined by ANOVA and pairwise comparison ($p < .01$ in all cases). Error bars show SEM of four replicates, except three for (A).

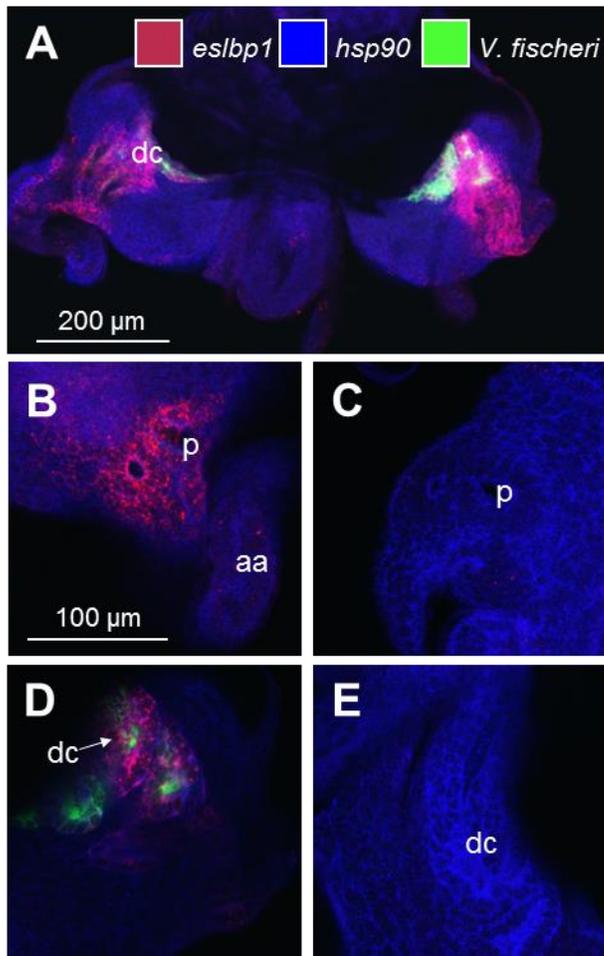


FIG 2-5 Localization of *eslbp1* transcript in 24 h light organs. (A) An overview of 24 h symbiotic light organ. (B) Micrograph of pore/anterior appendage region of 24 h symbiotic light organ. (C) Pore/anterior appendage region of 24 h aposymbiotic light organ. (D) Deep crypt region of 24 h symbiotic light organ. (E) Deep crypt region of 24 h aposymbiotic light organ. aa, anterior appendage; dc, deep crypt; p, pore. Red, *eslbp1* HCR probe; blue, *hsp90*; green, 16S ribosomal subunit.

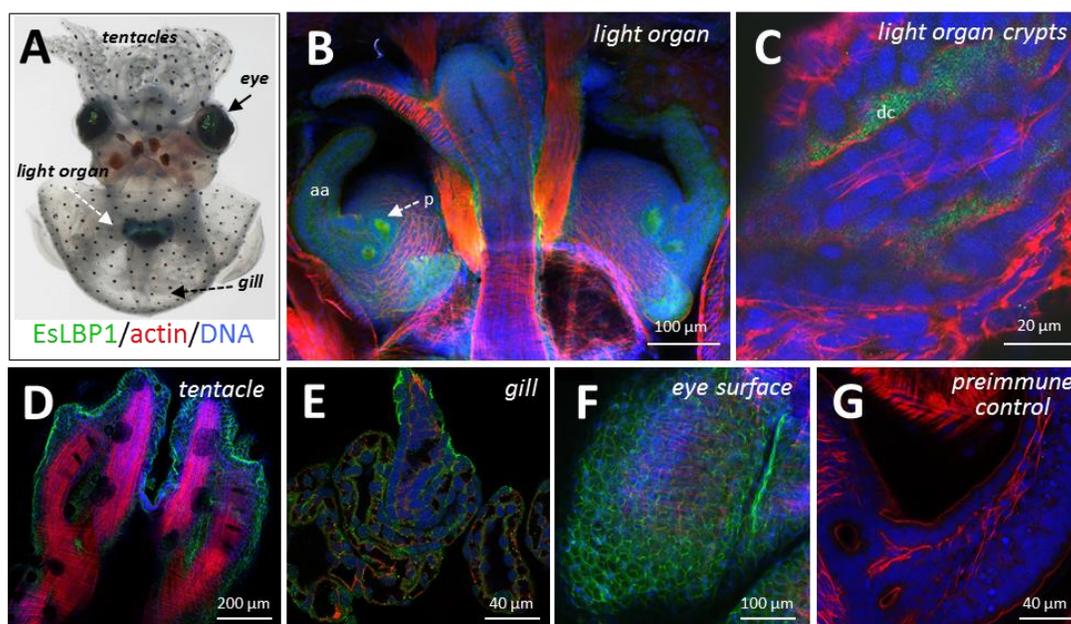


FIG 2-6 Localization of EsLBP1 protein in 18 h light organs. Juvenile *E. scolopes* were colonized with *V. fischeri* for 18 h and probed with chicken anti- EsLBP1 antibody. (A) *E. scolopes* hatchling under white light, ventral view. Labeling is provided for tissues shown in panels B-G. Light organ is transparent structure immediately ventral to ink sac. (B) Immunocytochemistry micrograph overview of an 18 h symbiotic light organ and closely associated viscera, highlighting anti-EsLBP1 staining. p, pore; aa, anterior appendage. (C) Higher-magnification micrograph of deep crypt spaces of symbiotic light organ. dc, deep crypt lumen. (D-F) High-magnification micrographs of non-light organ tissues: tentacles, gills, and outer surface of eye, respectively. (G) Pre-immune IgY control for EsLBP1 antibody at anterior appendage. Green, anti-LBP1 or pre-immune control (detected by goat anti-chicken FITC secondary antibody); red, actin (rhodamine-phalloidin); blue, DNA (TOTO-3).

light organs, the site of long-term colonization by *V. fischeri* (Fig. 2-6C).

To determine whether EsLBP1 protein is specific to the light organ, we also analyzed other epithelial tissues of the squid that interact with environmental bacteria, including tentacles, gills and eyes (Fig. 2-6D-F). We found that the protein is not exclusive to the light organ, but is abundant along the apical surfaces of many, or perhaps all, epithelia, but absent in deeper tissues such as muscle. Signal was not seen in light organs or other tissues treated with pre-immune serum (Fig. 2-6G).

DISCUSSION

In this study, we provide both structural and functional evidence consistent with EsLBP1 functioning as an LBP-like protein. Most notably, EsLBP1 binds Gram-negative bacterial LOS and LPS with nM or higher avidity under *in vitro* conditions, *i.e.*, when LPS/LOS is presented as part of supra-molecular assemblies containing LPS-rich lipid-water interfaces, as in aggregates of purified LPS/LOS. *eslbp1* gene expression is regulated by exposure to the peptidoglycan monomer TCT, which synergizes with LPS in the triggering of *V. fischeri*-induced morphogenesis of the host symbiotic tissues. The gene is expressed and the protein produced across the organ epithelia - from the point where *V. fischeri* initially gathers, along the path of its migration, to where it takes up permanent residence in the crypts. The protein is also abundant along the apical surfaces of other epithelial tissues, where colonization by bacteria does not occur.

Although EsLBP1 has only ~25% primary structure identity with mammalian LBPs in the N-terminal domain (Fig. 2-1A), and ~21% overall, several key structural features likely important for LPS-related functions are predicted to be conserved. These features include: 1) the

localization of the most cationic region of these proteins within the predicted outer “tip” of the N-terminal domain of these proteins (Fig. 2-1C-E), albeit with substitutions in some specific cationic residues conserved in mammalian LBP and BPI; 2) the extended “boomerang” configuration of the holo-proteins, possibly important for the coordinated binding of LPS-rich supra-molecular assemblies by the N-terminal domain, and delivery of LPS to specific host targets mediated by the C-terminal domain; and 3) the presence of conserved cysteines (140 and 175) within the N-terminal domain that form a disulfide linkage necessary for protein stability and function (58). Substitutions of specific cationic residues conserved in mammalian LBP and BPI are common in proteins of this family in many non-mammalian vertebrates and invertebrates, including the Atlantic cod *Gadus morhua* and mollusks *Crassostrea gigas* (oyster) and *Biomphalaria glabrata* (snail) (27, 59, 60), reinforcing the notion that overall charge of the region is critical. Remarkably, the much lower avidity of mammalian LBP and BPI for *F. tularensis* LPS compared to meningococcal LOS (55) is also manifest with EsLBP1 (Fig. 2-2B). These data suggest similar determinants of LPS interaction with the mammalian and squid proteins. In addition, they are consistent with a key role of electrostatic interactions between cationic regions of the proteins and anionic moieties at or closely neighboring the lipid A region of most LPS/LOS species that are less prominent in *F. tularensis* LPS (61). It is noteworthy that the quality of the 3D-model of EsLBP1 is strongest within the core of the boomerang structure that is unique to the LBP/BPI family. Regions of EsLBP1 that diverge from the human BPI 3D structure template are generally at the periphery of the molecule. Divergence at the N-terminus could reflect greater intrinsic mobility of the LPS-binding region to optimize interactions with diverse LPS-rich bacterial interfaces, and divergence at the C-terminus may correspond to functional regions that have evolved to recognize different (*e.g.*, host) targets.

The regulation of *lbp* gene expression by TCT and not LPS was unexpected, given that, in mammalian systems, LPS increases *lbp* expression (62, 63); we are not aware of experiments showing the effect of peptidoglycan treatment on mammalian *lbp* transcript or protein levels. Recent studies have indicated that LPS preparations are often contaminated with trace amounts of peptidoglycan, which can actually be the element that is active (64). In light of results in the squid system, a revisit of the MAMPs induction of genes encoding mammalian LBP/BPI may be fruitful. Regulation of an LBP by TCT is of added interest as it is the converse of regulation of expression of a well studied C-type lectin, RegIII γ . This protein is expressed in the intestinal epithelium of mice, binds peptidoglycan, and is preferentially bactericidal against Gram-positive bacteria (65, 66). However, the expression of the gene encoding RegIII γ is induced by the presence of Gram-negative bacteria or LPS (66, 67).

Crosstalk and synergy between MAMPs and their receptors have been observed elsewhere in mammalian systems. TCT and lipid A act synergistically in the case of nitric oxide synthase induction during *Bordetella pertussis* infection of hamster tracheal cells (68, 69); in this case, TCT and lipid A are thought to deliver their signal through parallel pathways. In contrast, a sequential “priming” effect of peptidoglycan on the LPS response has been noted in human blood; administration of staphylococcal peptidoglycan increases the response to subsequent LPS treatment, apparently through up-regulation of such factors as CD14 and TLR4 on monocytes (70). Our experimental results suggest a similar mechanistic relationship in the squid, wherein increased *eslbp1* gene expression induced by TCT promotes developmental responses induced by LPS and thus results in the LPS/peptidoglycan synergy in squid vibrio symbiosis that has been described (4). This synergy may be particularly important given the relatively low potency of the *V. fischeri* LPS (42), presumably related to its unusual structure (71, 72). TCT’s induction

of *eslbp1* is another example of peptidoglycan products from beneficial bacteria influencing host immune development. In mammals, the diverse effects attributed to the peptidoglycan of the microbiota include maturation of the gut-associated lymphoid tissues (reviewed in (73)).

Although the irreversible signal for light organ morphogenesis is delivered by MAMPs/host-cell interactions in the crypt spaces, cell death does not occur in these epithelia. Instead, at ~12 h, the MAMPS remotely trigger an irreversible program of cell death that results in a 4-5 day regression of the superficial ciliated epithelia, which are several cell layers away from the crypts (74). *eslbp1* is significantly up-regulated at this 12 h point (Fig. 2-4A) and thus its expression is plausibly timed to play a role in the transmission of this signal. Furthermore, the crypt epithelia have an abundance of this protein (Fig. 2-6C), a finding reminiscent of the observation of LBP in the mucus of mouse intestines (8).

Our studies of EsLBP1 show that it is not exclusively a light organ protein, but rather is abundant in most, if not all, epithelia (Fig.2-6D-F). These data suggest that the light organ has recruited LBP as protein to signal the presence of a mutualistic partner, or control its population, rather than to respond to a pathogen. The hypothesized signaling role for EsLBP1 is unproven, but is supported by EsLBP1's structural and functional properties that resemble more closely mammalian LBP rather than BPI. These features include its predicted near-neutral isoelectric point (27), and absence of bactericidal activity (Fig. 2-3) that are distinguishing features of mammalian LBP compared to BPI (11).

In mammals, LBP can either promote LPS-triggered inflammation or blunt it by promoting non-inflammatory clearance mechanisms (5, 10-13, 75). The former seems entirely CD14-dependent whereas the latter is largely CD14-independent. No CD14 has been detected among *E. scolopes* transcripts, though transcripts with a MD-2-related lipid-recognition (ML)

domain have been noted (29). The inability of EsLBP1, unlike mammalian LBP, to interact with monomeric LOS:sCD14 complexes (Fig. 2-2D) leaves open the possibility of an alternative target of EsLBP1:LPS complexes. In mammals, up to 100-fold increases in extracellular LBP levels can promote “silent uptake” of LPS rather than activation of inflammation (9, 10, 13). Thus, future studies in the squid will also need to investigate whether induced increases in EsLBP1 levels also have more complex effects on the evolution of the symbiotic response and relationship.

The presence of multiple proteins in the LBP/BPI family in *E. scolopes* predicted to have varying biochemical properties (28) raises the possibility that the squid, like mammals, express both LBP-like and BPI-like proteins whose expression and functions are differentiated in a way to best coordinate host responses to Gram-negative bacterial interaction, in this case leading to *E. scolopes-V. fischeri* symbiosis. If so, this system would provide the first example outside of mammals of the coexistence of both LBP- and BPI-like family members in a single species. Future studies of this family lipid (LPS)-binding proteins during *E. scolopes-V. fischeri* symbiosis should advance understanding of both the evolution and structure/function of the LBP/BPI family and its role in mutualism.

ACKNOWLEDGEMENTS

We thank T. Gioannini, A. Teghanemt, K. Nikolakakis and E. Ruby for technical assistance and advice on experimental design and on the manuscript. This work was supported by grants from the National Institutes of Health, R01 AI050661 and OD11024 to M.J.M.-N., R01 AI097157 to J.P.D., R01 AI059372 and VA Merit Award to J.P.W., T32 AI007414 to the University of Wisconsin Cellular and Molecular Parasitology Training Program, T32 AI55397 to U. of Wisc.

Microbes in Health and Disease Training Program and T32 GM008692 to U. of Wisc. Medical Scientist Training Program.

REFERENCES

1. **Patten DA, Collett A.** 2013. Exploring the immunomodulatory potential of microbial-associated molecular patterns derived from the enteric bacterial microbiota. *Microbiology* **159**:1535-1544.
2. **Eberl G, Boneca IG.** 2010. Bacteria and MAMP-induced morphogenesis of the immune system. *Curr Opin Immunol* **22**:448-454.
3. **McFall-Ngai M, Nyholm SV, Castillo MG.** 2010. The role of the immune system in the initiation and persistence of the *Euprymna scolopes*--*Vibrio fischeri* symbiosis. *Semin Immunol* **22**:48-53.
4. **McFall-Ngai MJ.** 2014. The importance of microbes in animal development: lessons from the squid-vibrio symbiosis. *Annu Rev Microbiol* **68**:177-194.
5. **Zweigner J, Schumann RR, Weber JR.** 2006. The role of lipopolysaccharide-binding protein in modulating the innate immune response. *Microbes Infect* **8**:946-952.
6. **Schultz H, Weiss JP.** 2007. The bactericidal/permeability-increasing protein (BPI) in infection and inflammatory disease. *Clin Chim Acta* **384**:12-23.
7. **Levy O, Canny G, Serhan CN, Colgan SP.** 2003. Expression of BPI (bactericidal/permeability-increasing protein) in human mucosal epithelia. *Biochem Soc Trans* **31**:795-800.
8. **Vreugdenhil AC, Snoek AM, Greve JW, Buurman WA.** 2000. Lipopolysaccharide-binding protein is vectorially secreted and transported by cultured intestinal epithelial cells and is present in the intestinal mucus of mice. *J Immunol* **165**:4561-4566.
9. **Hailman E, Vasselon T, Kelley M, Busse LA, Hu MC, Lichenstein HS, Detmers PA, Wright SD.** 1996. Stimulation of macrophages and neutrophils by complexes of lipopolysaccharide and soluble CD14. *J Immunol* **156**:4384-4390.
10. **Kitchens RL, Thompson PA.** 2005. Modulatory effects of sCD14 and LBP on LPS-host cell interactions. *J Endotoxin Res* **11**:225-229.
11. **Weiss J.** 2003. Bactericidal/permeability-increasing protein (BPI) and lipopolysaccharide-binding protein (LBP): structure, function and regulation in host defence against Gram-negative bacteria. *Biochem Soc Trans* **31**:785-790.
12. **Gioannini TL, Teghanemt A, Zhang D, Prohinar P, Levis EN, Munford RS, Weiss JP.** 2007. Endotoxin-binding proteins modulate the susceptibility of bacterial endotoxin to deacylation by acyloxyacyl hydrolase. *J Biol Chem* **282**:7877-7884.

13. **Hamann L, Alexander C, Stamme C, Zähringer U, Schumann RR.** 2005. Acute-phase concentrations of lipopolysaccharide (LPS)-binding protein inhibit innate immune cell activation by different LPS chemotypes via different mechanisms. *Infect Immun* **73**:193-200.
14. **Dunzendorfer S, Lee HK, Soldau K, Tobias PS.** 2004. TLR4 is the signaling but not the lipopolysaccharide uptake receptor. *J Immunol* **173**:1166-1170.
15. **Prohinar P.** 2003. Responses of *Escherichia coli* to the neutrophil bactericidal/permeability-increasing protein. PhD. University of Iowa.
16. **Iovine NM, Elsbach P, Weiss J.** 1997. An opsonic function of the neutrophil bactericidal/permeability-increasing protein depends on both its N- and C-terminal domains. *Proc Natl Acad Sci U S A* **94**:10973-10978.
17. **Han J, Mathison JC, Ulevitch RJ, Tobias PS.** 1994. Lipopolysaccharide (LPS) binding protein, truncated at Ile-197, binds LPS but does not transfer LPS to CD14. *J Biol Chem* **269**:8172-8175.
18. **Gazzano-Santoro H, Parent JB, Conlon PJ, Kasler HG, Tsai CM, Lill-Elghanian DA, Hollingsworth RI.** 1995. Characterization of the structural elements in lipid A required for binding of a recombinant fragment of bactericidal/permeability-increasing protein rBPI23. *Infect Immun* **63**:2201-2205.
19. **Theofan G, Horwitz AH, Williams RE, Liu PS, Chan I, Birr C, Carroll SF, Mészáros K, Parent JB, Kasler H.** 1994. An amino-terminal fragment of human lipopolysaccharide-binding protein retains lipid A binding but not CD14-stimulatory activity. *J Immunol* **152**:3623-3629.
20. **Iovine N, Eastvold J, Elsbach P, Weiss JP, Gioannini TL.** 2002. The carboxyl-terminal domain of closely related endotoxin-binding proteins determines the target of protein-lipopolysaccharide complexes. *J Biol Chem* **277**:7970-7978.
21. **Beamer LJ, Carroll SF, Eisenberg D.** 1997. Crystal structure of human BPI and two bound phospholipids at 2.4 angstrom resolution. *Science* **276**:1861-1864.
22. **Lamping N, Hoess A, Yu B, Park TC, Kirschning CJ, Pfeil D, Reuter D, Wright SD, Herrmann F, Schumann RR.** 1996. Effects of site-directed mutagenesis of basic residues (Arg 94, Lys 95, Lys 99) of lipopolysaccharide (LPS)-binding protein on binding and transfer of LPS and subsequent immune cell activation. *J Immunol* **157**:4648-4656.
23. **Reyes O, Vallespi MG, Garay HE, Cruz LJ, González LJ, Chinaea G, Buurman W, Araña MJ.** 2002. Identification of single amino acid residues essential for the binding of lipopolysaccharide (LPS) to LPS binding protein (LBP) residues 86-99 by using an Ala-scanning library. *J Pept Sci* **8**:144-150.
24. **Weiss J, Victor M, Elsbach P.** 1983. Role of charge and hydrophobic interactions in the action of the bactericidal/permeability-increasing protein of neutrophils on gram-negative bacteria. *J Clin Invest* **71**:540-549.

25. **Tobias PS, Soldau K, Iovine NM, Elsbach P, Weiss J.** 1997. Lipopolysaccharide (LPS)-binding proteins BPI and LBP form different types of complexes with LPS. *J Biol Chem* **272**:18682-18685.
26. **Inagawa H, Honda T, Kohchi C, Nishizawa T, Yoshiura Y, Nakanishi T, Yokomizo Y, Soma G.** 2002. Cloning and characterization of the homolog of mammalian lipopolysaccharide-binding protein and bactericidal permeability-increasing protein in rainbow trout *Oncorhynchus mykiss*. *J Immunol* **168**:5638-5644.
27. **Krasity BC, Troll JV, Weiss JP, McFall-Ngai MJ.** 2011. LBP/BPI proteins and their relatives: conservation over evolution and roles in mutualism. *Biochem Soc Trans* **39**:1039-1044.
28. **Beamer LJ, Fischer D, Eisenberg D.** 1998. Detecting distant relatives of mammalian LPS-binding and lipid transport proteins. *Protein Sci* **7**:1643-1646.
29. **Kremer N, Philipp EE, Carpentier MC, Brennan CA, Kraemer L, Altura MA, Augustin R, Häsler R, Heath-Heckman EA, Peyer SM, Schwartzman J, Rader BA, Ruby EG, Rosenstiel P, McFall-Ngai MJ.** 2013. Initial symbiont contact orchestrates host-organ-wide transcriptional changes that prime tissue colonization. *Cell Host Microbe* **14**:183-194.
30. **Chun CK, Troll JV, Koroleva I, Brown B, Manzella L, Snir E, Almabrazi H, Scheetz TE, Bonaldo MeF, Casavant TL, Soares MB, Ruby EG, McFall-Ngai MJ.** 2008. Effects of colonization, luminescence, and autoinducer on host transcription during development of the squid-vibrio association. *Proc Natl Acad Sci U S A* **105**:11323-11328.
31. **Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA, McWilliam H, Valentin F, Wallace IM, Wilm A, Lopez R, Thompson JD, Gibson TJ, Higgins DG.** 2007. Clustal W and Clustal X version 2.0. *Bioinformatics* **23**:2947-2948.
32. **Goujon M, McWilliam H, Li W, Valentin F, Squizzato S, Paern J, Lopez R.** 2010. A new bioinformatics analysis tools framework at EMBL-EBI. *Nucleic Acids Res* **38**:W695-699.
33. **Letunic I, Doerks T, Bork P.** 2012. SMART 7: recent updates to the protein domain annotation resource. *Nucleic Acids Res* **40**:D302-305.
34. **Biasini M, Bienert S, Waterhouse A, Arnold K, Studer G, Schmidt T, Kiefer F, Cassarino TG, Bertoni M, Bordoli L, Schwede T.** 2014. SWISS-MODEL: modelling protein tertiary and quaternary structure using evolutionary information. *Nucleic Acids Res* **42**:W252-258.
35. **Guex N, Peitsch MC, Schwede T.** 2009. Automated comparative protein structure modeling with SWISS-MODEL and Swiss-PdbViewer: a historical perspective. *Electrophoresis* **30 Suppl 1**:S162-173.
36. **Kiefer F, Arnold K, Künzli M, Bordoli L, Schwede T.** 2009. The SWISS-MODEL Repository and associated resources. *Nucleic Acids Res* **37**:D387-392.

37. **Arnold K, Bordoli L, Kopp J, Schwede T.** 2006. The SWISS-MODEL workspace: a web-based environment for protein structure homology modelling. *Bioinformatics* **22**:195-201.
38. **Benkert P, Biasini M, Schwede T.** 2011. Toward the estimation of the absolute quality of individual protein structure models. *Bioinformatics* **27**:343-350.
39. **Giardina PC, Giannini T, Buscher BA, Zaleski A, Zheng DS, Stoll L, Teghanemt A, Apicella MA, Weiss J.** 2001. Construction of acetate auxotrophs of *Neisseria meningitidis* to study host-meningococcal endotoxin interactions. *J Biol Chem* **276**:5883-5891.
40. **Post DM, Zhang D, Weiss JP, Gibson BW.** 2006. Stable isotope metabolic labeling of *Neisseria meningitidis* lipooligosaccharide. *J Endotoxin Res* **12**:93-98.
41. **Apicella MA, Griffiss JM, Schneider H.** 1994. Isolation and characterization of lipopolysaccharides, lipooligosaccharides, and lipid A. *Methods Enzymol* **235**:242-252.
42. **Foster JS, Apicella MA, McFall-Ngai MJ.** 2000. *Vibrio fischeri* lipopolysaccharide induces developmental apoptosis, but not complete morphogenesis, of the *Euprymna scolopes* symbiotic light organ. *Dev Biol* **226**:242-254.
43. **Cookson BT, Cho HL, Herwaldt LA, Goldman WE.** 1989. Biological activities and chemical composition of purified tracheal cytotoxin of *Bordetella pertussis*. *Infect Immun* **57**:2223-2229.
44. **Cloud KA, Dillard JP.** 2002. A lytic transglycosylase of *Neisseria gonorrhoeae* is involved in peptidoglycan-derived cytotoxin production. *Infect Immun* **70**:2752-2757.
45. **Kohler PL, Hamilton HL, Cloud-Hansen K, Dillard JP.** 2007. AtlA functions as a peptidoglycan lytic transglycosylase in the *Neisseria gonorrhoeae* type IV secretion system. *J Bacteriol* **189**:5421-5428.
46. **McFall-Ngai M, Montgomery MK.** 1990. The Anatomy and Morphology of the Adult Bacterial Light Organ of *Euprymna scolopes* Berry (Cephalopoda:Sepiolidae). *Biological Bulletin* **179**:332-339.
47. **Boettcher KJ, Ruby EG.** 1990. Depressed light emission by symbiotic *Vibrio fischeri* of the sepiolid squid *Euprymna scolopes*. *J Bacteriol* **172**:3701-3706.
48. **Bustin SA, Benes V, Garson JA, Hellemans J, Huggett J, Kubista M, Mueller R, Nolan T, Pfaffl MW, Shipley GL, Vandesompele J, Wittwer CT.** 2009. The MIQE guidelines: minimum information for publication of quantitative real-time PCR experiments. *Clin Chem* **55**:611-622.
49. **Heath-Heckman EA, Peyer SM, Whistler CA, Apicella MA, Goldman WE, McFall-Ngai MJ.** 2013. Bacterial bioluminescence regulates expression of a host cryptochrome gene in the squid-*Vibrio* symbiosis. *MBio* **4**.
50. **Pfaffl MW.** 2001. A new mathematical model for relative quantification in real-time RT-PCR. *Nucleic Acids Res* **29**:e45.

51. **Nikolakakis K, Lehnert EM, McFall-Ngai MJ, Ruby EG.** Using hybridization chain-reaction fluorescent in situ hybridization (HCR-FISH) to track gene expression by both partners during initiation of symbiosis. *Applied and Environmental Microbiology*:In Press.
52. **Choi HM, Beck VA, Pierce NA.** 2014. Next-generation in situ hybridization chain reaction: higher gain, lower cost, greater durability. *ACS Nano* **8**:4284-4294.
53. **Beamer LJ, Carroll SF, Eisenberg D.** 1998. The BPI/LBP family of proteins: a structural analysis of conserved regions. *Protein Sci* **7**:906-914.
54. **Eckert JK, Kim YJ, Kim JI, Gürtler K, Oh DY, Sur S, Lundvall L, Hamann L, van der Ploeg A, Pickkers P, Giamarellos-Bourboulis E, Kubarenko AV, Weber AN, Kabesch M, Kumpf O, An HJ, Lee JO, Schumann RR.** 2013. The crystal structure of lipopolysaccharide binding protein reveals the location of a frequent mutation that impairs innate immunity. *Immunity* **39**:647-660.
55. **Barker JH, Weiss J, Apicella MA, Nauseef WM.** 2006. Basis for the failure of *Francisella tularensis* lipopolysaccharide to prime human polymorphonuclear leukocytes. *Infect Immun* **74**:3277-3284.
56. **Weber JR, Freyer D, Alexander C, Schröder NW, Reiss A, Küster C, Pfeil D, Tuomanen EI, Schumann RR.** 2003. Recognition of pneumococcal peptidoglycan: an expanded, pivotal role for LPS binding protein. *Immunity* **19**:269-279.
57. **Giroir BP, Quint PA, Barton P, Kirsch EA, Kitchen L, Goldstein B, Nelson BJ, Wedel NJ, Carroll SF, Scannon PJ.** 1997. Preliminary evaluation of recombinant amino-terminal fragment of human bactericidal/permeability-increasing protein in children with severe meningococcal sepsis. *Lancet* **350**:1439-1443.
58. **Capodici C, Weiss J.** 1996. Both N- and C-terminal regions of the bioactive N-terminal fragment of the neutrophil granule bactericidal/permeability-increasing protein are required for stability and function. *J Immunol* **156**:4789-4796.
59. **Stenvik J, Solstad T, Strand C, Leiros I, Jørgensen T T.** 2004. Cloning and analyses of a BPI/LBP cDNA of the Atlantic cod (*Gadus morhua* L.). *Dev Comp Immunol* **28**:307-323.
60. **Gonzalez M, Gueguen Y, Destoumieux-Garzón D, Romestand B, Fievet J, Pugnère M, Roquet F, Escoubas JM, Vandenbulcke F, Levy O, Sauné L, Bulet P, Bachère E.** 2007. Evidence of a bactericidal permeability increasing protein in an invertebrate, the *Crassostrea gigas* Cg-BPI. *Proc Natl Acad Sci U S A* **104**:17759-17764.
61. **Barker JH, Kaufman JW, Zhang DS, Weiss JP.** 2014. Metabolic labeling to characterize the overall composition of *Francisella* lipid A and LPS grown in broth and in human phagocytes. *Innate Immun* **20**:88-103.
62. **Geller DA, Kispert PH, Su GL, Wang SC, Di Silvio M, Tweardy DJ, Billiar TR, Simmons RL.** 1993. Induction of hepatocyte lipopolysaccharide binding protein in models of sepsis and the acute-phase response. *Arch Surg* **128**:22-27; discussion 27-28.

63. **Wan Y, Freeswick PD, Khemlani LS, Kispert PH, Wang SC, Su GL, Billiar TR.** 1995. Role of lipopolysaccharide (LPS), interleukin-1, interleukin-6, tumor necrosis factor, and dexamethasone in regulation of LPS-binding protein expression in normal hepatocytes and hepatocytes from LPS-treated rats. *Infect Immun* **63**:2435-2442.
64. **Kaneko T, Goldman WE, Mellroth P, Steiner H, Fukase K, Kusumoto S, Harley W, Fox A, Golenbock D, Silverman N.** 2004. Monomeric and polymeric gram-negative peptidoglycan but not purified LPS stimulate the *Drosophila* IMD pathway. *Immunity* **20**:637-649.
65. **Vaishnava S, Yamamoto M, Severson KM, Ruhn KA, Yu X, Koren O, Ley R, Wakeland EK, Hooper LV.** 2011. The antibacterial lectin RegIIIgamma promotes the spatial segregation of microbiota and host in the intestine. *Science* **334**:255-258.
66. **Cash HL, Whitham CV, Behrendt CL, Hooper LV.** 2006. Symbiotic bacteria direct expression of an intestinal bactericidal lectin. *Science* **313**:1126-1130.
67. **Brandl K, Plitas G, Mihu CN, Ubeda C, Jia T, Fleisher M, Schnabl B, DeMatteo RP, Pamer EG.** 2008. Vancomycin-resistant enterococci exploit antibiotic-induced innate immune deficits. *Nature* **455**:804-807.
68. **Flak TA, Heiss LN, Engle JT, Goldman WE.** 2000. Synergistic epithelial responses to endotoxin and a naturally occurring muramyl peptide. *Infect Immun* **68**:1235-1242.
69. **Flak TA, Goldman WE.** 1999. Signalling and cellular specificity of airway nitric oxide production in pertussis. *Cell Microbiol* **1**:51-60.
70. **Hadley JS, Wang JE, Foster SJ, Thiemermann C, Hinds CJ.** 2005. Peptidoglycan of *Staphylococcus aureus* upregulates monocyte expression of CD14, Toll-like receptor 2 (TLR2), and TLR4 in human blood: possible implications for priming of lipopolysaccharide signaling. *Infect Immun* **73**:7613-7619.
71. **Phillips NJ, Adin DM, Stabb EV, McFall-Ngai MJ, Apicella MA, Gibson BW.** 2011. The lipid A from *Vibrio fischeri* lipopolysaccharide: a unique structure bearing a phosphoglycerol moiety. *J Biol Chem* **286**:21203-21219.
72. **Post DM, Yu L, Krasity BC, Choudhury B, Mandel MJ, Brennan CA, Ruby EG, McFall-Ngai MJ, Gibson BW, Apicella MA.** 2012. O-antigen and core carbohydrate of *Vibrio fischeri* lipopolysaccharide: composition and analysis of their role in *Euprymna scolopes* light organ colonization. *J Biol Chem* **287**:8515-8530.
73. **Wheeler R, Chevalier G, Eberl G, Gomperts Boneca I.** 2014. The biology of bacterial peptidoglycans and their impact on host immunity and physiology. *Cell Microbiol* **16**:1014-1023.
74. **Doino Lemus J, McFall-Ngai MJ.** 2000. Alterations in the proteome of the *Euprymna scolopes* light organ in response to symbiotic *Vibrio fischeri*. *Appl Environ Microbiol* **66**:4091-4097.

75. **Horwitz AH, Williams RE, Nowakowski G.** 1995. Human lipopolysaccharide-binding protein potentiates bactericidal activity of human bactericidal/permeability-increasing protein. *Infect Immun* **63**:522-527.

CHAPTER 3

O-antigen and Core Carbohydrate of *Vibrio fischeri* Lipopolysaccharide: Composition and Analysis of Their Role in *Euprymna scolopes* Light Organ Colonization

PREFACE

A version of this chapter was published as: Post DM*, Yu L*, Krasity BC*, Choudhury B, Mandel MJ, Brennan CA, Ruby EG, McFall-Ngai MJ, Gibson BW, Apicella MA. “O-antigen and Core Carbohydrate of *Vibrio fischeri* Lipopolysaccharide.” J Biol Chem. 2012 Mar 9;287(11):8515-30.

*These authors contributed equally to this work.

All supplemental material has been included here, and the figures and tables renumbered. Some sections have been renamed. The footnoted reference to Brennan, Mandel and Ruby, in prep, has been updated to cite the published manuscript. Initial work on the transposon mutant library and identification and characterization (Fig. 3-1) of the *waaL* mutant was by CAB and MJM.

Generation of the complement and associated strains, and more thorough characterization of the *waaL* mutant and complemented strain (Figs. 3-2 through 3-6) was by BCK, with advice from MJM, CAB, EGR and MJM-N. Purification and characterization of LPS derivatives was by DMP, LY, BC, BWG and MAA. BCK primarily wrote those portions of the manuscript pertaining to the *Euprymna-Vibrio* symbiosis in general, speculation on the importance of the O antigen in the symbiosis, and the experiments (Figs. 3-2 to 3-6) indicated above, with notable input from MJM-N.

ABSTRACT *Vibrio fischeri* exists in a symbiotic relationship with the Hawaiian bobtail squid, *Euprymna scolopes*, where the squid provides a home for the bacteria, and the bacteria in turn provide camouflage that helps protect the squid from night-time predators. Like other Gram-negative organisms, *V. fischeri* expresses lipopolysaccharide (LPS) on its cell surface. The structure of the O-antigen and the core components of the LPS and their possible role in colonization of the squid have not previously been determined. In these studies, an O-antigen ligase mutant, *waaL*, was utilized to determine the structures of these LPS components and their roles in colonization of the squid. WaaL ligates the O-antigen to the core of the LPS; thus, LPS from *waaL* mutants lacks O-antigen. Our results show that the *V. fischeri waaL* mutant has a motility defect, is significantly delayed in colonization, and is unable to compete with the wild-type strain in co-colonization assays. Comparative analyses of the LPS from the wild-type and *waaL* strains showed that the *V. fischeri* LPS has a single O-antigen repeat composed of yersiniose, 8-epi-legionaminic acid, and *N*-acetylfucosamine. In addition, the LPS from the *waaL* strain showed that the core structure consists of L-glycero-D-manno-heptose, D-glycero-D-manno-heptose, glucose, 3-deoxy-D-manno-octulosonic acid, *N*-acetylgalactosamine, 8-epi-legionaminic acid, phosphate, and phosphoethanolamine. These studies indicate that the unusual *V. fischeri* O-antigen sugars play a role in the early phases of bacterial colonization of the squid.

INTRODUCTION

One of the best studied examples of bacterial-eukaryotic host symbiosis is the relationship between *Vibrio fischeri* and the Hawaiian bobtail squid, *Euprymna scolopes*. *V. fischeri* is a luminescent bacterium that colonizes the light organ of the squid and provides it with camouflage from nocturnal predators. Colonization occurs early in life with populations of *V.*

fischeri filling the host crypts between 12 and 18 h post-hatching (1, 2). A remarkable characteristic of this interaction is that *V. fischeri* is culled by the squid from the entire microbiota of seawater of which *V. fischeri* is a relatively minor constituent (~0.1%). Alterations in the anatomy of the squid light organ, which serve to preserve this monobacterial state, are induced during initial colonization and completed within 96–120 h after hatching (3). These alterations include the loss of a ciliated field of epithelial cells that expedite the transit of the bacterial cells to the light organ during the colonization process (4). These changes in morphogenesis within the light organ are orchestrated by lipid A and peptidoglycan components released by *V. fischeri* (4, 5).

V. fischeri is a Gram-negative bacterium and expresses lipopolysaccharide (LPS) on its cell surface. LPS is composed of three regions as follows: lipid A, which anchors the structure to the outer membrane; the core, and the O-antigen, which typically consists of repeating saccharide units. A recent publication from our group showed that *V. fischeri* expresses a heterogeneous mixture of lipid A structures with varying lengths of acyl groups, ranging from tetra- to octa-acylated structures (6). In addition to the lipid A studies, our group also examined the whole LPS structure by silver-stained SDS-PAGE. These data demonstrated that unlike traditional LPS, which generates a ladder-like banding pattern on the gel due to the O-antigen repeat units, the *V. fischeri* LPS migrated as two low molecular weight bands, likely corresponding to the core and the core plus one O-antigen repeat unit as observed previously (7).

This study was initiated to elucidate the components of the *V. fischeri* LPS core and O-antigen and to determine what role they may play in colonization of the squid. A *waaL* mutant was utilized to help discern which components of the LPS were O-antigen and which were core sugars. WaaL is an enzyme that ligates the O-antigen to the lipid A-core of LPS as shown

in *Pseudomonas aeruginosa* (8, 9). Therefore, a *waaL* mutant, which eliminates the function of this ligase, is expected to express an LPS structure that lacks the O-antigen. A combination of experimental approaches, including mass spectrometry, GC-MS, and NMR were utilized to study the *V. fischeri* LPS structures from both wild-type and *waaL* mutant strains. In addition, motility and colonization studies were performed to determine whether the *V. fischeri* O-antigen plays a role in initiating colonization of the squid.

MATERIALS AND METHODS

Bacterial Strains and Culture Conditions. Strains and vectors used in this study are shown in Table 3-1. Wild-type strain ES114, which was previously isolated from *E. scolopes* (10), was used in this study. All *V. fischeri* strains were grown in either LBS medium (11) or seawater-based tryptone medium (12) made with Instant Ocean (Aquarium Systems, Mentor, OH). The *V. fischeri* ES114 *waaL* mutant strain, MB06859, was grown on selective medium containing 5 µg/ml erythromycin and was screened for lack of growth on media containing 100 µg/ml kanamycin.

Generation of *waaL* Mutant and Complement Strains. *V. fischeri* ES114 was mutagenized by conjugation with Tn5 delivery vector pMJM10 (13), a derivative of pEVS170 (14). The resulting transconjugants were selected for erythromycin resistance (transposon integration), arrayed in a 96-well format, and screened for kanamycin sensitivity (loss of donor plasmid). The resulting arrayed library was screened in 96-well format for motility defects in LBS 0.3% agar (13). A semi-arbitrarily primed PCR procedure was used to map transposon insertion junctions using a previously published protocol (15). High quality sequence of the

Table 3-1 Bacterial strains and plasmids

Strains	Description	Ref. or source
<i>V. fischeri</i>		
ES114	Sequenced wild-type <i>E. scolopes</i> light organ isolate	(16)
MB06859	ES114 <i>waaL</i> ::T _{nerm}	This work
BK111	MB06859 harboring pVSV105	This work
BK110	MB06859 harboring pBK01	This work
MB24439	ES114 <i>flaD</i> ::T _{nerm}	(13)
MJM1575	ES114 harboring pVSV103	This work
<i>E. coli</i>		
DH5 α - λ pir	Cloning vector	(17)
MJM534	CC118 λ pir, carries pEVS104 for <i>E. coli</i> to <i>V. fischeri</i> conjugation	(18)
Plasmids¹		
pVSV105	pES213-based plasmid used for complementation, CamR	(19)
pBK01	pVSV105 containing <i>waaL</i> ORF and upstream region	This work
pVSV103	pES213-based plasmid containing LacZ used for marking strains, KanR	(19)
pEVS104	pRK2013-based plasmid, helper for <i>E. coli</i> to <i>V. fischeri</i> conjugation, KanR	(18)
pVSV208	pES213-based plasmid containing <i>rfp</i> , CamR	(19)

¹ Abbreviations used are as follows: Kan, kanamycin; Cam, chloramphenicol.

junctions indicated that one strain, MB06859, contained a Tn5 insertion in the 5' half of *waaL* (VF_0151). This strain displayed growth comparable with the wild-type parent, stably maintained the transposon insertion, and was subsequently designated ES114 *waaL*. Next, a complemented strain was generated. The *V. fischeri waaL* gene and ~50 bp of upstream region, included to incorporate a ribosomal binding site, were PCR-amplified from bacterial chromosomal DNA. The primer sequences used were GCGCATGCATATGGCGATGATTAATAAGTAGATTT (forward primer) and GCGGTACCTTATGACCTGATATCTTTTGTCGAG (reverse primer), with underlined text indicating SphI and KpnI restriction sites, respectively. The amplified product was cut with the restriction enzymes indicated above and ligated into the cloning region of the vector pVSV105 (19). Both this construct, termed pBK01, and the parent vector pVSV105 were conjugated from *Escherichia coli* into the ES114*waaL* mutant, MB06859, as described previously (20), yielding strains designated BK110 (*waaL* complement) and BK111 (empty vector control), respectively.

***V. fischeri* Motility Studies.** Bacterial cultures were grown in LBS medium (11) to an A_{600} of 0.3. Then 3 μ l of these cultures were inoculated into plates containing seawater-based tryptone medium (12) with 0.3% agar, in triplicate. These plates were incubated at 28 °C for 11 h before photographing them and measuring the diameters of the motility fronts.

***E. scolopes* Light Organ Colonization Studies.** Adult *E. scolopes* were collected, maintained, and bred as described previously (21). Hatchling animals were collected as in previous experiments (22). To colonize, animals were exposed overnight to artificial seawater containing ~4000 colony-forming units (CFU)/ml of a *V. fischeri* strain (see Fig. 3-3 for strains used; a description of these strains is in Table 1). Luminescence was evaluated with a TD 20/20

luminometer at 24 and 44 h after the initial exposure to *V. fischeri*. In a confirmatory experiment, after luminescence was quantified at 24 h, animals were rinsed, homogenized, and plated on LBS medium as described previously (23) to determine the number of CFU of bacteria in the light organ.

In the competition experiment *V. fischeri* strain MJM1575, a wild-type strain carrying the plasmid pVSV103 (19), was used. The pVSV103 plasmid, which was introduced to ES114 using the same conjugation methods that were used for pVSV105 and pBK01, carries the *E. coli lacZ* gene, causing the strain to appear blue when exposed to X-gal. Animals in the competition experiment were exposed to ~7000 CFU/ml each of MJM1575 and an unlabeled strain, as indicated in Fig. 3-6. The precise ratio of the unlabeled strain to MJM1575 was determined by plating water from the inoculum on LBS medium containing 100 µg/ml of X-gal and counting blue (MJM1575) and white (unlabeled strains) colonies. The ratio of unlabeled strain to MJM1575 was in the range of 0.97 to 1.02 for all inoculations. At 48 h after the beginning of exposure to bacteria, homogenates of the animals were plated on the same medium as above, and blue and white colonies were counted to determine the ratio of the two strains in light organ tissues.

Visualization of *V. fischeri* in *E. scolopes* Light Organ. Using the protocol described above, animals were colonized with *V. fischeri* ES114 or ES114 *waaL* (MB06859), both of which had been transformed with the plasmid pVSV208 (19), enabling them to produce red fluorescent protein. After their luminescence was read 24 h into exposure, animals were fixed and then visualized by confocal microscopy according to a protocol described previously (24). Prior to visualization, animals were exposed overnight to Alexa Fluor 633 phalloidin from

Invitrogen (dissolved in methanol per manufacturer's recommendations) and then used at a 1:40 dilution in marine phosphate-buffered saline (PBS) (24) as a counterstain for actin.

Isolation and SDS-PAGE of LPS. Bacterial cultures were centrifuged, and cell pellets were washed once with PBS. Proteinase K-digested, phenol-extracted LPS was then isolated from the cells using a modified hot phenol/water method as described previously (25). SDS-PAGE analysis was performed as described by Lesse *et al.* (26). The gel was stained with ProQ Emerald glycostain per the manufacturer's instructions (Invitrogen).

Isolation of LPS O-antigen-Core Free of Lipid A. For NMR studies, the LPS from *V. fischeri* ES114 (wild-type) and ES114 *waaL* (~8 mg) was hydrolyzed in 1 ml of 1% acetic acid at 100 °C for 2 h. Each sample was then centrifuged at 4 °C for 30 min, and the supernatant was aspirated, buffered to pH 7.4, and subjected to further purification with a size exclusion column. The purified product was then repeatedly lyophilized and rehydrated in deuterium oxide three times prior to NMR analysis.

Compositional Analyses. LPS isolated from the wild-type and the *waaL* mutant *V. fischeri* strains was treated with 2% acetic acid at 100 °C for 2 h to liberate the polysaccharide (PS)⁴ from the lipid. The samples were centrifuged at 14,000 rpm for 5 min, and the supernatants containing the PS were lyophilized. The precipitates were suspended in water, lyophilized, and kept for fatty acid analysis. PS samples were further processed with 48% aqueous HF at 4 °C for 48 h, followed by removal of acid by cold water dialysis for 2 days to generate dephosphorylated PS. Intact LPS samples were also treated with anhydrous hydrazine to obtain *O*-deacylated LPS (*O*-LPS). Briefly, the samples were treated with anhydrous hydrazine at 40 °C for 1 h, followed by precipitation of the modified LPS by cold acetone (−70 °C). The precipitate was washed with cold acetone twice and lyophilized.

Composition analyses of intact LPS, PS, *O*-LPS, and dephosphorylated PS were done by GC-MS as their trimethylsilyl-methylglycoside derivatives. Briefly, the samples were methanolized with 1 M methanolic HCl at 80 °C for 18 h, followed by re-*N*-acetylation using 4:1:1 methanol/pyridine/acetic anhydride at 100 °C for 1 h. Finally, trimethylsilylation was done using Tri-Sil reagent at 80 °C for 30 min. GC-MS was done on a Resteck-5Sil (MS) column (15 m × 0.25 μm × 0.2 μm, length × inner diameter × film thickness) and loaded with a split injector using a 1:25 split. Samples were run at 100 °C for a 5-min hold, and then the temperature was ramped up at a rate of 4 °C/min until the temperature reached 260 °C. Data were collected in the electron impact mode at 70 eV and scanned within the range of 50–600 atomic mass units.

MALDI-MS Analyses of LPS. To generate LPS more amendable to mass spectrometric analyses, *O*-LPS samples were prepared by treating ~50 μg of LPS with 50 μl of anhydrous hydrazine followed by acetone precipitation as described previously (27). The PS was liberated from the lipid, and phosphate groups were removed from the PS by treatment with 48% aqueous HF for 16–24 h at 4 °C. HF was removed and neutralized by drying the sample with nitrogen gas under vacuum and over sodium hydroxide, as described previously (28). All samples were desalted by drop dialysis using 0.025-μm pore size nitrocellulose membranes (Millipore, Bedford, MA) and were subsequently lyophilized. Samples were reconstituted in 5–20 μl of HPLC grade H₂O; 1 μl was loaded onto the target, allowed to dry, and then overlaid with 1 μl of matrix (50 mg/ml 2,5-dihydroxybenzoic acid (Laser Biolabs, Sophia-Antipolis Cedex, France) in 70% acetonitrile). Samples were subsequently analyzed using matrix-assisted laser desorption ionization mass spectrometry (MALDI-MS) on an LTQ linear ion trap mass spectrometer coupled to a vMALDI ion source (MALDI-LIT) (Thermo Fisher, Waltham, MA). The vMALDI source uses a nitrogen laser that operates at 337.1-nm wavelength, 3-ns pulse duration, and 60-

Hz repetition rate. Data were collected in either the positive or negative ion mode using the automated gain control and the automatic spectrum filter settings. Tandem mass spectrometry (MS^n) data were collected using a precursor ion selection window of 2–3 m/z and normalized collision energy of 35–40%.

NMR Spectroscopy. Wild-type and *waaL* *V. fischeri* oligosaccharides were lyophilized and dissolved in 100% D_2O at ~0.5 mM concentration for NMR studies. 1H homonuclear two-dimensional DQF-COSY (29), TOCSY (30), ROESY (30), and NOESY (31) experiments and $^1H/^{13}C$ two-dimensional heteronuclear HMQC, HMBC (32), H2BC (33, 34) and selective-TOCSY-HSQC experiments (35) were collected at 25 and/or 35 °C on a Bruker Avance II 800 MHz spectrometer equipped with a cryoprobe. The $^{31}P/^1H$ COSY spectra (36) were acquired at 25 °C on a Bruker Avance II 500 MHz spectrometer. The 1H and ^{31}P chemical shifts were referenced to 2,2-dimethyl-2-silapentane-5-sulfonate and external 2% H_3PO_4 in D_2O (37), respectively. NMR spectra were processed with the NMRPipe package (38) and analyzed using NMRView (39).

RESULTS

Generation and Confirmation of *waaL* Mutant and Complement Strains. In an effort to find genes involved in *V. fischeri* motility, random mutagenesis was performed on the wild-type strain ES114. Motility was tested using a soft agar test where a thin metal replicating tool containing bacteria was inserted into the center of a plate, and the migration of the organisms from this central spot was mapped over time. One of the transconjugants, MB06859, isolated from the random mutagenesis showed a defect in its motility (Fig. 3-1), whereas the wild-type strain and a strain harboring a similar antibiotic resistance cassette in a different gene showed no

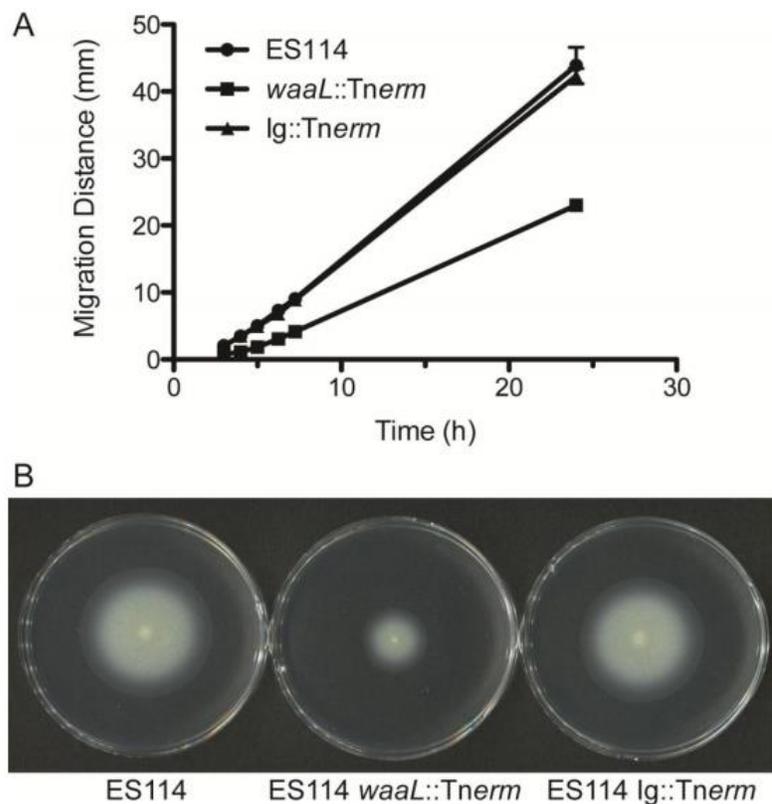


FIG 3-1. Motility testing of *V. fischeri* strain ES114 *waaL::Tnerm* demonstrates a diminished flagellar-dependent migration in soft agar. Strains of ES114 wild-type, ES114 *waaL::Tnerm* (MB06859) and ES114 *Ig::Tnerm* (control) were inoculated through LBS 0.3% agar using round toothpicks. Strain ES114 *Ig::Tnerm* harbors a similar antibiotic resistance cassette as the *waaL* strain and was used as a control to demonstrate that the motility defect was not related to the presence of the cassette. The migration distance was measured as the diameter of the outer ring during chemotaxis in soft agar (A), and plates were imaged at 24 h (B). Plotted are the mean and standard error for three replicates in a typical experiment.

motility defect. Sequencing of the transposon insertion site determined that the cassette inserted into the 5' end of the gene was predicted to be *waaL* (VF_0151). A BLASTP query of VF_0151 revealed that it had highest sequence homology with *waaL* from *Vibrionales* bacterium SWAT-3 with an expectation value of 3×10^{-35} . The second highest match for this search was with *waaL* from *Vibrio cholerae*, with an expectation value of 1×10^{-31} . To ensure that the phenotype observed in the *waaL* mutant was not due to a polar effect on a neighboring gene, we generated a complemented strain. The *V. fischeri waaL* gene was successfully cloned into the plasmid pVSV105, transformed into the *waaL* mutant strain, and subsequently designated as BK110. After generation of the complement strain, the plasmid DNA was isolated from BK110 and BK111 (the vector only control strain), and the cloning region was sequenced to verify the correct insertion (data not shown). In motility assays, expression of *waaL* in *trans* restored motility to the *waaL* mutant, indicating that the defects associated with transposon insertion in *waaL* are not merely polar effects. Colonies of the ES114 *waaL* mutant (MB06859) and the empty vector control (BK111) averaged 47 and 46% of wild-type swim diameter, respectively, whereas the complement strain, BK110, averaged 79% of wild-type swim diameter (Fig. 3-2). Complementation was also verified in the context of animal colonization (see below).

***E. scolopes* Light Organ Colonization Studies.** In an effort to determine how the motility defect of the *waaL* mutant impacts interactions with *E. scolopes*, light organ colonization studies were conducted. Luminescence was significantly higher for wild-type colonized animals than *waaL* mutant-colonized animals at 24 h, and complementation reversed the decrease in luminescence associated with the *waaL* mutant (Fig. 3-3A). To control for the *waaL* mutant's defect in motility, a symbiosis-associated trait (40), we exposed 10 squid in this experiment to *V. fischeri* MB24439, a strain that carries a mutation in the *flaD* gene and has

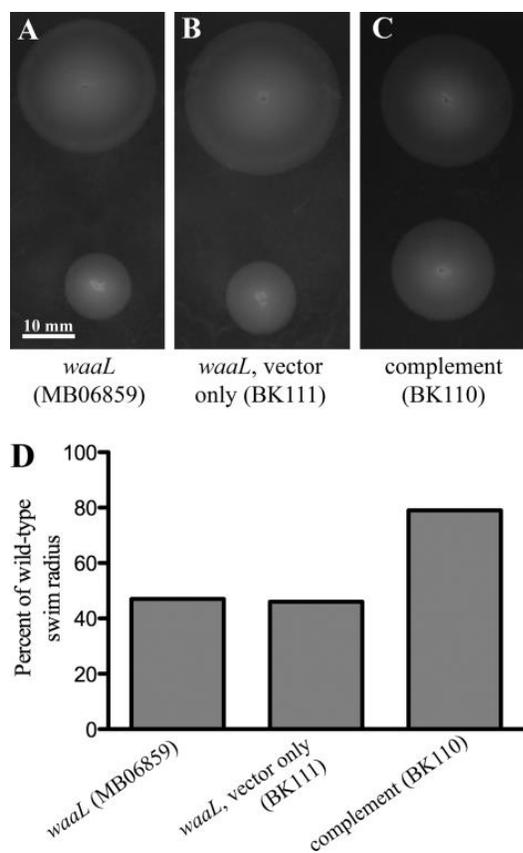


FIG 3-2 Motility studies of *waaL* mutant bacteria and complementation

constructs. (A)-(C), representative images of motility assays comparing wild-type *V. fischeri* (at top of each plate) to the strain indicated. (D) Comparison of diameter between each strain and the wild-type culture spotted on the same plates.

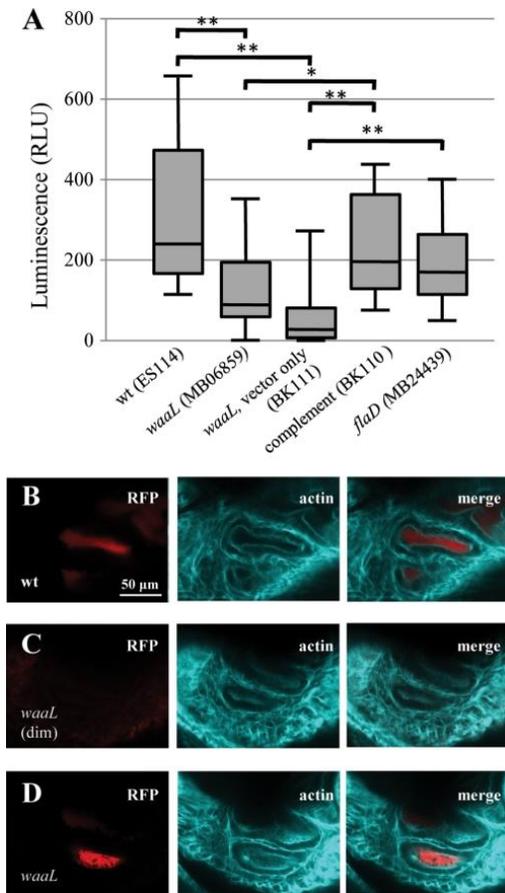


FIG 3-3 Animal colonization by *waaL* mutant. (A) Box-whisker plot of luminescence, in relative luminescence units (*RLU*), of animals colonized with different bacterial strains 24 h into exposure. $n = 10$ animals for each treatment except $n = 20$ for MB06859. Statistical comparison of treatments is by Mann-Whitney *U* test, following Kruskal-Wallis test for all data. Any two treatments not marked with an asterisk were not found to be significantly different (*, $p < 0.05$; **, $p < 0.01$). (B)-(D) Confocal microscopy of light organ crypts (actin cytoskeleton, shown in teal) in animals colonized with red fluorescent protein-labeled *V. fischeri* (red) at 24 h into exposure. (B) shows an animal exposed to wild-type bacteria, and (C) and (D) show animals exposed to the *waaL* mutant. The animal in (C) had no appreciable luminescence (*dim*) at the time of fixation (defined as relative luminescence units < 2), but the animal in (D) was noticeably luminescent.

similar motility behavior to the *waaL* mutant strain MB06859 (13), but it is not predicted to have the cell surface abnormalities of MB06859. Animals colonized with the *flaD* mutant did not show significantly different luminescence at 24 h than either wild-type- or *waaL* mutant-colonized animals (Fig. 3-3A). Luminescence readings for the same animals were taken at 44 h, at which point all animals had detectable luminescence (defined as >2 relative luminescence units), and a significant, but not highly significant, difference was found between colonizing strains; mean luminescence remained higher for wild-type-colonized animals than *waaL* mutant-colonized animals ($p = 0.037$ for comparison of wild-type-colonized and *waaL* mutant-colonized animals) (Fig. 3-4). These data suggest that *waaL* mutant bacteria are slow to colonize the *E. scolopes* light organ, a phenotype that may be attributable to their motility defect.

We verified that low animal luminescence at 24 h was not due to a defect in bacterial luminescence, and in fact it resulted from low levels of bacterial colonization. In wild-type-colonized animals (all of which had detectable luminescence) and in three of four *waaL* mutant-colonized animals that had detectable luminescence, fluorescently labeled bacteria could be detected in the light organ crypt spaces Fig. (3-3B,D). This result was not seen with those *waaL* mutant-colonized animals that had insignificant levels of luminescence at 24 h (3-3C). In a confirmatory experiment comparing animal colonization by either ES114 (wild-type) or MB06859 (*waaL* mutant) alone, wild-type-colonized animals were shown (via homogenization and plating at 24 h) to harbor significantly more *V. fischeri* CFUs than those colonized by the *waaL* mutant (Fig. 3-5).

Since it has previously been shown that a *V. fischeri* strain that shows a subtle defect or no defect in colonization when presented alone may be deficient when competed against a wild-type strain in colonization (41), we performed competition experiments using the strains

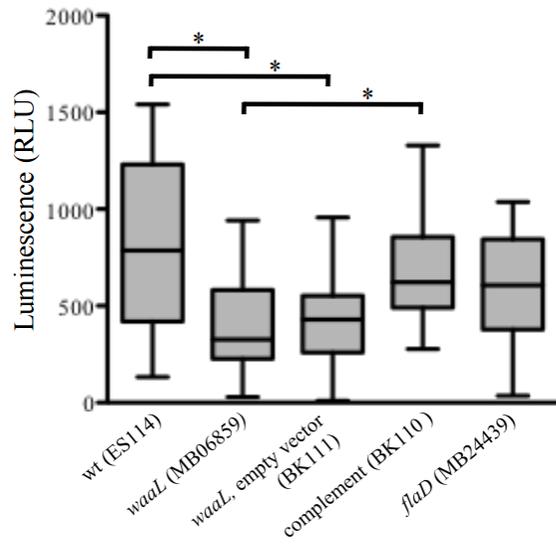


FIG 3-4 Box-whisker plot of luminescence, in relative luminescence units (RLU), of animals colonized with different bacterial strains, 44 h into exposure. N = 10 animals for each treatment except N=20 for MB06859. Statistical comparison of treatments is by Mann-Whitney U test, following Kruskal-Wallis test for all data. Significantly different treatments are marked with * ($p < .05$).

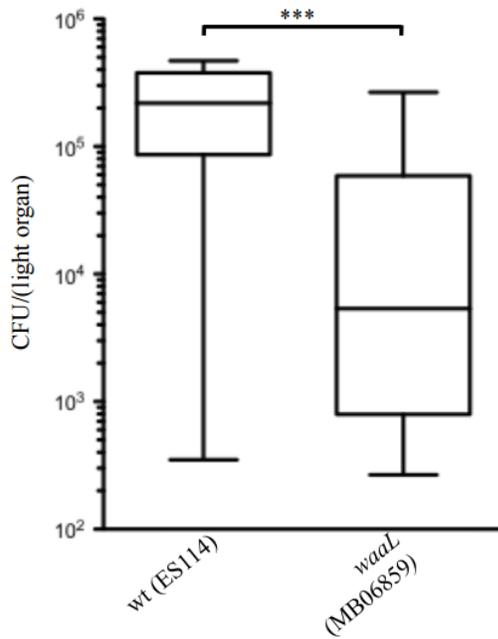


FIG 3-5 Box-whisker plot of *V. fischeri* CFU per light organ, determined by homogenizing and plating animals frozen 24 h into exposure. Animals were colonized with either wild-type or *waaL* mutant bacteria, N = 16 animals for each treatment. Statistical comparison of treatments is by Mann-Whitney U test, $p < .001$ (indicated by ***). Logarithmic scale.

generated in this study. For competition experiments (Fig. 3-6), in which two strains are co-incubated in the animal, a convention has been developed in which results for each strain are expressed in terms of relative competitive index (RCI) (23). Details for how this value is calculated are given in the legend of Fig. 3-6. For this experiment, evaluation of RCI values indicated that the *waaL* mutant and empty vector control, but not the complemented strain or the unlabeled wild-type strain, were outcompeted by the labeled wild-type strain. These results show that mutation of *waaL* renders *V. fischeri* less competitive than wild-type bacteria in colonizing the light organ.

SDS-PAGE of LPS. The LPS from *V. fischeri* strains ES114 wild-type, ES114 *waaL* (MB06859), ES114 *waaL* vector only control (BK111), and ES114 *waaL* complemented with *waaL* (BK110) were analyzed by SDS-PAGE (Fig. 3-7). These data showed that the wild-type LPS migrated as a major band containing one O-antigen subunit and a second faster migrating faint core band (Fig. 3-7, *lane 2*). The *waaL* LPS had a fast migrating lower major band consistent with a core region without O-antigen (Fig. 3-7, *lane 3*), clearly indicating that the *waaL* strain expresses a truncated LPS structure. The *waaL*-complemented strain, BK110, had a banding pattern that corresponded to the wild-type LPS (Fig. 3-7, *lane 4*), and the *waaL* vector only control (BK111) LPS migrated in a similar fashion as the *waaL* LPS (*lane 5*). These results demonstrate that the differences seen in the *waaL* strain were due to a mutation in *waaL*.

Composition Analysis of the LPS by GC-MS. Composition analysis, by GC-MS, of the wild-type LPS showed the presence of Glc, Kdo, FucNAc, L-glycero-D-mannoheptose (LD-Hep), GalNAc, and GlcNAc as the major sugars (Figs. 3-8 through 3-11). In addition, there were several peaks eluting at retention times 14.07, 14.98, and 15.13 min that had identical electron

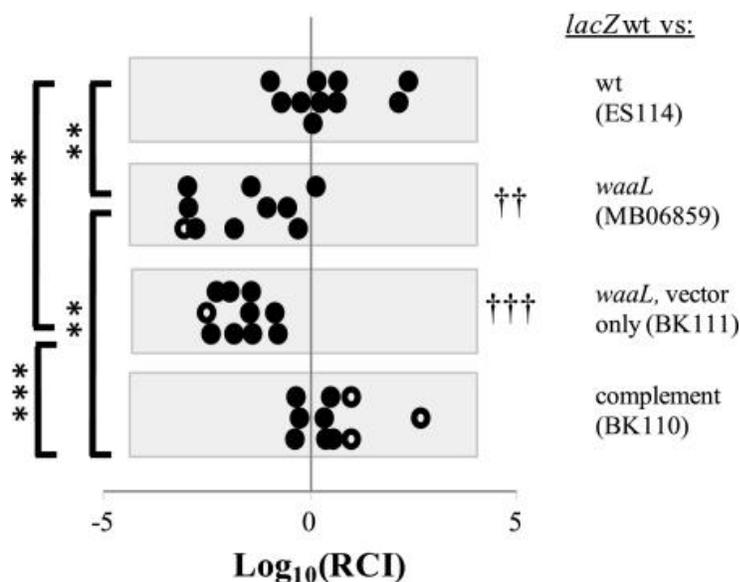


FIG 3-6 Colonization of the light organ competition assays of the various strains compared with the wild-type strain. Competition between wild type or *waaL* strains and MJM1575 (*lacZ*-labeled wild-type *V. fischeri*) is shown. The RCI is calculated as the ratio of (unlabeled strain/MJM1575) colonies from each animal plated, divided by the ratio of those strains in the original inoculum (20). Log₁₀(RCI) <0 indicates that MJM1575 is more prevalent than the unlabeled strain, and log₁₀(RCI) >0 indicates the opposite. *Closed circles* represent individual animals in which both strains were detected, and *open circles* indicate animals in which only one strain was detected; the value of log₁₀(RCI) for these points represents a limit of detection. $n = 10$ animals for each treatment. Treatments for which mean log₁₀(RCI) was significantly different from 0 by one sample *t* test are indicated by †† or ††† ($p < 0.01$ and $p < 0.001$, respectively). One-way analysis of variance and unpaired *t* tests were also used to compare log₁₀(RCI) between competitions, with significant differences shown in figure (**, $p < 0.01$; ***, $p < 0.001$).

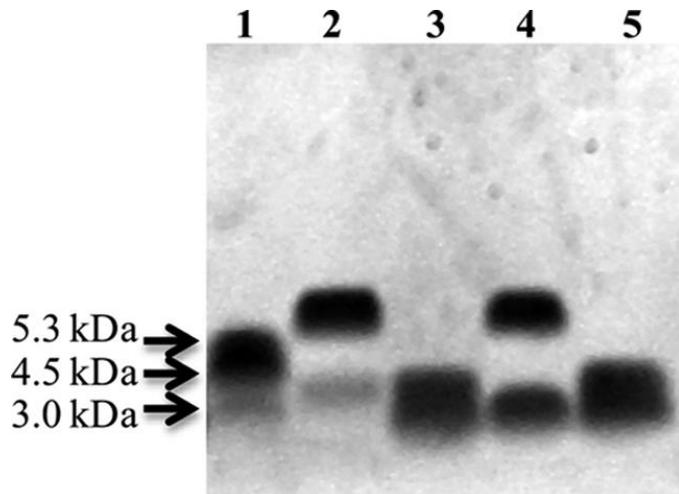


FIG 3-7 SDS-PAGE analyses of *V. fischeri* LPS. LPS was isolated from *Neisseria gonorrhoeae* strain PID2 and included as a molecular weight marker (*lane 1*). Lanes 2–5 are ES114 wild-type, ES114 *waaL* (MB06859), ES114 *waaL* complemented with *waaL* (BK110), and ES114 *waaL* vector only control (BK111), respectively. The gel was stained with ProQ Emerald glycostain.

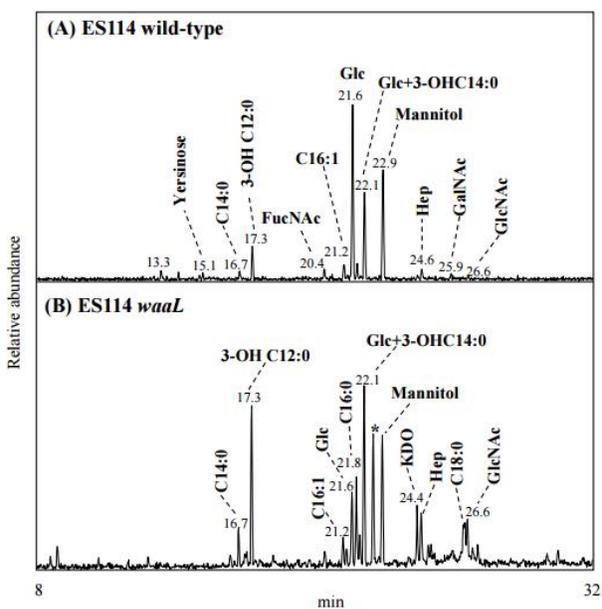


FIG 3-8 GC-MS analyses of the TMS derivatives of LPS isolated from (A) ES114 wild-type and (B) ES114 *waaL* *V. fischeri*. Mannitol was added as internal standard. The peak labeled with an * indicates an impurity peak in the sample.

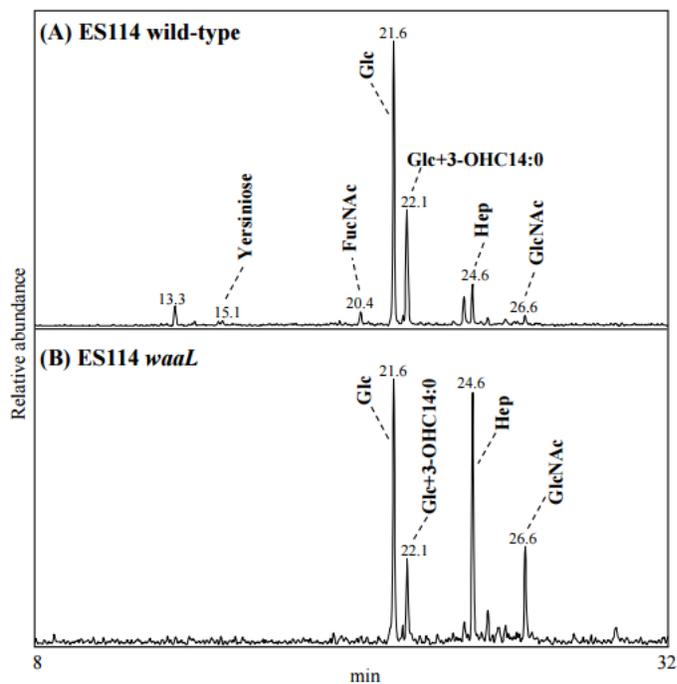


FIG 3-9 GC-MS analyses of the TMS derivatives of O-LPS isolated from (A) ES114 wild-type and (B) ES114 *waaL* *V. fischeri*.

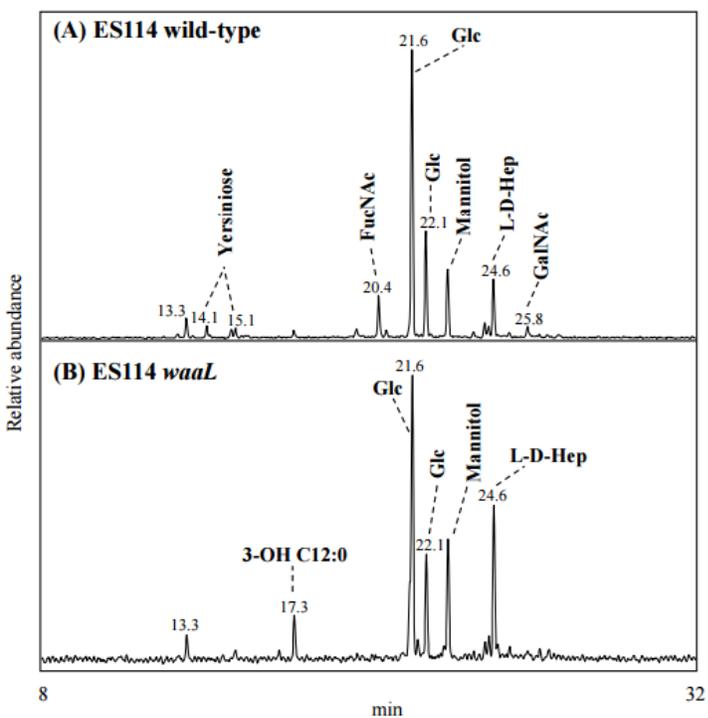


FIG 3-10 GC-MS analyses of the TMS derivatives of PS isolated from (A) ES114 wild-type and (B) ES114 *waaL* *V. fischeri*. These data demonstrate that the wild-type sample has FucNAc and Yersiniose whereas the PS isolated from ES114 *waaL* does not contain those sugars.

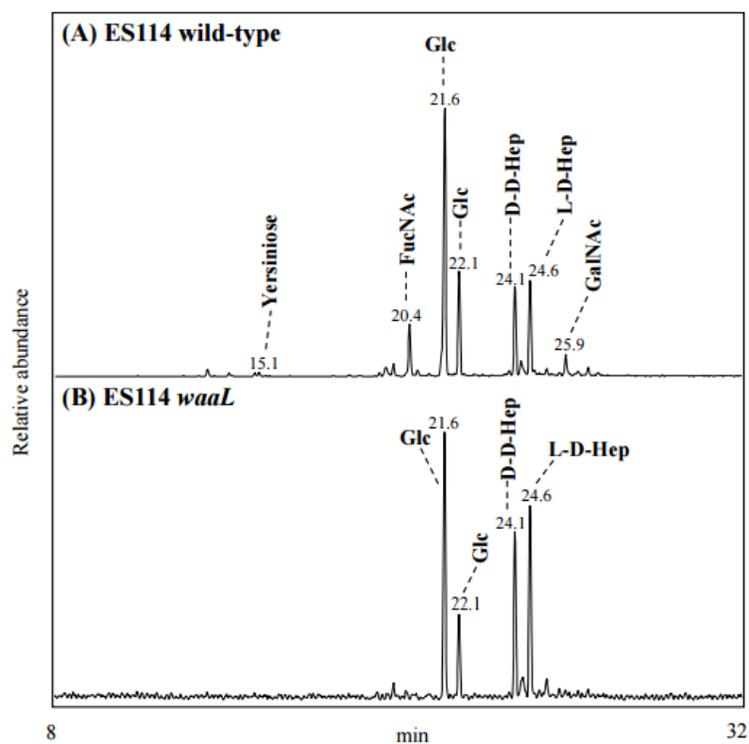


FIG 3-11 GC-MS spectra of TMS derivatives of HF-treated PS from (A) ES114 wild-type and (B) ES114 *waaL* *V. fischeri*. The HF treatment removed any phosphate groups from the structure. The appearance of DD-Hep in the HF treated PS, but not in the untreated PS demonstrates that a phosphate group is linked to this sugar. The assignment of LD-Hep versus DD-Hep was done by matching the GC-MS fragmentation pattern with previous reports of LD- and DD-Hep on similar phase columns.

impact fragmentation patterns, which we proposed to be from the sugar yersiniose (Yer) (Figs. 3-8 through 3-11). A small amount of ribose was also detected, which is most likely a contaminant from RNA co-extracted with the LPS. The composition analysis of the LPS from the *waaL* mutant strain showed the presence of Glc, Kdo, LD-Hep, GlcNAc, and GalNAc (Figs. 3-8 through 3-11); however, the amount of GalNAc detected appeared lower in the mutant compared with the wild-type LPS. In addition, both FucNAc and Yer were not clearly detected in the *waaL* LPS samples, suggesting that these sugars were absent from the *waaL* LPS. Both the wild-type and *waaL* LPS samples contained C14:0, 3-OH C12:0, 3-OH C14:0, and C16:0 fatty acids along with several unsaturated fatty acids, consistent with the recently elucidated *V. fischeri* lipid A structure (6). Analyses of the *O*-deacylated LPS from both strains indicated that 3-OH C14:0 was the only amide-linked fatty acid (Fig. 3-9), which was also consistent with the proposed *V. fischeri* lipid A structure (6).

The PS from both strains were liberated from the LPS by mild acid hydrolysis and further treated with HF to dephosphorylate the PS. The PS and dephosphorylated PS from both strains were analyzed by GC-MS. The wild-type PS consisted of Glc, DD-Hep, LD-Hep, FucNAc, and GalNAc (Figs. 3-10A and 3-11A). Small peaks in the wild-type sample were detected at 14.06, 14.16, 14.96, and 15.12 min and had fragmentation patterns consistent with those predicted for Yer. The composition of the *waaL* PS was determined to be Glc, DD-Hep, LD-Hep, and GalNAc (Figs. 3-10B and 3-11B). Kdo was also detected in both the wild-type and *waaL* samples. These data demonstrated that the wild-type PS contains FucNAc, whereas the *waaL* PS does not. In addition, the level of GalNAc was higher in the wild-type PS than the *waaL* PS (Fig. 3-11). The putative Yer peaks at 14.06, 14.16, 14.96, and 15.12 min were not detected in the *waaL* PS. The absence of GlcNAc in both PS samples suggests that the GlcNAc detected in the LPS samples

originates from the lipid A structure. The appearance of DD-Hep only after HF treatment clearly indicated the presence of a phosphate group on the DD-Hep residue in the PS from both strains (Fig. 3-11). A small amount of ribose, most likely originating from RNA, was detected in all PS samples.

MS Analyses of LPS. To make the LPS more amenable to MS analyses, the LPS samples were *O*-deacylated, using anhydrous hydrazine, to generate *O*-LPS as described previously (27). Initial evaluation of the *O*-LPS by MALDI-MS analyses demonstrated a clear difference between the wild-type and *waaL* samples (Fig. 3-12). The main glycoform observed in the wild-type *O*-LPS was at m/z 3761 (Fig. 3-12A). Minor glycoforms were detected at m/z 3588, 3096, 2722, and 2203. Examination of the *waaL* *O*-LPS showed that the major glycoforms were at m/z 2203 and 2722 (Fig. 3-12B). Both of the glycoforms detected in the *waaL* sample were also detected as minor glycoforms in the wild-type samples. These results demonstrated that the *waaL* *O*-LPS has a truncated structure compared with the wild-type *O*-LPS. Mass spectrometric analyses of the LPS from the *waaL* vector only control strain (BK111) and the *waaL* complement strain (BK110) showed spectra that were consistent with the *waaL* and wild-type data, respectively (Fig. 3-13), further confirming that the differences seen in the wild-type and *waaL* strains are due to a mutation in *waaL*.

To investigate the *waaL* and wild-type LPS structures further, the *O*-LPS samples were treated with HF to dephosphorylate the LPS and liberate the PS from the lipid A. The PS was subsequently analyzed by MALDI-MS in both the positive or negative ion mode (Fig. 3-13 and 3-14).

To determine the individual components of the LPS structures, multistage mass spectrometric (MS^n) analyses of the dephosphorylated *waaL* PS was carried out in the positive

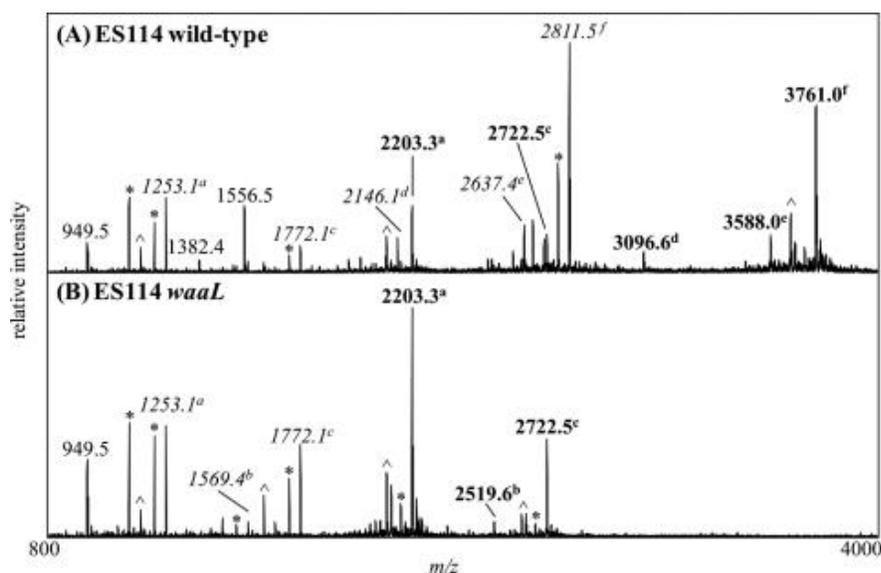


FIG 3-12 Negative ion vMALDI-LIT mass spectra of *O*-LPS from ES114 wild-type (A) and ES114 *waaL* (B) strains. Masses labeled in *boldface* correspond to the predominant glycoforms present in the samples. OS prompt fragments are labeled in *italic*. The *superscript letters a–f* correlate the OS prompt fragments with their intact *O*-LPS masses. Masses labeled with an * designate major masses minus water or CO₂. The neutral loss of phosphoric acid, H₃PO₄, from major masses are indicated with a ^ symbol.

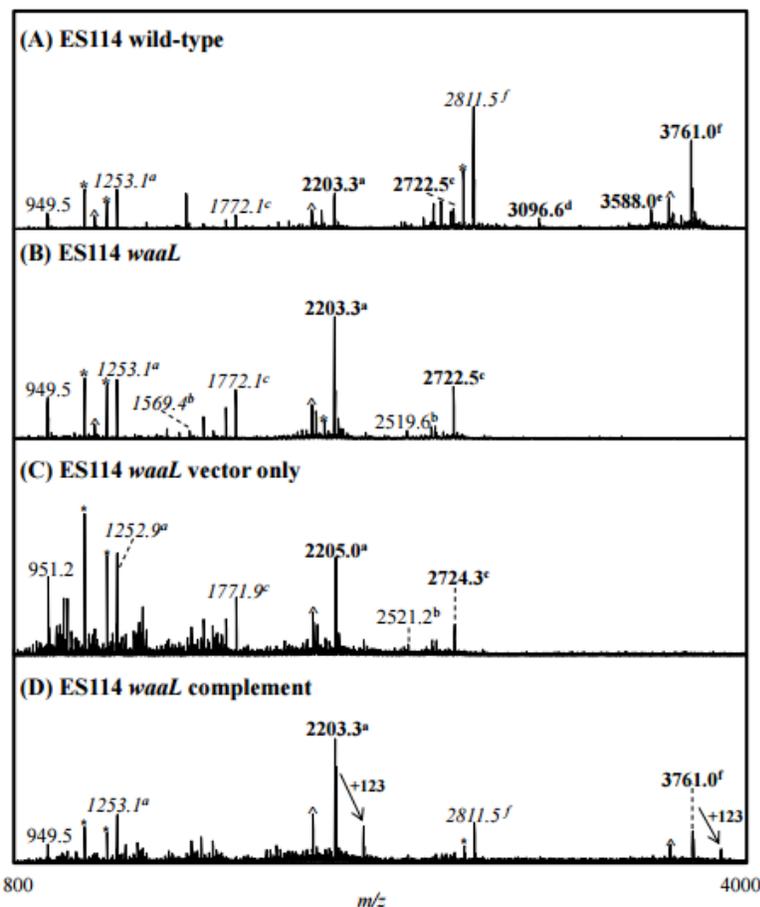


FIG 3-13 Negative ion vMALDI-LIT mass spectra of the *O*-deacylated LPS (*O*-LPS) from (A) ES114 wild-type, (B) ES114 *waaL* (MB06859) (C) ES114 *waaL* vector only control (BK111) and (D) ES114 *waaL* complement (BK110) strains. Masses labeled in bold correspond to the predominant glycoforms present in the samples. OS prompt fragments are labeled in italics. The a-f letters correlate the OS prompt fragments with their intact O-LPS masses. Masses labeled with an * designate major masses minus water or CO₂. The neutral loss of phosphoric acid, H₃PO₄, from major masses are indicated with a ^ symbol. The addition of 123 Da corresponds to the addition of one PEA group. The differences seen in the masses of the ES114 *waaL* vector only control glycoforms compared to the other strains are due to differences in the heterogeneity in the lipid A components of these structures.

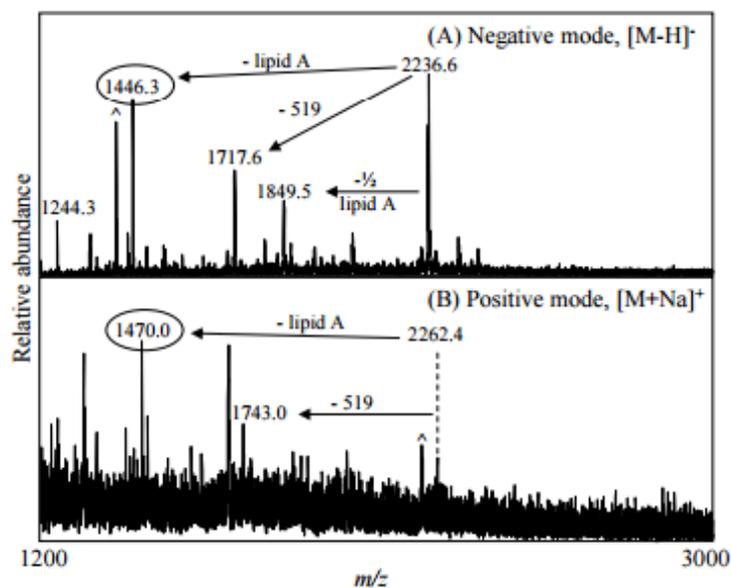


FIG 3-14 MALDI-MS spectra of HF treated ES114 *waaL* PS run in (A) negative ion mode and (B) positive ion mode. HF treatment removes any phosphate groups from the LPS and cleaves the PS from the lipid A. Ions circled indicate the dephosphorylated PS liberated from the *O*-LPS. Peaks labeled with a ^ indicate major ions which have lost CO₂.

ion mode. The PS corresponding to the major *waaL* glycoform was observed in its sodiated form at m/z 1470 (Fig. 3-14). MS/MS analyses of the sodiated monoisotopic mass $[M + Na]^+ = 1470$ Da generated one major fragment at m/z 951, corresponding to a loss of 519 Da (Fig. 3-16A). Based on composition analyses, the loss of 519 Da would correspond to the loss of one GalNAc (203 Da) and one 316-Da component. MS³ fragmentation of the m/z 951 peak resulted in two major fragments at m/z 789 and 731, corresponding to a loss of 162 Da (Glc) and 220 Da (Kdo), respectively. MS⁴ fragmentation of the precursor ion at m/z 731 generated one major fragment at m/z 569 (−162 Da), corresponding to the loss of Glc (Fig. 3-16C). MS⁵ fragmentation of the precursor ion at m/z 569 produced two major fragments at m/z 406 and 377, corresponding to the loss of one Glc (−162 Da) or one Hep (−192 Da), respectively (Fig. 3-16D). These fragments subsequently lost either Hep or Glc, resulting in the generation of a fragment at m/z 215 that corresponds to the sodiated monoisotopic mass of Hep. These data demonstrated that the dephosphorylated *waaL* PS or the “core” consists of two Hep, one Kdo, one GalNAc, two Glc, and one unknown 316-Da component. We compared these data with the *waaL* OS prompt fragments, which contain phosphate groups, to determine what the phosphate components are. Fig. 3-12B shows the negative ion MALDI-MS data from *waaL* O-LPS. The deprotonated OS prompt fragment observed at m/z 1772 corresponds to the intact structure seen at m/z 2722. The 326-Da mass difference between the PS ($M = 1447$ Da) and OS prompt fragment ($M = 1773$ Da) indicates that the phosphorylated *waaL* PS contained phosphate (80 Da) and two phosphoethanolamine (2 PEA, 246 Da) groups.

Similarly, MSⁿ analyses of the dephosphorylated wild-type PS were carried out in the negative ion mode. The major deprotonated ion in the wild-type PS sample was observed at m/z 2136 (Fig. 3-15). The expected dephosphorylated PS ion at m/z 2485, corresponding to the

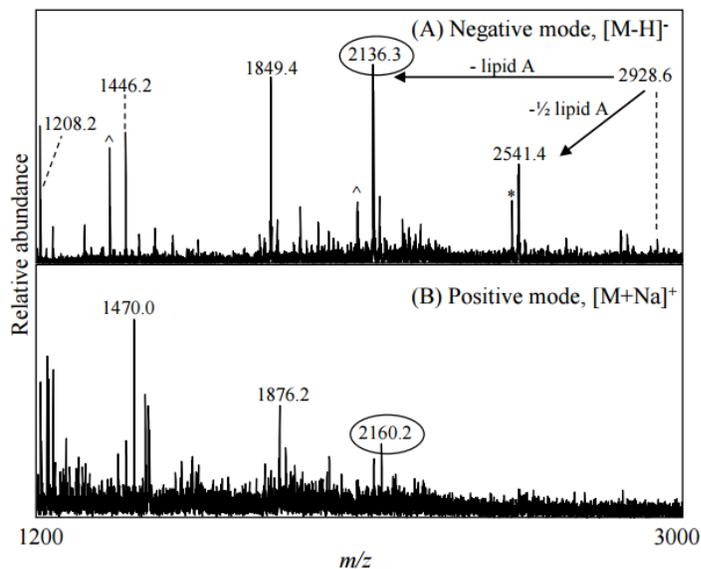


FIG 3-15 MALDI-MS spectra of HF treated ES114 wild-type PS run in (A) negative ion mode and (B) positive ion mode. HF treatment removes any phosphate groups from the LPS and cleaves the PS from the lipid A. Ions circled indicate the dephosphorylated PS liberated from the O-LPS. Peaks labeled with a ^ or with an * indicate major ions which have lost CO₂ or H₂O respectively.

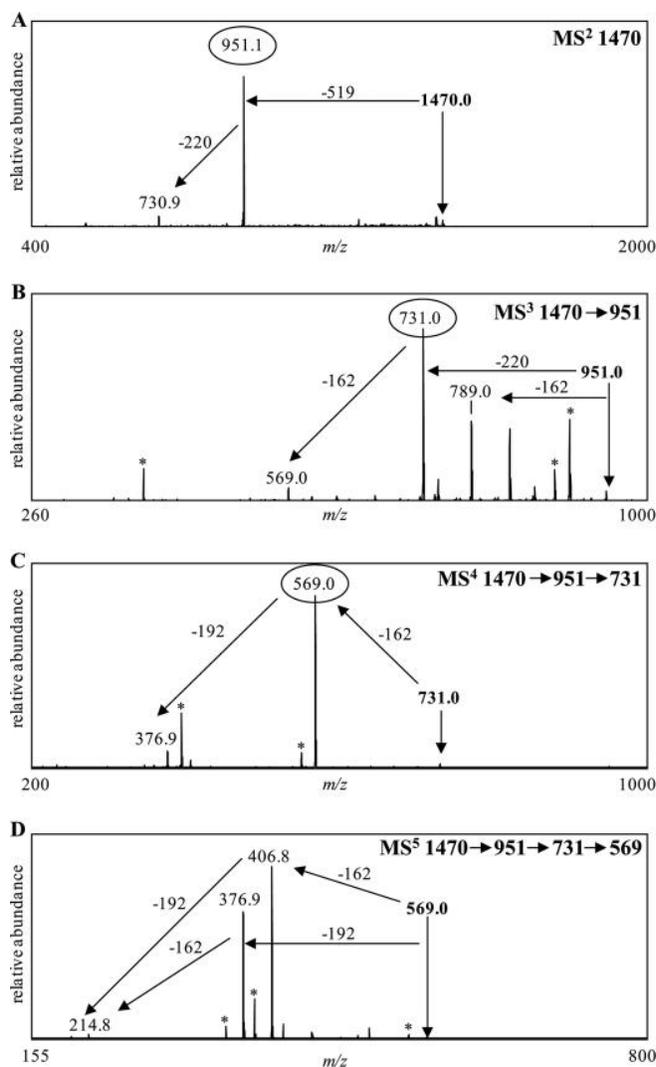


FIG 3-16 Positive-ion vMALDI-LIT mass spectra of HF-treated *waaL* O-LPS

(dephosphorylated PS). After HF treatment, the sodiated monoisotopic mass corresponding to the major glycoform present in *waaL* was observed at m/z 1470.0 (Fig. 3-14). MALDI- MS^n analyses were subsequently performed on this dephosphorylated PS. The parent ion at m/z 1470 was fragmented to yield the product ion at m/z 951 (A). Product ions at m/z 951 (B), m/z 731 (C), and m/z 569 (D) were sequentially fragmented to determine the composition of the parent mass. These analyses showed that the glycoform corresponding to the ion at m/z 1470 is composed of 1 Kdo, 2 Hep, 2 Hex, and two sugars with a combined mass of 519 Da. Masses labeled with an * designate major masses minus water or CO_2 .

major glycoform observed in the wild-type *O*-LPS sample (at m/z 2811) minus 326 Da from phosphate components, was not observed in our experiments. Instead the major ion observed in the dephosphorylated wild-type PS (at m/z 2136) corresponded to a structure that was 348 Da less than the structure observed in the wild-type *O*-LPS spectrum. Composition analyses suggested that the components missing from the wild-type PS after HF treatment were two Yer residues with a nominal mass of 174 Da each. It was previously reported that Yer is acid-labile, and this could account for the absence of these sugars after treatment with HF (42).

Subsequently, the major ion observed in the dephosphorylated wild-type PS at m/z 2136 was investigated by MS^{*n*} analyses (Fig. 3-17). MS/MS fragmentation of this precursor ion generated one major fragment ion at m/z 1208, corresponding to the loss of part of the core (−928 Da), consisting of Kdo, 2 Hep, and 2 Glc (Fig. 3-17A). Three minor fragment ions were also observed at m/z 1446, 1617, and 1820. MS³ fragmentation of the peak at m/z 1208 resulted in one major fragment ion at m/z 892, corresponding to a loss of the same unknown 316-Da moiety noted previously (Fig. 3-17B). MS⁴ fragmentation of the precursor ion at m/z 892 generated one major deprotonated fragment ion at m/z 518, corresponding to the loss of 374 Da (Fig. 3-17C).

Composition analyses suggested that the loss of 374 Da was most likely due to the loss of two FucNAc sugars that have a nominal mass of 187 Da each. The remaining 519-Da component was determined to consist of GalNAc and a 316-Da component.

These data indicated that the wild-type dephosphorylated PS consists of the core plus two FucNAc residues and one 316-Da component. The 316-Da component was determined to be 8-*epi*-legionaminic acid by NMR analyses (see below). In addition, comparisons of the dephosphorylated PS with the *O*-LPS indicated that two Yer sugars (174 Da each) were components of the wild-type *O*-antigen. GC-MS and NMR (see below) both confirmed the

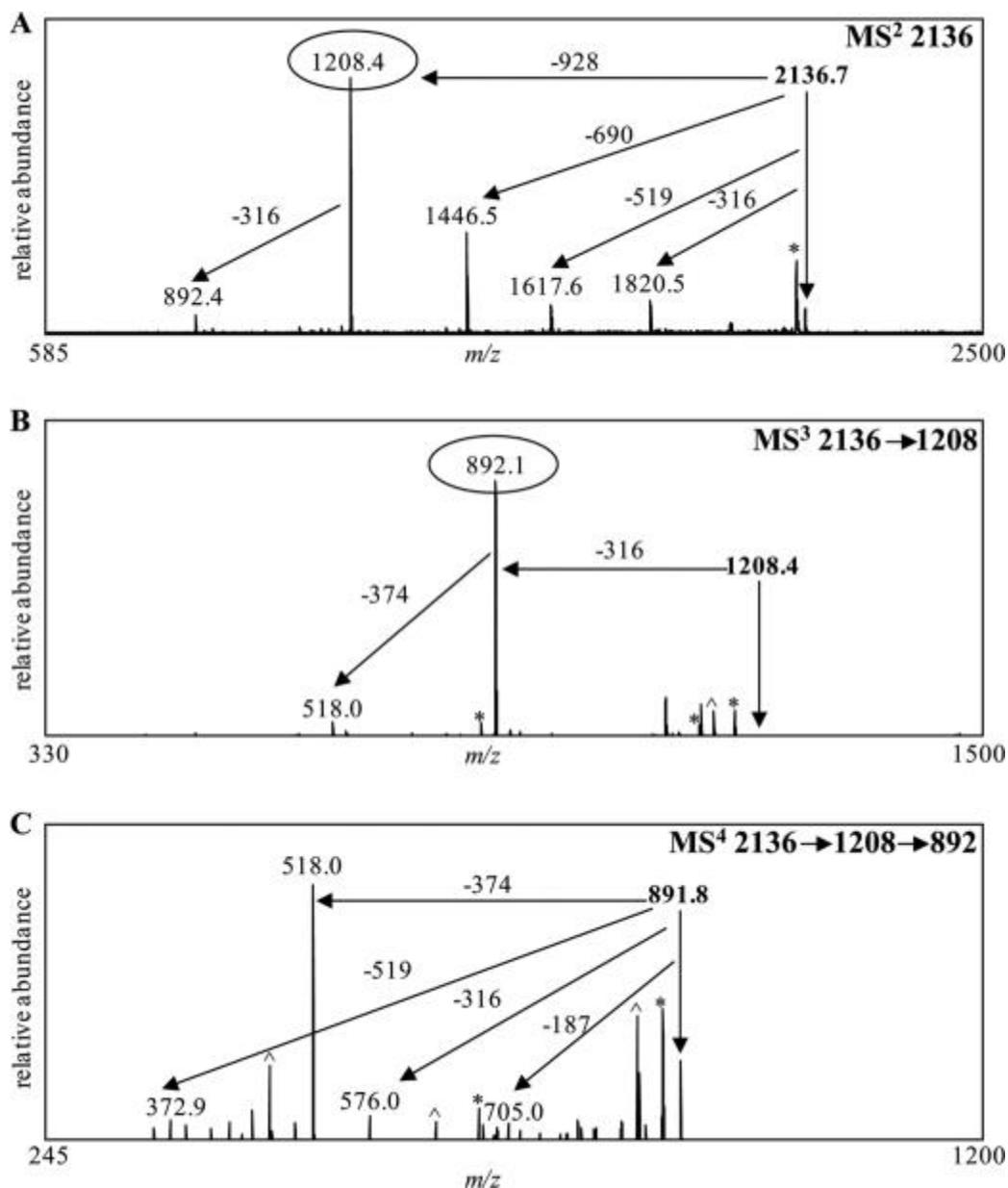


FIG 3-17 Negative ion vMALDI-LIT mass spectra of HF-treated wild-type *O*-

LPS. After HF treatment, the predominant deprotonated monoisotopic mass observed in the wild-type sample was at m/z 2136 (Fig. 3-15). MALDI- MS^n analyses were performed on the parent ion at m/z 2136 (A). Fragment ions at m/z 1208 and 892 were sequentially fragmented to determine the composition of the parent mass as shown in (B) and (C), respectively. Masses labeled with an * designate major masses minus water or CO_2 .

presence of Yer in the O-antigen of the wild-type LPS. To lend additional support to the presence of the Yer residues, we targeted an OS prompt fragment, at m/z 1556 (Fig. 3-18), generated from the wild-type O-LPS. MS/MS analysis of this m/z 1556 precursor showed a loss of 174 Da (Yer) at m/z 1382 (Fig. 3-18A). Fragmentation of this latter ion generated one major fragment ion at m/z 892, corresponding to the loss of 490 Da due one 316 Da sugar (8-epi-legionaminic acid) and one 174 Da sugar (Yer) (Fig. 3-18B). Further fragmentation of the ion observed at m/z 892 resulted in a loss of two consecutive FucNAc residues (187 Da, each) (Fig. 3-18C). These data further confirmed the presence of two FucNAc and two Yer residues on the wild-type LPS structure.

Assignment of Yer Residues F and G. Fig. 3-19A shows the HMQC spectrum of the wild-type *V. fischeri* PS. Using selective-TOCSY-HSQC experiments (35) with a TOCSY mixing time of 120 ms for coherence transfer, spin systems were identified for the sugar residues F and G at the 1st and 2nd positions (Fig. 3-19B) where the peaks are labeled hereafter with the 1st letter referring to the sugar subunit and the rest of the label referring to the position on that sugar residue (could be a proton position (*e.g.* H1), a carbon position (*e.g.* C1), or just a number (*e.g.* 1, referring to position 1)). By comparing the HMQC and HMBC spectra (Fig. 3-19C-D), the 2nd carbon of the residues F and G, whose cross-peaks are labeled as F2 and G2, was found to be connected to the H3 protons of a CH₂ group with ¹³C chemical shift at 32.16 ppm and ¹H chemical shifts at 1.76_{eq} and 1.68–1.70_{ax} ppm (Table 3-2). The 4th carbon of the residues F and G (labeled as FC4 and GC4) was a quaternary carbon with a chemical shift at 77.59 ppm (Table 3-2) and was found to be connected with the protons of H6 (a CH₃ group), H4² (another CH₃ group), H3 (a CH₂ group), H4¹ (a CHOH group), and H5 (another CHOH group) in the HMBC spectrum. The 5th carbon of the residues F and G, whose cross-peaks are labeled as F5

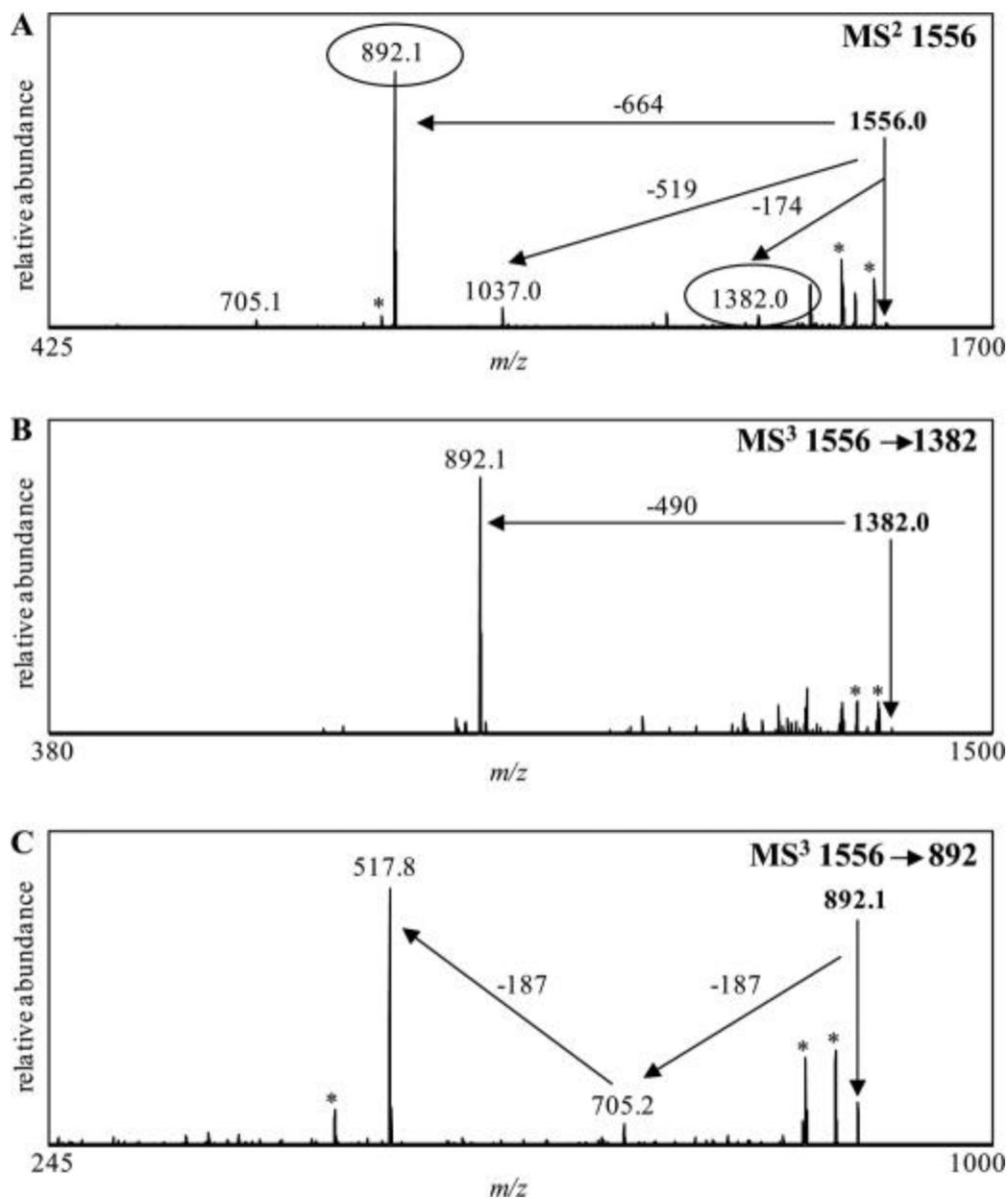


FIG 3-18 Negative-ion vMALDI-LIT mass spectra of the OS prompt fragment, from wild-type *O*-LPS, at m/z 1556. MALDI- MS^n analyses of the parent ion at m/z 1556 generated a number of fragment ions (A). Further fragmentation of these products at m/z 1382 (B) and 892 (C) was performed. Masses labeled with an * designate major masses minus water or CO_2 .

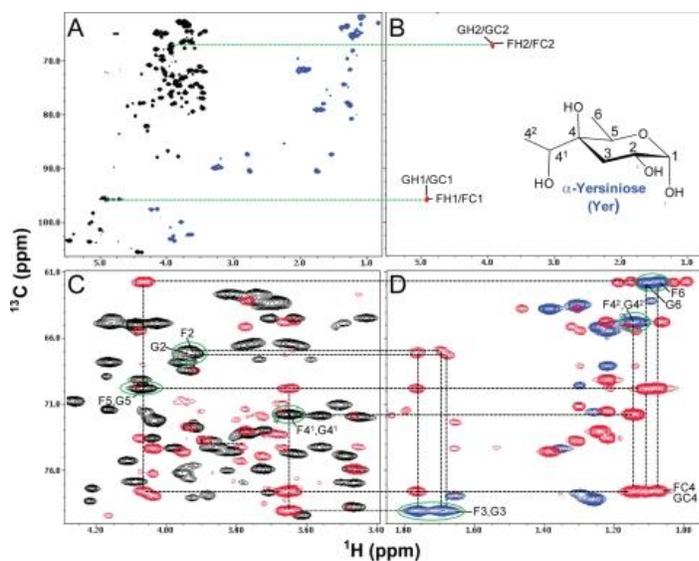


FIG 3-19 HMQC (A) and selective-TOCSY-HSQC (B) as well as HMQC and HMBC overlays (C and D) of the heteronuclear NMR spectra of the wild-type *V. fischeri* O-antigen oligosaccharide. These data illustrate the assignments of the spin systems for the sugar residues F and G of Yer. (A), HMQC spectrum is shown in *black* for the positive peaks and *blue* for the folded negative peaks. (B), selective-TOCSY-HSQC spectrum, shown in *red*, was obtained using a selective excitation pulse at the H1 positions of the sugar residues F and G with a selective TOCSY mixing time of 120 ms. (C) and (D), HMQC spectrum is shown in *black* for the positive peaks and *blue* for the folded negative peaks, and the HMBC spectrum is shown in *red*. The peaks are labeled with the following convention, *i.e.* the 1st letter in the labels refers to sugar subunit, and the rest of the labels refers to the position on that sugar residue. The *green circled peaks* in (C) and (D) indicate the HMQC cross-peaks derived from the sugar residues F and G of Yer. The *dotted lines* indicate the spin connectivities for these residues. F4¹ (or G4¹) and F4² (or G4²) indicate the first and second positions, respectively, of the group attached to the 4th carbon on the sugar residue F (or G). FC4 and GC4 indicate the cross-peaks derived from the C4 carbon of the residues F and G, respectively. The α -Yer structure and the position labeling are indicated in (B).

Sugar residues	$^{13}\text{C}/^1\text{H}$ (ppm)										
	1	2	3	4	5	6	7	8	9	NAc	NAc
G α -Yer-(1 \rightarrow 8)	95.75	67.28	32.16	77.5 9	69.8 3	14.8 9	71.78 <i>d</i>	17.90 <i>e</i>			
	4.91	3.93	1.76e q		4.07	1.11	3.65 ^{<i>d</i>}	1.15 ^{<i>e</i>}			
			1.68a x								
Minor form											
A α -8eLeg5Ac7Ac-(2 \rightarrow	176.1 7	103.4 2	43.38	71.2 4	55.3 6	75.7 2	55.67	71.50	21.2 4	25.03 ^{<i>a</i>}	24.78 ^{<i>b</i>}
			2.77e q	3.51	3.64	3.94	3.90	4.06	1.22	1.92 ^{<i>a</i>}	2.01 ^{<i>b</i>}
			1.76a x							176.70 <i>a</i>	176.81 <i>b</i>

^{*a*} For 5NAc.

^{*b*} For 7NAc.

^{*c*} For 2NAc.

^{*d*} For the first position of the group attached to the 4th carbon on the sugar ring, *i.e.* 4¹ position or the —CHOH— group.

^{*e*} For the second position of the group attached to the 4th carbon on the sugar ring, *i.e.* 4² position or the CH₃— group.

and G5, was connected with the protons H6, H3, and H4¹. Therefore, these spin systems as traced in Fig. 3-19 uniquely identify α -Yer as the sugar residue whose structure and numbering are indicated in Fig. 3-19B. These assignments were further confirmed in the conventional two-dimensional $^1\text{H}/^1\text{H}$ DQF-COSY, TOCSY, and NOESY spectra. The chiral center of the group attached at the 4th carbon (*i.e.* 4¹ position) was not determined from the current data. The ^1H and ^{13}C chemical shifts of the Yer residues assigned in this study were consistent with the previous report in which a series of Yer analogs were synthesized and tested by NMR (43).

Assignment of 8-Epi-legionaminic Acid Residues A and E Fig. 3-20A shows the DQF-COSY spectrum of the wild-type *V. fischeri* PS. Clearly, a CH₂ group at the 3rd position of the residues A and E (confirmed by the HMQC spectrum shown in Fig. 3-21) was assigned with the H_{ax} (labeled as AH3_{ax} and EH3_{ax}) and H_{eq} (labeled as AH3_{eq} and EH3_{eq}) protons that were well resolved from each other, and only the H_{ax} proton gave a cross-peak to its vicinal proton H4. The H9 protons gave COSY cross-peaks to the H8 proton. Moreover, the H3 protons gave TOCSY cross-peaks to H4, H5, and H6, whereas H9 protons gave TOCSY cross-peaks to H8 and H7 protons (Fig. 3-20B). From a series of NOESY experiments with different mixing times, it was found that only the H3 axial proton (not the equatorial proton) gave strong NOEs to H5 proton (Fig. 3-20C), indicating that H5 was in axial position or the 5-*N*-acetyl group was in equatorial position. These data indicated that the chiral center at the position 5 of this residue was different from that of the previously reported structure of *N*-acetyl-pseudaminic acid (44). These NMR assignments were consistent with the structure of 8-epi-legionaminic acid as shown in Fig. 3-20A. The spin systems of the residues A and E of 8-epi-legionaminic acid were further confirmed from a combination of the ¹H/¹³C two-dimensional heteronuclear HMQC, HMBC, and H2BC (33, 34) spectra (Fig. 3-21). For example, the 7th carbon of residue A (AC7) (with the attached

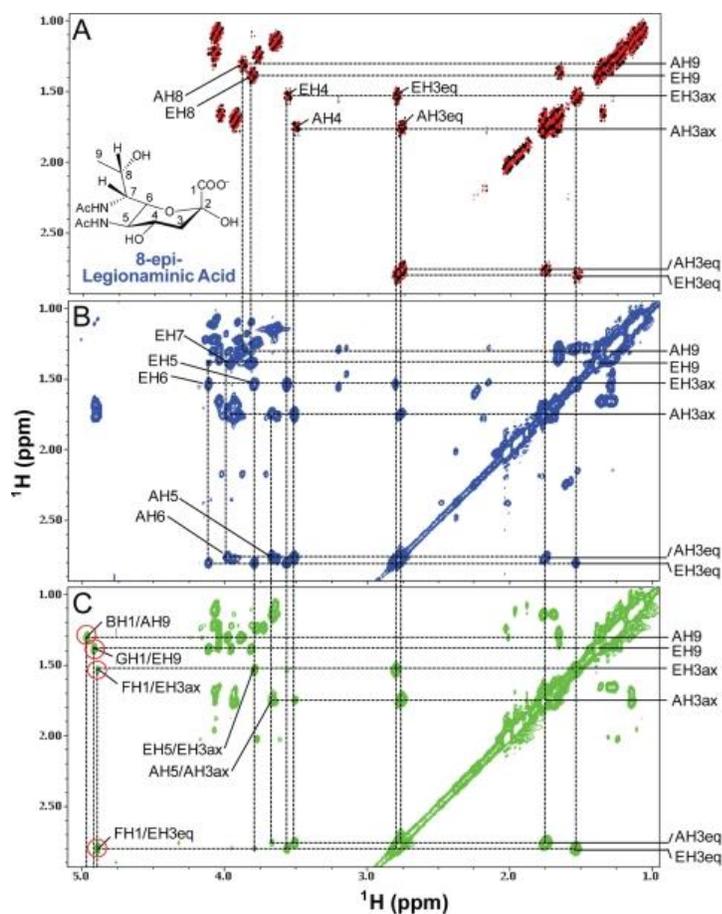


FIG 3-20 DQF-COSY (A), TOCSY (B), and NOESY (C) spectra of the wild-type *V. fischeri* O-antigen oligosaccharide. These data illustrate the assignments of the spin systems for the sugar residues A and E of 8-epi-legionaminic acids. The DQF-COSY spectrum (A) is shown in *black* and *red*, and the TOCSY (B) and NOESY (C) spectra are shown in *blue* and *green*, respectively. Some of the diagonal peaks of residues A and E are labeled along the *right side* of the panels. Some of the off-diagonal cross-peaks of the residues A and E are also labeled. The 1st letter in the labels refers to sugar subunit and the rest in the labels refers to the position on that sugar residue. The inter-glycosidic NOESY cross-peaks are indicated by *red circles* in (C). The TOCSY spectrum was collected with a mixing time of 103 ms, and the NOESY spectrum was acquired with a mixing time of 160 ms. The structure and position labeling of the 8-epi-legionaminic acid are indicated in (A).

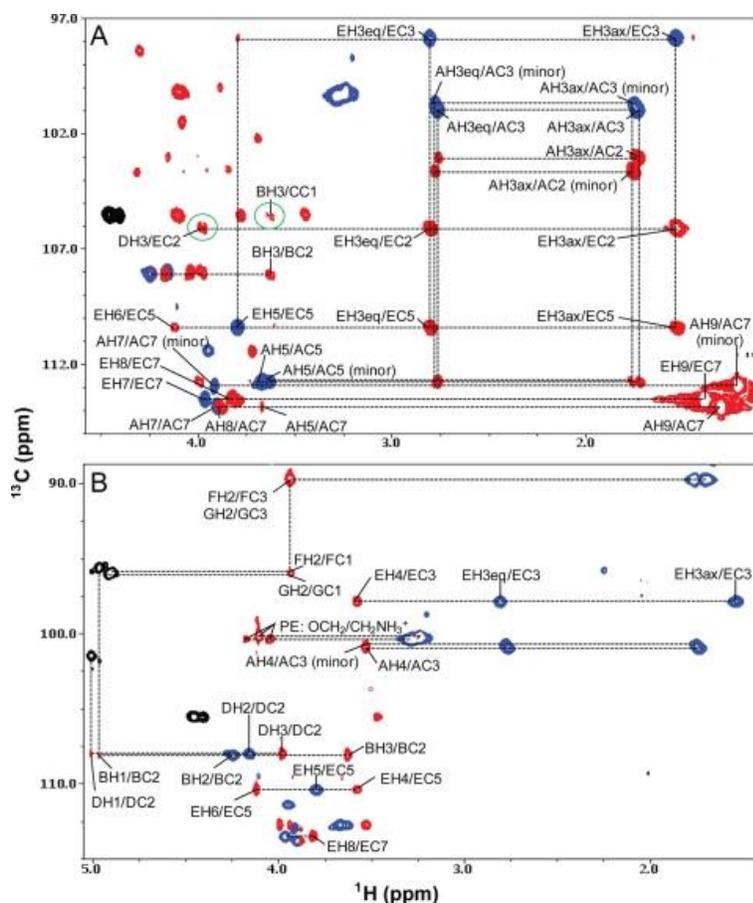


FIG 3-21 Overlays of HMQC and HMBC heteronuclear NMR spectra (A) and HMQC and H2BC heteronuclear NMR spectra (B) of the *V. fischeri* O-antigen

oligosaccharide. These data illustrate the assignments of the spin systems for the sugar residues

A and E of 8-epi-legionaminic acids, as well as for the sugar residues B, D, F, and G. In

both (A) and (B), the HMQC spectra are shown in *black* for the positive peaks and *blue* for the

folded negative peaks. The HMBC spectrum in (A) and the H2BC spectrum in (B) are shown

in *red*. (A), some of the cross-peaks belonging to the spin systems of the residues A and E are

labeled, and the observed inter-glycosidic cross-peaks are also indicated by *green circles*.

The *dotted lines* indicate the spin connectivities. The 1st letter in the labels refers to sugar

subunit, and the rest in the labels refers to the position on that sugar residue.

AcNH group) gave a strong HMBC cross-peak to a methyl group H9 (AH9/AC7) and medium/weak HMBC cross-peaks to H8 and H5 (AH8/AC7 and AH5/AC7, respectively) (Fig. 3-21A), whereas the 7th carbon of residue E gave a strong HMBC cross-peak to another methyl group H9 (EH9/EC7) and medium HMBC cross-peaks to H8 (EH8/EC7) (Fig. 3-21A). The 5th carbon of residue E gave $^2J_{\text{CH}}$ couplings to H4 (EH4/EC5) and H6 (EH6/EC5) in the H2BC spectrum (Fig. 3-21B), $^3J_{\text{CH}}$ couplings to H3_{eq} and H3_{ax} (EH3_{eq}/EC5 and EH3_{ax}/EC5, respectively), and $^2J_{\text{CH}}$ couplings to H6 (EH6/EC5) in the HMBC spectrum (Fig. 3-21A). Furthermore, the H3 protons of the residues A and E showed HMBC cross-peaks to the carbon atoms of C1 (data not shown), C2 (EH3_{ax}/EC2, EH3_{eq}/EC2, AH3_{ax}/AC2), and C5 (EH3_{ax}/EC5) (Fig. 3-21A). The 3rd carbon of the residues A and E clearly gave $^2J_{\text{CH}}$ H2BC connectivity to H4 proton (EH4/EC3 and AH4/AC3) (Fig. 3-21B).

Interestingly, during the NMR assignments, another set of peaks corresponding to 8-epi-legionaminic acid was identified (designated as residue A minor form, in ~40% quantity). As shown clearly in Fig. 3-21, residue A minor form exhibited slightly different chemical shifts from the major form of residue A. For example, the cross-peaks AH3_{ax}/AC3 (minor) and AH3_{ax}/AC2 (minor) were clearly shifted away from AH3_{ax}/AC3 and AH3_{ax}/AC2, respectively (Fig. 3-21A). Moreover, the detected HMBC cross-peak of AH9/AC7 (minor) showed a very large chemical shift change from the cross-peak of AH9/AC7 (Fig. 3-21A). Most importantly, the ^{13}C chemical shift of the methyl group A9 (minor) was very different and shifted downfield by 4.53 ppm from the corresponding peak A9 of the major form (Fig. 3-21A). The ^1H and ^{13}C chemical shifts of the 8-epi-legionaminic acid residue A minor form assigned in this study were

in good agreement with those reported previously for the chemically synthesized 8-epi-legionaminic acid (45).

Assignment of Other Sugar Residues. Fig. 3-22A-B shows the assignments in the methyl region and the C₁H and CHNHAc region of the HMQC spectrum of the O-antigen oligosaccharide of *V. fischeri*. Parts of the spin systems of residue B of GalNAc and the residues C and D of FucNAc are shown in Fig. 3-21 and 3-22. The complete assignments of the O-antigen residues of *V. fischeri* are listed in Table 3-2.

O-antigen Structure. The sugar residues F and G of Yer are linked via α 1–4 and α 1–8 glycosidic linkages to the 4th and 8th carbons of residue E of 8-epi-legionaminic acid as evidenced by the observed HMBC cross-peaks FH1/EC4 and GH1/EC8, respectively (Fig. 3-22C). These glycosidic linkage assignments were consistent with the observed strong NOEs between the H1 proton of residue F and the CH₂ group (at the 3rd position) of residue E (FH1/EH3_{ax} and FH1/EH3_{eq}) (Fig. 3-20C), and between the H1 proton of residue G and the H9 methyl protons of residue E (GH1/EH9) (Fig. 3-20C).

The presence of the HMBC cross-peak BH1/AC8 (Fig. 3-22C) demonstrated that residue B of GalNAc is linked via an α 1–8 glycosidic linkage to the 8th carbon of residue A of 8-epi-legionaminic acid. This assignment was also supported by the observed strong NOE between the H1 proton of residue B and the H9 methyl proton of residue A (BH1/AH9) (Fig. 3-20C).

Residue E of 8-epi-legionaminic acid is linked via an α 2–3 glycosidic linkage to the 3rd carbon of residue D of FucNAc as revealed by the detected DH3/EC2 cross-peak in the HMBC spectrum (Fig. 3-21A). The observed BH3/CC1 cross-peak in the HMBC spectrum demonstrated that residue C of FucNAc is connected to residue B via a β 1–3 glycosidic linkage (Fig. 3-21A). The observed HMBC cross-peak DH1/CC3 at a lower contour level indicated that residue D is

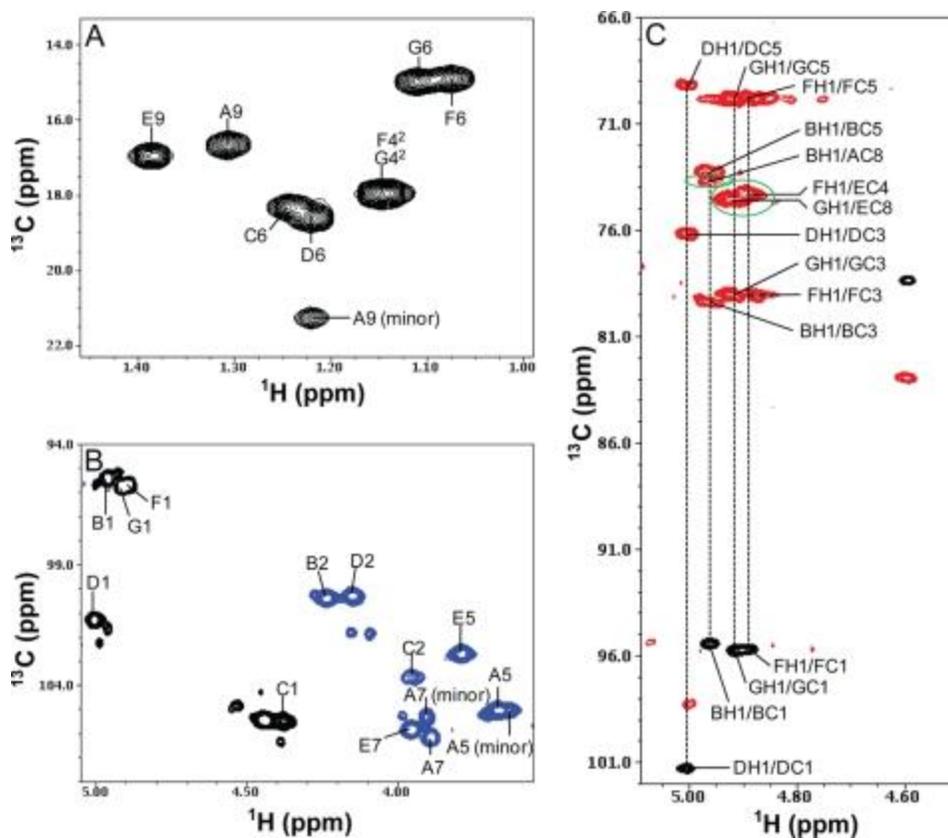


FIG 3-22 HMQC spectrum in the methyl region (A) and the C₁H and CHNHAc region (B), and the overlay of HMQC and HMBC heteronuclear NMR spectra (C) of the wild-type *V. fischeri* O-antigen oligosaccharide. The HMQC spectrum is shown in *black* for the positive peaks and *blue* for the folded negative peaks. The HMBC spectrum is shown in *red*. (C) illustrates the interglycosidic linkage assignments between residues A and B, residues E and F, and residues E and G. The cross-peaks are labeled, and the interglycosidic cross-peaks are *circled in green*. The 1st letter in the labels refers to sugar subunit, and the rest in the labels refers to the position on that sugar residue.

connected to residue C via a β 1–3 glycosidic linkage. This assignment was consistent with the observed strong inter-residue NOE between the H1 proton of residue D and the H3 proton of residue C.

Residues A and E had chemical shifts of 16.71 and 16.98 ppm for the C9 carbon, respectively (Table 3-2). However, residue A minor form had a chemical shift of 21.24 ppm for the same carbon. This large chemical shift perturbation ($\Delta\delta = \sim 4.5$ ppm) likely reflects the differences in the attachment at the 8th carbon. The 21.24 ppm of the C9 carbon of residue A minor form was consistent with having no attachment at the 8th carbon because the chemically synthesized 8-epi-legionaminic acid monosaccharide also has a C9 chemical shift at ~ 20.0 ppm (45). Therefore, the NMR data indicated that the oligosaccharide sample from the wild-type *V. fischeri* contains a major form ($\sim 60\%$) consisting of the O-antigen and the core component, and a minor form ($\sim 40\%$) composed of 8-epi-legionaminic acid (residue A) linked to the core. The assignments of sugar residues to the core or O-antigen components were based on differences observed, using NMR, MS, and GC-MS, in the wild-type and *waaL* LPS samples. Because attachment of the O-antigen to the LPS core is dependent on the O-antigen ligase WaaL, mutants that lack a functional O-antigen ligase would express an LPS consisting of only the core. The detection, by NMR and MS, of GalNAc (residue B) and one 8-epi-legionaminic acid component (residue A) in the *waaL* sample indicated that these were components of the core. Therefore, these data suggested that the O-antigen ligase of *V. fischeri* ligates the O-antigen to the 3rd carbon of residue B and that the O-antigen is composed of residues C–G. The final oligosaccharide structure of the O-antigen as well as some of the core components are shown in Fig. 3-23.

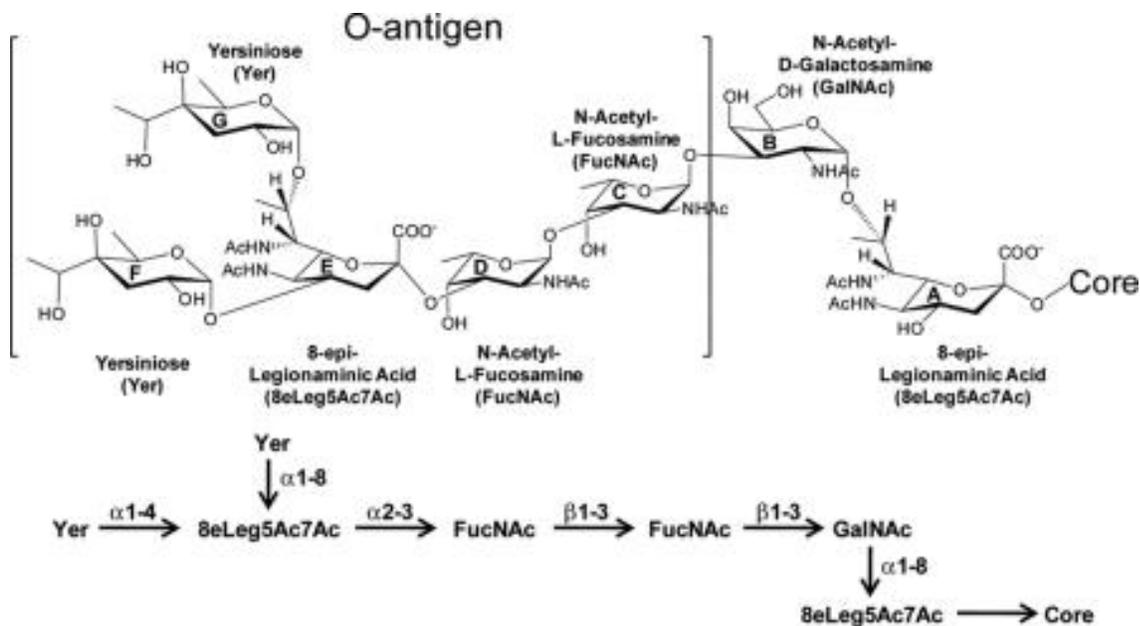


FIG 3-23 Determined structure of the O-antigen plus some of the core components of *V. fischeri*. The sugar subunits are labeled A–G. The linear representation of the oligosaccharide structure is shown in the *lower panel*.

PS from the *waaL* Mutant. Extensive NMR data were also collected on the PS sample from the *V. fischeri waaL* O-antigen ligase mutant. Figs. 3-24 and 3-25 show the one-dimensional ^{31}P NMR spectra and two-dimensional $^{31}\text{P}/^1\text{H}$ COSY spectra, respectively, of the PS samples from the wild-type and *waaL* mutant strains of *V. fischeri*. The one-dimensional ^{31}P NMR spectra of the wild-type and mutant samples were comparable, and both had two peaks with similar chemical shifts, although the peak centered at -0.8 ppm was much more heterogeneous in the mutant sample than the wild-type sample. The $^{31}\text{P}/^1\text{H}$ COSY spectra indicated that there were two PEA groups with their phosphate groups showing *J*-couplings to the ethanolamine protons as well as to the core residues in both the wild-type and mutant samples (Fig. 3-25). Therefore, these data clearly indicated that the PEA groups were attached to the core component. However, because of the sample heterogeneity, further details of the core structure were not determined in this NMR study.

DISCUSSION

In this study we compared the LPS structures of *V. fischeri* wild-type strain ES114 to the LPS O-antigen ligase mutant, ES114 *waaL*. Because an O-antigen ligase mutant cannot ligate the O-antigen to the LPS core, the LPS structure present in the *waaL* mutant strain should consist of only the core structure linked to the lipid A. By comparing the LPS structures of the *waaL* mutant to the wild type, it is possible to elicit which components of the LPS are part of the core structure and which are constituents of the O-antigen.

The core structure of *V. fischeri* strain ES114 was determined utilizing GC-MS, MALDI-MS, and NMR analyses of both the wild-type and *waaL* mutant strains. These data showed that the core consists of two PEA, one phosphate, two Hep (one LD-Hep and one DD-Hep), two Glc,

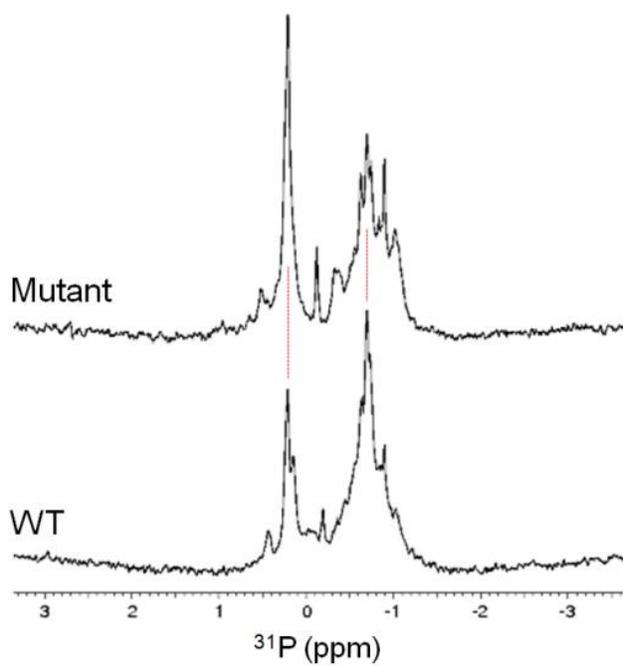


FIG 3-24 ^{31}P 1D NMR spectra of the ES114 wild-type and ES114 waaL mutant *V. fischeri* PS samples.

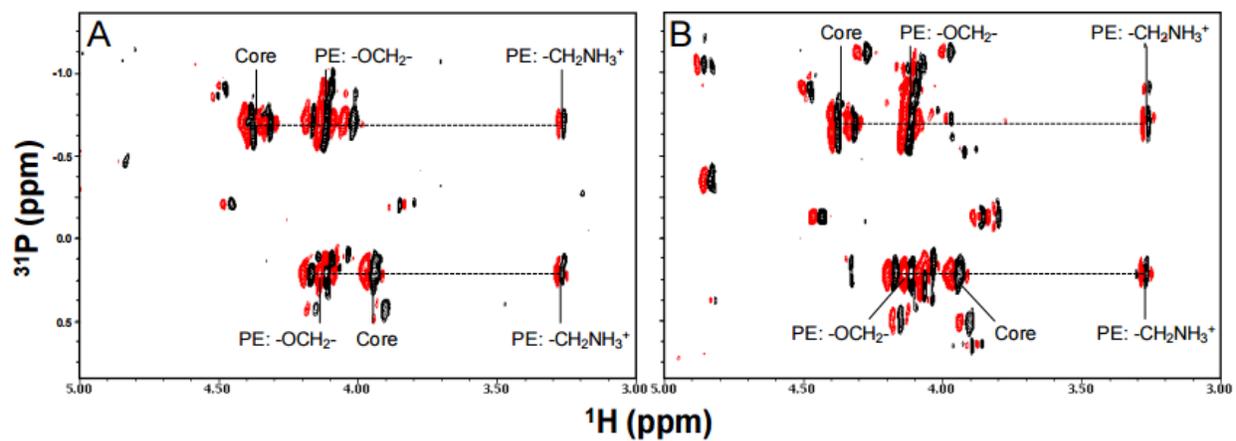


FIG 3-25 $^{31}\text{P}/^1\text{H}$ COSY spectra of the *V. fischeri* ES114 wild-type (A) and the ES114 waaL mutant (B) PS samples. The resonances derived from the PEA and the core components are indicated.

one Kdo, one GalNAc, and one 8-epi-legionaminic acid. Most of the components (Hep, Glc, Kdo, GalNAc, PEA, and phosphate) seen in the core are prototypical constituents of core components found in numerous Gram-negative bacteria, including other *Vibrio* spp. (6, 46-52). In addition to these typical components, the *V. fischeri* core also contains the unusual sugar 8-epi-legionaminic acid. Derivatives of this sugar have also been observed in the LPS, capsule, and/or flagellin from *Vibrio vulnificus*, *Rhizobium* sp. NGR234, *Aeromonas caviae*, *Helicobacter pylori*, and *Campylobacter jejuni* (44, 49, 53-56). Moreover, glycosylation of the flagella of *H. pylori* and *C. jejuni* with derivatives of legionaminic acid (57) was shown to be necessary for proper motility (55, 58).

Despite the heterogeneity of the oligosaccharide samples that prevented us from completely determining the core structure, NMR analyses of the wild-type PS sample did show that the core components of GalNAc and 8-epi-legionaminic acid are connected via an α 1-8 linkage (Fig. 3-23). MALDI-MSⁿ analyses suggested that Glc is the next sugar in the core structure, where it is linked to 8-epi-legionaminic acid at the reducing terminus. Because the core structures from a number of Gram-negative organisms have been determined, some predictions regarding the structural layout of the remainder of the core can be made. It is well established that Kdo links the LPS core to the lipid A. Previous studies from *Vibrio* spp. demonstrated that the LPS contained a single phosphorylated Kdo (50, 59, 60). The presence of a single phosphorylated Kdo has also been reported in the LPS from other Gram-negative bacteria, such as *Bordetella pertussis*, *Haemophilus ducreyi*, and *Haemophilus influenzae* (28, 50, 61, 62), suggesting that the phosphate we detected is attached to the Kdo in the core. Further information is needed to elucidate the structural arrangement of the remaining core components.

Using NMR, MS, and GC-MS, the *V. fischeri* O-antigen was determined to be a pentasaccharide consisting of two FucNAc, one 8-epi-legionaminic acid, and two Yer residues (Fig. 3-23). The two FucNAc residues are consecutively linked and serve as the attachment point between the core components and the remainder of the O-antigen. The two Yer residues are the terminal sugars of the O-antigen structure and are linked to the 4th and 8th positions of 8-epi-legionaminic acid. Thus, the O-antigen contains two unusual sugars, 8-epi-legionaminic acid, which was also present in the core, and Yer. Yer has been previously found in the O-antigen from other bacteria, including *Burkholderia brasiliensis*, *Legionella* spp., *Rhizobium* spp., and *Yersinia* spp (42, 44, 63, 64). Both MS and GC-MS analyses demonstrated that neither FucNAc nor Yer was found in any of the spectra from the *waaL* LPS indicating that FucNAc and Yer are O-antigen-specific components in *V. fischeri* ES114 LPS.

Analysis of the LPS samples from both the wild-type and *waaL* mutant by SDS-PAGE showed that the typical ladder-like banding pattern traditionally associated with the O-antigen repeat unit was absent in both *V. fischeri* strains. Instead, the wild-type LPS appeared as one predominant band containing core plus O-antigen, whereas the *waaL* LPS appeared as a faster migrating, lower molecular weight band containing only the core. The absence of the O-antigen ladder from the LPS, when examined by SDS-PAGE, has been previously observed for *V. fischeri* strain HMK as well as other *Vibrio* spp. (7, 50, 61, 65-67). In addition, both NMR and MS analyses (data not shown) suggested that the LPS from *V. fischeri* strain ES114 lacks repeating units of the O-antigen. It is unclear why some of the *Vibrio* spp. lack O-antigen repeat units. It is possible that these bacteria lack the O-antigen polymerase, Wzy; however, because the polymerase genes are not well conserved across bacterial species, identifying a *wzy* homologue has proven difficult.

Prior to this study, the role that the *V. fischeri* O-antigen plays in colonization of the squid *E. scolopes* was unknown. In this study, we examined the colonization properties of the ES114 wild-type, *waaL* mutant, and *waaL* complement strains. Our studies demonstrated that the *waaL* mutant strain of *V. fischeri* has an animal colonization defect that can be restored through complementation. When luminescence levels were low in animals colonized by the *waaL* mutant (Fig. 3-3A), it was due to decreased colonization by the symbiont, as visualized by confocal microscopy (Fig. 3-3C) or verified by plating animal light organs for CFUs. These data are consistent with the established finding that luminescence levels correlate with CFUs, with the exception of colonization by bacterial mutants defective in light production (68).

Defects in *V. fischeri* colonization of *E. scolopes* may be described in terms of initiation, accommodation, or persistence (69). Mutants in these three behaviors consequently fail to establish symbiosis with the animal (or are delayed in that establishment), fail to fully colonize the light organ at wild-type levels, or fail to persist indefinitely in the light organ, respectively. We detected significant differences in wild-type and *waaL* mutant luminescence 24 h into symbiosis (Fig. 3-3A). At 44 h, the difference remained significant, but all animals, including *waaL* mutant-colonized animals, had detectable luminescence (Fig. 3-4). Our results are consistent with a defect in the initiation phase. In animals exposed to both wild-type and *waaL* mutant bacteria, the wild-type bacteria appear to establish early, during the period when the *waaL* mutant is defective, and subsequently remain dominant (Fig. 3-6). From our experiments, we cannot rule out the possibility that the *waaL* mutant's defect in animal colonization is simply a result of a motility defect, as luminescence at 24 h was not significantly different between *waaL* mutant-colonized animals and those colonized by the *flaD* mutant (Fig. 3-3A). The *flaD* mutant is thought to be a “pure” motility mutant, without a predicted LPS

biogenesis defect (13). The precise structural effect of the *flaD* mutation is not known, but it could result in fewer or shorter flagellin filaments, without altering flagellar rotation (40, 70).

These results prompt the question of why a mutation affecting a surface antigen affects bacterial motility. In fact, mutation of *waaL* has been shown in several studies to cause defects in swarming behavior, without the swimming motility defects seen in our study of *V. fischeri* (Fig. 3-2A). For *Salmonella enterica* serovar Typhimurium, it has been suggested that the swarming phenotype is due to surfactant properties of the O-antigen (71). It is notable that whatever LPS was discharged from wild-type *V. fischeri* in our competition experiment did not rescue the *waaL* mutant's colonization defect (Fig. 3-6). In *Proteus mirabilis*, a mutation in *waaL* affects regulation of flagellar gene synthesis (*flhDC* genes), perhaps through the Rcs system. The resulting swarming defect can be reversed by overexpressing *flhDC*, without restoring the O-antigen itself (72). The Rcs system and *flhDC* have also been implicated in swarming and swimming deficiencies associated with mutation of LPS core synthesis genes in *E. coli* (73). Apart from flagellum-mediated motility, defects in O-antigen synthesis have been associated with reduction of social “gliding” motility in *Myxococcus xanthus* (74). The relevance of these data to the *V. fischeri* swimming defect is unclear at this time, as *V. fischeri* has not been reported to utilize swarming or gliding motility, and the data from *E. coli* mentioned above involve other portions of the LPS molecule than the O antigen. Finally, because *V. fischeri* has a sheathed flagellum (68), we speculate that the *waaL* mutant's altered LPS structure may affect flagellar function. Further experiments are necessary to verify this hypothesis.

ACKNOWLEDGEMENTS

Mass spectrometric instrumentation for this study was provided by the mass spectrometry core facility at the Buck Institute for Research on Aging. This work was supported, in whole or in part, by National Institutes of Health Grants PO1 AI44642 and RO1 AI24616 (to B. W. G. and M. A. A.), R01 AI50661 (to M. M.-N.), and R01 RR 12294 and IOS 0817232 (to E. G. R. and M. M.-N.). This work was also supported by University of Wisconsin Medical Scientist Training Program (to the M. M.-N. laboratory).

REFERENCES

1. **McFall-Ngai MJ, Ruby EG.** 1991. Symbiont recognition and subsequent morphogenesis as early events in an animal-bacterial mutualism. *Science* **254**:1491-1494.
2. **Nyholm SV, Stabb EV, Ruby EG, McFall-Ngai MJ.** 2000. Establishment of an animal-bacterial association: recruiting symbiotic vibrios from the environment. *Proc Natl Acad Sci U S A* **97**:10231-10235.
3. **Nyholm SV, McFall-Ngai MJ.** 2004. The winnowing: establishing the squid-vibrio symbiosis. *Nat Rev Microbiol* **2**:632-642.
4. **Foster JS, Apicella MA, McFall-Ngai MJ.** 2000. *Vibrio fischeri* lipopolysaccharide induces developmental apoptosis, but not complete morphogenesis, of the *Euprymna scolopes* symbiotic light organ. *Dev Biol* **226**:242-254.
5. **Koropatnick TA, Engle JT, Apicella MA, Stabb EV, Goldman WE, McFall-Ngai MJ.** 2004. Microbial factor-mediated development in a host-bacterial mutualism. *Science* **306**:1186-1188.
6. **Phillips NJ, Adin DM, Stabb EV, McFall-Ngai MJ, Apicella MA, Gibson BW.** 2011. The lipid A from *Vibrio fischeri* lipopolysaccharide: a unique structure bearing a phosphoglycerol moiety. *J Biol Chem* **286**:21203-21219.
7. **Pupo E, Phillips NJ, Gibson BW, Apicella MA, Hardy E.** 2004. Matrix-assisted laser desorption/ionization-time of flight-mass spectrometry of lipopolysaccharide species separated by slab-polyacrylamide gel electrophoresis: high-resolution separation and molecular weight determination of lipooligosaccharides from *Vibrio fischeri* strain HMK. *Electrophoresis* **25**:2156-2164.
8. **Abeyrathne PD, Lam JS.** 2007. WaaL of *Pseudomonas aeruginosa* utilizes ATP in in vitro ligation of O antigen onto lipid A-core. *Mol Microbiol* **65**:1345-1359.

9. **Abeyrathne PD, Daniels C, Poon KK, Matewish MJ, Lam JS.** 2005. Functional characterization of WaaL, a ligase associated with linking O-antigen polysaccharide to the core of *Pseudomonas aeruginosa* lipopolysaccharide. *J Bacteriol* **187**:3002-3012.
10. **Stabb EV, Reich KA, Ruby EG.** 2001. *Vibrio fischeri* genes hvnA and hvnB encode secreted NAD(+)-glycohydrolases. *J Bacteriol* **183**:309-317.
11. **Dunlap PV.** 1989. Regulation of luminescence by cyclic AMP in cya-like and crp-like mutants of *Vibrio fischeri*. *J Bacteriol* **171**:1199-1202.
12. **Boettcher KJ, Ruby EG.** 1990. Depressed light emission by symbiotic *Vibrio fischeri* of the sepiolid squid *Euprymna scolopes*. *J Bacteriol* **172**:3701-3706.
13. **Brennan CA, Mandel MJ, Gyllborg MC, Thomsgard KA, Ruby EG.** 2013. Genetic determinants of swimming motility in the squid light-organ symbiont *Vibrio fischeri*. *Microbiology Open* **2**:576-594.
14. **Lyell NL, Dunn AK, Bose JL, Vescovi SL, Stabb EV.** 2008. Effective mutagenesis of *Vibrio fischeri* by using hyperactive mini-Tn5 derivatives. *Appl Environ Microbiol* **74**:7059-7063.
15. **O'Toole GA, Pratt LA, Watnick PI, Newman DK, Weaver VB, Kolter R.** 1999. Genetic approaches to study of biofilms. *Methods Enzymol* **310**:91-109.
16. **Boettcher K, Ruby E.** 1994. Occurrence of plasmid dna in the sepiolid squid symbiont *Vibrio fischeri*. *Current Microbiology* **29**:279-286.
17. **Hanahan D.** 1983. Studies on transformation of *Escherichia coli* with plasmids. *J Mol Biol* **166**:557-580.
18. **Stabb EV, Ruby EG.** 2002. RP4-based plasmids for conjugation between *Escherichia coli* and members of the Vibrionaceae. *Methods Enzymol* **358**:413-426.
19. **Dunn AK, Millikan DS, Adin DM, Bose JL, Stabb EV.** 2006. New rfp- and pES213-derived tools for analyzing symbiotic *Vibrio fischeri* reveal patterns of infection and lux expression in situ. *Appl Environ Microbiol* **72**:802-810.
20. **Mandel MJ, Wollenberg MS, Stabb EV, Visick KL, Ruby EG.** 2009. A single regulatory gene is sufficient to alter bacterial host range. *Nature* **458**:215-218.
21. **Montgomery MK, McFall-Ngai MJ.** 1998. Late postembryonic development of the symbiotic light organ of *Euprymna scolopes* (Cephalopoda: Sepiolidae). *Biol Bull* **195**:326-336.
22. **Altura MA, Stabb E, Goldman W, Apicella M, McFall-Ngai MJ.** 2011. Attenuation of host NO production by MAMPs potentiates development of the host in the squid-vibrio symbiosis. *Cell Microbiol* **13**:527-537.
23. **Wollenberg MS, Ruby EG.** 2012. Phylogeny and fitness of *Vibrio fischeri* from the light organs of *Euprymna scolopes* in two Oahu, Hawaii populations. *ISME J* **6**:352-362.

24. **Lee PN, McFall-Ngai MJ, Callaerts P, de Couet HG.** 2009. Confocal immunocytochemistry of embryonic and juvenile Hawaiian bobtail squid (*Euprymna scolopes*) tissues. Cold Spring Harb Protoc **2009**:pdb.prot5320.
25. **Apicella MA.** 2008. Isolation and characterization of lipopolysaccharides. Methods Mol Biol **431**:3-13.
26. **Lesse AJ, Campagnari AA, Bittner WE, Apicella MA.** 1990. Increased resolution of lipopolysaccharides and lipooligosaccharides utilizing tricine-sodium dodecyl sulfate-polyacrylamide gel electrophoresis. J Immunol Methods **126**:109-117.
27. **Allen S, Zaleski A, Johnston JW, Gibson BW, Apicella MA.** 2005. Novel sialic acid transporter of *Haemophilus influenzae*. Infect Immun **73**:5291-5300.
28. **Phillips NJ, Apicella MA, Griffiss JM, Gibson BW.** 1992. Structural characterization of the cell surface lipooligosaccharides from a nontypable strain of *Haemophilus influenzae*. Biochemistry **31**:4515-4526.
29. **Rance M, Sørensen OW, Bodenhausen G, Wagner G, Ernst RR, Wüthrich K.** 1983. Improved spectral resolution in cosy 1H NMR spectra of proteins via double quantum filtering. Biochem Biophys Res Commun **117**:479-485.
30. **Bax A, Davis DG.** 1985. MLEV-17-based two-dimensional homonuclear magnetization transfer spectroscopy. J Magn Resn **65**:355-360.
31. **K. W.** 1986. NMR of Proteins and Nucleic Acids. John Wiley & Sons, Inc., New York.
32. **Bax A, Summers MF.** 1986. Proton and carbon-13 assignments from sensitivity-enhanced detection of heteronuclear multiple-bond connectivity by 2D multiple quantum NMR. J Am Chem Soc **108**:2093-2094.
33. **Nyberg NT, Duus JO, Sørensen OW.** 2005. Heteronuclear two-bond correlation: suppressing heteronuclear three-bond or higher NMR correlations while enhancing two-bond correlations even for vanishing 2J(CH). J Am Chem Soc **127**:6154-6155.
34. **Petersen BO, Vinogradov E, Kay W, Würtz P, Nyberg NT, Duus J, Sørensen OW.** 2006. H2BC: a new technique for NMR analysis of complex carbohydrates. Carbohydr Res **341**:550-556.
35. **Sato H, Kajihara Y.** 2005. An unambiguous assignment method by 2D selective-TOCSY-HSQC and selective-TOCSY-DQFCOSY and structural analysis by selective-TOCSY-NOESY experiments of a biantennary undecasaccharide. Carbohydr Res **340**:469-479.
36. **Chary KV, Rastogi VK.** 1993. An efficient 2D NMR technique HELCO for heteronuclear [³¹P-¹H] long-range correlation. J Magn Reson Series B102, 81-83 **B102**:81-83.
37. **Olsson U, Lycknert K, Stenutz R, Weintraub A, Widmalm G.** 2005. Structural analysis of the O-antigen polysaccharide from *Escherichia coli* O152. Carbohydr Res **340**:167-171.

38. **Delaglio F, Grzesiek S, Vuister GW, Zhu G, Pfeifer J, Bax A.** 1995. NMRPipe: a multidimensional spectral processing system based on UNIX pipes. *J Biomol NMR* **6**:277-293.
39. **Johnson BA, Blevins RA.** 1994. NMR View: A computer program for the visualization and analysis of NMR data. *J Biomol NMR* **4**:603–614.
40. **Millikan DS, Ruby EG.** 2004. *Vibrio fischeri* flagellin A is essential for normal motility and for symbiotic competence during initial squid light organ colonization. *J Bacteriol* **186**:4315-4325.
41. **Studer SV, Mandel MJ, Ruby EG.** 2008. AinS quorum sensing regulates the *Vibrio fischeri* acetate switch. *J Bacteriol* **190**:5915-5923.
42. **Mattos KA, Todeschini AR, Heise N, Jones C, Previato JO, Mendonça-Previato L.** 2005. Nitrogen-fixing bacterium *Burkholderia brasiliensis* produces a novel yersinirose A-containing O-polysaccharide. *Glycobiology* **15**:313-321.
43. **Zubkov VA, Gorshkova RP, Ovodov YS, Sviridov AF, Shashkov AS.** 1992. Synthesis of 3,6-dideoxy-4-C-(4(1)-hydroxyethyl)hexopyranoses (yersinioses) from 1,6-anhydro-beta-D-glucopyranose. *Carbohydr Res* **225**:189-207.
44. **Le Quéré AJ, Deakin WJ, Schmeisser C, Carlson RW, Streit WR, Broughton WJ, Forsberg LS.** 2006. Structural characterization of a K-antigen capsular polysaccharide essential for normal symbiotic infection in *Rhizobium* sp. NGR234: deletion of the rkpMNO locus prevents synthesis of 5,7-diacetamido-3,5,7,9-tetradexy-non-2-ulosonic acid. *J Biol Chem* **281**:28981-28992.
45. **Tsvetkov YE, Shashkov AS, Knirel YA, Zähringer U.** 2001. Synthesis and NMR spectroscopy of nine stereoisomeric 5,7-diacetamido-3,5,7,9-tetradexynon-2-ulosonic acids. *Carbohydr Res* **335**:221-243.
46. **Melaugh W, Phillips NJ, Campagnari AA, Tullius MV, Gibson BW.** 1994. Structure of the major oligosaccharide from the lipooligosaccharide of *Haemophilus ducreyi* strain 35000 and evidence for additional glycoforms. *Biochemistry* **33**:13070-13078.
47. **Banoub JH, El Aneed A, Cohen AM, Joly N.** 2010. Structural investigation of bacterial lipopolysaccharides by mass spectrometry and tandem mass spectrometry. *Mass Spectrom Rev* **29**:606-650.
48. **King JD, Kocíncová D, Westman EL, Lam JS.** 2009. Review: Lipopolysaccharide biosynthesis in *Pseudomonas aeruginosa*. *Innate Immun* **15**:261-312.
49. **Vinogradov E, Wilde C, Anderson EM, Nakhamchik A, Lam JS, Rowe-Magnus DA.** 2009. Structure of the lipopolysaccharide core of *Vibrio vulnificus* type strain 27562. *Carbohydr Res* **344**:484-490.
50. **Chatterjee SN, Chaudhuri K.** 2003. Lipopolysaccharides of *Vibrio cholerae*. I. Physical and chemical characterization. *Biochim Biophys Acta* **1639**:65-79.

51. **Hashii N, Isshiki Y, Iguchi T, Kondo S.** 2003. Structural analysis of the carbohydrate backbone of *Vibrio parahaemolyticus* O2 lipopolysaccharides. *Carbohydr Res* **338**:1063-1071.
52. **Hashii N, Isshiki Y, Iguchi T, Kondo S.** 2003. Structural characterization of the carbohydrate backbone of the lipopolysaccharide of *Vibrio parahaemolyticus* O-untypable strain KX-V212 isolated from a patient. *Carbohydr Res* **338**:2711-2719.
53. **Tabei SM, Hitchen PG, Day-Williams MJ, Merino S, Vart R, Pang PC, Horsburgh GJ, Viches S, Wilhelms M, Tomás JM, Dell A, Shaw JG.** 2009. An *Aeromonas caviae* genomic island is required for both O-antigen lipopolysaccharide biosynthesis and flagellin glycosylation. *J Bacteriol* **191**:2851-2863.
54. **Power PM, Jennings MP.** 2003. The genetics of glycosylation in Gram-negative bacteria. *FEMS Microbiol Lett* **218**:211-222.
55. **Schirm M, Soo EC, Aubry AJ, Austin J, Thibault P, Logan SM.** 2003. Structural, genetic and functional characterization of the flagellin glycosylation process in *Helicobacter pylori*. *Mol Microbiol* **48**:1579-1592.
56. **Thibault P, Logan SM, Kelly JF, Brisson JR, Ewing CP, Trust TJ, Guerry P.** 2001. Identification of the carbohydrate moieties and glycosylation motifs in *Campylobacter jejuni* flagellin. *J Biol Chem* **276**:34862-34870.
57. **McNally DJ, Aubry AJ, Hui JP, Khieu NH, Whitfield D, Ewing CP, Guerry P, Brisson JR, Logan SM, Soo EC.** 2007. Targeted metabolomics analysis of *Campylobacter coli* VC167 reveals legionaminic acid derivatives as novel flagellar glycans. *J Biol Chem* **282**:14463-14475.
58. **Goon S, Kelly JF, Logan SM, Ewing CP, Guerry P.** 2003. Pseudaminic acid, the major modification on *Campylobacter flagellin*, is synthesized via the Cj1293 gene. *Mol Microbiol* **50**:659-671.
59. **Brade H.** 1985. Occurrence of 2-keto-deoxyoctonic acid 5-phosphate in lipopolysaccharides of *Vibrio cholerae* Ogawa and Inaba. *J Bacteriol* **161**:795-798.
60. **Caroff M, Brisson J, Martin A, Karibian D.** 2000. Structure of the *Bordetella pertussis* 1414 endotoxin. *FEBS Lett* **477**:8-14.
61. **Cox AD, Brisson JR, Thibault P, Perry MB.** 1997. Structural analysis of the lipopolysaccharide from *Vibrio cholerae* serotype O22. *Carbohydr Res* **304**:191-208.
62. **Phillips NJ, Apicella MA, Griffiss JM, Gibson BW.** 1993. Structural studies of the lipooligosaccharides from *Haemophilus influenzae* type b strain A2. *Biochemistry* **32**:2003-2012.
63. **Holst O.** 2003. Lipopolysaccharides of *Yersinia*. An overview. *Adv Exp Med Biol* **529**:219-228.
64. **Sonesson A, Jantzen E, Bryn K, Tangen T, Eng J, Zähringer U.** 1994. Composition of 2,3-dihydroxy fatty acid-containing lipopolysaccharides from *Legionella israelensis*, *Legionella maceachernii* and *Legionella micdadei*. *Microbiology* **140** (Pt 6):1261-1271.

65. **Han TJ, Chai TJ.** 1992. Electrophoretic and chemical characterization of lipopolysaccharides of *Vibrio parahaemolyticus*. *J Bacteriol* **174**:3140-3146.
66. **Cox AD, Brisson JR, Varma V, Perry MB.** 1996. Structural analysis of the lipopolysaccharide from *Vibrio cholerae* O139. *Carbohydr Res* **290**:43-58.
67. **Iguchi T, Kondo S, Hisatsune K.** 1995. *Vibrio parahaemolyticus* O serotypes from O1 to O13 all produce R-type lipopolysaccharide: SDS-PAGE and compositional sugar analysis. *FEMS Microbiol Lett* **130**:287-292.
68. **Ruby EG, Asato LM.** 1993. Growth and flagellation of *Vibrio fischeri* during initiation of the sepiolid squid light organ symbiosis. *Arch Microbiol* **159**:160-167.
69. **Ruby EG.** 1996. Lessons from a cooperative, bacterial-animal association: the *Vibrio fischeri-Euprymna scolopes* light organ symbiosis. *Annu Rev Microbiol* **50**:591-624.
70. **Tambalo DD, Bustard DE, Del Bel KL, Koval SF, Khan MF, Hynes MF.** 2010. Characterization and functional analysis of seven flagellin genes in *Rhizobium leguminosarum* bv. *viciae*. Characterization of *R. leguminosarum* flagellins. *BMC Microbiol* **10**:219.
71. **Toguchi A, Siano M, Burkart M, Harshey RM.** 2000. Genetics of swarming motility in *Salmonella enterica* serovar typhimurium: critical role for lipopolysaccharide. *J Bacteriol* **182**:6308-6321.
72. **Morgenstein RM, Clemmer KM, Rather PN.** 2010. Loss of the waaL O-antigen ligase prevents surface activation of the flagellar gene cascade in *Proteus mirabilis*. *J Bacteriol* **192**:3213-3221.
73. **Girgis HS, Liu Y, Ryu WS, Tavazoie S.** 2007. A comprehensive genetic characterization of bacterial motility. *PLoS Genet* **3**:1644-1660.
74. **Bowden MG, Kaplan HB.** 1998. The *Myxococcus xanthus* lipopolysaccharide O-antigen is required for social motility and multicellular development. *Mol Microbiol* **30**:275-284.

CHAPTER 4

Novel Toll-like Receptors in the *E. scolopes* Light Organ

PREFACE

The results presented in this chapter are in preparation for submission to *Developmental and Comparative Immunology*, as part of a publication on the subject of *E. scolopes* Toll-like receptors with Bethany Rader (Southern Illinois University) as lead author. A final author list has yet to be compiled.

Of the work in this chapter, candidate TLR sequences were obtained from unpublished RNA-Seq databases produced by Natacha Kremer and Silvia Moriano-Gutierrez. Animals were reared by Eric Koch. RACE-PCR was carried out by Julian Cagnazzo and Benjamin Krasity, and full-length sequences amplified by BK. The exception is *estlr4*, which was not successfully amplified by PCR and the sequence included herein is taken from the RNA-Seq database (SM-G). Protein structure analysis, phylogenetic analysis, and qRT-PCR were carried out by BK. The writing of this chapter is entirely the work of BK.

ABSTRACT Toll-like receptors may play a role in the developmental response of the Hawaiian bobtail squid *Euprymna scolopes* to colonization by its beneficial bacterial partner *Vibrio fischeri*. The sequences of five previously unknown *E. scolopes* Toll-like receptors are presented, and their diverse positions in the phylogeny of protostomes considered. Examination of expression levels of these *estlrs* in four-week-old animals indicates variation in expression by tissue. In one case, an *estlr* is up-regulated in response to colonization of the light organ by *V. fischeri*. These receptors likely have distinct roles in host immunity.

INTRODUCTION

In the mutualistic relationship between the Hawaiian bobtail squid *Euprymna scolopes* and the bioluminescent marine bacterium *Vibrio fischeri*, the bacterium colonizes a specialized structure termed the light organ, wherein it provides light as its contribution to the symbiosis (reviewed in (1)). Numerous host responses are induced by bacterial microbe-associated molecular patterns (MAMPs), including lipopolysaccharide (LPS) and peptidoglycan (reviewed in (2)). An example of these changes is apoptosis and regression of the superficial epithelium of the light organ, a process by which the bacteria modify the organ in which they reside, promoting the regression of superficial structures used in the initial recruitment of colonizing bacteria. This process is driven by LPS and peptidoglycan from *V. fischeri* (3, 4). Details of host detection of MAMPs are mostly unknown, but it is suspected that lipopolysaccharide-binding proteins (LBPs) and peptidoglycan recognition proteins (PGRPs) play a role in the initial detection of extracellular MAMPs (Chapter 2, (5)). Signaling may proceed through Toll-like receptors, important players in immunity in both vertebrates and invertebrates (6, 7). Transcriptional

evidence for one Toll-like receptor, herein referred to as EsTLR1, has previously been reported in *E. scolopes*, along with numerous transcripts of putative NF- κ B pathway components (5).

Both LBPs and PGRPs have been reported to promote signaling through Toll or Toll-like receptors, through different mechanisms. Mammalian LBP binds LPS and promotes its delivery to CD14 (8). CD14 delivers LPS to MD-2 and TLR4, promoting receptor dimerization and signaling through Myd88, leading to NF- κ B signaling (9, 10). *Drosophila* PGRP-SA has been implicated in promoting Pro-Spätzle cleavage, producing the Toll ligand Spätzle and thereby activating NF- κ B signaling (11). Which, if either, of these strategies is used in *E. scolopes* is not known.

Our knowledge of TLR genes present throughout the eumetazoans has recently expanded, and there are numerous cases in the Lophotrochozoa, the superphylum including mollusks such as *E. scolopes*, of TLR expression being regulated by bacterial or MAMP challenge (reviewed in Chapter 1 and (12)). Thus, *E. scolopes* TLR genes may be involved in the immune response, including to the symbiont *V. fischeri*. Still unknown in every case, however, is the preferred ligand of lophotrochozoan TLRs. *Crassostrea gigas* TLRs, among the few to be biochemically characterized, show little to no evidence of MAMP binding (13). Examination of the phylogenetic context of lophotrochozoan TLRs may inform strategies for evaluating candidate endogenous and exogenous TLR ligands.

Here, we report on the identification of transcripts for five additional TLR homologues in *E. scolopes*, to add to the one originally reported *E. scolopes* TLR (5). Each of these was detected in mature light organs, four weeks post-hatching. While a detailed developmental trajectory for expression of these *tlr* genes has not yet been described, these data are circumstantial evidence for an immune role of these TLRs, instead of or in addition to any

developmental role they may play in the maturation of the light organ. Expression of these transcripts was investigated in three tissues of four-week-old animals: light organs, gills and eyes. Their phylogenetic relationships with each other and with other species' TLRs were also investigated.

MATERIALS AND METHODS

General procedures. Adult *E. scolopes* animals were collected from the sand flats of Oahu, Hawaii and transported and maintained as described in previous publications (14). Hatchlings were collected and raised to four weeks of age as described previously (15) before collection and stabilization in RNAlater (Thermo Fisher, Waltham, MA).

Cloning and sequencing of *estlr* transcripts. Transcripts from an RNA-Seq library of four-week-old *E. scolopes* light organs (N. Kremer *et al.*, unpublished data) for which the top blast hit was a Toll or Toll-like receptor were evaluated for the features of Toll-like receptors with SMART (16). Promising transcripts with such features as Toll/Interleukin 1 Receptor (TIR) domains were selected for PCR amplification and sequencing. In some cases, further sequence information was obtained from an RNA-Seq library of 24-hour-old animals (S. Moriano-Gutierrez *et al.*, unpublished data). RNA-Seq library generation was based on reference (17). In some cases, the 5' end of the ORF was apparently not present in the library (as indicated by the lack of a signal peptide) and primers for 5' RACE were designed. When the full-length sequence was believed to be known from RNA-Seq databases and RACE, primers were designed to amplify the full-length transcript (Table 4-1).

RNA for sequencing work was prepared from the light organs and gills of five ~4-week-old animals. Light organs and gills were dissected from animals and stabilized in RNAlater.

Table 4-1: Primers for RACE and amplification of *estlr* transcripts

Gene	Primer name	Use	Primer Sequence
<i>estlr2</i>	TLR2RACE1	5' RACE	GTCCATGTAGGTCCCTGTGAGATTC
<i>estlr2</i>	TLR2RACE2	5' RACE, nested	AGCGAGGGTAGATATCCCCAAGATTC
<i>estlr2</i>	TLR2fullF1	Full-length ORF amplification	ATCGCAGTGGATTTTTCTGG
<i>estlr2</i>	TLR2fullR1	Full-length ORF amplification	TGACTTGAAACCGACCATGT
<i>estlr3</i>	TLR3fullF1	Full-length ORF amplification	TCCATGAATGATCTTGAAGGTG
<i>estlr3</i>	TLR3fullR1	Full-length ORF amplification	CACACACATAAGCATACAAATAAAT
<i>estlr3</i>	TLR3fullF2	Full-length ORF amplification, nested	TGGGGATCTGTTAATTGCTACG
<i>estlr3</i>	TLR3fullR2	Full-length ORF amplification, nested	AAAAGCAGAATACAGAAAGGAAATG
<i>estlr5</i>	TLR5RACE1	5' RACE	CTCAAGCCCTTCAAAGCAGTGAAGTC
<i>estlr5</i>	TLR5RACE2	5' RACE, nested	CGGGAAGGTGAGATTAAGGATCTGGTT
<i>estlr5</i>	TLR5fullF1	Full-length ORF amplification	AAATCATTTTACACGGAACAACC
<i>estlr5</i>	TLR5fullR1	Full-length ORF amplification	TTTGATTCCGATTCCGATT
<i>estlr5</i>	TLR5fullF2	Full-length ORF amplification, nested	TTTTACACGGAACAACCTCCTT
<i>estlr5</i>	TLR5fullR2	Full-length ORF amplification, nested	CGATTCCGATTTATCGTCCAAG
<i>estlr6</i>	TLR6RACE1	5' RACE	GTTGTTCTCTAGCGAGAGTTCGACCAG
<i>estlr6</i>	TLR6RACE2	5' RACE, nested	CGATCCTCAGGACATTTAGGTGCTTG
<i>estlr6</i>	TLR6fullF1	Full-length ORF amplification	TCCTAGAACAAAGCAAACCAAA
<i>estlr6</i>	TLR6fullR1	Full-length ORF amplification	TCAGCCGACTTGCAAGAGTAT
<i>estlr6</i>	TLR6fullF2	Full-length ORF amplification, nested	AGCGAAGCGTATCATCACATT
<i>estlr6</i>	TLR6fullR2	Full-length ORF amplification, nested	CAACGAAATAACCACCACCAC

RNA extractions were performed with an RNeasy kit (Qiagen, Venlo, Netherlands) and DNA removed with Turbo DNase (Thermo Fisher). Single-stranded cDNA was prepared from 400-600 ng of RNA with MMLV reverse transcriptase (Clontech, Mountain View, CA) using 5' CDS primers (5'- (T)₂₅VN-3'). Light organ RNA for 5'-RACE was processed with a GeneRacer kit (Thermo Fisher) before the reverse-transcription step. PCR (RACE or otherwise) was performed with Platinum Taq High Fidelity (Thermo Fisher). PCR products were ligated into the plasmid pCR4-TOPO and transformed into TOP10 competent cells (Thermo Fisher); plasmids were obtained with QiaPrep MiniPrep kit (Qiagen) and the inserts sequenced.

EsTLR structure and phylogenetic analysis. Leucine-rich repeats were identified with LRRfinder (18); insignificant hits (<95% confidence) are not shown. Other features of TLR proteins (signal peptide, transmembrane region and TIR domain) were evaluated by SMART (16), PFAM (19), and SignalP 4.1 (20). Sequences for comparison in lophotrochozoan and ecdysozoan species were selected from the NCBI protein database to provide a broad sampling across these superphyla. Species sampled include *Toxocara canis*, *Caenorhabditis elegans*, *Ixodes scapularis*, *Litopenaeus vannamei*, *Drosophila melanogaster*, *Crassostrea gigas*, *Mytilus galloprovincialis*, *Lottia gigantea*, *Aplysia californica*, *Biomphalaria glabrata*, *Pinctada martensii*, *Haliotis discus discus*, *Azumapecten farreri*, *Hirudo medicinalis*, *Eisenia andrei*, *Helobdella robusta*, and *Capitella teleta*. Not all of these are annotated as Toll-like receptors in NCBI, but all are predicted to include key features of Toll-like receptors, including a region of LRRs, a transmembrane domain, and TIR domain. Human TLR4 was chosen as an outgroup to root the tree. Sequences selected are given in Table 4-2; EsTLR5* is the hypothesized “intact” version (Fig. 4-4C) rather than the truncated version indicated by this study’s PCR (Fig. 4-4B). Sequence alignment and phylogenetic tree construction were

Table 4-2: Accession numbers for Toll and TLR protein sequences used in this study

Species	Protein name per NCBI ¹	Shortened name	NCBI accession number
<i>Aplysia californica</i>	PREDICTED: protein toll-like	Ac_Tolllike	XP_005100757.1
<i>Aplysia californica</i>	PREDICTED: toll-like receptor 4-like	Ac_TLR4like	XP_005104953.1
<i>Aplysia californica</i>	PREDICTED: toll-like receptor 13-like	Ac_TLR13like	XP_005101868.1
<i>Azumapecten farreri</i>	Toll receptor	Af_Toll	ABC73693.1
<i>Biomphalaria glabrata</i>	toll-like receptor	Bg_TLR	AGB93809.1
<i>Caenorhabditis elegans</i>	TOL-1	Ce_TOL-1	CCD63554.1
<i>Capitella teleta</i>	hypothetical protein CAPTEDRAFT_189993	Ct_189993	ELU15098.1
<i>Capitella teleta</i>	hypothetical protein CAPTEDRAFT_209087	Ct_209087	ELU06716.1
<i>Crassostrea gigas</i>	Toll-like receptor 1	Cg_TLR1	AGY49095.1
<i>Crassostrea gigas</i>	Toll-like receptor 2	Cg_TLR2	AGY49096.1
<i>Crassostrea gigas</i>	Toll-like receptor 3	Cg_TLR3	AGY49097.1
<i>Crassostrea gigas</i>	Toll-like receptor 4	Cg_TLR4	AGY49098.1
<i>Drosophila melanogaster</i>	toll, isoform C	Dm_Toll	NP_001262995.1
<i>Drosophila melanogaster</i>	18 wheeler	Dm_Toll2	NP_476814.1
<i>Drosophila melanogaster</i>	MstProx	Dm_Toll3	NP_649719.2
<i>Drosophila melanogaster</i>	Toll-4	Dm_Toll4	NP_523519.2
<i>Drosophila melanogaster</i>	Toll-5	Dm_Toll5	AAF86227.1
<i>Drosophila melanogaster</i>	Toll-6, isoform A	Dm_Toll6	NP_524081.1
<i>Drosophila melanogaster</i>	Toll-7	Dm_Toll7	NP_523797.1
<i>Drosophila melanogaster</i>	tollo	Dm_Toll8	NP_524757.1
<i>Drosophila melanogaster</i>	Toll-9, isoform A	Dm_Toll9	NP_649214.1
<i>Eisenia andrei</i>	membrane pattern recognition receptor TLR	Ea_MPRR_TLR	AGS14315.2
<i>Euprymna scolopes</i>	Toll-like receptor [1]	Es_TLR1	AAV27971.1
<i>Euprymna scolopes</i>	[Toll-like receptor 2]	Es_TLR2	This study
<i>Euprymna scolopes</i>	[Toll-like receptor 3]	Es_TLR3	This study
<i>Euprymna scolopes</i>	[Toll-like receptor 4]	Es_TLR4	This study
<i>Euprymna scolopes</i>	[Toll-like receptor 5]	Es_TLR5	This study
<i>Euprymna scolopes</i>	[Toll-like receptor 6]	Es_TLR6	This study
<i>Haliotis discus discus</i>	pattern recognition receptor	Hd_PRR	AGJ03555.1
<i>Helobdella robusta</i>	hypothetical protein HELRODRAFT_163616	Hr_163616	XP_009025691.1
<i>Hirudo medicinalis</i>	Toll-like receptor 1	Hm_TLR1	ADK94453.1
<i>Homo sapiens</i>	toll-like receptor 4 isoform A	Hs_TLR4	NP_612564.1
<i>Ixodes scapularis</i>	toll, putative	Is_Toll	EEC14202.1
<i>Litopenaeus vannamei</i>	Toll protein	Lv_Toll1	ABK58729.1
<i>Litopenaeus vannamei</i>	toll2	Lv_Toll2	AEK86516.1
<i>Litopenaeus vannamei</i>	toll3	Lv_Toll3	AEK86517.1
<i>Lottia gigantea</i>	hypothetical protein LOTGIDRAFT_159050	Lg_159050	XP_009050960.1
<i>Lottia gigantea</i>	hypothetical protein LOTGIDRAFT_163508	Lg_167776	ESO90996.1
<i>Mytilus galloprovincialis</i>	Toll-like receptor a precursor	Mg_TLRA	AGG10802.1
<i>Mytilus galloprovincialis</i>	toll-like receptor b	Mg_TLRB	AFU48614.1
<i>Mytilus galloprovincialis</i>	Toll-like receptor l	Mg_TLRL	AGG10809.1
<i>Mytilus galloprovincialis</i>	Toll-like receptor r precursor	Mg_TLRR	AGI05194.1
<i>Pinctada martensii</i>	Toll-like protein	Pm_Tolllike	AHA85008.1
<i>Pinctada martensii</i>	Toll-like receptor4	Pm_TLR4	AHA85007.1
<i>Toxocara canis</i>	Protein toll	Tc_Toll	KHN70653.1

¹ Brackets indicate names proposed in this study.

performed in MEGA6 (21). Sequences were aligned with MUSCLE (22) on default settings and positions with less than 50% site coverage were eliminated prior to further analysis. Phylogeny was inferred with the Maximum Likelihood method, with the initial tree obtained by the Neighbor-Joining method and processed with a JTT matrix-based model (23). Bootstrapping was performed with 100 resamplings.

Quantitative reverse-transcription PCR. Gene-specific primers for EsTLR1,2,3,4,5, and 6 are shown, along with those for control genes, in Table 4-3. For EsTLR5, primers are upstream of the deletion which caused an unexpected, early stop codon. Light organs and gills were dissected from three and eyes from two 28- to 31-day-old animals per replicate, with four biological replicates used per tissue for aposymbiotic and symbiotic *E. scolopes*. RNA and cDNA were prepared as indicated above. Real-time PCR reactions were performed in duplicate with 4 μ L cDNA in a 20 μ L total reaction volume, using forward and reverse primers at 0.25 μ M and using Sso advanced SYBR green mix (Bio-Rad, Hercules, CA). No-template controls (NTC) were included, as well as no-reverse-transcriptase (NRT) controls to indicate any residual presence of chromosomal DNA. A CFX Connect qRT-PCR machine (Bio-Rad) was used with the protocol: 3 min at 94°C, 40 \times [15 s at 94°C, 20s at 59°C, 20 s at 68°C]. We used the comparative $\Delta\Delta C_q$ method to determine expression levels (24). *estlr* levels were normalized to the mean levels of control transcripts for the 40S ribosomal subunit and peptidyl-prolyl cis-trans isomerase A (PPIA). Transcript levels were log-transformed to provide for normality prior to statistical analysis. Results for each gene were analyzed by ANOVA followed by *post hoc* t test with the α value adjusted by a Bonferroni correction based on the total number of comparisons (apo- vs. symbiotic, tissue vs. tissue) across the experiment. Comparisons were made within a given *tlr* gene, between treatments and tissues, rather than between *tlr* genes.

Table 4-3: Primers for qRT-PCR of *estlr* transcripts and control genes

Gene	F primer name	F primer sequence	R primer name	R primer sequence
40S	40sF2	AATCTCGGCGTCCTTGAGAA	40SR2	GCATCAATTGCACGACGAGT
PPIA	PPIAqF1	TGTTGCTGAGGATCTGGATTT	PPIAqR1	GGTTCAATGCGACCAGTTTG
TLR1	TLR1qF1	GTTAACCCGAGAGACATTCCAG	TLR1qR1	AAAGGTTGTAATCGCCAAGAAA
TLR2	TLR2qF1	CCCGAACTCCAAACACTCTATC	TLR2qR1	CTTCCAATCGTATCAGGTGTGA
TLR3	TLR3qF1	TGACAGAGACTTTATCGCAGGA	TLR3qR1	GAGGACTTTACGGCTTTTCTCA
TLR4	TLR4qF1	CGCGGTTATCATGTTTGTCTT	TLR4qR1	TTGATGGCGTTTAGGATGTTC
TLR5 (set 1)	TLR5qF1	ATACCGTTGGTACATTCGGTTC	TLRqR1	TAATGTCCCCAAATTCTCCATC
TLR5 (set 2)	TLR5qF2	ATGCCACTGCAAGTCAACAAC	TLR5qR2	TCGGTGAGGTTTTCAAAGGATT
TLR6	TLR6qF1	ATCTGATGCTGCTGGAAAGAA	TLR6qR1	ACTGCGGCCCAAATACTTTAT

RESULTS

Four-week-old *E. scolopes* animals express six TLR-family genes. Analysis of an RNA-seq library of four-week-old *E. scolopes* yielded five promising candidate TLR sequences, which we propose to name EsTLR2-6, in addition to the previously published EsTLR (5), which we propose renaming EsTLR1. RACE-PCR was carried out for three of these sequences (EsTLR2, 5, and 6) to obtain the 5' end of the ORF; ultimately, signal peptide sequences were identified for each amino acid sequence, strongly suggesting the start codon was obtained for each; stop codons were also detected in each case. The entire ORFs for EsTLR2 (Fig. 4-1), EsTLR3 (Fig. 4-2) and EsTLR6 (Fig. 4-5) were successfully amplified by PCR and sequenced. Though attempts were made to amplify the full transcript for EsTLR4, they were not successful. This failure may be due to some combination of transcript length (it is the longest of the newly amplified EsTLRs) and rarity. However, the full-length ORF sequence is apparently present in an RNA-Seq database (S. Moriano-Gutierrez *et al.*, unpublished data) and is given in Fig. 4-3.

Full-length sequence was successfully amplified for EsTLR5, but with unexpected results that disrupt the ORF. 14 bp of sequence expected in EsTLR5 based on RNA-seq results (N. Kremer *et al.*, unpublished data) were not present in sequence amplified for this study, resulting in a frameshift and stop codon disrupting the amino acid sequence upstream of the transmembrane region (Fig. 4-4A). This apparent deletion was observed in PCR products based on cDNA from two entirely distinct RNA preparations, as well as an RNA-seq study (S. Moriano-Gutierrez *et al.*, unpublished data), and is thus unlikely to be a PCR, cloning or sequencing artifact. Resultant amino acid sequences for the PCR-amplified sequence (Fig. 4-4B) and a hypothetical transcript with the “missing” 14 bp inserted (Fig. 4-4C) are given. Fig. 4-4C

- A** ATCGCAGTGGATTTTTCTGGGTAAAGAAGAACTACATTAGTAGGCTTTTTTTTTTAATCAACAGTTACGTTATGCGAAAAAGGATACAAAG
 AAAAAAGGTGTC**ATGGAGCGATT**CAGCTTCAAACTTCGGTTATTTTTCTCC**TACTGTTATGCCTACAAATAACACCGTCACCTTCTTCAAAG**
TTGACCACCAACCGTGCATCAATCGCAACGGAGTAATGGTTTGTGCAAACCTCAATTACATCCCGATTTACCACGAAACACAACCAAGTTTGT
CCTTAATAAACTTTCGTGTAATCTCCTGAGCCGAAACACTCTCTCAAATCTGACACAACCTGCCATTACAATCATTACAACATAAAGATTCCA
TTTGGAAAGACGTTACAACAGACGTATTCACAGACATGGTACACTTGGGTACGTTGAGCATAGTFCGATCAAATTTGTTGCGAAATGCCCA
AAATGTTTCTTAGCCTAAACTTCACCAACATAACCACCCTAAGTGTACACCAGACCAAAATCAGCATAGTGGGCCTGATCTTTTGAAGGGT
TGCGTAACACCAGTCTACGTCGCTCTATTTGAGGAGATGTCGTTTACTAACTTCGATGGTGCACACCTCGTGGATCTAACAGCACTACAT
TGCTGGATTTGCTGTGAACGAAAACCTTGACGTCGAATCTTGGGGATATCTACCCTCGCTAAAGATTTTGAATCTCACAGGACCTCAATGG
ACCATCTGCTCTCGTTCAAATGAATCCCTTCGCACAGGTAATGCTATCCCGAACTCCAACACTCTATCTATCCTTATCTGATGGCTATC
TTATAAACGATGTTACATTCRAGGGGCTTCCGAATTTGACTCGCTTAATGATCCAATCCATTGACAATCTTCAGATCACACCTGATACGATTTG
GAAGCCAACTGCCAAAGTTAAAGAAGTTATACTTTATTTGGAGTGCAGGACGTTTTACTAAGCGTGGACAACAAGCGATAGTTAGTCTTCCGT
TAGAAGCCTCAAACATAAAATGTAACCTCAAGTCTCCAGTGAGTCTTCTATTTGCCCAATCTCAACACTCTTACATTGCTCGCTTCA
AAGTTGAGAAATTAATAAGTTTATTCGTATCCTATCAACATTACAACACCTGAAGCGTTAGAAATCTCAAACCTCCAAATCACTACTGTAC
CAGATCTATTTGTAACATAACCACACTAACCCAGCTGGACTTGCATGGCAATTTCATATATTCCTGGAAGACTAACGACTGTCAAGTCATTA
AACATCTAGAAGAAATTCGTTGGCCAGGAACGATATTAAGAAGATAAAACAAATCTTCCTTTTCAGAAACGATGGGAAAAATCCGAAATTA
AATGGGACCTTTTATAATAGTTACCGTTGCAACTGTGATTTGTTATGGTTCCGTGATTGGCTTCTGAGGACCAAAATTAATGCTCAACT
ACCCTAAGGGCTACCGTTGCAAGTTCTCTCTGGTTTATACATAAAATTCCTAAATTCGATCTGACATCAGAAATCTGTCTGACAGATTCC
CGCCGTCACTCTAATATCTTGCATCTGCTTAGCCGTACTGGCTCTCACCTTTCATCATTTGCAGCGTCTGCTAGCTAACAAACGTTGGT
CGGTCCTTTCTGGTCTTTCTATTCTGCAGCCGTCGGCGTAAGTATACACCGTTACCTGATAGCTGATAACGTCGACTCATCTACGATCGCTTCT
GCTACAGCAACCCGGATTGGAAATGGGTGATGAACAACTCCTGCCGATGATCGAGGAGACAGAGGGTTTTCAACTTTGTCTACATGACCCGG
ATTTCAAGGCTGGCTACACCATAGTAGATAACATCTTGAGAGTGTGCAAACCGCCGAAAGTCCGTGTTGATCTTGTCCGAGATTTTGCAA
GTAGCAGCTGGTGAAGTACGAGGCTAGCCTCGCCGAACAACGCTCTCTGGAAGACAAAAGAGATATGCTAGTCCGATCTTTTGAAGAGA
TTCCCTTCGAAGTCCAGTCGAGAAGGCTCGCCACTCTCATGAAAAGGATCACCTGCTTGGAAATGGAGTGACGATGAGCGCCGCCAGAACATAT
TCTGGAGAAGACTCACGGAGATTTTGAGAACCACCAATGCACACGGATATCTGATTATTGCGACATCATAACATTACTGAAACCAACTTTT
 TTTCTTTTGTGTTTCTTTTACAAAATGACAAATTATAATAGAAGTTAATAGAACGGGTTTTGCATAGCTCACTCCACATATAATGTTAT
 TCTAAGGTTGCCATCTCTGGAACAAAATATCATGGAAGTCTTCTCTACAACATGGTTCGGTTTCAAGTCA
- B** MERFSFKPSFIFLLLLLCLQITPPSPSSKVDHQPCINRNGVMVCANLNYIPILPRNNTLSLSLINFVNLLSRNTLSNLTQLPLQSLQLKSSIWKD
 VTTDVFTDMVHLRSLISRSKLFKFLPKMFLSLNFTNITLSVHQTKISIVGPDLLLEGLRNTSLRRLYLRRCLLNFDGAHLVDLTAQLLDL
 SVNENLHVESWGYLPSLKIINLNTGYMDHLLSFQNESLRTGNVYPELQTLYLSLSDGYLINDVTFKGLPNLRLMIQSIDNLQITPDTIGSQ
 PKLKKLYFIGVDVLLSVDNKAIVSSLESCLKKNVTLKSSSEFYCPNLNLTLSRFKVEKFNKFIIRLSTLQHLKRLIISNSKITTVPDSI
 CNITTLTQLDLHGNFIYSWKTNDQVMKHLEEFSLARNDIKKINKSSSFETMWNPKLKWDLNLSYRNCNCDLLWFRDWRLLRTRKIKMLNYPKG
 YRCKLSSGLYIKIPKFDLTSEFCRDRFRPSSLIILAICLAVLALTFIICSVVAYVKRWSVRFWCFLFCRRRKYTPADNVYIYDAFICYSN
 RDLEWVMNKLPMIEETEGFQLCLHDFKAGYITVDNILESQVTSRKKVLLSPDFASSWCKYEASLAEQRLLLEDKRDMLVPILLQEIPE
 VQSRRLATLMKRITCLEWSDDERGQNI FWRRLTEILRTPMHMTDI

FIG 4-1 Nucleotide sequence of *estlr2* transcript and derived amino acid sequence. (A)

PCR-amplified sequence of *estlr2* transcript following RACE-PCR. Bold and underlined: ORF, including stop codon. (B) Derived amino acid sequence of EsTLR2.

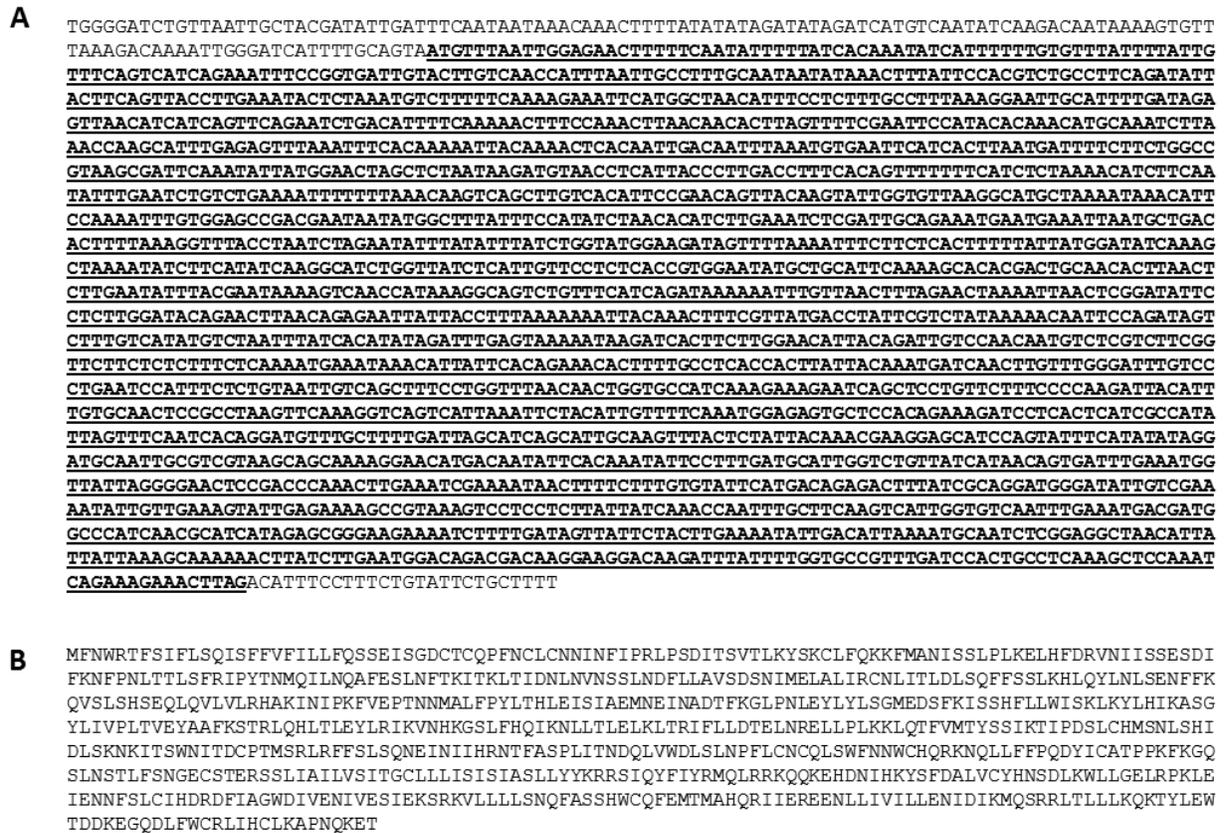


FIG 4-2. Nucleotide sequence of *estlr3* transcript and derived amino acid sequence. (A)

PCR-amplified sequence of *estlr3* transcript. Bold and underlined: ORF, including stop codon.

(B) Derived amino acid sequence of EsTLR3.

A

GATTTGACGACATAACAACAACAAAAAATTTGAAAAGTTTCAAAAAGATTTACGATACGAGTTCGGATTTCGCTTCTTATCATCA
GACGATAAATTTTATTCTTTAATATATCATTTATTTATTTATTTTGTGGAGTTTACGAGAGAACCAAGGGTTTATATTTTC
AAACATATATTACTCGGATATAATACATTTCTTTTGTCTGGCTAATATTACTGGAGAAAAGCAAAAAGCAAAAAGTAAAACTTGGTA
TTTTTCAAAGCACCGTAAACCTTAATACAGATTTCAATGGTGGTTGAAAAACAACATGACCCCTCGCCACCAGTTCGTTATTTCCATCT
TCTTTCATCACTGTTGTAAGCGGGCTTCTTCTCAGGACAAGAAGCAGTCTGTACCCGAATGCTCATTCCAACCGGACCAATGAATAATAAAA
ACCTCTTACAGATGACAGTGTGCTGCTGGCAACCCAAAGACATCGCAGTCATCATCAAACCATACACAGAGACCGGTTGCATTCTCCGGATGA
GCCAATTCGTTTTGAGTCCGCCGAAAGACAATGCTCATTACTTTTGTGGCTTGACAGTGGCCACCCGATCATCTTAGACACGGAGAAATG
TCACAACCATACCAGATAATATGCTTATTTACAAGTGCAGAAATGCTAATTTGATGGGGCTAACATGTTGAGTTTAAAAAGCTTACAATC
TGAAGTGTGACTTACTTCTTTACAGTCCAAACGCAACAAATATGGCTTGGAAATAAATCAACGTCGACTCCAGAAAGCAAGATTAGTGATG
TTGTCAGTTATACGGTACAGAGCAACATGTAATCAAACAATCGACCGTGGATACAGAAACGAACGAGACTTTCCCGCATATGGCCGAGATGA
GTATGACAGGGATGAACCTCACTGAACCTCCGCCAGATATGAACGTCGGTTTCCCTAACCTGCAATCGATCGAGCTGGATTGCAATGGTTTTTA
CACAGGTACCGTTTTTCCATACACAGAGAGCGTGTGCTCCTCCCGATGAACCTCAGCCGGACGCTAACATGCAAGATCATTATGCCAAAT
CGATCGGCATTGATGTCAGGCTAATATGTTCCGAGAATCCTTATGCTGAATAATAACTTCATCAAAGCCTAGACGATTTATTTGCCAAGC
GCACATTCAGTGTGATTTCCGTTGGGAACAACTACTGGATAAAATCTCCGATCGCGCTTCGATCGGGTCAATGGCTTACAAGTTCTGATCC
TGAGAAATAACAGGCTGACCCGAATCAGCGATAAGGTGTTCAATAAATGTCACAGAACTATTGCAGCTTGACCTAGAGCTGACACGATAACT
GCATACACCCCGGACGATTTAGAAAACCTCTCTAACTAAAGCAATGACCTTGTGAGTAAATAACTTGACCTGATAAACGACAGACGTTTT
ATGGGTTATCTTCTGTTGAAACCCCTAAAATAGACAAGAACAATCTAAGTAAATTTGATCGAATATCTTACCAATTCATTTACCAGCTCA
GAAATATCAATCTAAGTCAAGCAAAATTTAAAACCTTTGCTGTGTTATTTTCTGTTTACGCGGTTTACGAGCATCAACTGAGCCAAATGTTG
AGATGCTATCAACAAGATAAGCGATCTTTGACTTGTATCTCAGACAGTGAAGTAAATCGATAGCATATCAGATCCGGAATGCGCGAGAG
CTGATATATTTGAGAAACGAGAGACAAAAAGATAGTAAATCTTCAAGGTAACCTCATTAGCCGGCTAGACTTCAACAAATTTATCGTGCATCT
ACAAAAGATGGTTATCTGATCTCAATATTTTGAAGTGAATCTTCCGACAAACCCGATCAATGTTGACTGTCACTCCGCTTTTGG
ACTTCTTTGCCGAGCCCGTTAAATAAAGCGTTTACGGAACAGAACTTTTACAAGAACTGGTTGTCAATCGCCCGGAAATTTCCAG
ATCGCTCTTACTCTCCCTGCATAAAAAAGATCTGACTGCTCAGCAAAATGCTGGCTGCCCGTCAGACTGTGCTGCTGATAGCGGATGAG
TCTCAGGAATCGTCAATGTAAGTCTCTCGCAACAGCCGTTCTCATTCCAGACGAAATGCCGACGGTAGACTGGAACTTTGGCTGCAGT
ATAACAACATTAAGCGAAATGCGCCGAGGAGTATTTGCCAAAGATCGAGATATTGAACTGACGAAGAACTACTTACCAGCGTAAGCGTTT
CCATTTTGAAGCAATGAAGAAATCAAAATCTCTATCTGGATTCAAACCTCTTGACAACATTTGCCACAAGAGTTGAAGAGGTCACACTG
ATCTATTTAGTATCATCAAGATGAACCATTTACTTTGCAACTGTCAATACGATGGATGAAAGATGGCTGTTGAGAGTTCAAAGCAGAGCTCAG
GCTGGCATAAAATCGCTTGCAGCACCAGTAAAGATAAAGTGTTCAGTATTTGTAAGTGGAGGATGATGACTTTATCTGCATGGAAGAAGTTA
ACTTTAAGAAGTCTTCCCTTGGCTCTCCCTGGGCTTCGCTATCCCTTGGTATTAATCTGTCTTATCTTCTATTTACTACAGACTTG
AAGTCAAGTTTTGATATACATCTATTTAACCTACACCCGTTTGATTTGATCCGAAAAGATAGACGACGAGAGAAACCGATTTGATACGC
TGGTCACTGTTTCAAAATCAATGTTGGATGGGTTGCCGAGAAATGAAACCCGACCCCTGCAATCCCTGCGATTTAACGTGATCGAAATTAATC
GGGATTTCTTATTTGGCATGAGTATCCGCGAAAACATGAGAATGGCCGCTGTGTCGAGCAACCGCAGTCTAGTTATCTGTCCATGGATAGTT
TCAGAGACCCGATCATCAACTAGCACTGACGTTTACCCACGAGAAGGTTCTGCGGGAGCGCCGACCTACATGGTGTCTTCTTGCATCAAG
TCAAAAAGAGTATCATGATAACAGGACCTGAAGAAATACCTGACGAGCGGCTCGCTACATCCGTACCCGGGACAGCATGCTTAGCAAAAAGA
TGCTCTACTGCTGACCGGTTGCAACGACGCCAACAAACAGGCCCTGAAAATGAGGCGGAAAAAATCATACAAAATATCTTCTCCGAAATCT
TGAACACATCTTGTCCCAACAAGCTGCGTGGTTTACGACATTTTATCAGTTATCCAGATCAAGACTATCAGTTTGCAGACGAGTTCTCT
TTCGCGAAGTGCATTCTCGCGGTTATCATGTTGTCTTCTGACCCGGATTTTGGGTTAGGTGTCGCAAGGAGGAGAACATCTTAAACGCCA
TCAACCAATTCGGGACGAAACGCTAGTTATTTATCAGGAGAAATCACGTTGAGGACGAGTGGCAGTGTTCACGTTACGGGACCGGATCAGACCT
CCTTGAAGGAGACGTCGAACCTACCTGCTGTCATGCTCGACCCGAAAGGTTGACTTCGACAGCCTAGACCCCGGACGAGGCGCTTTCTTTCCA
CCCATGTTATATCTTTCAAGACGATCCGCTTTTTTGGCAGAAGTTTTATCGCTGATCCACCCGCGACCTTTAAGAAAACCGACTCTGACA
ATAACGCATACACGAACGAAACGATCAATTAATACGATCCCAACAGAGTCAATATTTACGTCAGCTGACGATCAAAATGAGGAAAACAAAAA
AACATTAAGAGAGTAAATAAAGCATAAACTAATATTTTAAATCTCATAAATATGATGAGTTTATTTGAGGAGCGCAGCAGAGCAAGA
ATTCTTAGCAGCGGGTTTCGAAAAAAGAGTCCGCTTGGCTTTATTTTCGCAATTTAAATAACAGCAAAATATTTGCAGATTTATGTCGATT
AATAACGTTTAGAGAACCGTATCGAAATATGAGAAGAAATAAATAATTTACTTTTATGTCAGTTTCGCTATTTACACTTCAATGTAAGGCCAA
ACGATTCCTTACTGTTATATAGCAGCGTGTGACCATGACCAATCATCTGGTATGCTTACTGCTGGCTGATCCTAACCATCAATGTAAGCAAA
TCAAATTAATTAATAAATAATATCTGATGTTTCCCTTGA

B

MTLATSSFLVAFITVIVSVGVPSSQDKNEVPECSFQTRPMNNKLNFTMNSVLLATQDIAVIKPYTETGCIILRMSQFVLKSAERQLLIYFVAC
RVAHPILDTEVNTIPDNIAYLQVQKCLIDGANMLKFEKLYNLKLVLYFFVTPTQIWLGNKSTSTPRSKI SDVVS YTVQSNMYSNNRPWIL
ETNEFPHMAEMSMTGMLNTELPDMNVRFNLSQIELDNSNGFTQVFPVTEKRVVLPMLNLSRTPNMQDHYAKSIGIDVKANMFRRLIMLNN
NFIKSLDDYFANGTLQLISVGNLIDKISDRAFDVIGLQVILRNRLTAISDKVFNKTELLQLDLELNSINWIHPDAFRKLSKQLTSL
ANNLTLINPETYGLSSLETLKLKNNLSKFDRTILPIHFTKLRNINLSQNKFKLPLVLFLLRGLSINLSQCEIAINKISDLDDLISDSEL
IDSIIIRSGRAEADIFEKRETKRIVNLQGNLISRLDFTNLSIYKMMVILINNFELILDNPINCDCHLVPLDDFAEARLNKRFDGTEYFY
KNWLQSPPEFQDRLLSLHKDLYCLEQMPGCPSDVCVYRYSVGIIVNCSRNRSVSPDEMPDGRLELWLQYNNIIEIAPRSYLPRIEIL
KLTKNLLTSVSVSILKQLKRIKILYLDNLLTLPQVEEVNLDLVIDMNHLLCNHNEWMKRLLRVQSRVRGWHKIACTSKDRVFSIVE
VEDDDFCIMEEVNFKVLALGLSLGFAILLVLIICYLLFYRLEVKLIYIYFNLHPFDCDPQKIDDEKEPIDTLVICSMSVGVVAENVGT
LESRLFNVEINRDFLIGMSIRENMRMAVCRSRSLVILMSDSFRDPIQLALFTTHEKVLRRPPTVMVFLHQVKSIIIDNQLKLYLSSGR
YIRTDGSMKQKMLYLLTGSNAANKQGLEMRKSYKIFPEILNTSCPTSCVVDIFISYPDQYQFATGVLPFLHSRHYVCLPDRDFR
VGVAKENILNAINQSRRLVIITENHVEDEWQLFTRTAIERSLKETSNYLLCMLDPKVDPSLDPETRAFSLTHVII FQDDPLFWQKFRS
IPPTFKKPSDNNAYTNETIINTIPTESVYSVT

FIG 4-3. Nucleotide sequence of *estlr4* transcript and derived amino acid sequence. (A)

Sequence of *estlr4* transcript obtained from RNASeq database (S. Moriano-Gutierrez *et al.*, unpublished data). Bold and underlined: ORF, including stop codon. (B) Derived amino acid sequence of EsTLR4.

A

TTTTACACGGAAACACCTCCTTAATAAGTTATTCGTTATGGATTTTCGGTTTAAAGCGTTGTCGAAATGGTTCGACTGGGATTTATCTATG
 GTTATGCATGTTAACAGCTGCAGCTGACAACCTGTTCTGTCCGAAAAATGCCACTGCAAGTCAACAACCTGTTCAATGCTTGCACCAAAAACT
 TACTGAGATCCCAAGGTATCCCAAGATGACCGAAAAATTTATCCTCGGTTTCAACAATTTGCCGTCAATCACCTCAAATCCTTTGAAAA
 CCTCACCGAGCTGAAATTTAGATCTGAGATCGAACAAGATCGCAAGTATCGGCAAAAATTTCTTTCCGCCACAGAGTAAGGTGAAATCTTT
 GACTCTGATCAATAATGATCTGATAGAAATCGATAAAGACGCATTTTCTGGCGCTATCTCTTTGGAAAACTTTACTTAAGTTCAAAATCGGCT
 TCATGCAGTTCACCTTTGGCGGGTTGACTTCCTTGCAAAAGCTCATAGTTGATACGAACCAGATCCTTAATCTCACCTTCCCGCCAGAGTT
 CACTGCTTTGAAGGGCTTGAGTTACATTAATGTAGCCAACAACAGAATCTCCTCCTTAAGCGACACGTCTTTCGTAATCTGCAAAAGTGCCAA
 TCTGCGTAGATTGCTACTGTCTCGAACTTCATCCAGTTTATTCATCCTGAAGCATTGCGCCTTTAAGGTCTTAGAATCCTTGAAGCTTTC
 CTTCAACCCACTGCAGGGCAACATCTTCGGATTGCGTTATCAAGCCTCAGTGGCGGTAATCTTCTTTCACTAGACATCAGCAACCTTTCTCT
 CTCCAACGTGTTACCGGCGAGTACATTTAGTGTGTAATAAATCAAACTTACGGACTCTCATATTTGGATCACAACCACTTCAACAACATCAA
 CAACAATGCTTTTCAGACCTTACCACCTTACTGACGCTGAACTGACGCTTTCGAGAATCCAGACGATTTCCAAAAATGCATTTTCATGGCCT
 TACACATTTGAACTACTTATTTTATCACAACGAAGTACTGACCTCACCTGGGATTTCCCTGAGAACTACACAACTCTACTTGGATCA
 CAACAAGATCCAGAAAATCCAAACGGTGTCTTCGGAAAGCTCATCAATCTAAAAGAGCTTACCTTGACTATAACAAGGTGCATGAATFCCA
 AAAAAATGCTTTTCGTTGGCTTGAAAACTCCAGAACTTCGACTTCCTCACAATCTATTAATGCAATCCCAAGTCTCTGTTCTCCAGTTT
 GGTGGACTTGAGTCTTTGGAGCTGAACAGTAATCCGAATACTACATTTGTGTCATGACACAACTTCTATACTTAAACATGGCCGACAATG
 ACATAACTTCTCTCTGCTGTAATAAGTGGCCTCCTTCAGCTTCACTATCTGTATCTGTCAAGCAACAACTCGGCAGCACAAATGACCC
 AAAACGATTTCAATCGCCTGTTCCGACCCACAACAACACTACAACACTCGACCTTCCGATAGCCGAATTCAGGGCACCTGTGCAAGACA
 TATTCAAGAACCTGGATGAACTCAAATCTTCAGCTCACAAAAACAGATCTCTTCATGGGATGGTGAACGTTCAAACCGATGAGCGATA
 AAATTGAGATCCTTGATATTCGAGAGAACCACATCACGACAATCAGCAAGTCAAGTGAATTTAGAATATCTGAACAGCTTGAAGTTTTAAAAATTA
 TGAAGAATCCTTTCCGCTGTAACGTGTCAGCTTCGGTGGTTCGGTACTGGCTTAAAAACAGACTGTCTTGATCCCAACTTGGCGGAGATGA
 AATGCAACAGTCTTAAAGCATGGCAGGAAAGCATCTTGTGCACTTTGACCGTTCAAAAAATGATTTGACTAATAACCCCTACTATATCT
 ATGGTGGTGTGCCATTGGCCTTGATAGCCTTAGTCACTGACTTCCGTTATGTATTCATACCGTTGGTACATTCGGTTCATGCTTACAGAC
 TAGGTAAGCGAATCAGATATGGATTTGCTGCGCGTGAAGGTTACGAAGAAATACCTGGGAATGACTTGGTATTTGATCTGAATATTAACATA
 CTGATCTGATCTTGAATGGGTTGATAGATAATGTATGAAACCTTTTGTATGATGGAGAATTTGGGGACATTAATTTGAGGGTGAAGTTAAAC
 TGTGCATTAAGGATCGAAATTTGAGGGCCGTCGGAAATGTCTGTATATTTGGAGAATAGAGAAAAGTCGTTTTTCGTTAATTTGTTCTTT
 CTCAGGATTTTAAAGGGTTACAGGTGCGAGTTTGTAGCTGATTACTGCTGTAATCATTTCAAATACAACCGTCAAAGATAATCACTGTTG
 CTGTTGGCGACCTTAGAGCGCAAGATATTTCCAACTCCTGAAACCCCTTTCGAAAAAGACACTTACCTTAAATGGGAAGACAAGACATAG
 CCAAAGAACACTTCAAGCAAAGTCTGTATGATATATAAAGAAGAAGAA TAATTGTAATACATCTAATAGTACTTGGACGATAAATCCGAAT
 CG

▶ AATTCATCACAGC ◀

B

MDPFFKALSCKGRGFLWLCLMTAAADNLFCPKKCHKSTTVQCLHQKLEI PQGIPKMTEKFLGFFNNLPSITSKSFENLTELKLLDLRSN
 KIASIGKNSFVHQSKVKSLLINNDLIEIDKDAFSGAISLEKLYLSSNRLHAVPPLAGLTSQKLIVDNQLNLNLTFFPEFTALKGLSYINVA
 NNRISSLSDTSFVNLQSANLRRLSLSRNFIQFIHPEAFAPLRSLESLSLNFNPLTGEHLRIALSSLSGGNLLSLDISNLSFSNVLPASTFQLL
 INTNLRLLILDHNFNNINNAFQTLPTLLTLKLTSCRIQTI SKNAFHGLTHLNYLFLSQNELVDLTWDFPEKLHKLYLDHNKIQKIPNGVFG
 KLINLKLHLDYKVKVHEFQKNAFVGLEKQLKLRPHNSINAI PSSLFSSLVGLSELELNSNSITAI PNTTFVSMTQLLYLNMADNDITSLPAV
 LMSGLLQLHYLYLSSNKLSTMTQDNFNLFAHKNLQTLDSLDSRIQGLSKDI FKNLDELKILQLTKNEISSWDGETFKPMSDKIEILDITR
 ENHITTTISKSNLEYLNSLVKIMKNPFACNCQLRWFDRDLKNTTVLIPNLAEKMCNSPKAWQGHVDFDRSKIDCTNYTPYYIYGGVAIGL
 VIALVMTSVMYSYRWYIRFYAYRLGKRIYRGAAREGYEIEIPGNDLVFDLNIYNTDSDLEWVVDNVMNPFDDGEFGDIKFEFEGFKLICKDRNL
 RAGPEMSVILENIEKSRFSLIVLSQDFKKGYRCEFELDYCCNHFKYNTSKIITVAVGDRAQDIPKLLKPLEKDTYKWKEDDKIAKEHFQKQ
 LYDILKKE

C

MDPFFKALSCKGRGFLWLCLMTAAADNLFCPKKCHKSTTVQCLHQKLEI PQGIPKMTEKFLGFFNNLPSITSKSFENLTELKLLDLRSN
 KIASIGKNSFVHQSKVKSLLINNDLIEIDKDAFSGAISLEKLYLSSNRLHAVPPLAGLTSQKLIVDNQLNLNLTFFPEFTALKGLSYINVA
 NNRISSLSDTSFVNLQSANLRRLSLSRNFIQFIHPEAFAPLRSLESLSLNFNPLTGEHLRIALSSLSGGNLLSLDISNLSFSNVLPASTFQLL
 INTNLRLLILDHNFNNINNAFQTLPTLLTLKLTSCRIQTI SKNAFHGLTHLNYLFLSQNELVDLTWDFPEKLHKLYLDHNKIQKIPNGVFG
 KLINLKLHLDYKVKVHEFQKNAFVGLEKQLKLRPHNSINAI PSSLFSSLVGLSELELNSNSITAI PNTTFVSMTQLLYLNMADNDITSLPAV
 LMSGLLQLHYLYLSSNKLSTMTQDNFNLFAHKNLQTLDSLDSRIQGLSKDI FKNLDELKILQLTKNEISSWDGETFKPMSDKIEILDITR
 ENHITTTISKSNLEYLNSLVKIMKNPFACNCQLRWFDRDLKNTTVLIPNLAEKMCNSPKAWQGHVDFDRSKIDCTNYTPYYIYGGVAIGL
 VIALVMTSVMYSYRWYIRFYAYRLGKRIYRGAAREGYEIEIPGNDLVFDLNIYNTDSDLEWVVDNVMNPFDDGEFGDIKFEFEGFKLICKDRNL
 RAGPEMSVILENIEKSRFSLIVLSQDFKKGYRCEFELDYCCNHFKYNTSKIITVAVGDRAQDIPKLLKPLEKDTYKWKEDDKIAKEHFQKQ
 LYDILKKE

FIG 4-4. Nucleotide sequence of *estlr5* transcript and derived amino acid sequence. (A)

PCR-amplified sequence of *estlr5* transcript following RACE-PCR. Bold and underlined: ORF, including stop codon. Arrowheads: site of presumed 14 bp deletion (inset). Italicized and underlined: alternate stop codon of hypothetical transcript, *estlr5**, without 14 bp deletion. (B) Derived amino acid sequence of EsTLR5. (C) Derived amino acid sequence of EsTLR5*, based on hypothetical insertion of 14 bp at indicated arrowheads in (A).

depicts an apparently complete TLR, EsTLR5* but it is important to note that the nucleotide sequence for EsTLR5* has not been detected in its entirety by either PCR or RNA-seq.

The derived amino acid sequences were analyzed with LRRFinder, SMART and PFAM for the domain features of TLR proteins (Fig. 4-6). For EsTLR5, the products of both the PCR-amplified transcripts and the predicted, but unobserved intact proteins (EsTLR5*) are represented.

EsTLR proteins form clades with different lophotrochozoan TLR proteins. Amino acid sequences for EsTLR1-6 (using the hypothesized intact EsTLR5*), Toll and TLR genes for diverse protostomes, and *Homo sapiens* TLR4 were aligned with Muscle and grouped into a maximum likelihood phylogenetic tree (Fig. 4-7). With the sole exception of *D. melanogaster* Toll 9, arthropod sequences grouped into two well-supported clades with no lophotrochozoan protein members, and the two nematode sequences formed a sister clade to the *Drosophila* Toll-2 group, though the node for this early ecdysozoan split was less well-supported.

The lophotrochozoan clade containing EsTLR1 is close to one of the main ecdysozoan clades, indicating, for example, that EsTLR1 and *D. melanogaster* Toll 2 share a more recent common ancestor than *D. melanogaster* Toll and Toll 2. The EsTLR1 clade includes proteins from all three molluscan groupings tested, but none of the annelid proteins included in this analysis. The other lophotrochozoan clades were entirely external to the ecdysozoan portion of the tree (except the poorly supported *D. melanogaster* Toll-9), and except for the most evolutionarily recent nodes, were rarely well supported. EsTLR2 and EsTLR3 are each other's closest relatives, branching off from a group including bivalve and annelid TLRs. EsTLR4 and EsTLR5 are each sister to a bivalve TLR, whereas EsTLR4's sibling is a gastropod protein. An alignment of TIR regions alone (not shown) had generally poor bootstrap values, but did notably

Protein	AAs	LRRs (cysteine cluster)	Diagram
EsTLR1	1191	25 (2)	
EsTLR2	696	6 (1)	
EsTLR3	698	6 (1)	
EsTLR4	1150	8 (2)	
EsTLR5	452	12 (0)	
EsTLR5 *	846	21 (1)	
EsTLR6	829	13 (1)	

FIG 4-6. Domain organization of EsTLR1-6. Length of EsTLR proteins, number of leucine rich repeats (LRR) and diagram of domains detected by LRRfinder/SMART/PFAM. “TLR5” is the derived amino acid sequence from the full-length sequence amplified for this study, whereas “TLR5*” is derived from a hypothetical transcript replacing 14 nucleotides observed in an RNA-Seq database (N. Kremer *et al.*, unpublished data) but apparently lost due to deletion in the animals from which RNA in this study’s PCR was derived. SP, signal peptide; circle, LRR; circle with ‘C’, C-terminal (cysteine cluster) LRR; TM, transmembrane helix; TIR, Toll-Interleukin 1 Receptor (TIR) domain.

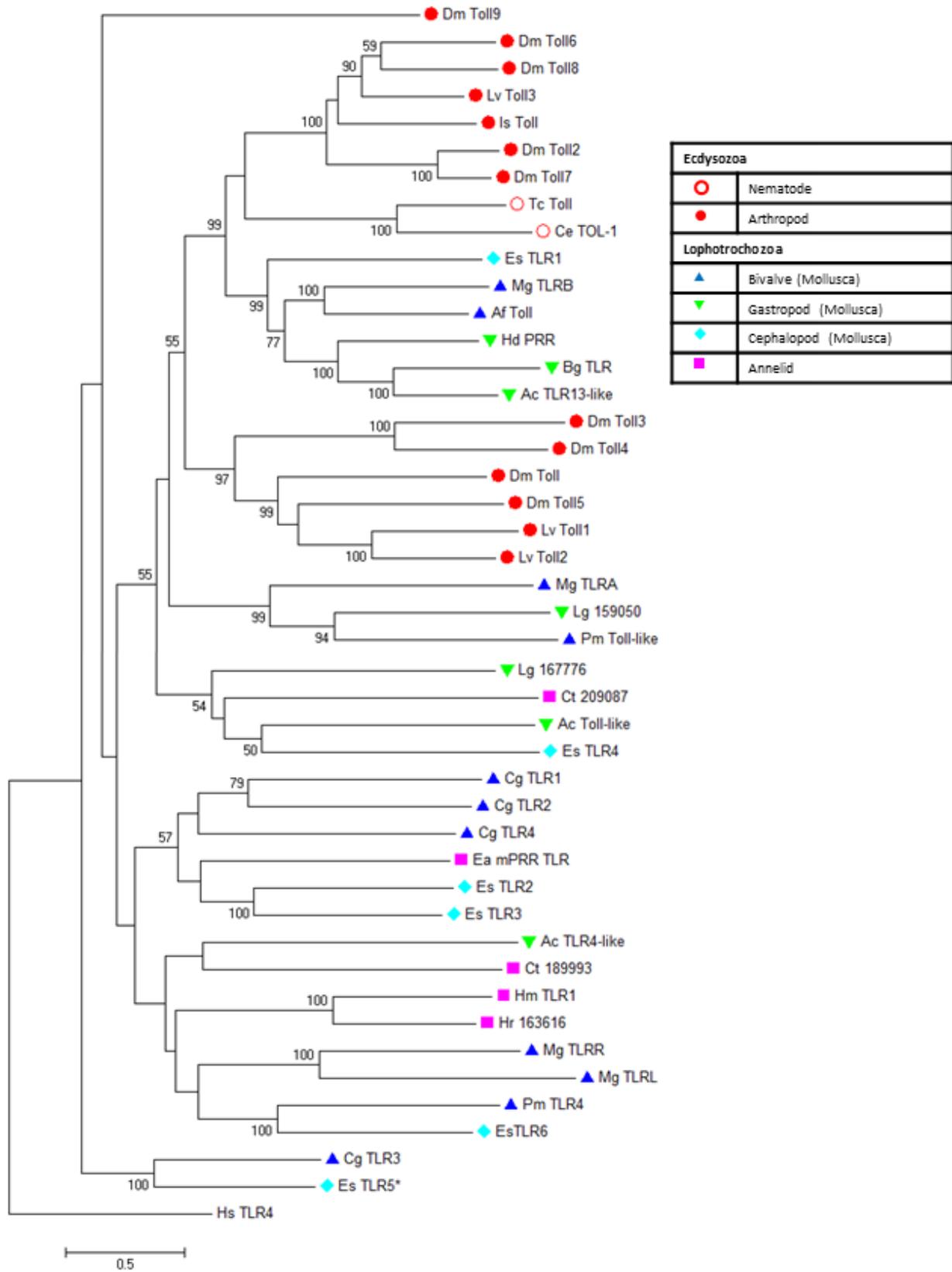


FIG 4-7. Phylogeny of Protostome Toll and Toll-like Receptor proteins. Maximum-likelihood tree for amino acid sequences of diverse protostome Toll and TLR proteins (see Methods and Table 2-2), including EsTLR1-4, EsTLR5*, and EsTLR6, rooted to *Homo sapiens* TLR4. Key indicates phyla of species from which protein sequences were obtained. Bootstrap values for 100 replicates are shown only if >49.

include all the same lophotrochozoan sequences of the EsTLR1 clade positioned among the ecdysozoan Tolls, similarly to the arrangement discussed above.

***estlr* transcripts are expressed differentially across *E. scolopes* tissues.** *estlr1-6* transcript levels were evaluated by qRT-PCR, normalized by comparison to 40S ribosomal subunit and PPIA transcript levels in light organs, gills and eyes taken from ~four-week-old aposymbiotic and symbiotic *E. scolopes* animals. Differential expression between tissues was detected in most cases, but only in one instance, an approximate doubling of *estlr2* transcript in symbiotic light organs compared to aposymbiotic, was a difference in apo- and symbiotic transcript levels of a given gene in a given tissue observed. Transcript levels were highest in the gills for *estlr2* and, apparently, *estlr6*, as *estlr6* levels were not measurable in other tissues. *estlr4* levels were highest in the light organ, and *estlr3* levels were indistinguishable between gills and light organs but lowest in eyes. *estlr1* levels were indistinguishable between all tissues, and *estlr5* was unmeasurable in all tissues. Transcript levels were termed unmeasurable for one or more of the following reasons: total failure to amplify, Ct values greater than the assay's linear dynamic range limit, or Ct values similar to those of no-reverse-transcriptase controls.

DISCUSSION

E. scolopes possesses genes for at least six distinct Toll-like receptors. Five of these (EsTLR1-4, EsTLR6) are predicted to possess unambiguous structural features of Toll-like receptors. The remaining protein, EsTLR5, warrants particular explanation. Its transcript detected in the mRNA (and therefore cDNA) used for our sequencing studies contains what

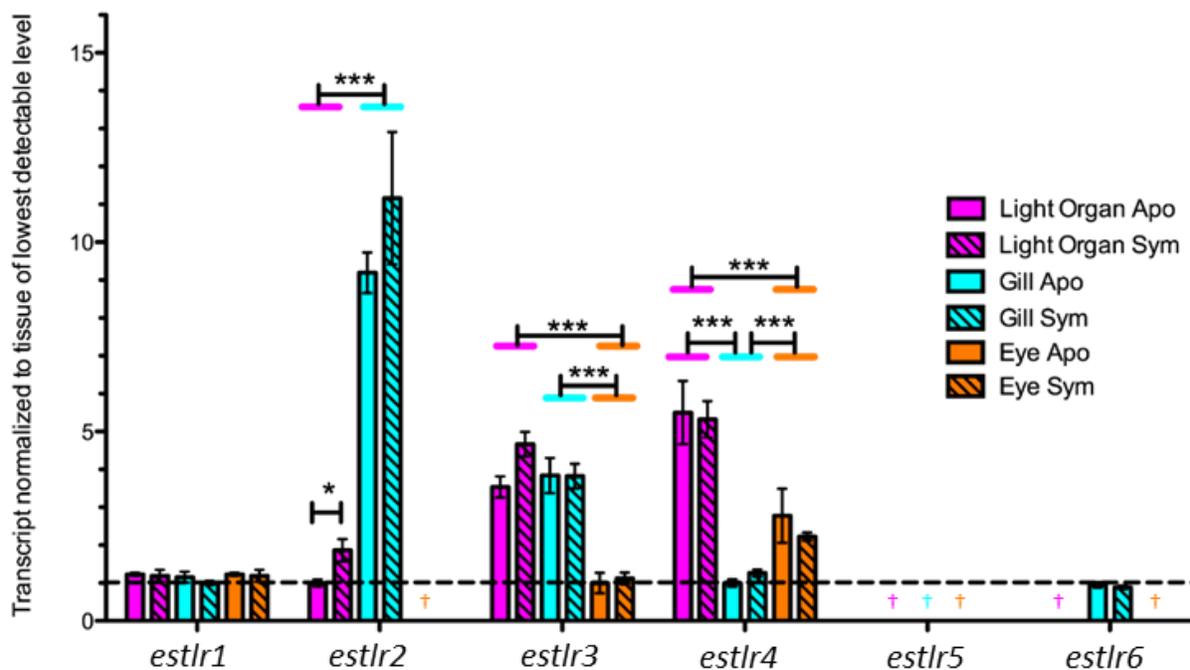


FIG 4-8. Expression of *estlr* genes in different four-week-old *E. scolopes* tissues. Transcript levels in light organs (purple), gills (cyan) and eye (orange) in apo- (solid) and symbiotic (hatched) animals indicated by qRT-PCR for each *estlr* transcript were normalized to 40S and *ppia* (peptidyl-prolyl cis-trans isomerase A) transcripts; each *estlr* is standardized separately, with the level of the tissue with lowest measurable transcript defined as 1. Significant comparisons for each gene are shown, either within a tissue between apo- and symbiotic animals or between different tissues (bars with color-coded labels). P-values: * $<.05$, *** $<.001$; †, transcript unmeasurable (see Results).

appears to be a 14 bp deletion in the LRR region (Fig. 4-4, Fig. 4-6). These 14 bp are present in one (N. Kremer *et al.*, unpublished data), but not both (S. Moriano-Gutierrez *et al.*, unpublished data) of two RNA-Seq databases from which candidate sequences for RACE-PCR and cloning were obtained. Restoration of these 14 bp yields an ORF coding for all of the important elements of a TLR (Fig. 4-4C, “EsTLR5*”), whereas the shorter transcript amplified simply yields a string of LRRs (Fig. 4-4B, EsTLR5) and a large, presumably untranslated 3' region. (We are not aware of examples of polycistronic messages for nuclear-encoded genes in mollusks.) We thus strongly suspect that wild populations of *E. scolopes* include animals with the truncated, presumably inactivated EsTLR5 sequenced in this study, as well as animals with an intact protein similar to the hypothesized EsTLR5* sequence. It is, however, important to consider that EsTLR5* has not been observed in its entirety, and a possible role for RNA editing (extensive in some cephalopod transcripts (25)) has not been investigated. The phylogenetic portion of this study included the restored EsTLR5*, as the truncated sequence would be lacking too many TLR elements to be meaningfully compared to intact TLRs from *E. scolopes* and other species.

The *E. scolopes* TLR proteins are, in some ways, markedly dissimilar, and the common ancestor of all six of these proteins is apparently quite distant. EsTLR1 and EsTLR4 have multiple cysteine clusters associated with their LRRs, whereas EsTLR2-3 and 5*-6 have a single, C-terminal (just prior to the transmembrane region) cysteine cluster, a pattern seen in vertebrate TLRs but much less common in invertebrates (26).

The grouping of lophotrochozoan TLRs has proven to be a challenging (27), more so than most of those examined in other metazoans. Indeed, the best-supported portion of our phylogenetic analysis [Fig. 4-7] consisted mainly of ecdysozoan sequences. The phylogenetic tree is perhaps most interesting for its placement of EsTLR1, as well as five other

lophotrochozoan TLR sequences, close to two clades of ecdysozoan Toll proteins. The similarity of EsTLR1 and MgTLR-b has been noted previously (28). Circumstantial evidence, including the transcription of *Haliotis discus discus* TLR in response to LPS and bacterial challenge (29) and *Azumpecten farreri* TLR in response to LPS (30), as well as the expression of MgTLR-b in a digestive gland and especially in hemocytes (28), implicates some of the other members of this lophotrochozoan clade in immune activity. Because of the apparent evolutionary proximity of this clade of lophotrochozoan TLRs to ecdysozoan Tolls, it may be instructive to compare their ligand-binding capabilities to *Drosophila* Tolls. Spätzle is related, and thought to bind its receptor similarly, to neurotrophins (31), which are found, for example, in *Aplysia*, wherein is a neurotrophin involved in development (32) that may be worth investigating in the context of immunity and TLR activation. We are not aware of transcriptional evidence of a similar gene in *E. scolopes*.

The other *E. scolopes* TLR genes cannot be grouped with great confidence into large families, though two of them have well-supported siblings in other molluscan TLRs. EsTLR5* is close to *Crassostrea gigas* TLR3, which is strongly implicated in immune function in hemocytes, but whose specific microbial ligand, if any, is unclear (13). EsTLR6 is sibling to a mostly uncharacterized *Pinctada martensii* TLR, and EsTLR2 and EsTLR3 are each other's own closest relations and thus may represent a duplication of a gene within *Euprymna* (or within the cephalopods). As there is no published *E. scolopes* genome (or any cephalopod genome), it is not possible to know what portion of the total EsTLR repertoire is represented in this study.

We also investigated the tissue tropism of *estlr* genes in four-week-old *E. scolopes* animals. Tissues selected included the light organ, site of a beneficial symbiosis with *V. fischeri*; the gills, a highly-perfused tissue with constant exposure to free-living bacteria that may also, in

older animals, be the site of bacterial cysts surrounded by hemocytes (33); and the eye, at which Toll-like receptors are involved in the defense against invading pathogens in humans (34). Regional specialization of immune function is known to occur within and between tissues in mammals (reviewed in (35)), likely reflecting differences in both beneficial and pathogenic microbial populations. We saw variation in expression level between tissues for most of the *estlr* transcripts (Fig. 4-8). There is no evidence of symbiosis in the light organ altering the level of any *estlr* in a distant tissue, and only one case, *estlr2*, was expression was significantly altered by symbiosis in the light organ itself. This expression pattern is especially interesting given that *estlr2* was much more highly expressed in the gills than light organs. It is possible that *estlr2* is highly expressed at any site with plentiful microbial exposure, irrespective of the beneficial or pathogenic nature of those microbes. It is also notable that *estlr5* levels were below the threshold of detection in all tissues examined, as this protein does not appear to be fully functional in all animals (Fig. 4-4, Fig. 4-6), though conditions such as pathogen exposure were not tested.

The widespread expression of these Toll-like receptors across tissues in older (four week) *E. scolopes* likely indicates a role in immunity in adult animals, distinct from a purely developmental role. Further characterization of the expression of these genes may be based on MAMP exposure studies. It also remains to be seen if these proteins have a developmental function, which may be investigated by examining timing and location of expression over the trajectory of development. In future functional studies, comparison of EsTLR1's ligand binding capability to that of the other EsTLRs, those more distantly related to the ecdysozoan proteins, may provide insight into the function of ancestral Toll/TLR proteins.

REFERENCES

1. **McFall-Ngai M, Heath-Heckman EA, Gillette AA, Peyer SM, Harvie EA.** 2012. The secret languages of coevolved symbioses: insights from the *Euprymna scolopes-Vibrio fischeri* symbiosis. *Semin Immunol* **24**:3-8.

2. **McFall-Ngai M, Nyholm SV, Castillo MG.** 2010. The role of the immune system in the initiation and persistence of the *Euprymna scolopes*-*Vibrio fischeri* symbiosis. *Semin Immunol* **22**:48-53.
3. **Foster JS, Apicella MA, McFall-Ngai MJ.** 2000. *Vibrio fischeri* lipopolysaccharide induces developmental apoptosis, but not complete morphogenesis, of the *Euprymna scolopes* symbiotic light organ. *Dev Biol* **226**:242-254.
4. **Koropatnick TA, Engle JT, Apicella MA, Stabb EV, Goldman WE, McFall-Ngai MJ.** 2004. Microbial factor-mediated development in a host-bacterial mutualism. *Science* **306**:1186-1188.
5. **Goodson MS, Kojadinovic M, Troll JV, Scheetz TE, Casavant TL, Soares MB, McFall-Ngai MJ.** 2005. Identifying components of the NF-kappaB pathway in the beneficial *Euprymna scolopes*-*Vibrio fischeri* light organ symbiosis. *Appl Environ Microbiol* **71**:6934-6946.
6. **Medzhitov R.** 2001. Toll-like receptors and innate immunity. *Nat Rev Immunol* **1**:135-145.
7. **Manfrulli P, Reichhart JM, Steward R, Hoffmann JA, Lemaitre B.** 1999. A mosaic analysis in *Drosophila* fat body cells of the control of antimicrobial peptide genes by the Rel proteins Dorsal and DIF. *EMBO J* **18**:3380-3391.
8. **Hailman E, Vasselon T, Kelley M, Busse LA, Hu MC, Lichenstein HS, Detmers PA, Wright SD.** 1996. Stimulation of macrophages and neutrophils by complexes of lipopolysaccharide and soluble CD14. *J Immunol* **156**:4384-4390.
9. **Park BS, Lee JO.** 2013. Recognition of lipopolysaccharide pattern by TLR4 complexes. *Exp Mol Med* **45**:e66.
10. **Fitzgerald KA, Rowe DC, Golenbock DT.** 2004. Endotoxin recognition and signal transduction by the TLR4/MD2-complex. *Microbes Infect* **6**:1361-1367.
11. **Michel T, Reichhart JM, Hoffmann JA, Royet J.** 2001. *Drosophila* Toll is activated by Gram-positive bacteria through a circulating peptidoglycan recognition protein. *Nature* **414**:756-759.
12. **Coscia M, Giacomelli S, Oreste U.** 2011. Toll-like receptors: an overview from invertebrates to vertebrates. *Invertebrate Survival Journal* **8**:210-226.
13. **Zhang Y, He X, Yu F, Xiang Z, Li J, Thorpe KL, Yu Z.** 2013. Characteristic and functional analysis of toll-like receptors (TLRs) in the lophotrocozoan, *Crassostrea gigas*, reveals ancient origin of TLR-mediated innate immunity. *PLoS One* **8**:e76464.
14. **McFall-Ngai M, Montgomery MK.** 1990. The Anatomy and Morphology of the Adult Bacterial Light Organ of *Euprymna scolopes* Berry (Cephalopoda:Sepiolidae). *Biological Bulletin* **179**:332-339.
15. **Koch EJ, Miyashiro T, McFall-Ngai MJ, Ruby EG.** 2014. Features governing symbiont persistence in the squid-vibrio association. *Mol Ecol* **23**:1624-1634.

16. **Letunic I, Doerks T, Bork P.** 2012. SMART 7: recent updates to the protein domain annotation resource. *Nucleic Acids Res* **40**:D302-305.
17. **Kremer N, Philipp EE, Carpentier MC, Brennan CA, Kraemer L, Altura MA, Augustin R, Häsler R, Heath-Heckman EA, Peyer SM, Schwartzman J, Rader BA, Ruby EG, Rosenstiel P, McFall-Ngai MJ.** 2013. Initial symbiont contact orchestrates host-organ-wide transcriptional changes that prime tissue colonization. *Cell Host Microbe* **14**:183-194.
18. **Offord V, Coffey TJ, Werling D.** 2010. LRRfinder: a web application for the identification of leucine-rich repeats and an integrative Toll-like receptor database. *Dev Comp Immunol* **34**:1035-1041.
19. **Finn RD, Bateman A, Clements J, Coggill P, Eberhardt RY, Eddy SR, Heger A, Hetherington K, Holm L, Mistry J, Sonnhammer EL, Tate J, Punta M.** 2014. Pfam: the protein families database. *Nucleic Acids Res* **42**:D222-230.
20. **Petersen TN, Brunak S, von Heijne G, Nielsen H.** 2011. SignalP 4.0: discriminating signal peptides from transmembrane regions. *Nat Methods* **8**:785-786.
21. **Tamura K, Stecher G, Peterson D, Filipski A, Kumar S.** 2013. MEGA6: Molecular Evolutionary Genetics Analysis version 6.0. *Mol Biol Evol* **30**:2725-2729.
22. **Edgar RC.** 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res* **32**:1792-1797.
23. **Jones DT, Taylor WR, Thornton JM.** 1992. The rapid generation of mutation data matrices from protein sequences. *Comput Appl Biosci* **8**:275-282.
24. **Pfaffl MW.** 2001. A new mathematical model for relative quantification in real-time RT-PCR. *Nucleic Acids Res* **29**:e45.
25. **Rosenthal JJ, Bezanilla F.** 2002. Extensive editing of mRNAs for the squid delayed rectifier K⁺ channel regulates subunit tetramerization. *Neuron* **34**:743-757.
26. **Leulier F, Lemaitre B.** 2008. Toll-like receptors--taking an evolutionary approach. *Nat Rev Genet* **9**:165-178.
27. **Halanych KM, Kocot KM.** 2014. Repurposed transcriptomic data facilitate discovery of innate immunity toll-like receptor (TLR) Genes across Lophotrochozoa. *Biol Bull* **227**:201-209.
28. **Toubiana M, Gerdol M, Rosani U, Pallavicini A, Venier P, Roch P.** 2013. Toll-like receptors and MyD88 adaptors in *Mytilus*: complete cds and gene expression levels. *Dev Comp Immunol* **40**:158-166.
29. **Elvitigala DA, Premachandra HK, Whang I, Nam BH, Lee J.** 2013. Molecular insights of the first gastropod TLR counterpart from disk abalone (*Haliotis discus discus*), revealing its transcriptional modulation under pathogenic stress. *Fish Shellfish Immunol* **35**:334-342.

30. **Wang M, Yang J, Zhou Z, Qiu L, Wang L, Zhang H, Gao Y, Wang X, Zhang L, Zhao J, Song L.** 2011. A primitive Toll-like receptor signaling pathway in mollusk Zhikong scallop *Chlamys farreri*. *Dev Comp Immunol* **35**:511-520.
31. **Lewis M, Arnot CJ, Beeston H, McCoy A, Ashcroft AE, Gay NJ, Gangloff M.** 2013. Cytokine Spatzle binds to the *Drosophila* immunoreceptor Toll with a neurotrophin-like specificity and couples receptor activation. *Proc Natl Acad Sci U S A* **110**:20461-20466.
32. **Kassabov SR, Choi YB, Karl KA, Vishwasrao HD, Bailey CH, Kandel ER.** 2013. A single *Aplysia* neurotrophin mediates synaptic facilitation via differentially processed isoforms. *Cell Rep* **3**:1213-1227.
33. **Small AL, McFall-Ngai MJ.** 1999. Halide peroxidase in tissues that interact with bacteria in the host squid *Euprymna scolopes*. *J Cell Biochem* **72**:445-457.
34. **Yu FS, Hazlett LD.** 2006. Toll-like receptors and the eye. *Invest Ophthalmol Vis Sci* **47**:1255-1263.
35. **Mowat AM, Agace WW.** 2014. Regional specialization within the intestinal immune system. *Nat Rev Immunol* **14**:667-685.

CHAPTER 5

Conclusions and Future Directions

PREFACE

The writing of this chapter is entirely the work of Benjamin Krasity. It references unpublished work carried out by Kirsten Guckes and Theresa Gioannini (*eslbp4* cloning and expression *in vitro*), Fangmin Chen (tracking expression of EsLBP2 and -4) and Bethany Rader (EsTLR1 characterization) and is inspired in part from conversations with Jerrold Weiss and Margaret McFall-Ngai.

In this study of microbe-associated molecular patterns (MAMPs) and pattern-recognition receptors (PRRs), Chapter 2 demonstrates the expression of a lipopolysaccharide-binding protein - bactericidal/permeability-increasing protein (LBP/BPI) family protein in *E. scolopes*. This protein was expressed at numerous epithelial surfaces in the squid, strongly suggesting a general role in host defense of various tissues, but was also strongly regulated by symbiosis in the light organ. Its isoelectric point and lack of bactericidal activity are strongly suggestive of LBP-like activity, rather than BPI-like, and we suspect it is responsible for carrying the LPS signal to host cells. To more thoroughly characterize this protein as an LBP, I propose:

1. Evaluating the ability of EsLBP1 to render LPS sensitive to deacylation by acyloxyacyl hydrolase (AOAH). Mammalian AOAH removes secondary acyl chains of LPS (1), reducing its immunoreactivity. Mammalian LBP promotes the activity of AOAH (2), raising an obvious question of whether recombinant EsLBP1 might have such an effect in synergy with either mammalian or *E. scolopes* (Appendix C) AOAH.

The possibility of an LBP/BPI dichotomy, outside of the well-characterized split in the vertebrates, is one of the most intriguing possibilities related to EsLBP1. If EsLBP1 is LBP-like, perhaps the more basic EsLBP2, -3, and -4 are BPI-like. The most obvious sign of BPI-like character would be antibacterial activity like that not seen in EsLBP1 (Fig. 2-3). *eslbp4* was cloned into the expression vector pBac-11 (Novagen/EMD Millipore, Billerica, MA); however, it appears the recombinant protein was unstable, as anti-His Western blots yielded little signal. It may be necessary to pay special attention to freeze-thaw issues, or find an alternate way to store the protein. Using the vector pBac-3 would allow more direct comparison to the experimental

results in Chapter 2, without an unnecessary cellulose-binding domain. To investigate a potential LBP/BPI dichotomy in *E. scolopes*, I propose:

2. Cloning one or more of *eslbp2*, *eslbp3*, and *eslbp4* into pBac-3, expressing the protein and comparing its biochemical properties, especially antibacterial activity, to those of EsLBP1. Synergy in bacterial killing with actinomycin D (3) may be more easily assayed than direct bactericidal activity. (Preliminary studies with EsLBP1 did not find an actinomycin D-sensitizing effect on PL2 *E. coli*.)
3. Comparing the expression patterns of different EsLBP proteins in different tissues, including the light organ. This work is already being undertaken by colleagues.

In Chapter 3 of this dissertation, the O-antigen of *V. fischeri* was characterized, along with its importance in the symbiosis. *V. fischeri* lacking the O-antigen exhibited a defect in colonization, though it could not be determined whether this phenotype was due to a factor other than its motility defect. One obvious experiment to bridge Chapters 2 and 3 of this thesis, and provide a possible additional mechanism for the relative inability of *waaL* mutant to colonize, is:

4. Testing the ability of EsLBP1 to bind *waaL* mutant *V. fischeri* LPS.

Previous studies have implicated the lipid A/core region of LPS in directly binding LBP, rather than the O-antigen (4); however, O-antigen presence can impact the antibacterial action of BPI (5). The O-antigen can affect bacterial susceptibility to antibodies (6) and complement (7).

Promising studies of the O-antigen, then, may relate to the relationship between the *V. fischeri* O-antigen and the *E. scolopes* complement system (8).

Chapter 4 of this dissertation reports five previously unknown Toll-like receptors in *E. scolopes*. In four week-old animals, these are expressed differentially across tissues. The morphogenetic program of the light organ is largely complete in these animals, suggesting that those EsTLR proteins expressed in the light organ have roles in the maintenance, rather than (or in addition to) the establishment of the symbiosis. EsTLR2, in particular, is implicated in the symbiosis, inasmuch as it is the only of these genes found to be regulated by symbiosis in the light organ. It is also likely, given their expression at diverse tissues, that EsTLRs mediate interactions with diverse bacteria, not all of them beneficial. To determine if these genes are involved in light organ development, and what bacterial or host factors they may respond to, I suggest:

5. Evaluating the expression of *estlr* genes over the trajectory of development and in response to MAMPs.
6. Evaluating the MAMP-binding properties of EsTLRs, either through a co-capture assay or by transfecting *estlr* genes into an NF- κ B reporter system (9).
7. Searching *E. scolopes* transcriptomic databases for possible molluscan equivalents of Spätzle or other endogenous TLR ligands. Such a protein may, if recombinantly expressed, cause transcriptional activity in insect larvae or cell lines (10).

Finally, there are still other “missing” components of the putative *E. scolopes* NF- κ B LPS-response pathway, between EsLBP and the EsTLRs. Mammalian LBP is dependent on

CD14 and MD-2 to deliver its signal to Toll-like receptor 4. Neither of these molecules has been reported in *E. scolopes*. The proteins predicted by *E. scolopes* transcriptomic libraries or the forthcoming *E. scolopes* genome, probed with more detailed searches such as for all MD2-related lipid recognition domains, may allow recognition of candidates, after which these proteins could be analyzed more thoroughly *in vivo* or *in vitro*. Alternatively, it may be possible to co-precipitate these factors with EsLBP1 and LPS, detecting and evaluating bound partners through means such as mass spectrometry or 2D gels. Such an experiment would likely require a more pure preparation of EsLBP1 than has been used thus far.

This work adds to our understanding of MAMP synergy in the *E. scolopes*/*V. fischeri* symbiosis. Upregulation of the *eslbp1* gene was not detected by lipid A, but rather by TCT, even though EsLBP1 does not itself bind TCT. These findings suggest that the host's response to LPS, as exerted through an LPS-binding protein, depends on the TCT which *V. fischeri* exports. The lipid A of *V. fischeri* is not an especially potent molecule (11), and the host's alkaline phosphatase activity is likely to reduce its activity further (12). Both host and microbe, it seems, conspire to make the effect of LPS subtle. The MAMPs of *V. fischeri* and the PRRs of *E. scolopes* have likely co-evolved to promote a detection system that is sensitive to bacterial presence without allowing the host to be harmed by endotoxic effects. The full extent of the involvement of host MAMP-binding molecules, including LBPs/BPIs, PGRPs, and TLRs, in this mutualistic relationship remains to be explored.

REFERENCES

1. **Munford RS, Hall CL.** 1986. Detoxification of bacterial lipopolysaccharides (endotoxins) by a human neutrophil enzyme. *Science* **234**:203-205.

2. **Gioannini TL, Teghanemt A, Zhang D, Prohinar P, Levis EN, Munford RS, Weiss JP.** 2007. Endotoxin-binding proteins modulate the susceptibility of bacterial endotoxin to deacylation by acyloxyacyl hydrolase. *J Biol Chem* **282**:7877-7884.
3. **Horwitz AH, Williams RE, Nowakowski G.** 1995. Human lipopolysaccharide-binding protein potentiates bactericidal activity of human bactericidal/permeability-increasing protein. *Infect Immun* **63**:522-527.
4. **Tobias PS, Soldau K, Ulevitch RJ.** 1989. Identification of a lipid A binding site in the acute phase reactant lipopolysaccharide binding protein. *J Biol Chem* **264**:10867-10871.
5. **Capodici C, Chen S, Sidorczyk Z, Elsbach P, Weiss J.** 1994. Effect of lipopolysaccharide (LPS) chain length on interactions of bactericidal/permeability-increasing protein and its bioactive 23-kilodalton NH₂-terminal fragment with isolated LPS and intact *Proteus mirabilis* and *Escherichia coli*. *Infect Immun* **62**:259-265.
6. **Hester SE, Park J, Goodfield LL, Feaga HA, Preston A, Harvill ET.** 2013. Horizontally acquired divergent O-antigen contributes to escape from cross-immunity in the classical bordetellae. *BMC Evol Biol* **13**:209.
7. **Caboni M, Pédrón T, Rossi O, Goulding D, Pickard D, Citiulo F, MacLennan CA, Dougan G, Thomson NR, Saul A, Sansonetti PJ, Gerke C.** 2015. An O Antigen Capsule Modulates Bacterial Pathogenesis in *Shigella sonnei*. *PLoS Pathog* **11**:e1004749.
8. **Castillo MG, Goodson MS, McFall-Ngai M.** 2009. Identification and molecular characterization of a complement C3 molecule in a lophotrochozoan, the Hawaiian bobtail squid *Euprymna scolopes*. *Dev Comp Immunol* **33**:69-76.
9. **Zhang Y, He X, Yu F, Xiang Z, Li J, Thorpe KL, Yu Z.** 2013. Characteristic and functional analysis of toll-like receptors (TLRs) in the lophotrochozoan, *Crassostrea gigas*, reveals ancient origin of TLR-mediated innate immunity. *PLoS One* **8**:e76464.
10. **An C, Jiang H, Kanost MR.** 2010. Proteolytic activation and function of the cytokine Spätzle in the innate immune response of a lepidopteran insect, *Manduca sexta*. *FEBS J* **277**:148-162.
11. **Foster JS, Apicella MA, McFall-Ngai MJ.** 2000. *Vibrio fischeri* lipopolysaccharide induces developmental apoptosis, but not complete morphogenesis, of the *Euprymna scolopes* symbiotic light organ. *Dev Biol* **226**:242-254.
12. **Rader BA, Kremer N, Apicella MA, Goldman WE, McFall-Ngai MJ.** 2012. Modulation of symbiont lipid A signaling by host alkaline phosphatases in the squid-vibrio symbiosis. *MBio* **3**:e00093-12.

Appendix A

NF- κ B Pathway Activation in *E. scolopes*

PREFACE

qRT-PCR for *esikB* in apo- and symbiotic animals was by Benjamin Krasity. Experiments with NF- κ B inhibitors were planned and carried out by BK and Julian Cagnazzo, with input from Margaret McFall-Ngai, and analyzed by BK. Cloning of *esikB* was by JC and BK. Cloning of *esrel* was by Eva Ziegelhoffer. Proteins were expressed by EZ. Western blotting and immunocytochemistry was by BK. This appendix was written by BK, with assistance in the Materials and Methods (pertaining to cloning and expression) from EZ.

INTRODUCTION

There are significant gaps in our understanding of how *E. scolopes* detects and responds to MAMPs. Earlier chapters of this thesis have focused on EsLBP1, a soluble protein capable of binding LPS (Chapter 2), and on Toll-like receptors, suspected membrane-bound proteins that may have a role in binding MAMPs or conveying endogenously produced developmental signaling molecules (Chapter 4). In this appendix, work is presented concerning the NF- κ B system of *E. scolopes*. NF- κ B proteins are transcription factors which, among other functions, are ultimate effectors at the mRNA transcription level for LPS molecules detected by vertebrate hosts (1).

Vertebrate LBP contributes to inflammatory signaling by interacting with CD14 and conveying a signal through MD-2 and membrane-bound TLR4. TLR4 signals through the Myd88 adapter protein, which interacts with IRAK proteins through its death domain, setting off a phosphorylation cascade through intermediates such as TRAF6 (2). “Classical” NF- κ B signaling ultimately proceeds through activation of the I κ B kinase (IKK) complex by means that are not yet fully elucidated. The IKK complex phosphorylates the NF- κ B inhibitor I κ B α , leading to its ubiquitination and proteasomal degradation. Until activated, NF- κ B is in a complex with I κ B and thus inactive. The degradation of I κ B α allows NF- κ B to translocate to the nucleus. NF- κ B dimers accumulate and lead to the expression of pro-inflammatory cytokines (reviewed in (1, 3)). There are alternative cell-surface receptors and signal transduction intermediates that may be involved in NF- κ B activation; the “alternative” pathway involves IKK α , but instead of involving the phosphorylation of I κ B, proceeds through the phosphorylation and processing of p100, which has both a DNA-binding Rel homology domain and the ankyrin repeats associated with the inhibitory function of I κ B (reviewed in (4)).

Many of these proteins have equivalents in *Drosophila* that have also been studied. The *Drosophila* equivalent of I κ B is Cactus, for example (5); the NF- κ B family transcription factor most implicated in dorsoventral patterning is termed Dorsal (6) and that most associated with the immune response to Gram-positive bacteria and fungi is Dif (7). In both cases, these *Drosophila* proteins function downstream of Toll receptors which respond to the endogenous ligand Spätzle, rather than directly binding a bacterial substrate (8).

NF- κ B activation may be associated with resistance to cell death in cancer (9), and mice deficient in IKK β have been shown to have increased intestinal apoptosis (10, 11). Despite these facts, in newborn mice, TLR4 activation has been shown to be involved in an increase in apoptosis in the small intestine (12), which likely indicates that NF- κ B can be pro- or anti-apoptotic depending on as yet unclear factors. The symbiotic bacterium *Bacteroides thetaiotamicron* promotes the export of p65 (an NF- κ B transcription factor) from the nucleus, reducing inflammation (13).

RACE-PCR has identified multiple probable components of the NF- κ B pathway in *E. scolopes*. The sequence for one TLR has been published (14), and transcripts for five more have since been identified (Chapter 4). Evidence of IRAK4, TRAF6, IKK γ , I κ B, Myd88 and two NF- κ B family transcription factors has also been published (14, 15). Full sequence is available for a Class I NF- κ B protein, EsRel (without ankyrin repeats), and a Class II NF- κ B protein (with ankyrin repeats, like p100 or p105) is suspected, but not yet described in full. Conspicuously absent thus far is an apparent homologue of CD14 or MD-2, though the MD-2-related lipid-recognition domain is present (NCBI GenBank AAN35173.1). Any number of processes relating to host response to *V. fischeri* colonization may depend on this NF- κ B machinery, including the LPS-driven apoptosis and regression mentioned above, as well as other MAMP-driven processes

such as mucus secretion (16), hemocyte trafficking (17), and nitric oxide synthesis attenuation (18).

Unpublished experiments¹ using a polyclonal antibody to human NF- κ B p65 suggested that *E. scolopes* EsRel localizes to the nucleus in response to *V. fischeri* colonization. This finding would likely indicate EsRel activation over the course of colonization, and a pro-apoptotic role in the light organ appendages (which undergo apoptosis and regression in response to *V. fischeri* colonization). This finding subsequently proved difficult to replicate, and may have depended on the lot of commercial antibody used. We sought to use an antibody produced in response to EsRel, specifically, to verify translocation to the nucleus. Additionally, if EsRel is activated by the NF- κ B classical pathway or something similar to it, the translocation of EsRel to the nucleus is likely preceded by degradation of the protein EsI κ B and followed by upregulation of the transcript *esikB*, as is seen in a mammalian NF- κ B autoregulatory feedback loop (19). An antibody to EsI κ B was also ordered, along with qRT-PCR probes to *esikB*, to monitor the hypothesized degradation and subsequent rebound of protein and transcript. Commercially available inhibitors of the NF- κ B pathway were also ordered, in order to determine if transcription driven by NF- κ B was essential for developmental phenotypes in *E. scolopes*. The inhibitors used were MG132, a proteasome inhibitor that prevents the degradation of I κ B α and therefore the translocation of NF- κ B to the nucleus (20, 21); JSH-23, which inhibits nuclear translocation of NF- κ B without affecting I κ B α levels (22); and BAY11-7082, an inhibitor of I κ B α phosphorylation via IKK α (23). Inhibitors with different mechanisms were chosen as the efficacy of NF- κ B inhibitors has not previously been thoroughly investigated in mollusks. These

¹ M. Goodson, *et al.*

experiments are incomplete, but are presented to aid to other researchers interested in the *E. scolopes* NF- κ B system.

MATERIALS AND METHODS

General procedures for animal experiments. Adult *E. scolopes* animals were collected from the sand flats of Oahu, Hawaii and transported and maintained as described in previous publications (24); experiments used newly hatched juveniles each kept in 4 mL (in a glass vial) or 0.5 mL (in 24-well tissue culture plates) of artificial seawater. Symbiotic animals were exposed to ~5000 CFU/mL of wild-type ES114 *V. fischeri* (25); aposymbiotic animals were not. Symbiotic status was verified with luminescence using a TD-20/20 luminometer (Turner Designs, Sunnydale, CA).

Use of MAMPs and NF- κ B inhibitors. TCT and *V. fischeri* lipid A were prepared as previously described (26, 27), with the water-phenol method used for lipid A (28). Lipid A solutions were sonicated before use as previously (27); preparations used in animal experiments were initially solubilized at 1 mg/mL in 10 mM PIPES buffer pH 6.3 before dilution. Lipid A was used at 10 ng/mL and TCT was used at 1 μ M in animal experiments; both were added directly to seawater at the time of bacterial inoculation. NF- κ B inhibitors MG132, JSH-23 and BAY11-7082 were obtained from Santa Cruz Biotechnology (Dallas, TX) and suspended in DMSO according to the manufacturer's recommendations. These compounds were added directly to seawater at the time of bacterial inoculation (with untreated animals receiving a corresponding volume of DMSO) in the case of animal experiments or, in the case of bacterial growth curves, to LBS medium at the time = 0 culture dilution.

qRT-PCR. Juvenile animals were collected indicated timepoints, or 24 h post-inoculation if not otherwise specified, and stabilized in RNAlater (Qiagen, Venio, Netherlands). Per treatment, four biological replicates of ~25 light organs were obtained by dissection. RNA extractions were performed with an RNeasy kit (Qiagen) and DNA removed with Turbo DNase (Life Technologies, Carlsbad, CA). Single-stranded cDNA was prepared from 400 ng of RNA with MMLV reverse transcriptase (Clontech, Mountain View, CA) using 5' CDS primers (5'-(T)₂₅VN-3') and diluted to a final volume of 120 μ L. Real-time PCR reactions were performed in duplicate with 4 μ L cDNA in a 20 μ L total reaction volume, using forward and reverse primers (Table A-1) at 0.25 μ M and using Sso advanced SYBR green mix (Bio-Rad, Hercules, CA). A CFX Connect qRT-PCR machine (Bio-Rad) was used with the protocol: 3 min at 94°C, 40 \times [15 s at 94°C, 20s at 59°C, 20 s at 68°C]. We used the comparative $\Delta\Delta C_q$ method to determine expression levels (29). *esikB* levels were normalized to the mean levels of control transcripts for the 40S ribosomal subunit and serine hydroxymethyltransferase (HMT), except in the experiment concerning the effects of MG132, wherein only serine HMT was used as a control gene. The animals used in timecourse and MAMP-treatment experiments (but not NF- κ B inhibitor experiments) are the same animals as were used in qRT-PCR experiments in Chapter 2.

***V. fischeri* growth curves in the presence of NF- κ B inhibitors.** ES114 (wild-type) *V. fischeri* were grown at 27°C in LBS medium until mid-log phase and diluted to $OD_{600} = 10^{-3}$ in 1 mL LBS in each well of a 24 well plate. Inhibitors were added over the range of 0-80 μ M (MG132), 0-30 μ M (JSH-23), or 0-1000 nM (BAY11-7082) and each concentration was evaluated in triplicate. Plates were incubated for 12 h, shaking, at 27°C in a TECAN GENios Pro absorbance reader, with OD_{600} readings taken at 10 min intervals.

Table A-1: Primers used for qRT-PCR experiments

Gene	Primer name	Primer sequence
EsIkB	qRTIkBF3	AGAGGACGAAGGAACATCCGAGTG
EsIkB	qRTIkBR3	TTTGTGCGTGTCCGCGAATAATAGC
Serine HMT	SerHMTqF	GTCCTGGTGACAAGAGTGCAATGA
Serine HMT	SerHMTqR	TTCCAGCAGAAAGGCACGATAGGT
40S ribosomal subunit	40SF2	AATCTCGGCGTCCTTGAGAA
40S ribosomal subunit	40SR2	GCATCAATTGCACGACGAGT

Measurement of early-stage apoptosis. MG132 (5 μ M final concentration), JSH-23 (10 μ M final concentration), or a corresponding volume of DMSO were added to the water containing *E. scolopes* at the time of bacterial inoculation. 24 h post-inoculation, animals were transferred to 1 μ g/mL acridine orange (30) in artificial seawater for 5 min, then anaesthetized in 2% ethanol in artificial seawater for 2 min. Mantles were opened and light organs evaluated for pycnotic nuclei with a LSM510 laser-scanning confocal microscope (Zeiss, Thornwood, NY).

Production of polyclonal antibodies to EsRel and EsI κ B. The *esrel* transcript (NCBI GenBank Accession No. AY956819.1) was amplified from *E. scolopes* light organ cDNA with primers NFKB-F: GACCCATGGATAACTTCAACAATCTTTATGA (containing an NcoI cut site) and NFKB-R2: GCAGTCGACATTTTCTGGTAGAAGAAAGGATG (containing a Sall cut site). The PCR product and the expression vector pET-22b(+) (EMD Millipore, Billerica, MA) were digested with restriction enzymes NcoI and Sall and ligated with T4 DNA ligase overnight (New England Biolabs, Ipswich, MA), adding a leader peptide and C-terminal His₆-tag to be expressed along with EsRel. This construct was transformed into DH5 α *E. coli*; plasmids were obtained with QiaPrep MiniPrep kit (Qiagen) and the inserts sequenced before plasmid transformation into One Shot BL21(DE3) *E. coli* (Thermo Fisher, Waltham, MA). Two SNPs resulting in amino acid substitutions relative to the published sequence were noted (P351L, K358N). The predicted mass of the recombinant EsRel was 53.7 kDa. To express the protein, cultures were grown overnight in LB-carbenicillin at 37°C, subcultured into 500 mL LB-carbenicillin and grown to mid-log phase at 37°C. Protein expression was induced for 3 h with 0.5 mM IPTG at 30°C, after which cultures were centrifuged at 4°C at 7,000 rpm [8,000 x g] for 10 min and the cell pellet stored at -20°C. The cell pellet was resuspended in 5 mL 1x PBS and

disrupted using a Sonifer cell disruptor 350 (Branson Ultrasonics, Danbury, CT) with 15 one-min continuous bursts at the microtip limit, interspersed with one min incubations on ice. The lysed cell mixture was centrifuged at 15,000 rpm at 4°C. The recombinant EsRel protein was found in the insoluble fraction, which was composed of a large volume of inclusion bodies and a small amount of cellular debris. The insoluble fraction was resuspended in 0.2 mL 1x PBS and added to 4 mL binding buffer: 20 mM sodium phosphate, 500 mM NaCl, 20 mM imidazole, 8 M urea, pH 7.4. This solubilized mixture was centrifuged a second time (at 15,000 rpm at 4°C) to remove the cellular debris. 1 mL of 50% Ni-NTA resin (Qiagen) was added to the solubilized inclusion body mixture and the mixture was rotated 1 h at 4°C. This slurry was then poured into a disposable gravity-flow column and purified via imidazole elution, with two washes in 20 mM imidazole followed by an elution in 45 mM imidazole. Urea was removed by dialysis using a Slide-A-Lyzer Dialysis Cassette (Thermo Scientific) and protein was concentrated with an Amicon Ultra-15 Centrifugal Filter Device (EMD Millipore). The recombinant EsRel was inoculated into rabbits by Covance (Princeton, NJ) for a 77 day antibody production protocol and serum collected at the conclusion of the protocol.

The production of the EsIkB antibody differed from the above in the following ways. The *esikB* transcript (NCBI accession number AY956818.1) was amplified from *E. scolopes* light organ cDNA with primers IkBcdsF: TTTCACTACGGACGAATCTCTTT and IkBcdsR: TCAATCAGACACGAGACCATTT. The purified PCR product was ligated into the plasmid pCR4-TOPO and transformed into TOP10 competent cells. One SNP resulting in an amino acid change relative to the published sequence (N22S) was noted. This *esikB* product was subcloned with amplification by primers IkBpET22bF2: ATTTACCATGGCGTATACTACGACTGATCGA (containing an NcoI cut site) and

IkBpET22bR: ATACACTCGAGGATTGGTTGGCCATTGATA (containing an XhoI cut site), digested with restriction enzymes NcoI and XhoI and ligation into pET-22b(+). This construct was again transformed into TOP10 cells and sequenced before being transformed into One Shot BL21(DE3) *E. coli*. The predicted mass of the recombinant EsIkB was 38.7 kDa. Due to lower yield for this protein, 2 L of culture was used for the over-expression, and the cell pellet was resuspended in 15 mL 1x PBS prior to sonication. The buffers used to purify recombinant EsIkB included 100 mM sodium phosphate, 10 mM Tris-Cl, 8M urea, with the binding buffer at pH 8.0, wash buffer at pH 6.3, and elution buffer at pH 5.9. Otherwise, protein expression and antibody production were performed as for EsRel.

Western blotting. The recombinant EsRel and EsIkB solutions described above were stored at -20°C after two-fold dilution in glycerol, then diluted 100-fold in PBS shortly before loading of the gel. To prepare total soluble *E. scolopes* protein, two intact *E. scolopes* juveniles were anaesthetized on ice before being homogenized in 100 μ L PBS with a ground glass mortar and pestle and centrifuged 20 min at 10,000 x g to remove any insoluble material.

Approximately 23 μ g of protein per well were loaded into a 4-12% Bis-Tris NuPAGE gel (Life Technologies) and run with MOPS buffer. Transfer to a nitrocellulose membrane was carried out with a Mini Trans-Blot system (Bio-Rad) according to manufacturer recommendations. The membrane was blocked overnight at 4°C in 4% milk in TTBS (tween-20 tris-buffered saline; 0.1 M Tris, 2.5 M NaCl, .05% tween-20, pH 7.5) and transferred into primary antibody incubation in 1% milk in TTBS for 1 h at room temperature with 1:250 anti-EsRel serum, 1:50 anti-EsIkB serum, or a corresponding volume of pre-immune serum. 3 x 10 min rinses in TTBS were followed with a secondary antibody incubation with 1:3000 goat-anti-rabbit-HRP antibody (Jackson ImmunoResearch, West Grove, PA) and StrepTactin (Life Technologies) for standard

detection, a further 3 x 10 min rinses in TTBS and 1 x 10 min rinse in TBS, then revelation with Pierce ECL Western Blotting Substrate (Life Technologies).

Immunocytochemistry. Juvenile animals, apo- and symbiotic, were collected at 24 h post-inoculation and fixed in 4% paraformaldehyde in marine PBS (mPBS) and processed for immunocytochemistry based on previous protocols (31). The primary antibody incubation was one of the following: anti-EsRel 1:50 diluted for 9 days, preimmune serum diluted 1:50 for 9 days, or anti-EsPGRP1 (32) diluted 1:1000 for 7 days; secondary antibody (goat anti-rabbit FITC), rhodamine phalloidin (for actin), and TOTO-3 staining (for DNA) were performed as previously described (31). Collection, fixation, permeabilization, blocking, and rinses of adult *E. scolopes* hemocytes were performed according to established protocols (33). The following specific antibodies and stains were used: rhodamine-phalloidin at a 1:40 dilution overnight (during the permeabilization step); anti-EsChit1 antibody (33) at 1:500, anti-EsRel at 1:125, or pre-immune serum at 1:125 for 3 h; and lastly goat-anti-rabbit-FITC (Jackson ImmunoResearch) at 1:40 for 45 min. All tissues prepared for immunocytochemistry were examined with an LSM510 laser-scanning confocal microscope (Zeiss, Thornwood, NY).

Statistics. Transcript levels in qRT-PCR experiments were log-transformed to provide for normality prior to statistical analysis. Quantitative comparisons between treatments in qRT-PCR experiments and apoptosis assays were made with ANOVA followed by *post hoc* pairwise comparisons with Tukey multiple comparisons of means.

RESULTS

***esikB* transcript regulation is similar to that of *eslbp1*.** The gene for IκB is, in mammals, itself a target of transcriptional regulation by NF-κB, forming an autoregulatory

feedback loop (19). We evaluated the level of transcript of *E. scolopes esikB*. The pattern of regulation was qualitatively very similar to that of *eslbp1* (Chapter 2), with significant up-regulation occurring 12 h into symbiosis (Fig. A-1) or upon 24 h of treatment with 1 μ M TCT (Fig. A-2). The magnitude of this increase was approximately 1.5-4 fold. 10 ng/mL *V. fischeri* lipid A did not affect transcript levels significantly.

NF- κ B inhibitors affect *V. fischeri* growth at high concentrations. The possibility of false-positive phenotypes in symbiotic animals treated with NF- κ B inhibitors, due to bacterial inhibition rather than any effect of inhibitor on host cellular machinery, motivated us to examine the effect of NF- κ B inhibitors on *V. fischeri* in culture. Concentrations tested were 0-80 μ M MG132, 0-30 μ M JSH-23, and 0-1000 nM BAY11-7082. An inhibitory effect was grossly apparent at high concentrations of JSH-23 and especially BAY11-7082 (Fig. A-2). Of the concentrations of inhibitors used animal experiments (5 μ M MG132, 10 μ M JSH-23, and 40 nM BAY11-7082) only growth curve for JSH-23 (Fig. A-2B) suggests the potential for some slight bacterial growth inhibition.

5 μ M MG132 does not affect *esikB* transcript levels in symbiotic *E. scolopes*.

Inhibiting the proteasomal degradation of I κ B has been used as a method of NF- κ B inhibition in other systems (21). MG132 does not appear to have this effect in *E. scolopes*, however, at a concentration of 5 μ M added directly to seawater. 24 h post-inoculation, symbiotic light organs, whether treated with MG132 or a corresponding volume of DMSO, had *esikB* transcript levels upregulated ~4-fold over DMSO-treated aposymbiotic light organs (Fig. A-3).

A proteasome inhibitor, but not another NF- κ B inhibitor, increases apoptosis in the appendages of *E. scolopes* light organs. NF- κ B-driven transcription can have a strong influence on cell survival, often increasing survival but sometimes implicated in cell death. A proteasome

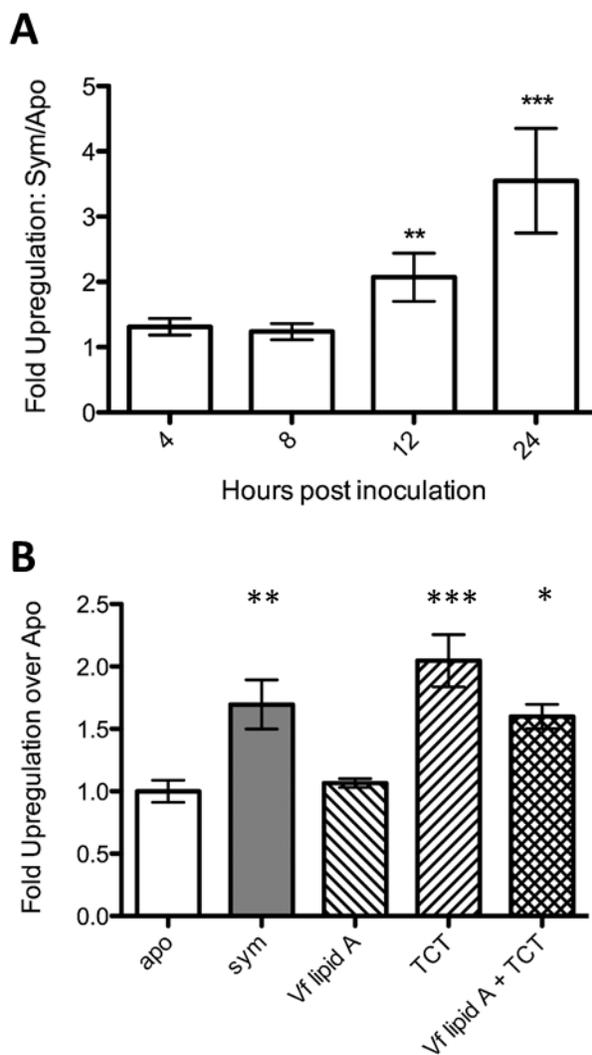


FIG A-1 *esikb* transcript induction by *V. fischeri* colonization or MAMPs. (A) *esikb* normalized transcript levels in *E. scolopes* light organs at 4, 8, 12, and 24 h post-inoculation given as a ratio of transcript levels in symbiotic light organs to those in aposymbiotic light organs. (B) Light organ *eslbp1* transcript levels at 24h for aposymbiotic, symbiotic, *V. fischeri* lipid A (10 ng/mL)- and/or TCT (1 μ M)-treated light organs. All statistical comparisons are shown relative to the corresponding aposymbiotic treatment. P values: *** < .001 < ** < .01 < * < .05. SEM of four replicates shown.

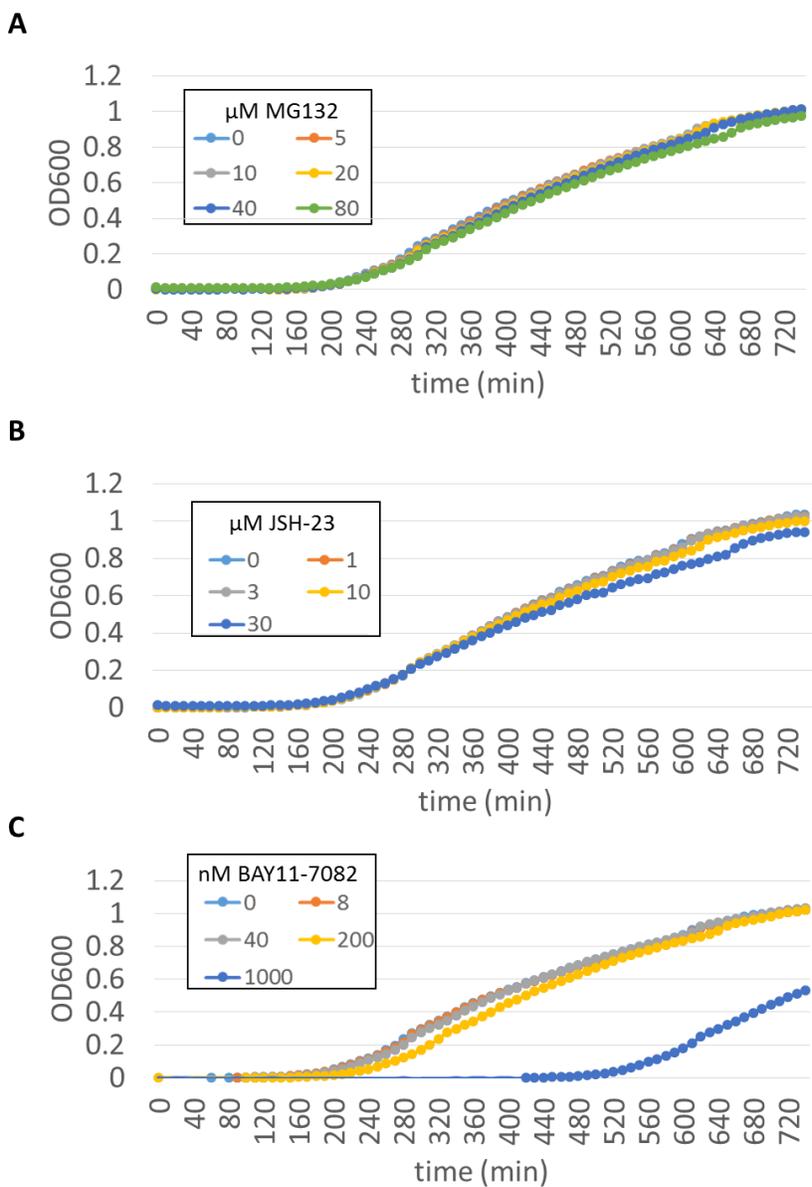


FIG A-2 *V. fischeri* growth rates in the presence of NF- κ B inhibitors. *V. fischeri* were grown in LBS medium from an OD₆₀₀ of 10^{-3} in the presence of indicated concentrations of MG132 (A), JSH-23 (B), or BAY 11-7082 (C) for 12 h.

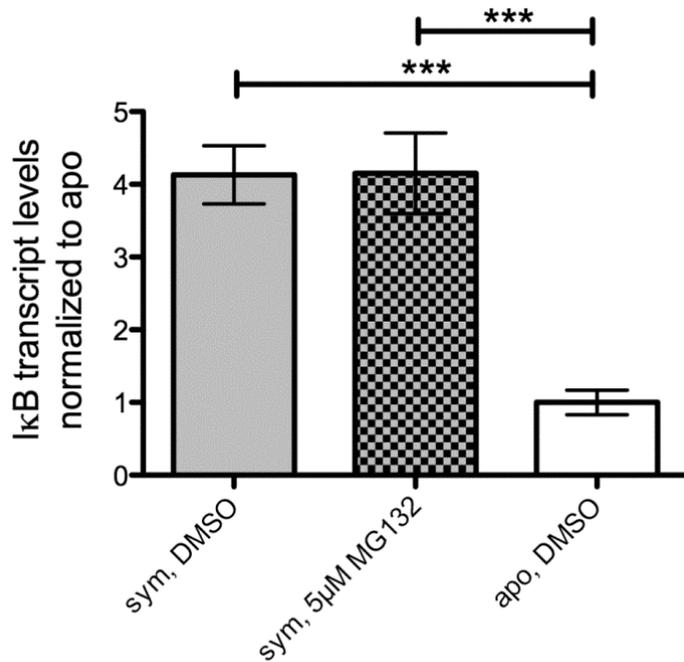


FIG A-3 Effect of 5 µM MG132 on normalized *esikB* transcript levels in light organs of 24 h post-inoculation symbiotic *E. scolopes*. Shown for comparison are apo- and symbiotic *E. scolopes* untreated with MG132, but instead with a corresponding volume of DMSO. P values: *** < .001. SEM of four replicates shown.

inhibitor, lactacystin, is known to increase apoptosis in *E. scolopes* light organs (34). We sought to verify this phenotype with another proteasome inhibitor, as well as test NF- κ B inhibitors working by other mechanisms. While MG132 increased apoptosis in aposymbiotic light organs at 24 h (Fig. A-4A) (symbiotic animals were included in this experiment as well, but in both inhibitor-treated and untreated light organs, pycnotic nuclei were too abundant to be counted reliably), no significant effect on apoptosis in apo- or symbiotic light organs was seen upon JSH-23 treatment (Fig. A-4B). 40 nM BAY11-7082 did not appear to affect apoptosis in symbiotic light organs, though the more sensitive comparison of treated and untreated aposymbiotic light organs was not conducted (data not shown).

Western blots indicate some nonspecific binding by anti-EsRel and anti-EsI κ B antibodies in whole-animal protein preparations. The Western blot for EsRel (Fig. A-5) shows a number of bands of greatly varying molecular weight in the total soluble *E. scolopes* protein preparation lane, with no band clearly dominant. It is not clear if one of these is *E. scolopes* EsRel (predicted molecular weight, 51.9 kDa). The Western for the recombinant protein prep shows a band at >50 kDa presumed to be the recombinant EsRel (predicted molecular weight, 53.7 kDa), with a lower molecular weight band of unknown identity also present. The Western blot for EsI κ B (Fig. A-6) shows a good deal of signal at high molecular weight in the total soluble *E. scolopes* protein preparation lane, as well as a distinct band that may represent EsI κ B (predicted molecular weight, 38 kDa); this band is slightly heavier than the 37 kDa marker, and lighter than the presumed recombinant, tagged EsI κ B (38.7 kDa).

An antibody to EsRel recognizes an antigen in *E. scolopes* hemocytes, but not whole light organs. Previous, unpublished experiments with a commercial anti-p65 antibody showed

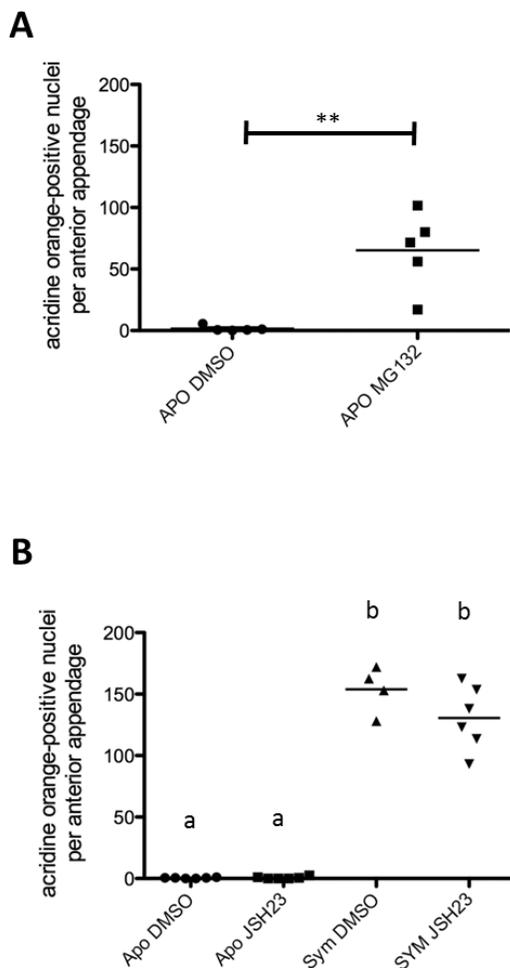


FIG A-4 Apoptosis in light organ anterior appendages in NF- κ B inhibitor-treated animals. (A) Apo animals treated with 5 μ M MG132 or corresponding volume DMSO and stained with acridine orange for pycnotic (early-stage apoptotic) nuclei. P values: ** < .01. (B) Apo- and symbiotic animals treated with 10 μ M JSH-23 or corresponding volume DMSO, stained as in A. “a” and “b” indicate statistically distinguishable groups of treatments ($p < .001$). Each data point is the average of two anterior appendages of an individual light organ and the horizontal line is the mean for each treatment.

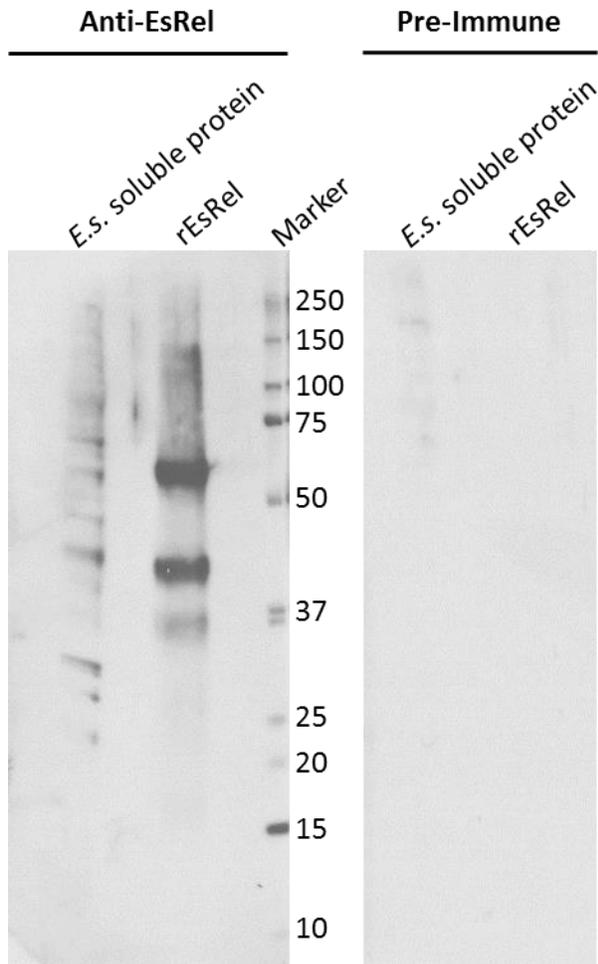


FIG A-5 Anti-EsRel Western blot for both *E. scolopes* total soluble protein (~23 μ g) and recombinant EsRel expressed in *E. coli*, run on 4-12% Bis-Tris gel. Primary antibody treatment was 1:250 anti-EsRel rabbit serum in 1% milk for 1 h at room temperature, or corresponding volume of pre-immune serum. Molecular weights in kDa.

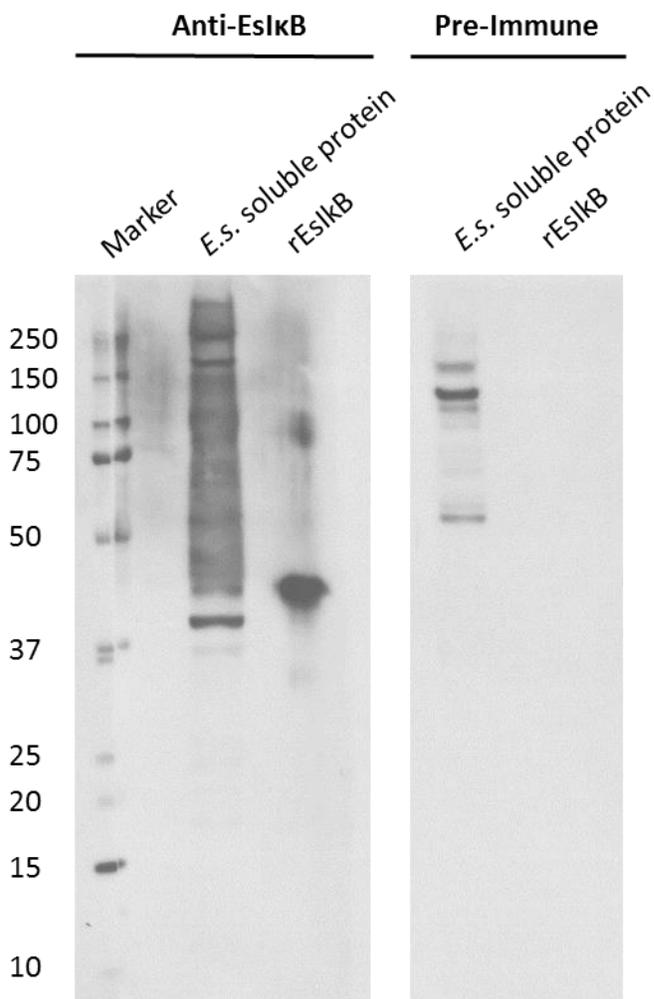


FIG A-6 Anti-EsIkB Western blot for both *E. scolopes* total soluble protein (~23 μ g) and recombinant EsIkB expressed in *E. coli*, run on 4-12% Bis-Tris gel. Primary antibody treatment was 1:50 anti-EsIkB rabbit serum in 1% milk for 1 h at room temperature, or corresponding volume of pre-immune serum. Molecular weights in kDa.

antibody signal in the nuclei of light organ appendage cells in symbiotic, but not aposymbiotic, animals. Nuclear translocation of EsRel (or the *E. scolopes* p105 homolog) would strongly suggest activation of the NF- κ B system. We inoculated rabbits to produce EsRel-specific polyclonal antibodies, but did not observe any nuclear translocation, or indeed, clear antibody-specific signal anywhere in the appendages or crypt-bordering cells in the *E. scolopes* light organ, whether symbiotic or aposymbiotic (Fig A-7B, D, E, F). Even increasing the gain from 500 to 700, at which setting signal could no longer be reliably distinguished from pre-immune background levels, did not suggest nuclear translocation in symbiotic light organs (Fig A-7C). Signal appeared to be concentrated primarily at the cell surface, less prominently in the cytoplasm, and least of all in the nucleus.

Constitutive activation of NF- κ B has been reported in human macrophages (35), making hemocytes, the immune blood cells of *E. scolopes*, obvious candidates for EsRel expression. Indeed, unlike juvenile light organ tissues, hemocytes taken from light organs reacted with the EsRel antibody (Fig A-8). In some cases, the staining pattern was strongly reminiscent of a hemocyte nucleus, but no nuclear stain was used in this experiment, so this thought is conjectural. Signal also appeared dimly throughout the cytoplasm and occasionally in sharp puncta.

DISCUSSION

An increase in *esikB* transcript in the light organ over the course of colonization or TCT treatment (Fig. A-1) is suggestive of activation of the NF- κ B system, and would be consistent with the observations of *ikB* behavior upon NF- κ B activation from other systems (19). This explanation is, however, only one possible interpretation of the *esikB* transcript levels. An

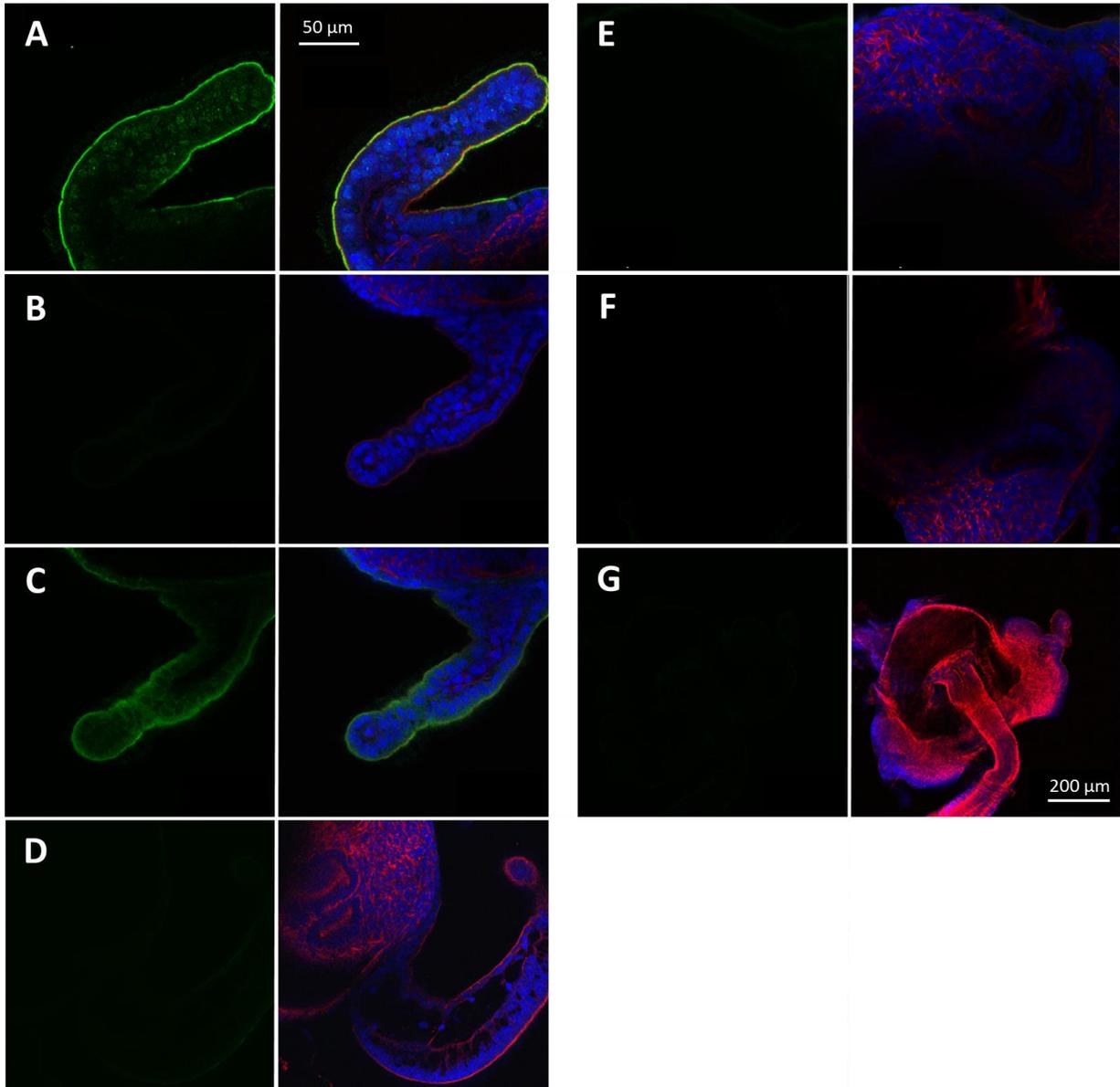


FIG A-7 Anti-EsRel antibody staining of light organs 24 h post-inoculation. Left panels show antibody staining (green) alone, right panels, merged with rhodamine phalloidin (actin, red) and TOTO-3 (DNA, blue). (A) Example of a nuclear-staining antibody, that to the peptidoglycan recognition protein EsPGRP1 (32). (B) Anterior appendage of a symbiotic light organ, stained with anti-EsRel antibody, gain 500 (corresponding to undetectable signal in preimmune-treated light organs, as in G). (C) As B, with gain increased from 500 to 700. At this gain setting, signal can no longer be reliably distinguished from pre-immune background levels. (D) Anterior appendage of an aposymbiotic light organ, stained with anti-EsRel antibody. Two crypt spaces are also visible in this image. (E) Crypt spaces of symbiotic light organ stained with anti-EsRel antibody. (F) Crypt spaces of apo light organ stained with anti-EsRel antibody. (G) Overview of symbiotic light organ treated with preimmune serum.

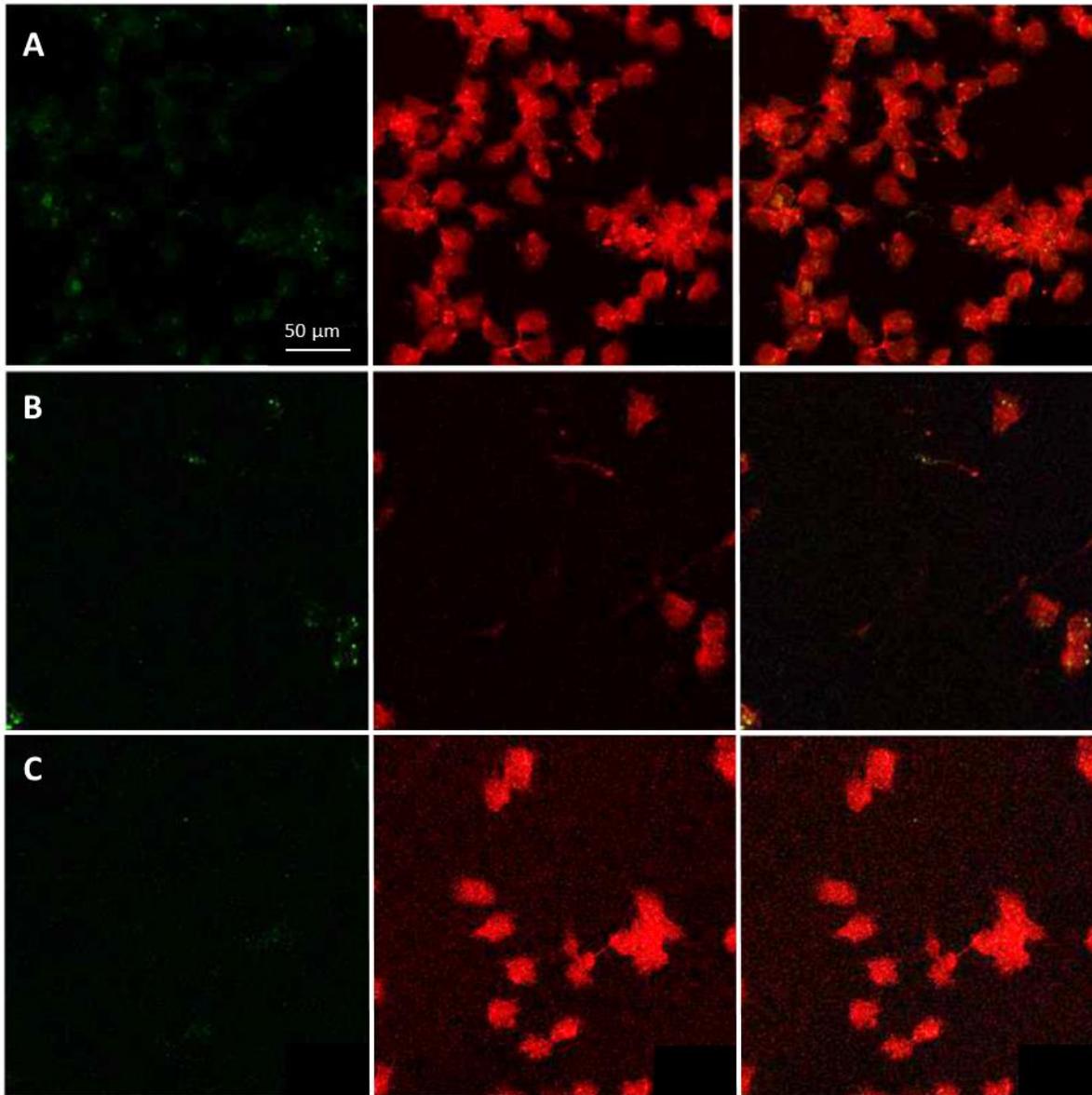


FIG A-8 Anti-EsRel antibody staining of adult *E. scolopes* hemocytes. Anti-EsRel, green; rhodamine phalloidin (actin stain), red. Rightmost panel merges both channels. (A) Anti-EsRel antibody-treated fixed hemocytes. (B) Example of a cytoplasm-staining antibody, that to the chitotriosidase EsChit1 (33). (C) Fixed hemocytes treated with pre-immune serum.

increased level of *esikB* transcript, and therefore protein, caused by some other transcription factor could reasonably be expected to reduce NF- κ B activity, rather than increase it. Other symbiotic bacteria have been known to downregulate NF- κ B activity (13). In either case, the upregulation of *esikB* is strong evidence of modulation of the NF- κ B system over the course of the colonization and early development of the light organ, and merits further investigation.

We were unable to affect *esikB* levels with NF- κ B inhibitors (Fig. A-3). The cause of this result is unclear and could range from an inappropriate method of delivery (the NF- κ B inhibitors may be poorly absorbed by host cells, or unstable in seawater or light) to differences in *E. scolopes* cellular machinery that result in these compounds having different effects in *E. scolopes* than are seen in mammals. It is also possible that inhibitor levels are suboptimal. A qRT-PCR experiment using MG132 at 20 μ M rather than 5 μ M is in process. A preliminary experiment evaluating the effect of 40 nM BAY11-7082 (which caused no animal fatalities, but did cause some treated animals to appear sluggish) on *esikB* transcript levels was completed; while normalized *esikB* levels appeared to increase, this was entirely due to higher, variable *serHMT* Ct values in inhibitor-treated animals, rather than lower *esikB* Ct values, and thus we regard this experiment with skepticism; these data are not shown here. JSH-23, of which 10 μ M appears to be near the highest concentration that can be used without greatly affecting *V. fischeri* growth (Fig. A-2B), has not yet been tested for its effect on *esikB* transcript levels.

The previously established (34) relationship between proteasome inhibitors and apoptosis in the *E. scolopes* light organ was confirmed (Fig. A4-A), but no such effect was seen with the proteasome-independent NF- κ B inhibitor JSH-23 (Fig. A4-B) or BAY11-7082 (data not shown). These findings provide some evidence that the apoptosis caused by proteasome inhibitors is due to other causes than NF- κ B inhibition; MG132 in other systems has been shown to cause

apoptosis by several mechanisms (36). However, until such time as evidence of NF- κ B regulation by a tested inhibitor is found (at either the transcript or protein level), this possibility will remain largely conjectural. It is possible that NF- κ B inhibition does alter levels of apoptosis in the *E. scolopes* light organ, independent of other activities of the proteasome, and the correct inhibitor to cleanly disrupt this effect has not yet been tested.

A thorough examination of NF- κ B activity in *E. scolopes* will require information regarding proteins such as EsRel and EsI κ B, rather than simply their transcripts. These efforts have been stymied in the past by a lack of useable antibodies. We created antibodies to recombinant EsRel and EsI κ B, but their sensitivity and specificity are not clear. The anti-EsI κ B Western blot (Fig. A-6) reveals a promising band for EsI κ B but requires further optimization; a Western blot eventually may allow visualization of light organ EsI κ B protein levels over colonization. In contrast, the anti-EsRel Western blot (Fig. A-5) revealed no obvious candidate band; a protein preparation of light organs alone may allow reduction of non-specific signal.

We did not see evidence of EsRel movement to the nucleus in *E. scolopes* light organ appendages over colonization, or indeed, clear signal anywhere in the light organ (Fig. A-7). Evidence was noted, however, of EsRel in hemocytes (Fig. A-8), though in light of Western blot results (Fig. A-5), the possibility that this represents off-target antibody staining must be entertained. Some of this signal was suggestive of nuclear staining, but this experiment remains to be repeated successfully with a nuclear stain included. Macrophages are sometimes visible in the light organ, including the blood sinus of the appendages, but no EsRel signal was noted in the light organ that could be attributable to hemocytes. This could be due to sheer coincidence, may relate to choice of tissue (light organs and circulating hemocytes), or else may relate to the distinction in age (24 h post-inoculation vs. adults) between the animals examined. Continuing

the timecourse of *esikB* transcript levels (Fig. A-1) beyond 24 h may provide clues regarding promising timepoints to sample light organs and hemocytes for ICC.

In summary, strong evidence exists for the regulation of NF- κ B activity over the course of *V. fischeri* colonization and early-stage development in the post-hatching *E. scolopes* light organ. There is no evidence of successful modulation of NF- κ B activity by commercially available pathway inhibitors, and it is suggested that the pro-apoptotic activity of proteasome inhibitors may be through mechanisms other than NF- κ B regulation. Evidence at the protein level for EsRel in the light organ was not found, but hemocytes were implicated as potentially important in the *E. scolopes* NF- κ B system.

REFERENCES

1. **Vallabhapurapu S, Karin M.** 2009. Regulation and function of NF-kappaB transcription factors in the immune system. *Annu Rev Immunol* **27**:693-733.
2. **Arancibia SA, Beltrán CJ, Aguirre IM, Silva P, Peralta AL, Malinarich F, Hermoso MA.** 2007. Toll-like receptors are key participants in innate immune responses. *Biol Res* **40**:97-112.
3. **Hinz M, Scheidereit C.** 2014. The I κ B kinase complex in NF- κ B regulation and beyond. *EMBO Rep* **15**:46-61.
4. **Sun SC.** 2011. Non-canonical NF- κ B signaling pathway. *Cell Res* **21**:71-85.
5. **Belvin MP, Jin Y, Anderson KV.** 1995. Cactus protein degradation mediates *Drosophila* dorsal-ventral signaling. *Genes Dev* **9**:783-793.
6. **Moussian B, Roth S.** 2005. Dorsoventral axis formation in the *Drosophila* embryo--shaping and transducing a morphogen gradient. *Curr Biol* **15**:R887-899.
7. **Lemaitre B, Hoffmann J.** 2007. The host defense of *Drosophila melanogaster*. *Annu Rev Immunol* **25**:697-743.
8. **Leulier F, Lemaitre B.** 2008. Toll-like receptors--taking an evolutionary approach. *Nat Rev Genet* **9**:165-178.
9. **Nogueira L, Ruiz-Ontañón P, Vazquez-Barquero A, Moris F, Fernandez-Luna JL.** 2011. The NF κ B pathway: a therapeutic target in glioblastoma. *Oncotarget* **2**:646-653.

10. **Siggers RH, Hackam DJ.** 2011. The role of innate immune-stimulated epithelial apoptosis during gastrointestinal inflammatory diseases. *Cell Mol Life Sci* **68**:3623-3634.
11. **Zaph C, Troy AE, Taylor BC, Berman-Booty LD, Guild KJ, Du Y, Yost EA, Gruber AD, May MJ, Greten FR, Eckmann L, Karin M, Artis D.** 2007. Epithelial-cell-intrinsic IKK-beta expression regulates intestinal immune homeostasis. *Nature* **446**:552-556.
12. **Richardson WM, Sodhi CP, Russo A, Siggers RH, Afrazi A, Gribar SC, Neal MD, Dai S, Prindle T, Branca M, Ma C, Ozolek J, Hackam DJ.** 2010. Nucleotide-binding oligomerization domain-2 inhibits toll-like receptor-4 signaling in the intestinal epithelium. *Gastroenterology* **139**:904-917, 917.e901-906.
13. **Kelly D, Campbell JI, King TP, Grant G, Jansson EA, Coutts AG, Pettersson S, Conway S.** 2004. Commensal anaerobic gut bacteria attenuate inflammation by regulating nuclear-cytoplasmic shuttling of PPAR-gamma and RelA. *Nat Immunol* **5**:104-112.
14. **Goodson MS, Kojadinovic M, Troll JV, Scheetz TE, Casavant TL, Soares MB, McFall-Ngai MJ.** 2005. Identifying components of the NF-kappaB pathway in the beneficial *Euprymna scolopes-Vibrio fischeri* light organ symbiosis. *Appl Environ Microbiol* **71**:6934-6946.
15. **Collins AJ, Schleicher TR, Rader BA, Nyholm SV.** 2012. Understanding the role of host hemocytes in a squid/vibrio symbiosis using transcriptomics and proteomics. *Front Immunol* **3**:91.
16. **Nyholm SV, Deplancke B, Gaskins HR, Apicella MA, McFall-Ngai MJ.** 2002. Roles of *Vibrio fischeri* and nonsymbiotic bacteria in the dynamics of mucus secretion during symbiont colonization of the *Euprymna scolopes* light organ. *Appl Environ Microbiol* **68**:5113-5122.
17. **Koropatnick TA, Kimbell JR, McFall-Ngai MJ.** 2007. Responses of host hemocytes during the initiation of the squid-*Vibrio* symbiosis. *Biol Bull* **212**:29-39.
18. **Altura MA, Stabb E, Goldman W, Apicella M, McFall-Ngai MJ.** 2011. Attenuation of host NO production by MAMPs potentiates development of the host in the squid-vibrio symbiosis. *Cell Microbiol* **13**:527-537.
19. **Sun SC, Ganchi PA, Ballard DW, Greene WC.** 1993. NF-kappa B controls expression of inhibitor I kappa B alpha: evidence for an inducible autoregulatory pathway. *Science* **259**:1912-1915.
20. **Chen Z, Hagler J, Palombella VJ, Melandri F, Scherer D, Ballard D, Maniatis T.** 1995. Signal-induced site-specific phosphorylation targets I kappa B alpha to the ubiquitin-proteasome pathway. *Genes Dev* **9**:1586-1597.
21. **Lin KI, Baraban JM, Ratan RR.** 1998. Inhibition versus induction of apoptosis by proteasome inhibitors depends on concentration. *Cell Death Differ* **5**:577-583.
22. **Shin HM, Kim MH, Kim BH, Jung SH, Kim YS, Park HJ, Hong JT, Min KR, Kim Y.** 2004. Inhibitory action of novel aromatic diamine compound on lipopolysaccharide-

- induced nuclear translocation of NF-kappaB without affecting IkappaB degradation. FEBS Lett **571**:50-54.
23. **Pierce JW, Schoenleber R, Jesmok G, Best J, Moore SA, Collins T, Gerritsen ME.** 1997. Novel inhibitors of cytokine-induced IkappaBalpha phosphorylation and endothelial cell adhesion molecule expression show anti-inflammatory effects in vivo. J Biol Chem **272**:21096-21103.
 24. **McFall-Ngai M, Montgomery MK.** 1990. The Anatomy and Morphology of the Adult Bacterial Light Organ of *Euprymna scolopes* Berry (Cephalopoda:Sepiolidae). Biological Bulletin **179**:332-339.
 25. **Boettcher KJ, Ruby EG.** 1990. Depressed light emission by symbiotic *Vibrio fischeri* of the sepiolid squid *Euprymna scolopes*. J Bacteriol **172**:3701-3706.
 26. **Koropatnick TA, Engle JT, Apicella MA, Stabb EV, Goldman WE, McFall-Ngai MJ.** 2004. Microbial factor-mediated development in a host-bacterial mutualism. Science **306**:1186-1188.
 27. **Foster JS, Apicella MA, McFall-Ngai MJ.** 2000. *Vibrio fischeri* lipopolysaccharide induces developmental apoptosis, but not complete morphogenesis, of the *Euprymna scolopes* symbiotic light organ. Dev Biol **226**:242-254.
 28. **Apicella MA, Griffiss JM, Schneider H.** 1994. Isolation and characterization of lipopolysaccharides, lipooligosaccharides, and lipid A. Methods Enzymol **235**:242-252.
 29. **Pfaffl MW.** 2001. A new mathematical model for relative quantification in real-time RT-PCR. Nucleic Acids Res **29**:e45.
 30. **Montgomery MK, McFall-Ngai M.** 1994. Bacterial symbionts induce host organ morphogenesis during early postembryonic development of the squid *Euprymna scolopes*. Development **120**:1719-1729.
 31. **Chun CK, Troll JV, Koroleva I, Brown B, Manzella L, Snir E, Almabrazi H, Scheetz TE, Bonaldo MeF, Casavant TL, Soares MB, Ruby EG, McFall-Ngai MJ.** 2008. Effects of colonization, luminescence, and autoinducer on host transcription during development of the squid-vibrio association. Proc Natl Acad Sci U S A **105**:11323-11328.
 32. **Troll JV, Adin DM, Wier AM, Paquette N, Silverman N, Goldman WE, Stadermann FJ, Stabb EV, McFall-Ngai MJ.** 2009. Peptidoglycan induces loss of a nuclear peptidoglycan recognition protein during host tissue development in a beneficial animal-bacterial symbiosis. Cell Microbiol **11**:1114-1127.
 33. **Schwartzman JA, Koch E, Heath-Heckman EA, Zhou L, Kremer N, McFall-Ngai MJ, Ruby EG.** 2015. The chemistry of negotiation: rhythmic, glycan-driven acidification in a symbiotic conversation. Proc Natl Acad Sci U S A **112**:566-571.
 34. **Kimbell JR, Koropatnick TA, McFall-Ngai MJ.** 2006. Evidence for the participation of the proteasome in symbiont-induced tissue morphogenesis. Biol Bull **211**:1-6.

35. **Pagliari LJ, Perlman H, Liu H, Pope RM.** 2000. Macrophages require constitutive NF-kappaB activation to maintain A1 expression and mitochondrial homeostasis. *Mol Cell Biol* **20**:8855-8865.
36. **Guo N, Peng Z.** 2013. MG132, a proteasome inhibitor, induces apoptosis in tumor cells. *Asia Pac J Clin Oncol* **9**:6-11.

Appendix B

Genome Walking in *E. scolopes*

PREFACE:

This appendix and the experiments described herein are entirely the work of Benjamin Krasity.

INTRODUCTION

Though work on the matter is ongoing, *Euprymna scolopes* does not have a published genome. In fact, no cephalopod has a published genome, though there are other mollusks, including *Lottia gigantea* and *Crassostrea gigas*, which do (1, 2). Almost all information on the sequences of genes and proteins in *E. scolopes* derives ultimately from transcriptomic databases (for prominent examples see (3, 4)). A general strategy has been to identify cDNA transcripts of interest and confirm them by PCR and sequencing the PCR product. If necessary, the 5' and 3' ends of incomplete transcripts may be more fully elucidated by RACE (rapid amplification of cDNA ends)-PCR. One result of the paucity of studies on *E. scolopes* genomic DNA is that, while multiple observational studies have provided plentiful information about the expression profiles of the *E. scolopes* light organ over the course of colonization, very little is known about the biochemical mechanics of transcriptional regulation in *E. scolopes*. (Another consequence is a lack of insight into *E. scolopes* RNA editing, including the processing of introns, though extensive RNA editing has been reported for other cephalopods (5).) Transcripts for transcription factors in the NF- κ B family have been identified in *E. scolopes* (6). These transcription factors are of interest for a hypothesized role in detection of *V. fischeri* microbe-associated molecular pattern (MAMP) signaling and/or the apoptotic process that occurs at the light organ surface upon colonization. For further information on the NF- κ B sequences in *E. scolopes* and the potential role NF- κ B in microbe-driven animal development, see Appendix A.

We sought to use genome walking techniques (reviewed in (7)) to investigate two *E. scolopes* genes for potential NF- κ B binding sites. These genes are *esikB*, believed to be an inhibitor of NF- κ B (see Appendix A), and *eslbp1*, an LBP/BPI-family protein (see Chapter 2). In mammals, *ikB* is the target of NF- κ B-driven transcription, part of a negative feedback loop by

means of which NF- κ B eventually inactivates itself (8). LBP is also believed to be regulated by NF- κ B in mammals (9).

MATERIALS AND METHODS

Genome walking methods. *E. scolopes* Genomic DNA was kindly provided by S. Nyholm (University of Connecticut, Storrs, CT) (10). A library of digested DNA was prepared with a Genome Walker kit (Clontech, Mountain View, CA). Genome walking primers were designed near the 5'-end of the known transcript sequences for these genes. In the case of *eslbp1*, primers were initially based on the published transcript sequence (NCBI GenBank accession number JF514880.1), but initial failure to obtain genome walking products motivated a redesign of these primers based on the longer (upstream of the ORF) transcript sequences for *eslbp1* in two transcriptomic databases¹. For *esikB*, primers designed in the 5'-untranslated region (UTR) of the published sequence (GenBank AY956818.1) were able to produce a genome walking product, and a second set of primers was designed based on this newfound sequence to walk further upstream of *esikB*. Promoter regions were amplified with kit primers and gene-specific primers with Advantage 2 polymerase (Clontech) according to manufacturer recommendations. Primers for traditional PCR were designed in the newly walked promoter sequence, facing towards the ORF to allow for amplification and confirmation of the sequence obtained; traditional PCR was carried out with Platinum Taq High Fidelity (Thermo Fisher, Waltham, MA) and the template was the EcoRV-digested genomic DNA prepared with the Genome Walker kit. All gene-specific primers used in this study are given in Table B-1. PCR products were ligated into the plasmid pCR4-TOPO and transformed into TOP10 competent cells (Thermo Fisher); plasmids were

¹ These databases are from light organs of four-week-old animal (Kremer, N. *et al.*) and from multiple tissues in one-day-old animals (Moriano-Gutierrez, S. *et al.*), unpublished.

Table B-1: Primers used for genome walking and related procedures

Gene	Primer name	Use	Primer sequence
<i>eslbp1</i>	LBP1WALK1	Outer RACE primer	AAATTCAGACCTCCGTTTGTGCGCCATA
<i>eslbp1</i>	LBP1WALK2	Nested RACE primer	AAAGCAAGTTCATTGGCGTAGTCGAGA
<i>eslbp1</i>	LBP1WALK3	Outer gene-specific walking primer; confirmatory traditional PCR	CCAAGTGTATTGCTTGGGTTTTACTCA
<i>eslbp1</i>	LBP1WALK4	Nested gene-specific primer	AAAACATGCCAGAGTGATACAAACCAC
<i>eslbp1</i>	LBP1WKF2	Confirmatory traditional PCR	GTCAAAAGTGGCACGCTCATACATTA
<i>esikB</i>	IkBWALK1	Outer gene-specific primer for first walking trial; confirmatory traditional PCR	CCCCTTCTTCGTCTTGATGGTGTAACA
<i>esikB</i>	IkBWALK2	Nested gene-specific primer for first walking trial	AATCCTTCACAGTCGCAATCCAAGTCG
<i>esikB</i>	IkBWALK3	Outer gene-specific primer for second walking trial	TACGACCTCGAAATTCATTCTCGTGT
<i>esikB</i>	IkBWALK4	Nested gene-specific primer for second walking trial	TGTCCCATATTTGAACATGGTAGCGTT
<i>esikB</i>	IkBWKF2	Confirmatory traditional PCR	ACCTTTCCCAAATGAGAGTGATGCAGT

obtained with a QiaPrep MiniPrep kit (Qiagen, Venlo, Netherlands) and the inserts sequenced. Obtained sequences were analyzed for possible transcription factor binding sites with PROMO (11, 12).

5' Rapid amplification of cDNA ends (RACE) for *eslbp1* transcript. To ensure that the sequence used for design of the *eslbp1* walking primers was indeed associated with the *eslbp1* transcript, and not the product of a sequence assembly error in the RNA-Seq database, primers within the EsLBP1 ORF were designed for a 5' RACE with the aim of connecting the ORF to the genome walking primer sites. To obtain cDNA, ~25 light organs were dissected from juvenile animals and stabilized in RNAlater (Thermo Fisher). RNA extractions were performed with an RNeasy kit (Qiagen, Venlo, Netherlands) and DNA removed with Turbo DNase (Thermo Fisher). RNA was processed for 5'-RACE with a GeneRacer kit (Thermo Fisher) and single-stranded cDNA was prepared from 400 ng of RACE-ready RNA with MMLV reverse transcriptase (Clontech, Mountain View, CA) using 5' CDS primers (5' - (T)25VN-3'). RACE was performed with Platinum Taq High Fidelity (Thermo Fisher) and sequenced as indicated above. All gene-specific primers used in this study are given in Table B-1.

RESULTS

The regions upstream of *esikB* and *eslbp1* contain putative NF- κ B binding elements.

The sequences upstream of the start codon obtained for *esikB* and *eslbp1* are given in Fig. B-1A and B, respectively; these are contigs of sequence obtained by genome walking in this study and the published transcript (*esikB*) or RACE product from this study (*eslbp1*). Genome walking must start from a known sequence, which in this case is cDNA from a transcriptomic library, and transcriptional start sites in *E. scolopes* are not well understood; the +1 position in the sequence

A.

AAAATAAAATAAATATTTTTTAAATTCAGGGTTTACAAGTTATAAATATAGTAAATTA**AGGGAATTGCCAT**AAGATCA
 AAGTATAGTATCTGTCCTGGTTGGTCAGTGGATAAGTTCTTCTAGATTGAAACAATTTTTATTTTTGTTCAAATCCA
 TGCCCTTTTTGGAGAGCTGAACCCTCCCCCCCCACATGAAAACTGGTCTTTTGTTTAACTTTGCCCTAGAACCTG
 TTGTCTTTGAAGATTAAGTAGAAATTAAAAAACAACAAAACTTTTAGATGTCAAACACTAACACATATGATTGCACG
 TTAATGTTGATGCTGGTGAAAGTAGTAATCTTTTTTTTTTTAGTTAAAAGTAAAACCTTTTCATAATTC AATTTTTATAT
 TTAATTTTTATTTAACACTTTAATGTGCAAACCTAAATTAATAAACATTCCTAGAAATTTTTTTAAAAGTTATCAAAC
 CATTTTTTTTTATTGTTTGCATACGTAATAATACTTTTTTTTTTTGCTCCAAATGCGATGAAACAATGTTTGTCTCCAA
 ATGTGACAAAATATTTCTTCAGACGCATAATTTGCTTCAAATCCATTA AAAACTTTTTGTTCCAGATGCAACAAAAC
 TTTTCGCTCCAGCTGCAATAGTTTTAAATAATTGAAAAATTTAGGAACGTATTTACATCCAAAATTTGTAGAAATATCG
 ATTAGCCAAATTTGAAAATATGGATCAATTAATAATTTACTAACGCTACCATATTTCAAATATGGGATACGAGAATGA
 AATTTTCGAGGTCGTAGAACAACCTCAAAAACGAAATATCGAATTA AAAAAAAAAAAAAATTAATCAGATAGTTATTTAAA
 AAAAAAAAAAGCAGTCTTTTTTTAGCTTTTTCAAAAATTTACAAAAAAAAAAACACTCATTACGAATATTTATTGATCGTT
 ATGAAACATCACATATAAATATAGATTTTTACTAAAATAAAAAACATATTTATCATAAATTTTCAGGTCAATAAGTAA
 AATGATTTTCGAA**AAAAATTCCCC**ACATTA AAAAGCGTGATAAGAAAATGAAATAGTTTTTTAGTGCCTGGAAAAAGTT
GGGTAATTTTTATTTGAGGCTGGAAATTTGCGACCACGCATGTGCGCTTATCGTCCGTTAG**AGGGAAAGTGGG**GAGG
TAATGACTGAATAGAGAGTGAAGAGAGTTGGAGAGAGTAAAGCAGCGTGAGCGAAGGAGGGACGAAC**TGGGTATTT**C
CAGCTTTTTATGTCGTTTTATTTTTCTTTAAAAAGAGAACATTTTTAGTCCCAAGTCATCATTCTGTATTCTCACTCG
TTGCTTTAAAGAAGCAACCCGACAAGATATTACGGGTGCGACATATAAAAAGGACAATCAAATATTTACCAAGAGA
TTTCAACTTTTACGGATTTATACTTACACGGGAGATTTTTTAAACGCGAAATCGACATTTTTCTTCTTCTTCTATT
TACCTATTTTTCTATTCACTTTAATTTTTCTTAATTTATTTTTTCACTACGGACGAATCTCTTTTTATTTGCAGCAT
TTAAATTTTTTTTTCTTAATAAAACAAAACAAAAAACAACAGGTTGACCTCTGATCTGTTTTTATAATTC AATTAG
GACATCGTTATTACCTTAGATCAACTATG

B.

GTCAAAAGTGGCAGCTCATA CATTAAATCAATTAATAAATAACATACTGACCCAACGGGTACCGGGTAGACTAGT
 GCATACATACACTCACAGATCATGGTATGCATATCCGTATTACTTTTTATATCACCTCCAAAAAGTGAAAGTAAATTA
 AGGCATTTATTAGGGGTATTTGTTTCACTCTGGCGTGTTTTTATAATCTAAATTTAGCAACGC**TGGGAAAAACCTGA**
 AAAAAGTGAAGTAAATTA AAACATTTGGTAGTAGTGGTTTGTTCCTCTGGCGTGTTTTTTTTATAATCTAAATTT
 AGCAACGC**TGGGAAAAACCTG**AAAAAAGTGAAGTAAATTA AAACATTTGGTAGTAGTGGTTTGTTCCTCTGGCG
 TGTTTTTTTTATAATCTAAATTTAGCAACGC**TGGGAAAAACCTG**AAAAAAGTGAAGTAAATTAAGACATTTGTTAGG
AGTGGTTTGTATCACTCTGGCATGTTTTTATAATCTAAATTTGCAACGC**TGGGAAAACCCCG**AAAAACGTCTATTA
TATACAAAAGA AACTCAGTGTAACATCAATTTATTGAGTAAAACCCAAGCAATACAGTTGGCAGTTTGTGTCTTTTAA
TGTATTAATACCAAGGAAAATATTTTTAAAAAATATCGGCTGCGAATATG

FIG B-1 The *esikB* and *eslbp1* promoters contain putative NF- κ B binding sites.

Contigs for *esikB* (A) and *eslbp1* (B) were formed from PCR sequences of this study and the published transcript (*esikB*) or RACE product from this study (*eslbp1*). Regions for which there is evidence of transcription are underlined and the final codon shown is the start codon. Putative NF- κ B binding sites (see also Table B-2) are bolded.

was assumed to be the most upstream of the published sequence or the sequence obtained from either RACE-PCR or any transcriptomic database ((4, 6) or Kremer, N. *et al.* unpublished or Moriano-Gutierrez, S. *et al.* unpublished). For *esikB*, 886 bp of sequence were obtained for which I am not aware of any evidence of transcription; for *eslbp1*, 399 bp of such sequence were obtained. These apparently untranscribed sequences were assumed to be the promoter regions and bases therein were assigned negative numbers for their positions. Hits for NF- κ B-family transcription factors identified by the PROMO program are given in Table B-2. Both the *esikB* and *eslbp1* promoters contain potential NF- κ B binding sites identified by this algorithm, both before and after the apparent transcriptional start site. None fit the form of the “canonical,” originally described NF- κ B consensus binding sequence GGRNNYYCC (13), and the sequences and positions of putative NF- κ B binding sites in these promoter regions are not identical. One putative NF- κ B binding site in the *eslbp1* promoter repeats identically three times; it is part of a larger, ~99 bp sequence that repeats almost identically.

RACE results. A total of 97 bp upstream of the *eslbp1* start codon were obtained by 5' RACE-PCR from within the *eslbp1* ORF. The sequence obtained by RACE overlapped exactly with that in transcriptomic databases, but was shorter. The RACEd sequence included the first primer used in the *eslbp1* genome walking, LBP1WALK3, extending 7 bp further, but did not include LBP1WALK4, the nested primer used in the *eslbp1* genome walking.

Table B-2: Putative NF- κ B binding sites in *esikB* and *eslbp1* promoters

Promoter of Gene:	Position	Sequence	TRANSFAC match
EsIkB	-829	AGGGAATTGCCA T	NF-kappaB [T00588/T00590], RelA [T00594], v-Rel [T00897]
EsIkB	+127	AAAAATTCCCC	NF-kappaB [T00588/T00590], RelA [T00594]
EsIkB	+255	AGGGAAAGTG	RelA [T00594]
EsIkB	+337	TGGGTATTTCC	NF-kappaB [T00590]
EsLBP1	-182	TGGGAAAAACCT G	NF-kappaB [T00588/T00590], RelA [T00594]
EsLBP1	-83	TGGGAAAAACCT G	NF-kappaB [T00588/T00590], RelA [T00594]
EsLBP1	+17	TGGGAAAAACCT G	NF-kappaB [T00588/T00590], RelA [T00594]
EsLBP1	+113	CTGGGAAAACCC CG	NF-kappaB [T00588/T00590], RelA [T00594], v-Rel [T00897]

DISCUSSION

The characteristics of cephalopod transcription start sites are not well established, though some limited investigation has begun (14). The reads in transcriptomic databases do not necessarily cover the entire transcript from which they are derived, and entries of different databases do not necessarily overlap entirely. While it may seem reasonable to assume that RACE-PCR reveals the entirety of the studied transcript, anecdotal data such as the results of this 5-RACE for *eslbp1* indicate that the transcript ligated to the GeneRacer oligo may be shorter than the maximal transcription product. In short, I am not aware of any reliable method to determine a transcriptional start site in *E. scolopes*, and therefore the +1 sites indicated in Table B-2 are simply the most upstream points for which there is evidence of transcription. In the cases of both *esikB* and *eslbp1*, however, it is apparent that at least some of the putative NF- κ B sites are downstream of the transcription start site, in the 5'-UTR, as these sites are seen in transcriptomic databases or RACE products. Others are upstream of any detected transcription products and are apparently in the promoter. There is precedent for NF- κ B responsive elements after the promoter; NF- κ B binding sites downstream of the start site, or even intronically located, have been shown to be important for regulation in mammalian genes and are suspected to be involved in control of human LBP expression (15-17).

The distribution of NF- κ B sites upstream of *esikB* and *eslbp1* are not obviously similar to those of human *lbp*. It is possible that the role of *eslbp1* in the symbiosis between *E. scolopes* and *V. fischeri* has influenced its regulation. Also notable are the differences in NF- κ B sites upstream of *esikB* and *eslbp1*, despite their qualitatively similar responses to MAMP and bacterial exposure (Chapter 2, Appendix A). Examining other transcription factor sites may yield further clues to the mechanisms of regulation of these two genes.

The detection of these candidate NF- κ B binding sites is only the first step in determining the activity of NF- κ B in *E. scolopes* promoters. The NF- κ B sites are a few among a great many potential binding sites detected by PROMO; there are hits for 322 unique TRANSFAC entries for the *eslbp1* region tested and 483 for *esikB*; in each case, four of these were noted to be some variation on NF- κ B (annotated as NF-kappaB [two entries], RelA, and v-Rel). Many of these are subtle variations on each other; for example, most of the hits for NF- κ B included in Table B-2 were overlapping hits for three different TRANSFAC entries, T00588 and T00590 (NF-kappaB), and T00594 (RelA). Nevertheless, the large number of different transcription factors implicated to potentially bind in these short DNA sequences renders apparent the importance of biochemical verification of these sites. NF- κ B binding sites are known to vary widely, in some cases, from the canonical GGRRNNYYCC (sometimes seen more specifically as GGGRNNYYCC) (13). Of note, the sequence GGGGAATTGCC, present at -828 upstream of *esikB*, binds both Dif and Rel in *Drosophila* (18).

The experiments that would enable investigation of interaction between the NF- κ B proteins and its putative DNA binding sites are simple in principle; an electrophoretic mobility shift assay would allow detection of such complexes, and variations on this experiment are made possible by antibodies to transcription factors. *esrel* has already been cloned into a bacterial expression vector and expressed with a His-tag (Appendix A), potentially enabling two different methods of protein detection by antibody (gene-specific or anti-His). The tools to investigate the biochemistry of putative NF- κ B binding sites in *E. scolopes* are thus at hand, for the motivated researcher.

Additionally, these experiments demonstrate the viability of genomic walking techniques in *E. scolopes*. These methods are especially important given the lack of a published genome in

E. scolopes, but may prove useful for investigating variation between individuals in promoter organization even after a generic genome is available.

REFERENCES

1. **Simakov O, Marletaz F, Cho SJ, Edsinger-Gonzales E, Havlak P, Hellsten U, Kuo DH, Larsson T, Lv J, Arendt D, Savage R, Osoegawa K, de Jong P, Grimwood J, Chapman JA, Shapiro H, Aerts A, Otilar RP, Terry AY, Boore JL, Grigoriev IV, Lindberg DR, Seaver EC, Weisblat DA, Putnam NH, Rokhsar DS.** 2013. Insights into bilaterian evolution from three spiralian genomes. *Nature* **493**:526-531.
2. **Zhang G, Fang X, Guo X, Li L, Luo R, Xu F, Yang P, Zhang L, Wang X, Qi H, Xiong Z, Que H, Xie Y, Holland PW, Paps J, Zhu Y, Wu F, Chen Y, Wang J, Peng C, Meng J, Yang L, Liu J, Wen B, Zhang N, Huang Z, Zhu Q, Feng Y, Mount A, Hedgecock D, Xu Z, Liu Y, Domazet-Lošo T, Du Y, Sun X, Zhang S, Liu B, Cheng P, Jiang X, Li J, Fan D, Wang W, Fu W, Wang T, Wang B, Zhang J, Peng Z, Li Y, Li N, Chen M, et al.** 2012. The oyster genome reveals stress adaptation and complexity of shell formation. *Nature* **490**:49-54.
3. **Chun CK, Scheetz TE, Bonaldo MeF, Brown B, Clemens A, Crookes-Goodson WJ, Crouch K, DeMartini T, Eyestone M, Goodson MS, Janssens B, Kimbell JL, Koropatnick TA, Kucaba T, Smith C, Stewart JJ, Tong D, Troll JV, Webster S, Winhall-Rice J, Yap C, Casavant TL, McFall-Ngai MJ, Soares MB.** 2006. An annotated cDNA library of juvenile *Euprymna scolopes* with and without colonization by the symbiont *Vibrio fischeri*. *BMC Genomics* **7**:154.
4. **Kremer N, Philipp EE, Carpentier MC, Brennan CA, Kraemer L, Altura MA, Augustin R, Häsler R, Heath-Heckman EA, Peyer SM, Schwartzman J, Rader BA, Ruby EG, Rosenstiel P, McFall-Ngai MJ.** 2013. Initial symbiont contact orchestrates host-organ-wide transcriptional changes that prime tissue colonization. *Cell Host Microbe* **14**:183-194.
5. **Rosenthal JJ, Bezanilla F.** 2002. Extensive editing of mRNAs for the squid delayed rectifier K⁺ channel regulates subunit tetramerization. *Neuron* **34**:743-757.
6. **Goodson MS, Kojadinovic M, Troll JV, Scheetz TE, Casavant TL, Soares MB, McFall-Ngai MJ.** 2005. Identifying components of the NF-kappaB pathway in the beneficial *Euprymna scolopes-Vibrio fischeri* light organ symbiosis. *Appl Environ Microbiol* **71**:6934-6946.
7. **Leoni C, Volpicella M, De Leo F, Gallerani R, Ceci LR.** 2011. Genome walking in eukaryotes. *FEBS J* **278**:3953-3977.
8. **Sun SC, Ganchi PA, Ballard DW, Greene WC.** 1993. NF-kappa B controls expression of inhibitor I kappa B alpha: evidence for an inducible autoregulatory pathway. *Science* **259**:1912-1915.

9. **Schumann RR.** 1995. Mechanisms of transcriptional activation of lipopolysaccharide binding protein (LBP). *Prog Clin Biol Res* **392**:297-304.
10. **Lee PN, McFall-Ngai MJ, Callaerts P, de Couet HG.** 2009. Preparation of genomic DNA from Hawaiian bobtail squid (*Euprymna scolopes*) tissue by cesium chloride gradient centrifugation. *Cold Spring Harb Protoc* **2009**:pdb.prot5319.
11. **Farré D, Roset R, Huerta M, Adsuara JE, Roselló L, Albà MM, Messeguer X.** 2003. Identification of patterns in biological sequences at the ALGGEN server: PROMO and MALGEN. *Nucleic Acids Res* **31**:3651-3653.
12. **Messeguer X, Escudero R, Farré D, Núñez O, Martínez J, Albà MM.** 2002. PROMO: detection of known transcription regulatory elements using species-tailored searches. *Bioinformatics* **18**:333-334.
13. **Wong D, Teixeira A, Oikonomopoulos S, Humburg P, Lone IN, Saliba D, Siggers T, Bulyk M, Angelov D, Dimitrov S, Udalova IA, Ragoussis J.** 2011. Extensive characterization of NF- κ B binding uncovers non-canonical motifs and advances the interpretation of genetic functional traits. *Genome Biol* **12**:R70.
14. **Vizoso M, Vierna J, González-Tizón AM, Martínez-Lage A.** 2011. The 5S rDNA gene family in mollusks: characterization of transcriptional regulatory regions, prediction of secondary structures, and long-term evolution, with special attention to Mytilidae mussels. *J Hered* **102**:433-447.
15. **Pierce JW, Lenardo M, Baltimore D.** 1988. Oligonucleotide that binds nuclear factor NF-kappa B acts as a lymphoid-specific and inducible enhancer element. *Proc Natl Acad Sci U S A* **85**:1482-1486.
16. **Xu Y, Kiningham KK, Devalaraja MN, Yeh CC, Majima H, Kasarskis EJ, St Clair DK.** 1999. An intronic NF-kappaB element is essential for induction of the human manganese superoxide dismutase gene by tumor necrosis factor-alpha and interleukin-1beta. *DNA Cell Biol* **18**:709-722.
17. **Schumann RR, Kirschning CJ, Unbehauen A, Aberle HP, Knope HP, Lamping N, Ulevitch RJ, Herrmann F.** 1996. The lipopolysaccharide-binding protein is a secretory class 1 acute-phase protein whose gene is transcriptionally activated by APRF/STAT/3 and other cytokine-inducible nuclear proteins. *Mol Cell Biol* **16**:3490-3503.
18. **Busse MS, Arnold CP, Towb P, Katrivesis J, Wasserman SA.** 2007. A kappaB sequence code for pathway-specific innate immune responses. *EMBO J* **26**:3826-3835.

APPENDIX C***Euprymna scolopes* Acyloxyacyl Hydrolase**

PREFACE:

Design of these experiments was by Benjamin Krasity and Jedediah Seltzer. RACE-PCR and confirmatory cloning were performed by JS. qRT-PCR was by BK, using primers designed by JS. This appendix was written by BK, with input from JS concerning Methods and Results.

INTRODUCTION

Acyloxyacyl hydrolase (AOAH) is an enzyme that, as studied in mammals, removes secondary acyl chains from the lipid A of lipopolysaccharide, creating a tetraacylated LPS molecule (1, 2) that is markedly less inflammatorily potent than hexaacylated LPS (3, 4). AOAH was first noted in neutrophils (2), but the primary site of LPS deacylation by AOAH appears to be the liver, where the enzyme is present in Kupffer cells (the macrophages of the liver) (5).

A candidate *aoah* transcript, apparently incomplete at the 5' end, was identified in an RNA-seq database of *Euprymna scolopes* transcripts (6). The fully octaacylated *V. fischeri* lipid A has four secondary acyl chains linked to the primary chains by ester groups (7); these are possible targets of AOAH activity. LPS is known to drive developmental changes in the host over light organ colonization (8-10), and in at least one case, a *V. fischeri* acyl group appears important in establishment of mutualism (11). The host is known to alter the phosphorylation of symbiont LPS (12), and the presence of AOAH suggests the possibility of host alteration of the acyl chains of LPS, as well.

MATERIALS AND METHODS

5' Rapid amplification of cDNA ends (RACE). ~25 light organs were dissected from juvenile animals and stabilized in RNAlater (Thermo Fisher, Waltham, MA). RNA was extracted with an RNeasy kit (Qiagen, Venlo, Netherlands) and DNase-treated with Turbo DNase (Thermo Fisher). RNA was prepared for 5'-RACE with a GeneRacer kit (Thermo Fisher) and reverse-transcribed with MMLV reverse transcriptase (Clontech, Mountain View, CA) using 5' CDS primers (5'-(T)₂₅VN-3'). RACE-PCR was performed with Platinum Taq High Fidelity enzyme (Thermo Fisher) according to the GeneRacer kit recommendations, with the gene-

specific primer AOA_H_5Rev3: GGACGTCCTGAAGGGTCAGATCCAT and, subsequently, the nested gene-specific primer AOA_H_5R2_NESTED: CCTCTCGGGTTACTATTCTTGCAGAA. PCR products were cloned into pCR4-TOPO and transformed into TOP10 competent cells (Thermo Fisher); plasmids were isolated with a QiaPrep MiniPrep kit (Qiagen, Venlo, Netherlands) and the inserts sequenced. Upon successful RACE-PCR, the entire *aoah* open reading frame (ORF) was cloned with the primers AOA_H_C_FWD1: GCAAACCGTCAACTTCGACGA and AOA_H_C_REV1: GCAGCGAGGATTTTGTGTGTT, then sequenced as above. The peptide sequence inferred from the cloned *E. scolopes aoah* was evaluated for various biochemical parameters with ExPASy ProtParam (13) and SignalP 4.1 (14) and aligned with human AOA_H (NCBI accession number: NP_001628.1) by ClustaW2 (15, 16).

Quantitative reverse-transcriptase PCR (qRT-PCR). qRT-PCR was used to compare levels of *aoah* transcript in the light organs of 24 h post-inoculation symbiotic and aposymbiotic animals. qRT-PCR was performed as described previously, with the 40s ribosomal subunit and Serine HMT transcript used as control genes (Chapter 2), using the AOA_H gene-specific primers AOA_H_qFWD1: GAGGTTTTGCCGTAATTGGA, and AOA_H_qREV1: AGTTGTGGCCAATCCATCTC. Transcript levels were log-transformed for normality and compared by t-test.

RESULTS AND DISCUSSION

Amino acid sequence is well-conserved between human and *E. scolopes aoah*. The *E. scolopes aoah* transcript was reverse-transcribed and sequenced (Fig. C-1A) and the predicted amino acid sequence of the protein product determined (Fig. C-1B). The predicted product was

A GCAAACCGTCAACTTCGACGAAAGAAAATAACGCCTCACCCAGTTGAACCAGCGCTACTGATTATAACAAATTAGAGAAACGCGAA
CACAAAACAAAACAAGACAACAACAATAAAAAAAAAAAAAAAAAATTACTATACGATTTTAAAAATTAATCTACGGATGGCTTTAAACG
CTAAAATGAGACTTGTGCTTCTTTCTTCGTGGTGACGTGTTATTTTACCGAAATCCCAAGCAATGTTTAAACAAAGGGACCAAT
GGAGGTTTTAAATGCATTGGTTGCACAGTTGTTTTGCGATAATGGATCAGTTATCGGTCGTCACAATGAATCGATGAGCCGCGC
AGCATATCGTATGTGTACTTTCTTACCTGACAAAAGGTTAACCGAATTGTGCAATGTCGTCGTTAGTGTTCTGCCACTTGGTTGG
ATCAAATACATTTCGATGGGCGCAACTCCTGACACAATATGTCAAACGCTCAATATATGCACAAGCGAGAAAGGAAAATATTGCCGA
ATTTTCACAAAGAACA AAAAGAACGATTTCTTCAGACAACAATCAAAAATAGACAATATTTAAAAATAGTCAAATAACCTGGC
CAATATACTCGGCAACTTGAATGAAATGGGTCCCAACATGGGCTTTCTTTGTAACTTCCGTTGATTAAGAAAATATGTTTTGCTA
TGAAACATTATTTCTCACATATGCCCGTATATGACAAAGATGGTGACTTGTTTAGTCAAAAATATCCACTTCGAGGTGCGTACTGG
CGAGGCAAGATTCGATGATTCGATGCTAAAATCCACCCTGGTGCCAAGACAATCAACTCTGATATCCATACCGATTCTAATTG
CAATGGAAATTTTCGGATCTGACCCTTCAGGACGTCCTTATGAAGATATTTCTGCAAGAATAGTAACCCGAGAGGTTTTGCCGTAA
TTGGAGATTCATAGCCGCTCATTACATATTCGGAAGAATGGCTCGATGCAACAAGATTACAAGGAATCGTTCAAAAATGTC
CTCTTTCTCATGAAAATGAAATGGATTGGCCACAACCTTTCATCACAACAGGTTTCATGAATAGCACTTGAATGTCGTCAAAGG
TTCAACAGATTCGATTTATTTATCGTCTTTGGAAGCGGAATCGTTGTATCCACAGAGACTTCCAGAATATCGGCATCAATGGTGCCA
GAAATTTGAAAGCAACAAAATATGTTAAAAGTTTATCCCGAATCAAAACAACGGATTATCCCTTGATTGTGAGTTATTCCTTAGTC
GGAAATGACGCTGCAACTGGCACATTGACAGCCTAAAGTATATGACCACCCCAAAGGAAATGTTTCATAACGCGATGGAACCTTT
GAAACTTGGATAACATTTACCAAAGGAAGTTTCGTTCTGATTAATGGTCTTTGTTGACGGTCGGATACTATATAATACGATGC
ACGAAAGAATACACCCGATTGGAAGGACCCGAAACGATGTCACTTACAACGACTTTTATGAGTATATGAACTGCATGGAGCTCAGT
ATGTGTAGCGGTTGGCTCACTTCGAATGAGACTCTTCGGAATCTTACCCAACAACGTCGGACCAACTGAGTAACGCTCTGGAAA
CATTATGAAAACCAACAACCTTCAAAAATTTCCGCCTCCATTATATCGGGAACTTTTTCACAGACTTGTAAATGCTTGGGAAA
AGAAAGCGCGCAGTGCCTGGCAAATTATTGAACCTGTCGATGGATTCACAACAACAGTTGAGTGAACCTCTTGTAAACGGATATA
GTCTGGAGATTTTAGAGAAGGAGGCACCGGAAATCTTGGACAGATCAACCCAAACAACGAGAAAATAAAAAAGATTTTTGGTAA
CCAAGGAGGATATTAATTACAACGACGCAACACAAAAATCCTCGCTGC

B MALNAKMRLVASFFVVTCLFLPNSQAMFNKGTNGGFKICIGCTVVFAIMDQLSVVHNESMSRAAY
 RMCTFLPKRLTEL CNVVVSGSATWLDQIHSMGATPDTICQTLNICTSEKGYCRI FTKNKND
 FFRQQIKIRQYFKNSQNNLANILGNLNEMGPNMGFLCKLPLIKKICFAMKHYF SHMPVYDKDGD
 LFSQKYPLRGAYWRGKDCDDSDAKIHPGAKT INSDIHTDSNCGIFGSDPSGRPYEDI FCKNSN
 PRGFAVIGDSIAAHLHIPEEWLDATKINKESFKNVLF LIENEMDWPQLSFTTGFMNSTWNVVKG
 STDSIYYRLWKRNR CIHRDFQNINGARS LKATKYVKSLSRNQTTDY PLIVSYSLVGN DVCNW
 HIDSLKYMTPKEMFHNAMETLKYLDTILPKGSFVLINGLVDGRILYNTMHERIHP IGRTRNDV
 TYNDFY EYMNCMELSMCSGWLTSNETLRNL TQQRADQLSNVLENI MKNHNNFKNFRLHY IGLKF
 HRLVNAWEKKGGS AWQIIEPVDGFHNNQLSE PLVTDIVWKILEKEAPE ILGQINPNNEKIKKIF
 GNQGGY

FIG C-1 Sequence of *E. scolopes* AOA H. (A) Reverse-transcribed and PCR-amplified sequence of *aoah* transcript in *E. scolopes*. Bold and underlined: ORF, including stop codon. (B) Derived amino acid sequence of *E. scolopes* AOA H protein.

582 amino acids long, with a molecular weight of 66.7 kDa and isoelectric point of 8.89. This protein sequence was aligned to that of human AOA (NP_001628.1) (Fig. C-2) and appears to preserve many of the important features of human AOA. The *E. scolopes* AOA was 43% identical overall to its human orthologue, and the catalytic triad of serine, aspartate and histidine (17, 18) was conserved entirely, though the environment of the catalytic serine differed somewhat, with the arrangement GXSSA replacing the typical GXSSG (17). Human AOA propeptide is cleaved into two subunits that are held together by one or more disulfide bonds (19); *E. scolopes* retains every cysteine of human AOA except that of the signal peptide (16 of 17). Both proteins contain a signal peptide. However, the subunit cleavage site at Ser156 of the human protein is altered in *E. scolopes*, raising questions of whether and how AOA subunits may be separated in *E. scolopes*.

Acylxyacyl hydrolase is not regulated in the light organ by symbiosis at the 24 h timepoint. While distinguishable from background in each case, transcript levels of *aoah* in apo- and symbiotic light organs 24 h post-inoculation were indistinguishable from each other (Fig. C-3). Given the expression of *aoah* in macrophages in mammals, hemocytes are obvious candidates for *aoah* expression in *E. scolopes*. It is possible that differences in transcript level exist in light organ hemocytes, but not within other tissues of the light organ, drowning out signal. This hypothesis may be tested by use of newly developed *in situ* hybridization methods (20). Transcripts strongly localized to *E. scolopes* hemocytes have been detected by *in situ* hybridization methods previously (21). A polyclonal antibody to human AOA large subunit (Santa Cruz Biotechnology, Dallas TX, catalogue number sc-135109) was tested against extracted adult *E. scolopes* hemocytes, but with negative results (not shown).

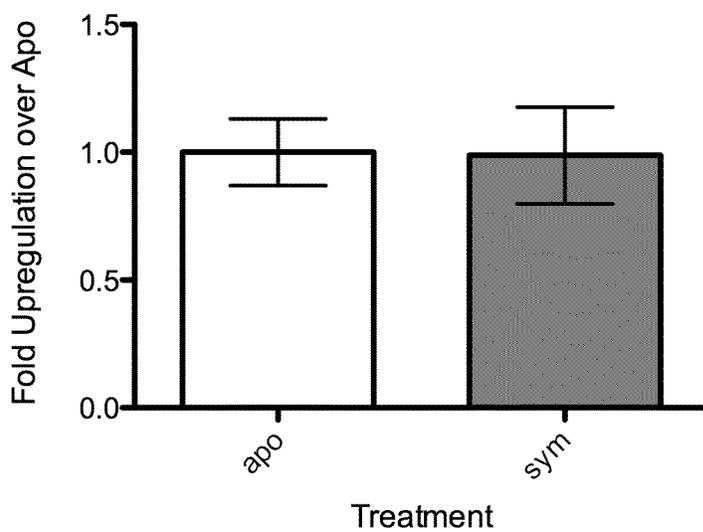


FIG C-3 Transcript levels of *aoah* in 24 h apo- and symbiotic light organs. Transcript levels were normalized to 40S and Serine HMT levels and the average aposymbiotic transcript level defined as 1. Error bars, SEM of four biological replicates. The difference between apo- and symbiotic treatments was not statistically significant.

REFERENCES

1. **Munford RS, Hall CL.** 1986. Detoxification of bacterial lipopolysaccharides (endotoxins) by a human neutrophil enzyme. *Science* **234**:203-205.
2. **Hall CL, Munford RS.** 1983. Enzymatic deacylation of the lipid A moiety of *Salmonella typhimurium* lipopolysaccharides by human neutrophils. *Proc Natl Acad Sci U S A* **80**:6671-6675.
3. **Kitchens RL, Munford RS.** 1995. Enzymatically deacylated lipopolysaccharide (LPS) can antagonize LPS at multiple sites in the LPS recognition pathway. *J Biol Chem* **270**:9904-9910.
4. **Kitchens RL, Ulevitch RJ, Munford RS.** 1992. Lipopolysaccharide (LPS) partial structures inhibit responses to LPS in a human macrophage cell line without inhibiting LPS uptake by a CD14-mediated pathway. *J Exp Med* **176**:485-494.
5. **Shao B, Munford RS, Kitchens R, Varley AW.** 2012. Hepatic uptake and deacylation of the LPS in bloodborne LPS-lipoprotein complexes. *Innate Immun* **18**:825-833.
6. **Kremer N, Philipp EE, Carpentier MC, Brennan CA, Kraemer L, Altura MA, Augustin R, Häslner R, Heath-Heckman EA, Peyer SM, Schwartzman J, Rader BA, Ruby EG, Rosenstiel P, McFall-Ngai MJ.** 2013. Initial symbiont contact orchestrates host-organ-wide transcriptional changes that prime tissue colonization. *Cell Host Microbe* **14**:183-194.
7. **Phillips NJ, Adin DM, Stabb EV, McFall-Ngai MJ, Apicella MA, Gibson BW.** 2011. The lipid A from *Vibrio fischeri* lipopolysaccharide: a unique structure bearing a phosphoglycerol moiety. *J Biol Chem* **286**:21203-21219.
8. **Foster JS, Apicella MA, McFall-Ngai MJ.** 2000. *Vibrio fischeri* lipopolysaccharide induces developmental apoptosis, but not complete morphogenesis, of the *Euprymna scolopes* symbiotic light organ. *Dev Biol* **226**:242-254.
9. **Altura MA, Stabb E, Goldman W, Apicella M, McFall-Ngai MJ.** 2011. Attenuation of host NO production by MAMPs potentiates development of the host in the squid-vibrio symbiosis. *Cell Microbiol* **13**:527-537.
10. **Heath-Heckman E.** 2014. Day/Night Cycles in the *Euprymna scolopes* – *Vibrio fischeri* Symbiosis. Ph.D. University of Wisconsin, Madison, WI.
11. **Adin DM, Phillips NJ, Gibson BW, Apicella MA, Ruby EG, McFall-Ngai MJ, Hall DB, Stabb EV.** 2008. Characterization of htrB and msbB mutants of the light organ symbiont *Vibrio fischeri*. *Appl Environ Microbiol* **74**:633-644.
12. **Rader BA, Kremer N, Apicella MA, Goldman WE, McFall-Ngai MJ.** 2012. Modulation of symbiont lipid A signaling by host alkaline phosphatases in the squid-vibrio symbiosis. *MBio* **3**.

13. **Gasteiger E, Gattiker A, Hoogland C, Ivanyi I, Appel RD, Bairoch A.** 2003. ExPASy: The proteomics server for in-depth protein knowledge and analysis. *Nucleic Acids Res* **31**:3784-3788.
14. **Petersen TN, Brunak S, von Heijne G, Nielsen H.** 2011. SignalP 4.0: discriminating signal peptides from transmembrane regions. *Nat Methods* **8**:785-786.
15. **Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA, McWilliam H, Valentin F, Wallace IM, Wilm A, Lopez R, Thompson JD, Gibson TJ, Higgins DG.** 2007. Clustal W and Clustal X version 2.0. *Bioinformatics* **23**:2947-2948.
16. **Goujon M, McWilliam H, Li W, Valentin F, Squizzato S, Paern J, Lopez R.** 2010. A new bioinformatics analysis tools framework at EMBL-EBI. *Nucleic Acids Res* **38**:W695-699.
17. **Staab JF, Ginkel DL, Rosenberg GB, Munford RS.** 1994. A saposin-like domain influences the intracellular localization, stability, and catalytic activity of human acyloxyacyl hydrolase. *J Biol Chem* **269**:23736-23742.
18. **Wei Y, Schottel JL, Derewenda U, Swenson L, Patkar S, Derewenda ZS.** 1995. A novel variant of the catalytic triad in the *Streptomyces scabies* esterase. *Nat Struct Biol* **2**:218-223.
19. **Munford RS, Hall CL.** 1989. Purification of acyloxyacyl hydrolase, a leukocyte enzyme that removes secondary acyl chains from bacterial lipopolysaccharides. *J Biol Chem* **264**:15613-15619.
20. **Nikolakakis K, Lehnert EM, McFall-Ngai MJ, Ruby EG.** Using hybridization chain-reaction fluorescent in situ hybridization (HCR-FISH) to track gene expression by both partners during initiation of symbiosis. *Applied and Environmental Microbiology*:In Press.
21. **Kimbell JR, Koropatnick TA, McFall-Ngai MJ.** 2006. Evidence for the participation of the proteasome in symbiont-induced tissue morphogenesis. *Biol Bull* **211**:1-6.