THE LIGAND AND FUNCTION OF THE REGIII FAMILY OF

BACTERICIDAL C-TYPE LECTINS

APPROVED BY SUPERVISORY COMMITTEE

Lora Hooper, Ph.D.

Vanessa Sperandio, Ph.D.

Kim Orth, Ph.D.

DEDICATION

I dedicate this work to my fiancée, Chris Morris, my parents, and my brother and sister for their unswerving love, support, and encouragement. Also, in the words of J.S. Bach, *Soli Deo Gloria.*

THE LIGAND AND FUNCTION OF THE REGIII FAMILY OF BACTERICIDAL C-TYPE LECTINS

by

HEATHER LYNN CASH

THESIS

Presented to the Faculty of the Graduate School of Biomedical Sciences

The University of Texas Southwestern Medical Center at Dallas

In Partial Fulfillment of the Requirements

For the Degree of

MASTER OF SCIENCE

The University of Texas Southwestern Medical Center at Dallas

Dallas, Texas

August, 2006

THE LIGAND AND FUNCTION OF THE REGIII FAMILY OF BACTERICIDAL C-TYPE LECTINS

Publication No. _____

Heather Lynn Cash, M.S.

The University of Texas Southwestern Medical Center at Dallas, 2006

Supervising Professor: Lora V. Hooper, Ph.D.

Beginning at birth, the intestines of humans and other mammals are colonized with a diverse society of resident bacteria that play a crucial role in host nutrient metabolism. To maintain this commensal relationship, resident microbes must be prevented from crossing the intestinal epithelium into host tissues where they can cause inflammation and sepsis. The innate immune system plays a crucial role in preventing bacterial incursions across gut epithelial surfaces. Mucosal epithelial cells produce a variety of secreted antimicrobial proteins that help to prevent bacterial attachment and encroachment at epithelial surfaces. Among these, Paneth cells are specialized small intestinal epithelial cells that have been

shown to produce and secrete antimicrobial proteins and peptides. To gain new insights into the adaptation of mucosal surfaces to microbial challenges, the Hooper lab has used DNA microarrays to screen for Paneth cell genes whose expression is modulated by intestinal microbes. This screen revealed that expression of two C-type lectins, RegIIIβ and RegIIIγ, is strongly induced following intestinal colonization with resident microbes.

Two features suggested that members of the RegIII family may have microbicidal functions. First, they are C-type lectin family members. Other C-type lectins, including the mannose binding lectin, have well-characterized innate immune functions and play critical roles in microbial killing by recruiting complement. Second, I have shown that the murine RegIII lectins localize to intestinal crypt cells, including Paneth cell secretory granules, and that they bind to luminal bacteria harvested from intestinal conditions. Based on these observations, we hypothesized that this family of proteins may perform an innate immune function, specifically antimicrobial defense.

The studies reported in this thesis characterize a family of C-type lectins. Specifically, we determined that these proteins interact with peptidoglycan by binding with high affinity to its glycan structure, representing a unique blend of peptidoglycan recognition and lectin function. Additionally, we have demonstrated that this binding results in the specific disruption of the Gram positive bacterial cell wall, where peptidoglycan is exposed, which is the first example of a family of directly bactericidal C-type lectins. We also present evidence for the regulation of these bactericidal proteins by colonization with an intestinal microflora. Therefore, the research presented in this thesis elucidates the function of three members of the RegIII family, in both mice and humans.

ACKNOWLEDGEMENTS

Foremost, I would like to express my deep gratitude to my mentor, Dr. Lora Hooper, for her guidance, friendship, and for all that she has taught me about conducting exceptional science. To the members of the Hooper lab, you have been the best colleagues, full of patience, encouragement, laughter, and brilliant ideas. But, more importantly, you have been loyal, steadfast friends; and for that, I thank you. In particular, thank you, Cecilia, for sharing a bay with me and working so diligently and competently on the Reg project. Also, thank you to the members of my committee, Dr. Vanessa Sperandio and Dr. Kim Orth, for their advice, ideas for experiments, and for always being excellent advocates.

Additionally, I would like to acknowledge all those who have helped with protocols and equipment – Angie Mobley, for help with flow cytometry; Tom Janeszewski, for assistance with electron microscopy; and the Molecular Pathology Core, for beautiful *in situ* experiments.

I would also like to particularly thank my roommate, Rochelle Hinman, for her quick humor, endless patience, and willingness to discuss our individual projects. To Kentner Singleton, Ann Davis, and the members of the Farrar lab: thank you for all of the helpful insights and discussions, and for making work a joy. Thank you to other members of my class, including Karen Lewis, Lillian Kuo, Olivia Lee, and Kristen Evaul for their friendship and support during classes, qualifying exams, and laboratory life.

I would like to acknowledge my mother and father for their support and sacrifices made during my education, and for teaching me what is truly important in life. I also thank my brother and sister, Joshua and Jessica, for your love and friendship.

Finally, to my dear fiancée, Chris Morris, for your love, complete understanding, and unwavering support.

TABLE OF CONTENTS

LIST OF PUBLICATIONS

Cash HL, Whitham CV, Behrendt CL, and Hooper LV. RegIIIβ is a peptidoglycan binding lectin with bactericidal activity, *manuscript in preparation*.

Cash HL, Whitham CV, Behrendt CL, and Hooper LV. Symbiotic bacteria direct expression of an intestinal bactericidal lectin, *manuscript submitted*.

Cash HL, Whitham CV, and Hooper LV. (2006) Refolding, purification, and characterization of human and murine RegIII expressed in *Escherichia coli*. *Protein Expression and Purification*, in press.

Cash HL and LV Hooper. (2005) Commensal bacteria shape intestinal immune development. *ASM News*. **71**(2): 77-83.

LIST OF FIGURES

LIST OF TABLES AND APPENDICES

LIST OF DEFINITIONS

Arg	Arginine
BHI	Brain-heart infusion media
bp	Base pair
BSA	Bovine serum albumin
cAMP	Cyclic adenosine monophosphate
cDNA	Complimentary DNA
CD	Circular dichroism
CFU	Colony forming unit
CRD	Carbohydrate recognition domain
DAP	Diaminopimelic acid
DC	Detergent compatable
DNA	Deoxyribonucleic acid
dNTP	Dideoxy-nucleotide triphosphate
dUTP	Dideoxyuridine triphosphate
dsRNA	Double-stranded RNA
DTT	Dithiothreitol
EDTA	Ethylenediamine-tetraacetic acid
GALT	Gut-associated lymphoid tissue
GlcNAc	N-acetylglucosamine
GST	Glutathione S-transferase
IB	Inclusion body

IBD	Inflammatory bowel disease
IFN	Interferon
IgA	Immunoglobulin A
IL	Interleukin
Ile	Isoleucine
iPGN	Insoluble peptidoglycan
IPTG	isopropyl-β-D-thiogalactopyranoside
ISRE	Interferon-stimulated response element
HIP/PAP	Hepatocellular induced protein/Pancreatitis associated protein
kD	Kilodalton
LB	Luria-Bertani broth
LPS	Lipopolysaccharide
Lys	Lysine
MALDI-TOF	Matrix assisted laser desorption/ionization time of flight
MAMPs	Microbe-associated molecular patterns
MASP	MBL-associated serine protease
MBL	Mannose binding lectin
MES	4-Morpholineethanesulfonic acid
MDP	Muramyl dipeptide
mRNA	Messenger RNA
MurNAc	N-acetylmuramic acid
NF-κB	Nuclear factor kappa beta

NOD	Nucleotide-binding oligomerization domain
OD	Optical density
ORF	Open reading frame
PAGE	Polyacrylamide gel electrophoresis
PBS	Phosphate buffered saline
PCR	Polymerase chain reaction
PGN	Peptidoglycan
PGRP	Peptidoglycan recognition protein
qPCR	Quantitative real-time PCR
RAG	Recombinase activating gene
RNA	Ribonucleic acid
RT-PCR	Reverse-transcriptase PCR
SCID	Severe combined immunodeficiency disease
SDS	Sodium dodecyl sulfate
SDS-PAGE	Sodium dodecyl sulfate polyacrylamide gel electrophoresis
SP	Surfactant protein
sPGN	Soluble peptidoglycan
SPR	Surface plasmon resonance
STAT	Signal transducer and activator of transcription
TLRs	Toll-like receptors
TNF-α	Tumor necrosis factor alpha
WGA	Wheat germ agglutinin

Chapter One:

INTRODUCTION

A. Overview – The Dynamic Host-Microbial Interactions within the Intestine

A primary theme of life involves the multifaceted interactions that occur between eukaryotic and prokaryotic organisms, many of which result in mutually beneficial relationships. Beginning at birth, the mammalian intestine acquires an array of bacterial species that aid in digestion and development, while also benefiting from a protected, nutrient-rich niche. This symbiosis is so effective that various systems and regulatory mechanisms within the intestine fail to fully develop in the absence of a luminal microflora [1]. Particularly, studies have demonstrated that intestinal colonization with a full microflora or a single commensal organism induces the development of the intestinal microvasculature and upregulates the expression of myriad genes involved in metabolism, the immune response, and intestinal barrier integrity [1, 2]. Despite the many

benefits attributable to a microflora, the sheer number of foreign cells in close proximity to underlying host tissues (approaching 10^{11} per mL of colonic contents [3]) underscores the necessity for maintaining established epithelial boundaries. Thus, the defense mechanisms utilized by the intestinal mucosal barrier play a vital role in establishing a productive colonization of the intestine, unhampered by potential bacterial invasion of underlying host tissues.

The primary mechanism utilized by tissues of the gastrointestinal tract to promote beneficial colonization while minimizing bacterial incursions is native physiology. For example, the reported pH of the stomach is ~2-3 [4], which allows for colonization by only a few organisms, including *Helicobacter pylori* [5, 6]. Likewise, the peristaltic movement of the upper small intestinal smooth muscles prevents permanent colonization by most pathogenic and even commensal organisms [7]. Additionally, the architecture of the small intestine is uniquely suited to maximize nutrient adsorption while corralling resident microbes away from underlying host tissues. Protruding into the lumen of the intestine are finger-like projections known as villi, which increase the adsorptive surface area and are covered with a continuously replenishing layer of epithelial cells. The progenitor stem cell responsible for the epithelial regeneration that is necessary for maintaining the intestinal barrier is situated at the base of each villus in the “crypts” [8, 9]. These crypts harbor the stem cell flanked by specialized epithelial cells such as Paneth cells and goblet cells (Figure 1.1). It has been suggested that these effector cells function to promote a relatively sterile microenvironment in the crypts, preventing potential damage to the progenitor cell [10]. Thus, luminal bacteria are separated from underlying host tissues by a single epithelial cell layer, prompting the need for sentinel innate and

adaptive immune effector cells including Paneth cells and intraepithelial lymphocytes among others [11].

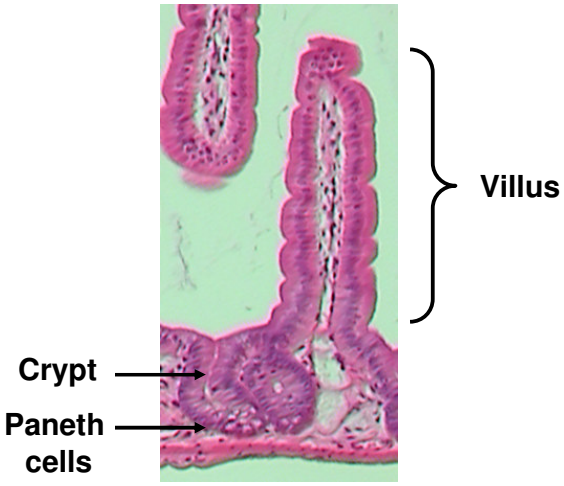

Figure 1.1: Basic small intestinal architecture: the crypt-villus unit.
Hematoxylin-eosin stain of cross-sections of paraffin-embedded mouse small intestine visualized using light microscopy at 40X magnification.

Investigations into the cross-talk between these host sentinel effector cells and luminal bacteria began making rapid strides with the advent of germ-free animal housing facilities [12]. Animals maintained in germ-free conditions are microbially sterile, having been born into a sterilized isolator, equipped with specialized air filtration and autoclaved contents. Using these animal models, researchers have begun to define the roles played by the commensal microflora and the signals by which the host and microflora communicate. Particularly, their studies indicate that the presence of a commensal microflora is necessary for both the appropriate development of blood vessels in the intestinal villi and in the acquisition of an intestinal adaptive immune system [1, 13]. Additionally, germ-free mice confirmed another crucial role for the microflora when studies determined that they require ~30% more nutrient intake than their colonized littermates, due to the lack of microorganisms to break down otherwise indigestible molecules [14].

The role of the intestinal microflora as "metabolic workhorses" is one that has been suggested by a variety of recent studies. Microarray profiling of changes in protein expression patterns of the commensal organism, *Bacteroides thetaiotaomicron*, indicates that this species upregulates many genes in response to a shift in environment from a liquid culture to the mammalian intestine [15]. As expected, many of these genes code for products involved in complex carbohydrate metabolism. While the human genome encodes one identified glycoside hydrolase, the *B. theta* genome contains 64 such enzymes to date, the transcription of many of which was significantly increased in the mammalian intestine. Further, this organism can selectively induce these enzymes to scavenge within its intestinal micro-environment [15]. Thus, the host diet and the enzymatic capabilities of one particular bacterial species intersect to provide required nutrients for both. As more members of the intestinal microflora are identified, transcriptional profiling may reveal similar symbiotic communications between host cells and these bacterial species.

The intestinal epithelium represents a unique interface between mammalian host tissues and the vast populations of luminal bacteria. As such, it must remain poised to employ an array of endogenous defense mechanisms to corral and to communicate with these beneficial microbes. Specifically, resident cells of the adaptive and innate immune defenses provide specialized and rapid responses, respectively, to bacterial encroachment, making both vital for maintaining intestinal homeostasis.

B. Mechanisms of Defense by the Adaptive Immune System

Taking sentinel positions in the interior of each villus are T- and B-cells, which are also present within large lymphoid structures known as Peyer's patches. Peyer's

patches form a portion of the gut-associated lymphoid tissue (GALT) and harbor both developing and differentiated adaptive immune cells. Research conducted by Knight and colleagues has demonstrated that intestinal bacteria are required for the development of GALT and for generating diversity in B cell antibodies [16]. Without colonizing bacteria and their associated antigens, intestinal lymphocytes are only exposed to a limited number of foreign antigens, hindering the development of antibody diversity. Therefore, the presence of the intestinal microflora appears to contribute not only to the formation of the tissues in which intestinal lymphocytes develop, but also may influence the ability of the immune system to mount a successful immune response [16].

In addition to the larger Peyer's patches, another form of GALT, designated "cryptopatches," punctuates the length of the small intestine. The immune cells contained within these small lymphoid tissues are relatively uncharacterized save for the expression of the lymphoid precursor, c-Kit, on their cell surfaces. Studies have suggested that intestinal CD8 intraepithelial lymphocytes may develop in cryptopatches [17-19]. As well as containing lymphoid cells, cryptopatches harbor dendritic cells (Figure 1.2), which may serve to present intestinal bacterial antigens to the developing lymphoid cells.

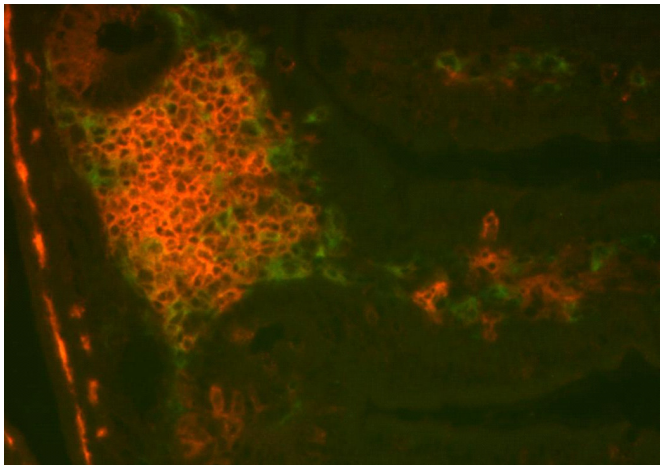

Figure 1.2: Cryptopatches harbor precursor lymphocytes and dendritic cells.

Prepared cryosections of conventionally colonized mouse small intestine were dual-stained with α-cKit (red) and α-IA/IE (green), markers of precursor lymphocytes and dendritic cells, respectively. The composite image was generated by overlaying images from the FITC and Cy3 fluorescent filters. This result is representative of two experiments**.**

Not only are the dendritic cells resident in the GALT engaged in direct interactions with other adaptive immune cells, but they also line the interior (lamina propria) of each small intestinal villus. Previous research has described dendritic cells as extending processes between epithelial tight junctions in order to sample bacterial particles within the lumen [20]. Based on the observation that no classic inflammatory response is generated in the healthy intestine, it has been hypothesized that this sampling could serve to "tolerize" circulating, sentinel T- and B-cells to the antigens found on resident microbes [21, 22]. Intriguingly, the intestines of germ-free mice display a marked reduction in the expression of the dendritic cell marker IA/IE (H. Cash, unpublished observations) within their villi, reflecting the lack of a luminal microflora to sample.

The combination of both B-cells and dendritic cells within the lamina propria could also serve to induce another B-cell dependent mechanism of intestinal barrier maintenance. Research conducted by Andrew MacPherson and colleagues has

demonstrated that dendritic cells loaded with commensal bacteria are crucial for inducing B-cell mediated immunoglobulin A (IgA) secretion [23]. This secretory IgA produced by intestinal B-cells and translocated across epithelial cells provides a non-specific method for agglutinating luminal microbes clustered near the epithelial surface, preventing them from penetrating into host tissues where they could cause inflammation and sepsis. Concomitantly, transcriptional profiling comparing the intestines of conventional mice with those of germ-free demonstrated that colonization increases the expression of the IgA transport protein, pIgR [2]. Thus, the commensal flora is responsible for induction of a molecule integrally involved in maintaining homeostasis.

Likewise, John Cebra and colleagues determined that commensal bacteria play crucial roles in driving B cell development in Peyer's patches, which are underdeveloped in germ-free mice [13]. Upon colonization of the intestine, commensal bacteria initiate a cascade of reactions including the increased production of IgA antibodies by B cells and the transient expansion of germinal center reactions between B and T cells in the Peyer's patches. As a result, germ-free mice generate reduced amounts of IgA as compared to mice that have an intestinal flora, and have decreased numbers of circulating B and T lymphocytes when tissues are compared microscopically [13]. Additionally, introduction of a single commensal bacterial species does not restore proper development, suggesting that a diverse repertoire of bacterial species and antigens are necessary to stimulate proper immune cell circulation [13]).

C. Mechanisms of Innate Immune Defense

Were the intestine to rely upon a targeted, adaptive immune response to develop for each encroachment of the epithelial border, the invading, pathogenic organism could have grown and disseminated to lethal levels throughout various tissues and organs before an effective response was mounted. Therefore, the gut employs effectors of the innate immune system that are rapid and diverse to swiftly eliminate encroaching pathogens. The speed and precision of the response can be attributed to the effector molecule diversity that is encoded in the germ-line of the host, eliminating the need for lengthy recombination processes [24]. This ability of the intestine to mount swift responses to bacteria or viruses that penetrate the mucosal barrier is mediated by the following systems, which serve to maintain the beneficial relationships between host and microbe in the gut.

i. Signaling pathways

The rapid-fire response of the innate immune system is characterized by its recognition of conserved patterns on foreign organisms to distinguish self from non-self [25]. These microbe-associated molecular patterns (MAMPs) range from cell wall components such as peptidoglycan (PGN) and lipopolysaccharide (LPS) to nucleic acids such as double-stranded RNA (dsRNA) and methylated DNA. Toll-like receptors (TLRs) are among the best-known molecular sensors of MAMPs, having been the focus of extensive investigation since their initial identification in *Drosophila* [26-28]. Currently, eleven TLRs have been identified in mammals, each binding a unique MAMP ligand. [24, 29, 30]. Interaction between each TLR and its ligand initiates a signaling cascade, which results in a significant induction of NF-κB-mediated genes, including IFN-γ along

with other signaling and effector molecules [24, 26]. Recent work has shown that this signaling mostly employs the adapter protein MyD88, which interacts with the intracellular domains of some TLRs [31]. However, TLRs may also signal using a MyD88-independent pathway, although the latter is not as well characterized [24, 32]. The importance of some of these receptors in mounting an initial response to pathogens has been underscored by elegant studies in which individual TLRs were ablated from the mouse genome, demonstrating that TLR2 and 4 knock-out mice are unable to generate an effective inflammatory response to deal with pathogenic threats [30, 33].

While all TLRs are membrane-bound proteins, some, such as TLRs 3 and 5, are found on intracellular vesicles, and others, such as TLRs 2 and 4, are expressed on the cell surface [34]. Their location is ideally suited for interactions with the surrounding environment or foreign substances that have entered the cellular cytoplasm. Surprisingly, the healthy intestine does not exhibit signs of inflammation, despite the massive bacterial load, the presence of numerous MAMPs, and the expression of TLRs [7, 35]. Recent studies have determined that intestinal epithelial cells display no detectable surface TLRs, such as TLR 4, but rather they are located on the basolateral cell surface where they would be positioned to interact with invading microorganisms [35]. Alternative localization of these receptors could represent a mechanism of preventing a detrimental immune assault on the commensal flora of the lumen and a persistent, harmful state of inflammation in intestinal tissues. However, intriguing research conducted in the laboratory of Ruslan Medzhitov has demonstrated that although TLRs may not be expressed on the surfaces of intestinal epithelial cells, ablation of TLR signaling by eliminating the adapter molecule Myd88 prevents wound healing in the large intestine

[32]. Thus, detection of the commensal microflora is required for basal levels of production of unidentified protective factors by epithelial cells. Additionally, sensing of penetrating microorganisms in the event of epithelial damage may result in an increase in these protective factors as well as swift activation of the adaptive immune system.

In addition to the membrane-bound TLRs, two cytoplasmic proteins, NOD1 and NOD2, also trigger potent inflammatory responses through the NF-κB pathway [36, 37]. Initial research demonstrated that NOD2 specifically recognizes muramyl dipeptide (MDP), the fundamental component of bacterial peptidoglycan, which would be released by invading, intracellular bacteria [38]. Less is known about the precise ligands of NOD1, but some evidence suggests that this receptor also responds to bacterial PGN [39]. The cell types expressing these proteins are more restricted than those producing TLRs, being primarily confined to macrophages and other innate immune effector cells, including intestinal Paneth cells [37, 40]. Thus, the NOD proteins are ideally situated to sense and respond to invading bacteria.

Interestingly, mutations in NOD2 that ablate its recognition of MDP and subsequent signaling have been genetically linked to patients with Crohn's disease, a debilitating intestinal disorder [41, 42]. Recent work by Kobayashi et al. has characterized mice deficient in NOD2, confirming that this protein is essential for detection of bacterial MDP and that it is required for robust expression of cryptdins 4 and 10, which are antimicrobial peptides secreted by intestinal Paneth cells [43]. These findings support those of Wehkamp et al. who described a decrease in the expression of α-defensins, the human homologues, in Crohn's disease patients [44, 45]. Therefore, the NOD proteins probably play an integral role in sensing breaches of the epithelial borders

in the intestine by subsequently signaling for production and release of known Paneth cell antimicrobial factors.

ii. The Paneth cell antimicrobial arsenal

First described near the beginning of the twentieth century, Paneth cells are easily distinguished from other intestinal epithelial cells by their distinct morphology and darkly staining granules. Ayabe et al. have convincingly demonstrated that these pyramidal cells secrete their granular contents apically into the crypt space in response to bacterial products [46]. Surrounding the prominent granules and the nucleus are many ridges of endoplasmic reticulum, belying the amount of protein translation undertaken by these cells. These specialized cells primarily function as manufacturers of an arsenal of antimicrobial proteins and peptides that can be stored in their large granules and then targeted apically into the intestinal lumen [47-49].

One of the most well known constituents of Paneth cell granules is lysozyme, first described by Alexander Fleming in 1922 as being a component of mucous secretions [50]. This potent muralytic enzyme was later shown to cleave bacterial peptidoglycan at the β-1,4-linked N-acetylmuramic acid (MurNAc) and N-acetylglucosamine (GlcNAc) residues that alternate to form the backbone of peptidoglycan [51]. Lysozyme is widespread throughout body tissues and fluids, produced particularly at interfaces between host cells and the environment such as the respiratory and gastrointestinal tracts.

In addition to this and other proteins, Paneth cell granules harbor myriad antimicrobial peptides. Defensins (or cryptdins in mice) are small cationic peptides that are produced and secreted abundantly by Paneth cells [46, 52-54]. They are estimated to comprise ~20% of the total granular contents, and demonstrate potent, non-specific

microbicidal activity [46]. All three forms of defensins (alpha, beta, and theta) are stabilized extracellularly by disulfide bonds formed between conserved cysteine residues [47, 55]. In order to achieve full functionality, the pro-peptide form of each defensin must be processed by phospholipase A_2 in mice or trypsin in humans to release the active form [56, 57]. While the precise mechanism of defensins' antimicrobial activity remains undefined, one model suggests that interactions of multiple defensins on a bacterial cell surface could result in pore formation, resulting in release of bacterial cytoplasmic contents [47, 55, 58]. Thus, a primary theme in the function of antimicrobial proteins and peptides is the disruption of the target cell wall.

D. C-type lectins – innate immune effectors

Yet another mechanism of innate immune defense is mediated by members of the calcium-dependent (C-type) lectin family of proteins, which bind particular carbohydrate structures. Carbohydrates are the most abundant form of post-translational modification of newly processed and folded proteins [59]. Experimental analysis has implicated carbohydrates in influencing such cellular processes as protein trafficking, adhesion, cell-cell contact, and immune recognition [60, 61]. Such diverse functions imply the existence of proteins equipped to decode the information stored within these modifications – proteins that were later designated as lectins [59].

Drickamer categorized the known lectins into two groups, C- and S-type lectins, for their use of calcium or thiol groups to coordinate carbohydrate ligands, respectively [59]. C-type lectins are classically defined as carbohydrate-binding proteins, which utilize a Ca^{2+} ion to coordinate the hydroxyl groups of the carbohydrate ring within a conserved

binding pocket [59, 62]. The presence of eighteen virtually invariant residues within this carbohydrate recognition domain (CRD) has allowed for the identification of many putative lectins by homology searches [59, 63]. Among these, multiple examples lectins from various organisms exist that bind their ligands using a calcium-independent mechanism, including the pancreatic protein, lithostatine, a member of the Reg family (see section F; [64]). In many cases, a biological carbohydrate ligand has yet to be identified; for instance, most of the ~130 putative C-type lectins found in a homology search in the *Caenorhabditis elegans* genome have no described function [65].

However, lectin functions that have been characterized are quite diverse, ranging from intra-cellular trafficking and mediating cell-cell contact to innate immune defense. Within the latter category, the collectin family, including mannose binding lectin (MBL) and the surfactant proteins (SP-A and SP-D), has been most extensively studied [66-69]. Early research grouped these lectins into the mannose-binding group, which contains lectins that rely primarily upon the orientation of the 3 and 4 carbon hydroxyl groups for binding. Because mannose, N-acetylglucosamine (GlcNAc), fucose, and glucose all have similar chemical structures at these positions, lectins that recognize mannose moieties often cross-react with the other listed carbohydrates [70]. Interestingly, these carbohydrates typically coat the surfaces of infectious agents, while the penultimate sugars on mammalian cells are usually sialic acid or galactose [68, 71]. The carbohydrates bound by these proteins and their presence as secreted proteins in either serum or lung surfactant led to the hypothesis that they may play a primary role in innate immune recognition of pathogens.

Studies of the structure of MBL reveal that is composed of an extended N-terminal collagen region followed by a α-helical domain, finally terminating in a canonical CRD. The collagen domain allows for multimerization of protein monomers, which results in the formation of a trimeric complex and a dramatic increase in carbohydrate ligand affinity (Figure 1.3, [68]). In support of the model describing MBL as an innate immune effector molecule, the crystal structure of MBL multimers reveals a 45Å spacing between joined CRDs, corresponding nicely to the predominant spacing of mannose-type carbohydrates on pathogen surfaces [68, 71]. Additionally, MBL has been shown to opsonize bacteria and to recruit complement through the lectin pathway of complement activation via the MBL-associated serine proteases (MASPs) 1 and 2 [72-74]. Genotypic studies in humans have identified mutations in MBL that reduce its efficacy in mediating clearance of invading pathogens, underscoring the importance of this protein as a primary serum factor involved in initial responses to microbes [75, 76].

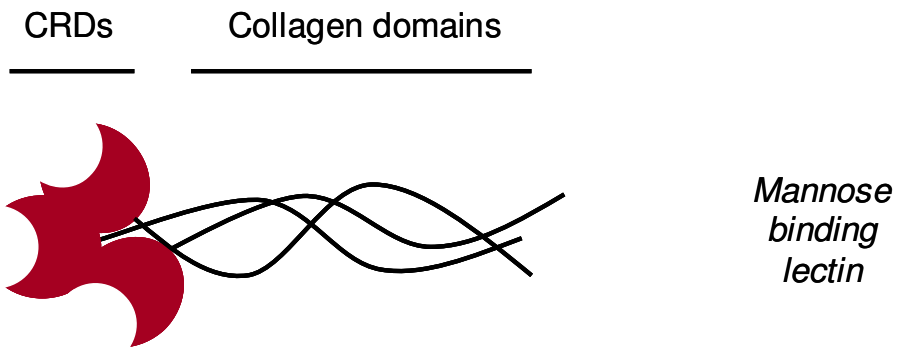

Figure 1.3: Domain organization of MBL.

N-terminal collagen domains allow protein monomers to form trimeric complexes that increase affinity for carbohydrate ligands [68].

E. Peptidoglycan recognition proteins (PGRPs)

The molecule peptidoglycan (PGN) is produced by all known bacterial species, and is a unique polymer of GlcNAc and MurNAc residues cross-linked by tetrapeptide crossbridges (Figure 5.7A, [77]). Because PGN is an exclusively bacterial molecule, it is an example of an ideal ligand to stimulate an innate immune response, and, indeed, PGN and its derivatives are potent inducers of an inflammatory response [78, 79]. A portion of this response can be attributed to the signaling molecules described above, but another family of innate immune effector proteins highly conserved from insects to humans, the peptidoglycan recognition proteins (PGRPs), also mediates recognition of this MAMP. Initially, these proteins were isolated by Yoshida et al. from the silkworm *Bombyx mori* [80], but were subsequently identified as being highly conserved in *Drosophila*, *Anopheles*, and mammals including cows, pigs, camels, rats, mice, and humans [81, 82].

A distinguishing feature of this family that facilitated identification of novel members in various organisms is the distinctive PGRP domain, which bears homology to T7 lysozyme, an N-acetylmuramoyl-L-alanine amidase [83-85]. Based on genetic sequence and protein size and structure, the family has been divided into two major groups, the short PGRPs (PGRP-S), which are secreted, ~19kD proteins, and the long PGRPs (PGRP-L), which have extended transcripts and contain at least one transmembrane domain [83]. Due to their small size, rendering them more amenable to purification, the PGRP-S members from many organisms have been well-characterized, and many have ascribed functions. Recently, much research has been conducted on the PGRP-L proteins, particularly from *Drosophila*, suggesting a role in signaling for these molecules that are generally expressed at the cell surface or within vacuoles [86-89].

The functions of PGRPs are more diverse in insects, where these proteins act in concert with TLRs to generate immune responses to pathogens [87, 90, 91]. Thus far, five major functions have been ascribed to PGRPs in insects, with two of those being conserved in mammals. In *Drosophila*, PGRP-LC signals for expression of antimicrobial peptides through the initiation of the Imd pathway, while PGRP-S and SA also induce antimicrobial effects – either the prophenol-oxidase pathway or activation of Toll [83, 92].

In mammals, the signaling pathways are not conserved; instead, the directly antimicrobial functions of these proteins participate in thwarting the spread of pathogens. For example, research from the laboratories of Roman Dziarski and Michael Selsted has elegantly demonstrated that bovine, murine, and human PGRP-Ss have directly bactericidal activity [93-95]. These findings are consistent with localization of this protein to neutrophil granules, and with the evidence demonstrating that it is secreted [96]. Mice that lack PGRP-S are more susceptible to intraperitoneal infection with *Bacillus subtilis*, but remain resistant to *E. coli*. This result was attributed to the inability of the neutrophils from PGRP-S$^{-/-}$ mice to kill either *B. subtilis* or *Micrococcus luteus*, supporting the previous results describing the release of this protein by neutrophil granules [93]. Studies investigating both the *Drosophila* and human forms of PGRP-L have suggested this protein functions as an N-acetylmuramoyl-L-alanine amidase in both organism, while its scavenger receptor function appears to be specific for the *Drosophila* isoform [97, 98]. Despite their varied functions, all reveal a role in initiating or participating in the innate immune response to foreign invasion of host tissues.

Strikingly, affinity measurements of various PGRPs and solutions of the crystal structures of certain members have revealed a marked specificity for certain PGNs above others [99-101]. Although the glycan structure of PGN remains remarkably constant across bacterial species, the amino acids comprising the peptide moiety vary according to species, with the most diversity occurring at the third amino acid [77]. The majority of bacteria, including most Gram-negative organisms as well as *Bacillus* and *Clostridia* species, produce PGN containing diaminopimelic acid (DAP), an alternate amino acid, at position three. By contrast, many Gram-positive organisms generate PGN with L-lysine substituted as the third amino acid [77]. Using surface plasmon resonance (SPR), Kumar and colleagues measured the affinity of human PGRP-S for DAP-PGN as being ~300 nM, while this same protein bound Lys-PGN with only ~100 μM affinity [100]. The inverse was found to be true for human PGRP-Iα [100].

The solution of the crystal structure of the *Drosophila* PGRP-LCx and LCa ectodomains by Chang et al. provides a structural explanation for the binding specificity of these protein domains in binding DAP-PGN, a preference that had been described previously [96]. Stacking interactions between DAP and the MurNAc carbohydrate moiety within the binding groove induce proper orientations in phenylalanine and tryptophan residues, which contribute to LCa and LCx domain interactions [99]. Together, these studies indicate that various members of the PGRP family are highly specific for PGN from different organisms – an attribute that shapes a more targeted innate immune response in both insects and mammals.

F. The RegIII Family

Analysis of cDNAs from rat-derived regenerating pancreas islets identified a protein, designated "regenerating factor" or *Reg*, due to its elevated expression solely in these regenerating islets [102]. After confirming transcribed mRNA, this group predicted the resulting protein sequence as containing 165 amino acids with a 21 residue leader peptide. Subsequent homology searches and genome sequencing demonstrated that this *Reg* gene, later designated *Reg*I, was conserved among rats, hamsters, mice, cows, and humans [103-105]. As the number of genes and proteins displaying homology to this original family member increased, the *Reg* family was grouped together according to homology into *Reg*I – IV (Table 1.1). Though these proteins can be found in distinct tissues, they share many similarities. Almost all are localized to the small intestine, pancreas, or liver; their genetic locus structure including six exons separated by five introns is highly conserved; and their gene products invariably encode for ~170 amino acid putative C-type lectins.

Table 1.1: Comparison of the Reg Family of C-type Lectins

Superfamily group	Species	Chromosome localization	Length of amino acid
Reg I	Mouse Reg I	6	165
	Rat Reg	4q33-q34	165
	Human Reg/PSP/PTP	2p12	166
RegII	Mouse Reg II	6	173
	Human Reg II	2p12	175
RegIII	Rat PAP	4q33-q34	175
	Rat peptide 23	4q33-q34	175
	Hamster INGAP		
	Bovine PTP	N/A	175
	Human HIP/PAP	2p12	175
	Human RegIII	2p12	175
	Mouse RegIIIa, b, g, d	6	175, 174 (g)
RegIV	Human Reg IV	1q12-q21	158
	Mouse Reg IV	2	158

Modified from ref. [105].

Research conducted by Narushima et al. identified three members of the *Reg*III family (α, β, and γ) simultaneously due to their genetic proximity to each other (within 75 kb) on mouse chromosome 6c [104]. Further analysis determined that an additional *Reg*III gene, *Reg*IIIδ, is located within the same mouse genetic locus as well as mouse *Reg*I and *Reg*II. This study also analyzed the promoter region of *Reg*IIIδ, revealing interferon-stimulated response elements (ISREs) and multiple STAT binding sites in the promoters of this gene [106]. Later studies also described IL-6 response elements in the promoter of the rat form of *PAP* or *Reg*IIIβ, as well as significant transcriptional induction in response to IFN-γ and TNF-α [107].

Interestingly, though the family is conserved in many mammals, the four proteins identified in mice correspond to only two in humans and one in rats and hamsters [106, 108, 109]. This suggests not only gene duplication within the *Reg* family as a whole, but

also swift generation of new, highly homologous protein-coding sequences within the *Reg* gene subsets as well. Strikingly, ClustalW analysis of human and mouse RegIII proteins indicates that RegIIIβ and RegIIIγ are more homologous to the human HIP/PAP than to the other mouse proteins RegIIIα and RegIIIδ (Figure 1.4). This supports the idea of diversification of these proteins from a common ancestor gene. The rapid diversification of this family and the presence of innate cytokine response elements in the promoters suggest that these genes and their products could be involved in innate immunity [107].

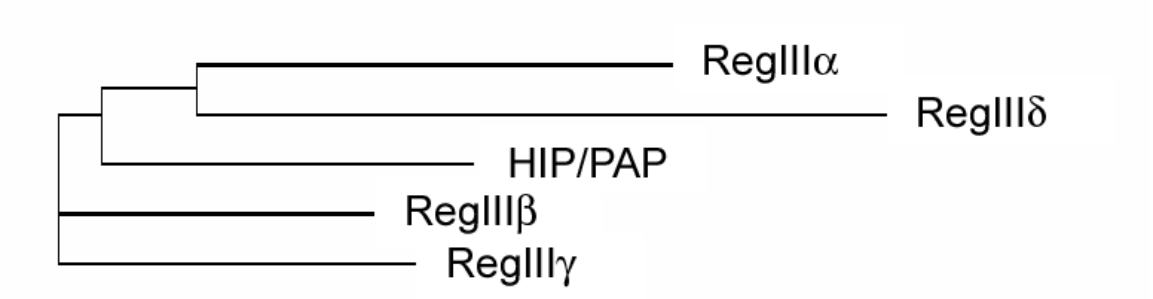

Figure 1.4: ClustalW analysis of the RegIII proteins.

A cladogram generated from aligning the sequences of the four murine RegIII proteins with the human protein, HIP/PAP, reveals that HIP/PAP is more closely related to RegIIIβ and RegIIIγ than these proteins are to the other mouse RegIII proteins.

Corresponding to this proposed innate immune function, many groups swiftly observed that all of the RegIII proteins are small ~16kD proteins consisting primarily of a putative CRD attached to an N-terminal signal peptide [104-106, 108, 110]. CRDs are the carbohydrate-binding portion of the lectin protein family, and the presence of a CRD in the RegIII proteins' primary sequence suggests that they have target carbohydrate ligands. The structure of these proteins, secreted CRDs without other domains, represented a previously unidentified form of C-type lectin. In fact, according to the Drickamer lectin classification, the RegIII proteins are the only members of the type VII

family, which consists of secreted CRDs, in mammals [63]. Because C-type lectins have diverse functions in many systems, initial efforts at determining the function of the RegIII proteins have centered on verifying these proteins' abilities to bind carbohydrates.

The human RegIII protein, HIP/PAP is overexpressed in the liver of hepatocellular carcinoma patients, in the pancreas of those with pancreatitis, and in the colons of patients with inflammatory bowel disease (IBD; [111, 112]). With so many links to disease, this protein has been extensively studied compared to the other Reg proteins. Research conducted by Christa et al. utilized both a GST-fusion protein approach and a transgenic mouse expression system to generate recombinant protein for carbohydrate binding studies, where they observed only moderate binding of the recombinant GST-HIP/PAP fusion protein to the disaccharide lactose [110, 113]. Subsequent studies by this group have suggested multiple activities for HIP/PAP, including binding to RIIα regulatory subunit of cAMP-dependent protein kinase and the ability to initiate liver regeneration in liver-resected transgenic mice [114, 115]. However, these results are not definitive because HIP/PAP is a secreted protein and would therefore have no opportunity to interact with a cytoplasmic kinase. Additionally, this family of proteins is primarily expressed in the small intestine under non-disease conditions, with limited expression in the pancreas and nervous system, but none detected in the liver [104, 106, 111, 116].

To date, no biochemical analyses of the mouse RegIII proteins have been undertaken, despite the significant upregulation of both RegIIIβ and RegIIIγ in multiple mouse models [117, 118]. Particularly, in the mouse model of congenital diarrhea, animals lacking NH3, the major ion exchanger in the intestine, display a 5-fold increase

in both RegIIIβ and RegIIIγ transcripts, which the authors speculate is an attempt to maintain intestinal homeostasis [118]. Interestingly, Keilbaugh et al. observed a significant induction of these same transcripts in the colons of immunocompromised SCID mice colonized with a defined microflora, but not in the colons of their germ-free counterparts [117]. Although it has been speculated that Reg family members could be acting as mitogens to stimulate epithelial growth and regeneration in response to damage [115, 119, 120], no biochemical evidence conclusively supports these claims. However, the observation that the RegIII proteins, both mouse and human, are induced or ectopically expressed under inflammatory or immunocompromised conditions does imply a function for these proteins in maintaining homeostasis or responding to infectious agents. Generating a ready source of purified, recombinant RegIII proteins, followed by extensively characterizing their functions were the goals of the research presented in Chapters 4-6.

Chapter Two:

OVERALL OBJECTIVE AND SYNOPSIS

The intestines of humans and other mammals are colonized with a quantitatively vast and diverse society of resident bacteria that play a crucial role in host nutrient metabolism. To maintain this commensal relationship, resident microbes must be prevented from crossing the intestinal epithelium into host tissues, for the consequences of such a breach can be dire, including inflammation and sepsis. Therefore, the innate immune system plays a crucial role in preventing bacterial incursions across gut epithelial surfaces by limiting bacterial attachment and encroachment at epithelial surfaces. Previous research has demonstrated that Paneth cells, specialized small intestinal epithelial cells, produce and secrete a variety of antimicrobial proteins and peptides [10, 46, 47, 52, 57]. In order to gain new insights into the adaptation of mucosal surfaces to microbial challenges, we employed a DNA microarray strategy to screen for Paneth cell

genes whose expression is modulated by intestinal microbes. This screen revealed that expression of two C-type lectins, RegIIIβ and RegIIIγ, is strongly induced following intestinal colonization with resident microbes (43.9 ± 18.8 and 31.7 ± 13.5 fold, respectively).

On initial inspection, several features suggested that members of the RegIII family may have microbicidal functions. First, they are members of the C-type lectin superfamily of proteins. Other C-type lectins, including the mannose binding lectin, have well-characterized innate immune functions and play critical roles in microbial killing by recruiting complement. Second, our microarray screen identified RegIII transcripts in Paneth cells, which are known innate immune effector cells. Lastly, these and other RegIII family members have been described as being upregulated in response to inflammation and infection. These data combined with an analysis of these proteins' amino acid sequence revealing an N-terminal secretion sequence suggested that they could be targeted to the gut lumen where they would likely encounter luminal bacteria. Based on these observations, we hypothesized that this family of proteins may perform an innate immune function, specifically antimicrobial defense.

In order to elucidate the function of the RegIII proteins, protein purification and biochemical analyses of each of the three murine RegIII proteins and the human homolog HIP/PAP were undertaken. The data generated in these studies indicate that our purification procedure results in functionally active RegIII proteins as determined by a variety of biochemical techniques. Purified, recombinant RegIII proteins bound specifically to mannan and chitin, which are polysaccharides of mannose and GlcNAc, respectively. This is the first example of a successful purification scheme for the RegIII

proteins that results in quantities sufficient for subsequent functional assays. Additionally, the carbohydrate binding profiles for RegIIIβ, RegIIIγ, and HIP/PAP are among the first to describe potential ligands for these C-type lectins.

Subsequent analysis of RegIII binding using a combination of flow cytometry and pulldown techniques demonstrated that purified recombinant RegIIIγ binds specifically to Gram positive bacteria by interacting with exposed peptidoglycan (PGN) on the bacterial cell surface. This binding occurs via lectin-like recognition of the carbohydrate portion of PGN, as confirmed by binding of the RegIII proteins to chitin, which mimics the carbohydrate backbone of PGN. Following binding to PGN, RegIIIβ, RegIIIγ and HIP/PAP exhibit direct bactericidal activity against target Gram positive but not Gram negative organisms. This killing activity is connected to carbohydrate binding, for both RegIIIγ and HIP/PAP were inhibited with soluble polysaccharides whose structures mimicked those carbohydrates identified in the pull-down assay. These data represent the first example of a directly bactericidal C-type lectin, and also demonstrate a novel mechanism of PGN recognition by effector proteins.

While the RegIII family has two members in humans and one in rats and hamsters, this same family has diversified to incorporate four members in mice. This diversification is characteristic of other innate immune effecter protein families such as the defensin family, and suggests that this diversification occurred to combat encountered infectious agents. Purified, recombinant RegIIIβ was analyzed using the assays outlined for RegIIIγ in order to investigate the function of this additional RegIII family member. The results of these analyses indicate that RegIIIβ binds to the same population of intestinal bacteria as its family member, RegIIIγ, and that this recognition is mediated by

specific binding to exposed PGN. These data reveal an overlapping, but perhaps not completely redundant, binding function for each RegIII protein in the murine intestine, and will be instrumental in the future development of targeted gene deletions of each protein in mice.

Through extensive biochemical characterization of the functional characteristics of the RegIII proteins, the research described in this thesis reveals new insights into how the intestinal epithelium utilizes specific effector proteins to protect internal tissues from encroachments by the gut's vast commensal populations. A more complete understanding of this cross-talk could aid in understanding why these communications sometimes fail, as in Crohn's disease and IBD.

Chapter Three:

MATERIALS AND METHODS

A. Vectors, strains, and supplies

The expression vector pET3a was from Novagen. *E. coli* BL21-CodonPlus (DE3)-RIL and *E. coli* BL21-CodonPlus (DE3)-RILP competent cells were from Stratagene. Oligonucleotides and restriction enzymes were supplied by Invitrogen. Other DNA modifying enzymes and isopropyl-β-D-thiogalactopyranoside (IPTG) were from Roche Molecular Biochemicals. Luria Broth and Brain Heart Infusion Broth were purchased from VWR. Sephacryl S-100 high resolution gel filtration medium and size exclusion chromatography standards were from GE Healthcare. Goat anti-rabbit horseradish peroxidase was from Jackson Immuno-Research. Chitooligosaccharides were purchased from Seikagaku. Peptidoglycan purified from *Bacillus subtilis* was purchased

from Sigma-Aldrich. SYBR Green was from Stratagene; other real-time PCR reagents were from Invitrogen. All other chemicals and reagents were from Sigma.

Table 3.1: Primers used in this study

Primer Name	Oligonucleotide (5' - 3')	Reference
actin-F	GAAGTACCCCATTGAACATGGC	[49]
actin-R	GACACCGTCCCCAGAATCC	[49]
18SFORWARD	CATTCGAACGTCTGCCCTATC	[49]
18SREVERSE	CCTGCTGCCTTCCTTGGA	[49]
Reg3g-2F	ATTGCGAGGCATATGGAAGTTGCCAA GAAAGATGCCCCAT	This study
Reg3g-2R	CTATGGGGATCCCTAGGCCTTGAATTT GCAGACATAGGGT	This study
hPAP-F	ATTGCGAGGCATATGGAAGAACCCCA GAGGGAACTGC	This study
hPAP-R2	CTATGGTGATCACTAGTCAGTGAACTT GCAGACATAGGGTAA	This study
hPAPm1-F	ATTGCGAGGCATATGGAAGAACCACA AAGAGAACTGC	This study
Reg3g-F	TTCCTGTCCTCCATGATCAAAA	This study
Reg3g-R	CATCCACCTCTGTTGGGTTCA	This study
Reg3b-2F	ATTGCGAGGCATATGGAAGACTCCCTG AAGAATATACCCTCCG	This study
Reg3b-2R	CTATGGGGATCCTTAACCAGTAAATTT GCAGACATAGGGCAAC	This study

Reg3b-1F	TACTGCCTTAGACCGTGCTTTCTG	This study
Reg3b-1R	GACATAGGGCAACTTCACCTCACA	This study
RegIIIbISH-F2	ATTGCGAGGAAGCTTATGGGAATGGA GTAACAAT	This study
RegIIIbISH-F3	ATTGCGAGGAAGCTTGGAGTAACAAT GACGT	This study
RegIIIgISH-F1	ATTGCGAGGAAGCTTTGGAGCAATGCT	This study
RegIIIgISH-R1	CTATGGGGATCCGAATTTGCAGACATA GGGTAACTCTAA	This study
Reg3a-F	ATTGCGAAGCATATGGAAGACTTCCAG AAGGAAGTGCCC	This study
Reg3a-R	CTATGGGCATGCCTACTGCTTGAACTT GCAGACAAATGGTAATG	This study
RegIIIa-4F	ACTGTGGGAGTCTGACAGCAAGT	This study
RegIIIa-4R	AGCCTTGCCGGGTGAATTCTTGT	This study
RegIIIaISH-F1	ATTGCGAGGAAGCTTGAGTGGAGTAA CTCC	This study
RegIIIaISH-R1	CTATGGTGATCAGCTTGAACTTGCAGA CAAATGGTAATGT	This study
RegIIIaISH-F1	ATTGCGAGGAAGCTTGAGTGGAGTAA CTCC	This study
RegIIIaISH-R1	CTATGGTGATCAGCTTGAACTTGCAGA CAAATGGTAATGT	This study
Reg3Delta - F	GACACCTGGCATTTCTGCTCACTT	This study
Reg3Delta - R	AGAGCTGCTCCACTTCCATCCATT	This study

B. Construction of the mouse RegIIIγ expression vector

A 474 bp amplicon was generated by RT-PCR from mouse small intestinal RNA using the specific primers 5′-ATTGCGAGGCATATGGAAGTTGCCAAGAAAGATGCCCCAT-3′ (forward primer) and 5′- CTATGGGGATCCCTAGGCCTTGAATTTGCAGACATAGGGT-3′ (reverse primer). The forward primer contains an *Nde*I restriction site (underlined) for cloning into pET3a. The reverse primer incorporates the native stop codon followed by an engineered *Bam*HI site (underlined). The resulting amplicon contained a methionine start codon in place of the signal sequence, and thus encoded the mature secreted protein. PCR products and vector were digested with *Nde*I and *Bam*HI, gel-purified, and ligated. The recombinant plasmid (pET3a-RegIIIγ) was sequenced to confirm the absence of mutations, and was transformed into *E. coli* BL21-CodonPlus (DE3)-RIL for protein expression.

C. Construction of the RegIIIβ expression vector

A 474 bp amplicon was generated by RT-PCR from mouse small intestinal RNA using the specific primers 5′-ATTGCGAGGCATATGGAAGACTCCCTGAAGAATATACCCTCCG-3′ (forward primer) and 5′-CTATGGGGATCCTTAACCAGTAAATTTGCAGACATAGGGCAAC-3′ (reverse primer). The forward primer contains an *Nde*I restriction site (underlined) for cloning into pET3a. The reverse primer incorporates the native stop codon followed by

an engineered *Bam*HI site (underlined). The resulting amplified sequence contained a methionine start codon in place of the signal sequence, and thus encoded the mature secreted protein. PCR products and vector were digested with *Nde*I and *Bam*HI, gel-purified, and ligated. The recombinant plasmid (pET3a-RegIIIβ) was sequenced to confirm the absence of mutations, and subsequently transformed into *E. coli* BL21-CodonPlus (DE3)-RIL for recombinant protein expression.

D. Construction of the RegIIIα expression vector

A bp amplicon was generated by RT-PCR from mouse small intestinal RNA using the specific primers 5′-ATTGCGAAGCATATGGAAGACTTCCAGAAGGAAGTGCCC-3′ (forward primer) and 5′-CTATGGTGATCACTACTGCTTGAACTTGCAGACAAATGGTAATG-3′ (reverse primer). The sequence for RegIIIα is unique among the mouse RegIII family in that it contains an endogenous *Bam*HI site at ~360 bp into the amino acid coding sequence. Therefore, the *Nde*I restriction site was maintained in the forward primer (underlined), while the reverse primer incorporates the native stop codon followed by an engineered *Bcl*I site (underlined). The nucleotide sequence recognized by *Bcl*I (TGATCA) generates a cohesive end compatible for ligation to DNA digested with *Bam*HI. PCR products and vector were digested with *Nde*I and *Bcl*I, gel-purified, and ligated into pET3a that had been digested with *Nde*I and *Bam*HI. The resulting sequence contains a methionine start codon in place of the signal sequence, and thus encodes the mature secreted protein. The recombinant plasmid (pET3a-RegIIIα) was sequenced to

confirm the absence of mutations, and subsequently transformed into *E. coli* BL21-CodonPlus (DE3)-RIL for recombinant protein expression.

E. Construction of HIP/PAP expression strains

A 474 bp amplicon was generated by RT-PCR from human small intestinal RNA (Ambion) using the specific primers 5′-ATTGCGAGGCATATGGAAGAACCCCAGAGAGGAACTGC-3′ (forward primer) and 5’- CTATGGTGATCACTAGTCAGTGAACTTGCAGACATAGGGTAA-3′ (reverse primer). The forward primer contained an *Nde*I restriction site (underlined) for cloning into pET3a. The reverse primer incorporates the native stop codon followed by an engineered *Bcl*I site (underlined). The resulting amplicon lacks the HIP/PAP signal sequence and thus encodes the mature secreted protein [8]. The PCR product was digested with *Nde*I and *Bcl*I, ligated into *Nde*I/*Bam*HI-digested pET3a, and the resulting plasmid (pET3a-HIP/PAP) sequenced to confirm the absence of mutations.

A second expression construct (pET3a-HIP/PAP*mut*) was generated with silent mutations engineered into the 5’ end of the HIP/PAP coding sequence. Mutations were introduced by redesigning the forward primer that was used to generate the wild-type HIP/PAP construct:

5’-ATTGCGAGGCATATGGAAGAACCACAAAGAGAAACTGC-3′ (mutant bases are underlined; also see Figure 1). A 474 bp amplicon was generated by PCR with this mutant primer and the *HIP/PAP*-specific reverse primer above, using pET3a-HIP/PAP as template. The amplicon was cloned into pET3a as described for pET3a-HIP/PAP. The resulting plasmid was sequenced to confirm incorporation of the silent mutations and the

absence of additional mutations. Both pET3a-HIP/PAP and pET3a-HIP/PAP*mut* were transformed into *E. coli* BL21-CodonPlus (DE3)-RILP for protein expression.

F. Expression and purification of RegIIIγ

E. coli BL21-CodonPlus (DE3)-RIL harboring pET3a-RegIIIγ were grown at 37^{o}C in 500 ml of LB broth supplemented with 0.1 mg/ml ampicillin to an absorbance of 0.6-1.0 (mid-log phase) at 600 nm. Protein expression was induced by the addition of 0.4 mM IPTG, and the culture was incubated for another 3 hours at 37°C with good aeration. Cells were collected by centrifugation at 6,500*g* for 15 minutes at 4°C, and the pellet resuspended in 1/20 culture volume (25 ml) of Inclusion Body (IB) Wash Buffer (20mM Tris-HCl, 10 mM EDTA, 1% Triton X-100, pH 7.5). The cells were divided into five equal 5 ml aliquots and ruptured by sonication in two 1 minute bursts at setting 4 using a Misonix XL-2020 Sonicator fitted with a 4.8 mm tapered probe. The lysate was centrifuged at 10,000*g* for 10 minutes, and the insoluble fraction was resuspended in 50 ml of IB Wash Buffer using a Dounce homogenizer. Centrifugation and resuspension were repeated, and the final insoluble inclusion body preparation was collected by centrifugation at 10,000*g* for 10 minutes followed by dispersion in 10 ml of Resuspension Buffer (7M guanidine-HCl, 0.15M reduced glutathione, 0.1M Tris-HCl, 2 mM EDTA, pH 8.0) and rotation for 2 hours at room temperature. Four milliliters of the resuspended inclusion bodies were added dropwise to a total of 200 ml of RegIIIγ Refolding Buffer (0.5M arginine-HCl, 0.6mM oxidized glutathione, 50 mM Tris-HCl, pH 8.0) and left to stand for 24 hours. The solution was clarified by centrifugation at 10,000*g* for 30 minutes, and dialyzed overnight against 10 volumes of Dialysis Buffer 1 (25 mM Tris-

HCl, 25 mM NaCl, 2 mM $CaCl_2$, pH 7), followed by a second overnight dialysis against 10 volumes of Dialysis Buffer 2 (25 mM MES, 25 mM NaCl, 2 mM $CaCl_2$, pH 6). The dialysate was centrifuged at 10,000*g* for 30 minutes, and RegIIIγ was captured by passage over a 5 ml column of SP-Sepharose Fast Flow cation exchange resin (Sigma) equilibrated in Dialysis Buffer 2. After washing in 10 column volumes of Dialysis Buffer 2, the protein was batch-eluted in 5 column volumes of 0.6 M NaCl in Dialysis Buffer 2. Fractions containing protein were identified by the Bradford method [121] using the Bio-Rad Protein Assay Dye Reagent Concentrate (Bio-Rad).

G. Expression and purification of RegIIIβ

The expression and purification procedure developed for RegIIIβ was based upon that used for RegIIIγ, but modified to maximize protein recovery. Solubilized RegIIIβ inclusion bodies were refolded by dilution into a total of 200 mL of RegIIIβ Refolding Buffer (50 mM Tris-HCl pH 8.0, 0.5 M guanidine-HCl, 2 mM $MgCl_2$, 2 mM $CaCl_2$, 0.6 M arginine-HCl, 0.6 mM oxidized glutathione). This buffer was identified in a screen performed for all of the murine RegIII proteins, and the results for each protein of the screen are listed in Appendix A. Subsequent dialysis and ion-exchange chromatography steps were performed as described for RegIIIγ.

H. Expression and purification of RegIIIα

The expression and purification protocol developed for RegIIIα was similar to that of the other RegIII proteins, with minor differences. The refolding of recombinant

RegIIIα was performed by dilution of 4 mL of solubilized inclusion bodies into 200 mL of RegIIIα Refolding Buffer (50 mM Tris-HCl pH 8.0, 0.5M guanidine-HCl, 250 mM NaCl, 10 mM KCl, 2 mM $MgCl_2$, 2mM $CaCl_2$, 0.6 mM arginine-HCl, 0.6 mM oxidized glutathione). Following clarification by centrifugation, refolded RegIIIα was dialyzed against two changes of 10 column volumes of Column Loading Buffer (25 mM Tris-HCl pH 8.5, 25 mM NaCl). The dialysate was centrifuged at 10,000*g* for 30 minutes, and RegIIIα was captured by passage over a 5 ml column of Q-Sepharose Fast Flow anion exchange resin (Sigma) equilibrated in Column Loading Buffer. After washing in 10 column volumes of Column Loading Buffer, the protein was batch-eluted in 5 column volumes of 0.4 M NaCl in 25 mM Tris-HCl pH 7.5.

I. Expression and purification of HIP/PAP

Recombinant HIP/PAP was expressed from pET3a-HIP/PAP*mut*. The expression and purification protocol was similar to that of RegIIIγ, with some significant changes. First, the IPTG concentration used for protein induction was 1 mM, and induction proceeded for 2 hours. HIP/PAP-containing inclusion bodies were refolded in HIP/PAP Refolding Buffer (50 mM Tris-HCl pH 8.0, 10 mM KCl, 240 mM NaCl, 2 mM $MgCl_2$, 2 mM $CaCl_2$, 0.5M guanidine-HCl, 400 mM sucrose, 500 mM arginine-HCl, 1 mM reduced glutathione, 0.1 mM oxidized glutathione). Subsequent dialysis and ion-exchange chromatography steps were performed as described for RegIIIγ.

J. Characterization of recombinant proteins

Purity of recombinant protein preparations was evaluated by SDS-PAGE through 15% gels. N-terminal sequencing on an ABI494 sequencer (PE Biosystems) and Matrix Assisted Laser Desorption/Ionization Time of Flight (MALDI-TOF) mass spectrometry using a Micromass spectrometer in the UT Southwestern Protein Chemistry Technology Center was performed for RegIIIγ, RegIIIβ, and HIP/PAP.

K. Column assays

i. Size exclusion chromatography

Size exclusion chromatography was performed using a 1.5 x 63 cm Sephacryl S-100 column, at a flow rate of 22 ml/hour. The column was equilibrated in 25 mM Tris-HCl pH 7.5, 25 mM NaCl, 2 mM $CaCl_2$. 1 ml of protein at 2 mg/ml was applied and eluted in equilibration buffer, and eluted fractions were monitored for protein at 280 nm. Molecular weights were determined in comparison to the standards provided in the Amersham Low Molecular Weight Calibration Kit (GE Healthcare).

ii. Carbohydrate columns

Ten milliliter mannose-Sepharose or mannan-Sepharose columns were run in 25 mM MES pH 6.0, 25 mM NaCl and were eluted in 25 mM MES pH 6.0, 25 mM NaCl containing either 10 mM $CaCl_2$ or 20 mM $CaCl_2$. Collected fractions were analyzed for protein content by spectrophotometry at 280 nm.

L. Circular dichroism

Circular dichroism (CD) analysis was performed on an Aviv 62DS spectropolarimeter with a 1 mm cell length. Spectra of purified RegIIIγ and HIP/PAP were recorded in 25 mM Tris-HCl pH 7.5, at a protein concentration of 10 μM. Three spectra were recorded for each condition from 190 to 260 nm in 1 nm increments, averaged, and the background spectrum of buffer without protein was subtracted from the protein-containing spectra. CD spectra were initially analyzed by the software accompanying the spectropolarimeter. Analysis of spectra to extrapolate secondary structures was performed by Dichroweb [122] (website is found at http://www.cryst.bbk.ac.uk/cdweb/html/home.html) using the K2D and Selcon 3 analysis programs [123, 124].

M. Polyclonal antibody generation and Western blot analysis

Purified RegIIIγ was submitted to Cocalico Biologicals for polyclonal antibody generation in rabbits. Protein extracts for Western blot analysis were generated from mouse small intestine (jejunum). A 2 cm piece of freshly isolated intestinal tissue was flushed, lyophilized overnight, and pulverized under liquid N_2. The pulverized tissue was resuspended in 1 ml of Extraction Buffer (8 M urea, 1% SDS, 0.15 M Tris-HCl pH 7.5) and lysed by passing the suspension through an 18 gauge needle 3-5 times, followed by 3-5 passages through a 21 gauge needle. Total protein was quantitated with the Bio-Rad Detergent Compatible (DC) protein assay (Bio-Rad). Tissue protein and recombinant RegIIIγ were subjected to SDS-PAGE electrophoresis in a 15% gel and transferred to PVDF (Millipore). Membranes were blocked with 5% nonfat milk and incubated with

polyclonal antiserum or preimmune serum followed horseradish peroxidase-conjugated goat anti-rabbit IgG (Jackson ImmunoResesarch). Immunoreactivity was detected using the Pierce SuperSignal West Pico Chemiluminescent detection kit.

N. Carbohydrate binding assays

Yeast mannan (Sigma) or dextran (Sigma) were coupled to Sepharose 6B using a previously published protocol [125]. Mannose-agarose was purchased from Sigma. 25 µl of each resin was washed extensively in Binding Buffer (25 mM MES pH 6.0, 25 mM NaCl, 1% BSA), and 50 µg recombinant protein was added to each resin in a total volume of 1 ml of Binding Buffer. After rotation for 2 hours at 4°C, the beads were washed twice with 1 ml of Wash Buffer (25 mM MES pH 6.0, 25 mM NaCl). Bound protein was eluted by boiling the beads in SDS-PAGE buffer (10% glycerol, 5% β-mercaptoethanol, 2% SDS, 62.5 mM Tris-HCl, 0.003% bromophenol blue, pH 6.8) and resolved by SDS-PAGE through a 15 % acrylamide gel.

O. Germ-free mice

Germ-free NMRI/KI mice [126] were maintained in plastic gnotobiotic isolators on a 12 hour light cycle and given free access to an autoclaved chow diet. 8-12 week old males were inoculated with an unfractionated microbiota isolated from ileum/cecum of conventionally-raised NMRI/KI mice. The resulting mice were designated "conventionalized" mice. Mice were killed 10 days later, and small intestinal tissues were

prepared for laser capture microdissection as described below or snap frozen for RNA or protein extraction alongside age- and gender-matched germ-free controls.

For monocolonization experiments, age-matched germ-free C57/b6 and RAG1$^{-/-}$ mice were inoculated with stationary phase cultures of *Bacteroides thetaiotaomicron* (strain VPI-5482) and *Listeria innocua* (strain CLIP 11262) by spreading the organisms on fur as previously described [126]. Mice were killed 10 days later, and the distal 1 cm of intestine was used to determine CFU/ml ileal contents. All mice were colonized to ~10^8 CFU/ml.

P. Real-time PCR analysis

Total RNA was isolated from small intestinal tissues using the Qiagen RNeasy RNA isolation kit. RNAs were treated with deoxyribonuclease (Roche) prior to random-primed cDNA synthesis and real-time PCR analysis. SYBR Green-based real-time PCR was performed using primers specific for each RegIII transcript (see Table 3.1). Control experiments established that amplicons were derived from cDNA and not from genomic DNA or primer-dimers. Signals were normalized to 18S rRNA (forward primer: 5'-CATTCGAACGTCTGCCCTATC; reverse primer: 5'-CCTGCTGCCTTCCTTGGA). Normalized data were used to quantitate relative levels of RegIIIγ using ΔΔCt analysis [127].

Q. Electron microscopy

Paneth cell electron microscopy was performed as described [128] with tissues isolated from a male C57/B6 mouse housed in a barrier facility. Freshly isolated intestine was divided into thirds, and the distal third was fixed overnight at 4°C in 4% paraformaldehyde, 0.1% gluteraldehyde in 0.1M phosphate buffer. Tissues were dehydrated in an ethanol gradient at decreasing temperatures (0 to -35°C), then embedded in Lowicryl HW20 under UV illumination for two days. Embedded tissues were then sectioned using a microtome, floated onto grids, and visualized using the JEOL 1200 EX electron microscope.

For electron microscopy of bacteria, mid-log phase cultures of *L. monocytogenes* (strain EGD-e) were washed and resuspended in 900 μL of 25 mM MES pH 6.0, 25 mM NaCl. 300 μL of resuspended bacteria were added to each of three reactions containing either buffer alone, 10 μM RegIIIγ, or 10 μM HIP/PAP. Reactions were incubated for 2 hours at 37°C before the bacteria were pelleted. Bacteria were washed once with assay buffer, and each pellet was resuspended in 500 mL of 5% gluteraldehyde in 0.1M phosphate buffer. Bacterial pellets were cryo-embedded, sectioned, and positioned on grids as for the mouse tissues described above. Images were obtained using a JEOL 1200 EX electron microscope and AMT410 digital imaging equipment and software.

R. Flow cytometry

Recombinant RegIIIγ and RegIIIβ were labeled with AlexaFluor555 and AlexaFluor 488, respectively, using the Molecular Probes AlexaFluor555 or 488 MicroScale Protein Labeling kit. Intestinal bacteria were harvested from conventional

C57/b6 mice after excising complete small intestines. The contents from each intestine were flushed with sterile 25 mM MES pH 6, 25 mM NaCl into buffer containing protease inhibitors (Roche) incubated on ice. Large particulate matter was allowed to settle to the bottom of the tube, the cleared supernatant was removed to a fresh tube, and formaldehyde was added to each supernatant to a final concentration of 2%. After adding formaldehyde, the samples were incubated for fifteen minutes on ice. Additional particulate matter was separated by centrifugation at 500 rpm for 2 minutes, and the supernatant was retained. The supernatants were centrifuged at 4000 rpm for 5 minutes to pellet intestinal bacteria. The pellets were washed in Standard Assay buffer + 0.5% BSA and resuspended in 1 mL of the same buffer. 100 μL of resuspended bacteria were incubated with 30 μM RegIIIγ-AlexaFluor555 on ice for 40 minutes. Bacteria were washed twice with reaction buffer and subsequently incubated with or without RegIIIβ – AlexaFluor488 or WGA-AlexaFluor488 (Molecular Probes). Bacteria were washed three times in buffer with no BSA, then resuspended in 2 mL of buffer without BSA and analyzed using a FACScalibur flow cytometer and CellQuest Pro analysis software.

Mid-log phase cultures of *E. faecalis*, *L. innocua*, *S. typhimurium* and *P. aeruginosa* were pelleted at 4000 rpm for 5 minutes and fixed in 2% formaldehyde for 15 minutes. Following fixation, cells were washed in Standard Assay buffer + 0.5% BSA, and resuspended in 1 mL of the same buffer. 25 μL of each bacterial species was incubated with or without 30 μM RegIIIγ for 40 minutes on ice, followed by staining with rabbit anti-RegIIIγ and goat anti-rabbit-Cy3 (Biomeda). Following antibody staining, bacterial pellets were washed three times in buffer without BSA, resuspended in 2 mL of buffer without BSA, and analyzed as for intestinal bacteria.

S. Peptidoglycan binding assays

Peptidoglycan from *Bacillus subtilis* (Sigma) was pelleted and washed in Standard Assay buffer (25 mM MES pH 6.0, 25 mM NaCl). Recombinant RegIIIα, RegIIIβ, RegIIIγ or HIP/PAP were added in the indicated amounts in a final volume of 200 μL. Reactions were incubated for 2 hours at 4°C, then the peptidoglycan was pelleted at 6000*g* for 5 minutes and washed 2X in Standard Assay buffer. Bound protein was eluted by boiling the peptidoglycan in 2X SDS-PAGE buffer (10% glycerol, 5% β-mercaptoethanol, 2% SDS, 62.5 mM Tris-HCl, 0.003% bromophenol blue, pH 6.8) and resolved by SDS-PAGE through 15 % acrylamide gels.

For inhibition experiments, soluble peptidoglycan was generated by sonication of insoluble peptidoglycan as described [129]. Recombinant RegIIIβ, RegIIIγ or HIP/PAP was pre-incubated for 2 hours at 4°C with soluble peptidoglycan fragments from *Bacillus subtilis*. The entire pre-incubation mixture was then applied to pelleted, washed insoluble peptidoglycan. Reactions containing protein, insoluble, and soluble peptidoglycan were incubated for 30 minutes at 4°C, then the peptidoglycan was pelleted at 6000*g* for 5 minutes and washed 3X in Standard Assay buffer. Bound protein was eluted by boiling the peptidoglycan in 2X SDS-PAGE buffer (10% glycerol, 5% β-mercaptoethanol, 2% SDS, 62.5 mM Tris-HCl, 0.003% bromophenol blue, pH 6.8) and resolved by SDS-PAGE through 15 % acrylamide gels. Bound recombinant RegIIIβ, RegIIIγ or HIP/PAP were detected using Western blotting as described above.

T. Assays for microbicidal activity

Target microorganisms were grown to mid-log phase in brain-heart infusion (BHI) broth and resuspended in 25 mM MES pH 6, 25 mM NaCl. Initial bacterial concentrations ranged from 10^5 to 10^6 CFU/ml. After incubation for 2 hours at 37°C, viable bacteria were quantitated by dilution plating on BHI agar plates. Inhibition assays were performed by pre-incubating lectin for 10 minutes with chitooligosaccharides or sPGN prior to addition of bacteria.

U. *In vivo* bactericidal assays via enema infusion

Male germ-free NMRI mice were mono-colonized with *L. innocua* for >4 weeks before the beginning of the experiment. Pre-infusion fecal contents were collected for each mouse and comparable colonization levels were verified by dilution plating of the fecal contents. After administration of general anesthetic/sedative, recombinant RegIIIγ or buffer was infused anally into the intestinal tract using a syringe pump with a set flow rate of 1 mL/hour for 2 hours. The mice were sacrificed upon completion of infusion, the intestinal tracts were excised, and remaining colony forming units (CFUs) were quantified in triplicate by dilution plating for the ascending colon, descending colon, and rectum.

V. *In situ* hybridization vector construction

In order to generate *in situ* probes that would be specific for each *Reg*III transcript, we used ClustalW analysis to identify regions within the open reading frames

(ORFs) with little or no homology to each other. Particularly, a ~150 bp region near the 3′ end was only ~50% homologous according to sequence alignments. The following primers were designed to amplify each transcript specifically from this region using non-homologous nucleotides at the 3′ ends of the primers.

i. RegIIIα

A 156 bp amplicon was generated using RT-PCR from mouse small intestinal tissue using the specific primers 5′-ATTGCGAGGAAGCTTGAGTGGAGTAACTCC-3′ (forward primer) and 5′-CTATGGTGATCAGCTTGAACTTGCAGACAAATGGTAATGT-3′ (reverse primer). The forward primer contains a *Hind*III site (underlined) for cloning into pBluescript. As in the expression construct, the reverse primer incorporates an engineered *Bcl*I site (underlined), since the endogenous *Bam*HI site present in the RegIIIα sequence is contained within the region that was amplified. The resulting sequence spans the region from 439 bp – 595 bp from the start codon of the ORF. PCR products and vector were digested with *Hind*III and *Bcl*HI, gel-purified, and ligated into pBluescript that had been linearized by *Hind*III and *Bam*HI double digest. The recombinant plasmid (pBlue-RegIIIα) was sequenced to confirm the absence of mutations, and subsequently transformed into *E. coli* DH5α for plasmid preparation and storage.

ii. RegIIIβ

The specific primers 5′-ATTGCGAGGAAGCTTATGGGAATGGAGTAACAAT-3′ (forward primer) and 5′-CTATAGGGATCCGCAGACATAGGGCAACTTCA-3′ (reverse primer) were used to amplify a 157 bp sequence by RT-PCR from mouse small

intestinal tissue. The forward primer contains a *Hind*III site (underlined) for cloning into pBluescript. Likewise, the reverse primer incorporates an engineered *Bam*HI site (underlined). The resulting amplified sequence spans the region from 388 bp – 544 bp from the start codon of the ORF. PCR products and vector were digested with *Hind*III and *Bam*HI, gel-purified, and ligated. The recombinant plasmid (pBlue-RegIIIβ) was sequenced to confirm the absence of mutations, and subsequently transformed into *E. coli* DH5α for plasmid preparation and storage.

iii. RegIIIγ

A 152 bp amplicon was amplified by RT-PCR from mouse small intestinal tissue using the specific primers 5′-ATTGCGAGGAAGCTTTGGAGCAATGCT-3′ (forward primer) and 5′-CTATGGGGATCCGAATTTGCAGACATAGGGTAACTCTAA-3′ (reverse primer). The forward and reverse primers contain engineered *Hind*III (underlined) and *Bam*HI sites (underlined), respectively, for cloning into pBluescript. The resulting amplified sequence spans the region from 396 bp – 548 bp from the start codon of the ORF. PCR products and vector were digested with *Hind*III and *Bam*HI, gel-purified, and ligated. The recombinant plasmid (pBlue-RegIIIγ) was sequenced to confirm the absence of mutations, and subsequently transformed into *E. coli* DH5α for plasmid preparation and storage.

W. *In situ* hybridization probe preparation

Plasmid DNA constructed as outlined above was purified from 50 mL cultures using the GenElute HP Plasmid Maxiprep kit (Sigma) following manufacturer's

instruction. The nucleic acid concentration was calculated by measuring absorbances and purity was verified by elecrophoresis through a 1% agarose gel. Five micrograms of plasmid DNA containing each *Reg* sequence was digested with either *Bam*HI (sense) or *Hind*III (antisense), followed by phenol-chloroform extraction and analysis on a 1% agarose gel. Subsequently, 200 ng of linearized plasmid DNA from each digestion reaction was radiolabeled with ^{35}S-dUTP by *in vitro* transcription using the Ambion Maxiscript kit (Ambion, TX). The T3 enzyme was added to generate sense probes, and the T7 enzyme was added for the antisense probes. Unincorporated radio-nucleotides were separated by centrifugation through G-50 spin columns (Roche) according to manufacturer's instructions. Pre- and post-column samples were quantified by scintillation counting and analyzed for probe synthesis by resolving through denaturing 5% polyacrylamide gels.

X. Determination of dissociation constants (K_d)

Purified recombinant RegIIIγ or HIP/PAP (200 μg) was labeled with 1 mCi of ^{125}I (GE Healthcare) using Iodobeads (Pierce). Incorporated and free counts were separated by passage over D-Salt Excellulose desalting column (Pierce). For the binding assay, increasing amounts of labeled protein were added to 10 μg of *Bacillus subtilis* peptidoglycan (Sigma), and were incubated for 1 hour at 25^{o}C. The peptidoglycan was pelleted at 6000*g* for 5 minutes, the supernatant was removed and the unbound (free) counts were quantitated. The pellet was washed 2X in Wash Buffer (25 mM MES pH 6.0, 25 mM NaCl), and the bound counts were quantitated. Nonspecific binding was determined by performing the assay in the presence of 25 μg of unlabeled lectin. Specific

binding was the difference between total binding and nonspecific binding. The apparent dissociation constant, K_d, was determined by non-linear regression analysis using GraphPad Prism software.

Chapter Four:

PURIFICATION AND CHARACTERIZATION OF HUMAN AND MURINE REGIII PROTEINS

A. Introduction

In order to begin delineating the roles of the RegIII proteins in intestinal biology and human disease, we realized that an abundant source of purified recombinant mouse and human RegIII proteins would be needed for biochemical and functional characterization. Human HIP/PAP has been purified previously from the milk of transgenic mice engineered to express the protein in mammary gland [113]. Although this approach yielded quantities of protein sufficient for crystallographic analysis [130], the approach is technically challenging, time-consuming, and expensive. We therefore wished to develop a system for the rapid expression and purification of recombinant

RegIII proteins that is simple, high yield, and readily adaptable to other Reg family members.

In this chapter, we describe a new method for high level expression of mouse RegIIIγ, RegIIIβ, RegIIIα, and HIP/PAP using a bacterial expression system. Initial problems with low HIP/PAP expression were solved by introducing silent mutations, designed to relax local mRNA secondary structure, into the 5' end of the gene. Subsequently, we present details of a procedure for the refolding and purification of these four RegIII proteins from bacterial inclusion bodies. This simple protocol yields milligram quantities of each protein, and the following data represent the first example of Reg protein production from a bacterial expression system. Finally, we provide evidence suggesting that RegIIIγ, RegIIIβ, and HIP/PAP bind oligomerized but not monomeric mannose.

B. Results

Expression of recombinant mouse RegIIIγ, RegIIIβ, RegIIIα and human HIP/PAP in E. coli

The open reading frame corresponding to mature mouse RegIIIγ (lacking the N-terminal signal peptide) was ligated into the bacterial expression vector pET3a to yield pET3a-RegIIIγ. This construct allows IPTG inducible expression from a T7 promoter. I chose to express RegIIIγ in the *E. coli* expression strain BL21-CodonPlus(DE3)-RIL, which is genetically modified to express tRNAs corresponding to specific Arg, Ile, and

Leu codons that are normally rare in *E. coli*. To assess protein expression levels, IPTG was added to log-phase cultures, and pre- and post-induction cell lysates were analyzed by SDS-PAGE. Coomassie blue staining of gels revealed robust induction of RegIIIγ expression (Figure 4.1A).

Using standard cloning techniques, the coding sequences of the mature forms of RegIIIβ and RegIIIα were placed into the pET3a vector to generate the expression plasmids pET3a-RegIIIβ and pET3a-RegIIIα. In attempting to clone the RegIIIα open reading frame, we identified an endogenous *Bam*HI restriction site positioned ~360 bp into the sequence. Since no isoschizomers for *Bam*HI have been identified, we utilized a compatible ends approach, utilizing *Bcl*I as an enzyme that would generate nucleotide overhangs compatible with *Bam*HI. Sequencing of pET3a-RegIIIα indicated that the ligation was successful, and had generated a nucleotide sequence not recognized by either restriction enzyme, as predicted. After transformation of each plasmid into *E. coli* BL21-CodonPlus(DE3)-RIL, we assayed for recombinant protein induction upon the addition of IPTG. Log-phase cultures were treated with 0.4 mM IPTG, pre- and post-induction cell lysates were analyzed for each protein as done for RegIIIγ, and the results demonstrate robust induction from both constructs (Figure 4.1A).

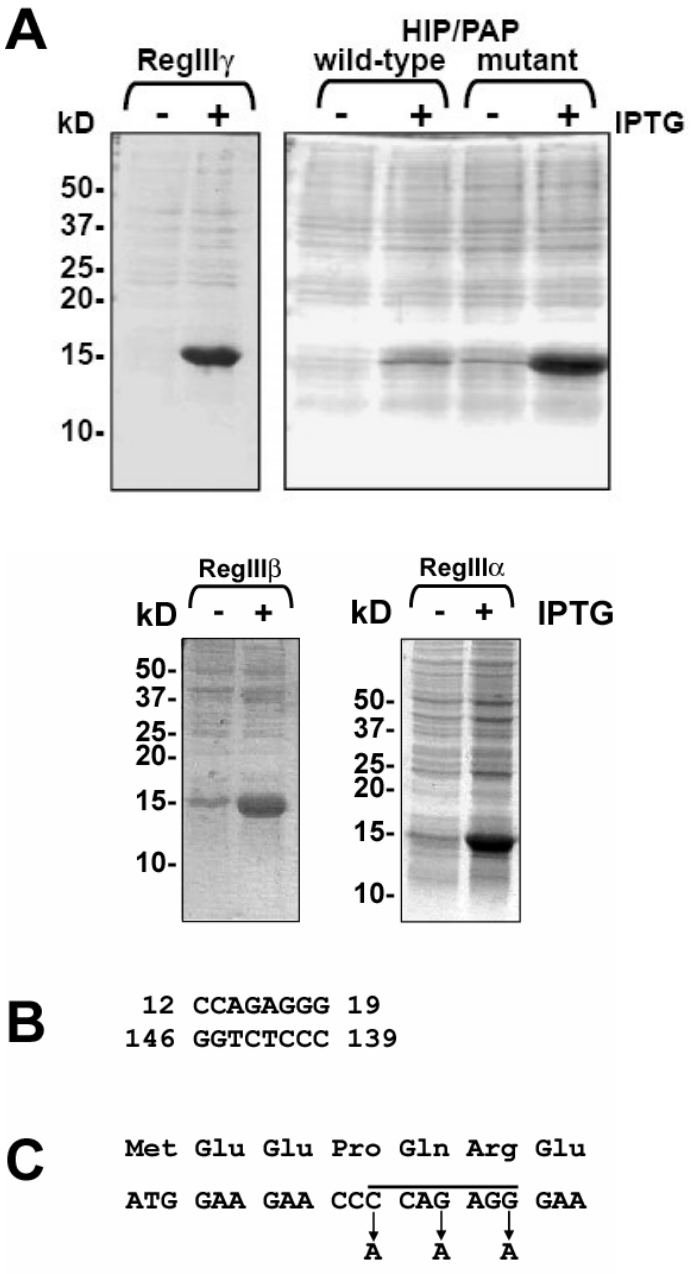

Figure 4.1:Expression of recombinant RegIIIα, RegIIIβ, RegIIIγ, and HIP/PAP in *E. coli*.

(A) Expression of RegIIIα, β, and γ constructs (pET3a-RegIIIα, β, or γ) were induced by addition of 0.4 mM IPTG. HIP/PAP expression constructs (wild-type = pET3a-HIP/PAP; mutant = pET3a-HIP/PAP*mut*) were induced by the addition of 1 mM IPTG. Total *E. coli* lysates from pre- and post-induction cultures were analyzed by electrophoresis through a 15% SDS-PAGE gel followed by Coomassie blue staining. **(B)** Predicted stem structure involving residues 12-19 of the HIP/PAP mRNA coding region. The stem was predicted by analyzing the mature HIP/PAP coding sequence using the web-based RNA secondary structure prediction algorithm at www.genebee.msu.su. **(C)** Positions of the silent mutations incorporated into the forward primer used to generate pET3a-HIP/PAP*mut*. The residues corresponding to the predicted stem are indicated by a line.

The mature human HIP/PAP open reading frame was also cloned into pET3a to yield pET3a-HIP/PAP. In contrast to the robust expression observed with the mouse RegIII proteins, we could detect very little HIP/PAP expression following the addition of IPTG to growing cultures (Figure 4.1A). These results are consistent with those of other investigators who have expressed HIP/PAP in *E. coli* [111]. Attempts to improve protein levels by altering induction conditions (IPTG concentration, induction time, induction temperature) were unsuccessful. We were also unsuccessful at improving protein expression by using an *E. coli* strain, BL21-CodonPlus(DE3)-RILP, that harbors an additional rare Pro tRNA. We hypothesized that translation initiation from the pET3a-HIP/PAP construct might be impaired, resulting in poor induction of recombinant HIP/PAP expression. Analysis of the mature HIP/PAP coding sequence (including the engineered start codon), using the RNA secondary structure prediction algorithm at www.genebee.msu.su, revealed the presence of a predicted stem involving nucleotides 12-19, with a free energy of -17.5 kcal/mol. By comparison, the mature RegIIIγ coding sequence contains a predicted stem encompassing nucleotides 8-11. This stem has a free energy of -7.1 kcal/mol, indicating a much less stable structure. Based on this analysis, we reasoned that the presence of a stable stem structure close to the 5' end of the HIP/PAP mRNA could interfere with ribosome binding and subsequent translation. To test this idea, we engineered 3 silent mutations into the forward primer used to amplify the HIP/PAP coding sequence (Figure 4.1B), and cloned the resulting amplicon into pET3a to generate pET3a-HIP/PAP*mut*. The mutations were designed to abolish the predicted stem by substituting A for G or C, thus reducing the stability of the base pair

interactions. Indeed, re-analysis of the altered sequence via the web-based algorithm confirmed the absence of the predicted stem. Consistent with this, the mutant HIP/PAP construct resulted in a remarkable increase in protein expression relative to the wild-type construct (Figure 4.1A). Thus, all further HIP/PAP expression/purification studies were done using recombinant protein derived from pET3a-HIP/PAP*mut*.

Purification of mouse proteins RegIIIγ, RegIIIβ, RegIIIα and human HIP/PAP

E. coli BL21 expression strains such as BL21-CodonPlus(DE3)-RIL and BL21-CodonPlus(DE3)-RILP lack the ability to generate disulfide bonds between cysteine residues in proteins. As the three murine RegIII proteins and HIP/PAP all contain 3 predicted disulfide bonds, we expected that the recombinant proteins would be misfolded and targeted to bacterial inclusion bodies. As shown in Figure 4.2, all four proteins were absent from the soluble fraction of *E. coli* lysates and were found in purified inclusion bodies. In each case, the recombinant proteins represented the vast majority of inclusion body protein.

The recombinant RegIII proteins from isolated inclusion bodies could be solubilized in 7 M guanidine hydrochloride under reducing conditions. However, each protein required refolding and reoxidation prior to purification. In the case of RegIIIγ, we refolded the protein using an approach identical to that used previously to obtain native Angiogenin-4 [131]. This procedure involves the dropwise addition of the solubilized inclusion body protein to a solution containing 0.5 M arginine and oxidized glutathione at pH 8. It has been proposed that arginine inhibits protein aggregation during refolding, while the oxidized glutathione promotes the formation of disulfide bonds [132]. This

buffer resulted in good recovery of RegIIIγ (27% of total inclusion body protein) following removal of arginine by dialysis (Table 4.1).

Attempts to refold RegIIIβ, RegIIIα, or HIP/PAP using the RegIIIγ refolding buffer resulted in a large amount of aggregation and the recovery of negligible soluble protein. Therefore, we screened a variety of refolding conditions for each protein, including cations (Ca^{2+}, Mg^{2+}), chelator (EDTA), salt (NaCl, KCl), pH, and additives such as L-arginine and sucrose. As detailed in Appendix 1, we obtained the best recoveries of soluble RegIIIβ by dilution into a solution of cations, guanidine, arginine, and a mixture of reduced and oxidized glutathione. Both RegIIIα and HIP/PAP recoveries were improved using a solution containing the additives listed for RegIIIβ as well as KCl and NaCl for both, and sucrose for HIP/PAP only. Following dialysis, the yields of soluble RegIIIβ, RegIIIα, and HIP/PAP were 29%, 11%, and 78% of total inclusion body protein, respectively (Table 4.1).

RegIIIγ, RegIIIβ, and HIP/PAP all have a basic predicted pI (8.5 for RegIIIγ, 7.5 for RegIIIβ, and 7.6 for HIP/PAP, as calculated by the algorithm found at the ExPASy website (http://au.expasy.org/tools/pi_tool.html)). Thus, we predicted that each of these recombinant proteins would bind to a cation exchange resin. Following refolding, the proteins were dialyzed into a low ionic strength buffer at pH 6 and were bound to SP-Sepharose. RegIIIγ was eluted in 0.6 M NaCl, and yielded ~16 mg/liter of culture. Likewise, RegIIIβ was eluted using 0.6 M NaCl, resulting in a yield of ~10 mg/liter of culture. Following elution with 0.6 M NaCl, HIP/PAP yielded ~24 mg/liter of culture (Table 4.1).

In contrast to its related family members, the predicted pI of RegIIIα is notably lower (6.2 vs 7.5-8.5). Therefore, we expected that it would be efficiently captured on Q-Sepharose at a basic pH. Upon refolding, RegIIIα was dialyzed against a low ionic strength buffer at pH 8.5 and bound to Q-Sepharose. We used 0.6 M NaCl in 25 mM Tris pH 7.5 to completely elute captured RegIIIα, which generated yields of ~5 mg/liter of culture (Table 4.1).

Table 4.1: Yields of recombinant mouse RegIII proteins from overexpressing *Escherichia coli*

Lectin	Solubilized inclusion bodies		Ion exchange eluate	
	Total protein (*mg*)	Protein yield (%)	Total protein (*mg*)	Protein yield (%)
Gamma	34.0	25.0	7.9	6.0
HIP/PAP	28.0	16.0	12	6.8
Beta	13.9	8.8	4.0	2.5
Alpha	6.3	7.2	0.7	0.8

Results are derived from 500 mL cultures of *E. coli* expressing the recombinant proteins. Total protein was estimated by the method of Bradford [121].

Characterization of Recombinant Proteins

The purities of the recombinant proteins were assessed initially by SDS-PAGE. Following elution from SP-Sepharose, all four proteins migrated as single species (Figure 4.2). To further confirm their identities, RegIIIγ and HIP/PAP were analyzed by MALDI-TOF mass spectrometry. RegIIIγ yielded a single peak corresponding to a molecular mass of 16.6 kD, in good agreement with the predicted molecular mass of the mature protein (16.5 kD). Likewise, analysis of purified HIP/PAP demonstrated a single peak indicating

a molecular mass of 16.4 kD, in agreement with its predicted molecular mass (16.7 kD). Furthermore, Edman N-terminal sequencing of recombinant RegIIIγ yielded MEVAK for RegIIIγ, the expected amino terminus of the mature [Met^{-1}] protein. Similar analysis of purified RegIIIβ resulted in the sequence, MEDSL, the expected amino terminal sequence of the mature [Met^{-1}] protein. Finally, analysis of the HIP/PAP N-terminus yielded the sequence MEEPQ, which corresponds to the predicted N-terminus of the recombinant mature [Met^{-1}] protein. At this stage in our investigations, the N-terminal sequence of RegIIIα has not been confirmed.

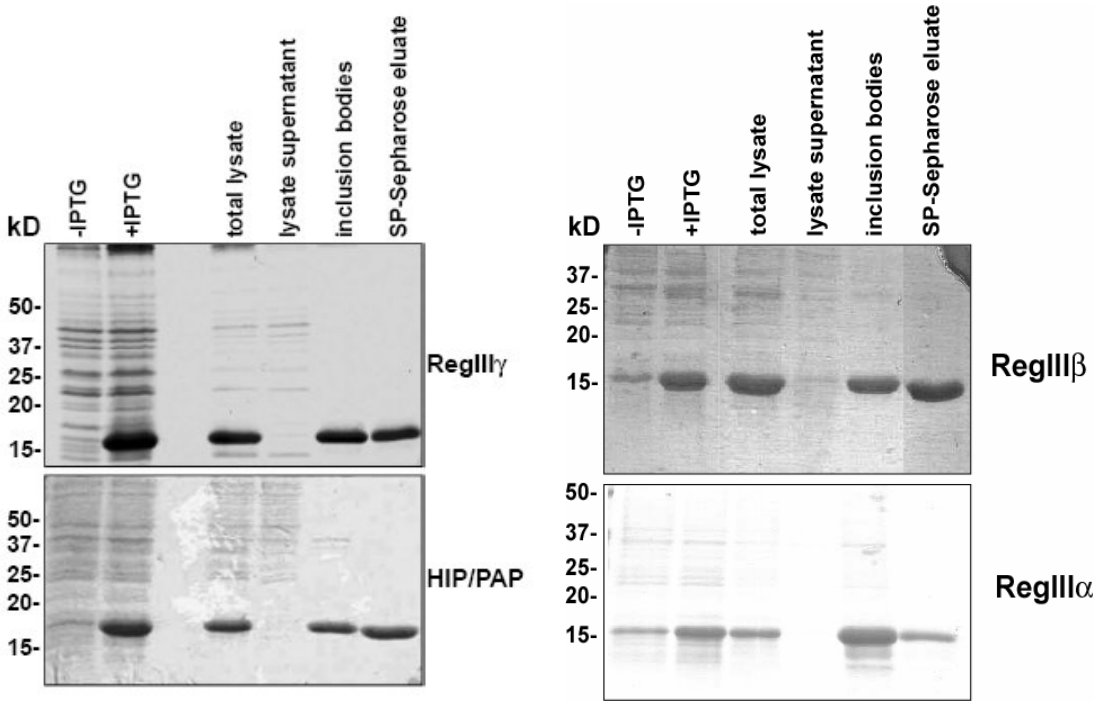

Figure 4.2: SDS-PAGE analysis of samples taken during the purification of the RegIII proteins.

E. coli cells overexpressing recombinant RegIII proteins were collected before and after induction with IPTG (+ and – IPTG). The lanes containing total lysate and lysate supernatant samples were loaded with 10 mg total protein. Inclusion body and SP-Sepharose column eluate sample lanes contain 5 mg of protein. Proteins were resolved on a 15% polyacrylamide gel and stained with Coomassie brilliant blue.

Size-exclusion chromatography was performed to determine whether the recombinant proteins formed oligomers in solution. Chromatography through Sephacryl S-100 revealed that RegIIIγ elutes at a molecular mass of 20 kD (Figure 4.3). While this corresponds to a molecular weight slightly greater than the predicted molecular mass of monomeric protein (16 kD), dimer would likely elute at a molecular weight in excess of 32 kD. Our results therefore suggest that the majority of RegIIIγ remains a monomer in solution. Similarly, HIP/PAP exhibits a major elution peak at 26 kD, suggesting that it also is predominantly a monomer in solution. Interestingly, both HIP/PAP and RegIIIγ displayed small elution peaks at a volume corresponding to ~38 kD, which could represent dimeric complexes. Interestingly, preliminary analysis indicates that RegIIIβ elutes at a molecular mass of 38 kD, suggesting that the majority of this protein forms dimers in solution (Figure 4.3). These data represent the first to describe a potential for multimerization of RegIII proteins. As described in later chapters, this ability to self-complex may facilitate execution the RegIII functions of ligand binding and bactericidal activity.

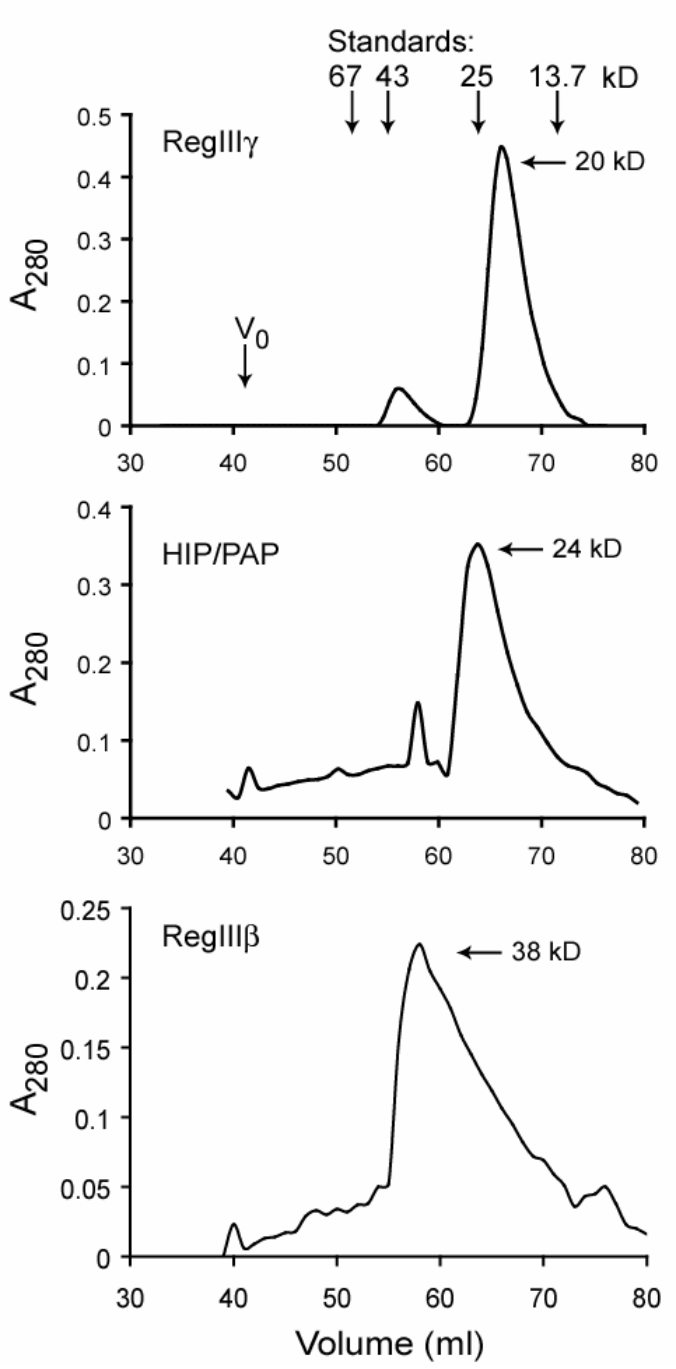

Figure 4.3: Size exclusion chromatography of RegIIIγ, HIP/PAP, and RegIIIβ.

Two milligram samples of each purified recombinant protein were applied to a Sephacryl S-100 column and eluted with detection at 280 nm. Positions of void volume (V_0) and standards are indicated: albumin (67 kD), ovalbumin (43 kD), chymotrypsinogen (25 kD), and ribonuclease (13.7 kD).

Previous crystallographic analysis of HIP/PAP has elucidated a secondary structure that is composed of 9 β-strands and 2 α-helices [130]. This structure is very similar to those of virtually all other C-type lectin CRDs characterized [62]. We performed circular dichroism spectroscopy to characterize the secondary structures of purified recombinant RegIIIγ and HIP/PAP. The results for both proteins reveal maximal negative ellipticity in the range of 205-215 nm (Figure 4.4), and the spectra were similar overall to those derived from other C-type lectin family members, including langerin [133], surfactant protein A [134], and mannose binding lectin [135]. Indeed, analysis of the spectra by Dichroweb [122] using the K2D and Selcon 3 analysis programs [123, 124] indicate that RegIIIγ and HIP/PAP are both predominantly comprised of β-sheet structure, while α-helix structure is not as prevalent (Figure 4.4). Our findings are thus consistent with the secondary structures revealed by the HIP/PAP crystal structure as well as those other C-type lectin CRDs. These results indicate that purified recombinant RegIIIγ and HIP/PAP have acquired their expected secondary structures and are thus correctly refolded.

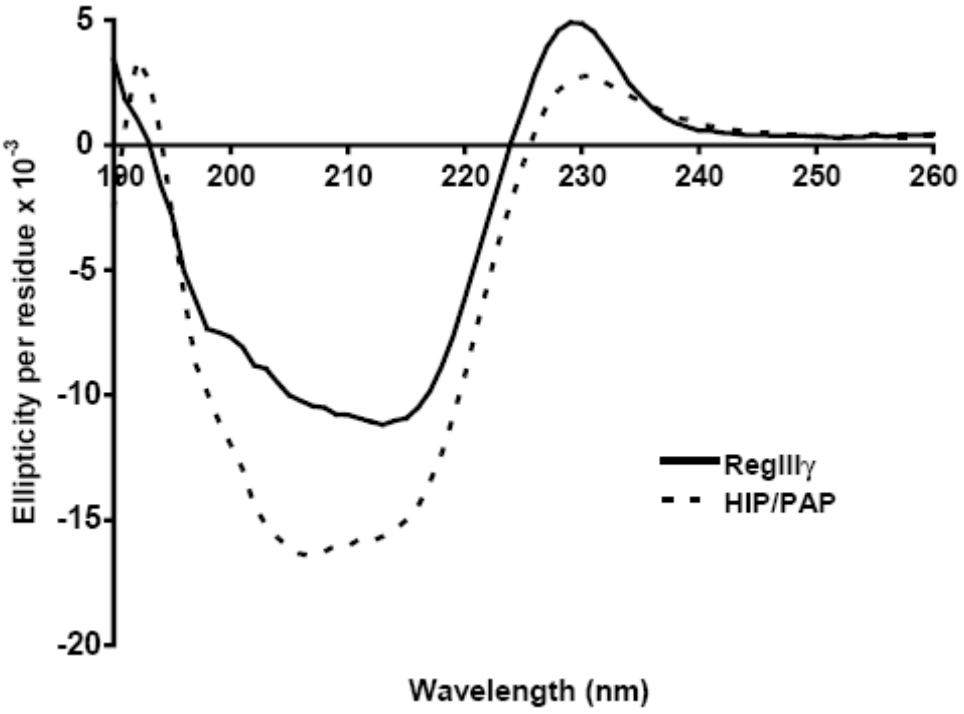

Figure 4.4: Circular dichroism spectra of RegIIIγ and HIP/PAP.

Results show that RegIIIγ and HIP/PAP are composed predominantly of β-sheet structure, which is consistent with correct refolding.

Recombinant mammalian proteins expressed in *E. coli* generally lack the post-translational modifications present on their endogenous counterparts. Such differences can pose difficulties for functional and biochemical analysis of recombinant proteins. Previous studies have demonstrated the existence of an O-glycosylated form of another Reg family member, human RegIα [64]. Analysis of the RegIIIγ primary sequence indicates that the protein does not harbor a consensus sequence for N-glycosylation (Asn-Xaa-Ser/Thr). However, there is at least one potential O-glycosylation site as determined by the NetOGlyc algorithm at Expasy (http://www.cbs.dtu.dk/services/NetOGlyc/) [136]. To determine whether the endogenous protein is postranslationally modified by glycosylation or another modification, we compared the molecular weight of the

endogenous proteins with that of recombinant RegIIIγ. Western blot analysis using a polyclonal antibody raised against purified RegIIIγ, but that recognizes all three murine RegIIIs, revealed that the recombinant protein migrates at the same molecular weight as protein from intestinal tissue homogenates (Figure 4.5). The lower band detected in the endogenous sample is identical in molecular weight to a cleaved form that we observe following exposure of recombinant RegIIIγ to exogenous proteases, suggesting that RegIIIγ is processed *in vivo* by endogenous intestinal proteases. Although the polyclonal antibody recognizes all three mouse proteins, the recombinant forms of RegIIIα, RegIIIβ, and RegIIIγ all migrate at the same molecular weight when analyzed by SDS-PAGE (Figure 4.2). However, we cannot discount the possibility of an alternatively spliced form of the RegIII proteins. In summary, these results remain consistent with the conclusion that RegIIIγ is not modified by glycosylation or other posttranslational modifications.

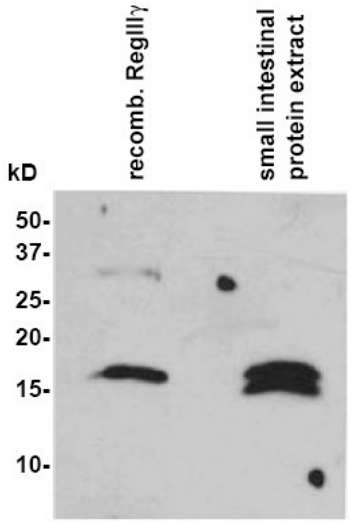

Figure 4.5: Western blot comparison of recombinant and endogenous mouse RegIIIγ.

Total mouse small intestinal protein was prepared as described in Materials and methods. Five nanograms of purified recombinant RegIIIγ and 20 μg of total intestinal protein were separated by SDS-PAGE and transferred to PVDF. The blot was probed with antiserum raised against recombinant RegIIIγ. *Courtesy of Cecilia Whitham*.

Characterization of the carbohydrate binding activity of the RegIII proteins

The primary amino acid sequences of the four RegIII proteins are composed almost entirely of a conserved C-type lectin carbohydrate recognition domain (CRD). According to the classification scheme proposed by Drickamer and Fadden, they are members of the type VII C-type lectin subfamily [63]. Although members of the other subfamilies have well-characterized carbohydrate ligands, only preliminary evidence exists for the ligands of type VII lectins, including RegIIIα, RegIIIβ, RegIIIγ and HIP/PAP [110]. To gain initial insight into the carbohydrate binding specificity of these proteins, we covalently coupled various mono- and polysaccharides to Sepharose 6B resin and assayed for binding of purified RegIII proteins. None of the immobilized monosaccharides tested (glucose, galactose, N-acetylglucosamine, N-acetylgalactosamine, mannose, fucose) supported binding of any of the lectins (data not shown). In addition, we did not observe binding of any of the proteins to the disaccharide lactose (data not shown), in contrast to a previous report demonstrating that lactose is a ligand for GST-tagged HIP/PAP [110]. However, RegIIIβ, RegIIIγ and HIP/PAP bound to immobilized mannan (Figure 4.6 and Figure 6.5), a polysaccharide composed of polymerized mannose. By contrast, none of the proteins bound to dextran, a polysaccharide composed of α1,6- and α1,3-linked glucose, suggesting that these lectins are specific for mannose polysaccharides. Moreover, we did not detect binding of either RegIIIγ or HIP/PAP to mannose (Figure 4.6), indicating that these two proteins bind polymerized but not monomeric mannose. Intriguingly, we did not observe any binding by RegIIIα to any mono- or polysaccharide tested, notably mannan and chitin (data not shown). This suggests two possibilities, either

a unique ligand for this RegIII family member exists or this murine protein is nonfunctional, both of which warrant further investigation.

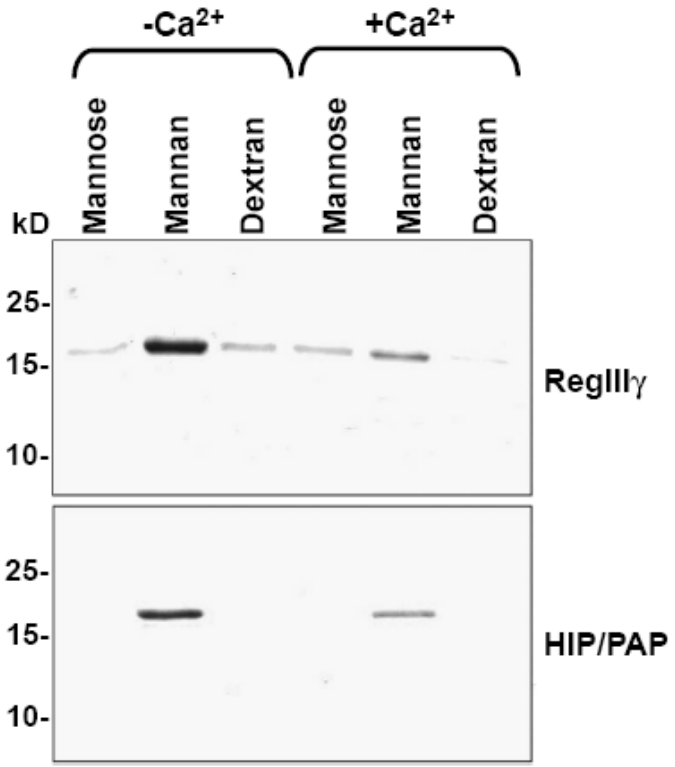

Figure 4.6: RegIIIγ and HIP/PAP bind to immobilized polysaccharides.

50 μg of either RegIIIγ or HIP/PAP was bound to 25 μL of immobilized mono- or polysaccharide for 2 hours at 4°C. After washing, bound proteins were released by boiling the Sepharose beads in 2X SDS-PAGE buffer followed by electrophoresis through 15% polyacrylamide gels and Coomassie blue staining.

The C-type lectin family includes members whose ligand binding is calcium-dependent. However, previous studies have shown that at least one other Reg family member, RegIα, does not bind Ca^{2+} [64]. Furthermore, crystallographic analysis of RegIα revealed significant alterations in the polypeptide loop that binds Ca^{2+} in other C-type lectins [64]. We therefore asked whether RegIIIβ, RegIIIγ or HIP/PAP binding to mannan is influenced by Ca^{2+}. Our results revealed that 10 mM $CaCl_2$ reduces binding of

both RegIIIγ and HIP/PAP to mannan (Figure 4.6). These results thus suggest that carbohydrate ligand binding to RegIIIγ and HIP/PAP is indeed modulated by Ca^{2+}, but in a way that is distinct from other C-type lectins.

We next wished to determine whether the entire purified RegIIIβ, RegIIIγ or HIP/PAP protein populations are capable of binding to mannan. Therefore, we applied purified recombinant protein to a column of mannose- or mannan-conjugated Sepharose beads. As expected, each protein passed through the mannose-Sepharose column (Figure 4.7). In contrast, RegIIIγ and HIP/PAP bound quantitatively to the mannan-Sepharose column, and RegIIIβ was retarded, eluting in the wash fractions (Figure 4.7C). These date indicate that each purified protein population is refolded to a functionally active state in its entirety. Furthermore, we eluted both RegIIIγ and HIP/PAP with $CaCl_2$, confirming that both lectin-carbohydrate interactions are inhibited by Ca^{2+} (Figure 4.7A and B).

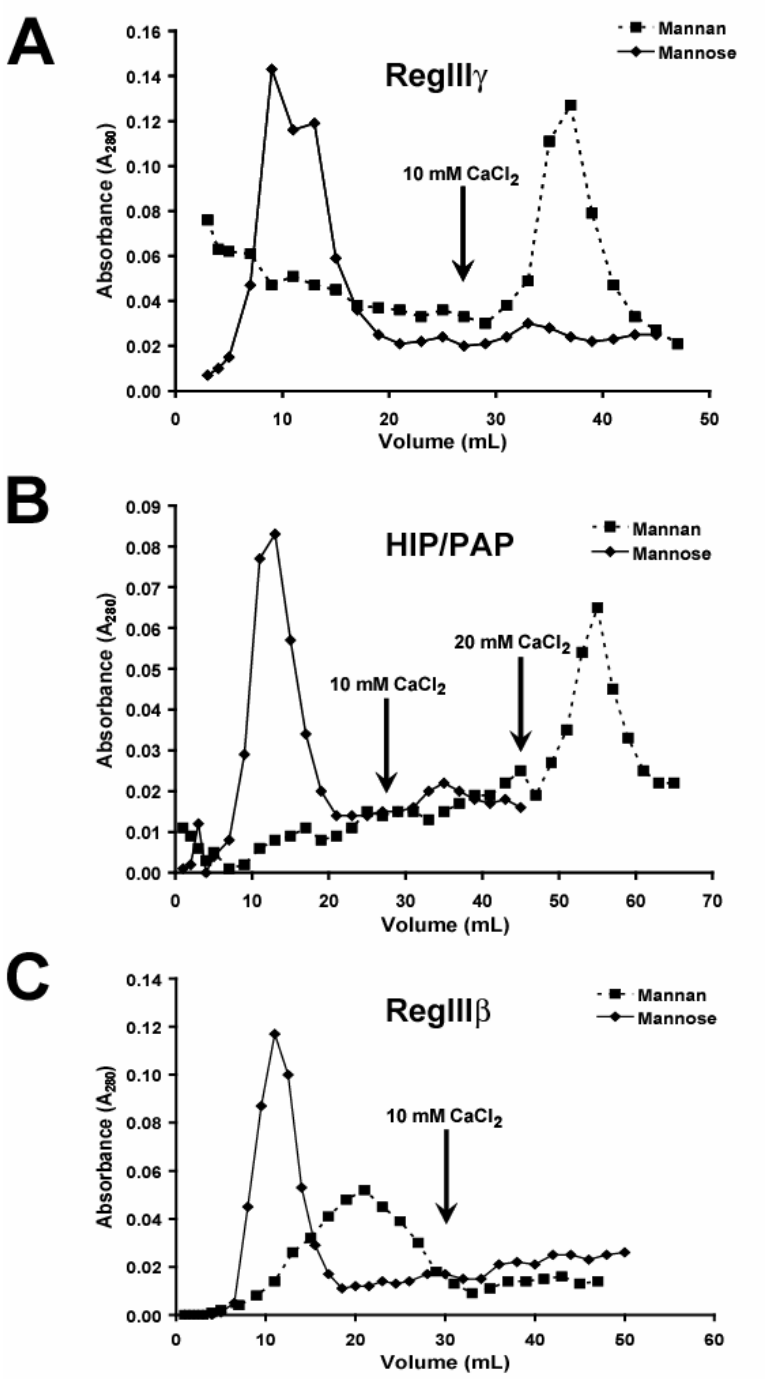

Figure 4.7: Chromatography of the RegIII proteins on mannose- and mannan-Sepharose.

(A) 0.25 mg of RegIIIγ was applied to a 10 mL mannose- or mannan-Sepharose column. Two milliliter fractions were collected and protein was detected by spectrophotometry at 280 nm. Protein bound to mannan-Sepharose was eluted with 10 or 20 mM $CaCl_2$ as indicated. **(B and C)** HIP/PAP and RegIIIβ, respectively, were analyzed according to the procedure described for RegIIIγ above.

C. Discussion

A number of prior studies have suggested that the Reg protein family plays critical roles in intestinal biology. Several members of the Reg family, including those described in this chapter, are highly expressed in the intestine as compared to other tissues [104, 108, 111]. Additionally, dysregulated expression of these proteins, particularly RegIIIγ and HIP/PAP, has been associated with intestinal diseases such as inflammatory bowel disease [112, 137, 138]. Based on their initial identification in regenerating tissue, certain family members have been ascribed a role in tissue repair [109]. However, prior to these investigations, the exact functions of Reg family proteins were still poorly understood. Moreover, although all are members of the C-type lectin family, their carbohydrate ligands remained unidentified. As a first step in elucidating the biological functions and ligands of the RegIII family, we developed a simple protocol for purification of four RegIII proteins using a bacterial overexpression system.

Two distinct heterologous expression systems have been used previously to obtain pure Reg proteins. HIP/PAP has been isolated from lactating transgenic mice expressing the recombinant protein under the control of a mammary gland-specific promoter [113]. A purification protocol involving a GST-tagged form of HIP/PAP has also been published, although this method yielded only microgram quantities of protein from large volume bacterial cultures [110]. Additionally, RegIV has been purified following overexpression in a high density fermentation system using the yeast *Pichia pastoris* [139]. However, these methods either require specialized equipment and techniques that can be expensive and time consuming or do not produce quantities of protein sufficient for biochemical analyses. Therefore, in this chapter, we detail a relatively simple method

for obtaining milligram quantities of four RegIII proteins without the use of protein tags such as GST or Histidine. By overexpressing the proteins in *E. coli*, isolating inclusion bodies, and refolding the proteins to their native conformations, we obtained milligram quantities of RegIIIα, RegIIIβ, RegIIIγ and HIP/PAP. This method is likely to be easily adaptable to other Reg family members by screening for refolding conditions specific to each protein.

Initial attempts to induce production of HIP/PAP in *E. coli* resulted in very low expression levels, in contrast to the robust expression seen with the three murine RegIII proteins. In order to improve HIP/PAP expression, we incorporated silent mutations designed to alleviate predicted RNA secondary structures at the ribosome binding site. This resulted in greatly improved HIP/PAP expression, suggesting that a similar approach could be employed to increase expression of other Reg family members that exhibit low expression levels in *E. coli*. In fact, an analogous approach has been used previously to amplify *E. coli* expression of porcine liver cytochrome P-450 reductase [140].

Although refolding denatured proteins can result in improper conformations, three lines of evidence indicate that our procedures yield correctly refolded proteins. First, unfolded proteins tend to aggregate. However, our size exclusion chromatography experiments revealed that RegIIIβ, RegIIIγ, and HIP/PAP eluted as a single peak, either monomeric or dimeric, suggesting that they were not entirely misfolded. Second, circular dichroism analysis shows that the refolded proteins have secondary structure, and that this structure is predominantly β-sheet which is typical of C-type lectins. Third, the fact that RegIIIβ, RegIIIγ and HIP/PAP bind specifically to mannan-Sepharose indicates that they are refolded to a functionally active state.

The carbohydrate binding data presented here are the first to suggest a potential ligand for Reg family members. To date, investigations have not identified ligands for any of the Group VII lectins. Our results suggest that the binding specificities of RegIIIβ, RegIIIγ and HIP/PAP are similar to that of mannose binding lectin, which binds to mannose residues on bacterial surfaces and initiates recruitment of complement components that carry out microbial killing [68]. However, there are some key differences. Although these three proteins interact with mannan, they do not bind monomeric mannose, suggesting a requirement for a highly oligomerized ligand. Second, whereas mannose binding lectin requires Ca^{2+} for ligand binding [69], the binding of RegIIIβ, RegIIIγ and HIP/PAP to mannan is Ca^{2+}-independent. This observation is consistent with the fact that other Reg family members do not bind Ca^{2+}, and have been shown to lack the Ca^{2+} binding site seen in other C-type lectins [64].

While RegIIIβ, RegIIIγ, and HIP/PAP all have similar predicted pI's and exhibit similar refolding and binding activities, RegIIIα was distinct from its family members. Interestingly, ClustalW analysis predicts HIP/PAP to be more closely related to the mouse proteins, RegIIIβ and RegIIIγ, than RegIIIα is to any of the other three proteins (Figure 4.8). In keeping with this primary sequence divergence, the cloning and purification of RegIIIα required unique procedures. Initial cloning of this protein had to be modified to account for the presence of *Bam*HI and *Bgl*II restriction sites within the coding sequence. Once refolded, the low predicted pI dictated purification using an anion exchange resin. However, most intriguingly, purified RegIIIα did not bind to any of the ligands tested, including those bound by its family members. At this stage of our investigations, it is possible that either the protein is not functionally refolded or that we

simply have not identified this protein's biological ligand. As discussed in Chapter 7, this divergence in RegIII carbohydrate ligands could prove useful in future structure-function analyses of these proteins.

Previous studies by other investigators have suggested that RegIIIγ and HIP/PAP are secreted from gut epithelial cells [111, 141]. If so, then these proteins are likely targeted to the intestinal lumen, which is inhabited by large populations of resident microbes. The binding to mannan suggests that, like mannose binding lectin, RegIIIβ, RegIIIγ and HIP/PAP could bind to surface carbohydrate moieties on microbes. The ability to produce milligram quantities of these proteins in an *E. coli* overexpression system allowed this hypothesis to be tested and facilitated the further analyses of the ligands and functions of the RegIII family that are detailed in the following chapters.

Chapter Five:

REGIIIγ AND HIP/PAP ARE BACTERICIDAL LECTINS

A. Introduction

The mammalian intestine is home to a vast consortium of symbiotic bacteria. Members of this complex microflora provide metabolic assistance for their hosts by breaking down dietary substances, such as plant polysaccharides, that would be otherwise indigestible [142]. In addition to their essential nutritional contributions, indigenous gut microbes facilitate proper development of the intestinal vasculature and the resident adaptive immune system [13, 16, 23], in return inhabiting a protected, nutrient rich environment. However, maintaining the mutually beneficial nature of this relationship

requires strict sequestration of resident bacteria in the intestinal lumen, since microbial incursions across epithelia can elicit inflammation and sepsis.

Epithelial antimicrobial proteins are evolutionarily ancient innate immune effectors. As key elements of the front-line mucosal defense, they likely play an important role in maintaining mutually-beneficial host-microbial relationships by restricting contact between resident microbes and mucosal surfaces. This idea is underscored by the fact that deficiencies in antimicrobial peptide expression are associated with inflammatory bowel disease [44, 143], a chronic inflammatory disorder thought to be triggered by resident gut microbes. Though cationic antimicrobial peptides such as defensins are well characterized, the full repertoire of gut antimicrobial mechanisms remains undefined. Many antimicrobial proteins, including defensins, are contained within Paneth cells, key effectors of small intestinal antimicrobial defense. These specialized epithelial cells are located at the crypt base and harbor abundant cytoplasmic secretory granules, which eject their contents into the intestinal lumen. To gain new insights into how intestinal surfaces cope with microbial challenges, we used DNA microarrays to identify Paneth cell antimicrobial factors whose expression is altered by bacteria.

As elaborated in the Synopsis, this initial microarray analysis comparing Paneth cells from germ-free and conventionalized mice identified two transcripts of the *RegIII* family that were significantly upregulated in the presence of an intestinal microflora. This protein family constitutes a unique group of mammalian C-type lectins, in that it is comprised of a diverse group of secreted proteins that are composed primarily of a single C-type lectin carbohydrate recognition domain (CRD). Several RegIII family members

are expressed predominantly in small intestine, including mouse RegIIIγ (Figure 6.1) and human HIP/PAP [111, 112]. Inflammatory stimuli such as bacteria [112, 144] or mucosal damage [145] increase gastrointestinal expression of mouse RegIIIγ. Furthermore, HIP/PAP expression increases in the mucosa of patients with inflammatory bowel disease [112, 138], a disorder characterized by increased mucosal adherence of resident bacteria and chronic intestinal inflammation [146]. Although mitogenic functions have been suggested for Reg proteins in other tissues [119], the biological functions of intestinal RegIII proteins and their role in IBD have remained poorly defined.

Using the ready supply of purified RegIII proteins detailed in the previous chapter and other tools of molecular biology, we now provide evidence that resident gut bacteria drive intestinal epithelial expression of RegIIIγ, that this protein binds peptidoglycan, and that it has direct antimicrobial activity. These functions are conserved in the human form, HIP/PAP, suggesting a primitive mechanism of lectin-mediated innate immunity that is directly responsive to intestinal colonization.

B. Results

RegIII *genes are induced by colonization with intestinal microbes*

One of the most prominent responses uncovered by our microarray analysis was a 31-fold increase in the abundance of *RegIIIγ* transcript in Paneth cells from conventionalized as compared to germ-free mice. Increased expression of *RegIIIγ* mRNA was confirmed by real-time PCR (Figure 5.1A), and correlated with increased protein expression as determined by Western blot analysis of total small intestinal protein (Figure

5.1B). Similar to the lower molecular weight species described in Chapter 4, the lower band observed in the protein sample from conventionalized mice is likely the result of proteolytic cleavage at a trypsin-like site near the N-terminus similar to that described previously for the Reg family [147].

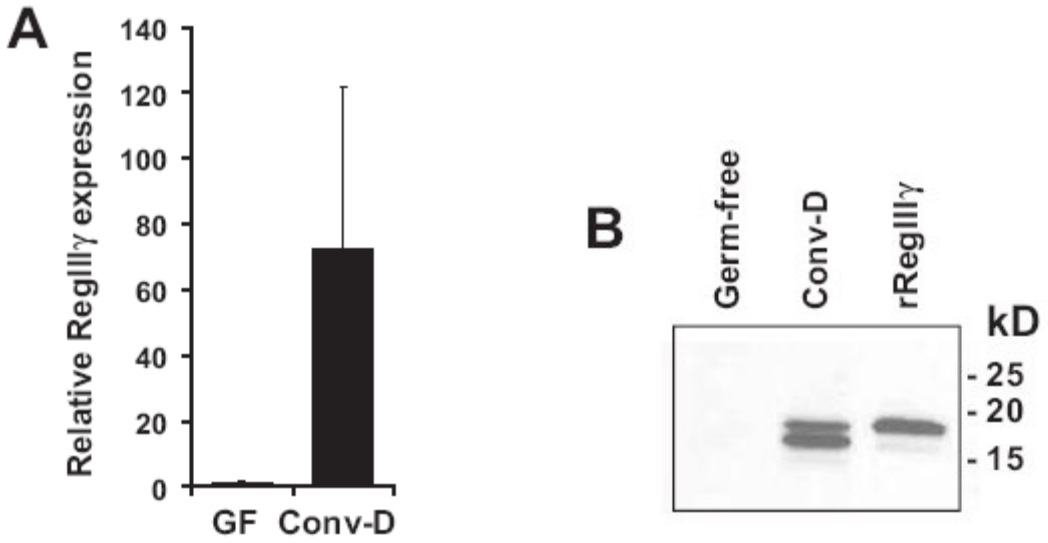

Figure 5.1: RegIIIγ is induced by resident intestinal microbes.

(A) *RegIIIγ* mRNA expression was analyzed in Paneth cells that were harvested by laser-capture microdissection from germ-free and conventionalized small intestines. Q-PCR analysis was performed on RNAs from microdissected Paneth cells from 3 mice per group. Values were normalized to GAPDH expression and mean ± s.d. is plotted. Results are expressed relative to one of the germ-free samples. GF = germ-free; Conv-D = conventionalized. **(B)** RegIIIγ protein expression in the small intestine was detected by immunoblot with anti-RegIIIγ antiserum (see Chapter 3). Results are representative of two independent experiments. rRegIIIγ = recombinant RegIIIγ. *Courtesy of Cecilia Whitham.*

The RegIII proteins are secreted into the intestinal lumen

All RegIII proteins described to date contain a 12 amino acid N-terminal leader peptide that targets these proteins for secretion through the canonical secretory apparatus. Before secretion from the cell, this signal sequence is removed to produce the mature form of these proteins. The mature ~16kD protein consists solely of a predicted CRD

(Figure 5.2), which includes three disulfide bonds to provide additional stability for folding outside of the cellular cytoplasm. Because of the existence of the leader peptide predicting secretion from the cell, its small size, and the presence of disulfide bonds similar to other known antimicrobial proteins, we hypothesized that RegIIIγ could be present in Paneth cell secretory granules and therefore secreted into the intestinal lumen.

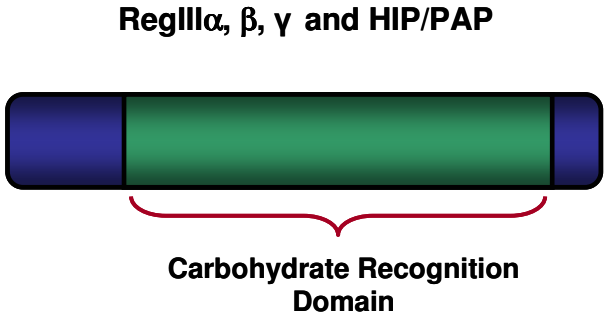

Figure 5.2: Schematic of RegIII protein primary amino acid domains.

All described RegIII proteins encode an N-terminal secretion sequence, which is cleaved to release the mature form. The majority of each protein contains a predicted carbohydrate recognition domain, which classifies the RegIII proteins as Type VII (secreted CRD) lectins according to the Drickamer classification [63].

To test this hypothesis, we employed immunogold electron microscopy using the polyclonal antiserum raised against recombinant RegIIIγ (Chapter 4). Imaging of embedded Paneth cells revealed that RegIIIγ is present exclusively in their secretory granules (Figure 5.3). Because RegIIIγ and RegIIIβ (Pap) are quite homologous, our polyclonal antibody cross-reacts with both RegIII family members. Therefore, we cannot discount the possibility of localizing RegIIIβ to Paneth cell granules in addition to RegIIIγ. Regardless, these findings correlate with our real-time analysis of Paneth cell transcripts and the mRNA expression patterns observed using *in situ* hybridization, as described in Chapter 6 (Figure 6.2). Previous research has confirmed that granule contents are released at the apical cell surface [148], indicating that RegIIIγ is targeted to

the gut lumen which harbors large resident bacterial populations. Additionally, immunoblotting of isolated intestinal contents confirmed the presence of RegIIIγ in the lumen (Figure 5.3B). Since other members of the C-type lectin family, such as the mannose binding lectin (MBL), bind to microbial surface carbohydrates and trigger innate immune responses [68], we hypothesized that secreted RegIIIγ might similarly bind intestinal bacteria.

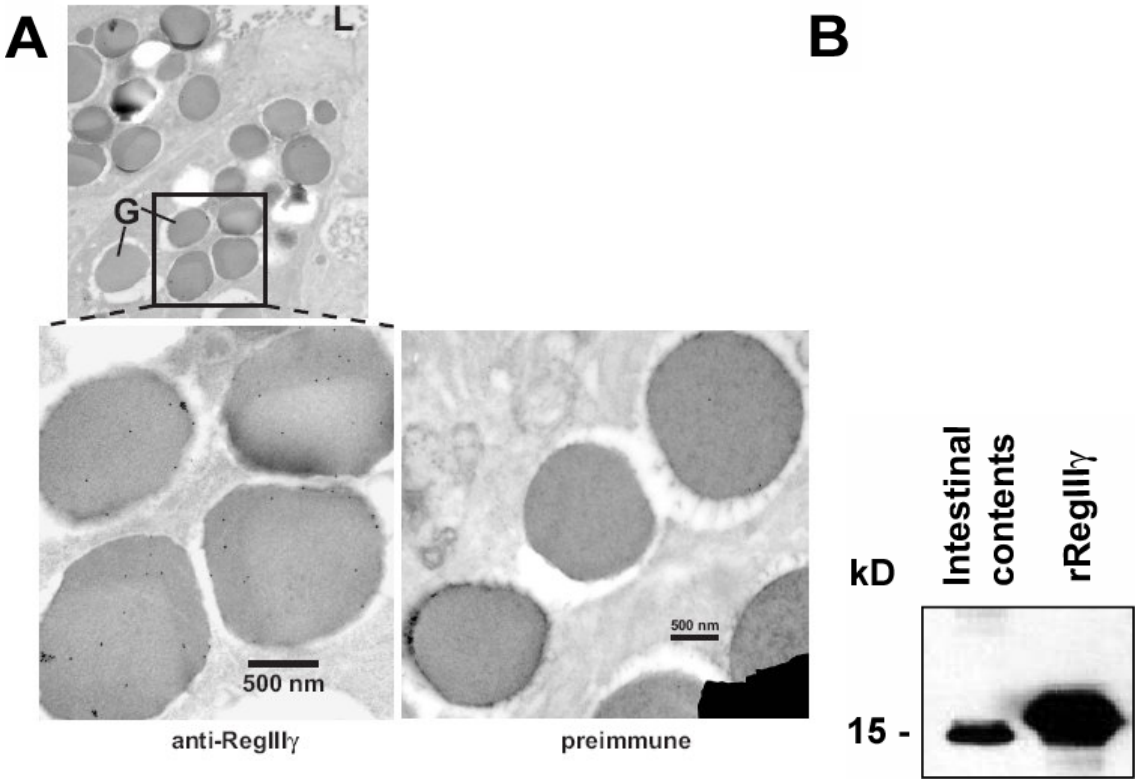

Figure 5.3: RegIIIγ localizes to Paneth cell secretory granules and is secreted into the intestinal lumen.

(A) Sections were cut from blocks as described in Materials and Methods, incubated with anti-RegIIIγ and gold-labeled goat anti-rabbit IgG, and visualized by electron microscopy. G, granule; L, gut lumen. Bar = 500 nm. Controls were incubated with pre-immune serum and gold-labeled goat anti-rabbit IgG. Representative photos of Paneth cell granules are shown above. **(B)** RegIIIγ can be detected in the intestinal lumen of conventional mice. Intestinal contents were isolated and fixed in 2% formaldehyde as described in Materials and methods. 1.5 μg of total intestinal contents or 0.6 ng of recombinant RegIIIγ were detected by Western blot using anti-RegIIIγ antiserum. Results are representative of two separate mice and the experiment was repeated twice.

RegIIIγ binds intestinal bacteria and is specific for Gram positive organisms

To test the idea that secreted RegIIIγ could bind to members of the intestinal microflora, we conjugated recombinant RegIIIγ to a fluorophore, which would enable us to look for binding to a mixed microbial population harvested from the small intestines of conventionally-raised mice. Flow cytometry revealed that RegIIIγ bound to a sub-population of intestinal bacteria (Figure 5.4A).

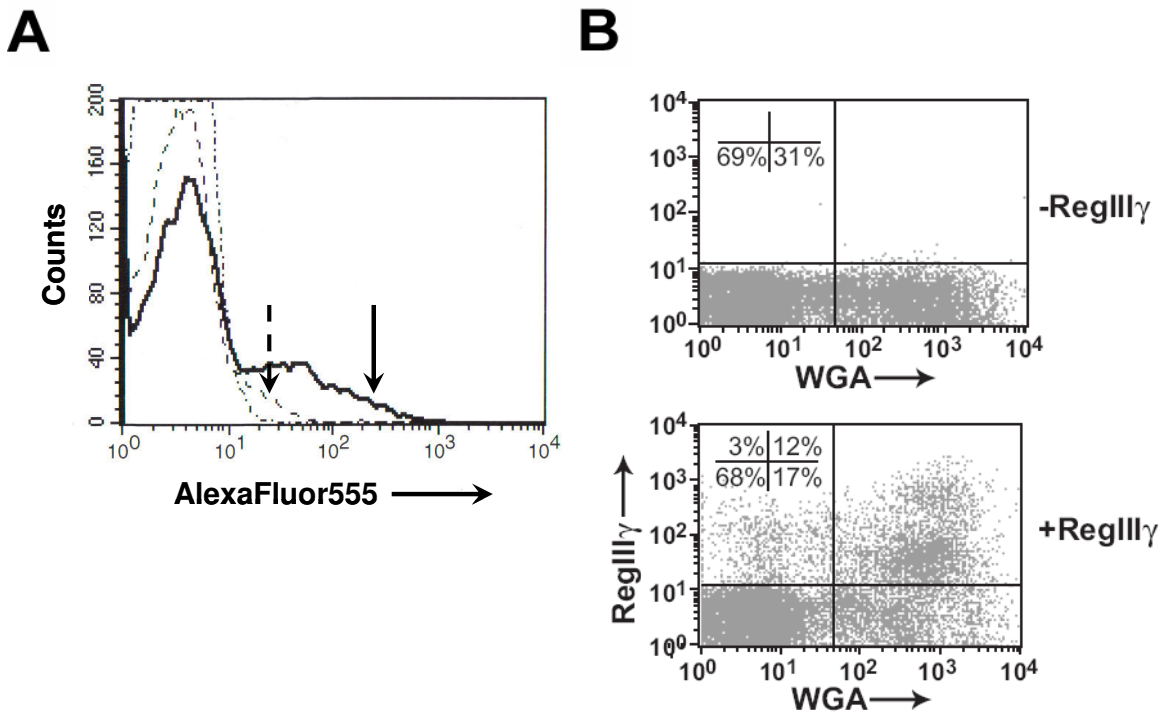

Figure 5.4: RegIIIγ binds to Gram positive intestinal bacteria.

(A) Flow cytometry reveals binding of AlexaFluor555-conjugated RegIIIγ to intestinal bacteria recovered from conventional mouse small intestine. Results are representative of independent experiments with 3 mice. (Dashed arrow = BSA-AlexaFluor555; Solid arrow = RegIIIγ-AlexaFluor555) **(B)** Dual-color flow cytometry analysis with WGA-AlexaFluor488 and RegIIIγ-AlexaFluor555 shows preferential binding to the WGA-positive bacterial population. Results are representative of 3 independent experiments.

Given that intestinal microbial communities consist of both Gram-positive and Gram-negative species [149], we asked whether RegIIIγ bound preferentially to one of these groups. Wheat germ agglutinin is a plant lectin that binds to surface peptidoglycan on Gram-positive bacteria, thus distinguishing between Gram-positive and Gram-negative populations [150]. Dual staining with fluorochrome-conjugated RegIIIγ and WGA revealed that RegIIIγ preferentially bound to the WGA-positive bacterial population (Figure 5.4B). Furthermore, we determined that RegIIIγ bound to pure preparations of cultured Gram-positive bacteria, including *Listeria innocua* and *Enterococcus faecalis*, with comparatively little binding to preparations of cultured Gram-negative bacteria, including *Salmonella typhimurium* and *Pseudomonas aeruginosa* (Figure 5.5).

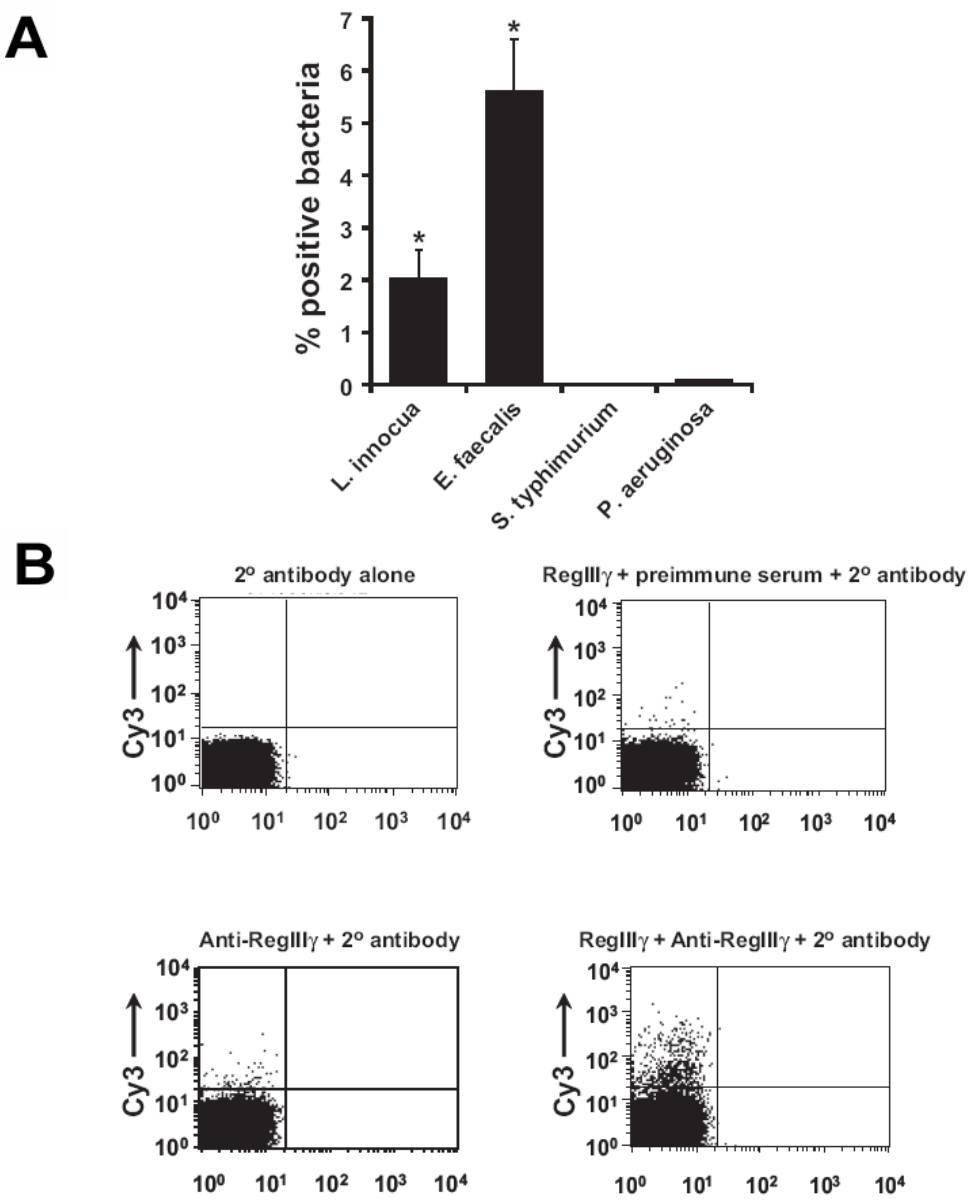

Figure 5.5: RegIIIγ binds preferentially to cultured Gram positive bacteria.

(A) Pure cultures of Gram positive bacteria (*Listeria innocua* and *Enterococcus faecalis*) and Gram negative bacteria (*Salmonella typhimurium* and *Pseudomonas aeruginosa*) were incubated with RegIIIγ followed by detection with anti-RegIIIγ antiserum and goat anti-rabbit-Cy3, and analyzed by flow cytometry. Results are representative of three experiments. Asterisks indicate statistically significant differences between Gram-positive species and *S. typhimurium* ($p<0.05$). **(B)** Purified, recombinant RegIIIγ was incubated with pure preparations of cultured *L. innocua, E. faecalis, S. typhimurium,* or *P. aeruginosa*. Each species was also incubated with secondary antibody alone (goat anti-rabbit-Cy3), RegIIIγ pre-immune serum followed by secondary antibody, or anti-RegIIIγ followed by secondary antibody. Shown are representative dot plots for *L. innocua*. All reactions were done in triplicate in each experiment and the experiment repeated three times.

RegIIIγ and HIP/PAP bind preferentially to PGN

The preference of RegIIIγ for Gram-positive bacteria suggested that this protein binds PGN, a molecule that is exposed on the Gram-positive bacterial surface but buried in the periplasmic space of Gram-negative bacteria. To test this idea, we performed pull-down assays using insoluble cell wall peptidoglycan [81]. Purified RegIIIγ was completely removed from solution by incubation with PGN, and was retained in the PGN-bound fraction even after extensive washing (Figure 5.6A). Human HIP/PAP is 65% identical to RegIIIγ and exhibited a similar PGN binding activity (Figure 5.6A). The specificity of both interactions was confirmed by using soluble PGN (sPGN) to compete for binding to insoluble PGN (iPGN; Figure 5.6B). These results demonstrate that RegIIIγ and HIP/PAP bind PGN.

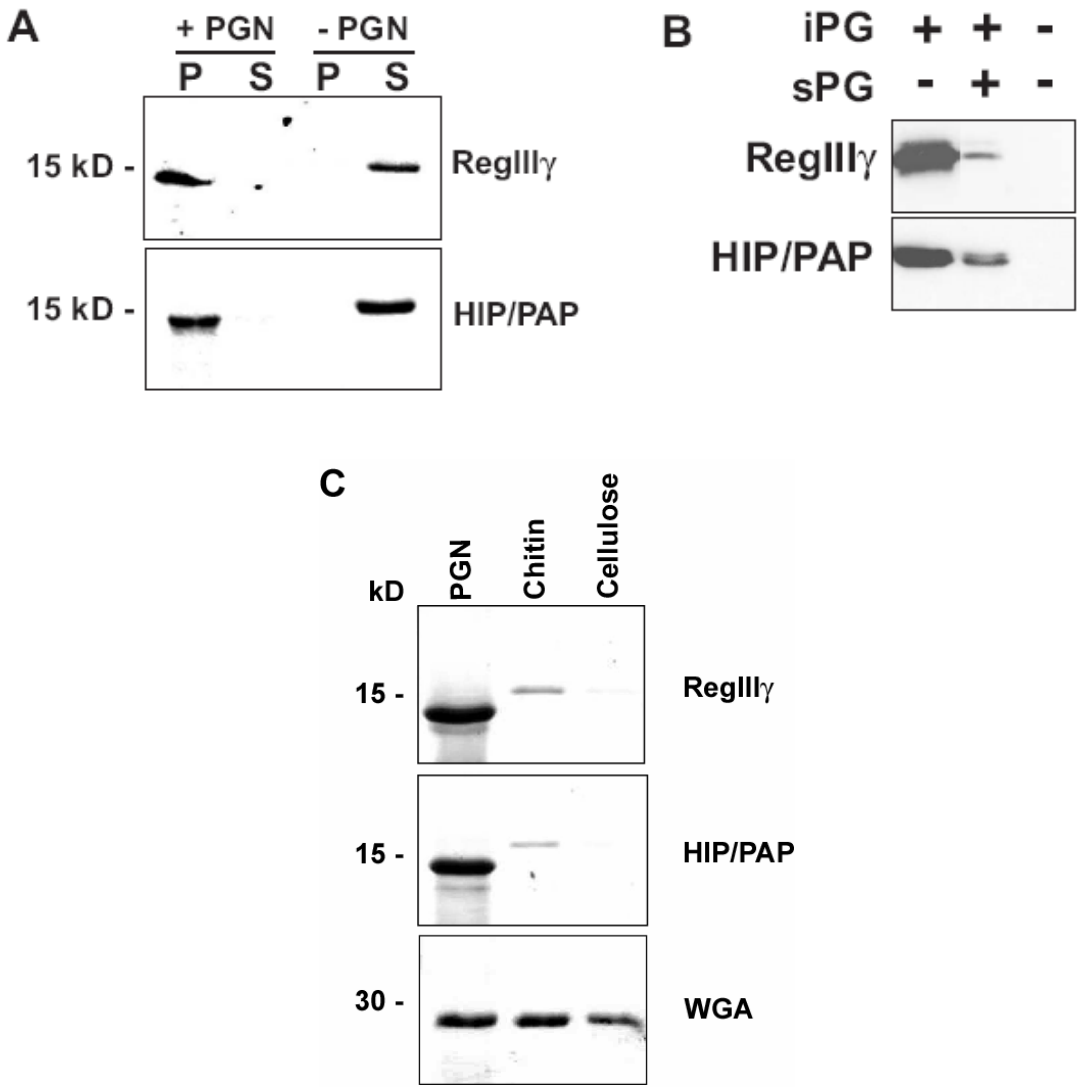

Figure 5.6: Mouse RegIIIγ and human HIP/PAP bind peptidoglycan.

(A) PGN pulldown assays were performed by adding 10 μg of purified recombinant RegIIIγ or HIP/PAP to 50 μg of insoluble *Bacillus subtilis* PGN and pelleted. Pellet (P) and supernatant (S) fractions were analyzed by SDS-PAGE and Coomassie Blue staining. **(B)** Lectin binding to insoluble peptidoglycan is competed by soluble PGN. Pull-down assays were performed with 10 μg of purified recombinant lectins, 2 μg of insoluble *B. subtilis* PGN, and with or without 100 μM soluble *B. subilis* PGN as indicated. **(C)** 30 μg of each lectin was incubated with 100 μg of the indicated polysaccharide for 1 hr. at 4°C. Protein associated with the pellet fraction was analyzed by SDS-PAGE and Coomassie Blue staining.

To determine the range of RegIIIγ and HIP/PAP binding to carbohydrates with β-1,4 linkages, we tested chitin and cellulose next to PGN in our pull-down assay. Using

comparable microgram amounts of insoluble polysaccharide, more RegIII protein was associated with the pellet fraction of PGN than either of the other β-1,4-linked carbohydrates, suggesting a binding preference for PGN (Figure 5.6C). In contrast, comparable quantities of WGA was pulled down by each polysaccharide, indicating that this binding pattern was specific to the RegIII proteins. Additionally, binding affinities for peptidoglycan recognition proteins (PGRPs) have been described using a variety of methods, including Scatchard plots and surface plasmon resonance (SPR) [100, 101, 151] (see Table 5.1). In order to determine the affinity of RegIIIγ and HIP/PAP for *B. subtilis* PGN, we iodinated each protein and performed the pull-down assay either in the presence or absence of unlabeled Reg protein. The results (listed in Table 5.1) indicate that both RegIIIγ and HIP/PAP bind iPGN with low nanomolar affinity, similar to that seen with other PGRPs. Taken together, these data strongly suggest that we have identified PGN as the biological ligand of these two RegIII proteins.

Table 5.1: Comparison of determined dissociation constants for PGRPs and RegIII proteins

Name	Organism	Affinity (K_d)	Technique	Reference
PGRP-S	Mouse	13 nM	Scatchard	Liu et al. 2000.
"	Human	189 mM (Lys)	SPR	Kumar et al. 2005
"	Human	~100 nM (DAP)*	SPR	Kumar et al. 2005
"	Human	~3.7 mM	SPR	Cho et al. 2005.
PGRP-Iα	Human	62 mM	SPR	Kumar et al. 2005.
PGRP-LC	Drosophila	N/A	SPR	Cho et al. 2005.
RegIIIγ	Mouse	~40 nM	Scatchard	This study
HIP/PAP	Human	~37 nM	Scatchard	This study

* This PGRP displays preferential binding to DAP-PGN found in most Gram negatives

Interestingly, preliminary data suggests that both proteins recognize the PGN isolated from *B. subtilis, Streptomyces* or *Micrococcus luteus* comparably, while binding to *S. aureus* PGN was significantly reduced (H. Cash, unpublished observations). Recent studies have described the ability of *S. aureus* to O-acetylate its PGN, which provides the basis for its resistance to lysozyme [51]. Further investigation will elucidate whether this modification accounts for the reduced binding of the RegIII proteins to this species' PGN as well.

The RegIII proteins recognize the carbohydrate backbone of PGN

PGN consists of extended glycan chains composed of alternating N-acetylglucosamine (GlcNAc) and N-acetylmuramic acid (MurNAc) residues cross-linked by short peptides (Figure 5.7A). Because RegIIIγ and HIP/PAP contain predicted CRDs, we predicted that they would recognize the carbohydrate moiety of peptidoglycan. In order to test this hypothesis, we utilized chitin, which is a β1,4-linked GlcNAc polymer that is virtually identical to the peptidoglycan carbohydrate backbone (Figure 5.7A). As shown in Figure 5.7B, purified recombinant RegIIIγ and HIP/PAP bound to chitin immobilized on Sepharose beads. In contrast, neither lectin bound dextran-Sepharose, uncoupled Sepharose, or monomeric GlcNAc-Sepharose (data not shown). We also did not detect binding to a number of other immobilized monosaccharides and disaccharides, including glucose, galactose, lactose, and trehalose [152]. Although the C-type lectin family includes members that bind their ligands in a calcium-dependent manner, we found that RegIIIγ and HIP/PAP do not require calcium for binding to peptidoglycan, chitin, and mannan ([152]; and data not shown). The results of our biochemical and

glycoarray analyses demonstrate that both lectins bind highly polymeric GlcNAc (chitin), but do not bind monomeric GlcNAc. This binding preference suggests that RegIIIγ and HIP/PAP are pattern recognition proteins that recognize the microbe-associated molecular pattern represented by the peptidoglycan carbohydrate backbone. However, the data presented in Figure 5.6C also indicate that the peptidoglycan peptide moiety influences the avidity of lectin binding. Delineating the contributions of each portion of the PGN molecule is an intriguing future direction, which would provide more specific information on the structural interactions of these proteins with their ligands.

C-type lectins that bind GlcNAc-containing saccharides frequently also bind mannose-containing saccharides [153]. GlcNAc and mannose have a similar arrangement of their 3- and 4-hydroxyl groups, which make key contacts with C-type lectin binding sites [62]. As predicted, both RegIIIγ and HIP/PAP bound to mannan, a polymer of mannose residues (Figure 5.7B) [152]. However, neither lectin bound to monomeric mannose [152], highlighting the preference of both proteins for polysaccharides.

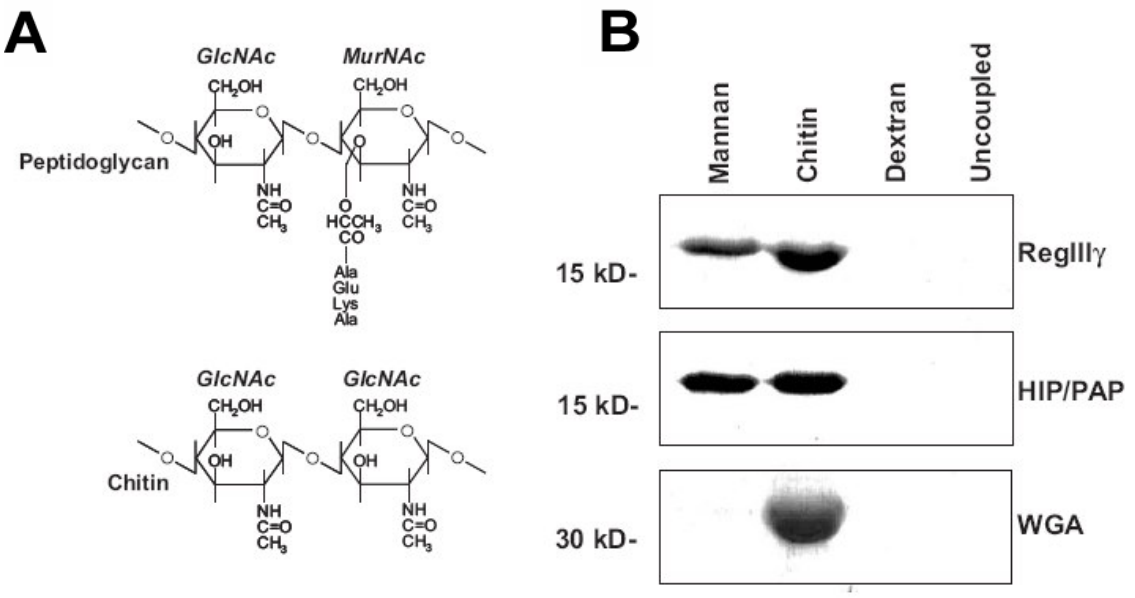

Figure 5.7: The RegIII lectins interact with PGN through its carbohydrate backbone.

(A) Comparison of PGN and chitin structures. The structure of a typical Gram-positive PGN is depicted. **(B)** Lectin binding to immobilized polysaccharides was tested by adding 50 μg of each lectin to 25 μL of immobilized polysaccharide for 2 hours at 4°C. After washing, bound proteins were released by boiling the Sepharose beads in SDS-PAGE sample buffer followed by electrophoresis through 15% polyacrylamide gels and Coomassie blue staining. Note that WGA shows the appropriate specificity for chitin.

In order to screen through a variety of standard ligands characterized for lectins, we also utilized a glycoarray approach [154]. These spotted arrays contain ~200 various carbohydrate structures composed of a variety of mono-, di-, tri-, tetra-, and pentasaccharides and linkages. Results from multiple experiments indicated that none of our recombinant lectins bound appreciably to any of the array ligands, with the exception of ligand 127. Because this ligand is one of the few bacterial carbohydrates represented on the array (muramyl dipeptide (MDP), the basic component of PGN), the observed increase in fluorescence is due to the cross-reactivity of the polyclonal antiserum raised to a recombinant *E. coli* protein probably contaminated with trace MDP (Figure 5.8). These data are consistent with our observations that the RegIII proteins interact with highly polymeric carbohydrate structures rather than smaller saccharide chains. However,

with further optimization, this array represents a powerful tool for identification of alternate ligands for the RegIII proteins.

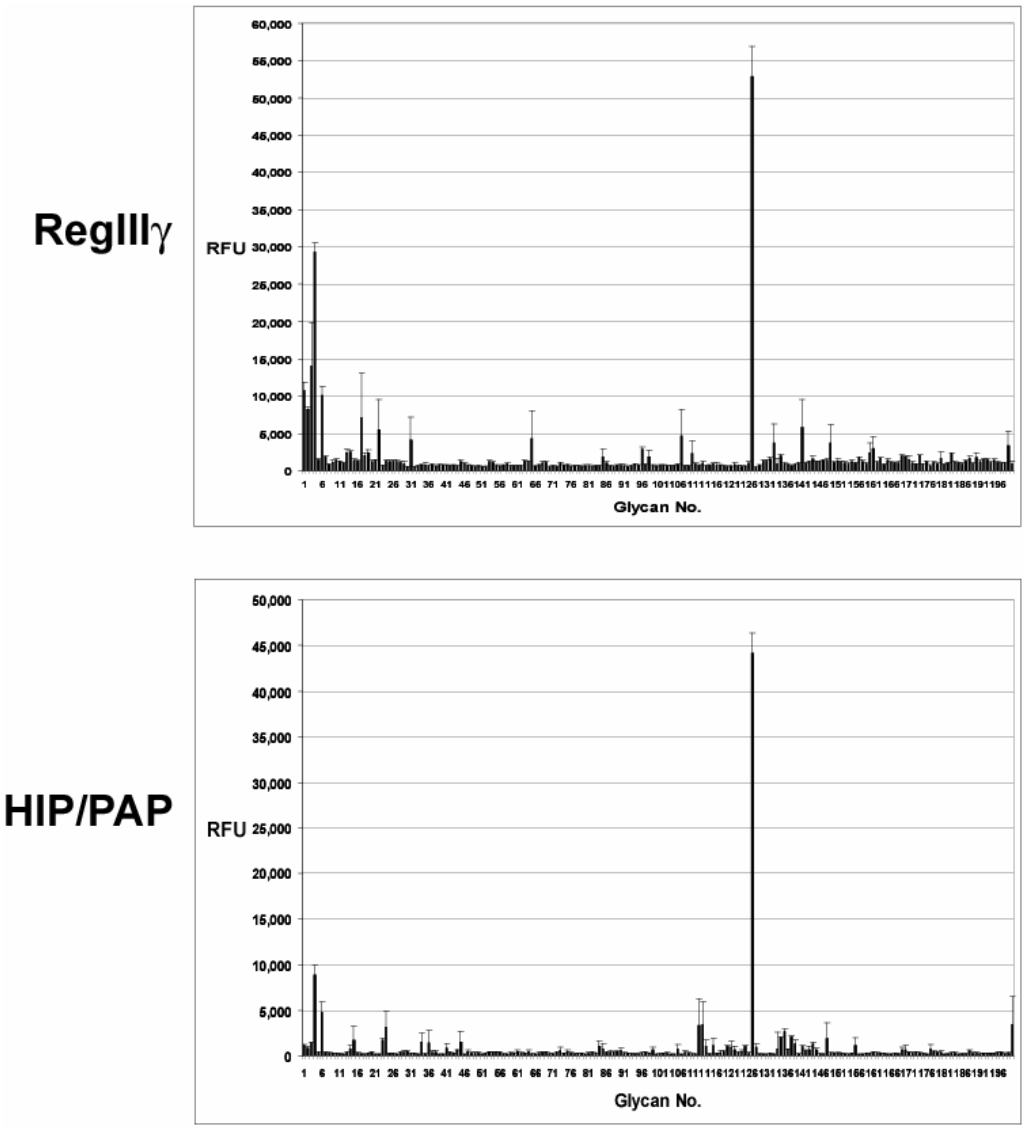

Figure 5.8: Glycoarray analysis of RegIIIγ and HIP/PAP.

A total of 200 glycans were screened for binding of RegIIIg or HIP/PAP as described in Materials and methods. Error bars represent mean s.d. replicate determinations from a single experiment. Note that ligand 127 (muramyl dipeptide) stands out for ligand binding. RFU = relative fluorescence units.

RegIIIγ and HIP/PAP are bactericidal, targeting Gram positive organisms

Our results reveal a carbohydrate ligand preference that is remarkably similar to that of mannose binding lectin (MBL), a C-type lectin with an established role in innate immunity. As a serum protein, MBL recognizes invading microbes by binding to surface mannose residues [68] or to peptidoglycan [155]. This binding triggers the lectin-activated complement pathway, which is initiated by recruitment of the serine proteases MASP1 and 2 via interactions with the MBL collagenous domain. In contrast to MBL, RegIIIγ and HIP/PAP consist solely of secreted CRDs that lack collagenous domains required for complement recruitment. We therefore postulated that RegIIIγ and HIP/PAP might be directly antimicrobial, without requiring additional factors to kill targeted microbes. We therefore added purified RegIIIγ and HIP/PAP to Gram-positive enteric microbes including *Listeria monocytogenes*, *Listeria innocua*, and *Enterococcus faecalis*, and observed a dose-dependent reduction in the viability of each organism (Figure 5.9A). The number of colony forming units (CFU) of each microbe declined by 99% after a 2 hour exposure to 5 μM HIP/PAP (Figure 5.9A). A similar decline in the viability of *L. monocytogenes* and *L. innocua* was observed after a 2 hour exposure to 5 μM RegIIIγ (Figure 5.9A). The viability of *E. faecalis* declined by ~80% after a 2 hour exposure to 10 μM RegIIIγ (Figure 5.9A). Thus, the effective antibacterial concentrations of both lectins are similar to those of other intestinal antimicrobial proteins [54, 156].

As expected, neither RegIIIγ nor HIP/PAP was bactericidal toward the Gram-negative enteric organisms *Escherichia coli* or *Bacteroides thetaiotaomicron* (Figure 5.9A). This is consistent with our observation of preferential binding to Gram-positive bacteria and the fact that peptidoglycan is buried in the periplasmic space of Gram-

negative bacteria. Additionally, neither lectin reduced the viability of fungal microorganisms, including *Saccharomyces cervisiae* and *Candida albicans* (data not shown).

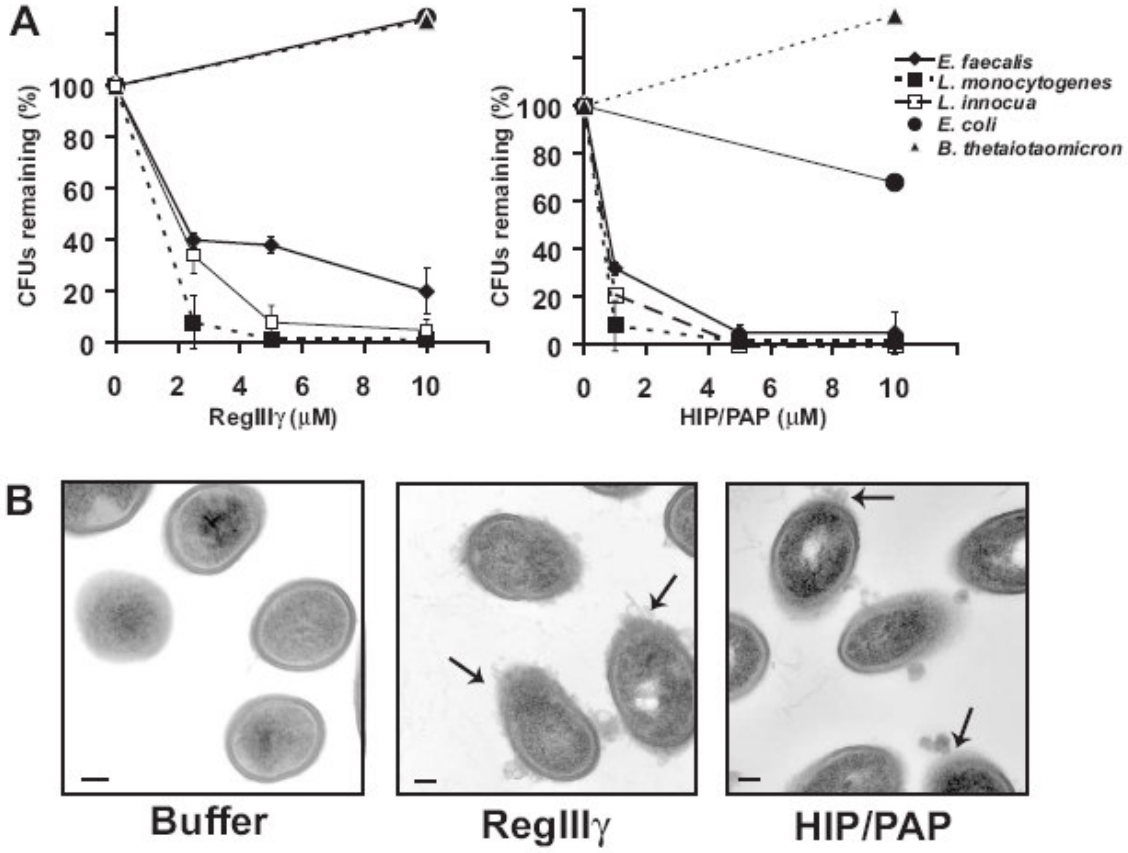

Figure 5.9: Mouse RegIIIγ and human HIP/PAP display antibacterial activity against Gram-positive bacteria.

(A) Percentage of CFUs remaining after exposure to purified RegIIIγ and HIP/PAP. *Listeria innocua, Listeria monocytogenes, Enterococcus faecalis, Escherichia coli* K12, and *Bacteroides thetaiotaomicron* were grown to mid-log phase and incubated with purified lectins. Initial bacterial concentrations ranged from 10^5 to 10^6 CFU/mL. After incubation for 2 hr. at 37°C, viable bacteria were quantified by dilution plating. Assays were done in triplicate. Mean ± s.d. is plotted. **(B)** Transmission electron microscopy of *L. monocytogenes* following a 2 hr. exposure at 37°C to 10 μM purified recombinant lectins. Arrows indicate examples of damaged cell surfaces and cytoplasmic leakage. Bar = 100 nm.

We used transmission electron microscopy to visualize morphological changes in *Listeria monocytogenes* cells following exposure to RegIIIγ and HIP/PAP. Our images revealed evidence of cell wall damage and cytoplasmic leakage (Figure 5.9B). These results are remarkably similar to those obtained with cationic antimicrobial peptides such as human β-defensin-3 [156], which kill bacteria by cell wall permeabilization. Our findings indicate that lectin-mediated bacterial killing also involves cell wall damage.

RegIIIγ and HIP/PAP bactericidal activities were inhibited with soluble peptidoglycan and chitin fragments, linking peptidoglycan binding to antibacterial function. Addition of 35 μM soluble peptidoglycan (sPGN) to antibacterial assays inhibited the bactericidal activity of both lectins (Figure 5.10). 10 mM chitotetraose, a 4-sugar acid hydrolysis fragment of chitin, also fully inhibited the antibacterial activity of both RegIIIγ and HIP/PAP (Figure 5.10). Consistent with the preference of RegIIIγ and HIP/PAP for polymeric sugars, 10 mM monomeric GlcNAc or chitobiose ($GlcNAc_2$) were less inhibitory. These results demonstrate that a soluble oligosaccharide that mimics the peptidoglycan saccharide backbone is sufficient to inhibit lectin antimicrobial activity. These findings are consistent with a model in which lectin binding to surface peptidoglycan carbohydrate precedes microbial killing.

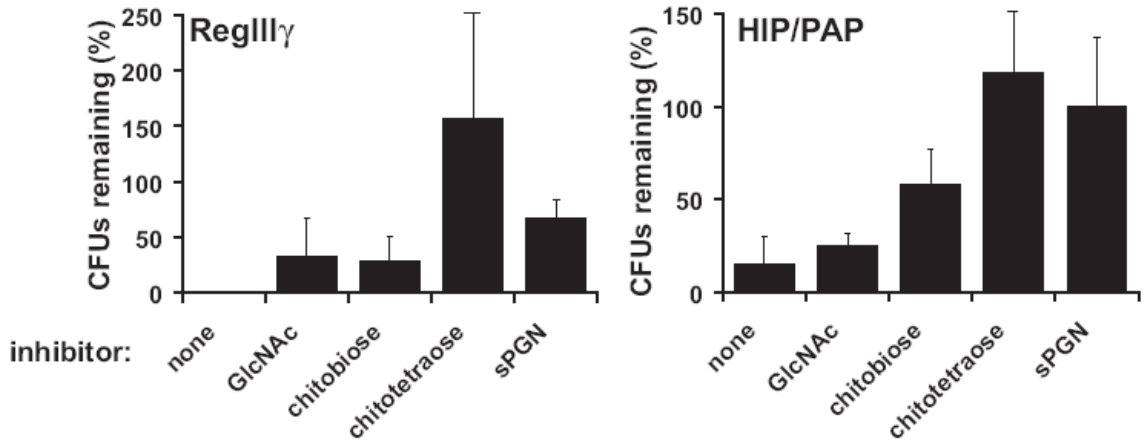

Figure 5.10: Lectin bactericidal activity is inhibited by chitooligosaccharides and soluble peptidoglycan (sPGN).

10 mM GlcNAc, chitobiose ($GlcNAc_2$), or chitotetraose ($GlcNAc_4$), or 35 μM sPGN were added to antibacterial assays performed on *L. innocua* as in Figure 4.9. Each % CFU was calculated relative to a no-lectin control assay containing an identical amount of chitooligosaccharide or sPGN. Results are representative of multiple experiments.

In order to test the bactericidal activity of RegIIIγ *in vivo*, we initially performed an enema experiment using germ-free mice that had been mono-colonized with a known target organism, *L. innocua*. Recombinant RegIIIγ was concentrated to 63 μM and infused into the gastrointestinal tracts of two male NMRI mice, while buffer was infused concomitantly into two control mice. Following sacrifice of the mice and excision of the distal small intestines, large intestines, and rectums, we quantified the number of viable CFUs in each intestinal segment by performing dilution plating on luminal contents. As shown in Figure 5.11, infusion of recombinant RegIIIγ did not result in appreciable depletion of luminal bacteria. However, the function of RegIIIγ's antibacterial activity *in vivo* would be more accurately assessed using a targeted genetic deletion approach.

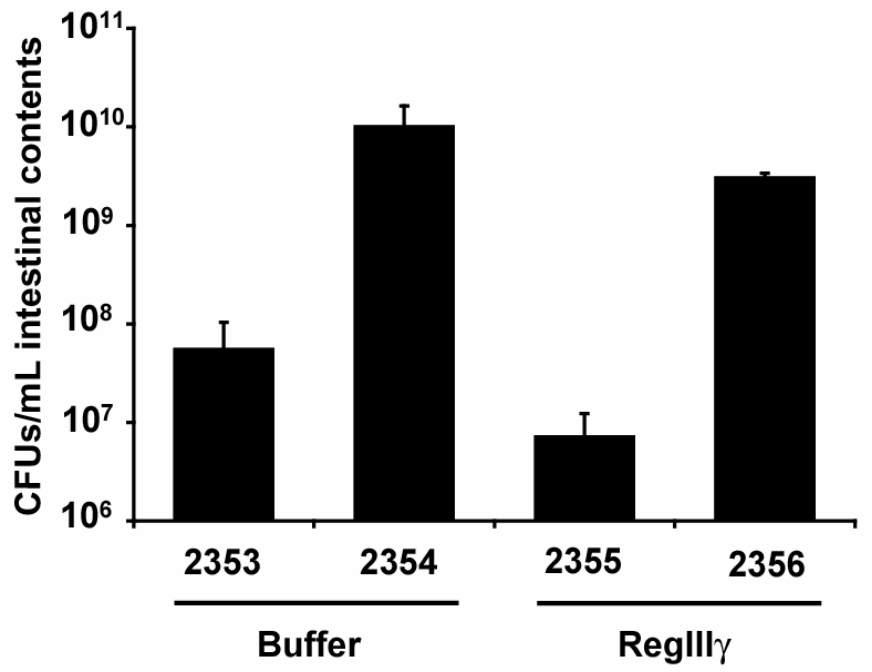

Figure 5.11: Results of infusion of recombinant RegIIIγ into mouse intestines.

Recombinant RegIIIγ or buffer was infused by enema into the GI tracts of male NMRI mice mono-colonized with *L. innocua* for 2 hours as described in Materials and methods. Remaining, viable CFUs were quantified for each mouse in triplicate by dilution plating. Mean ± s.d. is plotted for bacteria recovered from the rectum of each mouse.

RegIIIγ expression is triggered by intestinal bacteria

Given RegIIIγ's bactericidal activity, we predicted that its expression patterns would reflect microbial colonization levels in the mouse small intestine. Q-PCR analysis of *RegIIIγ* mRNA levels along the cephalocaudal axis of conventionalized small intestines revealed increasing expression toward the distal region (ileum; Figure 5.12A), concomitant with increasing microbial densities. By contrast, germ-free mice showed minimal *RegIIIγ* expression throughout the small intestine (Figure 5.12A). We also assayed for changes in *RegIIIγ* mRNA expression during postnatal intestinal development. *RegIIIγ* mRNA levels rose dramatically during the weaning period

(postnatal days 17-22) and remained high into adulthood (≥P28) in conventionally-raised but not germ-free mice (Figure 5.12B). Weaning is associated with dramatic changes in gut microflora composition as well as withdrawal of maternal IgA antibodies. The antibacterial activity of RegIIIγ suggests that its expression is elicited as part of a compensatory response to maintain mucosal homeostasis in the face of changing microbial ecology and withdrawal of passive immunity.

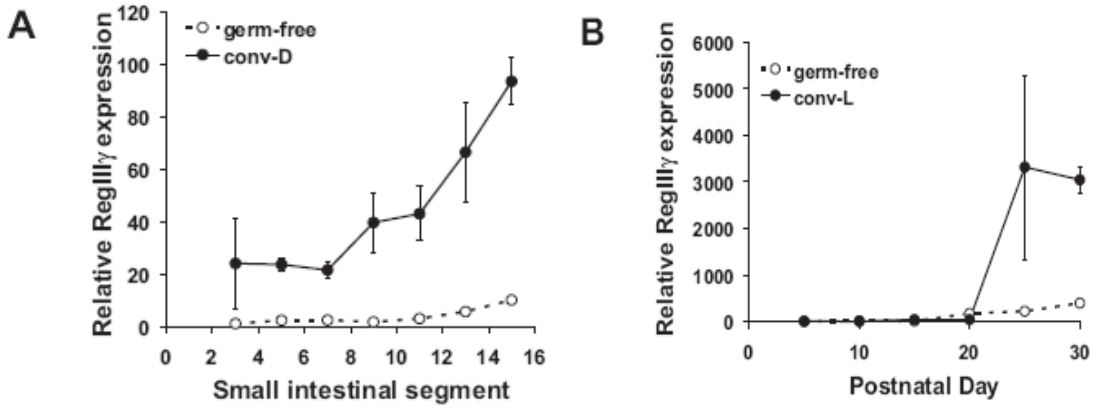

Figure 5.12: RegIIIγ expression mimics levels of intestinal colonization.

(A) RegIIIγ expression along the cephalocaudal axis of the small intestine. Small intestines from adult germ-free or conventionalized (conv-D) NMRI mice were divided into 16 equal segments and *RegIIIγ* mRNA expression was determined in specific segments using Q-PCR. Results are representative of experiments in two sets of mice. **(B)** *RegIIIγ* mRNA increases during the weaning period (P17-22) in developing, conventionally raised NMRI mice. Assays were performed on pooled mid-small intestinal RNAs (n-3 mice/time point). *Figure 4.11B courtesy of Cecilia Whitham.*

Because conventional microflora are composed of diverse microbial societies, we asked whether single enteric bacterial species are sufficient to drive small intestinal *RegIIIγ* expression. As expected, a mixed microbial community recovered from a conventional mouse elicited a ~20-fold increase in *RegIIIγ* expression when introduced into germ-free wild-type C57/b6 mice. In contrast, colonization with the Gram-negative symbiont *Bacteroides thetaiotaomicron* elicited only a 2.5-fold increase in expression,

while the non-invasive Gram-positive *Listeria innocua* had no effect on *RegIIIγ* mRNA levels (Figure 5.13). These results indicate that neither organism alone was sufficient to stimulate *RegIIIγ* expression to conventional levels in wild-type mice. However, bacteria that are normally strictly compartmentalized in the intestinal lumen show increased mucosal adherence and invasion in mice that lack mucosal IgA [21]. We therefore postulated that mucosal defenses such as secretory IgA might be sufficient to sequester *B. thetaiotaomicron* and *L. innocua* in the gut lumen, accounting for the inability of these single species to stimulate RegIIIγ expression. Indeed, we found that *B. thetaiotaomicron* and *L. innocua* trigger a 52- and 41-fold increase, respectively, in *RegIIIγ* mRNA expression following colonization of germ-free RAG1-deficient mice, which lack mature lymphocytes and are therefore IgA-deficient [157] (Figure 5.13). Wild-type and RAG1-deficient mice were colonized to virtually identical levels (~10^8 cfu/ml ileal contents), indicating that differences in *RegIIIγ* mRNA expression did not result from differences in total microbial numbers. Our findings thus support a model in which increased bacterial-epithelial contact drives *RegIIIγ* expression as a mechanism to limit potential microbial penetration and maintain mucosal surface integrity.

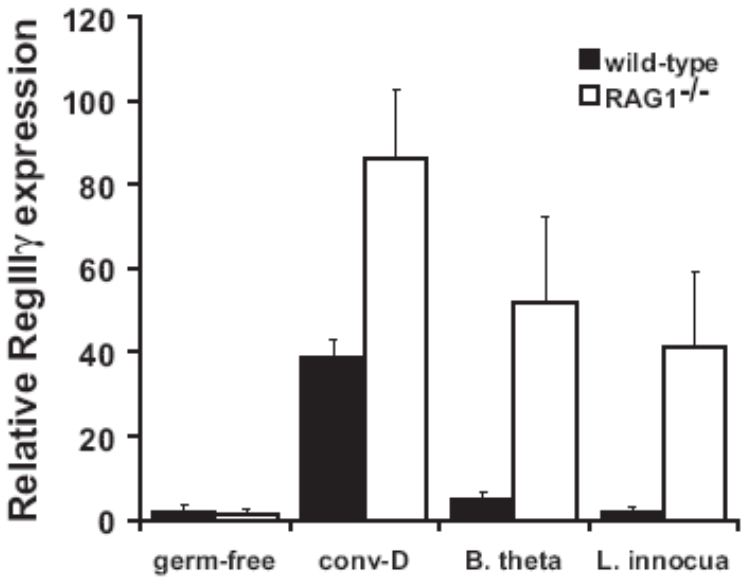

Figure 5.13: RegIIIγ expression is triggered by intestinal bacteria.

RegIIIγ expression is induced by single Gram-positive or Gram-negative bacterial species in immunodeficient mice. Q-PCR determinations were done on cDNAs from mid-small intestine. Each point represents the average value from ≥3 different mice. All Q-PCR determinations in the figure were performed in triplicate (mean ± s.d. plotted) and were normalized to 18S rRNA. *Courtesy of Cecilia Whitham*.

C. Discussion

The data presented in this chapter demonstrate that RegIIIγ and HIP/PAP represent a new family of inducible antibacterial proteins. These effectors seek out their microbial targets by interacting with exposed peptidoglycan on the bacterial cell surface. Because they lack domains necessary for complement recruitment and are directly bactericidal, these proteins reveal a new function for mammalian C-type lectins. Some of the significant implications of these findings are that these antibacterial lectins could represent an evolutionarily primitive form of lectin-mediated innate immunity, and that

the lectin-mediated complement pathway may have evolved from a directly antimicrobial precursor. In support of these hypotheses, simple model organisms such as *Drosophila* and *Caenorhabditis elegans* harbor a number of genes encoding putative C-type lectins that consist solely of a simple CRD with an N-terminal secretion signal [65, 158]. Our results suggest that these proteins could function in innate antimicrobial defense. Furthermore, the human and mouse Reg families encompass multiple members, many of which are expressed in the gut [159]. It seems likely that other members of this extensive protein family are also antimicrobial, but may exhibit preferences for different microbial targets.

Our initial microarray screen, validated by Q-PCR data, demonstrates that the expression of RegIIIβ and RegIIIγ increases dramatically upon intestinal colonization. Such modulation by bacterial load within the intestinal lumen represents another example of a transcriptional change elicited within host cells in response to the presence of beneficial microbes. Indeed, induction of RegIII transcript is independent of bacterial cell wall architecture, membership in the normal flora of a conventional mouse, or inherent pathogenic ability (Figure 5.13). Rather, production and secretion of RegIIIγ appears to be governed by mucosal proximity of luminal bacteria or increased adherence and attachment.

The mechanism of RegIIIγ and HIP/PAP binding to PGN is distinct from the one described for other peptidoglycan recognition proteins (PGRPs, [100, 101, 160]). Our results reveal an interaction with bacterial PGN that is lectin-like in its specificity for the carbohydrate backbone structure of the PGN molecule instead of the tetrapeptide crossbridge (Figure 5.7A). This was further substantiated by the inhibitory ability of

soluble carbohydrate fragments added to the bactericidal assays. Thus, these data support a model in which the RegIII lectins bind to the carbohydrate moieties of PGN, followed by target cell lysis. Further investigation is required to determine if the specificity of the RegIII proteins for Gram-positive targets is due to a structural difference in their PGN or a cell wall architecture difference, in that their PGN is exposed on the cell surface.

Our attempts to validate the function of RegIIIγ *in vivo* resulted in ambiguous results with a variety of potential explanations. First, the concentration of RegIIIγ infused into the intestine was not sufficient to cause marked changes in the number of organisms because of proteolytic degradation in the intestine, steric inability to contact potential target bacteria, or due to dilution of the protein below effective concentrations in the large volume of the intestine. Secondly, RegIIIγ is produced and secreted by Paneth cells situated in the base of small intestinal crypts, which have been described as harboring few bacterial cells relative to the intestinal lumen. Thus, the RegIII proteins may function best in the microenvironment of the crypt, where relatively few organisms are encountered. Finally, infusion of the protein through the rectum and colon could preclude any RegIIIγ reaching the small intestine, where we have shown its expression to be exclusively localized. Therefore, the altered environments of the lower GI tract could inhibit RegIIIγ bactericidal function. Future studies of the RegIII family will involve targeted genetic deletion strategies in mice to more specifically delineate the roles played by each RegIII protein in maintaining mucosal-epithelial barrier integrity.

Preliminary analysis of target bacteria exposed to RegIIIγ revealed a disruption of the bacterial cell walls, which is in accordance with recognition of surface carbohydrate components by the RegIII proteins. Additionally, these data are similar to those seen

upon exposure of bacteria to β-defensins [58, 161], suggesting that the RegIII lectins may employ a destructive method similar to the pore-forming ability ascribed to the defensin family. Alternatively, these proteins may act as amidases, cleaving the chemical bonds of PGN to destabilize the cell wall, an action which has been attributed to members of the PGRP protein family [86, 97, 98]. Future research will focus on distinguishing between these two hypothetical mechanisms of RegIII-mediated bactericidal activity.

The discovery of directly antibacterial C-type lectins points to a previously uncharacterized mucosal defense mechanism that helps to sequester the gut microflora and preserve intestinal homeostasis. Our results suggest that RegIIIγ expression is triggered by increased microbial-epithelial contact at mucosal surfaces. Enhanced expression of Reg proteins such as HIP/PAP in IBD patients may therefore be a compensatory response that limits mucosal penetration by gut microbes. Since Reg proteins exhibit increased expression in IBD mucosa [112, 138] while α-defensin expression decreases [44, 143], these two groups of antimicrobial proteins are probably regulated by distinct mechanisms. Further investigation will be required to decipher the host and microbial signaling mechanisms that regulate antimicrobial lectin expression. These studies will contribute to a better understanding of IBD pathogenesis and provide new insights into how symbiotic host-microbial relationships are maintained.

Chapter Six:

CHARACTERIZATION OF REGIIIβ LIGAND BINDING AND FUNCTION

A. Introduction

RegIIIβ is one of four members of the RegIII protein multigene family in mice, which has only two members in humans. Due to their homology to regenerating factor I (or lithostatine), these proteins have been attributed mitogenic functions in promoting pancreatic or liver regeneration [115, 120]. However, all RegIII family members contain a canonical carbohydrate recognition domain (CRD), which also classifies them as putative C-type lectins. In contrast to other well-known members of the lectin superfamily, the RegIII proteins are small, consisting almost entirely of their CRD accompanied by an N-terminal secretion signal peptide [62, 106]. Prior investigation has localized the expression of two members of this family, RegIIIβ and RegIIIγ, almost exclusively to the small intestine, where they are upregulated under

immunocompromised conditions [104, 117]. These data and those presented in the previous chapters describe an innate immune effector function for members of the RegIII family.

In this chapter, we present data, which demonstrate that RegIIIβ binds intestinal bacteria by specifically interacting with exposed peptidoglycan, suggesting a role as a component of intestinal innate immunity. To perform the following assays, we purified RegIIIβ using bacterial overexpression and ion exchange chromatography, and characterized its ligand binding. Additionally, we demonstrated that this protein binds to intestinal bacteria and that it binds the same population of bacteria as its family member, RegIIIγ. These studies describe the ligand and function of a previously uncharacterized type VII lectin, while also providing additional insight into the defense of host tissues by innate immune effectors.

B. Results

Both RegIIIβ and RegIIIγ display similar expression patterns

Initial identification of the mouse C-type lectin RegIII family defined the small intestine and pancreas as the two tissues responsible for the vast majority of all RegIII protein production [104, 106]. Subsequently, published reports have implicated two proteins within this family, RegIIIγ and RegIIIβ, as demonstrating increased expression in a variety of intestinal diseases [117, 118]. As described in the Project Overview, a microarray screen comparing Paneth cell transcripts from germ-free mice and

conventionalized mice revealed that these two proteins were the most significantly upregulated transcripts in conventionalized mice compared to germ-free. In conjunction with these findings, real-time quantitative PCR analysis confirmed that RegIIIβ transcript levels are most highly expressed in the small intestine, with some slight expression in the rectum (Figure 6.1A). This robust, almost exclusive, expression in the small intestine mimics that of its closely homologous family member, RegIIIγ (Figure 6.1B).

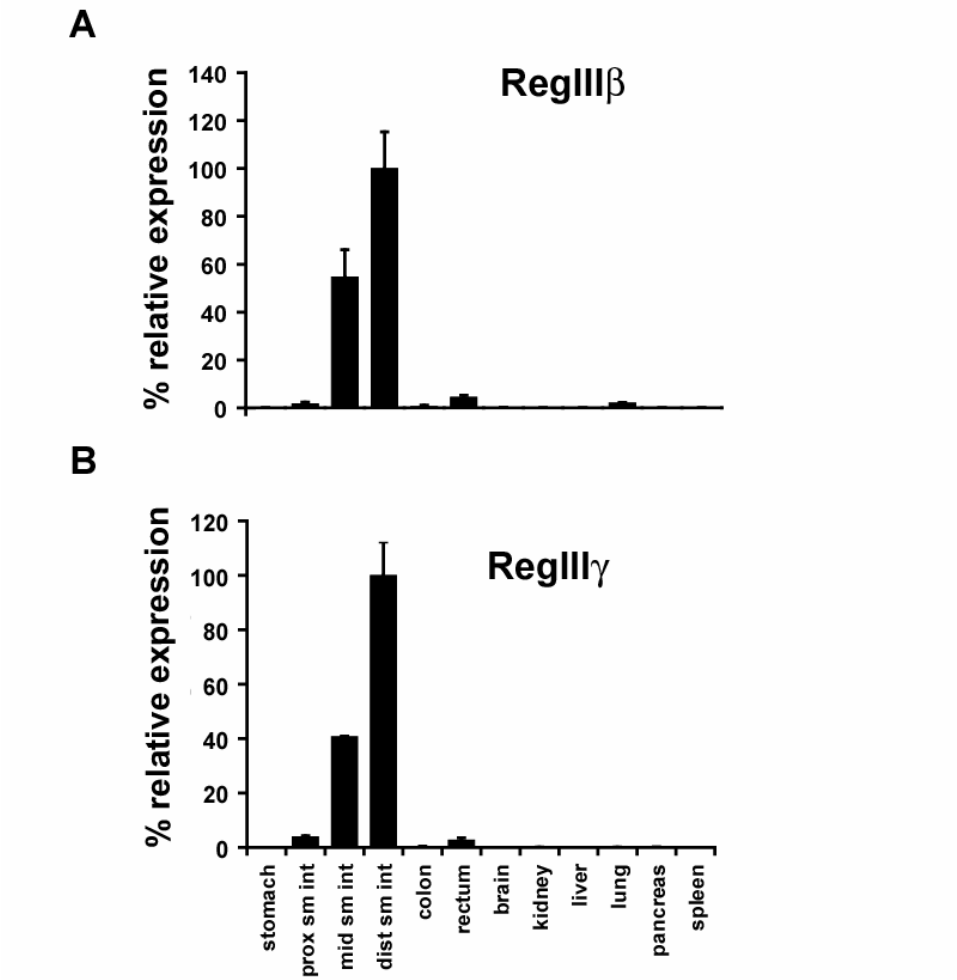

Figure 6.1: Expression of RegIIIβ and RegIIIγ transcripts in various tissues.

RegIIIβ **(A)** or *RegIIIγ* **(B)** mRNA expression was assessed in whole tissues harvested from conventionally raised NMRI mice. Q-PCR analysis was performed on RNAs isolated from the tissues listed above. Values were normalized to 18S rRNA, and mean ± s.d. is plotted for each tissue. Results are expressed relative to the brain sample and are representative of >3 independent experiments.

In order to determine the cellular compartment responsible for RegIII production, we utilized *in situ* hybridization, a technique which allowed us to distinguish between the closely related RegIII transcripts. As shown in Figure 6.2, anti-sense probes generated to both RegIIIβ and RegIIIγ hybridized to endogenous transcripts produced by the epithelial cells at the base of intestinal crypts, including Paneth cells. The lack of expression in the epithelial cells at the tips of the villi is not attributable to the degradation of cellular RNA during the fixation process, for as RegIIIβ expression levels increase along the length of the small intestine, epithelial cells lining the villus begin to display positive hybridization (see Figure 6.8). Conversely, sense-oriented radioprobes generated to RegIIIβ, RegIIIγ, and LFABP did not demonstrate any hybridization to similarly prepared tissue (Figure 6.2).

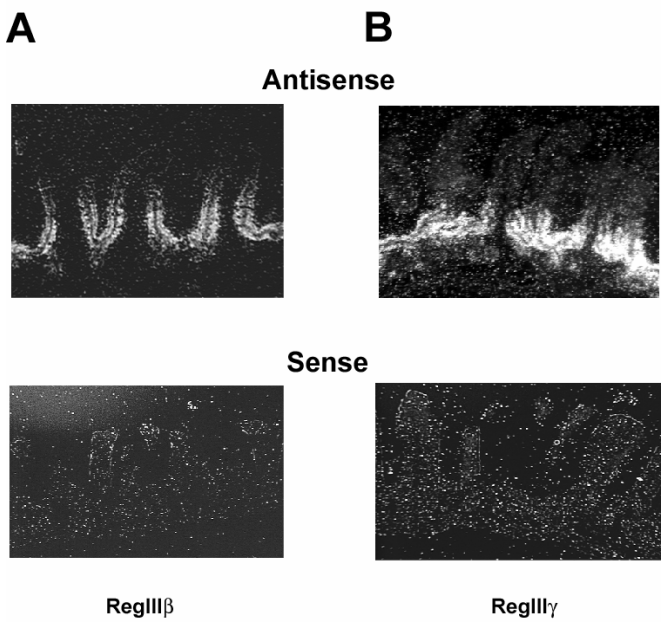

Figure 6.2: *In situ* hybridization localizes RegIII transcripts to intestinal crypts.

Radio-labeled RNA probes specific for either RegIIIβ **(A)** or RegIIIγ **(B)** were generated using an *in vitro* transcription reaction. Whole intestines were isolated from C57/B6 mice, fixed, and embedded in paraffin. Probes were allowed to hybridize with preserved RNAs for 1 week, the reaction was halted, and slides were visualized using dark field microscopy.

Additionally, we utilized real-time RT-PCR analysis to confirm the results obtained from the microarray by comparing intestinal RegIIIβ and RegIIIγ expression in germ-free and conventionally colonized mice. Both transcripts were significantly upregulated in response to colonization with a full microflora (Figure 6.3), demonstrating that expression of both proteins can be induced by addition of mixed microbial populations.

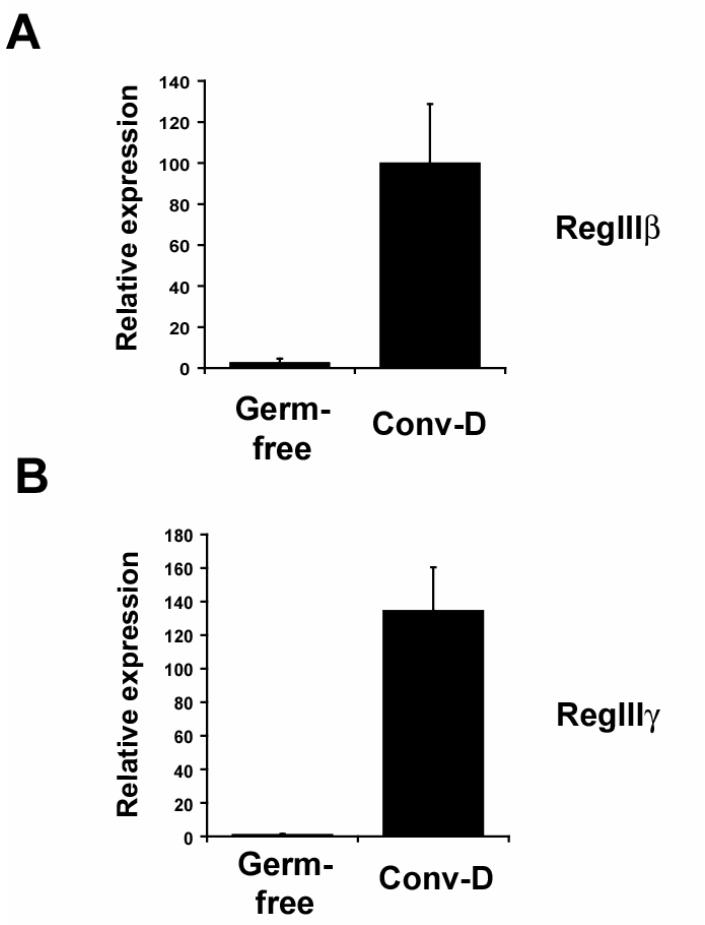

Figure 6.3: Comparison of RegIIIβ and RegIIIγ expression in germ-free and conventionalized mice.

Q-PCR determinations were performed on cDNAs synthesized from RNAs isolated from mid-small intestines of either germ-free or conventionalized (Conv-D) mice. Each point represents the average value from 3 mice. Each sample was analyzed in triplicate (mean ± s.d. plotted) and was normalized to 18S rRNA values.

RegIIIβ binds to peptidoglycan by recognizing the carbohydrate backbone

Because both RegIIIβ and RegIIIγ transcripts are localized to the same intestinal cell type and are regulated similarly by microbial colonization, we hypothesized that RegIIIβ would display an overlapping function with the PGN-binding capacity ascribed to RegIIIγ. Purified, recombinant RegIIIβ was incubated with PGN isolated from *B. subtilis* and bound (pellet) and free (supernatant) fractions were subsequently analyzed. RegIIIβ was cleared from the supernatant fraction after incubation with PGN, while peanut lectin, which recognizes galactose-containing carbohydrates, was detected solely in the supernatant, demonstrating the specificity of the RegIII lectin interaction (Figure 6.4A).

To further confirm the specificity of RegIIIβ's binding to PGN, we inhibited binding by adding soluble peptidoglycan from the same bacterial species. Upon incubation with both soluble and insoluble PGN (sPGN and iPGN, respectively), the amount of RegIIIβ associated with the iPGN pellet fraction was dramatically decreased compared to incubation in the absence of sPGN (Figure 6.4B). These results indicate that RegIIIβ specifically recognizes and binds to bacterial PGN, and that this binding can be competed by solubilized PGN.

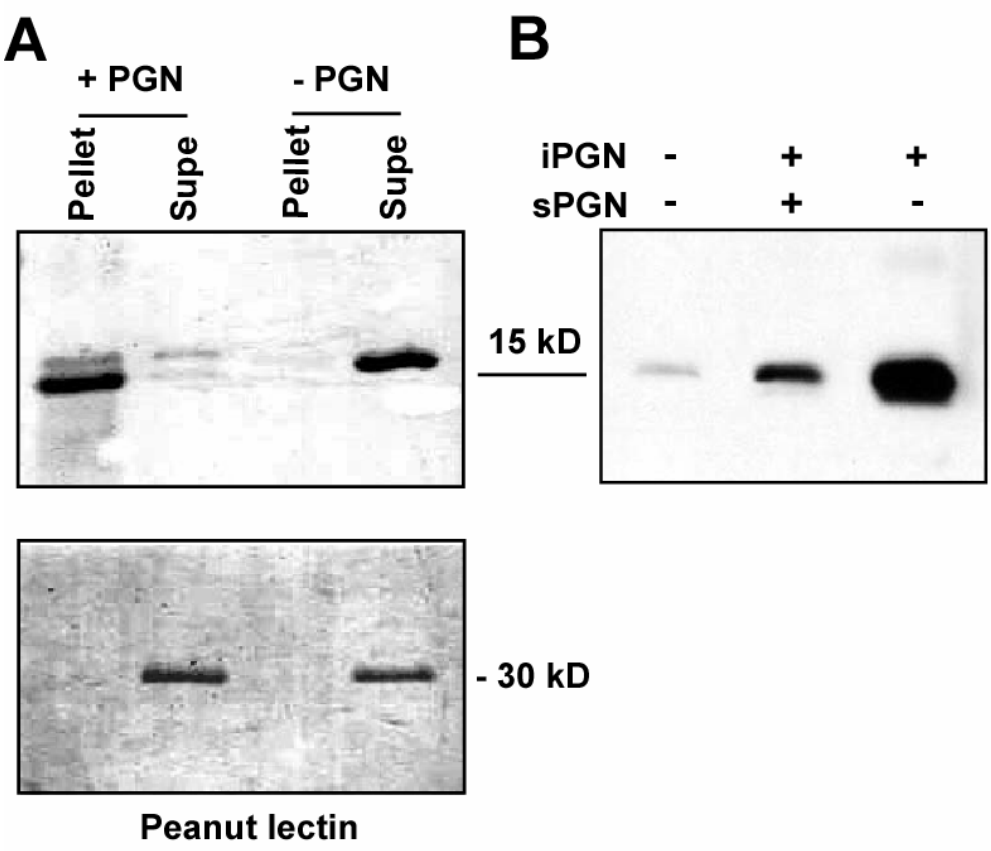

Figure 6.4: RegIIIβ binds to peptidoglycan.

(A) PGN pulldown assays were performed by adding 10 μg of purified recombinant RegIIIβ or peanut lectin to 50 μg of insoluble *Bacillus subtilis* PGN and pelleted. Pellet (P) and supernatant (S) fractions were analyzed by SDS-PAGE and Coomassie Blue staining. **(B)** RegIIIβ binding to insoluble peptidoglycan is competed by soluble PGN. Pull-down assays were performed with 10 μg of purified recombinant lectins, 2 μg of insoluble *B. subtilis* PGN, and with or without 100 μM soluble *B. subilis* PGN as indicated.

RegIIIβ is a member of the Type VII family of C-type lectins according to the Drickamer classification [63]. Because it is primarily composed of a carbohydrate recognition domain, we predicted that it interacts with PGN through the carbohydrate backbone. When RegIIIβ was incubated with a variety of polysaccharides coupled to Sepharose beads, it specifically bound to both mannan and chitin, two GlcNAc polysaccharides whose chemical structures mimic that of PGN without the peptide crossbridges (Figure 6.5). This binding is similar to that described for other C-type

lectins, particularly wheat germ agglutinin (WGA) and mannose binding lectin. In contrast to mannose binding lectin, the observed binding was independent of the presence of calcium, which has been described for other lectins [64, 162]. Additionally, in contrast to WGA binding characteristics, RegIIIβ bound to mannan oligosaccharides as well as chitin, revealing a carbohydrate binding pattern distinct from WGA. This interaction between RegIIIβ and mannan had been previously described by comparing the elution profile of mannose-Sepharose and mannan-Sepharose, on which the entire population of RegIIIβ protein was retarded (Figure 4.7). Finally, no binding was observed to dextran-sepharose, a β-1,3 linked polymer of glucose, or to uncoupled Sepharose, demonstrating that the conditions under which the assay was performed allowed for specific interactions (Figure 6.5).

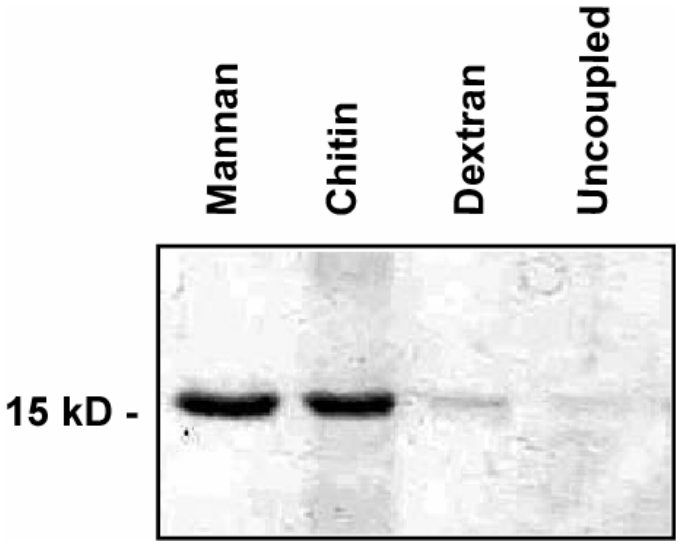

Figure 6.5: RegIIIβ binds to carbohydrates that mimic PGN.

(A) 50 μg of RegIIIβ was bound to 25 μL of immobilized polysaccharides for 2 hours at 4°C. Bound proteins were analyzed by electrophoresis through a 15% polyacrylamide gel followed by staining with Coomassie brilliant blue.

RegIIIβ binds intestinal bacteria

The intestine harbors both Gram positive and Gram negative organisms, as well as fungi and viruses. Because PGN is exposed on the surface of Gram positive organisms, we predicted that RegIIIβ could bind to a sub-population of intestinal bacteria. Using recombinant RegIIIβ directly conjugated to AlexaFluor 488 (Beta-FITC), we used FACS analysis to determine if RegIIIβ labeled intestinal bacteria. As shown in Figure 6.6A, Beta-FITC labeled ~40% of gated events, a percentage significantly above the background binding demonstrated by BSA-FITC (~3%). Interestingly, when intestinal bacteria were co-stained with Beta-FITC and Gamma-Cy3, again ~40% of gated events were double-positive, indicating that RegIIIγ binds the same sub-population of bacteria as RegIIIβ (Figure 6.6B). In addition, a percentage of the analyzed bacteria were labeled with only RegIIIβ. This confirms the results of labeling with either RegIIIβ or RegIIIγ alone, since RegIIIβ bound twice as many intestinal bacteria as did RegIIIγ (Fig. 6.6B). Therefore, RegIIIβ binds to bacteria isolated from intestinal conditions, but appears to have a wider range of microbial targets.

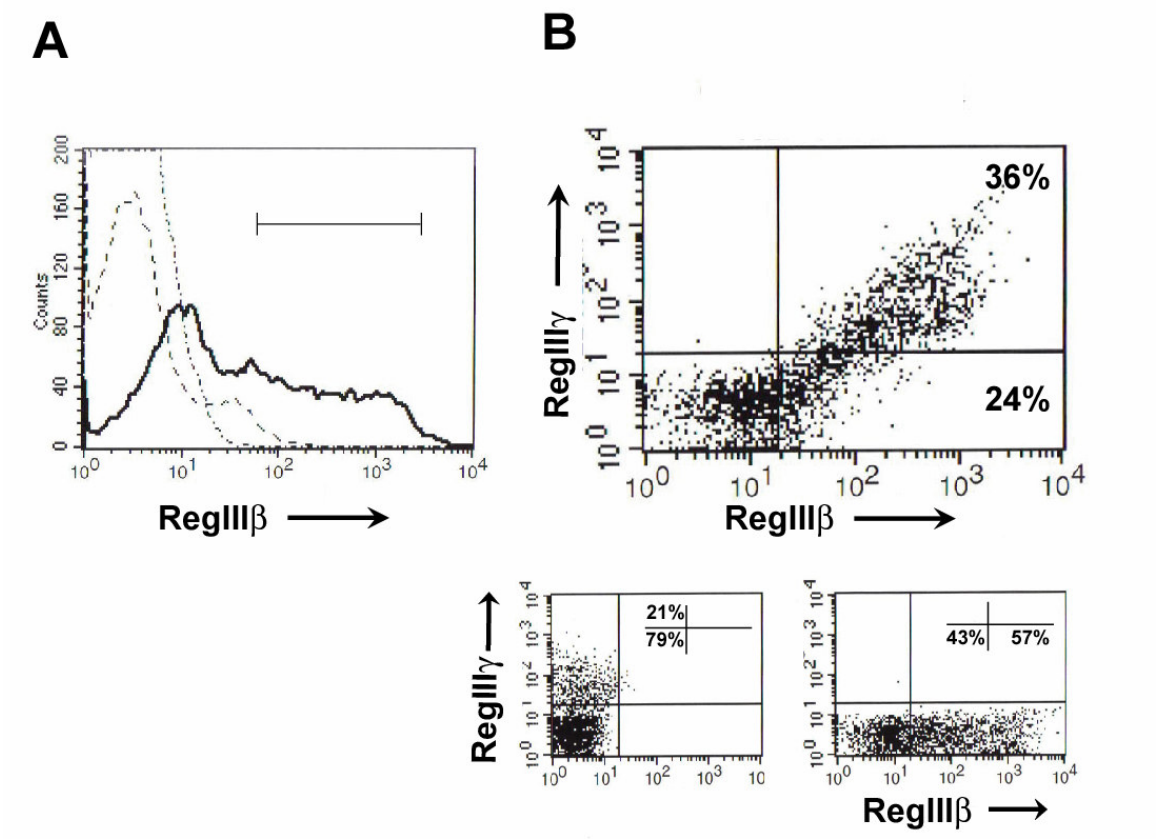

Figure 6.6: RegIIIβ and RegIIIγ bind the same population of intestinal bacteria.

(A) RegIIIβ binds to intestinal bacteria. FITC-RegIIIβ was incubated with mixed bacteria isolated from *in vivo* conditions as described in Materials and methods. Bacteria bound by RegIIIβ are indicated by a shift along the x-axis. (Buffer = dashed line, FITC-BSA = broken line, FITC-RegIIIβ = solid line). **(B)** RegIIIβ and RegIIIγ bind the same population of intestinal bacteria. FITC-RegIIIβ and Cy3-RegIIIγ were co-incubated with intestinal bacteria, and the fluorescence of bacterial cells was monitored by FACS. The lower panels display percentages of bacteria bound by each protein alone. The data displayed are representative of results obtained with 3 mice.

Expression of RegIIIβ increases in conjunction with bacterial load and is induced by intestinal bacteria

While the large intestine harbors approximately 10^{14} bacterial cells, the bacterial load in the small intestine is lower, but increases in the distal ileum as compared to the duodenum [163]. Therefore, we predicted that the expression of RegIIIβ, which we

demonstrated to have antimicrobial properties, would increase in conjunction with elevated numbers of luminal bacteria. Using real-time PCR analysis, we determined that RegIIIβ transcript levels do increase to their highest levels in the most distal small intestinal segment, which borders the ileo-cecal junction in mice (Figure 6.7A). This pattern is not recapitulated in germ-free mice of the same strain, indicating that RegIIIβ expression is attributable to the presence of bacteria. The increase in RegIIIβ transcript from the proximal to distal end of the small intestine was also confirmed by performing *in situ* hybridization on a cross-section of the entire length of small intestine. Distal small intestinal tissues demonstrated a marked increase in RegIIIβ probe hybridization as compared to proximal tissues, while still maintaining crypt expression (Figure 6.7A).

Furthermore, we wished to investigate the signals responsible for eliciting the increase in RegIIIβ expression upon colonization with a full microflora. In order to do so, we colonized germ-free mice with individual bacterial species, *B. theta* (Gram negative) or *L. innocua* (Gram positive) and assayed for an increase in RegIIIβ transcript levels using real-time PCR. As displayed in Figure 6.7B, single colonization with either a Gram positive or Gram negative organism was not sufficient to induce expression of RegIIIβ to the levels seen in mice harboring a full microbial flora. However, in RAG -/- mice, which do not produce any IgA, colonization with *B. theta* or *L. innocua* is sufficient to restore RegIIIβ transcript levels to those seen in RAG mice harboring a full microflora (Figure 6.7B). These results suggest that both Gram positives and Gram negatives are capable of inducing this bactericidal protein under conditions where bacterial-epithelial contact may be increased.

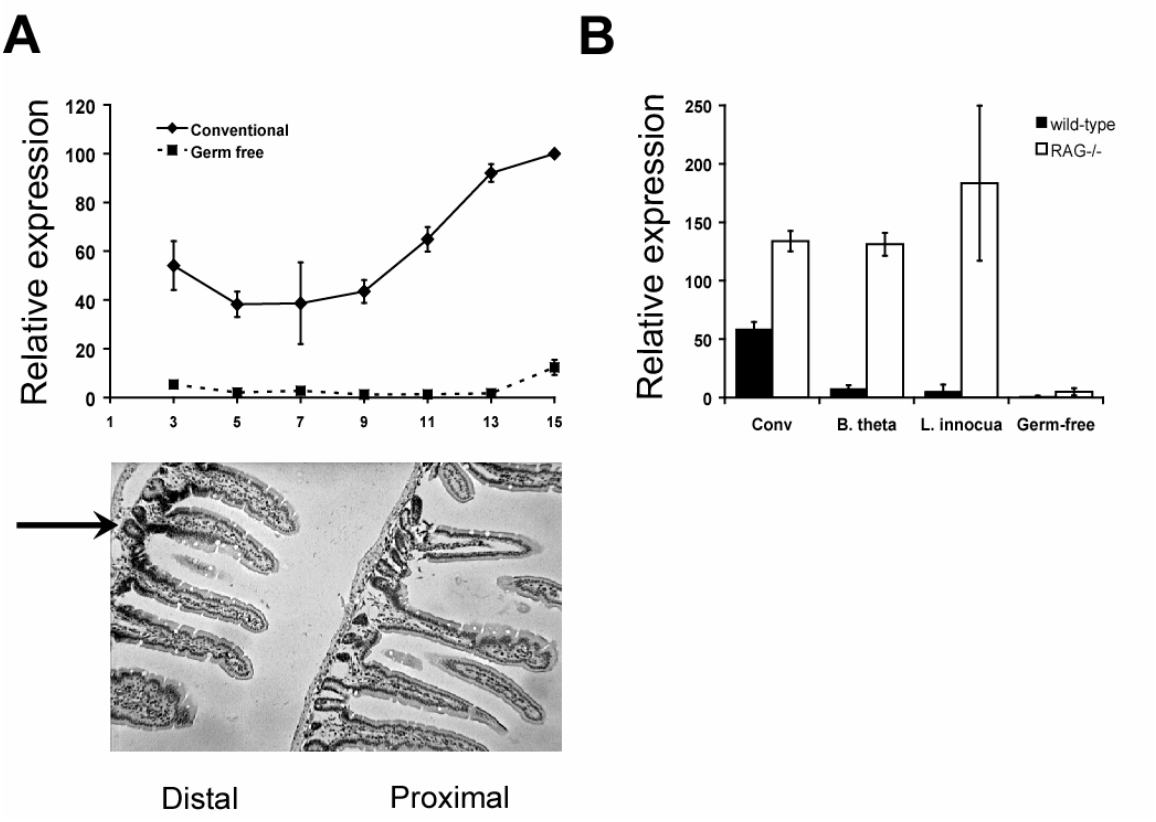

Figure 6.7: RegIIIβ expression is induced by bacterial colonization.

(A) Upper panel: RegIIIβ expression along the cephalocaudal axis of the small intestine. Small intestines from adult germ-free or conventionalized (conv-D) NMRI mice were divided into 16 equal segments and *RegIIIβ* mRNA expression was determined in specific segments using Q-PCR. Results are representative of experiments in two sets of mice. Lower panel: *In situ* hybridization performed on a cross-section through a whole intestinal "roll". Cells exhibiting increased RegIIIβ transcript are marked by black staining using phase contrast microscopy (arrow). **(B)** Q-PCR determinations were done on cDNAs from mid-small intestine. Each point represents the average value from ≥3 different mice. All Q-PCR determinations in the figure were performed in triplicate (mean ± s.d. plotted) and were normalized to 18S rRNA. *Courtesy of Cecilia Whitham*.

C. Discussion

Many studies have highlighted a potential role for the RegIII proteins in intestinal homeostasis, since members of this family are upregulated under immunocompromised or inflammatory conditions in the intestines of mice and humans [117, 118].

Interestingly, the genomes of mice encode four known RegIII proteins, while only two RegIII proteins have been identified in humans, which is suggestive of the rapid diversification seen in families of proteins involved in innate defense [47, 164]. In support of their role in innate immunity, the work detailed in the previous chapter has demonstrated that peptidoglycan is the ligand recognized by RegIIIγ and HIP/PAP, and that both proteins are directly bactericidal. Here we present the expression profile, purification, and characterization of the closely related mouse family member, RegIIIβ, also demonstrating that it binds to intestinal bacteria by interacting with exposed peptidoglycan. Such research is invaluable in gaining a more comprehensive understanding of the defenses employed at the intestinal mucosal barrier and in designing strategies for targeted deletions of the *RegIII* genes in mice.

The mRNA expression studies of *RegIIIβ* indicate that its transcript also is regulated by colonization of the intestine with a full microflora, which confirms the results of the microarray screen performed comparing colonized versus germ-free Paneth cells (Introduction). We further characterized the expression of this RegIII transcript by determining its tissue expression pattern. Similar to its closely related family member, RegIIIγ, RegIIIβ transcript localized primarily to the small intestine, particularly the distal region. In contrast to previous research characterizing their expression [104], we observed no expression of either transcript in the pancreas. However, this could be due to the differences in housing facilities of the mice used, where exposure to more pathogens could elicit expression of these proteins in other tissues.

Within the intestine, the expression of RegIIIβ and RegIIIγ appears to be confined to cells of the intestinal crypts, confirming the previously reported localization of RegIIIγ

to the granules of Paneth cells. The use of *in situ* hybridization allowed us to distinguish between these two highly homologous transcripts, since the use of an antibody against either protein target would result in cross-reaction with any present RegIII proteins. In fact, the immunogold localization of RegIIIγ to Paneth cell granules presented in Chapter 5 could have also detected RegIIIβ present in those cells. Other epithelial cells lining the crypts could be secreting these proteins via the classical secretion pathway, since both proteins contain N-terminal secretion sequences. Alternatively, the transcript alone may be expressed but not translated by cells other than Paneth cells, although this seems unlikely due to the high levels of expression observed in all cell types producing RegIII transcript. Comprehensively, these data confirm previous observations that RegIII transcripts and proteins are produced and secreted by cells lining intestinal crypts, suggesting that these effector proteins function within the crypt microenvironment.

As a member of the type VII C-type lectin family, RegIIIβ binds to highly structured carbohydrates with similar chemical structures, including mannan, chitin, and bacterial peptidoglycan. These carbohydrates all contain b-1,4-linked mannose or GlcNAc residues, which differ from each other only by the position of the 3-carbon hydroxyl group. This pattern of polysaccharide binding is consistent with that observed for the other RegIII family members, RegIIIγ and HIP/PAP. However, the affinities of the murine RegIII proteins are lower than that calculated for human HIP/PAP, suggesting that the two RegIII proteins may cooperate in the mouse intestine to increase binding affinity to exposed peptidoglycan on target bacteria (data not shown). Future experiments combining both proteins in the microbial killing assay, affinity assay, and binding assay

would serve to address this hypothesis, and provide insights into the diversification of this family in mice.

The flow cytometry data presented in this chapter indicate that recombinant RegIIIβ, in addition to binding the same bacteria recognized by RegIIIγ, also binds to an additional subpopulation of bacteria. Because the population of bacteria bound by RegIIIγ has been previously confirmed to be primarily Gram positive, the dual staining of both proteins indirectly confirms the binding of RegIIIβ to Gram positive organisms. The previously described pull down experiments using peptidoglycan also lend support to the hypothesis that RegIIIβ, similar to RegIIIγ, binds to exposed PGN on the surfaces of Gram positive organisms. However, RegIIIβ also recognizes a population of bacteria distinct from that of RegIIIγ. Determining the composition of this alternate bacterial population through further experimentation will provide an enhanced understanding of the range of bacteria targeted by the RegIII protein family.

Although no direct evidence exists for the ability of RegIIIβ to directly target bacteria, three lines of evidence suggest that this protein will perform such a function. Firstly, this protein's primary amino acid sequence is highly homologous to both RegIIIγ and HIP/PAP, and the mRNA transcript displays the same tissue localization as that of RegIIIγ. Second, we have identified PGN as the biological ligand for RegIIIβ and have demonstrated that this protein binds intestinal bacteria, both characteristics similar to the known antimicrobial protein, RegIIγ (Chapter 5). Finally, the regulation of RegIIIβ expression by the intestinal microflora mimics that described for RegIIIγ. Therefore, we hypothesize that further testing of RegIIIβ will reveal a directly bactericidal function.

However, this protein could target a slightly different population of intestinal bacteria from those killed by its family member, RegIIIγ.

In summary, RegIIIβ represents another C-type lectin family member that uniquely recognizes the glycan structures of PGN. Characterizing the redundancy between the highly homologous RegIIIβ and RegIIIγ is vital to a better understanding of the mechanisms employed by the murine intestine to protect the epithelial barrier. Additionally, this information will assist in the development of a targeted knock-out strategy to better address the function of these proteins *in vivo*. Because these proteins have overlapping functions, generating a targeted deletion of either protein might result in a compensatory upregulation of the transcript of the other protein, masking a potential phenotype. Therefore, future genetic studies should focus on characterizing both single deletions of each protein as well as a dually targeted RegIII locus.

Chapter Seven:

DISCUSSION AND FUTURE DIRECTIONS

The underlying goal of this thesis was to determine the function of a family of C-type lectins whose expression is induced by bacterial colonization of the intestine. Much recent research has focused on describing the interactions between the normal microbial flora of the mammalian intestine and the host cells of the mucosal epithelium [1, 2, 13, 15, 16, 22, 23, 32, 57, 137, 149, 165-168]. Such studies have provided a foundation for understanding the molecular basis of complex pathological conditions that affect millions of patients, such as inflammatory bowel disease and Crohn's disease [42, 45, 112]. However, investigations into mechanisms of immune recognition and homeostasis within the intestine, such as the studies presented here, are required for new insights into the defenses employed to maintain homeostasis in the intestine.

Prior research investigating the function of the RegIII proteins was hampered by the lack of a ready source of pure protein with which to perform functional assays. To address this, we developed an overexpression and purification strategy in *Escherichia coli*, which we have utilized to successfully purify four RegIII proteins to homogeneity. Because members of this protein family contain six conserved cysteine residues, which form three disulfide bonds, they are shuttled to inclusion bodies, and are contained in the insoluble fraction. By isolating inclusion bodies and refolding the denatured recombinant proteins, we were able to generate milligram quantities of properly refolded RegIII proteins (Chapter 3).

Our initial characterizations of these proteins using a variety of biochemical and biophysical techniques determined that they are refolded to a functional conformation. Size exclusion chromatography indicates that none of the recombinant proteins are aggregated. Interestingly, we observed that RegIIIγ and HIP/PAP elute primarily as monomers in solution, while RegIIIβ elutes as a dimer at pH 6. The small fraction of RegIIIγ and HIP/PAP that exist as dimers and the predominance of RegIIIβ dimer suggest that these proteins could cooperatively interact to perform the binding and bactericidal functions identified in this study.

Additionally, we performed circular dichroism to confirm that RegIIIγ and HIP/PAP display spectra consistent with the formation of secondary structure. Indeed, our analysis revealed that both recombinant proteins are primarily composed of β-strands, consistent with the existing crystal structure of HIP/PAP, which shows that this protein is primarily composed of β-strands. More conclusive, however, was our determination that

our refolding procedure yielded proteins that specifically bound to the highly oligomerized polysaccharides described below.

The data presented in this thesis are the first to describe and fully characterize a carbohydrate ligand for the RegIII lectins. Although previous research described binding of GST-tagged HIP/PAP to lactose, we observed no binding of any RegIII protein to this carbohydrate or any other mono- or- disaccharide using our pull-down methods. Instead, we identified interactions of RegIIIβ, RegIIIγ, and HIP/PAP to highly oligomerized polysaccharides, such as mannan and chitin. While the precise chemical structures of these polymers have not been determined, the fundamental monosaccharides composing each are mannose and N-acetyl glucosamine (GlcNAc), respectively. Mannose and GlcNAc have fairly conserved chemical structure, save for the inversion of a hydroxyl group at the 3-carbon position, suggesting that the orientation at this carbon is not necessary for binding of the RegIII proteins. Additionally, they do not bind to cellulose, a polymer of glucose, demonstrating that they do not recognize all β-1,4-linked polysaccharides.

However, the predominant binding interaction that we observed was the specific recognition of PGNs isolated from a variety of species. This binding is specific, high affinity, and correlates with the observed susceptibility of Gram-positive bacteria to RegIII-mediated binding and killing. Because RegIIIβ, RegIIIγ, and HIP/PAP also bind the complex carbohydrates mannan and chitin, which are structural mimics of PGN without the tetrapeptide moiety, our data indicate that these proteins recognize the carbohydrate backbone of PGN. Since previously described PGRPs interact with the

tetrapetide portion of the PGN molecule [100, 101], the RegIII lectin-like binding to the carbohydrate backbone of PGN represents a unique mechanism of PGN recognition.

As evidenced in Figure 7.1 below, incubation of recombinant RegIIIβ, RegIIIγ, or HIP/PAP with insoluble peptidoglycan results in the formation of a lower molecular weight band at ~14 kD. N-terminal sequencing of the HIP/PAP lower band revealed that a trypsin-like cleavage event had occurred, releasing an undecapeptide from the N-terminus of the protein. The Ile-Arg residues between which the digestion occurs are completely conserved across all Reg proteins, and correlates with trypsin cleavage that has been described for rat Reg proteins [147].

Interestingly, after purifying RegIIIα to homogeneity, we observed no binding of this RegIII lectin to any of the carbohydrates recognized by its family members, particularly mannan, chitin, or PGN (Figure 7.1). The expression of this protein is also unique in that it is not regulated by intestinal colonization, nor is the RegIIIα transcript produced to the extent of RegIIIβ or RegIIIγ (data not shown). Strikingly, ClustalW analysis of the amino acid sequences of the four RegIII proteins investigated in this thesis indicates that the human homologue, HIP/PAP, is more closely related to RegIIIβ and RegIIIγ than is RegIIIα (Figure 1.4). Therefore, RegIIIα not only demonstrates that our procedure for purification does not yield proteins with identical carbohydrate ligands, but it also serves as an excellent tool for subsequent structure-function studies of the RegIII family members.

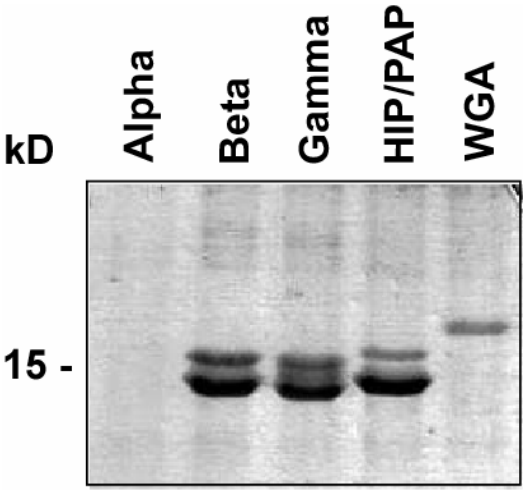

Figure 7.1: Peptidoglycan pull-down assay with all four RegIII proteins.

25 μg of each purified lectin was incubated with PGN isolated from *Bacillus subtilis* for 30 minutes at 4°C. The washed PGN was boiled and bound proteins were separated by electrophoresis through a 15% polyacrylamide gel followed by staining with Coomassie brilliant blue.

The structures of multiple lectins co-crystallized with their ligands have been solved, as well as a preliminary crystal structure of HIP/PAP [64, 69, 130, 135]. The RegIII proteins all form two surface loops, which appear to be similar in structure to the carbohydrate binding loops found on other lectins. As mentioned above, we intend to use the observation that RegIIIα does not bind to PGN in order to begin analyzing the residues required for PGN recognition. As displayed in Figure 7.2, a sequence alignment of the two loop regions predicted to interact with carbohydrates identifies residues that are conserved in the RegIIIs that bind PGN, but are distinct in RegIIIα. Using this information, our future analysis will employ site-directed mutagenesis at these residues, predicting that such changes will ablate ligand binding in RegIIIβ, RegIIIγ, and HIP/PAP, or introduce the ability to bind PGN into RegIIIα. This data and the elucidation of the

crystal structures of these proteins with and without their ligands would be useful in understanding how these lectins bind and coordinate their carbohydrate ligands.

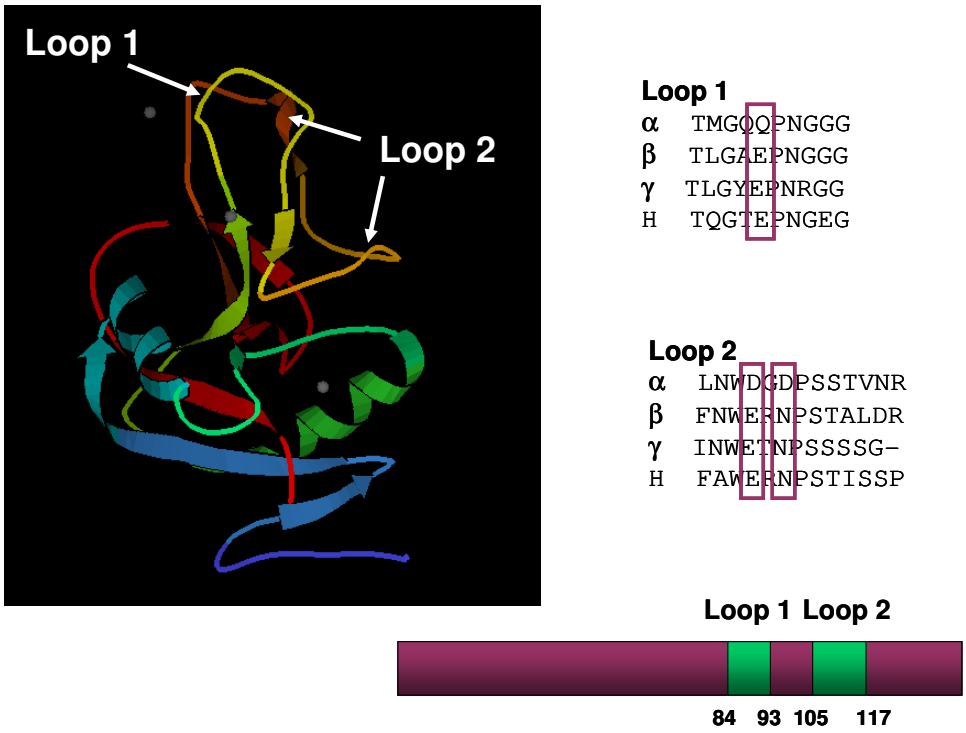

Figure 7.2: Proposed site-directed mutagenesis of the RegIII family using sequence alignments.

At left is the crystal structure for HIP/PAP solved by Abergel, et al. [130] as viewed using the Jmol software on the Protein Data Bank website (www.rcsb.org/pdb), with the two predicted carbohydrate coordinating loops demarcated. At right are the sequence alignments of the two loop regions of the three murine RegIII proteins and HIP/PAP, with residues unique to RegIIIα highlighted in red. The positions of the two loops within the protein are indicated in the schematic at bottom left.

Our results indicate that binding of these proteins to their carbohydrate ligands is not calcium-dependent, which is contrary to the binding requirements of MBL, but is consistent with the crystal structure of the related protein RegIα that lacks a Ca^{2+} binding pocket. Interestingly, RegIII binding to mannan and chitin was inhibited by the addition

of 10-20 mM calcium, as determined by elution of bound RegIIIγ and HIP/PAP from mannan columns by buffer containing this cation. While RegIIIβ did not bind as tightly to mannan and so could not be eluted by buffer containing calcium, the entire population of this protein was retarded by mannan-Sepharose. Such retardation is typical of other C-type lectins, including MBL, some of which demonstrate only millimolar affinities for their carbohydrate ligands [153, 169]. The prevention of recombinant RegIIIγ and HIP/PAP ligand binding by inclusion of this cation represents the first identification of a lectin family inhibited by calcium. Currently, we do not understand the mechanism or relevance of this observed calcium inhibition, but the future determinations of the crystal structures of the RegIII proteins, particularly co-crystallized with ligands, could elucidate the means by which the RegIII proteins coordinate carbohydrates.

Additionally, further studies could be undertaken to identify other potential RegIII carbohydrate ligands and to determine the affinities with which these proteins bind to these various carbohydrate structures. With some technical modifications, the glycoarray approach presented in Chapter 5 could yield alternate carbohydrate ligands for the RegIII family or assist in identifying a ligand for RegIIIα. Initially, the affinities of each protein for small PGN fragments will be assessed using either radioactive binding assays or surface plasmon resonance analysis. Our preliminary analysis of the affinities of RegIIIγ and HIP/PAP for insoluble PGN suggests that these proteins bind with ~20-30 nM affinity (Chapter 5) – a result that we predict will be confirmed in future experiments. The affinities of these proteins for different types of PGNs will also be analyzed using these techniques, perhaps revealing a preference for one PGN type, similar to the PGRP family [100, 101, 160]. Finally, comparisons between the affinities for PGN and other

identified carbohydrate ligands will be determined to confirm the preference for binding to PGN described in Chapter 5. Based on the data outlined in this thesis, we would predict that any other ligands identified for RegIIIβ, RegIIIγ, or HIP/PAP would be complex polysaccharides with structures and linkages similar to those found in both PGN and chitin.

Aside from PGN binding, we have also ascribed a directly bactericidal function to two of the RegIII proteins. This antibacterial activity is specific for Gram-positive bacteria, results in disruption of the bacterial cell wall, and can be inhibited by structured chitin-based oligosaccharides. Thus, we propose a model in which RegIII protein binding to exposed bacterial PGN results in eventual target cell destruction by destabilization of the cell wall (Figure 7.3).

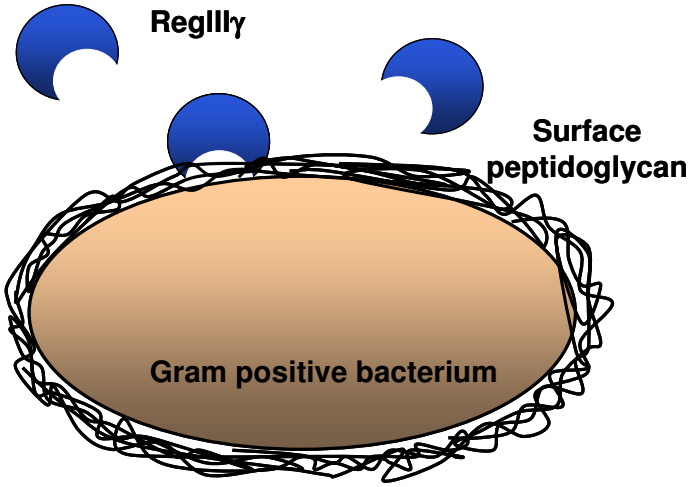

Figure 7.3: Model of RegIII binding to bacterial surface peptidoglycan.

Recognition of exposed peptidoglycan by secreted RegIII CRDs results in cell wall perturbation and lysis.

Our model of RegIII function suggests disruption of the target bacterial cell wall, eventually resulting in lysis. In support of this hypothesis, the electron micrographs of bacteria after exposure to RegIIIγ or HIP/PAP are moderately similar to those taken of *S. aureus* exposed to human β-defensin 3 [161], suggesting that the RegIII proteins could mediate their bactericidal activity by forming pores in the bacterial cell wall. Initial experiments assaying for release of glucose-6-phosphate dehydrogenase indicate that this enzyme is not released by bacteria exposed to RegIII proteins (data not shown). However, RegIII disruption of bacterial cell walls could result in small pores, only allowing for leakage of ions, which would destabilize the osmotic pressure inside the bacterial cell. Alternatively, the RegIII proteins could function as amidases similar to lysozyme, actually cleaving the peptidoglycan coating the outer layer of Gram-positive bacteria. Further experiments monitoring soluble PGN release by fluorescence, radioactivity, or presence of digested fragments will be required to delineate the precise mechanism of RegIII-mediated bacterial cell wall degradation.

Because the RegIII proteins are classified as type VII lectins (secreted CRDs), they may represent a more primitive form of lectin-mediated innate immunity. Other well-characterized lectins such as MBL or the surfactant proteins contain collagenous N-terminal domains that allow for multimerization or recruitment of additional effector proteins. However, the RegIII proteins lack additional domains that would allow for the formation of multimers. This observation led to our initial hypothesis regarding RegIII function as bactericidal molecules, and the data presented in chapter five demonstrate that RegIIIγ and HIP/PAP display bactericidal activity. Not only do these results reveal a

unique function for C-type lectins, but they also suggest a novel form of lectin-mediated innate immunity that does not require additional effector proteins.

The observation that these small lectins solely recognize highly structured polysaccharides lends itself well to a model of self and non-self discrimination. Each CRD of other well-characterized lectins, such as mannose binding lectin or the surfactant proteins, can interact with mono- or disaccharide carbohydrates on the surfaces of target cells. As mentioned in the Introduction, these proteins are capable of forming multimers, and this multivalency allows for specific recognition of carbohydrate ligands found on foreign cells [68]. In contrast, the RegIII proteins consist of secreted CRDs, and our results do not conclusively demonstrate the formation of multimers in solution. Therefore, their primary interaction with complex carbohydrate molecules specific to foreign organisms could serve as a mechanism for preventing recognition of polysaccharides found on the exterior of host cells.

The second half of this thesis focused on characterizing the other prominently expressed RegIII in the mouse small intestine, RegIIIβ. In contrast to the human genome that encodes only two RegIII proteins, the mouse genome contains four RegIII coding sequences [106, 108]. To begin investigating the diversification of this family specifically in mice, we compared the mRNA expression patterns of the mouse RegIII transcripts. Of these, we localized RegIIIβ and RegIIIγ mRNA expression predominantly to the small intestine, while RegIIIδ is expressed primarily in the pancreas, and RegIIIα has a more broad tissue distribution, including the lung (Chapter 6 and data not shown). Additionally, by modifying the purification scheme used for RegIIIγ and HIP/PAP and

utilizing a variety of biochemical assays, we were able to demonstrate that RegIIIβ is a lectin that binds to intestinal bacteria by recognizing exposed peptidoglycan.

While both mouse RegIIIβ and RegIIIγ exhibit similar binding preferences, they are not entirely redundant. Our flow cytometric analysis revealed that RegIIIβ binds a subpopulation of intestinal bacteria distinct from the subset co-bound with RegIIIγ, suggesting that this protein could be more promiscuous in its bacterial targets. In addition to having targets distinct from each other, these two proteins could also act cooperatively to confer increased affinity and perhaps specificity for shared targets, both of which hypotheses would provide a molecular explanation for the diversification of this family. The latter is preliminarily supported by data demonstrating that both RegIIIβ and RegIIIγ bind to the same population of intestinal bacteria, and that the affinities of both of these proteins for PGN are decreased compared to that of HIP/PAP.

The studies summarized above, regarding RegIIIβ expression and binding, suggest that this RegIII protein may also exhibit directly bactericidal activity similar to its closely related family members, RegIIIγ and HIP/PAP. Although no experimental evidence to date confirms this hypothesis, two sets of data prompted us to hypothesize a bactericidal function. Firstly, the expression patterns of RegIIIβ are similar to those reported in this thesis for RegIIIγ. Additionally, the binding function that we have identified for RegIIIβ mimics that described for both RegIIIγ and HIP/PAP. Since this protein transcript is upregulated in response to intestinal colonization and the protein binds intestinal bacteria through their peptidoglycan, we propose a directly bactericidal function for this protein. This activity may be targeted against a unique population of bacteria from those susceptible to RegIIIγ, an idea preliminarily supported by the larger

subset of bacteria bound by this protein. Alternatively, both RegIIIβ and RegIIIγ could act in concert to facilitate more swift or diverse bacterial killing.

The combination of expression and functional data presented in this thesis suggests a model whereby the presence of molecular patterns from commensal bacteria within the lumen triggers production of the RegIII proteins. While the precise signals governing expression of these proteins has not been characterized, preliminary data supporting this hypothesis indicate that host cellular sensing of intestinal microbes is MyD88-dependent, implicating Toll-like receptor signaling pathways (S. Vaishnava, unpublished observations). Correspondingly, under germ-free conditions, no molecular patterns are present to induce RegIII production. Conversely, mice deficient in secretory IgA ($RAG^{-/-}$ mice) or bacteria with increased ability to adhere or penetrate the epithelium elicit a more robust RegIII response in order to maintain mucosal barrier integrity (Figure 7.4). Therefore, we propose that the RegIII lectins exert their bactericidal effects on microorganisms neighboring the epithelial surface in order to prevent adherence to or invasion of the intestinal epithelium.

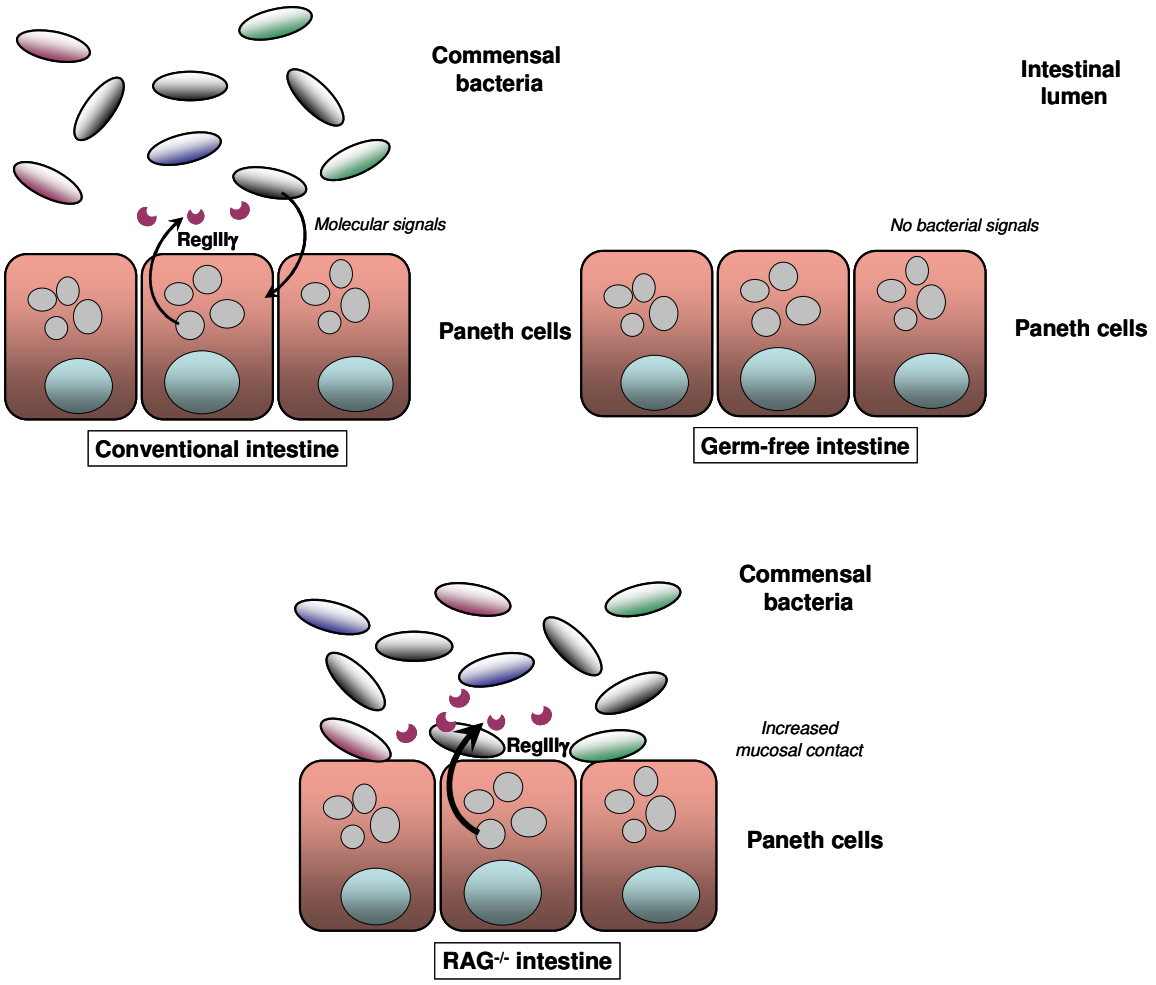

Figure 7.4: RegIII expression is triggered by intestinal bacteria.

Top left: Molecular signals released by intestinal bacteria induce RegIII production and secretion by Paneth cells. **Top right:** In the absence of intestinal colonization (germ-free conditions), no bacterial signals are present to drive RegIII expression. **Bottom:** Immunocompromised mice ($RAG^{-/-}$ mice - lack secretory IgA) or highly invasive bacteria drive increased RegIII production in order to maintain the integrity of the mucosal epithelium.

In conclusion, the newly identified bactericidal function of the RegIII proteins may play an important role in maintaining the integrity of the intestinal epithelial barrier. We initially observed that the transcripts of two RegIII proteins, RegIIIβ and RegIIIγ, were increased in response to colonization of the intestine with a full microflora. Subsequently, we were able to overexpress and purify four RegIII lectins and begin to characterize their binding to carbohydrate ligands (Chapter 4). Our studies were the first to identify PGN as the carbohydrate ligand of the RegIII lectins, and to describe the bactericidal function of these proteins (Chapters 5 and 6). Finally, our research has allowed us to gain a more comprehensive understanding of the regulation of these proteins by intestinal bacteria (Chapters 5 and 6). Using the results of these studies, we have now developed a model of both RegIII binding to bacterial targets and the molecular signals responsible for triggering RegIII expression. Through biochemical and expression analyses, we have initiated an investigation of the ligands and function of the RegIII proteins, which has yielded novel insights into the crosstalk between the intestinal epithelium and the luminal microflora.

Appendix - FoldIt screen results for the murine RegIII proteins

Fold It Screen Results - Alpha											
Buffer	**Supe**	**Pellet**	**Total Protein**	**% folded**	**Signif.**	**Buffer**	**Salt**	**Guan.**	**Cat/Che**	**Additives**	**Redox**
1	0.03	0.19	0.22	**13.67**	NO						
2	0.04	0.47	0.51	**8.06**	NO						
3	0.06	0.38	0.44	**13.88**	NO						
4	0.05	0.15	0.20	**25.84**	YES	Tris	264, 11	no	Mg, Ca	Suc, Arg	DTT
5	0.04	0.19	0.23	**18.08**	YES	MES	264,11	no	Mg, Ca	Suc	GSH, GSSG
6	0.03	0.23	0.26	**11.67**	NO						
7	0.07	0.06	0.13	**50.88**	YES	Tris	10.56,0.44	550	Mg, Ca	Arg	DTT
8	0.08	0.19	0.27	**30.21**	YES	MES	264, 11	no	EDTA	Arg	GSH, GSSG
9	0.02	0.20	0.23	**10.55**	NO						
10	0.03	0.34	0.37	**8.17**	NO						
11	0.03	#NUM!			NO						
12	0.05	0.24	0.29	**16.47**	YES	MES	264, 11	550	EDTA	Arg	DTT
13	0.01	0.25	0.26	**3.24**	NO						
14	0.00	0.28	0.28	**-1.58**	NO						
15	0.04	0.23	0.27	**13.76**	NO						
16	0.13	0.60	0.73	**17.62**	YES	Tris	264, 11	550	Mg, Ca	Suc, Arg	GSH, GSSG

Fold It Screen Results - Beta

Buffer	Supe	Pellet	Total Protein	% folded	Signif.	Buffer	Salt	Guan.	Cat/Che	Additives	Redox
1	0.01	0.08	0.09	**11.91**	NO						
2	0.03	0.16	0.18	**14.55**	NO						
3	0.07	0.04	0.12	**64.48**	YES	MES	10.56, 0.44	550	EDTA	Suc, Arg	GSH, GSSG
4	0.04	0.03	0.06	**55.06**	YES	Tris	264, 11	no	Mg, Ca	Suc, Arg	DTT
5	0.02	0.09	0.11	**21.54**	NO						
6	0.03	0.03	0.06	**42.20**	YES	Tris	10.56, 0.44	550	EDTA	Suc	DTT
7	0.07	0.02	0.09	**76.61**	YES	Tris	10.56,0.44	550	Mg, Ca	Arg	DTT
8	0.04	0.12	0.16	**24.28**	NO						
9	0.02	0.09	0.11	**15.07**	NO						
10	0.01	0.29	0.30	**4.81**	NO						
11	0.03	0.03	0.06	**52.57**	YES	Tris	10.56, 0.44	no	Mg, Ca	Arg	GSH, GSSG
12	0.04	0.37	0.40	**8.66**	NO						
13	0.02	0.05	0.07	**32.56**	YES	Tris	264, 11	550	EDTA	none	GSH, GSSG
14	0.02	0.23	0.25	**8.92**	NO						
15	0.02	0.09	0.12	**20.79**	NO						
16	0.06	0.03	0.08	**69.26**	YES	Tris	264, 11	550	Mg, Ca	Suc, Arg	GSH, GSSG

BIBLIOGRAPHY

1. Stappenbeck, T.S., L.V. Hooper, and J.I. Gordon, *Developmental regulation of intestinal angiogenesis by indigenous microbes via Paneth cells.* Proc Natl Acad Sci U S A, 2002. **99**(24): p. 15451-5.

2. Hooper, L.V., et al., *Molecular analysis of commensal host-microbial relationships in the intestine.* Science, 2001. **291**(5505): p. 881-4.

3. Savage, D.C., *Microbial ecology of the gastrointestinal tract.* Annu Rev Microbiol, 1977. **31**: p. 107-33.

4. Guyton, A.C.a.J.E.H., *Textbook of Medical Physiology*. 10th ed. 2000, Philadelphia: W. B. Saunders.

5. Blaser, M.J., *Hypotheses on the pathogenesis and natural history of Helicobacter pylori-induced inflammation.* Gastroenterology, 1992. **102**(2): p. 720-7.

6. Marshall, B.J. and J.R. Warren, *Unidentified curved bacilli in the stomach of patients with gastritis and peptic ulceration.* Lancet, 1984. **1**(8390): p. 1311-5.

7. Hao, W.L. and Y.K. Lee, *Microflora of the gastrointestinal tract: a review.* Methods Mol Biol, 2004. **268**: p. 491-502.

8. Bjerknes, M. and H. Cheng, *The stem-cell zone of the small intestinal epithelium. I. Evidence from Paneth cells in the adult mouse.* Am J Anat, 1981. **160**(1): p. 51-63.

9. Leedham, S.J., et al., *Intestinal stem cells.* J Cell Mol Med, 2005. **9**(1): p. 11-24.

10. Ouellette, A.J., *Defensin-mediated innate immunity in the small intestine.* Best Pract Res Clin Gastroenterol, 2004. **18**(2): p. 405-19.

11. Ismail, A.S. and L.V. Hooper, *Epithelial cells and their neighbors. IV. Bacterial contributions to intestinal epithelial barrier integrity.* Am J Physiol Gastrointest Liver Physiol, 2005. **289**(5): p. G779-84.

12. Reyniers, J.A., *The production and use of germ-free animals in experimental biology and medicine.* Am J Vet Res, 1957. **18**(68): p. 678-87.

13. Jiang, H.Q., et al., *Interactions of commensal gut microbes with subsets of B- and T-cells in the murine host.* Vaccine, 2004. **22**(7): p. 805-11.

14. Wostmann, B.S., et al., *Dietary intake, energy metabolism, and excretory losses of adult male germfree Wistar rats.* Lab Anim Sci, 1983. **33**(1): p. 46-50.

15. Sonnenburg, J.L., et al., *Glycan foraging in vivo by an intestine-adapted bacterial symbiont.* Science, 2005. **307**(5717): p. 1955-9.

16. Rhee, K.J., et al., *Role of commensal bacteria in development of gut-associated lymphoid tissues and preimmune antibody repertoire.* J Immunol, 2004. **172**(2): p. 1118-24.

17. Lambolez, F., et al., *Characterization of T cell differentiation in the murine gut.* J Exp Med, 2002. **195**(4): p. 437-49.

18. Oida, T., et al., *Role of gut cryptopatches in early extrathymic maturation of intestinal intraepithelial T cells.* J Immunol, 2000. **164**(7): p. 3616-26.

19. Suzuki, K., et al., *Gut cryptopatches: direct evidence of extrathymic anatomical sites for intestinal T lymphopoiesis.* Immunity, 2000. **13**(5): p. 691-702.

20. Rescigno, M., et al., *Dendritic cells express tight junction proteins and penetrate gut epithelial monolayers to sample bacteria.* Nat Immunol, 2001. **2**(4): p. 361-7.

21. Macpherson, A.J., et al., *A primitive T cell-independent mechanism of intestinal mucosal IgA responses to commensal bacteria.* Science, 2000. **288**(5474): p. 2222-6.

22. Macpherson, A.J. and N.L. Harris, *Interactions between commensal intestinal bacteria and the immune system.* Nat Rev Immunol, 2004. **4**(6): p. 478-85.

23. Macpherson, A.J. and T. Uhr, *Induction of protective IgA by intestinal dendritic cells carrying commensal bacteria.* Science, 2004. **303**(5664): p. 1662-5.

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

25. Medzhitov, R. and C. Janeway, Jr., *Innate immune recognition: mechanisms and pathways.* Immunol Rev, 2000. **173**: p. 89-97.

26. Belvin, M.P. and K.V. Anderson, *A conserved signaling pathway: the Drosophila toll-dorsal pathway.* Annu Rev Cell Dev Biol, 1996. **12**: p. 393-416.

27. Lemaitre, B., et al., *The dorsoventral regulatory gene cassette spatzle/Toll/cactus controls the potent antifungal response in Drosophila adults.* Cell, 1996. **86**(6): p. 973-83.

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

29. Medzhitov, R., P. Preston-Hurlburt, and C.A. Janeway, Jr., *A human homologue of the Drosophila Toll protein signals activation of adaptive immunity.* Nature, 1997. **388**(6640): p. 394-7.

30. Poltorak, A., et al., *Defective LPS signaling in C3H/HeJ and C57BL/10ScCr mice: mutations in Tlr4 gene.* Science, 1998. **282**(5396): p. 2085-8.

31. Medzhitov, R., et al., *MyD88 is an adaptor protein in the hToll/IL-1 receptor family signaling pathways.* Mol Cell, 1998. **2**(2): p. 253-8.

32. Rakoff-Nahoum, S., et al., *Recognition of commensal microflora by toll-like receptors is required for intestinal homeostasis.* Cell, 2004. **118**(2): p. 229-41.

33. Takeuchi, O., K. Hoshino, and S. Akira, *Cutting edge: TLR2-deficient and MyD88-deficient mice are highly susceptible to Staphylococcus aureus infection.* J Immunol, 2000. **165**(10): p. 5392-6.

34. Akira, S. and K. Takeda, *Toll-like receptor signalling.* Nat Rev Immunol, 2004. **4**(7): p. 499-511.

35. Ortega-Cava, C.F., et al., *Strategic compartmentalization of Toll-like receptor 4 in the mouse gut.* J Immunol, 2003. **170**(8): p. 3977-85.

36. Inohara, N., et al., *Nod1, an Apaf-1-like activator of caspase-9 and nuclear factor-kappaB.* J Biol Chem, 1999. **274**(21): p. 14560-7.

37. Ogura, Y., et al., *Nod2, a Nod1/Apaf-1 family member that is restricted to monocytes and activates NF-kappaB.* J Biol Chem, 2001. **276**(7): p. 4812-8.

38. Girardin, S.E., et al., *Nod2 is a general sensor of peptidoglycan through muramyl dipeptide (MDP) detection.* J Biol Chem, 2003. **278**(11): p. 8869-72.

39. Viala, J., et al., *Nod1 responds to peptidoglycan delivered by the Helicobacter pylori cag pathogenicity island.* Nat Immunol, 2004. **5**(11): p. 1166-74.

40. Ogura, Y., et al., *Expression of NOD2 in Paneth cells: a possible link to Crohn's ileitis.* Gut, 2003. **52**(11): p. 1591-7.

41. Bonen, D.K., et al., *Crohn's disease-associated NOD2 variants share a signaling defect in response to lipopolysaccharide and peptidoglycan.* Gastroenterology, 2003. **124**(1): p. 140-6.

42. Ogura, Y., et al., *A frameshift mutation in NOD2 associated with susceptibility to Crohn's disease.* Nature, 2001. **411**(6837): p. 603-6.

43. Kobayashi, K.S., et al., *Nod2-dependent regulation of innate and adaptive immunity in the intestinal tract.* Science, 2005. **307**(5710): p. 731-4.

44. Wehkamp, J., et al., *NOD2 (CARD15) mutations in Crohn's disease are associated with diminished mucosal alpha-defensin expression.* Gut, 2004. **53**(11): p. 1658-64.

45. Wehkamp, J., et al., *Reduced Paneth cell alpha-defensins in ileal Crohn's disease.* Proc Natl Acad Sci U S A, 2005. **102**(50): p. 18129-34.

46. Ayabe, T., et al., *Secretion of microbicidal alpha-defensins by intestinal Paneth cells in response to bacteria.* Nat Immunol, 2000. **1**(2): p. 113-8.

47. Ouellette, A.J. and M.E. Selsted, *Paneth cell defensins: endogenous peptide components of intestinal host defense.* Faseb J, 1996. **10**(11): p. 1280-9.

48. Porter, E.M., et al., *The multifaceted Paneth cell.* Cell Mol Life Sci, 2002. **59**(1): p. 156-70.

49. Sherman, M.P., et al., *Paneth cells and antibacterial host defense in neonatal small intestine.* Infect Immun, 2005. **73**(9): p. 6143-6.

50. Fleming, A., *On a remarkable bacteriolytic element found in tissues and secretions.* Proc R Soc London B Ser, 1922. **93**: p. 306-317.

51. Bera, A., et al., *Why are pathogenic staphylococci so lysozyme resistant? The peptidoglycan O-acetyltransferase OatA is the major determinant for lysozyme resistance of Staphylococcus aureus.* Mol Microbiol, 2005. **55**(3): p. 778-87.

52. Jones, D.E. and C.L. Bevins, *Defensin-6 mRNA in human Paneth cells: implications for antimicrobial peptides in host defense of the human bowel.* FEBS Lett, 1993. **315**(2): p. 187-92.

53. Porter, E.M., et al., *Localization of human intestinal defensin 5 in Paneth cell granules.* Infect Immun, 1997. **65**(6): p. 2389-95.

54. Porter, E.M., et al., *Broad-spectrum antimicrobial activity of human intestinal defensin 5.* Infect Immun, 1997. **65**(6): p. 2396-401.

55. Bevins, C.L., *Paneth cell defensins: key effector molecules of innate immunity.* Biochem Soc Trans, 2006. **34**(Pt 2): p. 263-6.

56. Ghosh, D., et al., *Paneth cell trypsin is the processing enzyme for human defensin-5.* Nat Immunol, 2002. **3**(6): p. 583-90.

57. Ayabe, T., et al., *Activation of Paneth cell alpha-defensins in mouse small intestine.* J Biol Chem, 2002. **277**(7): p. 5219-28.

58. Brogden, K.A., *Antimicrobial peptides: pore formers or metabolic inhibitors in bacteria?* Nat Rev Microbiol, 2005. **3**(3): p. 238-50.

59. Drickamer, K., *Two distinct classes of carbohydrate-recognition domains in animal lectins.* J Biol Chem, 1988. **263**(20): p. 9557-60.

60. Florman, H.M. and P.M. Wassarman, *O-linked oligosaccharides of mouse egg ZP3 account for its sperm receptor activity.* Cell, 1985. **41**(1): p. 313-24.

61. Feizi, T. and R.A. Childs, *Carbohydrates as antigenic determinants of glycoproteins.* Biochem J, 1987. **245**(1): p. 1-11.

62. Drickamer, K., *C-type lectin-like domains.* Curr Opin Struct Biol, 1999. **9**(5): p. 585-90.

63. Drickamer, K. and A.J. Fadden, *Genomic analysis of C-type lectins.* Biochem Soc Symp, 2002(69): p. 59-72.

64. Bertrand, J.A., et al., *Crystal structure of human lithostathine, the pancreatic inhibitor of stone formation.* Embo J, 1996. **15**(11): p. 2678-84.

65. Drickamer, K. and R.B. Dodd, *C-Type lectin-like domains in Caenorhabditis elegans: predictions from the complete genome sequence.* Glycobiology, 1999. **9**(12): p. 1357-69.

66. Reid, K.B., *Functional roles of the lung surfactant proteins SP-A and SP-D in innate immunity.* Immunobiology, 1998. **199**(2): p. 200-7.

67. Quesenberry, M.S. and K. Drickamer, *Role of conserved and nonconserved residues in the Ca(2+)-dependent carbohydrate-recognition domain of a rat mannose-binding protein. Analysis by random cassette mutagenesis.* J Biol Chem, 1992. **267**(15): p. 10831-41.

68. Ezekowitz, R.A., *Role of the mannose-binding lectin in innate immunity.* J Infect Dis, 2003. **187 Suppl 2**: p. S335-9.

69. Weis, W.I., K. Drickamer, and W.A. Hendrickson, *Structure of a C-type mannose-binding protein complexed with an oligosaccharide.* Nature, 1992. **360**(6400): p. 127-34.

70. Ng, K.K., K. Drickamer, and W.I. Weis, *Structural analysis of monosaccharide recognition by rat liver mannose-binding protein.* J Biol Chem, 1996. **271**(2): p. 663-74.

71. Sheriff, S., C.Y. Chang, and R.A. Ezekowitz, *Human mannose-binding protein carbohydrate recognition domain trimerizes through a triple alpha-helical coiled-coil.* Nat Struct Biol, 1994. **1**(11): p. 789-94.

72. Kuhlman, M., K. Joiner, and R.A. Ezekowitz, *The human mannose-binding protein functions as an opsonin.* J Exp Med, 1989. **169**(5): p. 1733-45.

73. Ma, Y.G., et al., *Human mannose-binding lectin and L-ficolin function as specific pattern recognition proteins in the lectin activation pathway of complement.* J Biol Chem, 2004. **279**(24): p. 25307-12.

74. Matsushita, M. and T. Fujita, *Activation of the classical complement pathway by mannose-binding protein in association with a novel C1s-like serine protease.* J Exp Med, 1992. **176**(6): p. 1497-502.

75. Garred, P., et al., *Association of mannose-binding lectin gene heterogeneity with severity of lung disease and survival in cystic fibrosis.* J Clin Invest, 1999. **104**(4): p. 431-7.

76. Super, M., et al., *Association of low levels of mannan-binding protein with a common defect of opsonisation.* Lancet, 1989. **2**(8674): p. 1236-9.

77. Schleifer, K.H. and O. Kandler, *Peptidoglycan types of bacterial cell walls and their taxonomic implications.* Bacteriol Rev, 1972. **36**(4): p. 407-77.

78. Dziarski, R., A.J. Ulmer, and D. Gupta, *Interactions of CD14 with components of gram-positive bacteria.* Chem Immunol, 2000. **74**: p. 83-107.

79. Dziarski, R., et al., *Soluble CD14 enhances membrane CD14-mediated responses to peptidoglycan: structural requirements differ from those for responses to lipopolysaccharide.* Infect Immun, 2000. **68**(9): p. 5254-60.

80. Yoshida, H., K. Kinoshita, and M. Ashida, *Purification of a peptidoglycan recognition protein from hemolymph of the silkworm, Bombyx mori.* J Biol Chem, 1996. **271**(23): p. 13854-60.

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

82. Kang, D., et al., *A peptidoglycan recognition protein in innate immunity conserved from insects to humans.* Proc Natl Acad Sci U S A, 1998. **95**(17): p. 10078-82.

83. Dziarski, R., *Peptidoglycan recognition proteins (PGRPs).* Mol Immunol, 2004. **40**(12): p. 877-86.

84. Girardin, S.E. and D.J. Philpott, *Mini-review: the role of peptidoglycan recognition in innate immunity.* Eur J Immunol, 2004. **34**(7): p. 1777-82.

85. Steiner, H., *Peptidoglycan recognition proteins: on and off switches for innate immunity.* Immunol Rev, 2004. **198**: p. 83-96.

86. Zaidman-Remy, A., et al., *The Drosophila amidase PGRP-LB modulates the immune response to bacterial infection.* Immunity, 2006. **24**(4): p. 463-73.

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

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

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

90. Mellroth, P., et al., *Ligand-induced dimerization of Drosophila peptidoglycan recognition proteins in vitro.* Proc Natl Acad Sci U S A, 2005. **102**(18): p. 6455-60.

91. Garver, L.S., J. Wu, and L.P. Wu, *The peptidoglycan recognition protein PGRP-SC1a is essential for Toll signaling and phagocytosis of Staphylococcus aureus in Drosophila.* Proc Natl Acad Sci U S A, 2006. **103**(3): p. 660-5.

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

93. Dziarski, R., et al., *Defect in neutrophil killing and increased susceptibility to infection with nonpathogenic gram-positive bacteria in peptidoglycan recognition protein-S (PGRP-S)-deficient mice.* Blood, 2003. **102**(2): p. 689-97.

94. Lu, X., et al., *Peptidoglycan recognition proteins are a new class of human bactericidal proteins.* J Biol Chem, 2006. **281**(9): p. 5895-907.

95. Tydell, C.C., et al., *Bovine peptidoglycan recognition protein-S: antimicrobial activity, localization, secretion, and binding properties.* J Immunol, 2006. **176**(2): p. 1154-62.

96. Cho, J.H., et al., *Human peptidoglycan recognition protein S is an effector of neutrophil-mediated innate immunity.* Blood, 2005. **106**(7): p. 2551-8.

97. Mellroth, P., J. Karlsson, and H. Steiner, *A scavenger function for a Drosophila peptidoglycan recognition protein.* J Biol Chem, 2003. **278**(9): p. 7059-64.

98. Wang, Z.M., et al., *Human peptidoglycan recognition protein-L is an N-acetylmuramoyl-L-alanine amidase.* J Biol Chem, 2003. **278**(49): p. 49044-52.

99. Chang, C.I., et al., *Structure of tracheal cytotoxin in complex with a heterodimeric pattern-recognition receptor.* Science, 2006. **311**(5768): p. 1761-4.

100. Kumar, S., et al., *Selective recognition of synthetic lysine and meso-diaminopimelic acid-type peptidoglycan fragments by human peptidoglycan recognition proteins I{alpha} and S.* J Biol Chem, 2005. **280**(44): p. 37005-12.

101. Swaminathan, C.P., et al., *Dual strategies for peptidoglycan discrimination by peptidoglycan recognition proteins (PGRPs).* Proc Natl Acad Sci U S A, 2006. **103**(3): p. 684-9.

102. Terazono, K., et al., *A novel gene activated in regenerating islets.* J Biol Chem, 1988. **263**(5): p. 2111-4.

103. Bartoli, C., et al., *A gene homologous to the reg gene is expressed in the human pancreas.* FEBS Lett, 1993. **327**(3): p. 289-93.

104. Narushima, Y., et al., *Structure, chromosomal localization and expression of mouse genes encoding type III Reg, RegIII alpha, RegIII beta, RegIII gamma.* Gene, 1997. **185**(2): p. 159-68.

105. Zhang, Y.W., L.S. Ding, and M.D. Lai, *Reg gene family and human diseases.* World J Gastroenterol, 2003. **9**(12): p. 2635-41.

106. Abe, M., et al., *Identification of a novel Reg family gene, Reg IIIdelta, and mapping of all three types of Reg family gene in a 75 kilobase mouse genomic region.* Gene, 2000. **246**(1-2): p. 111-22.

107. Dusetti, N.J., et al., *Pancreatitis-associated protein I (PAP I), an acute phase protein induced by cytokines. Identification of two functional interleukin-6 response elements in the rat PAP I promoter region.* J Biol Chem, 1995. **270**(38): p. 22417-21.

108. Nata, K., et al., *Molecular cloning, expression and chromosomal localization of a novel human REG family gene, REG III.* Gene, 2004. **340**(1): p. 161-70.

109. Okamoto, H., *The Reg gene family and Reg proteins: with special attention to the regeneration of pancreatic beta-cells.* J Hepatobiliary Pancreat Surg, 1999. **6**(3): p. 254-62.

110. Christa, L., et al., *The human HIP gene, overexpressed in primary liver cancer encodes for a C-type carbohydrate binding protein with lactose binding activity.* FEBS Lett, 1994. **337**(1): p. 114-8.

111. Christa, L., et al., *HIP/PAP is an adhesive protein expressed in hepatocarcinoma, normal Paneth, and pancreatic cells.* Am J Physiol, 1996. **271**(6 Pt 1): p. G993-1002.

112. Ogawa, H., et al., *Increased expression of HIP/PAP and regenerating gene III in human inflammatory bowel disease and a murine bacterial reconstitution model.* Inflamm Bowel Dis, 2003. **9**(3): p. 162-70.

113. Christa, L., et al., *High expression of the human hepatocarcinoma-intestine-pancreas/pancreatic-associated protein (HIP/PAP) gene in the mammary gland of lactating transgenic mice. Secretion into the milk and purification of the HIP/PAP lectin.* Eur J Biochem, 2000. **267**(6): p. 1665-71.

114. Demaugre, F., et al., *HIP/PAP, a C-type lectin overexpressed in hepatocellular carcinoma, binds the RII alpha regulatory subunit of cAMP-dependent protein kinase and alters the cAMP-dependent protein kinase signalling.* Eur J Biochem, 2004. **271**(19): p. 3812-20.

115. Lieu, H.T., et al., *HIP/PAP accelerates liver regeneration and protects against acetaminophen injury in mice.* Hepatology, 2005. **42**(3): p. 618-26.

116. Lasserre, C., et al., *HIP/PAP gene, encoding a C-type lectin overexpressed in primary liver cancer, is expressed in nervous system as well as in intestine and pancreas of the postimplantation mouse embryo.* Am J Pathol, 1999. **154**(5): p. 1601-10.

117. Keilbaugh, S.A., et al., *Activation of RegIIIbeta/gamma and interferon gamma expression in the intestinal tract of SCID mice: an innate response to bacterial colonisation of the gut.* Gut, 2005. **54**(5): p. 623-9.

118. Woo, A.L., et al., *In vivo evidence for interferon-gamma-mediated homeostatic mechanisms in small intestine of the NHE3 Na+/H+ exchanger knockout model of congenital diarrhea.* J Biol Chem, 2002. **277**(50): p. 49036-46.

119. Livesey, F.J., et al., *A Schwann cell mitogen accompanying regeneration of motor neurons.* Nature, 1997. **390**(6660): p. 614-8.

120. Watanabe, T., et al., *Pancreatic beta-cell replication and amelioration of surgical diabetes by Reg protein.* Proc Natl Acad Sci U S A, 1994. **91**(9): p. 3589-92.

121. Bradford, M.M., *A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding.* Anal Biochem, 1976. **72**: p. 248-54.

122. Lobley, A., L. Whitmore, and B.A. Wallace, *DICHROWEB: an interactive website for the analysis of protein secondary structure from circular dichroism spectra.* Bioinformatics, 2002. **18**(1): p. 211-2.

123. Bohm, G., R. Muhr, and R. Jaenicke, *Quantitative analysis of protein far UV circular dichroism spectra by neural networks.* Protein Eng, 1992. **5**(3): p. 191-5.

124. Sreerama, N. and R.W. Woody, *Protein secondary structure from circular dichroism spectroscopy. Combining variable selection principle and cluster analysis with neural network, ridge regression and self-consistent methods.* J Mol Biol, 1994. **242**(4): p. 497-507.

125. Taylor, M.E. and K. Drickamer, *Structure-function analysis of C-type animal lectins.* Methods Enzymol, 2003. **363**: p. 3-16.

126. Hooper, L.V., et al., *Molecular analysis of commensal host-microbial relationships in the intestine.* Science, 2001. **291**(5505): p. 881-4.

127. Stappenbeck, T.S., et al., *Laser capture microdissection of mouse intestine: characterizing mRNA and protein expression, and profiling intermediary metabolism in specified cell populations.* Methods Enzymol, 2002. **356**: p. 167-96.

128. Hooper, L.V., et al., *Angiogenins: a new class of microbicidal proteins involved in innate immunity.* Nat Immunol, 2003. **4**(3): p. 269-73.

129. Rosenthal, R.S. and R. Dziarski, *Isolation of peptidoglycan and soluble peptidoglycan fragments.* Methods Enzymol, 1994. **235**: p. 253-85.

130. Abergel, C., et al., *Crystallization and preliminary crystallographic study of HIP/PAP, a human C-lectin overexpressed in primary liver cancers.* Acta Crystallogr D Biol Crystallogr, 1999. **55 (Pt 8)**: p. 1487-9.

131. Holloway, D.E., et al., *High-level expression of three members of the murine angiogenin family in Escherichia coli and purification of the recombinant proteins.* Protein Expr Purif, 2001. **22**(2): p. 307-17.

132. De Bernardez Clark, E., E. Schwarz, and R. Rudolph, *Inhibition of aggregation side reactions during in vitro protein folding.* Methods Enzymol, 1999. **309**: p. 217-36.

133. Stambach, N.S. and M.E. Taylor, *Characterization of carbohydrate recognition by langerin, a C-type lectin of Langerhans cells.* Glycobiology, 2003. **13**(5): p. 401-10.

134. Taneva, S., D.R. Voelker, and K.M. Keough, *Adsorption of pulmonary surfactant protein D to phospholipid monolayers at the air-water interface.* Biochemistry, 1997. **36**(26): p. 8173-9.

135. Weis, W.I. and K. Drickamer, *Trimeric structure of a C-type mannose-binding protein.* Structure, 1994. **2**(12): p. 1227-40.

136. Julenius, K., et al., *Prediction, conservation analysis, and structural characterization of mammalian mucin-type O-glycosylation sites.* Glycobiology, 2005. **15**(2): p. 153-64.

137. Dieckgraefe, B.K., et al., *Analysis of mucosal gene expression in inflammatory bowel disease by parallel oligonucleotide arrays.* Physiol Genomics, 2000. **4**(1): p. 1-11.

138. Dieckgraefe, B.K., et al., *Expression of the regenerating gene family in inflammatory bowel disease mucosa: Reg Ialpha upregulation, processing, and antiapoptotic activity.* J Investig Med, 2002. **50**(6): p. 421-34.

139. Li, A., et al., *Expression of a novel regenerating gene product, Reg IV, by high density fermentation in Pichia pastoris: production, purification, and characterization.* Protein Expr Purif, 2003. **31**(2): p. 197-206.

140. Kimura, S. and T. Iyanagi, *High-level expression of porcine liver cytochrome P-450 reductase catalytic domain in Escherichia coli by modulating the predicted local secondary structure of mRNA.* J Biochem (Tokyo), 2003. **134**(3): p. 403-13.

141. Lechene de la Porte, P., et al., *Immunocytochemical localization of pancreatic stone protein in the human digestive tract.* Pancreas, 1986. **1**(4): p. 301-8.

142. Hooper, L.V., T. Midtvedt, and J.I. Gordon, *How host-microbial interactions shape the nutrient environment of the Mammalian intestine.* Annu Rev Nutr, 2002. **22**: p. 283-307.

143. Wehkamp, J., et al., *Reduced Paneth cell {alpha}-defensins in ileal Crohn's disease.* Proc Natl Acad Sci U S A, 2005. **102**(50): p. 18129-34.

144. Syder, A.J., et al., *The impact of parietal cells on Helicobacter pylori tropism and host pathology: An analysis using gnotobiotic normal and transgenic mice.* Proc Natl Acad Sci U S A, 2003. **100**(6): p. 3467-72.

145. Pull, S.L., et al., *Activated macrophages are an adaptive element of the colonic epithelial progenitor niche necessary for regenerative responses to injury.* Proc Natl Acad Sci U S A, 2005. **102**(1): p. 99-104.

146. Swidsinski, A., et al., *Mucosal flora in inflammatory bowel disease.* Gastroenterology, 2002. **122**(1): p. 44-54.

147. Graf, R., et al., *A family of 16-kDa pancreatic secretory stress proteins form highly organized fibrillar structures upon tryptic activation.* J Biol Chem, 2001. **276**(24): p. 21028-38.

148. Ayabe, T., et al., *Secretion of microbicidal alpha-defensins by intestinal Paneth cells in response to bacteria.* Nat Immunol, 2000. **1**(2): p. 113-8.

149. Eckburg, P.B., et al., *Diversity of the human intestinal microbial flora.* Science, 2005. **308**(5728): p. 1635-8.

150. Holm, C. and L. Jespersen, *A flow-cytometric gram-staining technique for milk-associated bacteria.* Appl Environ Microbiol, 2003. **69**(5): p. 2857-63.

151. Liu, C., et al., *Mammalian peptidoglycan recognition protein binds peptidoglycan with high affinity, is expressed in neutrophils, and inhibits bacterial growth.* J Biol Chem, 2000. **275**(32): p. 24490-9.

152. Cash, H.L., C.V. Whitham, and L.V. Hooper, *Refolding, purification, and characterization of human and murine RegIII proteins expressed in Escherichia coli.* Protein Expr Purif, 2006. **in press**.

153. Drickamer, K., *Engineering galactose-binding activity into a C-type mannose-binding protein.* Nature, 1992. **360**(6400): p. 183-6.

154. Blixt, O., et al., *Printed covalent glycan array for ligand profiling of diverse glycan binding proteins.* Proc Natl Acad Sci U S A, 2004. **101**(49): p. 17033-8.

155. Nadesalingam, J., et al., *Mannose-Binding Lectin Recognizes Peptidoglycan via the N-Acetyl Glucosamine Moiety, and Inhibits Ligand-Induced Proinflammatory Effect and Promotes Chemokine Production by Macrophages.* J Immunol, 2005. **175**(3): p. 1785-94.

156. Harder, J., et al., *Isolation and characterization of human beta -defensin-3, a novel human inducible peptide antibiotic.* J Biol Chem, 2001. **276**(8): p. 5707-13.

157. Mombaerts, P., et al., *RAG-1-deficient mice have no mature B and T lymphocytes.* Cell, 1992. **68**(5): p. 869-77.

158. Dodd, R.B. and K. Drickamer, *Lectin-like proteins in model organisms: implications for evolution of carbohydrate-binding activity.* Glycobiology, 2001. **11**(5): p. 71R-9R.

159. Miyashita, H., et al., *Human REG family genes are tandemly ordered in a 95-kilobase region of chromosome 2p12.* FEBS Lett, 1995. **377**(3): p. 429-33.

160. Lim, J.H., et al., *Structural basis for preferential recognition of diaminopimelic acid-type peptidoglycan by a subset of peptidoglycan recognition proteins.* J Biol Chem, 2006. **281**(12): p. 8286-95.

161. Harder, J., et al., *Isolation and characterization of human beta -defensin-3, a novel human inducible peptide antibiotic.* J Biol Chem, 2001. **276**(8): p. 5707-13.

162. Lima, M.E., et al., *Purification of a lectin from the marine red alga Gracilaria cornea and its effects on the cattle tick Boophilus microplus (Acari: Ixodidae).* J Agric Food Chem, 2005. **53**(16): p. 6414-9.

163. Savage, D.C., *Factors involved in colonization of the gut epithelial surface.* Am J Clin Nutr, 1978. **31**(10 Suppl): p. S131-S135.

164. Lynn, D.J., et al., *Evidence of positively selected sites in mammalian alpha-defensins.* Mol Biol Evol, 2004. **21**(5): p. 819-27.

165. Bevins, C.L., *Events at the host-microbial interface of the gastrointestinal tract. V. Paneth cell alpha-defensins in intestinal host defense.* Am J Physiol Gastrointest Liver Physiol, 2005. **289**(2): p. G173-6.

166. Hooper, L.V., *Bacterial contributions to mammalian gut development.* Trends Microbiol, 2004. **12**(3): p. 129-34.

167. Macpherson, A.J., M.B. Geuking, and K.D. McCoy, *Immune responses that adapt the intestinal mucosa to commensal intestinal bacteria.* Immunology, 2005. **115**(2): p. 153-62.

168. Simon, M.T., et al., *HIP/PAP stimulates liver regeneration after partial hepatectomy and combines mitogenic and anti-apoptotic functions through the PKA signaling pathway.* Faseb J, 2003. **17**(11): p. 1441-50.

169. Kawasaki, N., T. Kawasaki, and I. Yamashina, *Isolation and characterization of a mannan-binding protein from human serum.* J Biochem (Tokyo), 1983. **94**(3): p. 937-47.

VITAE

Heather Lynn Cash was born on October 9, 1981 in Ft. Worth, Texas, the daughter of Karla VanHoy Cash and Michael Devon Cash. She has a younger sister, Jessica, and a younger brother, Joshua, who are twins. After completing her schooling at Covenant Christian Academy, Huntsville, Alabama in 1999, she enrolled at Samford University in Birmingham, Alabama. While there, she was the recipient of the Wilcox Scholarship for Biology and Chemistry, as well as the James Brakefield Award for Excellence in Biology honoring her senior coursework, and she completed an independent research project investigating mosquitofish genetics. During the summers of 1999-2001, she was employed by Research Genetics. She graduated summa cum laude with a Bachelor of Science in biology from Samford University in May, 2003. In August, 2003, she entered the Graduate School of Biomedical Sciences at the University of Texas Southwestern Medical Center at Dallas, joining Lora Hooper's laboratory. During her studies, she conducted research and coursework in the Department of Molecular Microbiology. She is engaged to Christopher Ivan Morris of Coatsville, Pennslyvania, and they will be married on October 14, 2006.

Permanent Address: 6726 McEachern Lane
Owens Cross Roads, Alabama 35763