MOLECULAR BASIS OF PEPTIDOGLYCAN RECOGNITION

BY A BACTERICIDAL GUT LECTIN

APPROVED BY SUPERVISORY COMMITTEE

Lora V. Hooper, Ph.D.

Kim Orth, Ph.D.

Johann Deisenhofer, Ph.D.

Margaret Phillips, Ph.D.

Dedicated to Mom and Papa

for their love and support

MOLECULAR BASIS OF PEPTIDOGLYCAN RECOGNITION BY A BACTERICIDAL GUT LECTIN

by

REBECCA ELIZABETH LEHOTZKY

DISSERTATION

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

DOCTOR OF PHILOSOPHY

The University of Texas Southwestern Medical Center at Dallas

Dallas, Texas

February, 2010

ACKNOWLEDGMENTS

First, I want to thank my mentor, Lora Hooper. I am incredibly fortunate to have been able to perform my dissertation work in her laboratory. She has been an amazing role model every step of the way, and I can't thank her enough for her continued support and guidance. I also wish to thank my dissertation committee members, Kim Orth, Hans Deisenhofer, and Meg Phillips for their all of their advice and encouragement. I owe sincere thanks to Kevin Gardner and Carrie Partch. Our collaboration was critical to the success of this project, and they've taught me more than I ever expected to know about NMR.

All of the members of the Hooper lab have contributed to making it a wonderful place to come to work each day. Sohini, thank you for being a fantastic bay mate and for the example you've set with your work ethic. Thank you also for your help and advice on this project. To Anisa and Clare, thank you so much for your friendship and support, especially during some of the more trying times of graduate school. Cassie and Shipra have both been in the lab since the time I joined, and I've greatly enjoyed getting to know them, both in and out of the lab. Thank you to Darcy and Kelly, for all you do to keep things in the lab running smoothly, as well as giving me someone to talk to about the latest sports news. I also thank Breck, Xiaofei, Jamaal, Kari, and Charmaine. Thank you for your insight and for helping make the Hooper lab what it is.

A number of friends have been constant sources of support and encouragement. I am eternally grateful to Amanda, Catie, and Vicky. I don't know what I would do without you guys. Thank you to Alaina P., Alaina E., Lori, Amanda, and everyone else from our Wednesday group. I'll always remember the time spent knitting and chatting.

The influence and support of many people have contributed to the path I'm on today. Thank you to Greg Beitel, for giving me my first lab job. Thank you to Andreas Matouschek for giving me my first opportunity to perform research and to Sumit Prakash for guiding me through it. My godparents, Peter and Debbie, have provided endless love and encouragement. I cannot adequately express how grateful I am to my parents for their infinite love and support. I would not be where I am without their sacrifices. Finally, thank you to my siblings, Andrew, Erica, and Joseph, for always being there to provide support and inspiration. You guys rock.

MOLECULAR BASIS OF PEPTIDOGLYCAN RECOGNITION

BY A BACTERICIDAL GUT LECTIN

REBECCA ELIZABETH LEHOTZKY, Ph.D.

The University of Texas Southwestern Medical Center at Dallas, 2010

LORA V. HOOPER, Ph.D.

The mammalian gut is densely populated by varied microbial species. This relationship is mutually beneficial as long as bacteria remain corralled in the gut lumen. The epithelium is protected by the secretion of antimicrobial proteins by specialized epithelial cells in the intestinal crypts. This molecular arsenal includes the RegIII family. RegIII proteins are novel in that they are C-type lectins that directly kill Gram-positive bacteria and thus play a vital role in antimicrobial protection of the mammalian gut. RegIII proteins bind their bacterial targets via interactions with cell wall peptidoglycan, but lack the canonical sequences that support calcium-dependent carbohydrate binding in

other C-type lectins. Given these novel functions and the lack of structural clues, nothing was known about the molecular mechanisms by which RegIII family members recognize and bind to peptidoglycan. Furthermore, the question of how RegIII proteins specifically recognize target microbes in the presence of soluble peptidoglycan shed by bacteria *in vivo* still remained.

In this dissertation, I have used NMR spectroscopy as an unbiased approach to study the molecular basis for peptidoglycan recognition by HIP/PAP, a human RegIII lectin. I have shown that HIP/PAP recognizes the peptidoglycan carbohydrate backbone, showing that ligand recognition by RegIII family members is unique compared to other peptidoglycan recognition proteins. This work also shows that HIP/PAP recognizes peptidoglycan in a calcium-independent manner via a conserved 'EPN' motif that is critical for bacterial killing. While EPN sequences govern calcium-dependent carbohydrate recognition in other C-type lectins, the unusual location and calcium-independent functionality of the HIP/PAP EPN motif suggest that this sequence is a versatile functional module that can support both calcium-dependent and calcium-independent carbohydrate binding. Further, these studies show that HIP/PAP binding affinity for carbohydrate ligands depends on carbohydrate chain length, supporting a binding model in which HIP/PAP molecules "bind and jump" along the extended polysaccharide chains of peptidoglycan, reducing dissociation rates and increasing binding affinity. I propose that dynamic recognition of highly multivalent carbohydrate epitopes in native peptidoglycan is an essential

mechanism governing high affinity interactions between HIP/PAP and the bacterial cell wall.

TABLE OF CONTENTS

LIST OF PUBLICATIONS

Partch, C.L.[*], **Lehotzky, R.E.[*]**, Hooper, L.V., and Gardner, K.H. A Conserved Glutamate Switch Mediates Carbohydrate Binding by a Bactericidal Gut Lectin (in preparation).

Lehotzky, R.E.[*], Partch, C.L.[*], Mukherjee, S., Cash, H.L., Goldman, W.E., Gardner, K.H., and Hooper, L.V. (2010) Molecular Basis for Peptidoglycan Recognition by a Bactericidal Lectin. *PNAS* (in press).

Mukherjee, S.*, Partch, C.L.*, **Lehotzky, R.E.**, Whitham, C.V., Chu, H., Bevins, C.L., Gardner, K.H., and Hooper, L.V. (2009) Regulation of C-type lectin antimicrobial activity by a flexible N-terminal prosegment. *Journal of Biological Chemistry* **284**(8):4881-4888.

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

AAA – ATPases associated with a variety of cellular activities

AMP – antimicrobial protein

Ang – angiogenin

ATP – adenosine triphosphate

BHI – brain heart infusion

CARD – caspase-recruitment domain

CCP – complement control protein

CRAMP – cathelin-related antimicrobial peptide

CRD – carbohydrate recognition domain

CRS – cryptdin-related sequence

CTL – C-type lectin

CTLD – C-type lectin-like domain

DAP – diaminopimelic acid

DC-SIGN – dendritic cell-**S**pecific intercellular adhesion molecule-3-grabbing non-integrin

DEFA – α-defensin gene

DEFB – β-defensin gene

Defcr – mouse α-defensin gene (cryptdin)

DNase - deoxyribonuclease

ECP – eosinophil cationic protein

EDTA – ethylenediaminetetraacetic acid

EGF – epidermal growth factor

GMDP – GlcNAc-MurNAc dipeptide

GNBP – Gram-negative binding protein

hCAP18 – human cationic protein 18 kDa

HIP/PAP – hepatocarcinoma-intestine-pancreas/pancreatic associated protein

HNP – human neutrophil peptide

HSQC – heteronuclear single quantum coherence

IBD – inflammatory bowel disease

IL – interleukin

Imd – immune deficiency

IPTG – isopropyl-β-D-thiogalactopyranoside

ITC – isothermal titration calorimetry

GlcNAc – N-acetylglucosamine

LRR – leucine-rich repeat

MAMP – microbe-associated molecular pattern

MBP – mannose-binding protein

MBTH – 3-methyl-2-benzothiazolinone hydrazone

MDP – muramyl dipeptide

MES – 2-(*N*-morpholino)ethanesulfonic acid

MManR – macrophage mannose receptor

MTri-DAP – DAP-type muramyl tripeptide

MurNAc – N-acetylmuramic acid

MyD88 – myeloid differentiation primary response gene

NK cells – natural killer cells

NMR – nuclear magnetic resonance

NOD – nucleotide-binding oligomerization domain

PCR – polymerase chain reaction

PGLYRP – human peptidoglycan recognition protein

PGRP – peptidoglycan recognition protein

PPO – pro-phenol-oxidase

R_1 – spin lattice relaxation

R_2 – spin-spin relaxation

Reg – regenerating

Rel/NF-κB – Rel homology domain/nuclear factor kappa-light-chain-enhancer of activated B cells

RNase – ribonuclease

RT-PCR – real-time polymerase chain reaction

SAB – standard assay buffer

SDS-PAGE – sodium dodecyl sulfate-polyacrylamide gel electrophoresis

sPGN – solubilized peptidoglycan

$sPLA_2$ – secretory phospholipase A_2

TCA – trichloroacetic acid

TCT – tracheal cytotoxin

Th – helper T cell

T_m – melting temperature

Tri-DAP – DAP-type tripeptide

Tri-Lys – Lys-type tripeptide

TLR – Toll-like receptor

WT – wild-type

CHAPTER I

INTRODUCTION

Microbes and the mammalian gut

The mammalian gut harbors vast populations of microbial species. Approximately 100 trillion bacteria are present in the mammalian gut, meaning microorganisms outnumber host cells 10:1 [1]. This microbial presence is mutually beneficial. Microbes are provided a safe, nutrient-rich environment in which to multiply. Meanwhile, microbes provide nutritional benefits to their hosts by aiding in the metabolism of otherwise indigestible polysaccharides and the production of nutrients such as vitamin K [2]. Other physiological benefits for the host include immune system development [3], blood vessel development [4], and regulation of fat storage [5].

However, for this relationship to remain beneficial, it is critical that microbes remain corralled in the gut lumen. Breach of the epithelial layer can lead to inflammation and sepsis. Due to the dense bacterial populations in the gut, the epithelium is always under attack [6]. One way the epithelium protects against these constant and varied bacterial challenges is through the secretion of antimicrobial proteins [6, 7].

Antimicrobial proteins

Antimicrobial proteins (AMPs) are a component of the innate immune system and are present in virtually all multicellular organisms [6]. They are most highly expressed in tissues prone to microbial attack, such as the skin and the epithelium of the lungs and intestines [8]. In the mammalian gut, AMP production is performed largely by Paneth cells, which are specialized epithelial cells located at the bases of intestinal crypts (Fig. 1-1) [9].

AMPs function by targeting microbe-associated molecular patterns (MAMPs). These MAMPs include essential cell wall components and are thus unlikely to undergo significant structural alterations that would confer resistance to AMP activity [6]. AMP mechanisms of action include pore formation, enzymatic attack on cell walls or sequestration of essential nutrients [8]. These microbicidal activities target a broad spectrum of microorganisms, including both Gram-positive and Gram-negative bacterial species as well as fungi, protozoa, and some enveloped viruses [10].

Due to the acute nature of microbial attack at epithelial surfaces, host defenses must be able to respond quickly. Because of this, AMP expression in vertebrates can be either constitutive or rapidly induced by microbial macromolecules or cytokines via the rel/NF-κB family. Many previously studied AMPs are expressed as larger precursors that are activated by subsequent

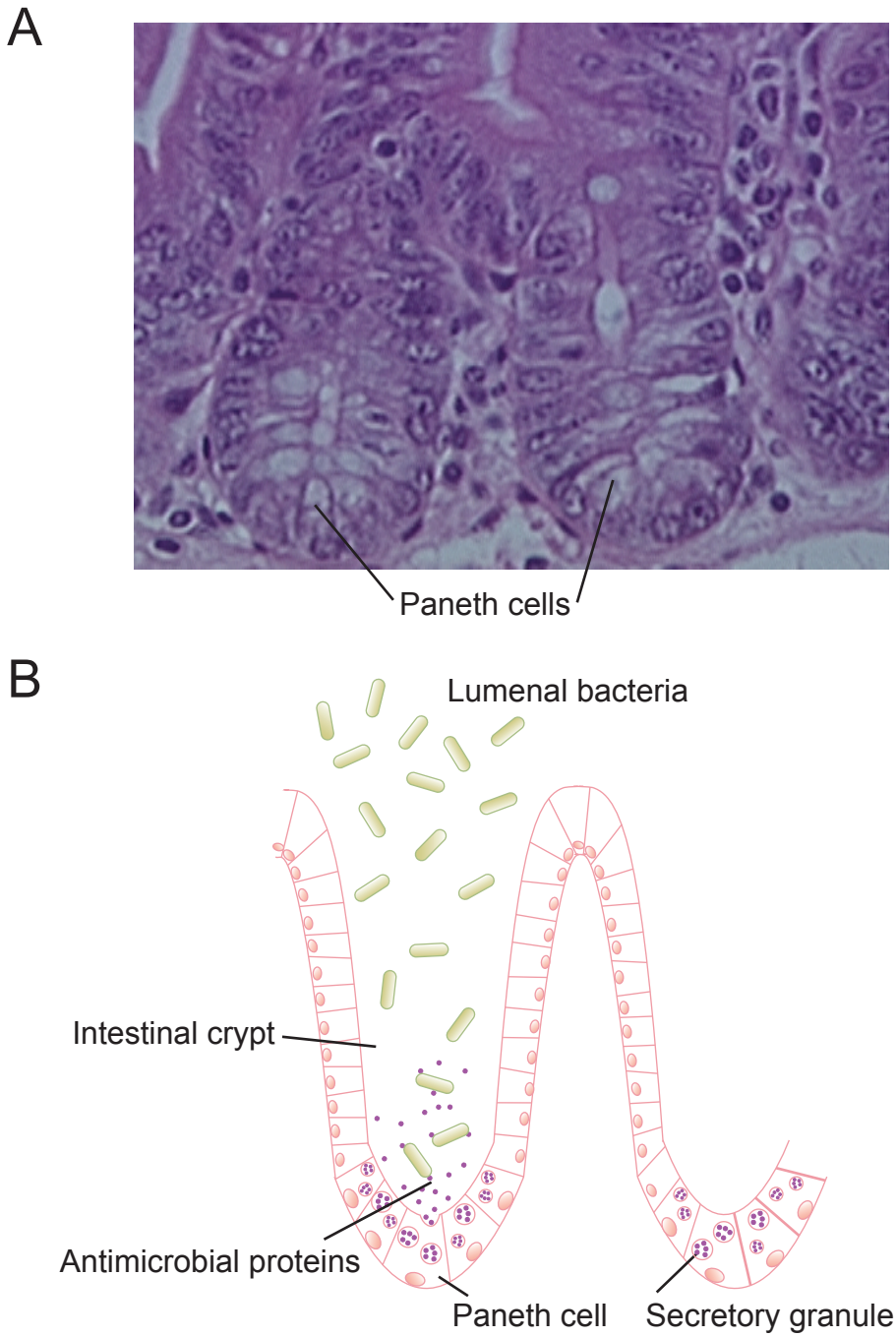


Figure 1-1: Paneth cells in the mammalian gut. Paneth cells are located at the bases of intestinal crypts in the mammalian gut. **(A)** Hematoxylin and eosin staining of mouse small intestinal tissue. **(B)** Paneth cells secrete antimicrobial proteins (AMPs) in response to bacteria to protect the epithelial barrier.

proteolytic processing, thus providing an additional level of temporal and spatial control over the expression of antimicrobial activity [8].

AMP deficiencies have been linked to a number of pathologies. Knockout mouse models, for example, exhibit decreased resistance to infection [11, 12]. Additionally, reduced AMP expression has been linked to inflammatory bowel disease (IBD), a chronic intestinal inflammation thought to be triggered by the local flora in genetically predisposed individuals [13, 14].

AMPs that function through enzymatic attack on microbial cell walls

Enzymatically active AMPs function by attacking crucial microbial cell components. One target is the microbial cell wall. Lysozyme is a 14 kDa glycosidase that specifically targets the β-1,4 glycoside bonds between N-acetylglucosamine (GlcNAc) and N-acetylmuramic acid (MurNAc) sugars that make up the backbone of peptidoglycan, a critical bacterial cell wall component. Lysozyme is only effective against Gram-positive bacteria, in which peptidoglycan is found on the cell surface, unless cofactors are present to help the enzyme breach the Gram-negative membrane and reach the peptidoglycan buried in the periplasmic space [15].

Secretory phospholipase A_2 ($sPLA_2$) is a highly basic enzyme. This characteristic allows it to penetrate the bacterial cell wall to reach the cell

membrane, where it compromises membrane integrity by hydrolyzing phospholipids [16, 17].

Several members of the ribonuclease family have been shown to have potent antimicrobial activity. RNase 7 expressed in the skin has antibacterial activity against a broad spectrum of species [18]. Eosinophil cationic protein (ECP) and Angiogenin-4 (Ang-4) also have both ribonuclease and antimicrobial activity, although whether the two functions are related remains unclear [19, 20].

Defensins

Defensins are small (2-6 kDa), cationic peptides that disrupt microbial cell membranes [7]. They are the major family of membrane disruption peptides in mammals, as well as one of the most diverse and highly expressed proteins in the gut. They are expressed in a variety of cells, including macrophages, neutrophils, and epithelial cells. Defensins target microbial membranes via electrostatic interactions with the negatively charged phospholipid groups of the bacterial membrane [6]. This mechanism provides specificity since bacterial membranes feature unique positioning of the phospholipid head groups on the outer leaflet, which are not present on eukaryotic membranes [21]. Once defensins reach a critical bound concentration on the microbial membrane, they form transient pores which promote osmotic lysis [22-24].

Beyond size and charge, defensins share several structural characteristics. All lack glycosyl or acyl side-chain modifications and feature tertiary structures made up of turn-linked β-strands [25]. While defensins display a high degree of sequence variability, they contain highly conserved cysteine residues that form intramolecular disulfide bonds. Classification of defensins into three subfamilies – α-defensins, β-defensins, or θ-defensins – is based on the arrangement and spacing of these disulfide bonds [10].

α-Defensins

Expression of α-defensins (also known as 'cryptdins' in mice) is highly restricted to Paneth cells and neutrophils in humans (and Paneth cells alone in mice). Genome analysis has shown that humans have 11 α-defensin genes. Six of these genes (DEFA1-6) are expressed while 5 (DEFA7P-11P) are most likely pseudogenes. Mice harbor 26 cryptdin genes (Defcr1-26) and one known pseudogene (Defcrps1). Additionally, mice but not humans have a diverse family of 20 cryptdin-related sequence (CRS) peptides [7]. These peptides contain four intramolecular disulfide bridges and can also form covalent dimers through an additional disulfide bridge. The ability to form homo- and heterodimers allows for high combinatorial diversity [26].

α-Defensins are known to have broad antimicrobial activity against both Gram-positive and Gram-negative bacterial species, as well as fungi, protozoa,

and enveloped viruses [25]. The high degree of α-defensin expression by Paneth cells suggests they form part of a constitutive barrier in the intestine [7]. However, some gut pathogens are thought to have evolved mechanisms to subvert this barrier. For example, oral infection with *Salmonella typhimurium* has been shown to result in decreased cryptdin 1 expression in the small intestine within 24 hours [27]. Microbial evolution is thought to be the driving force behind the relatively rapid evolution of defensin genes in mammals through duplication and diversification.

Recent studies have indicated that α-defensins may have roles that extend beyond their antimicrobial activity. One report showed that a combination of HNP1-3 (the neutrophil-expressed products of DEFA1-3) inhibited endothelial cell adhesion, migration, and proliferation in culture. This mixture also inhibited capillary-like tube formation in a three-dimensional fibrin gel [28]. A second study showed that mouse cryptdin 3 induced IL-8 in cultured human intestinal epithelial cells. These data suggest that cryptdins can act as paracrine regulators of innate inflammatory responses [29]. The physiological relevance of these studies has yet to be established and will present a significant challenge given the number of closely related defensins *in vivo*.

β-Defensins

β-Defensin expression is observed in enterocytes throughout the mammalian gut [30]. Twenty-eight β-defensin genes have been identified in the human genome thus far (DEFB1-28), including eight that are known to be expressed and several that are most likely pseudogenes [7, 31]. Mice, meanwhile, have 14 β-defensin genes that are known to be expressed in epithelial organs including the lungs, skin, and intestine. β-Defensins have been shown to have broad antimicrobial effects *in vitro* [7]. *In vivo* studies in mDEFB1 knockout mice indicated that these animals have difficulty clearing *Haemophilus influenza* from the lung [32] and have higher numbers of staphylococci in their bladders [33], although no intestinal pathology has been identified. This is likely due to overlap in function of the β-defensins, and further studies will be necessary to determine whether mDEFB1 deletion has more subtle effects on commensal populations or susceptibility to specific pathogenic microbes.

Like α-defensins, β-defensins are thought to have functions in addition to their antibacterial activity. Studies have indicated that β-defensins are chemotactic for dendritic cells and T cells [34]. Reports have also suggested that they may induce the release of multiple cytokines and chemokines including CD80, CD86, CD40, IL-6, and IL-10 [35, 36].

θ-Defensins

θ-Defensins are formed from α-defensin paralog precursors. A nonapeptide is excised from the precursor and spliced to an identical or similar peptide to form a cyclic molecule stabilized by three disulfide bonds. They are the only known cyclic polypeptides in animals and are expressed in several primate species, including monkeys and orangutans, but not in humans [25]. While humans express mRNA for θ-defensin orthologs, stop codons in the θ-defensin precursors prevent translation [37].

Cathelicidins

Cathelicidins are cationic α-helical peptides that are related by their conserved 14 kDa N-terminal 'cathelin' (cathepsin L inhibitor)-like domain [6]. Enzymatic cleavage of this conserved domain activates the highly variable C-terminal region [10]. Humans and mice each have only a single cathelicidin gene which encode the proteins LL-37/hCAP18 and cathelin-related antimicrobial peptide (CRAMP), respectively [6]. While they were originally identified in neutrophils [38, 39], cathelicidins are now known to also be expressed in other granulocytes such as NK-cells and mast cells [10], as well as epithelial cells of the colon [40], lung [41], skin [42], and urinary tract [43].

Cathelicidins exhibit microbicidal activity against both Gram-positive and Gram-negative bacteria as well as fungi. They function by first associating with

the microbial membrane via electrostatic interactions and then inserting in and disrupting the membrane [41]. However, cathelicidins perform a number of functions unrelated to their direct microbicidal activity. LL-37 has been shown to be chemotactic for monocytes, macrophages, T cells, and neutrophils *in vitro* [7]. There is also evidence that LL-37 induces Th1 cytokine secretion by dendritic cells [44] and promotes angiogenesis through the formyl peptide receptor-like 1 expressed on endothelial cells [45].

C-type lectins

C-type lectins have only recently been shown to have direct antimicrobial activity, but their role in immunity has long been established. Lectins are Metazoan carbohydrate binding proteins [46]. C-type lectins are so named because of their dependence on Ca^{2+} for ligand binding [47]. This binding is mediated through a conserved carbohydrate recognition domain (CRD). The CRD is part of a larger domain family called the C-type lectin-like domain (CTLD). CTLDs have sequence and structural similarities to CRDs, but may bind proteins rather than carbohydrates and are often Ca^{2+}-independent. CTLD-containing proteins are further grouped according to the organization of the structural domains in the proteins where they are found [46] (Fig. 1-2).

Many of the CTLD-containing protein subfamilies are known to be involved in immunity. The selectins (Group IV) are proteins composed of a type

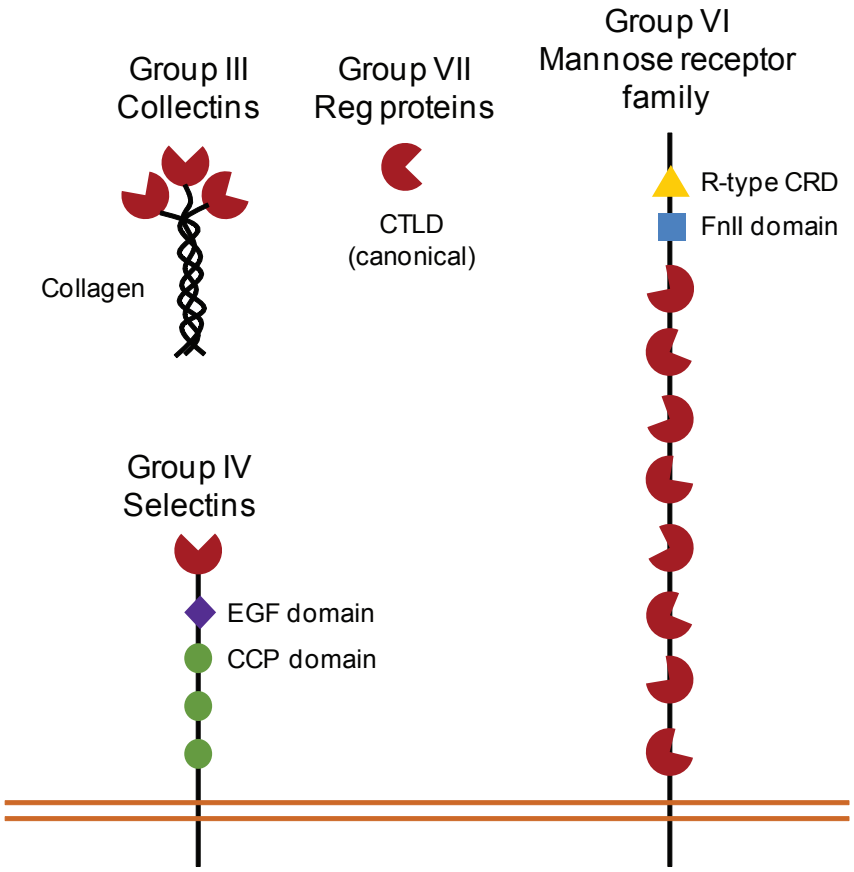


Figure 1-2: Domain organization of CTLD-containing proteins. CCP, complement control protein; EGF, epidermal growth factor; Fn,fibronectin; CTLD, C-type lectin-like domainl; CRD, carbohydrate recognition domain.

I transmembrane protein containing an N-terminal CRD and an epidermal growth factor (EGF)-like domain. These domains are then linked to the cell membrane by 2-9 complement control protein (CCP) domains [46]. Selectins facilitate selective cell adhesion between leukocytes and epithelial cells during normal leukocyte trafficking and also as part of inflammatory responses [48, 49].

Multi-CTLD endocytic receptors (Group VI), such as the macrophage mannose receptor (MManR), function by leading to direct phagocytosis of target microbes. Group VI members are type I transmembrane proteins that include an N-terminal ricin-like domain, a fibronectin type 2 domain and 8 or 10 extracellular CTLDs. Interestingly, most CTLDs of these proteins do not contain recognized Ca^{2+}-binding motifs [46].

One of the structurally simplest family members is also one of the best studied. The collectins (Group III) are soluble homo-oligomers made up of polypeptides containing N-terminal collagenous domains and C-terminal CRDs [46]. Collectins frequently mediate pathogen neutralization via the complement pathway. Mannose-binding protein (MBP) is the best understood member of this group, both structurally and functionally. MBP is known to recognize a number of Gram-negative species such as *Escherichia coli* and *Mycobacterium tuberculosis* as well as some fungi [50, 51].

C-type lectin-like domain (CTLD)

Tertiary structure is highly conserved in members of the CTLD family (Fig. 1-3). The core of the protein is composed of two β-sheets and two α helices. Extending from this core is a region known as the long loop, the structure of which forms two smaller loops (designated 'Loop 1' and 'Loop 2'). Four conserved cysteines form disulfide bridges that anchor the long loop and the domain as a whole. The long loop region is responsible for coordinating Ca^{2+} as well as ligand binding [46].

CTLDs have four potential Ca^{2+}-binding sites that exhibit different levels of occupancy in crystallographic studies depending on crystallization conditions and the specific CTLD being studied. Sites 1, 2, and 3 are in the long loop region, although site 3 is thought to be a crystallographic artifact in rat MBP-A (Fig. 1-3). A fourth site, which is not observed in MBP-A, contributes to the salt bridge formation between α2 and the β1/β5 sheet in other CTLD family members [46].

While CRDs contain a high degree of homology, their binding specificities vary significantly. In the case of MBP, broad ligand specificity is critical for the protein to be able to recognize a wide spectrum of microbes. MBP specificity includes ligands with equatorial stereochemistry of the 3- and 4-OH groups of D-mannose, such as GlcNAc and L-fucose. Conversely, MBP does not recognize galactose-type ligands, which have axial 4-OH groups. This type of specificity is common in C-type lectins, as most family members can be

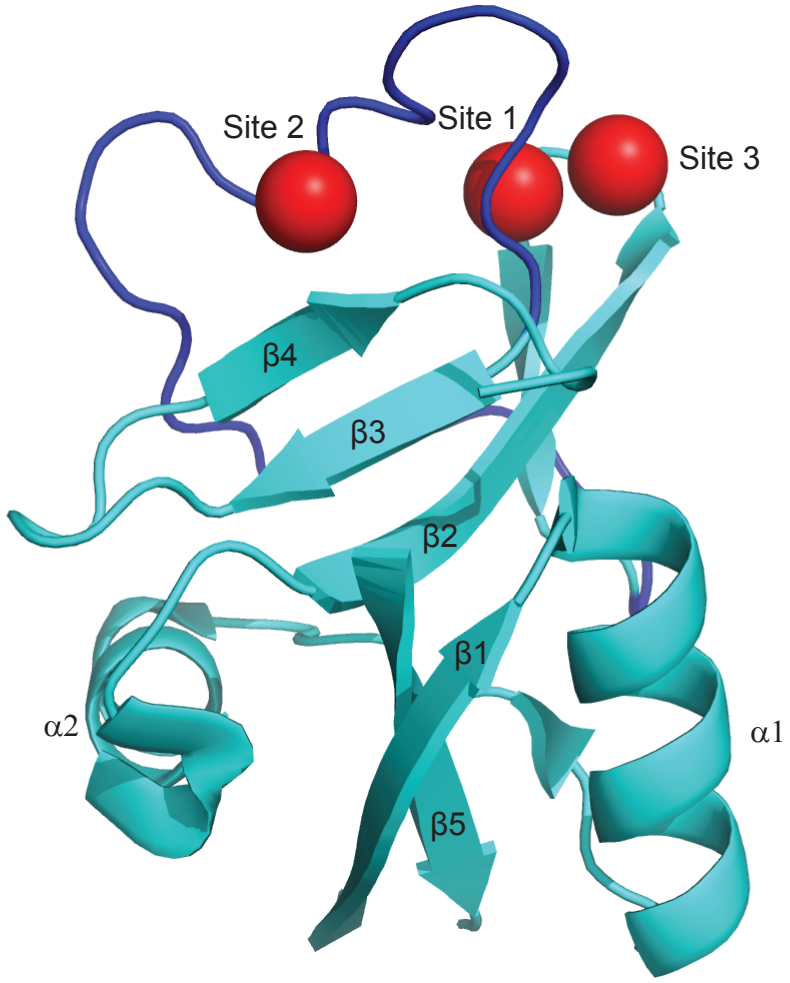


Figure 1-3: Structural characteristics of CTLD. Rat MBP-A structure (2msb), with long loop region in dark blue. Ca^{2+} ions are depicted by red spheres.

categorized as binding either mannose-type ligands or galactose-type ligands [49]. Specificity can be predicted by the presence of an 'EPN motif' or 'QPD motif', respectively, in loop 2 of the long loop region [52]. A conserved 'ND' is known to also be important for Ca^{2+} coordination, but does not affect ligand specificity [46].

Like other C-type lectins, MBP exhibits relatively low affinity for monosaccharide ligands (~2 mM) [53]. Affinity is greatly enhanced for oligomeric ligands, with affinities measured in nM range [54]. This effect is due to the clustering of binding sites as well as lectin clustering, and provides a mechanism for allowing C-type lectins to preferentially target microbial cells rather than host cells. Mannose residues are more tightly packed on the cell surface of microbes than on mammalian cells, thus allowing for stronger multivalent binding [49].

The C-type lectin EPN motif

The EPN motif is directly involved in the coordination of both Ca^{2+} and saccharide. In MBP-A, this motif corresponds to residues 185-187 and also includes Ca^{2+}-binding site 2 (Fig. 1-4). Ca^{2+} is coordinated by electron pairs from the carbonyl groups from the E185 and N187 side chains. The second pair of electrons from each side chain carbonyl forms hydrogen bonds with the 3- and 4-OH groups of the target mannose. The mannose hydroxyl groups form additional

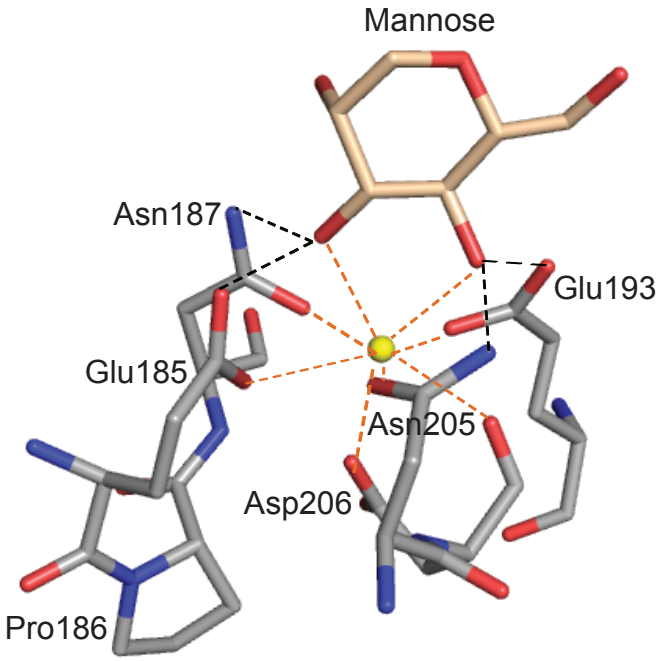


Figure 1-4: Ca^{2+} and mannose coordination by EPN motif (2msb). MBP-A binding site residues (gray) in complex with mannose (wheat). Coordination bonds are in orange. H-bonds are in black. Oxygen and nitrogen atoms are in red and blue, respectively. The Ca^{2+} ion is depicted as a yellow sphere.

coordination bonds with the Ca^{2+} ion. N205, E193 and D206 each contribute an additional coordination bond to the Ca^{2+} via their side chains, and D206 also forms an additional coordination bond using an electron pair from its backbone carbonyl group [55].

The hydrogen bonds formed with E185 and N187 determine the binding specificity of MBP-A. Work by Kurt Drickamer illustrated the importance of this tripeptide motif when it was shown that mutating these residues to produce a QPD motif in MBP resulted in a protein with galactose-type binding affinity as opposed to the native mannose-type affinity [52].

Reg family

Reg proteins make up Group VII of C-type lectins. They are approximately 16 kDa and are composed of a single CTLD and an N-terminal secretion signal. The family is further organized into subgroups (I, II, III, and IV) based on primary sequence. Previous work in our lab showed that some RegIII family members are expressed in the Paneth cells of conventionalized, but not germ-free (microbiologically sterile) mice. RegIIIγ, and the human ortholog HIP/PAP, were shown to be directly microbicidal against Gram-positive bacteria [56] (Fig. 1-5A). This is a unique function among C-type lectins. Furthermore, we now know that this activity is regulated by the cleavage of an N-terminal prosegment by intestinal trypsin [57]. While N-terminal proteolytic processing

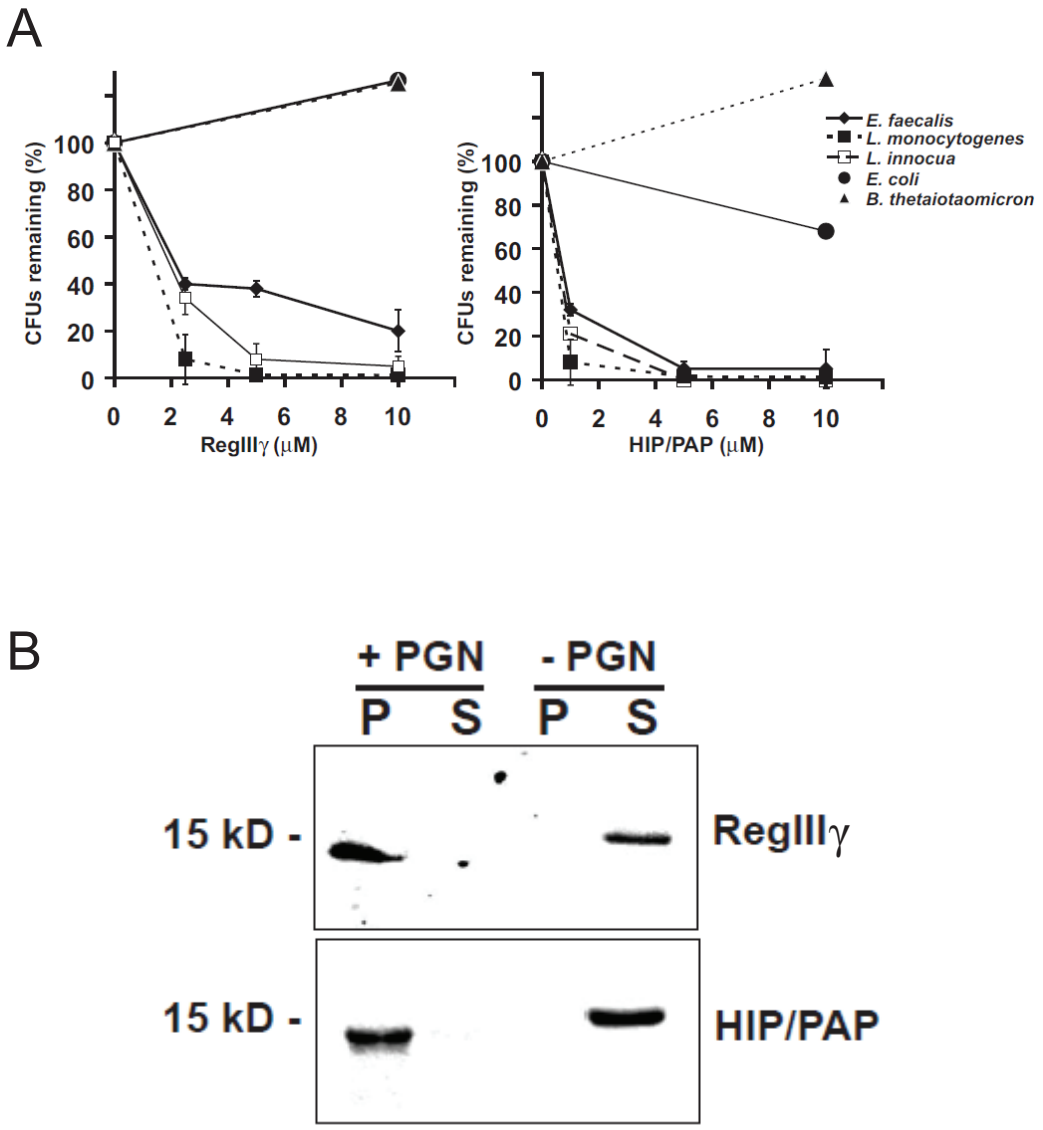


Figure 1-5: RegIIIγ and HIP/PAP target bactericidal activity against Gram-positive microbes through peptidoglycan recognition (56). **(A)** RegIIIγ and HIP/PAP each exhibit dose-dependent killing against Gram-positive, but not Gram-negative species. **(B)** RegIIIγ and HIP/PAP each localize to the peptidoglycan pellet fraction in pull-down assays. Both proteins remain soluble in the absence of peptidoglycan. Peptidoglycan-associated proteins have a reduced molecular weight due to N-terminal cleavage from proteases in the peptidoglycan preparation.

has been demonstrated in the defensin family, it represents a unique regulatory mechanism among C-type lectins. Pull-down assays revealed that RegIIIγ and HIP/PAP target bacteria through recognition of peptidoglycan (Fig. 1-5B), which is a major structural component of the bacterial cell wall [56].

Peptidoglycan recognition

Peptidoglycan is critical for bacterial survival. It forms a mesh-like covering on the surface of Gram-positive bacterial species, and is buried in the periplasmic space of Gram-negatives, between the inner and outer membranes. It contributes to cell shape and protects against osmotic lysis [58]. It also acts as a scaffold for other cell envelope components, including proteins [59] and teichoic acids [60].

Since peptidoglycan is found specifically in microbial cell walls, it is a useful MAMP for targeting by the host immune system [61]. Several families of immune proteins have evolved to exploit this essential feature of microbe physiology.

Peptidoglycan structure

Peptidoglycan is composed of long carbohydrate chains of repeating GlcNAc-MurNAc disaccharides connected via β-1,4 linkages (Fig. 1-6A). In all Gram-negative and some Gram-positive species, such as *Bacillus subtilis*, these

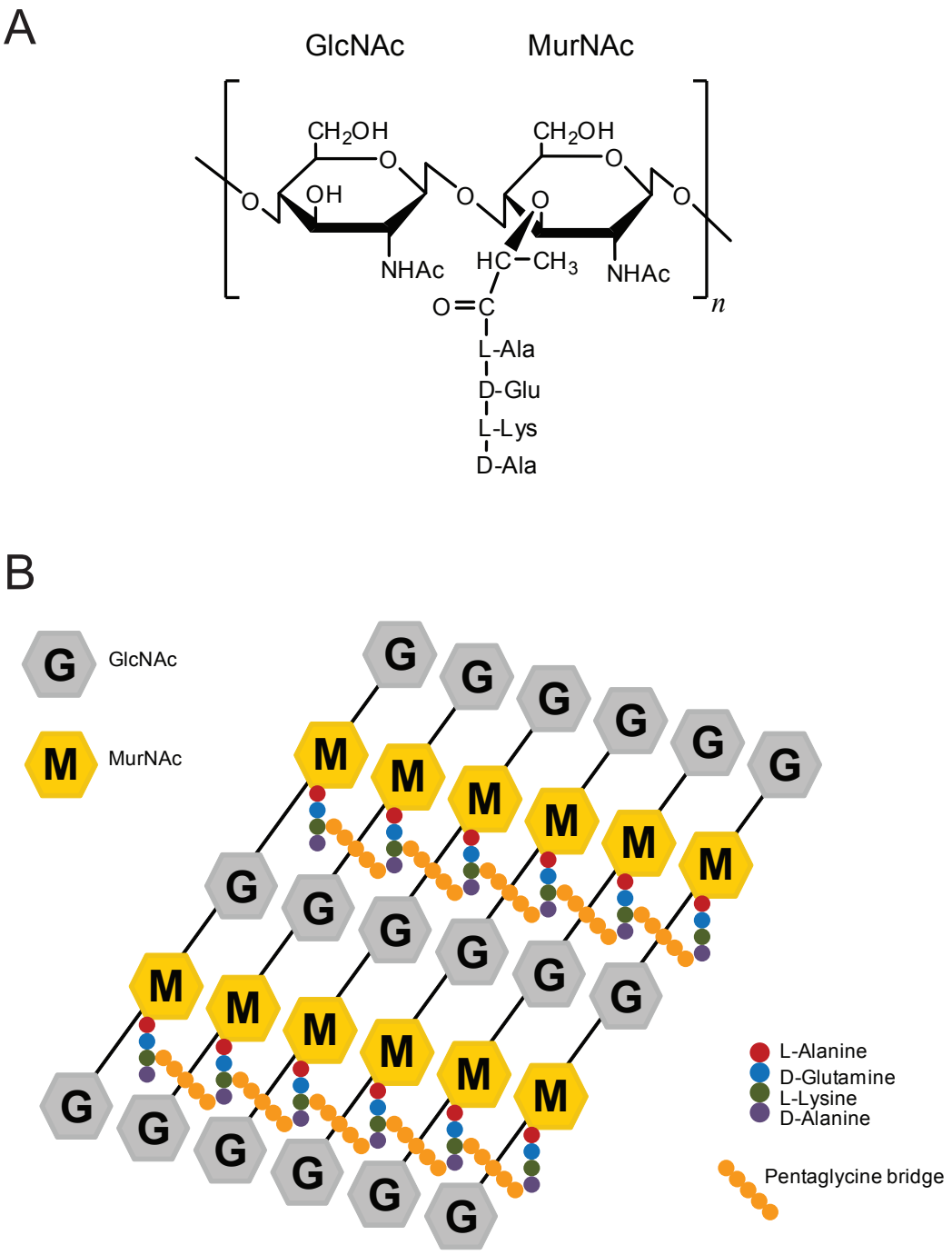


Figure 1-6: Molecular structure of peptidoglycan. **(A)** Structural unit of Lys-type peptidoglycan. **(B)** Cross-linking of Lys-type peptidoglycan structure units via pentaglycine bridges.

saccharide chains lack reducing termini and instead end with a 1,6-anhydroMurNAc residue [58].

The lactoyl group on each MurNAc residue is substituted by a pentapeptide that is processed to a tetrapeptide in mature peptidoglycan. The sequence of this peptide is highly conserved amongst all bacterial species. It is most often L-Ala-D-Glu-(*meso*DAP or L-Lys)-D-Ala-D-Ala. The identity of the amino acid in the third position is dictated by the species of bacteria. Most Gram-positive species have L-Lys, whereas most Gram-negative species, but also Bacilli, have *meso*DAP (diaminopimelic acid) [58]. DAP is structurally similar to lysine, differing only by the presence of a side chain carboxyl group, and is a precursor of lysine synthesis in bacteria [62].

Peptidoglycan chains are cross-linked to each other through their tetrapeptides, but the nature of this cross-linking varies. Cross-linking most often occurs between the side chain amino group of the third amino acid of one chain and the D-Ala carboxyl group of another. This link may be direct, as is most often the case in Gram-negative species, or via a short peptide bridge as is frequently seen in Gram-positive species. The composition of this peptide linker varies, but is often a pentaglycine bridge [58] (Fig. 1-6B).

Peptidoglycan recognition proteins (PGRPs)

PGRPs are important components of the innate immune effector protein arsenal. This family of proteins is highly conserved in evolution, and its members have been identified in insects and mammals, but not plants or nematodes. Each member of the family contains at least one, and sometimes two, PGRP domains, which are homologous to bacteriophage and bacterial type 2 amidases [63].

Insect PGRPs

Drosophila have 13 PGRP genes that are expressed as at least 19 proteins [64]. These proteins are divided into two subfamilies. Short PGRPs (PGRP-S) are small extracellular proteins. Long PGRPs (PGRP-L) include intracellular, extracellular, and transmembrane proteins [64, 65].

Insect PGRPs can exert their antimicrobial activity both directly and through indirect signaling pathways. Three PGRP-S proteins (PGRP-SA, PGRP-SD, and PGRP-SC1) directly recognize peptidoglycan and activate proteases that cleave Spaetzle, an extracellular protein found in insect haemolymph, which subsequently activates the Toll pathway [66-68]. This ultimately leads to the activation of the transcription factors Dorsal and Dif, which correspond to mammalian NF-κB, and expression of antimicrobial peptides specific for Gram-positive bacteria and fungi [69]. This pathway is triggered predominantly by Lys-

type peptidoglycan [70], even though PGRP-SA will also recognize DAP-type ligands. This is likely due to the fact that DAP-type peptidoglycan specifically triggers L,D-carboxypeptidase activity in PGRP-SA [71]. Toll activation is facilitated by Gram-negative binding protein-1 (GNBP-1) [72, 73], which cleaves peptidoglycan to produce reducing ends that are recognized by PGRP-SA [74]. PGRP-SA and PGRP-SC1 have each been shown to promote phagocytosis, although the mechanism for this activity has yet to be elucidated [67].

PGRP-L proteins lead to the activation of the *Imd* pathway [75-77]. Peptidoglycan recognition is mediated by PGRP-LC, which preferentially recognizes DAP-type ligands. This pathway leads to the activation of the Relish transcription factor and ultimately to the expression of antimicrobial proteins that are predominantly Gram-negative-specific [63]. PGRP-LC may also play a role in phagocytosis of Gram-negative microbes, but the current evidence is still inconclusive [77]. Also, PGRP-LE has been shown to activate the pro-phenol-oxidase cascade, which generates additional antimicrobial molecules [78].

Like PGRP-SA, PGRP-SC1 and PGRP-LB have each been shown to exert direct enzymatic cleavage on peptidoglycan. Both act as N-acetylmuramyl-L-alanine amidases and target the amide bond that joins MurNAc and L-Ala, resulting in removal of the peptidoglycan peptide moiety [79, 80].

Mammalian PGRPs

The mammalian PGRP family contains four members (PGLYRP-1-4). All are secreted proteins [63]. Until recently, little was known about their functions or mechanisms of activity.

Biochemical assays have suggested that PGLYRP-1, PGLYRP-3, and PGLYRP-4 are secreted as disulfide-linked dimers. Each has the ability to form homodimers, although when PGLYRP-3 and PGLYRP-4 were coexpressed, only PGLYRP-3:4 heterodimers were observed [81]. Each of these proteins has been shown to have direct bactericidal activity, which is Ca^{2+}-dependent in PGLYRP-1. Human PGLYRPs target specific Gram-positive bacteria, including *Listeria monocytogenes*, *Staphylococcus aureus*, as well as several *Bacillus* and *Lactobacillus species*, although there are variations in specificity in other mammals. While direct evidence is still lacking, kinetic studies suggest the mechanism of this microbicidal activity is through the inhibition of peptidoglycan synthesis. PGLYRPs have also been shown to have synergistic effects with bacteriolytic enzymes and AMPs *in vitro*, which likely has physiological benefits, since these proteins are often expressed at the same surfaces [63].

Unlike the other PGLYRP family members, PGLYRP-2 functions as an N-acetylmuramoyl-L-alanine amidase like PGRP-SC1 and PGRP-LB [82, 83]. In mammals, this may have the additional effect of reducing the proinflammatory activity of peptidoglycan, which often requires polymeric peptidoglycan [63].

All PGLYRPs that have been studied bind peptidoglycan. However, studies have shown that PGLYRPs also bind fungi, which have no peptidoglycan, indicating that this protein family may target other ligands as well, possibly through hydrophobic domains found on the opposite side of the molecule from the peptidoglycan binding site. Each PGLYRP has at least one PGRP domain, but PGLYRP-3 and PGLYRP-4 have two. These PGRP domains have significant variability. The 37-43% identity of these domains suggests that specific molecular recognition characteristics, specificities, and binding affinities likely differ among family members [63]. All crystallized PGRPs, including PGLYRP-1, C-terminal PGLYRP-3, and PGRP-SA, have required muramyl-tripeptide as a minimum for ligand recognition [71, 79, 84-87], which is also the smallest fragment cleaved by the amidase activity of PGLYRP-2 [83]. Both human and insect PGRPs show specificity for DAP or Lys-type peptidoglycan, distinguishing between the two by specific recognition of the third amino acid of the peptide and through specificity for the type of peptide cross-linking [63].

Nucleotide-binding oligomerization domain (Nod) proteins

NOD proteins are present in the mammalian cell cytoplasm, facilitating detection of invasive microbes [61]. Two members of this family, Nod1 and Nod2, recognize peptidoglycan-derived MAMPs. This recognition leads to interaction with the RICK serine-threonine kinase, which subsequently promotes

NF-κB activation [88, 89]. This activation is MyD88-independent, thus representing a TLR-independent proinflammatory mechanism [61]. NOD gene mutations have been linked to a number of inflammatory disorders, including IBD [89].

Nod1 and Nod2 each contain C-terminal leucine-rich repeats (LRR), a Nod domain, and an N-terminal caspase-recruitment domain (CARD). The main structural difference between Nod1 and Nod2 is that Nod2 contains a second CARD domain [89]. Nod1 expression is ubiquitous in adult tissues, whereas Nod2 expression is found primarily in monocytes [90, 91].

Recent studies have identified exactly which peptidoglycan moieties are necessary for Nod activation. Nod1 is specific for DAP-type peptidoglycan, allowing its targets to include Gram-negative species and Bacilli. The minimum molecule required for Nod1 activation is DAP-type tripeptide (L-Ala-D-Glu-*meso*DAP) [92-94]. However, Nod2 is activated by the muramyl dipeptide (MurNAc-L-Ala-D-Glu) portion of peptidoglycan, and thus does not display specificity towards DAP-type or Lys-type bacterial species [90, 94-96]. Furthermore, Nod studies have relied on downstream effects of Nod activation as an indicator of peptidoglycan recognition, and thus direct binding of peptidoglycan by Nod proteins has not been definitively established.

Goal of this dissertation

This work focuses on the recognition of peptidoglycan by the RegIII family of bactericidal C-type lectins. This family represents a novel group of antimicrobial proteins, since neither peptidoglycan binding nor directly microbicidal activity had previously been attributed to C-type lectins. Therefore, the molecular mechanisms by which RegIII proteins recognize and bind to peptidoglycan were unknown prior to this work.

The main goals of my work were to (1) identify the peptidoglycan binding site on HIP/PAP, the human ortholog of mouse RegIIIγ, and (2) determine what peptidoglycan moieties are required for this recognition. As I will detail in the chapters of my thesis, these studies reveal that HIP/PAP recognizes the peptidoglycan carbohydrate moiety via a conserved EPN motif. This motif functions in a Ca^{2+}-independent manner and is in a distinct location compared to the EPN sequences found in Ca^{2+}-dependent lectins. This suggests that EPN motifs are independently evolved, versatile functional modules that can support both Ca^{2+}-dependent and Ca^{2+}-independent carbohydrate binding.

The discovery that HIP/PAP targets peptidoglycan via the carbohydrate backbone makes it unique compared to other peptidoglycan binding proteins. My work has additionally shown that HIP/PAP binding affinity for saccharide ligands is critically dependent on carbohydrate chain length. This is consistent with observations for other lectins, which display greatly enhanced affinity for ligands

with densely clustered epitopes. This finding suggests that the highly multivalent carbohydrate epitopes found in native peptidoglycan are essential for efficient function of HIP/PAP *in vivo* and provides a molecular explanation for why HIP/PAP activity is not inhibited by soluble peptidoglycan fragments in the gut lumen.

This work has therefore shown that the HIP/PAP-peptidoglycan interaction is a novel mechanism of peptidoglycan recognition and provides initial insight into how bactericidal C-type lectins target and kill bacteria.

CHAPTER II

IDENTIFICATION OF THE PEPTIDOGLYCAN BINDING SITE ON HIP/PAP

Introduction

While previous work in our lab had shown that HIP/PAP recognizes peptidoglycan [56], nothing was known about how this interaction takes place. In other C-type lectin (CTL) family members, ligand binding depends on conserved Loop 2 motifs. CTLs that target mannose-type carbohydrates, such as N-acetylglucosamine (GlcNAc), rely on EPN and ND motifs for ligand recognition [46]. However, these motifs are not present in Loop 2 of RegIII family members, meaning that a more unbiased approach would be necessary to identify the HIP/PAP peptidoglycan binding site.

We addressed this question using solution nuclear magnetic resonance spectroscopy (NMR). This system provided a quantitative approach to screen for ligand binding. It also allowed us to detect binding of relatively low-affinity interactions. Importantly, it allowed us to identify ligand binding effects on specific amino acid residues in an unbiased manner.

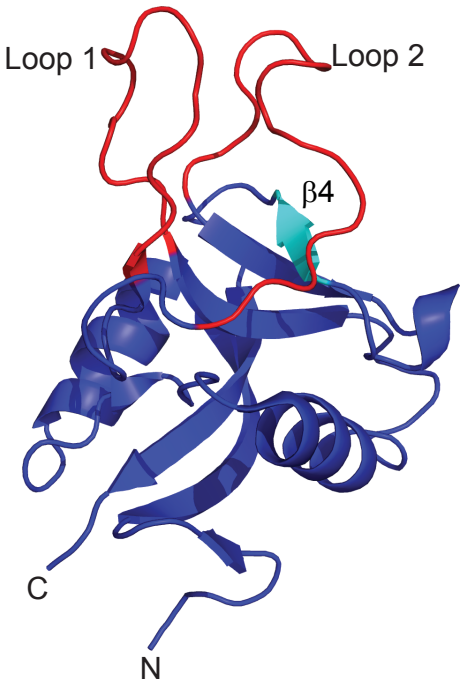
Using this approach, we have identified a Loop 1 EPN motif necessary for HIP/PAP ligand binding. The presence of this motif in other RegIII family members suggests this recognition mechanism is conserved. Further, mutation of this site resulted in significantly reduced peptidoglycan binding and bactericidal activity, establishing that this motif is critical for HIP/PAP peptidoglycan recognition and suggesting that the microbicidal activity of HIP/PAP is a multistep process that first requires peptidoglycan recognition.

Results

RegIII family members lack canonical carbohydrate binding motifs

The three-dimensional structure of HIP/PAP exhibits a C-type lectin fold with its characteristic "long loop" structure [46, 97]. This loop exhibits two distinct subdomains, that have been designated "Loop 1" (residues 107-121) and "Loop 2" (residues 130-145) (Fig. 2-1A). In other C-type lectins, such as mannose-binding protein (MBP) and DC-SIGN, a tripeptide motif in Loop 2 (EPN or QPD) is critical for Ca^{2+}-dependent binding to saccharides, interacting with both Ca^{2+} and carbohydrate to coordinate the sugar 3-OH [55] (Fig. 2-1B). A second motif (ND) located in the β4 strand also makes contacts with both Ca^{2+} and the 4-OH of the carbohydrate [52, 55]. Sugar binding and specificity is dictated by positioning of the hydrogen-bond donors and acceptors in the residues

A



B

		Loop 2	β4	saccharide specificity
HIP/PAP	128	MNYFAWERNPSTISSP-G-H----CASLSRSTAFLRWKDYNC		GlcNAc/Man
RegIIIγ	128	MNYI-WETNPSSSS---GNH----CGTLSRASGFLKWRENYC		GlcNAc/Man
MBP-C	203	LTYTNWNEG**EPN**NA---GSDED--CVLLLKN--GQ-W**ND**VPC		GlcNAc/Man
DC-SIGN	316	KQY--WNRG**EPN**NV---GEED---CAEFS----GNGW**ND**DKC		GlcNAc/Man
ASGP-R	231	TGFKNWRPE**QPD**DWYGHGLGGGEDCAHFTDDGR---W**ND**DVC		Gal
Aggrecan	2339	MQFENWRPN**QPD**NFFAAGED----CVVMIWHEKGE-W**ND**VPC		Gal

Figure 2-1: HIP/PAP lacks canonical C-type lectin carbohydrate binding motifs. (A) NMR structure of HIP/PAP (PDB code 2GO0.pdb) (97), with Loops 1 and 2 of the long loop region indicated in red, and the β4 strand in blue. **(B)** Alignment of the long loop region of RegIII family members with other members of the C-type lectin family. MBP-C and DC-SIGN each harbor Loop 2 EPN motifs that govern sugar ligand binding and confer selectivity for mannose (Man) and GlcNAc (52). Asialoglycoprotein receptor (ASGP-R) and aggrecan contain Loop 2 QPD motifs that confer selectivity for Gal and GalNAc. Despite their selective binding to GlcNAc and Man polysaccharides (56), RegIII lectins lack the Loop 2 EPN motif.

flanking the conserved proline, and is determined by the orientations of the sugar 3- and 4-OH groups. Selective binding is conferred by the Loop 2 tripeptide motif, with the EPN motif binding GlcNAc or mannose (which have equatorial 3- and 4-OH groups), and the QPD motif binding GalNAc or galactose (which have axial 4-OH groups) [52] (Fig. 2-1B). While RegIIIγ and HIP/PAP selectively recognize GlcNAc or mannose polysaccharides, they do so in a Ca^{2+}-independent manner [56, 98]. Furthermore, both proteins lack the Loop 2 EPN and the β4 ND motifs (Fig. 2-1B), suggesting a novel mode of carbohydrate recognition.

Using nuclear magnetic resonance to study HIP/PAP-peptidoglycan interactions

I initially attempted to study HIP/PAP-peptidoglycan binding using various standard biochemical assays. However, none of these assays were ideal due to several unique challenges presented by this interaction. First, I attempted studies of this interaction using peptidoglycan pull-down assays. Previous work in the lab had relied heavily on peptidoglycan pull-down assays, which depends on the insolubility of native peptidoglycan. While this assay allowed the initial discovery of the HIP/PAP-peptidoglycan interaction, it is not quantitative and thus not well suited for more detailed characterization of the binding interaction. Additionally, the dependence on ligand insolubility precludes binding studies using soluble, structurally-defined peptidoglycan analogs.

Second, I performed pilot studies using isothermal titration calorimetry (ITC). ITC has an advantage over pull-down assays in that it can be used with soluble, structurally-defined ligands. However, ITC became problematic as it provides poor readouts of relatively low-affinity binding interactions. As I will explain further below, HIP/PAP displays markedly reduced binding affinities for soluble ligands, thus making it difficult for me to detect binding using ITC. ITC is also low throughput and has large sample requirements that were difficult to achieve given the reduced binding affinities and the low solubility of soluble PGN analogs.

Finally, I undertook the study of this system using solution nuclear magnetic resonance (NMR) spectroscopy. NMR offered two key advantages over the other approaches. First, NMR is a very sensitive quantitative method, and is thus ideal for the detection and quantification of low affinity binding interactions. Second, chemical shift perturbations of backbone amide resonances, observed in $^{15}N/^{1}H$ HSQC spectra, are sensitive indicators of changes in chemical environment at specific amino acid residues. Thus, NMR also offered an unbiased approach for identifying HIP/PAP residues involved in peptidoglycan binding. This was especially important for understanding the HIP/PAP-peptidoglycan interaction, as the lack of canonical C-type lectin binding motifs suggested a novel mode of carbohydrate recognition.

Peptidoglycan induces conformational changes in a Loop 1 EPN motif

Initial NMR experiments displayed good peak dispersion for HIP/PAP, making it well suited for this system. Uniformly labeled $^{15}N/^{13}C$-HIP/PAP was used to determine backbone assignments (Fig. 2-2) using standard methods such as CBCA(CO)NH and HNCACB. We titrated solubilized *Staphylococcus aureus* peptidoglycan (sPGN) into ^{15}N-labeled HIP/PAP and monitored chemical shift and linewidth changes in $^{15}N/^{1}H$ HSQC spectra. As we had assigned the backbone chemical shifts of 98% of the HIP/PAP $^{15}N/^{1}H$ resonances, we were able to map these changes to specific sites within the protein.

During titrations with sPGN, we did not observe significant chemical shift perturbations in Loop 2, where most C-type lectins have been shown to bind their carbohydrate ligands. However, we detected small but significant ($\Delta\delta$>0.04 ppm), dose-dependent chemical shift changes in five backbone amides within Loop 1: G112, T113, E114, W117, and E118 (Fig. 2-3A, B). Loop 1 has not previously been ascribed any function in ligand recognition by C-type lectins.

Inspection of the HIP/PAP primary sequence revealed that HIP/PAP E114 is part of a Loop 1 EPN sequence that is conserved in HIP/PAP, RegIIIγ, and RegIIIβ (Fig. 2-3C), which each recognize *Bacillus subtilis* PGN in pull-down assays (Fig. 2-3D). However, E114 is not conserved in RegIIIα, which does not bind PGN (Fig. 2-3D). Furthermore, this EPN sequence is absent from Loop 1 of C-type lectins that recognize mannose and GlcNAc in a Ca^{2+}-dependent manner

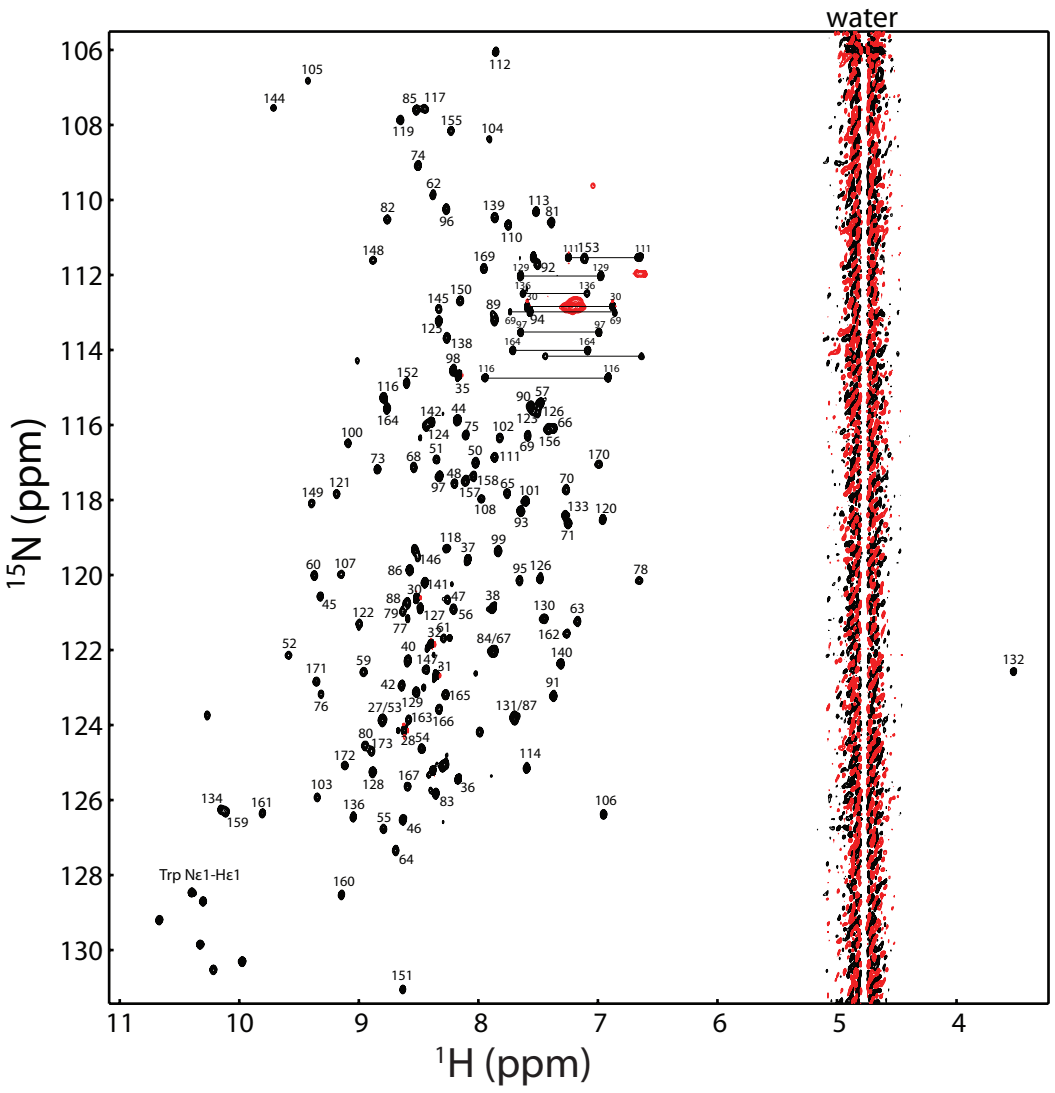


Figure 2-2: ^{15}N-HIP/PAP HSQC with backbone assignments.

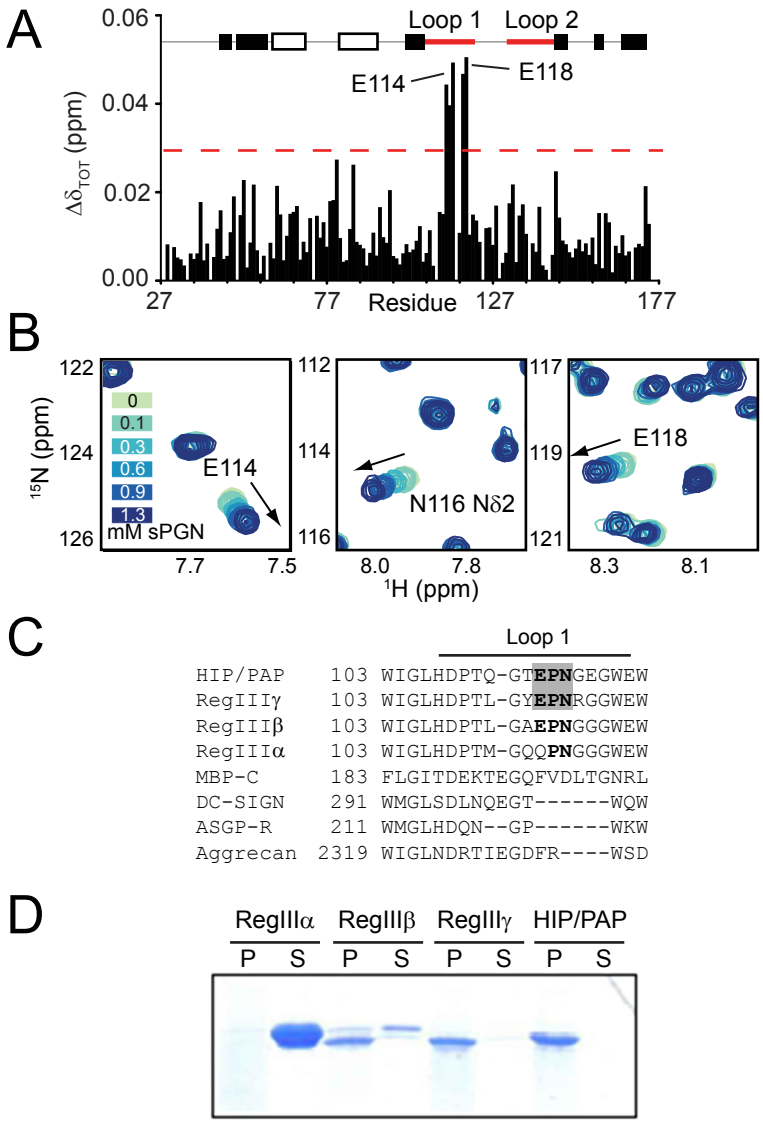


Figure 2-3: Peptidoglycan induces HIP/PAP Loop 1 conformational changes which encompass a conserved EPN motif. **(A)** Quantification of HIP/PAP chemical shift changes from the $^{15}N/^{1}H$ HSQC in the presence of 1.3 mM solubilized *Staphylococcus aureus* peptidoglycan (sPGN). The red dotted line indicates chemical shift perturbations greater than the mean + 2σ (0.029 ppm). **(B)** $^{15}N/^{1}H$ HSQC spectra of 100 μM ^{15}N-HIP/PAP titrated with increasing concentrations of sPGN. Nδ2, sidechain amide. **(C)** E114 is part of a HIP/PAP Loop 1 EPN motif that is conserved in RegIIIγ and RegIIIβ but is lacking in Ca^{2+}-dependent C-type lectins and RegIIIα. **(D)** Comparison of RegIII family members binding to Bacillus subtilis PGN by pull-down assay. RegIIIβ, RegIIIγ, and HIP/PAP each recognize PGN whereas RegIIIα does not.

(Fig. 2-3C). Cross-peaks from the N116 side-chain amide were significantly perturbed with titration of sPGN (Fig. 2-3B), consistent with this group serving as a critical hydrogen-bond donor to the sugar 3'-OH group in other lectins [46]. Because proline residues lack backbone amides, P115 is not detectable in the $^{15}N/^{1}H$ HSQC spectrum.

In C-type lectins that require Ca^{2+} to bind carbohydrate, the glutamic acid and asparagine residues of EPN function to coordinate Ca^{2+} [55]. Notably, the peptide bond between the E and P residues of the EPN motif of MBP undergoes a reversible switch from a *trans* to *cis* conformer in the presence of Ca^{2+}, orienting the protein backbone for Ca^{2+} coordination and carbohydrate binding [99]. Since RegIII lectins do not require Ca^{2+} to bind peptidoglycan, chitin, or mannan [56, 98], we sought to determine whether the HIP/PAP E-P bond also isomerizes in the presence of Ca^{2+}. The HIP/PAP $^{15}N/^{1}H$ HSQC spectrum revealed no perturbation of either the E114 or N116 cross-peaks upon Ca^{2+} addition, indicating that the E-P peptide bond does not isomerize in the presence of Ca^{2+} (Fig. 2-4A). To determine the E-P peptide bond configuration, we performed C(CO)NH TOCSY on $^{13}C/^{15}N$-labeled HIP/PAP to obtain the P115 $^{13}C\beta$ and $^{13}C\gamma$ chemical shifts. The difference between these chemical shifts reports on the configuration of the peptide bond preceding the proline, and was measured at 5.64 ppm in the absence of Ca^{2+}, indicative of the *trans* conformer [100] (Fig. 2-4B). Collectively, these

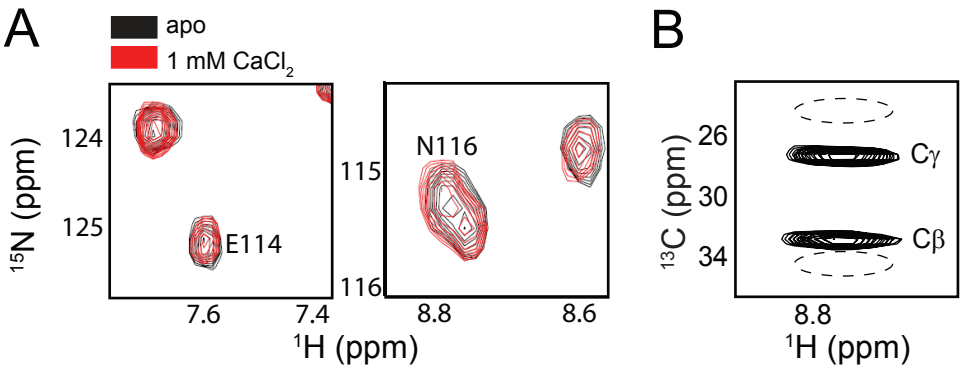


Figure 2-4: The HIP/PAP E-P bond maintains a *trans* conformation in the presence or absence of $CaCl_2$. (A) $^{15}N/^{1}H$-HSQC spectra of 100µM ^{15}N-HIP/PAP in the absence and presence of 1 mM $CaCl_2$. The lack of chemical shift perturbations indicates no isomerization of the E-P peptide bond in the presence of Ca^{2+}. **(B)** C(CO)NH TOCSY strip of $^{13}C/^{15}N$-HIP/PAP P115. The separation between the Cγ and Cβ peaks is 5.64 ppm, which is within the expected range for a *trans* peptide bond (100). Dashed circles indicate the cross-peak locations predicted for a *cis* peptide bond.

findings suggested that the EPN sequence might be important for ligand recognition in the RegIII lectins, but with a novel Ca^{2+}-independent functionality.

We fitted the Loop 1 chemical shift perturbations in the presence of sPGN to extract a binding affinity of 0.40 mM (Table 1). This K_d is several orders of magnitude higher (~10^4-fold) than the K_d of 26 nM for insoluble native peptidoglycan [56], suggesting a molecular basis for selective binding of HIP/PAP to bacterial cell surfaces despite the presence of soluble peptidoglycan fragments from the intestinal microbiota.

Since N-terminal cleavage of HIP/PAP is necessary for bactericidal activity, we wanted to investigate whether this modification also impacts peptidoglycan recognition. Titration of sPGN fragments into cleaved ^{15}N-HIP/PAP revealed a binding affinity of 0.78 mM, indicating that ligand binding is not significantly impacted by N-terminal processing.

HIP/PAP E114 is essential for peptidoglycan binding and bactericidal activity

To test the functional importance of HIP/PAP Loop 1 residues in peptidoglycan binding, we introduced conservative Glu (E) to Gln (Q) mutations into E114 and E118, the two Loop 1 residues with the most pronounced chemical shift perturbations upon sPGN titration (Fig. 2-3A). Binding of these mutants to sPGN was assessed by monitoring changes in their $^{15}N/^{1}H$ HSQC spectra upon sPGN titration. Both mutants retained the same general chemical shift patterns of

wild-type HIP/PAP, indicating that the overall protein structure remained intact (Fig. 2-5). The E118Q mutation had only a minor effect on sPGN binding affinity while the E114Q mutation weakened the interaction by >5-fold (Fig. 2-6A, Table 1). Thus, E114 plays a key role in peptidoglycan binding, suggesting the importance of the Loop 1 EPN motif in HIP/PAP peptidoglycan recognition.

We next assessed whether the reduced peptidoglycan binding of the E114 mutant correlated with HIP/PAP antimicrobial activity. HIP/PAP-E114Q showed a >6-fold reduction in bactericidal activity against *Listeria monocytogenes* (Fig. 2-6B), indicating that E114 is also critical for HIP/PAP bactericidal function. Surprisingly, we detected a similar reduction in bactericidal activity in HIP/PAP-E118Q (Fig. 2-6B). Thus, while E118 is not essential for peptidoglycan binding it is important for antimicrobial function. Together, these data establish the importance of Loop 1 residues, including E114, for HIP/PAP bactericidal function.

Table 1: Binding of PGN analogs to HIP/PAP mutants

HIP/PAP variant	sPGN K_d (mM)[1]	Chitopentaose K_d (mM)	Tri-DAP K_d (mM)
WT	0.4 ± 0.2	5.0 ± 1.6	6.4
E114Q	2.1 ± 1.6[2]	NB	8.4
E118Q	0.6 ± 0.3	nd	nd
E114Q/E118Q	2.6 ± 1.0[3]	NB	8.0

[1]K_ds were calculated based on chemical shift changes of eight Loop 1 residues (±SEM), [2]P<0.0002, [3]P<0.00003; NB, no binding detected; nd, not determined.

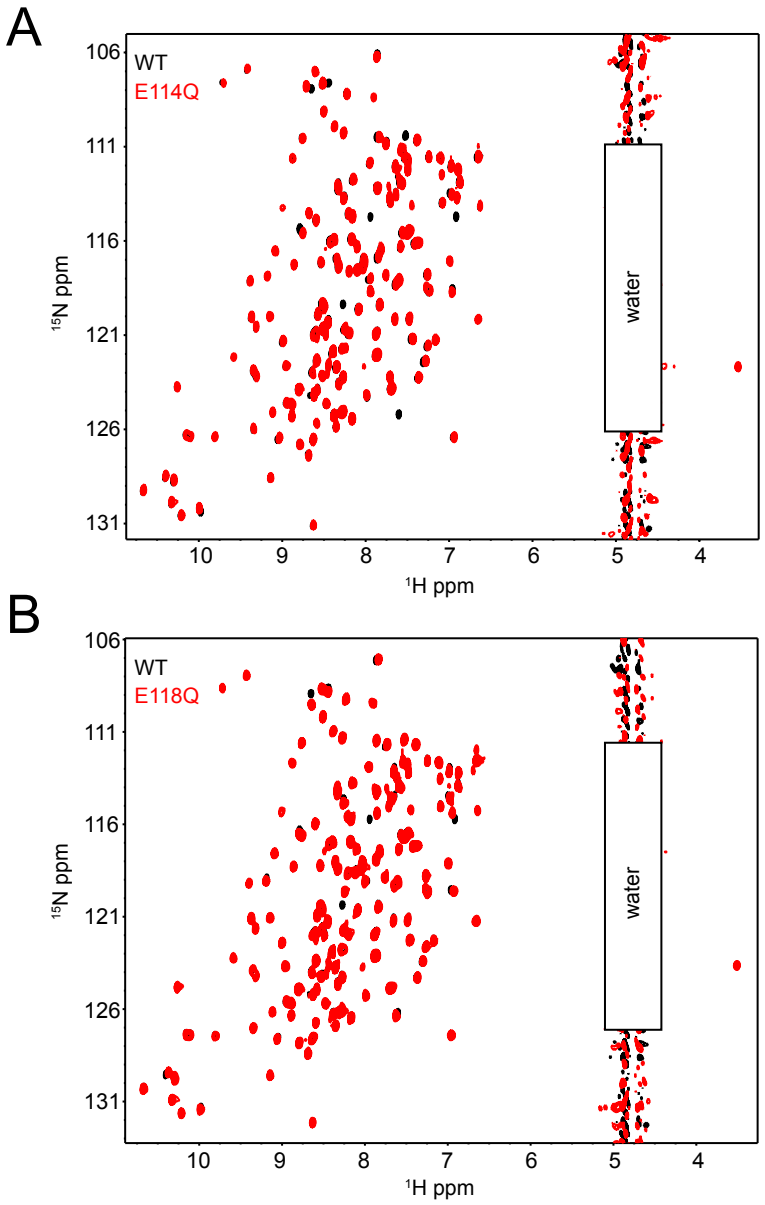


Figure 2-5: Characterization of HIP/PAP mutant proteins. $^{15}N/^{1}H$ HSQC spectra overlays of wild-type HIP/PAP, and E114Q and E118Q mutants, establishing that the mutations do not alter overall protein structure.

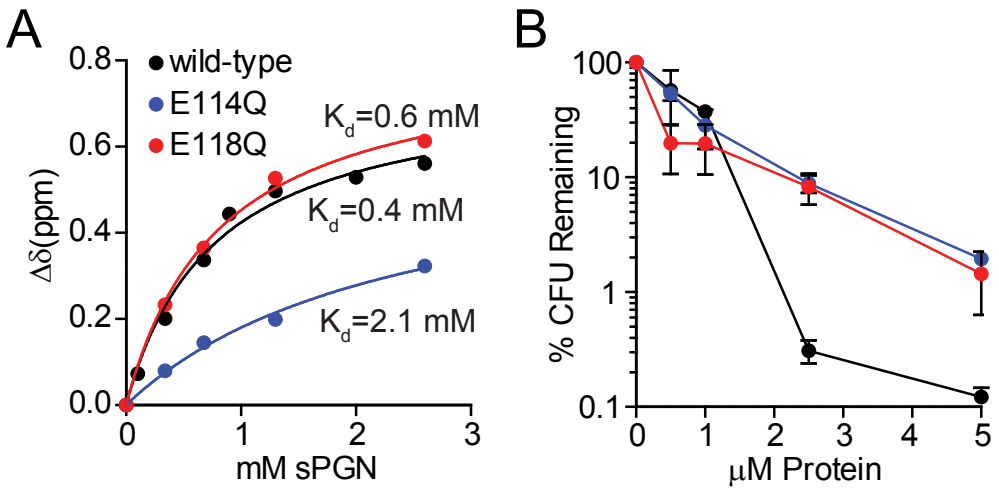


Figure 2-6: HIP/PAP E114 is essential for peptidoglycan binding and bactericidal activity. **(A)** *Staphylococcus aureus* sPGN titrated into 100 μM ^{15}N-HIP/PAP Loop 1 mutants. HIP/PAP-E114Q has a reduced affinity for sPGN, as evidenced by a decrease in the dose-dependent chemical cross-peak perturbation of T113 in the E114Q mutant. K_ds were derived from chemical shift changes at eight Loop 1 residues and averaged (values ± SEM are also reported in Table 1). **(B)** Point mutations in HIP/PAP Loop 1 attenuate bactericidal activity. *Listeria monocytogenes* was grown to mid-log phase and incubated with purified wild-type or mutant HIP/PAP. After incubation for 2 hours at 37°C, viable bacteria were quantified by dilution plating. Assays were done in triplicate. Mean ± SEM is plotted.

Discussion

In this chapter, I used solution NMR to identify the site of peptidoglycan recognition on HIP/PAP. This method highlighted a novel Loop 1 EPN motif that correlates with peptidoglycan recognition ability in RegIII family members. While CTLs such as MBP recognize mannose-type ligands in a Ca^{2+}-dependent manner via a Loop 2 EPN motif [46], HIP/PAP is unique in that it does not require Ca^{2+} for ligand binding. The importance of this motif was illustrated by the finding that HIP/PAP-E114Q has significantly reduced ability to bind peptidoglycan and kill target microbes.

In Ca^{2+}-dependent ligand binding by other mannose-specific CTLs, the Loop 2 EPN motif controls ligand specificity based on the stereochemistry of the target ligand's 3- and 4-OH groups [55]. Since HIP/PAP also recognizes mannose-type ligands via an EPN tripeptide, this motif may also determine carbohydrate specificity in RegIII family members. The fact that this interaction is Ca^{2+}-independent could be due to the lack of a neighboring ND motif, which acts in Ca^{2+} coordination in other CTLs, or because the E-P bond already has a *trans* conformation. Given that the EPN motif takes part in both Ca^{2+}-dependent and Ca^{2+}-independent binding, and due to the variation in sequences adjacent to these motifs, it is likely that this site evolved independently for each binding mechanism. While other CTLD-containing proteins, such as dectin-1 and

langerin, can also recognize saccharides without Ca^{2+}, this binding does not involve an EPN motif [101, 102].

NMR binding assays showed that N-terminal cleavage necessary for HIP/PAP microbicidal activity does not affect ligand recognition, thus indicating that binding and antibacterial function are independent molecular events. Since mutation of E114 abrogated both peptidoglycan binding and bactericidal activity, this suggests that peptidoglycan recognition is a critical first step for the microbial killing activity of HIP/PAP.

CHAPTER III

DELINEATION OF PEPTIDOGLYCAN MOIETIES REQUIRED FOR RECOGNITION BY HIP/PAP

Introduction

Previous work has shown that HIP/PAP specifically recognizes peptidoglycan [56]. However, as directly bactericidal C-type lectins (CTLs), RegIII family members are a novel group of peptidoglycan recognition proteins, and little was known about how this interaction was taking place at a molecular level. The goal of this work was therefore to identify which peptidoglycan moieties are required for this recognition.

Peptidoglycan structure is fundamentally similar in all studied bacterial species. It is composed of extended carbohydrate chains of β1,4-linked N-acetylglucosamine (GlcNAc) and N-acetylmuramic acid (MurNAc) residues and terminated by a 1,6-anydroMurNAc residue. Each MurNAc residue has a D-lactoyl group that is substituted by a short peptide. The peptide sequence is generally L-Ala-γ-D-Glu-L-Lys-D-Ala in mature PGN from Gram-positive species, whereas the L-Lys is substituted by *meso*DAP in most Gram-negative species, Bacilli, and Mycobacteria. These peptides are cross-linked to each other

either directly or via short peptides to join the carbohydrate chains and form a protective layer around the cell [58].

Since HIP/PAP is a CTL, I predicted that peptidoglycan is recognized via its carbohydrate backbone. My previous finding that the HIP/PAP Loop 1 EPN tripeptide is necessary for ligand binding further supported this idea, as the EPN motif has been shown to directly mediate binding of mannose-type ligands in other CTLs. However, previous work using pull-down assays showed that HIP/PAP does not bind chitin as efficiently as it binds peptidoglycan [56], indicating a possible role for the peptidoglycan peptide in ligand recognition.

To address this question, I used NMR to quantitatively screen for HIP/PAP binding to a range of peptidoglycan analogs. Using this approach, I found that the HIP/PAP-peptidoglycan interaction takes place through the carbohydrate backbone. Further, I found no significant differences in binding affinity for DAP-type ligands vs. Lys-type ligands, suggesting that the specificity of HIP/PAP for Gram-positive bacteria is due to the accessibility of the peptidoglycan on the microbial cell surface. Unexpectedly, I also found evidence for two distinct HIP/PAP conformational states, highlighting a potential regulatory mechanism for RegIII family activity.

Results

HIP/PAP E114 is essential for binding to the peptidoglycan carbohydrate moiety

Because the EPN motif is important for carbohydrate binding in other C-type lectins, we postulated that residues of the HIP/PAP Loop 1 EPN motif specifically recognize the peptidoglycan carbohydrate moiety. To test this idea we examined HIP/PAP binding to chitooligosaccharides, β1,4GlcNAc polysaccharides that mimic the linear structure of the peptidoglycan carbohydrate backbone (Fig. 3-1). Titration of chitopentaose $(\beta 1,4GlcNAc)_5$ into ^{15}N-labeled HIP/PAP yielded chemical shift changes in several of the same Loop 1 residues (Fig. 3-2A) and along the same trajectory (Fig. 3-2B) as for sPGN. These changes included perturbations in both the E114 backbone amide and the N116 side-chain amide peaks (Fig. 3-2B). This suggests that similar chemical moieties in sPGN and chitopentaose elicit similar conformational changes in the protein, particularly at residues that are components of the Loop 1 EPN motif. The chitopentaose K_d was 5.0 mM, indicating a moderate affinity interaction that is consistent with the millimolar K_ds typically observed for C-type lectin interactions with monosaccharides and short oligosaccharides [103, 104]. Furthermore, the moderate K_d for a soluble peptidoglycan analog is consistent with HIP/PAP's need to selectively recognize the bacterial cell surface without competitive inhibition by soluble peptidoglycan fragments from the intestinal microbiota.

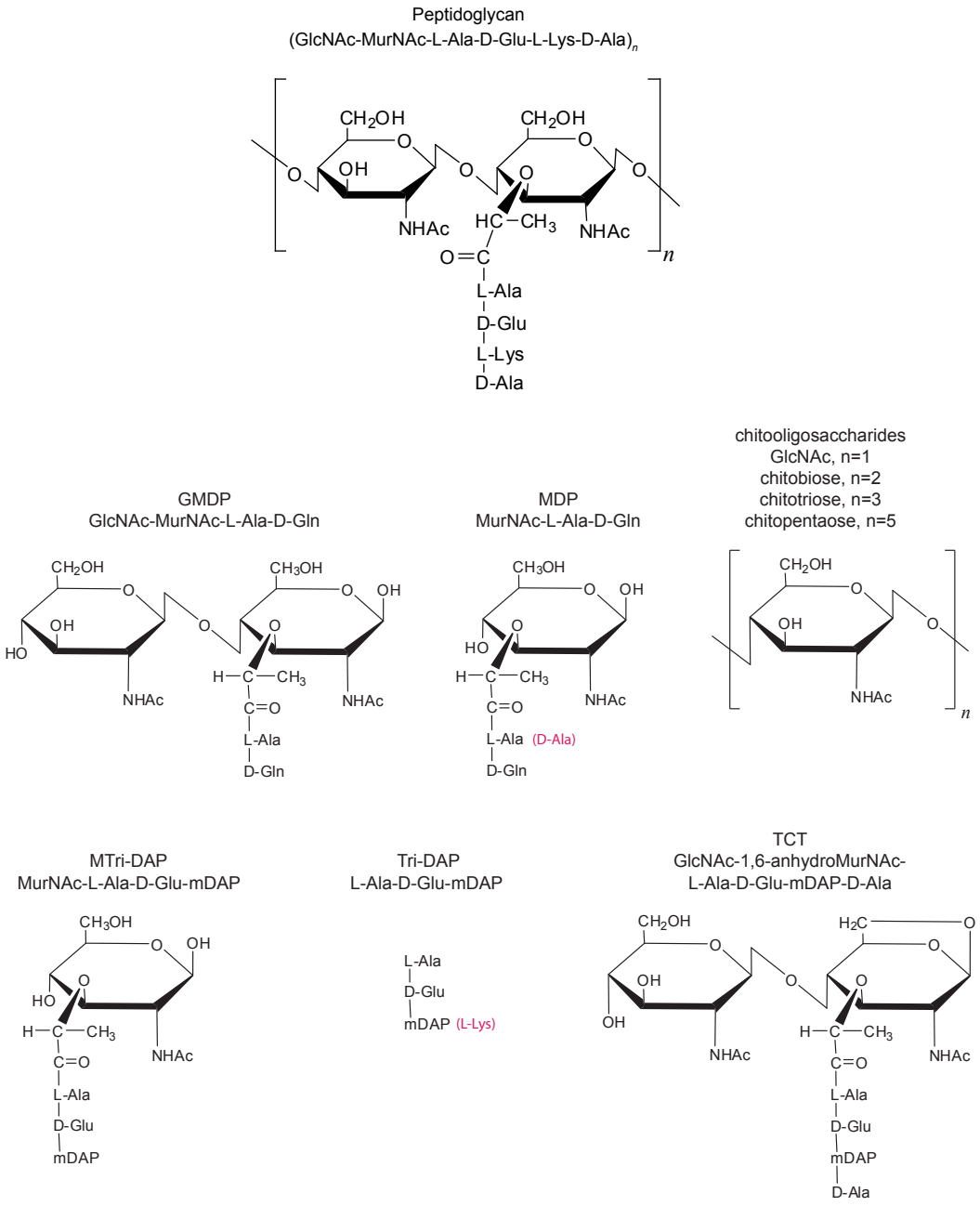


Figure 3-1: Structures of peptidoglycan derivatives used in solution NMR binding studies. mDAP indicates *meso*-diaminopimelic acid.

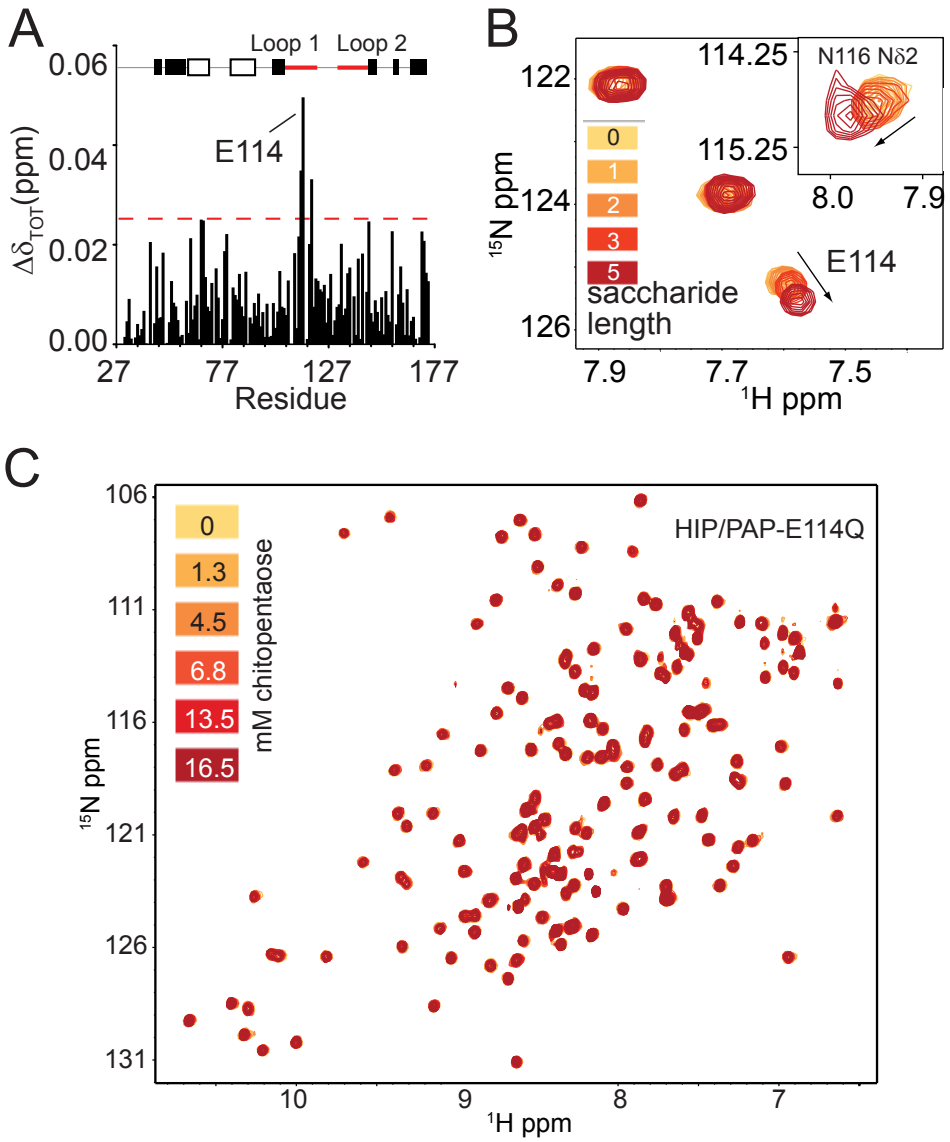


Figure 3-2: Carbohydrate recognition by the HIP/PAP Loop 1 EPN motif depends on saccharide chain length. (A) Quantification of HIP/PAP chemical shift changes from the ^{15}N/^{1}H HSQC recorded in the presence of 20 mM chitopentaose. The red dotted line indicates chemical shift perturbations greater than the mean + 2σ (0.024 ppm). **(B)** Overlay of residue E114 backbone amide and N116 sidechain amide chemical shifts with addition of 20 mM GlcNAc, chitobiose, chitotriose or chitopentaose. **(C)** Mutation of HIP/PAP E114 abolishes chitopentaose binding. Superimposed ^{15}N/^{1}H HSQC spectra of ^{15}N-labeled HIP/PAP-E114Q with titrated chitopentaose, showing an absence of chemical shift perturbations in the mutant protein.

Addition of GlcNAc monosaccharides up to 20 mM produced no detectable chemical shift perturbations (Fig. 3-2B). Chitobiose $(\beta 1,4\text{-GlcNAc})_2$ also failed to elicit significant chemical shift changes, while chitotriose $(\beta 1,4\text{-GlcNAc})_3$ titration up to 20 mM produced chemical shift perturbations that were too small to allow accurate determination of binding affinity (Fig. 3-2B). We were unable to examine binding of chitohexaose or other longer chitooligosaccharide chains due to their reduced solubility at millimolar concentrations. Binding was specific for chitooligosaccharides, as titration of cellopentaose $(\beta 1,4\text{-Glc})_5$ did not elicit any significant chemical shift changes. This is consistent with the results of pull-down assays with insoluble cellulose, the source of cellopentaose [56]. These results show that Loop 1 residues, including E114 and N116, recognize soluble analogs of the peptidoglycan carbohydrate backbone, and demonstrate that binding affinity is dictated by polysaccharide chain length.

We next tested whether E114 is essential for carbohydrate ligand binding. Chitopentaose titration up to 20 mM into the HIP/PAP-E114Q mutant yielded no detectable chemical shift perturbations (Fig. 3-2C), indicating a lack of binding. This demonstrates that E114 is critical for binding saccharides with structures that mimic the peptidoglycan carbohydrate backbone, and establishes the importance of the HIP/PAP EPN motif in carbohydrate ligand recognition.

Carbohydrate is the principal determinant of the HIP/PAP-peptidoglycan interaction

The peptidoglycan carbohydrate backbone is cross-linked by peptide moieties. To determine whether peptide might contribute to HIP/PAP recognition of peptidoglycan we examined binding of chemically defined peptidoglycan structural analogs that contain peptide. GMDP (GlcNAc-MurNAc-L-Ala-D-Gln) consists of a disaccharide moiety corresponding to the basic repeating saccharide subunit of peptidoglycan but with a truncated peptide (Fig. 3-1). Surprisingly, GMDP titrations produced a pattern of chemical shift changes that were similar to those produced by sPGN, occurring at the same Loop 1 residues (Fig. 3-3) and yielding a K_d of 11 mM (Table 2).

Table 2: HIP/PAP binding to soluble PGN analogs

Analog	K_d (mM)	Chemical shift trajectory
sPGN*	0.4 ± 0.2	Type I
sPGN (12 saccharide fraction)	1.4 ± 0.4	Type I
sPGN (5 saccharide fraction)	4.9 ± 1.3**	Type I
GlcNAc	NB	
chitobiose	NB	
chitotriose	nd	Type I
chitopentaose	5.0 ± 1.6	Type I
cellopentaose	NB	
GMDP	11.0 ± 4.5	Type II
MDP	8.6 ± 3.4	Type II
MDP-D,D	28 ± 5	Type II
MTri-DAP	5.5	Type II
Tri-DAP	6.4	Type II
Tri-Lys	8.8	Type II
TCT	7.1	Type II

*average chain length = 36 saccharides
**Projected Kd based on titration to 3.2 mM
NB, no binding detected; nd, not determined; ±SEM reported

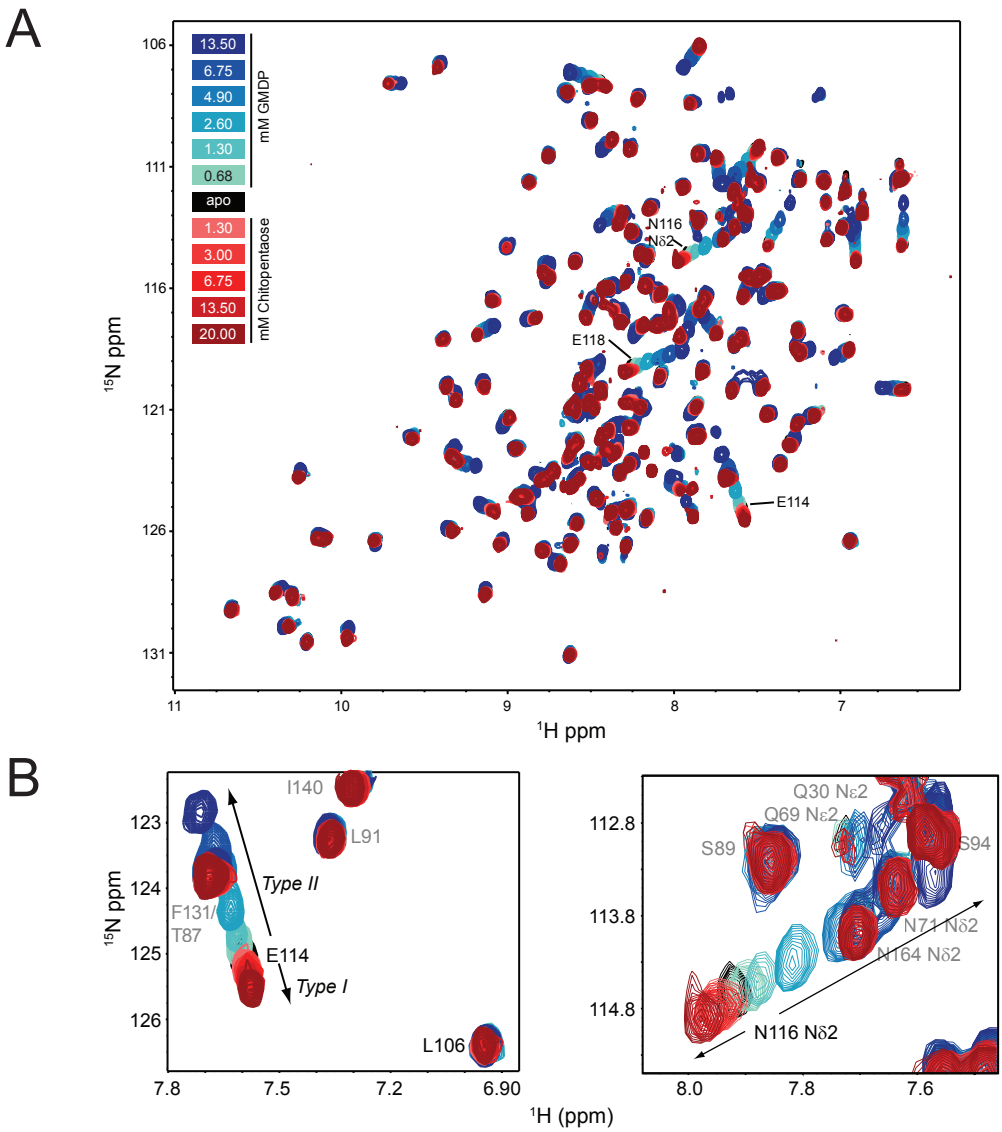


Figure 3-3: Chitopentose and GMDP induce distinct chemical shift trajectories in Loop 1 residues. **(A)** Superimposed $^{15}N/^{1}H$ HSQC spectra of ^{15}N-labeled HIP/PAP with titrated GMDP and chitopentaose. Several switch residues are indicated, showing that chitopentaose and GMDP elicit chemical shift changes on the same vector but with opposing trajectories. **(B)** Zoomed views of cross-peaks representing the E114 backbone amide and the N116 side-chain amide (N116 Nδ2). Type I and Type II chemical shift trajectories are indicated. sPGN induces only Type I trajectories (Fig. 2-3B), indicating that carbohydrate is the principal determinant of the HIP/PAP interaction with solubilized native peptidoglycan.

While GMDP induced chemical shift changes along the same vector as sPGN and chitopentaose, they were on the opposite trajectory (Fig. 3-3). This behavior indicates a fast interconversion (microseconds or faster) of Loop 1 between two conformations in the apo state, with ligand binding shifting the equilibrium more strongly towards either conformer. To distinguish between the opposing chemical shift changes, we have designated the trajectory observed for chitopentaose as "Type I" and that observed for GMDP as "Type II" (Fig. 3-3B). Because the disaccharide chitobiose did not produce detectable chemical shift changes in the NMR binding assay, these alterations suggested that HIP/PAP recognizes the GMDP peptide moiety. To further test this idea, we assessed binding of muramyl dipeptide (MDP), a monosaccharide that has the same dipeptide structure as GMDP. MDP produced a similar chemical shift pattern as seen for GMDP, with perturbations occurring along the Type II trajectory. MDP bound with a K_d of 8.6 mM, similar to that of GMDP (Table 2). An MDP variant (MDP-D,D), in which the chirality of the α-carbon of the first amino acid (Ala) is altered from the usual L-form to a D-form, bound with ~3-fold reduced affinity (28.2 mM; Table 2). This indicates sensitivity to peptide conformation, and demonstrates that HIP/PAP recognizes the dipeptide. Together, our results reveal a Loop 1 conformational switch that encompasses the EPN motif and is differentially regulated by peptidoglycan carbohydrate and peptide moieties.

We also examined binding of peptidoglycan analogs with longer peptide chains that more closely resemble the native peptidoglycan structure. Muramyl tripeptide (MTri-DAP) is a disaccharide tripeptide in which the third amino acid in the chain is diaminopimelic acid (DAP; Fig. 3-1). MTri-DAP and the related analog Tri-DAP, which consists solely of peptide, each induced Type II shifts in the same Loop 1 residues as GMDP, as well as significant chemical shift perturbations over a wider range of residues (Fig. 3-4). Additionally, we observed evidence of cooperative binding, but were unable to determine accurate K_ds for low affinity binding sites due to ligand solubility limits. The high affinity binding sites were found to have K_ds of 5.5 mM and 6.4 mM for MTri-DAP and Tri-DAP, respectively (Table 2).

Variations in peptidoglycan structure occur mainly at the third amino acid position of the stem peptide. This position is most frequently occupied by DAP or L-Lys. Tri-Lys bound to HIP/PAP with a K_d of 8.8 mM, suggesting that the presence of DAP or L-Lys in the third amino acid position does not significantly impact peptidoglycan recognition by HIP/PAP. This is consistent with our previous findings regarding the specificity of HIP/PAP bactericidal activity, which is directed against Gram-positive organisms harboring either DAP-type (e.g., *Listeria monocytogenes*) and Lys-type (e.g., *Staphylococcus aureus*, *Enterococcus faecalis*) peptidoglycan. Finally, we measured HIP/PAP binding to tracheal cytotoxin (TCT), a naturally-occurring fragment of DAP-type

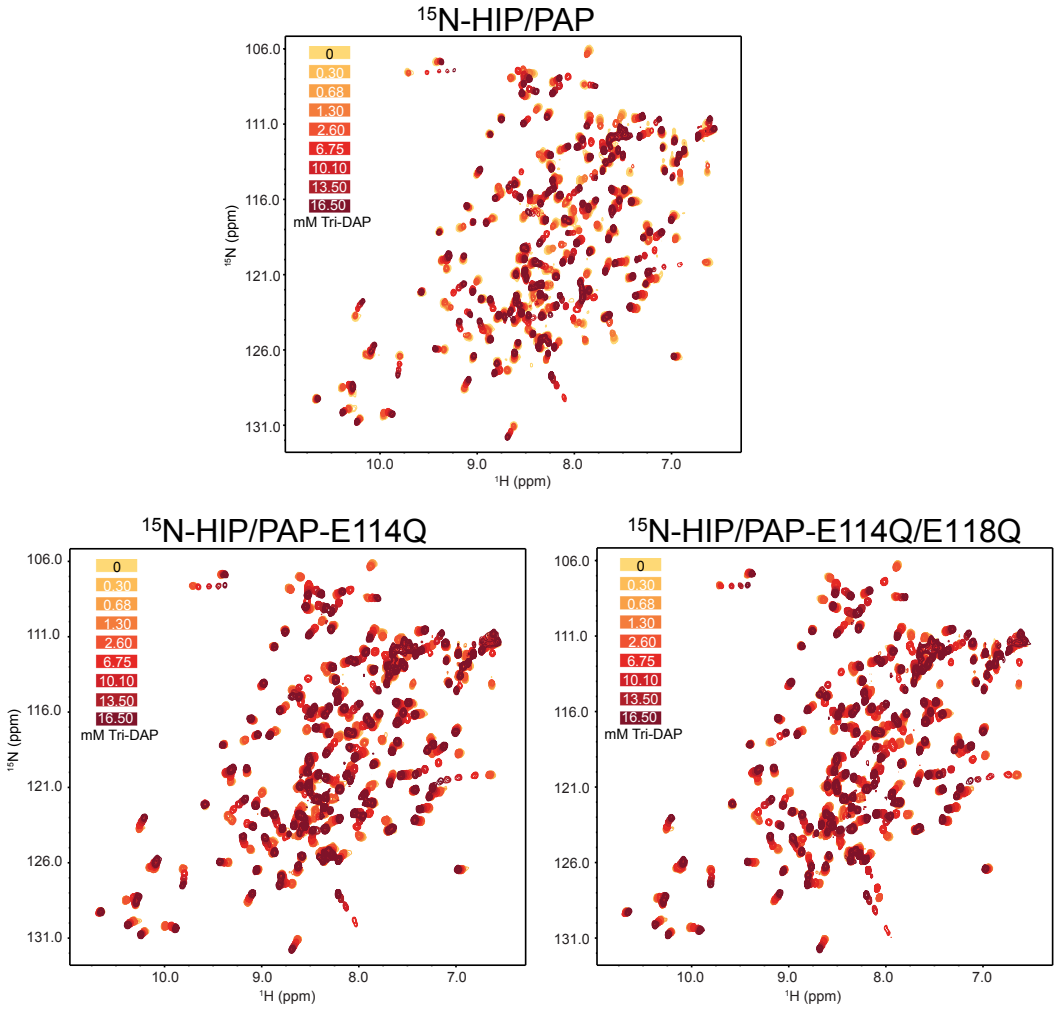


Figure 3-4: HIP/PAP E114 and E118 are not essential for recognition of the peptidoglycan peptide moiety. Tri-DAP was titrated into ^{15}N-labeled wild-type HIP/PAP, and HIP/PAP-E114Q and HIP/PAP-E114Q/E118Q mutants. Similar dose-dependent chemical shift perturbations were observed in all three proteins. As E114 is essential for recognition of chitopentaose (Fig. 3-2C), these results demonstrate that E114 is necessary for HIP/PAP recognition of peptidoglycan carbohydrate but not the peptide moiety.

peptidoglycan that is a GlcNAc-MurNAc disaccharide-tetrapeptide containing an anhydro form of MurNAc [105]. TCT bound HIP/PAP with a K_d of 7.1 mM, which is similar to that of MTri-DAP and Tri-DAP. This suggests that the fourth amino acid in the peptide chain does not alter binding affinity.

These results identify Type II chemical shift trajectories as a hallmark indicator of peptide interactions with HIP/PAP. sPGN titration elicits only Type I chemical shift trajectories (Fig. 2-3B), demonstrating that carbohydrate interactions predominate in HIP/PAP interactions with sPGN.

Titration of the pure peptide ligand Tri-DAP into the HIP/PAP-E114Q mutant yielded a K_d of 8.4 mM, similar to the K_d of 6.4 mM calculated for the wild-type protein (Table 1). Likewise, binding to HIP/PAP-E114Q/E118Q produced a K_d of 8.0 mM. This establishes that HIP/PAP E114 and E118 are dispensable for binding to the peptidoglycan peptide moiety, and suggests that the peptide interacts with HIP/PAP at a distinct site and elicits conformational changes at these Glu residues through an indirect mechanism.

To further explore the relative importance of HIP/PAP interactions with peptide, we examined $^{15}N/^{1}H$ HSQC spectra obtained in the presence of chemically-defined peptide-containing ligands, comparing them to spectra obtained during titrations of the pure saccharide ligand chitopentaose (Fig. 3-5). These comparisons revealed several peptide-specific chemical shift changes in the region of the $^{15}N/^{1}H$ HSQC spectrum corresponding to Asn and Gln sidechain

amides. Low crosspeak intensities in triple resonance datasets precluded us from making definitive assignments of these sidechain resonances; however, the peptide-specific alterations in the $^{15}N/^{1}H$ HSQC spectrum represent another set of unique hallmarks of peptide interactions with HIP/PAP. Notably, none of these peptide-specific alterations was evident upon sPGN titration, even at saturating ligand concentrations (Fig. 3-5). This supports the idea that peptide interactions are not predominant in HIP/PAP binding to sPGN, which includes fragments with longer saccharide chains (Fig. 3-6). Rather, both our NMR results and our mutagenesis studies identify carbohydrate as the principal determinant of the HIP/PAP-peptidoglycan binding interaction.

Analysis of sPGN fragment sizes

With the knowledge that the peptidoglycan peptide does not significantly improve binding affinity, I wanted to better understand the greater than 12-fold stronger binding affinity of sPGN compared to chitopentaose. My chitooligosaccharide titration data suggested that binding affinity depends on saccharide chain length and led to the hypothesis that the uncharacterized sPGN preparation must contain ligand populations that are longer than five saccharides.

To assess fragment size, the sPGN preparation was loaded onto a Sephacryl S-100 size-exclusion column (Fig 3-6A). Individual fractions were assayed for carbohydrate content by absorbance at 218 nm. Peak fractions were

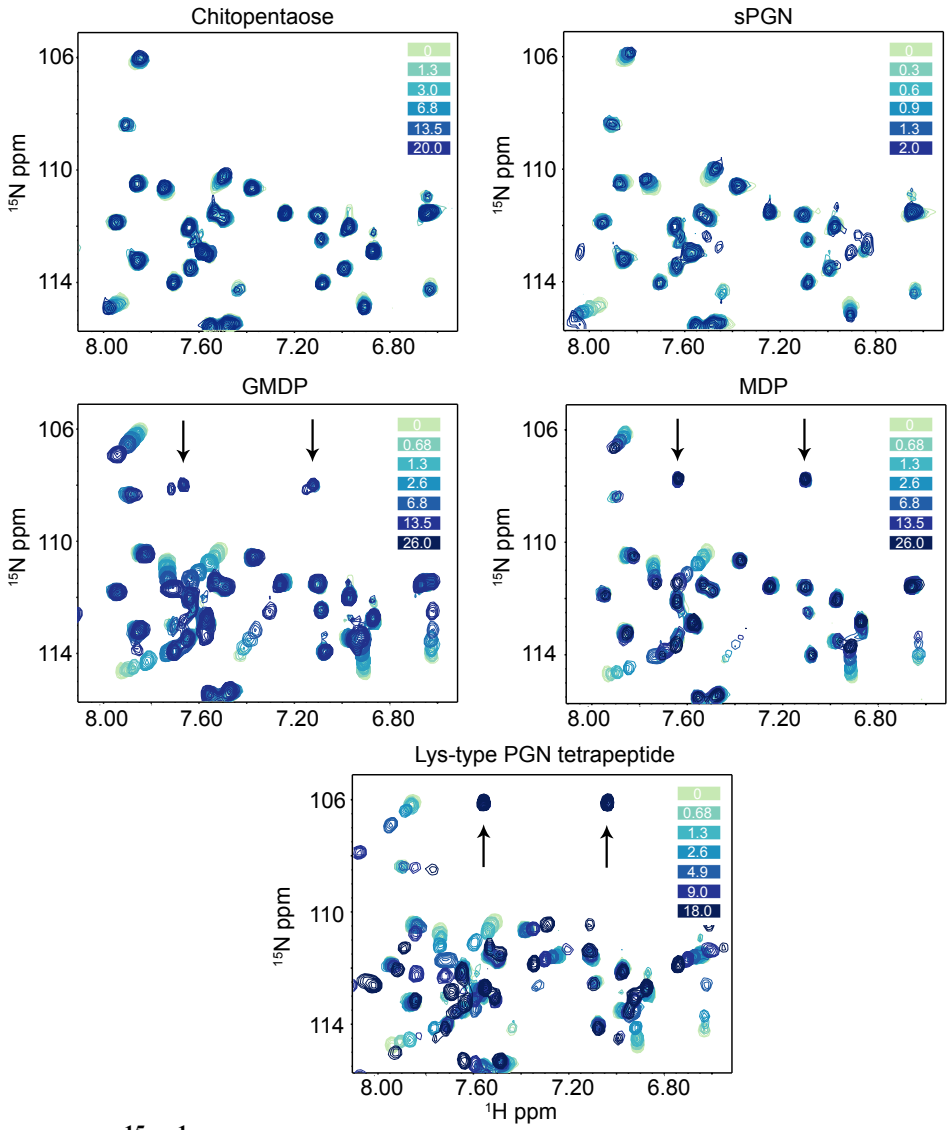


Figure 3-5: $^{15}N/^{1}H$ HSQC overlays showing that peptide-specific cross-peaks are not induced by sPGN. We examined $^{15}N/^{1}H$ HSQC spectra obtained in the presence of the peptide-containing ligands GMDP, MDP, and Lys-type tetrapeptide (L-Ala-D-Glu-L-Lys-D-Ala), comparing them to spectra obtained with titration of the pure saccharide ligand chitopentaose. These comparisons revealed the appearance of several new peptide-specific cross-peaks (indicated by arrows) on the slow exchange timescale in the region of the $^{15}N/^{1}H$ HSQC spectrum corresponding to Asn and Gln side-chain amides. Importantly, these peptide-specific cross-peaks are absent from $^{15}N/^{1}H$ HSQC spectra obtained in the presence of saturating sPGN, supporting the conclusion that carbohydrate is the principal determinant of the HIP/PAP-peptidoglycan interaction.

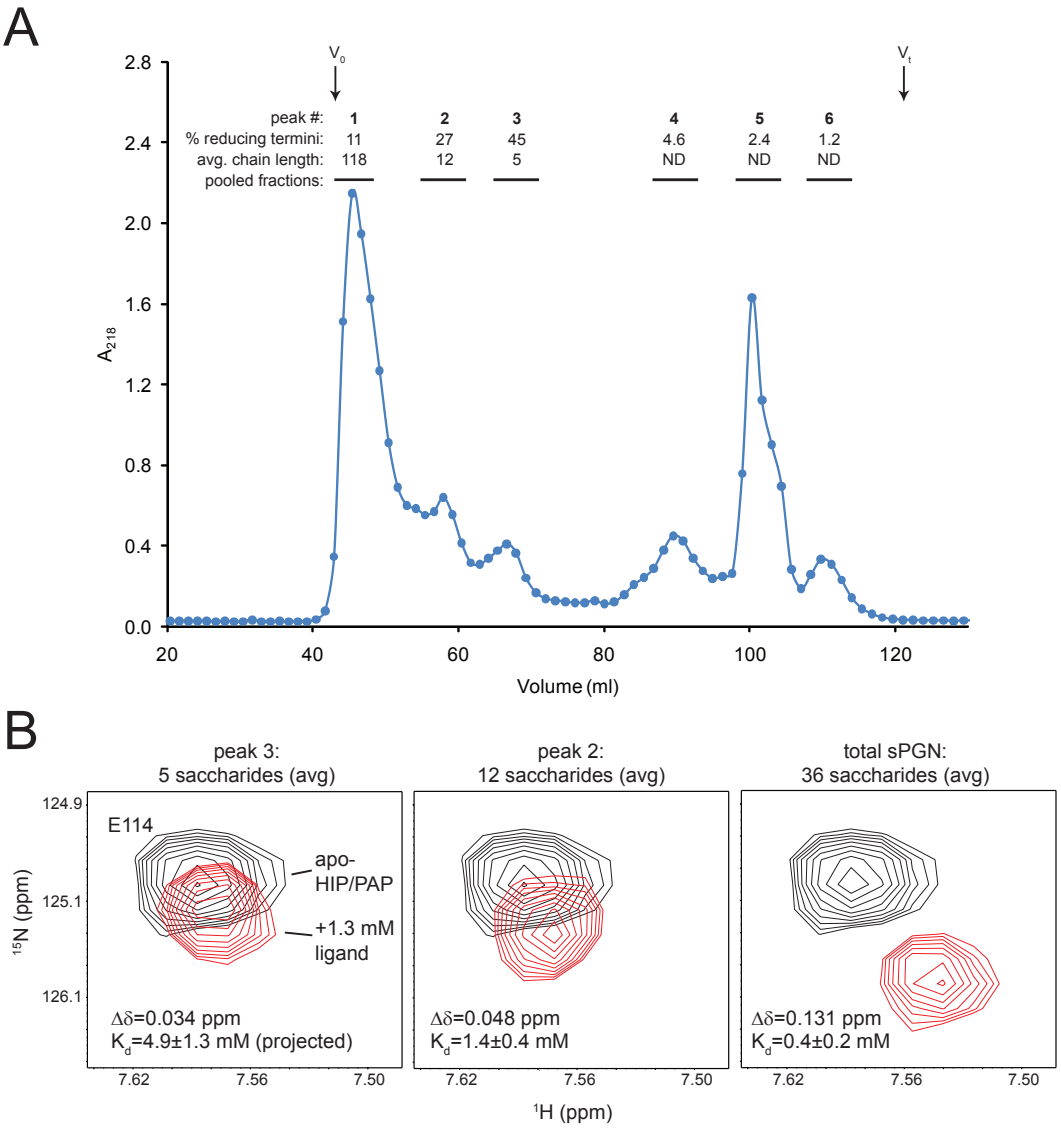


Figure 3-6: Analysis of sPGN fragment size. **(A)** Fractionation of sPGN on a Sephacryl S-100 column. The percentage of reducing termini and average calculated chain length are given for each peak. V_0, void volume; V_t, total column bed volume; ND, not determined. **(B)** $^{15}N/^{1}H$ HSQC of ^{15}N-HIP/PAP E114 indicating chemical shift perturbations in the presence of titrated sPGN fractions.

pooled as indicated (Fig. 3-6A) and assayed for moles of total carbohydrate by absorbance at 218 nm and moles of reducing termini by the 3-methyl-2-benzothiazolinone hydrazone (MBTH) method [106]. A reducing terminus is the chemically reactive end of a glycan chain [107]. Since each saccharide chain has a reducing end, we were able to calculate average carbohydrate chain length from the ratio of total carbohydrate content to moles of reducing termini. Peaks 2 and 3 corresponded to average chain lengths of 12 and five, respectively. Peaks 4-6 contained very low concentrations of reducing termini, suggesting fragments in these fractions contained anhydro groups or otherwise altered termini.

With the sPGN separated by size, we wanted to investigate whether we could detect a difference in binding affinity for peaks corresponding to different molecular weights. Peaks 2 and 3 were each pooled, concentrated, and titrated into ^{15}N-HIP/PAP for analysis by NMR (Fig. 3-6B). By deriving a binding curve from the chemical shift perturbations at E114, we determined a K_d of 1.4 mM for the 12 saccharide fraction. Due to insufficient material, a definitive K_d for the five saccharide fraction could not be calculated. However, our data indicate a K_d of greater than 3 mM, with a projected K_d of 4.9 mM (as determined by NMRView), which is in agreement with our calculated K_d of 5.0 mM for chitopentaose.

These data show that the average sPGN fragment size is significantly larger than chitopentaose and suggests a mechanism by which the binding affinity

of HIP/PAP for sPGN is increased. Further, this supports the idea that affinity of the HIP/PAP-peptidoglycan interaction is governed by carbohydrate chain length.

Discussion

In these studies, solution NMR was used to determine that peptidoglycan recognition by HIP/PAP takes place through the carbohydrate backbone. This finding shows that the Gram-positive bactericidal activity displayed by HIP/PAP is due to the accessibility of peptidoglycan on the microbial cell surface, rather than molecular variation of peptidoglycan peptide composition across bacterial species. Finally, carbohydrate and peptide moieties differentially promote switching between two potential HIP/PAP conformational states. While full understanding of this finding awaits additional studies, these results could point to an additional regulatory mechanism for HIP/PAP microbicidal activity.

The binding affinity of HIP/PAP for peptidoglycan is critically dependent on the length of the saccharide moiety. This is not surprising, given the physiological context of HIP/PAP activity. HIP/PAP functions in intestinal crypts in which bacterial are continually shedding peptidoglycan fragments. If HIP/PAP bound with high affinity to soluble peptidoglycan fragments, this would hamper the ability of HIP/PAP to bind microorganisms and exert its antibacterial activity.

Several lectins have been shown to preferentially bind longer polysaccharides. For example, MBP recognizes oligosaccharide ligands with nM affinities, while binding monosaccharide ligands with only mM affinity [49]. Likewise, mucin-binding lectins target native polymeric ligands with K_ds that are approximately 10^6-fold higher than what is observed for corresponding monosaccharides [104]. This behavior can be explained by a binding model known as the "internal diffusion model". When a protein targets ligands featuring densely clustered epitopes, the protein has the opportunity to 'bind and jump' to a neighboring epitope before fully dissociating. This activity is known as the "recapture effect" and leads to a lower apparent dissociation rate, and thus a significant increase in binding affinities for polymeric ligands [108] (Fig. 3-7).

Several observations suggest that carbohydrate plays a much more significant role in peptidoglycan recognition than peptide *in vivo*. First, the internal diffusion model provides a means for markedly increased binding affinities between HIP/PAP and intact microbes in the gut. While peptidoglycan carbohydrate backbones are cross-linked to form a vast two-dimensional surface, the peptidoglycan peptide moieties are less accessible and their dispersion means they would be unable to benefit from the recapture effect. Second, solubilized peptidoglycan induced Type I cross-peak perturbations in our NMR studies, despite the presence of peptides. This indicates that in the presence of longer saccharide chains, HIP/PAP will preferentially bind via a carbohydrate-specific

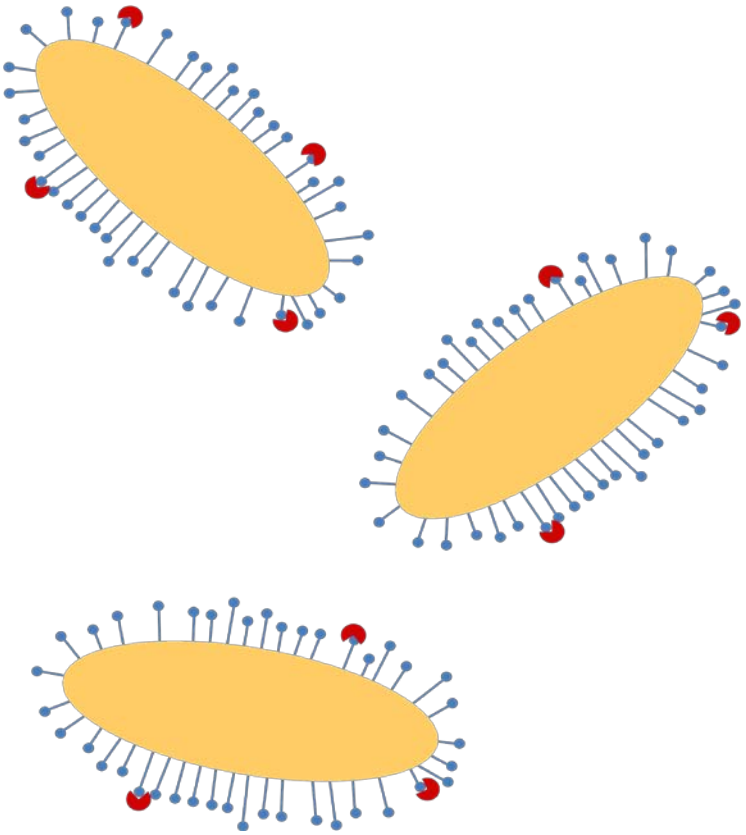


Figure 3-7: Bind-and-jump model of ligand binding. Binding affinities are greatly enhanced by the clustering of epitopes. Rather than fully dissociating, a protein has the opportunity to 'bind and jump' to a neighboring epitope, thus decreasing the apparent dissociation rate and strengthening apparent binding affinity. In this model, microbial cells are depicted in yellow. Carbohydrate epitopes are in blue. Proteins are in red.

mechanism. Third, peptide-specific cross-peaks were absent from NMR spectra collected on HIP/PAP in the presence of saturating concentrations of sPGN. Finally, microbial killing assays presented in Chapter II showed that mutating HIP/PAP E114 significantly decreased bactericidal activity. In this chapter, I have shown that the same mutant no longer recognizes chitopentaose, while the ability to recognize peptide ligands is unchanged. This supports the idea that carbohydrate is the major determinant of peptidoglycan recognition by HIP/PAP.

Since HIP/PAP targets peptidoglycan through the carbohydrate backbone, this makes the interaction unique relative to peptidoglycan recognition proteins (PGRPs). PGRPs preferentially bind smaller peptidoglycan fragments, and this interaction involves direct recognition of the peptidoglycan peptide. This allows PGRPs to discriminate between Gram-positive and Gram-negative species [109, 110]. HIP/PAP specifically targets Gram-positive microbes, but does not differentiate between DAP-type and Lys-type peptidoglycan. This suggests that HIP/PAP specificity is due to the location of peptidoglycan on the Gram-positive cell surface. Since peptidoglycan in buried in the periplasmic space of Gram-negative species, it may be inaccessible to HIP/PAP.

NMR has revealed that carbohydrate and peptide peptidoglycan moieties promote conformational switching in HIP/PAP. The area subject to switching corresponds to the HIP/PAP Loop 1 region and includes the EPN motif. Given

the importance of this site for HIP/PAP function, it is possible that this activity serves to modulate HIP/PAP activity. For example, the peptidoglycan peptide could increase ligand binding affinity by allosterically adapting Loop 1 peptidoglycan recognition, although our measurement of binding affinities for sPGN fragments suggest that this is unlikely. Alternatively, peptide recognition could assist HIP/PAP in the release of peptidoglycan, allowing the protein to target additional microbes. Better understanding of the functional significance of this observation will require further investigation.

CHAPTER IV

ASSESSMENT OF THE INFLUENCE OF PH ON HIP/PAP FUNCTION

Introduction

Several observations suggest that pH plays an important role in the regulation of HIP/PAP function. Early work to purify HIP/PAP and RegIIIγ showed that both proteins exhibit optimum stability at pH 6 or less. Furthermore, functional assays have shown that HIP/PAP bactericidal efficiency is highly pH dependent, with optimal activity found in assays performed at pH 4.5-5.5 (Fig. 4-1).

These observations are consistent with the physiological conditions under which HIP/PAP is active in the gut. While lumenal pH in the small intestine is normally above 7 [111], measurements at the epithelial surface of the jejunum have reported a pH of approximately 6 [112]. This difference is maintained by the presence of Na^+/H^+ antiporters at the intestinal surface, which pump hydrogen ions out of the cell where they are trapped by negatively charged mucopolysaccharide chains [113].

My aim in this study was therefore to look more closely at the effects of pH on HIP/PAP structure and function, including how HIP/PAP detects changes in pH and how pH influences HIP/PAP binding activity at a molecular level. This was primarily done using solution NMR and site-directed mutagenesis.

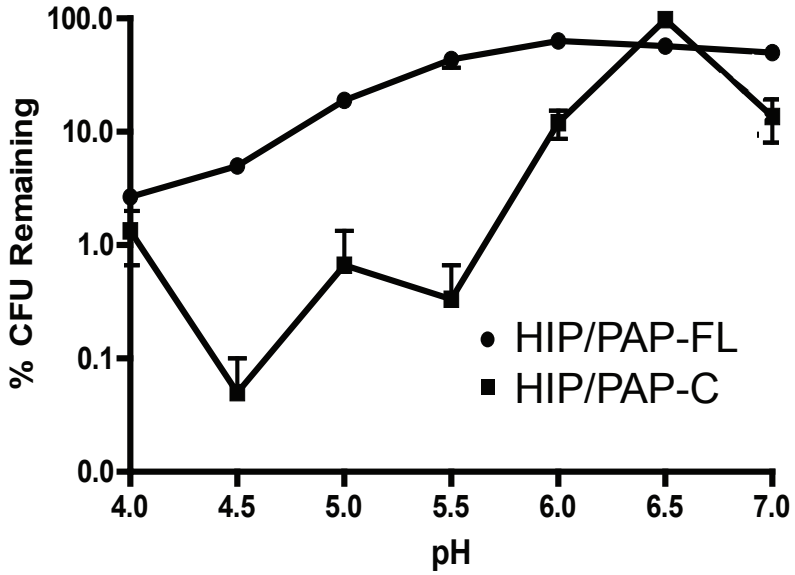


Figure 4-1: HIP/PAP bactericidal activity is pH-dependent. HIP/PAP-C microbicidal activity against *L. monocytogenes* is most efficient at pH 4.5-5.5. HIP/PAP-FL exhibits reduced activity over the same range (courtesy of Dr. Sohini Mukherjee).

In this chapter, we provide further evidence for two distinct HIP/PAP conformational states, potentially mediated by a glutamate switch. We also show that HIP/PAP E114 is critically involved in pH sensing by the protein.

Results

HIP/PAP stability is not compromised by pH adjustments within the HIP/PAP active range

Previous results had shown HIP/PAP function is closely tied to pH. To determine whether this effect was due solely to changes in the stability of the protein, we performed thermal shift assays. Thermal shift assays are based on the fact that the less stable a protein is, the lower the temperature that is required to denature it. This assay utilizes SYPRO Orange dye (Molecular Probes), which fluoresces in regions with low dielectric constants, such as the hydrophobic regions that are revealed as a protein unfolds. A real-time PCR machine is used to perform thermal denaturation of samples containing dye and the protein of interest. Fluorescence is measured as temperature increases and the protein denatures, allowing association of the dye with the newly exposed protein hydrophobic domains. Melting temperature (T_m) can be estimated by identifying the midpoint between minimal and maximal fluorescence on this curve [114].

We used the thermal shift assay to determine whether HIP/PAP stability is significantly affected by changes in pH (Fig. 4-2). T_m differences were modest

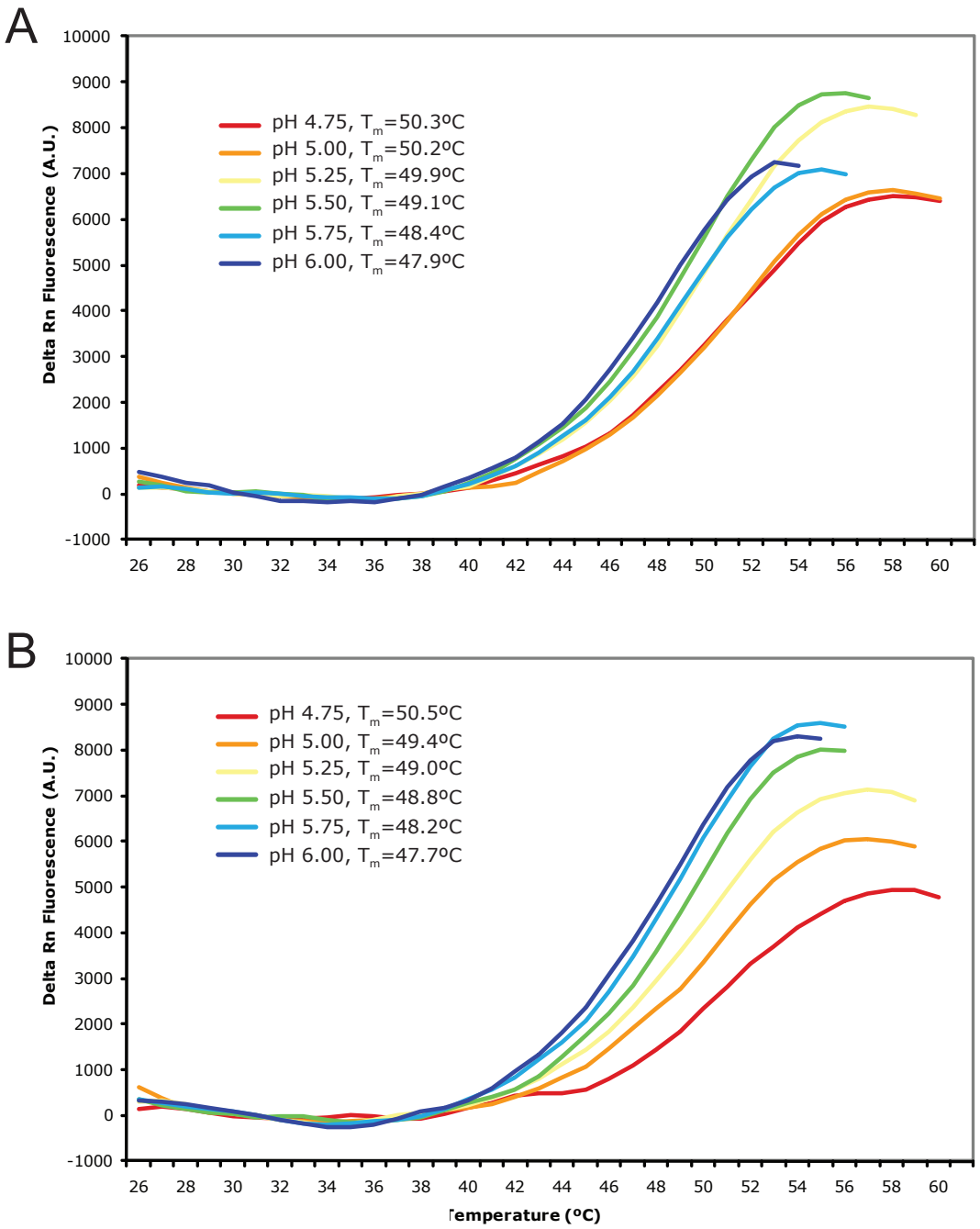


Figure 4-2: Change in pH over the HIP/PAP active range does not significantly affect protein stability. Thermal shift assay comparing stability of **(A)** HIP/PAP-FL and **(B)** HIP/PAP-C from pH 4.75-6.00.

over the pH range in which HIP/PAP is known to be active, ranging from 47.9-50.3°C for full-length HIP/PAP and 47.7-50.5°C for cleaved HIP/PAP. These data indicate that within this range, pH does not significantly affect HIP/PAP stability, suggesting that pH might instead be involved in the regulation of HIP/PAP functional activity.

Changes in pH induce switch-like behavior in HIP/PAP

Due to the importance of pH for efficient HIP/PAP function, we wanted to determine the molecular mechanism underlying this effect. We used NMR to look for structural effects at pH 4.5 vs. pH 5.5 (Fig. 4-3). Interestingly, increasing pH resulted in chemical shift perturbations corresponding to the same residues affected by peptidoglycan binding (residues 110-120). More surprisingly, these chemical shift changes were also along the same trajectory. Reducing pH relative to our standard assay conditions produced Type II shifts identical to those seen in response to the titration of peptide-containing ligands. At pH 5.5, where binding to peptidoglycan is maximal, apo HIP/PAP cross-peak perturbations are 83% in the Type I direction when compared to HIP/PAP saturated with Type I (carbohydrate) or Type II (peptide) ligands.

Together, these results provide additional support for the idea of two distinct HIP/PAP conformational states that can be induced by either ligand or pH. Further evidence can be garnered from the HIP/PAP crystal structure

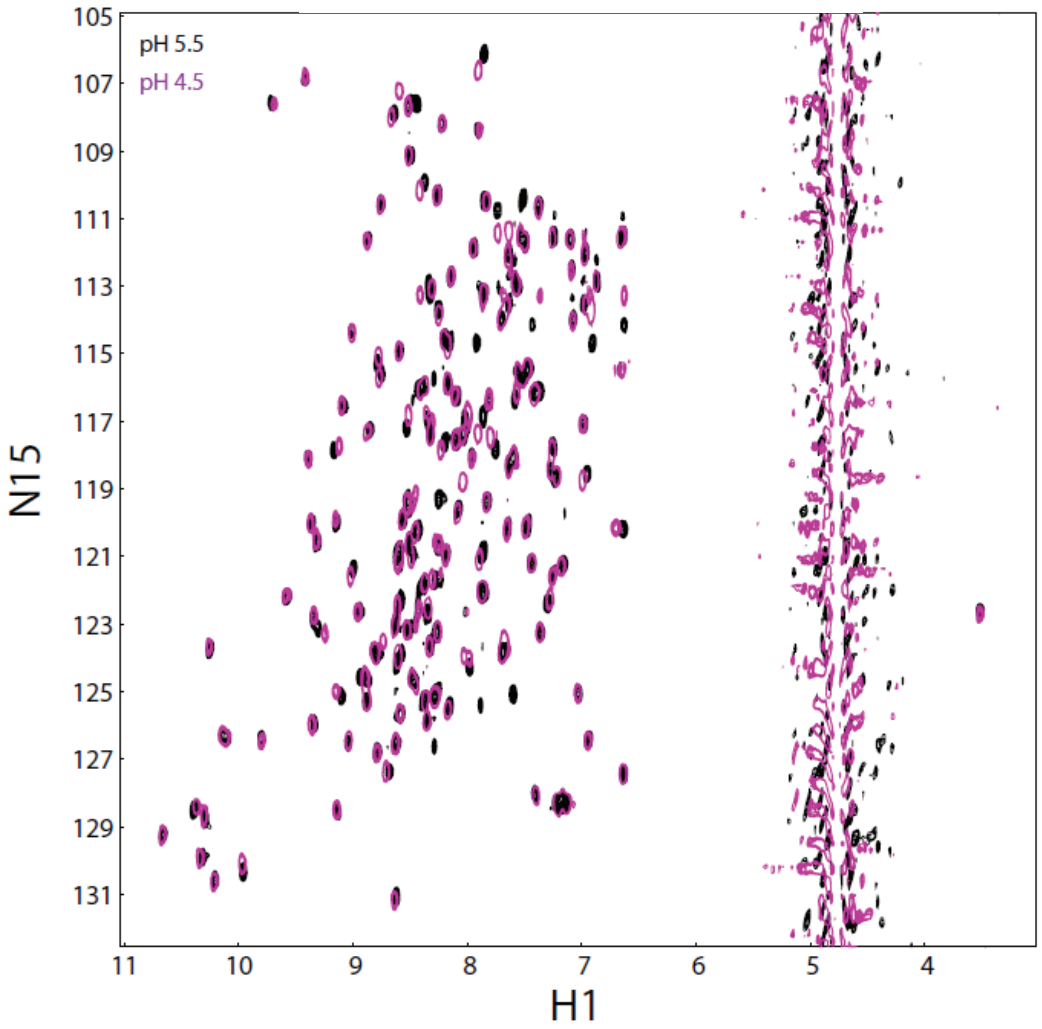


Figure 4-3: HIP/PAP exhibits pH-dependent peak shift perturbations. Reducing the buffer pH of apo-HIP/PAP results in peak shift changes of the same residues and in same trajectory (Type II) as seen in the presence of peptide-containing peptidoglycan analogs.

(1UV0.pdb). The crystal structure contains a gap in Loop 1 corresponding to P115 and N116 (Fig. 4-4A). While it was originally hypothesized that this was because of flexibility in the loop, my results suggested that it might instead be due to the presence of more than one conformational state in the crystal lattice. To investigate this idea, we used NMR to assess HIP/PAP dynamics by measuring ^{15}N relaxation rates, which are affected by the fast (ps-ns) and slow (μs-ms) motions of the backbone amides within a protein (Fig. 4-4B). Specifically, we determined rate constants for spin lattice relaxation (R_1) and spin-spin relaxation (R_2). The ratio of these values (R_1/R_2) for Loop 1 residues is indistinguishable from those determined for all other residues that make up the main body of the protein. This indicates that Loop 1 residues are not significantly more flexible than other portions of the protein and supports the idea that Loop 1 has multiple conformational states.

HIP/PAP-E114Q mutation abolishes pH-dependent peak shift perturbations

While it was known that pH exerts conformational effects on HIP/PAP, it was not known which residue or residues might be responsible for this activity. HIP/PAP binding studies highlighted the importance of E114 in ligand recognition. Interestingly, apo ^{15}N-HIP/PAP-E114Q cross-peaks were shifted towards the Type II trajectory when compared to ^{15}N-HIP/PAP-WT (Fig. 4-5). Additionally, ^{15}N-HIP/PAP-E114Q no longer exhibited pH-dependent switch-like

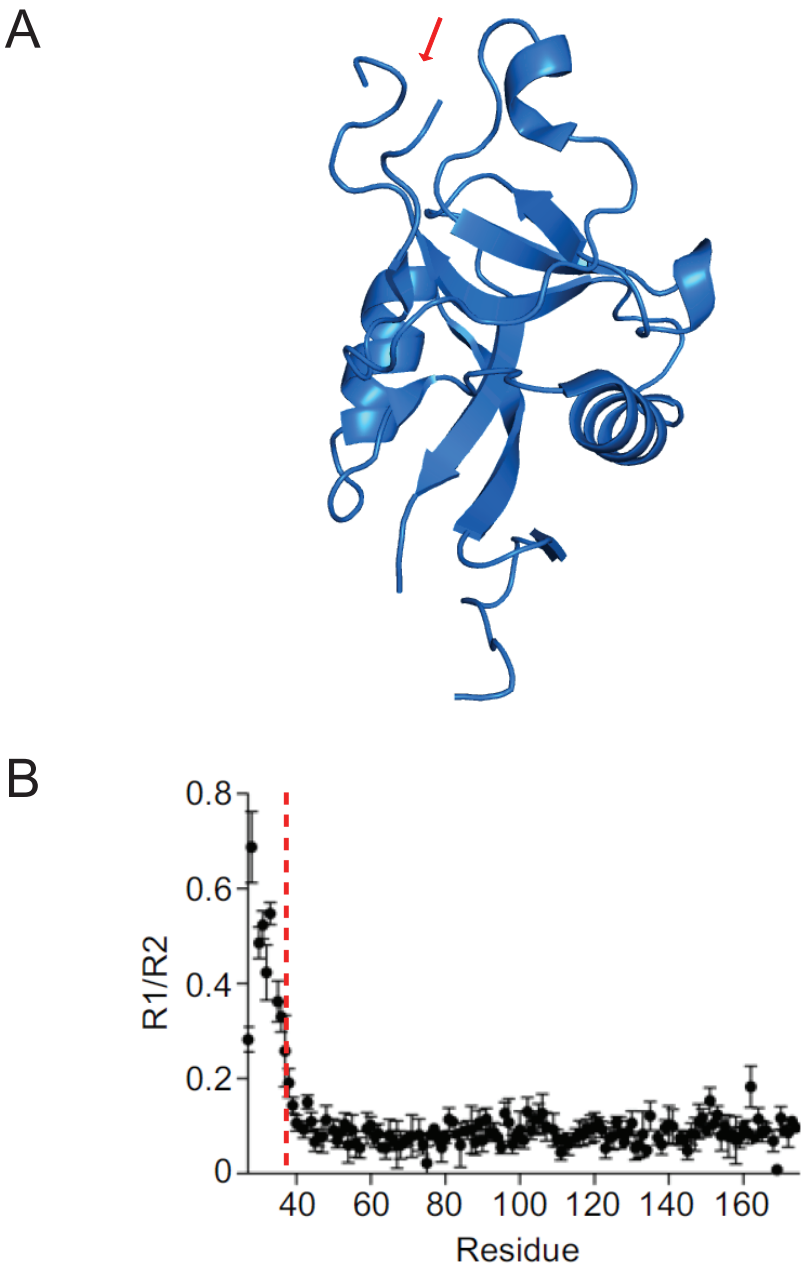


Figure 4-4: HIP/PAP Loop 1 is inflexible. **(A)** The HIP/PAP crystal structure (1UV0.pdb) contains a Loop 1 gap corresponding to P115 and N116 (highlighted by red arrow), suggesting the presence of multiple conformational states in the crystal lattice. **(B)** ^{15}N R1 and R2 relaxation rates for full-length HIP/PAP, displayed as R1/R2, show that the HIP/PAP Loop 1 switch residues have distict tertiary structure, rather than forming a flexible loop. The N-terminal trypsin cleavage between R37 and I38 is indicated by a dashed red line.

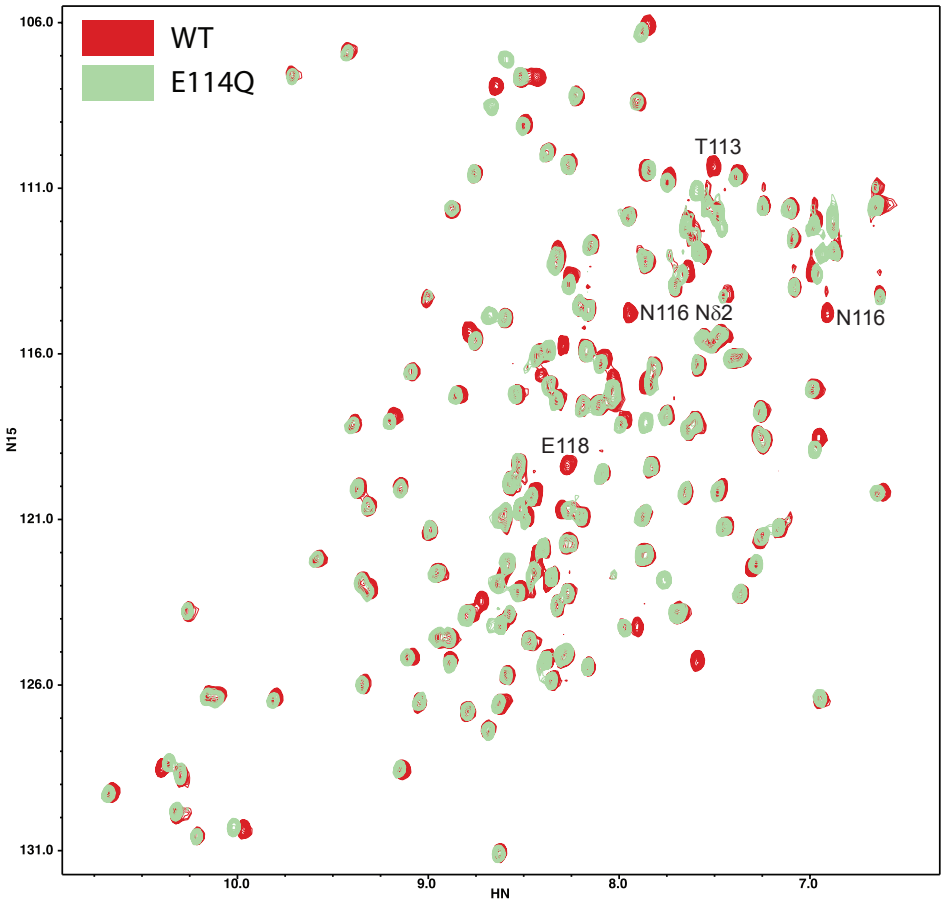


Figure 4-5: HIP/PAP-E114Q mutation causes peak shifts in the Type II direction. Overlay of ^{15}N-HIP/PAP and ^{15}N-HIP/PAP-E114Q HSQCs.

behavior. This suggested that the E114 side chain carboxylate group could be functioning as a HIP/PAP pH sensor.

In the HIP/PAP three dimensional structure, E114 and N116 are in close proximity. This makes it possible that with increased pH, the deprotonated E114 side chain could form a hydrogen bond with the N116 side chain amide. To determine whether N116 also plays a role in the pH response, I performed site-directed mutagenesis to generate a HIP/PAP-N116L mutant. We used NMR to compare ^{15}N-HIP/PAP-N116L with ^{15}N-HIP/PAP-WT (Fig. 4-6A). As with the HIP/PAP-E114Q mutant, HIP/PAP-N116L Loop 1 residues were shifted along the Type II trajectory. However, HIP/PAP-N116L still responded to changes in pH (Fig. 4-6B), indicating that N116 is not directly involved in pH sensing.

Discussion

These studies provide additional evidence for the presence of two distinct conformational states of HIP/PAP. Data presented here and in Chapter III indicate that these two states can be induced either by ligand or pH. Other C-type lectins, such as mannose-binding protein (MBP), display switch-like behavior in the presence of Ca^{2+}. In MBP, Ca^{2+} triggers isomerization of the E-P bond from a *cis* to a *trans* conformation, positioning the EPN motif for ligand binding [99]. However, we know that at pH 5.5, HIP/PAP maintains a *trans* conformer in the presence or absence of Ca^{2+}. It is possible that pH plays a role in priming the

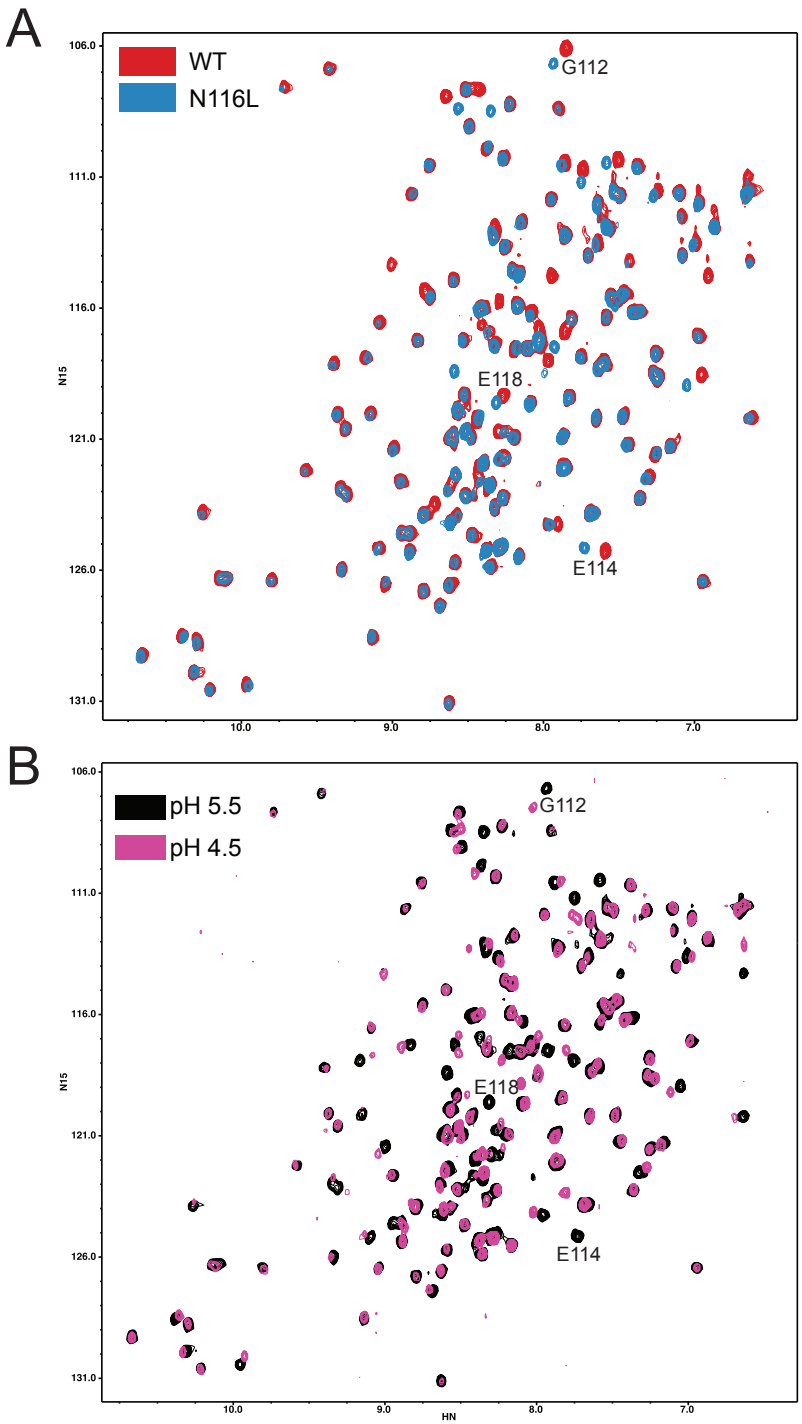


Figure 4-6: HIP/PAP-N116L mutation causes Type II peak shifts. (A) Overlay of ^{15}N-HIP/PAP and ^{15}N-HIP/PAP-N116L HSQCs. **(B)** Overlay of ^{15}N-HIP/PAP-N116L HSQCs at pH 5.5 and pH 4.5.

Loop 1 EPN motif for ligand binding in HIP/PAP. However, further experiments will be necessary to test this hypothesis. The possibility of two HIP/PAP conformational states also raises additional questions. Identifying point mutants that 'lock' HIP/PAP into one conformation or the other would allow additional analysis regarding how each state differs in function.

These results also highlight E114 as a key residue in HIP/PAP pH detection, potentially by acting as a glutamate switch. Glutamate switches have been identified in other protein families, such as the AAA+ ATPases. These proteins are found in both prokaryotic and eukaryotic species and function in a variety of cellular processes, such as protein degradation, membrane-fusion events, transcriptional regulation, and DNA replication and repair. AAA+ proteins often function as molecular motors in multicomponent assemblies to facilitate these complex reactions [115].

In many AAA+ ATPases, glutamate switches regulate ATPase activity in response to ligand binding by altering the orientation of a conserved glutamate side chain that facilitates polarization of a water molecule necessary for targeting the γ-phosphate during hydrolysis. The glutamate side chain is inactivated by a change in orientation that leads to a hydrogen bond with a nearby asparagine residue [115]. pH may be promoting this same type of behavior in the HIP/PAP EPN motif.

The positions of E114 and N116 in HIP/PAP Loop 1 allow for the possibility that the side chains of these amino acids may form a hydrogen bond when E114 is deprotonated as pH increases. This idea is supported by both the E114Q and N116L mutants. Since each mutant is unable to form this hydrogen bond, this could lead to both proteins behaving as if they were at a lower pH. Repeating pH titrations over a higher pH range will provide additional insight into this hypothesis, since we would expect to see additional perturbations of the protein as pH increases and the protein is unable to form a hydrogen bond. Triple resonance studies to determine the pK_a of the E114 side chain will allow us to better understand the role of pH in this potential mechanism. Additionally, given the effect of pH on HIP/PAP structure, more functional studies, including peptidoglycan pull-down assays, will show how these changes affect HIP/PAP activity.

HIP/PAP regulation by pH may serve as a mechanism to control HIP/PAP function *in vivo*. pH is lower near the gut epithelium than in the lumen due to the Na^+/H^+ antiporters at the cell surface [113]. By strictly limiting HIP/PAP activity to low pH regions, this could ensure that bactericidal activity specifically protects the epithelial surface from microbial attack, while shielding harmless microbes that remain in the gut lumen from innate immune destruction.

CHAPTER V

MATERIALS AND METHODS

Expression and purification of recombinant HIP/PAP

HIP/PAP was purified as previously described [57, 98]. Recombinant pET3a-pro-HIP/PAP was transformed into BL21-CodonPlus (DE3)-RILP cells. These cells were used to inoculate 500 mL LB broth containing 100 μg/mL ampicillin. Upon reaching log phase, the culture was induced with 0.4 mM IPTG and incubated for an additional three hours. Cells were then collected by centrifugation and stored overnight at -20°C. The cell pellet was resuspended in 25 mL Inclusion Body Wash Buffer 1 (20 mM Tris pH 7.5, 10 mM EDTA pH 8, 1% Triton X-100, Roche Complete EDTA-free protease inhibitors). Cells were lysed by sonication and lysates left on ice for 1 hour prior to pelleting by centrifugation. The pellet was thoroughly dispersed in Inclusion Body Wash Buffer 2 (20 mM Tris pH 7.5, 10 mM EDTA pH 8, 1% Triton X-100, 0.5 M NaCl) using a Dounce homogenizer. The centrifugation was repeated, and the pellet once again resuspended in Inclusion Body Wash Buffer 2. Following pelleting by centrifugation, the pellet was rususpended in 10 mL Resuspension Buffer (7 M guanidine-HCl, 0.15 M reduced glutathione, 0.1 M Tris pH 8, 2 mM EDTA pH 8) using a Dounce homogenizer and left to rotate overnight on a LabQuake rotator. Following centrifugation, 4 mL of inclusion bodies were

added dropwise to 200 mL HIP/PAP Refolding Buffer (50 mM Tris pH 8, 10 mM KCl, 2 mM $MgCl_2$, 2 mM $CaCl_2$, 240 mM NaCl, 500 mM guanidine-HCl, 400 mM sucrose, 500 mM arginine-HCl, 1 mM reduced glutathione, 0.1 mM oxidized glutathione) and allowed to stand at room temperature for 24 hours. The solution was centrifuged and the refolding supernatant was transferred to 45 mm diameter dialysis tubing and dialyzed against two changes of Dialysis Buffer 1 (25 mM Tris pH 7, 2 mM $CaCl_2$, 25 mM NaCl) and one change of Dialysis Buffer 2 (25 mM MES pH 5.5, 2 mM $CaCl_2$, 25 mM NaCl) at 4°C. The solution was collected from the dialysis tubing and centrifuged to remove precipitates. The supernatant was then run over a 5 mL SP-Sepharose column equilibrated in Dialysis Buffer 2. The column was washed with 25 mL Dialysis Buffer 2 + 0.15 M NaCl and eluted in 15 mL Dialysis Buffer 2 + 0.4 M NaCl. Protein was then dialyzed against Standard Assay Buffer (25 mM MES pH 5.5, 25 mM NaCl) and stored at -20^{o}C for a maximum of 1 week prior to assay. HIP/PAP in which the inhibitory N-terminal prosegment was removed was expressed from a pET3a construct and purified as described for full length HIP/PAP, except protease inhibitors were not added to Inclusion Body Wash Buffer 1 and 0.5% Triton X-100 was included in the HIP/PAP Refolding Buffer.

Cloning, expression, and purification of recombinant HIP/PAP mutant proteins

Point mutations were introduced into full length HIP/PAP using the QuikChange II Site Directed Mutagenesis Kit (Stratagene) and primers harboring the desired mutations (Table 3). Expression and purification of mutant proteins was carried out as for the wild-type proteins.

Table 3: Primers used for HIP/PAP mutagenesis

Primer Name	Sequence (5'→3')
HPapE88Q-F	CACACAGGGCACCCAGCCCAATGGAGA
HPapE88Q-R	TCTCCATTGGGCTGGGTGCCCTGTGTG
HP-E92Q-F	CCGAGCCCAATGGACAAGGTTGGGAGTGG
HP-E92Q-R	CCACTCCCAACCTTGTCCATTGGGCTCGG
HP-E88-92-Q(2)-F	CCCAGCCCAATGGACAAGGTTGGGAGTGG
HP-E88-92-Q(2)-R	CCACTCCCAACCTTGTCCATTGGGCTGGG

Preparation of solubilized peptidoglycan (sPGN)

Native peptidoglycan was isolated from *Staphylococcus aureus* by adapting a previously published protocol [116]. An overnight culture of *S. aureus* culture in BHI media was heated to 100°C for 20 minutes to kill the bacteria. The bacteria were then pelleted by centrifugation at 6500 g for 10 minutes and washed twice with cold saline (150 mM NaCl). Bacteria were then thoroughly washed with acetone on a Buchner funnel and dried at 37°C. Dried bacteria were

resuspended in 25 mL ice cold water to a final concentration of 100 mg/mL. The suspension was divided among 4 15 mL conical tubes and each aliquot was sonicated using a Misonix XL-2020 Sonicator on setting 4 with a 4.8 mm tapered probe for 2 x 1 minute with 30 seconds for cooling between pulses. Unbroken cells were removed by centrifugation at 1500 g for 10 minutes at 4°C. Bacterial cell walls were isolated by centrifugation at 6500 g for 30 minutes. The cell wall pellet was washed three times with water and resuspended to a total volume of 250 mL 50 mM phosphate buffer (pH 7.6) containing 0.2% toluene, 100 µg/mL RNase (Sigma) and 50 µg/mL DNase (Sigma). The resulting suspension was incubated for 18 hours at 37°C with slow mixing. Trypsin was added to a final concentration of 200 µg/mL and the 18 hour incubation was repeated. The cell walls were once again collected by centrifugation at 6500 g for 30 minutes and washed 4X with milliQH_2O and lyophilized. The dried cell walls were resuspended in 5% TCA (1 g in 250 mL) and incubated for 18 hours at 22°C. The insoluble material was collected by centrifugation at 6500 g for 30 minutes, and the pellet was washed 2X with 5% TCA. The pellet was heated to 90°C for 15 minutes and then washed 3X each with milliQH_2O and acetone. The final pellet was air dried at 37°C. The insoluble peptidoglycan pellet was resuspended to 10 mg/mL in 50 mM Tris-HCl pH 8, and sonicated for 30 minutes using the parameters described above. The suspension was heated to 90°C for 30 min for sterilization. After cooling to 37°C, lysostaphin (Sigma) was added to a final

concentration of 50 μg/mL. Following overnight incubation at 37°C, the preparation was heated to 95°C for 5 minutes to denature the lysostaphin and then centrifuged for 10 min at 10,000 g to remove remaining insoluble peptidoglycan. The soluble fraction was lyophilized, resuspended in 1 mL milliQH_2O, and desalted on a Sephadex G-10 column (GE Healthcare).

Ligand preparation

N-acetyl-chitooligosaccharides (Seikagaku), cellopentaose (Seikagaku), GMDP (Calbiochem), MDP (Sigma), MDP-DD (InvivoGen), GlcNAc (Sigma), MTri-DAP (InvivoGen), Tri-DAP (InvivoGen), and Tri-Lys (InvivoGen) were obtained from commercial sources. Tracheal cytotoxin (TCT) was obtained from our collaborator, Dr. William E. Goldman (University of North Carolina, Chapel Hill). TCT was isolated from *Bordetella pertussis* and purified as previously described [105].

NMR spectroscopy

To prepare samples for NMR experiments, *Escherichia coli* BL21-CodonPlus (DE3)-RILP cells were transformed with expression plasmids were grown in M9 minimal media containing 1 g/L of $^{15}NH_4Cl$ for uniformly ^{15}N labeled samples, further substituting unlabeled glucose with 3 g/L of $^{13}C_6$-glucose for uniformly $^{15}N/^{13}C$-labeled samples. HIP/PAP was purified as previously

described [98]. Backbone ^{15}N, ^{13}C, and ^{1}H chemical shift assignments for HIP/PAP were generated from triple resonance NMR data recorded at 25°C using a cryoprobe-equipped Varian Inova 600 mHz spectrometer. Backbone assignments were made from a 300 μM sample of $^{15}N/^{13}C$-labeled protein in 25 mM MES pH 5.5, 25 mM NaCl using standard methods (10). All spectra were processed using NMRPipe [117] and analyzed with NMRView [118].

Peptidoglycan binding assays

Bacillus subtilis peptidoglycan (Fluka) was pelleted and resuspended in Standard Assay Buffer (SAB; 25 mM MES pH 5.5, 25 mM NaCl) with EDTA-free protease inhibitors (Roche) and pre-incubated at 4°C with rocking for 2 hours. Peptidoglycan was then pelleted and washed three times with SAB. For each protein of interest, 20 μg protein was added to 100 μg peptidoglycan in SAB to a final volume of 300 μL. Samples were incubated overnight with rocking at 4°C. Peptidoglycan was pelleted at 16100g for 5 minutes and washed 3 times in SAB. 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% SDS-polyacrylamide gels. Gels were stained with Coomassie Blue.

Antimicrobial assays

Antimicrobial assays were done as previously described using HIP/PAP and HIP/PAP mutants engineered to lack the inhibitory N-terminal pro-segment [56, 57]. *Listeria monocytogenes* was exposed to the indicated lectin concentrations at 37^{o}C for 2 hr, and surviving bacteria were quantified by dilution plating as described.

Gel filtration of solubilized peptidoglycan

sPGN was loaded onto a 1.5 x 70 cm Sephacryl S-100 column for analysis by size-exclusion chromatography. Individual fractions were assayed for carbohydrate content by absorbance at 218 nm.

Determination of carbohydrate reducing termini in solubilized peptidoglycan

Concentrations of carbohydrate reducing termini in solubilized peptidoglycan preparations were determined using a previously published method [106]. Two hundred microliter aliquots of GlcNAc standards and sPGN samples were transferred to 2-mL screw cap tubes. Each sample was mixed with 200 µL 0.5 N NaOH. Equal volumes of 3 mg/mL 3-methyl-2-benzothiazolinone hydrazone (MBTH) and 1 mg/mL DTT were combined and 200 µL of the fresh mixture was added to each tube. Reactions were incubated at 80°C for 15 minutes and then immediately mixed with 400 µL of a solution containing 0.5%

$(FeNH_4(SO_4)_2)\cdot 12H_2O$, 0.5% sulfamic acid, and 0.25 N HCl. Samples were allowed to cool to room temperature and the absorbance of each was determined at 620 nm.

Thermal shift assay

Thermal shift assays were performed using an Applied Biosystems 7500 RT-PCR machine. Samples containing SYPRO Orange dye (Molecular Probes) and HIP/PAP-FL or HIP/PAP-C protein were set up in citrate buffer due to temperature-dependent pH changes in MES [119]. Reactions varied in pH from 4.75-6.0 in increments of 0.25. Temperature was increased in increments of 1°C. Fluorescence was measured at each temperature step.

CHAPTER VI

DISCUSSION AND FUTURE DIRECTIONS

Discussion

The mammalian gut epithelium is under constant microbial attack due to the dense bacterial load in the lumen. Epithelial cells produce a variety of antimicrobial proteins to cope with this continual threat [6]. Work in our laboratory led to the discovery of RegIII family of C-type lectins as a novel group of antimicrobial proteins.

Previous work has shown that the intestinal C-type lectins RegIIIγ (mouse) and HIP/PAP (human) are directly bactericidal against Gram-positive bacteria [56]. It was further shown that these proteins target bacteria by binding to cell wall peptidoglycan. Both the peptidoglycan binding activity and the direct bactericidal activity represent novel biological functions for C-type lectins, and thus the mechanisms underlying lectin-mediated bacterial recognition and killing were previously unknown.

In this thesis I have used solution NMR to gain insight into the molecular basis for peptidoglycan recognition by HIP/PAP, a bactericidal C-type lectin. Surprisingly, we found that Ca^{2+}-independent recognition of the peptidoglycan carbohydrate backbone involves an EPN motif similar to that required for Ca^{2+}-dependent recognition of mannose and GlcNAc-containing saccharides by C-type

lectins such as MBP [55]. Further, the presence of an EPN tripeptide correlates with peptidoglycan binding activity among the members of the mouse and human RegIII lectin family. Mutation of the Glu residue (E114) of this tripeptide revealed the functional importance of this site in HIP/PAP saccharide binding, peptidoglycan recognition, and bacterial killing. These studies have thus uncovered a unique mechanism of ligand recognition by a C-type lectin that plays a critical role in innate immune defense of the intestine.

EPN function in HIP/PAP

Ca^{2+}-dependent saccharide recognition by mannose-specific lectins such as MBP and DC-SIGN is critically dependent on a Loop 2 EPN motif that coordinates Ca^{2+} and forms hydrogen bonds with sugar hydroxyls [52, 55]. The EPN sequence is an essential determinant of specificity for GlcNAc and mannose, which are related by the orientations of their 3- and 4-OH groups [52]. Given that HIP/PAP also selectively recognizes GlcNAc and mannose polysaccharides [56], we propose that the HIP/PAP Loop 1 EPN motif may be a key determinant of saccharide specificity. The loss of Ca^{2+} dependence in the Loop 1 EPN-carbohydrate interaction could be due to the absence of additional Ca^{2+} coordination sites such as the ND dipeptide, or the constitutive *trans* conformer of the EPN proline. Determination of a high resolution structure of HIP/PAP in complex with a peptidoglycan structural analog will be required to elucidate the

exact mechanism by which the HIP/PAP EPN motif supports Ca^{2+}-independent carbohydrate binding. Nevertheless, our findings suggest that the EPN tripeptide is a flexible motif that plays multiple roles in lectin-carbohydrate interactions. The lack of homology between the sequences immediately surrounding the Loop 1 and Loop 2 EPN motifs further suggests that the motif evolved independently in the RegIII subfamily, rather than arising from a common C-type lectin ancestor.

A limited number of other CTLD-containing proteins also recognize carbohydrates in a Ca^{2+}-independent manner via mechanisms that are distinct from HIP/PAP. Dectin-1 binds to oligomers of the fungal cell wall component β1,3-glucan. Unlike HIP/PAP, dectin-1 does not have a Loop 1 EPN motif, and binding to β-glucan involves hydrophobic interactions with the β3 strands of dectin-1 dimers [101]. The C-type lectin langerin also provides an example of Ca^{2+}-independent carbohydrate binding. However, the langerin Ca^{2+}-independent site lies in Loop 2, and is secondary to a canonical Ca^{2+}-dependent site that binds mannose via a mechanism similar to that of MBP and DC-SIGN [102].

The "bind and jump" model of peptidoglycan recognition

HIP/PAP carries out its bactericidal functions in the intestinal lumen, an environment rich in soluble peptidoglycan fragments derived from resident bacteria. Because HIP/PAP bactericidal activity requires recognition of intact bacteria [56], high affinity recognition of soluble peptidoglycan fragments would

be counterproductive by competitively inhibiting cell wall binding. Our finding that HIP/PAP binds to its native polymeric ligand with a much higher affinity than to soluble fragments provides a molecular explanation for the ability of HIP/PAP to selectively bind intact bacteria in the lumenal environment.

A low affinity for soluble analogs is a characteristic feature of lectins that bind to linear polysaccharides having clustered and repeated epitopes. For example, mucin-binding lectins bind to their native extended ligands with affinities that are up to 10^6-fold higher than for the corresponding monovalent carbohydrates, which are bound with millimolar K_ds [104]. Thermodynamic studies have elucidated the mechanism by which mucin-binding lectins achieve selective high-affinity binding to extended native ligands, showing that the lectins "bind and jump" from carbohydrate epitope to epitope, resulting in a gradient of decreasing microaffinity constants. This "recapture effect" results in a decreased off-rate and thus an increased apparent binding affinity [108]. Thus, a hallmark feature of these interactions is a direct relationship between binding affinity and carbohydrate chain length. Our finding that the strength of HIP/PAP binding to chitooligosaccharides depends critically on saccharide chain length strongly suggests that HIP/PAP utilizes a similar binding mechanism. Further, it indicates that direct carbohydrate binding is more important than peptide recognition *in vivo*, since peptide binding affinities would not be strengthened by this model. We therefore propose that dynamic recognition of the multivalent carbohydrate

backbone of native peptidoglycan is an essential mechanism governing selective high affinity interactions between HIP/PAP and the bacterial cell wall.

HIP/PAP utilizes a novel mechanism of peptidoglycan recognition

The molecular basis of HIP/PAP recognition of peptidoglycan differs fundamentally from that of the peptidoglycan recognition protein family (PGRPs) (Fig.6-1). *Drosophila* PGRPs bind with high affinity to soluble PGN fragments shed by bacteria, resulting in either direct enzymatic cleavage or initiating signaling cascades that govern expression of antibacterial peptides. Peptidoglycan recognition by PGRPs involves moderate- to high-affinity interactions with the peptide moiety, allowing discrimination between Gram-positive and Gram-negative bacteria and the initiation of immune responses specific for either group [109, 110]. Mammalian PGLYRPs similarly recognize the peptidoglycan peptide moiety, but in contrast to many *Drosophila* PGRPs, carry out direct bacterial killing [63]. In the case of PGLYRP2, microbicidal activity is due to N-acetylmuramoyl-L-alanine amidase activity, which cleaves peptidoglycan peptides from the carbohydrate backbone and damages the integrity of the bacterial cell wall [82, 83]. While the bactericidal mechanisms of the other PGLYRP family members have yet to be uncovered, evidence suggests they act to inhibit peptidoglycan synthesis [63].

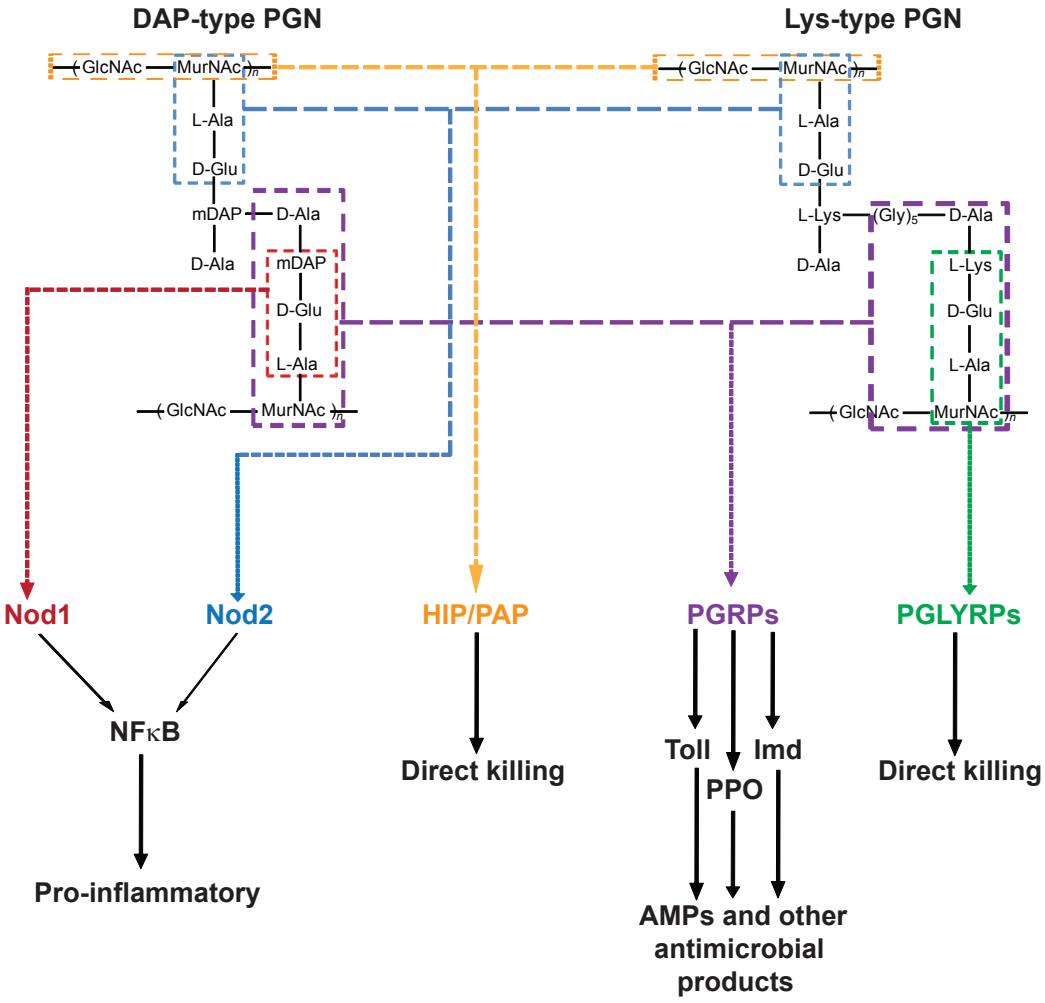


Figure 6-1: Ligand recognition and immune response by peptidoglycan recognition proteins. PGN, peptidoglycan; PPO, pro-phenol-oxidase.

HIP/PAP recognizes peptidoglycan through a unique mechanism. In contrast to both *Drosophila* and mammalian PGRPs, it interacts predominantly with the carbohydrate moiety of peptidoglycan, and likely accomplishes selective high affinity binding to the bacterial cell wall through multivalent interactions with the highly clustered presentation of carbohydrate epitopes. Additionally, this suggests that HIP/PAP specificity for Gram-positive bacteria is due to the accessibility of peptidoglycan on the cell surface, rather than the composition of the peptidoglycan peptide.

The influence of pH on HIP/PAP recognition of peptidoglycan

While the pH in the lumen of the small intestine is typically neutral [111], the microenvironment at the epithelium is acidic [112]. Work in our laboratory has shown that HIP/PAP bactericidal activity is maximal below pH 6. Our finding that pH modulates HIP/PAP conformational switching thus suggests a mechanism by which microbicidal activity can be specifically targeted towards encroaching bacteria and not harmless commensals in the lumen.

Function of some other lectins has been shown to be highly regulated by pH. Cation-dependent mannose 6-phosphate receptor (CD-MPR), a P-type lectin family member, relies on pH to regulate ligand binding affinity and trafficking by altering the protonation states of key residues. As is predicted for HIP/PAP, CD-MPR has two conformational states, although each is triggered by the presence or

absence of ligand, but not pH [120]. Conformational changes of HIP/PAP seem to be produced by both ligand binding and pH. We have identified E114 as a key residue in HIP/PAP pH sensing. E114 may be acting as a glutamate switch by forming a pH-dependent hydrogen bond with N116, similar to the ligand-dependent switch activity identified in AAA+ ATPases [115].

Future Directions

Biochemical and structural basis for the HIP/PAP-peptidoglycan interaction

The identification of EPN-mediated carbohydrate binding in a CTLD-containing protein raises questions about how this recognition is coordinated at a molecular level. These questions will only be answered by the crystallization of the HIP/PAP-ligand complex. Although we are planning to attempt co-crystallization of HIP/PAP with chitopentaose, obtaining crystallized complexes is likely to be challenging given the 5 mM K_d of the interaction. In the meantime, additional information about these interactions will be garnered by ligand binding and killing assays with the HIP/PAP-N116L mutant, which is part of the conserved Loop 1 EPN motif.

Several experiments will provide greater insight into the role of pH in HIP/PAP activity. Peptidoglycan pull-down assays over a pH range will provide more information about the physiological relevance of the observed pH effects. Determination of the HIP/PAP E114 side chain pK_a will help investigate the idea

that these effects are mediated through a pH-dependent E114-N116 hydrogen bond.

Additional work will be necessary to confirm the presence of the HIP/PAP two-state switch and what effect this has on function. Identification of point mutants that produce each conformation state would allow functional comparison of the two states.

Molecular mechanism of HIP/PAP bactericidal activity

A major remaining question in understanding HIP/PAP function is the molecular mechanism of bacterial killing. Given the binding specificity of HIP/PAP for peptidoglycan, early work was performed to determine whether HIP/PAP might, like lysozyme and PGLYRP2, act through enzymatic degradation of the cell wall. While we have been unable to obtain experimental evidence to support this hypothesis, it remains a formal possibility. However, several lines of evidence have shown that peptidoglycan binding is necessary, but not sufficient for bactericidal activity. First, killing, but not peptidoglycan binding, is inhibited by the presence of an N-terminal prosegment [57]. Also, mutation of the post-cleavage HIP/PAP N-terminal I38 abolishes killing while binding ability is unaffected [57]. Previous work has shown that HIP/PAP microbicidal activity can also be inhibited by the presence of chitooligosaccharides [56]. Finally, work in this dissertation has shown that

mutation of the HIP/PAP EPN motif reduces both binding and killing efficiency. Together, these studies show that peptidoglycan binding is a crucial step of a multi-step antibacterial mechanism.

The fact that HIP/PAP bactericidal activity is blocked prior to secretion by an N-terminal prosegment suggested that HIP/PAP might also have the potential to harm host cells. We now know that cleaved HIP/PAP results in hemoglobin release from erythrocytes and lactate dehydrogenase release from epithelial cells. Since killing activity affects both host and microbial cells, it was apparent that HIP/PAP must be able to recognize a structure common to both cell types. Indeed, fluorescence resonance energy transfer (FRET) analysis has shown that HIP/PAP binds lipid, which comprises the membrane bilayer of both prokaryotic and eukaryotic cells. Future studies will test the hypothesis that HIP/PAP bactericidal activity involves membrane disruption.

These findings suggest a model of HIP/PAP microbicidal activity involving two distinct steps. We hypothesize that binding of peptidoglycan by HIP/PAP is a critical first step in this process, allowing the protein to associate with bacteria, penetrate the cell wall, and access the cell membrane. This activity would then allow HIP/PAP to disrupt the membrane, resulting in microbial cell death (Fig. 6-2).

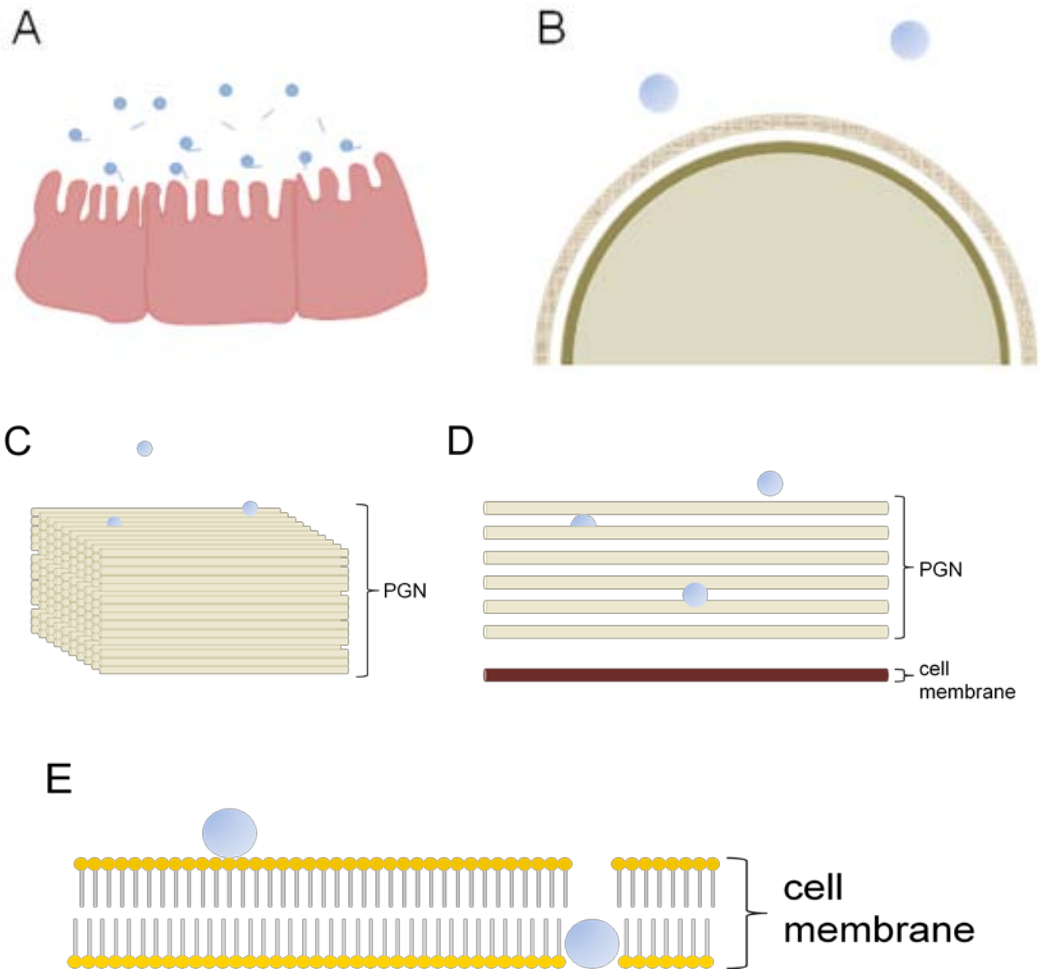


Figure 6-2: Proposed model for HIP/PAP bactericidal activity. **(A)** HIP/PAP is secreted into the gut lumen by Paneth cells (56) where it is N-terminally cleaved by trypsin (57). HIP/PAP **(B)** targets encroaching Gram-positive bacteria by **(C)** binding to the peptidoglycan carbohydrate moieties of the cell wall (56). **(D)** The low binding affinity of HIP/PAP for individual peptidoglycan epitopes allows it to 'bind and jump' from epitope to epitope through the cell wall until it reaches the cell membrane. **(E)** HIP/PAP kills the target microbe by disrupting the cell membrane.

Functional characterization of other Reg family members

Other members of the Reg family are expressed in the gut and may act similarly in epithelial defense [56, 121-125]. Future studies will investigate binding targets and microbicidal capabilities of these proteins.

RegI will be one target of this work. It is expressed throughout the small intestine of both mice and humans and shares 45% sequence identity with RegIIIγ, as well as a predicted N-terminal trypsin site. Additionally, the human RegI crystal structure has been solved and reveals a long loop structure similar to HIP/PAP (Fig. 6-3). In contrast to RegIII proteins, RegI lacks an EPN motif (Fig. 6-3). Thus, many questions regarding RegI ligand binding and biological function remain. Studies are underway to determine the biological function of this family. These results will guide our work to identify RegI ligands.

The absence of a RegI EPN motif leaves two possibilities. RegI might bind peptidoglycan, but through a mechanism distinct from RegIII proteins. Alternatively, RegI might recognize an entirely distinct ligand. Binding screens are underway to identify the nature of the RegI ligand.

```
mRegIIIγ   HDPTLGYEPNRGGWE
mRegIIIβ   HDPTLGAEPNGGGWE
hHIP/PAP   HDPTQGTEPNGEGWE
hRegI      HDPKK----NR-RWH
```

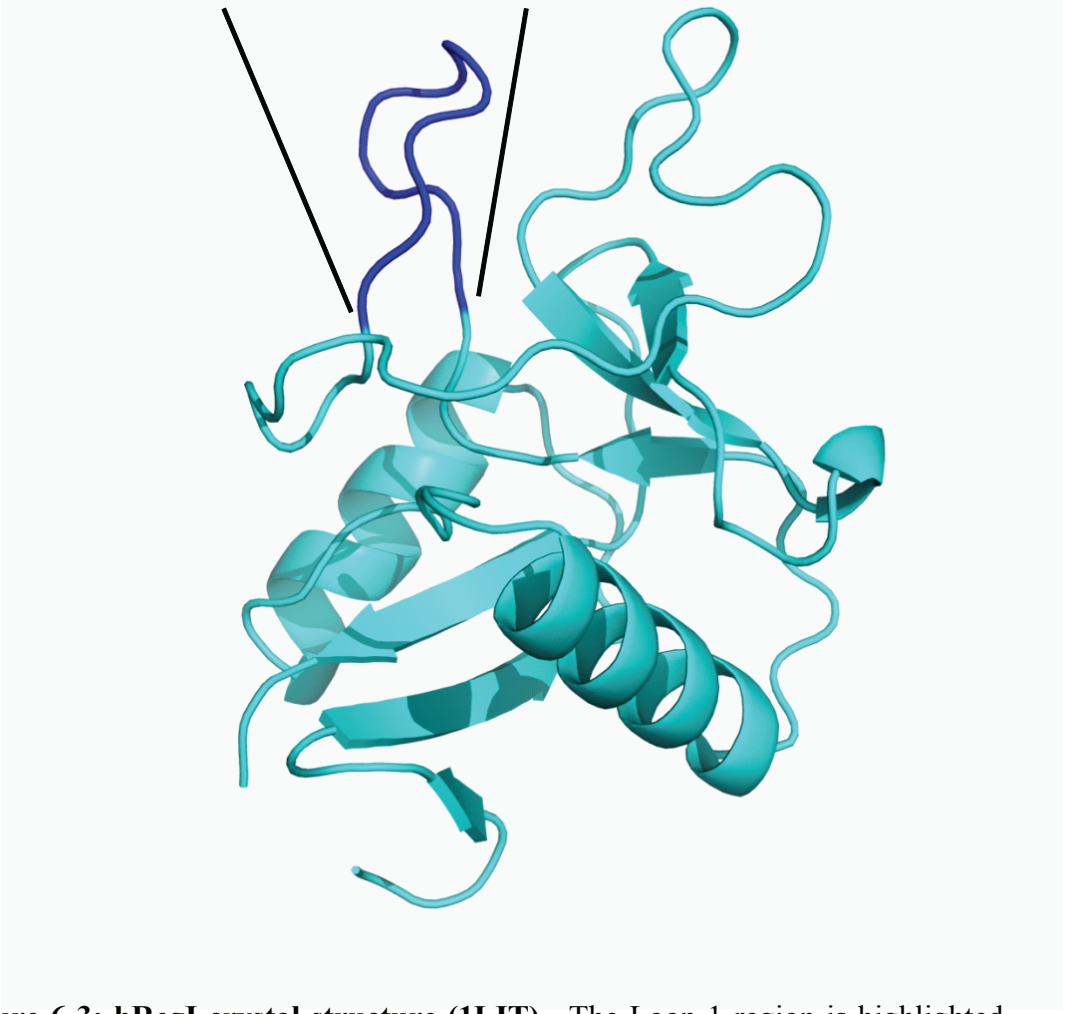


Figure 6-3: hRegI crystal structure (1LIT). The Loop 1 region is highlighted in dark blue.

BIBLIOGRAPHY

1. Turnbaugh, P.J., et al., *The human microbiome project.* Nature, 2007. **449**(7164): p. 804-10.

2. 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.

3. Mazmanian, S.K., et al., *An immunomodulatory molecule of symbiotic bacteria directs maturation of the host immune system.* Cell, 2005. **122**(1): p. 107-18.

4. 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.

5. Backhed, F., et al., *The gut microbiota as an environmental factor that regulates fat storage.* Proc Natl Acad Sci U S A, 2004. **101**(44): p. 15718-23.

6. Mukherjee, S., S. Vaishnava, and L.V. Hooper, *Multi-layered regulation of intestinal antimicrobial defense.* Cell Mol Life Sci, 2008. **65**(19): p. 3019-27.

7. Eckmann, L., *Defence molecules in intestinal innate immunity against bacterial infections.* Curr Opin Gastroenterol, 2005. **21**(2): p. 147-51.

8. Ganz, T., *The Role of Antimicrobial Peptides in Innate Immunity.* Integr Comp Biol, 2003. **43**: p. 300-304.

9. Lievin-Le Moal, V. and A.L. Servin, *The front line of enteric host defense against unwelcome intrusion of harmful microorganisms: mucins, antimicrobial peptides, and microbiota.* Clin Microbiol Rev, 2006. **19**(2): p. 315-37.

10. Lai, Y. and R.L. Gallo, *AMPed up immunity: how antimicrobial peptides have multiple roles in immune defense.* Trends Immunol, 2009. **30**(3): p. 131-41.

11. Nizet, V., et al., *Innate antimicrobial peptide protects the skin from invasive bacterial infection.* Nature, 2001. **414**(6862): p. 454-7.

12. Wilson, C.L., et al., *Regulation of intestinal alpha-defensin activation by the metalloproteinase matrilysin in innate host defense.* Science, 1999. **286**(5437): p. 113-7.

13. 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.

14. 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.

15. Ganz, T., *Antimicrobial polypeptides.* J Leukoc Biol, 2004. **75**(1): p. 34-8.

16. Beers, S.A., et al., *The antibacterial properties of secreted phospholipases A2: a major physiological role for the group IIA enzyme that depends on the very high pI of the enzyme to allow penetration of the bacterial cell wall.* J Biol Chem, 2002. **277**(3): p. 1788-93.

17. Koprivnjak, T., et al., *Role of charge properties of bacterial envelope in bactericidal action of human group IIA phospholipase A2 against Staphylococcus aureus.* J Biol Chem, 2002. **277**(49): p. 47636-44.

18. Harder, J. and J.M. Schroder, *RNase 7, a novel innate immune defense antimicrobial protein of healthy human skin.* J Biol Chem, 2002. **277**(48): p. 46779-84.

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

20. Rosenberg, H.F., *Recombinant human eosinophil cationic protein. Ribonuclease activity is not essential for cytotoxicity.* J Biol Chem, 1995. **270**(14): p. 7876-81.

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

22. Bechinger, B. and K. Lohner, *Detergent-like actions of linear amphipathic cationic antimicrobial peptides.* Biochim Biophys Acta, 2006. **1758**(9): p. 1529-39.

23. Huang, H.W., *Action of antimicrobial peptides: two-state model.* Biochemistry, 2000. **39**(29): p. 8347-52.

24. Kagan, B.L., et al., *Antimicrobial defensin peptides form voltage-dependent ion-permeable channels in planar lipid bilayer membranes.* Proc Natl Acad Sci U S A, 1990. **87**(1): p. 210-4.

25. Selsted, M.E. and A.J. Ouellette, *Mammalian defensins in the antimicrobial immune response.* Nat Immunol, 2005. **6**(6): p. 551-7.

26. Hornef, M.W., et al., *Increased diversity of intestinal antimicrobial peptides by covalent dimer formation.* Nat Immunol, 2004. **5**(8): p. 836-43.

27. Salzman, N.H., et al., *Enteric salmonella infection inhibits Paneth cell antimicrobial peptide expression.* Infect Immun, 2003. **71**(3): p. 1109-15.

28. Chavakis, T., et al., *Regulation of neovascularization by human neutrophil peptides (alpha-defensins): a link between inflammation and angiogenesis.* FASEB J, 2004. **18**(11): p. 1306-8.

29. Lin, P.W., et al., *Paneth cell cryptdins act in vitro as apical paracrine regulators of the innate inflammatory response.* J Biol Chem, 2004. **279**(19): p. 19902-7.

30. O'Neil, D.A., et al., *Expression and regulation of the human beta-defensins hBD-1 and hBD-2 in intestinal epithelium.* J Immunol, 1999. **163**(12): p. 6718-24.

31. Schutte, B.C., et al., *Discovery of five conserved beta -defensin gene clusters using a computational search strategy.* Proc Natl Acad Sci U S A, 2002. **99**(4): p. 2129-33.

32. Moser, C., et al., *beta-Defensin 1 contributes to pulmonary innate immunity in mice.* Infect Immun, 2002. **70**(6): p. 3068-72.

33. Morrison, G., et al., *Characterization of the mouse beta defensin 1, Defb1, mutant mouse model.* Infect Immun, 2002. **70**(6): p. 3053-60.

34. Yang, D., et al., *Beta-defensins: linking innate and adaptive immunity through dendritic and T cell CCR6.* Science, 1999. **286**(5439): p. 525-8.

35. Funderburg, N., et al., *Human -defensin-3 activates professional antigen-presenting cells via Toll-like receptors 1 and 2.* Proc Natl Acad Sci U S A, 2007. **104**(47): p. 18631-5.

36. Niyonsaba, F., et al., *The human beta-defensins (-1, -2, -3, -4) and cathelicidin LL-37 induce IL-18 secretion through p38 and ERK MAPK activation in primary human keratinocytes.* J Immunol, 2005. **175**(3): p. 1776-84.

37. Cole, A.M., et al., *Retrocyclin: a primate peptide that protects cells from infection by T- and M-tropic strains of HIV-1.* Proc Natl Acad Sci U S A, 2002. **99**(4): p. 1813-8.

38. Agerberth, B., et al., *The human antimicrobial and chemotactic peptides LL-37 and alpha-defensins are expressed by specific lymphocyte and monocyte populations.* Blood, 2000. **96**(9): p. 3086-93.

39. Romeo, D., et al., *Structure and bactericidal activity of an antibiotic dodecapeptide purified from bovine neutrophils.* J Biol Chem, 1988. **263**(20): p. 9573-5.

40. Hase, K., et al., *Cell differentiation is a key determinant of cathelicidin LL-37/human cationic antimicrobial protein 18 expression by human colon epithelium.* Infect Immun, 2002. **70**(2): p. 953-63.

41. Bals, R., et al., *The peptide antibiotic LL-37/hCAP-18 is expressed in epithelia of the human lung where it has broad antimicrobial activity at the airway surface.* Proc Natl Acad Sci U S A, 1998. **95**(16): p. 9541-6.

42. Braff, M.H., et al., *Keratinocyte production of cathelicidin provides direct activity against bacterial skin pathogens.* Infect Immun, 2005. **73**(10): p. 6771-81.

43. Chromek, M., et al., *The antimicrobial peptide cathelicidin protects the urinary tract against invasive bacterial infection.* Nat Med, 2006. **12**(6): p. 636-41.

44. Davidson, D.J., et al., *The cationic antimicrobial peptide LL-37 modulates dendritic cell differentiation and dendritic cell-induced T cell polarization.* J Immunol, 2004. **172**(2): p. 1146-56.

45. Koczulla, R., et al., *An angiogenic role for the human peptide antibiotic LL-37/hCAP-18.* J Clin Invest, 2003. **111**(11): p. 1665-72.

46. Zelensky, A.N. and J.E. Gready, *The C-type lectin-like domain superfamily.* FEBS J, 2005. **272**(24): p. 6179-217.

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

48. Lasky, L.A., *Selectin-carbohydrate interactions and the initiation of the inflammatory response.* Annu Rev Biochem, 1995. **64**: p. 113-39.

49. Weis, W.I., M.E. Taylor, and K. Drickamer, *The C-type lectin superfamily in the immune system.* Immunol Rev, 1998. **163**: p. 19-34.

50. van de Wetering, J.K., L.M. van Golde, and J.J. Batenburg, *Collectins: players of the innate immune system.* Eur J Biochem, 2004. **271**(7): p. 1229-49.

51. Wu, H., et al., *Surfactant proteins A and D inhibit the growth of Gram-negative bacteria by increasing membrane permeability.* J Clin Invest, 2003. **111**(10): p. 1589-602.

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

53. Weis, W.I. and K. Drickamer, *Structural basis of lectin-carbohydrate recognition.* Annu Rev Biochem, 1996. **65**: p. 441-73.

54. Lee, R.T., et al., *Multivalent ligand binding by serum mannose-binding protein.* Arch Biochem Biophys, 1992. **299**(1): p. 129-36.

55. 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.

56. Cash, H.L., et al., *Symbiotic bacteria direct expression of an intestinal bactericidal lectin.* Science, 2006. **313**(5790): p. 1126-30.

57. Mukherjee, S., et al., *Regulation of C-type lectin antimicrobial activity by a flexible N-terminal prosegment.* J Biol Chem, 2009. **284**(8): p. 4881-8.

58. Vollmer, W., D. Blanot, and M.A. de Pedro, *Peptidoglycan structure and architecture.* FEMS Microbiol Rev, 2008. **32**(2): p. 149-67.

59. Dramsi, S., et al., *Covalent attachment of proteins to peptidoglycan.* FEMS Microbiol Rev, 2008. **32**(2): p. 307-20.

60. Neuhaus, F.C. and J. Baddiley, *A continuum of anionic charge: structures and functions of D-alanyl-teichoic acids in gram-positive bacteria.* Microbiol Mol Biol Rev, 2003. **67**(4): p. 686-723.

61. Dziarski, R. and D. Gupta, *Peptidoglycan recognition in innate immunity.* J Endotoxin Res, 2005. **11**(5): p. 304-10.

62. Hutton, C.A., T.J. Southwood, and J.J. Turner, *Inhibitors of lysine biosynthesis as antibacterial agents.* Mini Rev Med Chem, 2003. **3**(2): p. 115-27.

63. Dziarski, R. and D. Gupta, *Mammalian PGRPs: novel antibacterial proteins.* Cell Microbiol, 2006. **8**(7): p. 1059-69.

64. 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.

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

66. Bischoff, V., et al., *Function of the drosophila pattern-recognition receptor PGRP-SD in the detection of Gram-positive bacteria.* Nat Immunol, 2004. **5**(11): p. 1175-80.

67. 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.

68. 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.

69. 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.

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

71. Chang, C.I., et al., *A Drosophila pattern recognition receptor contains a peptidoglycan docking groove and unusual L,D-carboxypeptidase activity.* PLoS Biol, 2004. **2**(9): p. E277.

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

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

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

75. 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.

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

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

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

79. Kim, M.S., M. Byun, and B.H. Oh, *Crystal structure of peptidoglycan recognition protein LB from Drosophila melanogaster.* Nat Immunol, 2003. **4**(8): p. 787-93.

80. 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.

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

82. Gelius, E., et al., *A mammalian peptidoglycan recognition protein with N-acetylmuramoyl-L-alanine amidase activity.* Biochem Biophys Res Commun, 2003. **306**(4): p. 988-94.

83. 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.

84. Chang, C.I., et al., *Structure of the ectodomain of Drosophila peptidoglycan-recognition protein LCa suggests a molecular mechanism for pattern recognition.* Proc Natl Acad Sci U S A, 2005. **102**(29): p. 10279-84.

85. Guan, R., et al., *Crystal structure of the C-terminal peptidoglycan-binding domain of human peptidoglycan recognition protein Ialpha.* J Biol Chem, 2004. **279**(30): p. 31873-82.

86. Guan, R., et al., *Crystal structure of human peptidoglycan recognition protein S (PGRP-S) at 1.70 A resolution.* J Mol Biol, 2005. **347**(4): p. 683-91.

87. Reiser, J.B., L. Teyton, and I.A. Wilson, *Crystal structure of the Drosophila peptidoglycan recognition protein (PGRP)-SA at 1.56 A resolution.* J Mol Biol, 2004. **340**(4): p. 909-17.

88. Dziarski, R., *Recognition of bacterial peptidoglycan by the innate immune system.* Cell Mol Life Sci, 2003. **60**(9): p. 1793-804.

89. Inohara, N., Y. Ogura, and G. Nunez, *Nods: a family of cytosolic proteins that regulate the host response to pathogens.* Curr Opin Microbiol, 2002. **5**(1): p. 76-80.

90. Bertin, J., et al., *Human CARD4 protein is a novel CED-4/Apaf-1 cell death family member that activates NF-kappaB.* J Biol Chem, 1999. **274**(19): p. 12955-8.

91. 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.

92. Chamaillard, M., et al., *An essential role for NOD1 in host recognition of bacterial peptidoglycan containing diaminopimelic acid.* Nat Immunol, 2003. **4**(7): p. 702-7.

93. Girardin, S.E., et al., *Nod1 detects a unique muropeptide from gram-negative bacterial peptidoglycan.* Science, 2003. **300**(5625): p. 1584-7.

94. Girardin, S.E., et al., *Peptidoglycan molecular requirements allowing detection by Nod1 and Nod2.* J Biol Chem, 2003. **278**(43): p. 41702-8.

95. 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.

96. Inohara, N., et al., *Host recognition of bacterial muramyl dipeptide mediated through NOD2. Implications for Crohn's disease.* J Biol Chem, 2003. **278**(8): p. 5509-12.

97. Ho, M.R., et al., *Human pancreatitis-associated protein forms fibrillar aggregates with a native-like conformation.* J Biol Chem, 2006. **281**(44): p. 33566-76.

98. 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. **48**(1): p. 151-9.

99. Ng, K.K., S. Park-Snyder, and W.I. Weis, *Ca2+-dependent structural changes in C-type mannose-binding proteins.* Biochemistry, 1998. **37**(51): p. 17965-76.

100. Schubert, M., et al., *A software tool for the prediction of Xaa-Pro peptide bond conformations in proteins based on 13C chemical shift statistics.* J Biomol NMR, 2002. **24**(2): p. 149-54.

101. Brown, J., et al., *Structure of the fungal beta-glucan-binding immune receptor dectin-1: implications for function.* Protein Sci, 2007. **16**(6): p. 1042-52.

102. Chatwell, L., et al., *The carbohydrate recognition domain of Langerin reveals high structural similarity with the one of DC-SIGN but an additional, calcium-independent sugar-binding site.* Mol Immunol, 2008. **45**(7): p. 1981-94.

103. Collins, B.E. and J.C. Paulson, *Cell surface biology mediated by low affinity multivalent protein-glycan interactions.* Curr Opin Chem Biol, 2004. **8**(6): p. 617-25.

104. Dam, T.K., et al., *Binding studies of alpha-GalNAc-specific lectins to the alpha-GalNAc (Tn-antigen) form of porcine submaxillary mucin and its smaller fragments.* J Biol Chem, 2007. **282**(38): p. 28256-63.

105. Cookson, B.T., A.N. Tyler, and W.E. Goldman, *Primary structure of the peptidoglycan-derived tracheal cytotoxin of Bordetella pertussis.* Biochemistry, 1989. **28**(4): p. 1744-9.

106. Horn, S.J. and V.G.H. Eijsink, *A reliable reducing end assay for chito-oligosaccharides.* Carbohydrate Polymers, 2004. **56**: p. 35-39.

107. Varki, A., et al., *Essentials of Glycobiology*. 1st ed. 1999, Plainview, NY: Cold Spring Harbor Laboratory Press.

108. Dam, T.K. and C.F. Brewer, *Effects of clustered epitopes in multivalent ligand-receptor interactions.* Biochemistry, 2008. **47**(33): p. 8470-6.

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

110. 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.

111. Ewe, K., et al., *Inflammation does not decrease intraluminal pH in chronic inflammatory bowel disease.* Dig Dis Sci, 1999. **44**(7): p. 1434-9.

112. Sanderson, I.R., *The physicochemical environment of the neonatal intestine.* Am J Clin Nutr, 1999. **69**(5): p. 1028S-1034S.

113. Bachmann, O., et al., *The Na+/H+ exchanger isoform 2 is the predominant NHE isoform in murine colonic crypts and its lack causes NHE3 upregulation.* Am J Physiol Gastrointest Liver Physiol, 2004. **287**(1): p. G125-33.

114. Siebold, C., et al., *High-resolution structure of the catalytic region of MICAL (molecule interacting with CasL), a multidomain flavoenzyme-signaling molecule.* Proc Natl Acad Sci U S A, 2005. **102**(46): p. 16836-41.

115. Zhang, X. and D.B. Wigley, *The 'glutamate switch' provides a link between ATPase activity and ligand binding in AAA+ proteins.* Nat Struct Mol Biol, 2008. **15**(11): p. 1223-7.

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

117. Delaglio, F., et al., *NMRPipe: a multidimensional spectral processing system based on UNIX pipes.* J Biomol NMR, 1995. **6**(3): p. 277-93.

118. Johnson, B.A., *Using NMRView to visualize and analyze the NMR spectra of macromolecules.* Methods Mol Biol, 2004. **278**: p. 313-52.

119. Fukada, H. and K. Takahashi, *Enthalpy and heat capacity changes for the proton dissociation of various buffer components in 0.1 M potassium chloride.* Proteins, 1998. **33**(2): p. 159-66.

120. Olson, L.J., et al., *Structural insights into the mechanism of pH-dependent ligand binding and release by the cation-dependent mannose 6-phosphate receptor.* J Biol Chem, 2008. **283**(15): p. 10124-34.

121. 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.

122. Bodeker, H., et al., *PAP I interacts with itself, PAP II, PAP III, and lithostathine/regIalpha.* Mol Cell Biol Res Commun, 1999. **2**(3): p. 150-4.

123. 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.

124. 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.

125. 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.