UNDERSTANDING THE ACTIVATION AND MICROBIAL-BINDING PROPERTIES OF X-TYPE LECTIN FAMILY MEMEBERS

by

JONATHAN ADAMS VIOLA

(Under the Direction of J. Michael Pierce)

ABSTRACT

Intelectins are a novel class of X-type lectins with roles in innate immunity. The first X-type lectin was identified from *Xenopus laevis* oocytes and embryos. This X. laevis cortical granule lectin (XCGL-1 or XL-35) crosslinks the jelly-coat protein surrounding the oocyte to prevent polyspermy. Following the discovery of XCGL-1, two human intelectin (hIntL-1 and -2) members were identified. The goal of this dissertation was two-fold. First, we sought to determine the glycan-binding properties of members of the X-type lectin family. Using glycan microarrays, we observed specific interactions between recombinant hIntL-1 and -2 and pathogenic microbial glycans isolated from *Proteus vulgaris*, *Streptococcus pneumoniae*, *Proteus mirabilis*, and *Klebsiella pneumoniae*. Whereas hIntL-1 showed binding to microbial glycans containing β-galactofuranose (β-Gal*f*) and glycerol phosphate (GroP) moieties, hIntL-2 showed minimal binding to glycans with β-Gal*f* and only in specific glycans did it bind GroP. XCGL-1 interacted with the majority of the *S. pneumoniae* glycans on the array, but not with the affinity of hIntL-1 or -2. Molecular dynamics were employed to validate and rationalize the interactions between Gal*f* and hIntL-1 and -2. A simulation of β-Gal*f* and hIntL-1 indicated a stable interaction with β-Gal*f* remaining bound to hIntL-1. In contrast, hIntL-2 was unable to

bind β-Gal*f* for the duration of the simulation. Amino acid sequence alignment of the three lectins showed multiple key ligand-binding residue substitutions, suggesting these residues are critical for glycan specificity and affinity. Second, we investigated the signaling mechanisms for IL-13 induced hIntL-1 activation using a colon cancer cell line. Here, IL-13 activated both the MAPK and JAK/STAT pathways to induce the expression and secretion of hIntL-1. Pharmacologically inhibiting either the MAPK or JAK/STAT pathway prevented IL-13-induced secretion of hIntL-1, suggesting hIntL-1 activation is co-dependent on both pathways. Together, these findings examine the unique activation and differential glycan-binding properties of intelectins, and further support their role in the immune system as microbial recognition molecules.

INDEX WORDS: lectin, microbial glycans, intelectin, IL-13, MAPK, JAK/STAT

UNDERSTANDING THE ACTIVATION AND MICROBIAL-BINDING PROPERTIES OF X-TYPE LECTIN FAMILY MEMEBERS

by

JONATHAN ADAMS VIOLA

B.S., RENSSELAER POLYTECHNIC INSTITUTE, 2010

A Dissertation Submitted to the Graduate Faculty of The University of Georgia in Partial Fulfillment of the Requirements for the Degree

DOCTOR OF PHILOSPOHY

ATHENS, GEORGIA

2017

UNDERSTANDING THE ACTIVATION AND MICROBIAL-BINDING PROPERTIES OF X-TYPE LECTIN FAMILY MEMEBERS

by

JONATHAN ADAMS VIOLA

Major Professor:	Michael Pierce
Committee:	Lance Wells
	Michael Tiemeyer
	Shelley Hooks

Electronic Version Approved:

Suzanne Barbour
Dean of the Graduate School
The University of Georgia
December 2017

DEDICATION

I would like to dedicate this dissertation to my parents, Michael and Mary. They have fostered an interest in science since I was a young child and have supported me throughout my career. Thank you to my wife, Lindsay, for your love and kindness.

ACKNOWLEDGEMENTS

First, I would like to thank my advisor, Professor J. Michael Pierce. He has helped me tremendously in the way of becoming a better scientist. To Dr. Jin-Kyu Lee, who originally introduced me to the intelectin project and laid the groundwork for me to succeed. Thank you to the Pierce Lab members, Dr. Huabei Guo, Pam Kirby, and DJ Bernsteel, you have been here since I rotated and have provided me technical assistance, thoughtful discussion, and friendship. Thank you to the Moremen lab for expressing recombinant proteins. Lastly, thanks to Professor Michael Adang for allowing me to use his laboratory space for I^{125} experiments.

TABLE OF CONTENTS

LIST OF TABLES

Page

LIST OF FIGURES

Page

CHAPTER 1

LITERATURE REVIEW AND INTRODUCTION

ROLE OF GLYCOSYLATION

The surface of animal cells is coated with a layer of diverse glycans termed the glycocalyx [1]. These glycans are constructed from different monosaccharide building blocks. Within the mammalian system, a set of 10 different monosaccharides can be utilized and over 700 monosaccharides are found in the microbial systems [2, 3]. In addition to monosaccharide diversity, a further layer of complexity is added due to different possible linkages between the monosaccharides. The linkages between two monosaccharides can have important functional consequences. For example, both amylose and cellulose are polymers of glucose but differ in glyosidic linkages. Whereas the glucose monomers of starch are linked by a β1-4 linkage, cellulose is linked by an α1,4 linkage. The differences in the stereochemistry between the α and β1-4 bonds are so substantial that the coiled-spring confirmation of amylose changes into a linear chain formation for cellulose. Since humans only possess amylase, the enzyme that catalyzes the hydrolysis of α1,4 glyosidic bonds, cellulose is unable to serve as a source of energy.

Glycan structures in the mammalian system are generally located on specific proteins via either an asparagine residue (N-linked glycosylation) or on a serine or threonine residue (O-linked glycosylation). N-linked glycosylation occurs within a consensus sequence of N-X-S/T where X is an amino acid except for proline [4]. In

addition to the N- and O-linked glycosylation, other glycan classes exist including: C-mannosylation, glycosphingolipids, and glycosaminoglycans. Although the possible combinations of different monosaccharides and linkages are seemingly endless, specific restraints are imposed on the glycan diversity due to the nature of glycan biosynthesis mechanisms, e.g., specific glycan structures are not found in nature due to a lack of specific glycosyltransferase(s) that would be required to build them.

Numerous functional roles of both mammalian and microbial glycosylation have been determined and additional functions likely remain to be discovered. In the simplest sense, the glycans that compose the glycocalyx of mammalian cells serve a structural support function- similar to the functional roles of cellulose, hemicellulose and pectin present in a plant cell wall has on the structural integrity of a plant cell [5]. In addition, specific proteoglycans – a type of glycosylated protein found on the mammalian cell surface – play an important role in maintaining structural integrity. In one example, a missense mutation in the proteoglycan aggrecan resulted in numerous skeletal defects in patients with this mutation [6, 7]. Specifically, this phenotype was caused by a mutation in the C-type lectin domain of aggrecan which prevented the association with components in the extracellular matrix (ECM).

The molecular recognition of glycans is another biological role of glycosylation that mediates both intracellular and intercellular events. Glycans are recognized by molecules termed glycan-binding proteins (GBP), which can be further divided into either lectins or sulfated glycosaminoglycan (GAG)-binding proteins. Lectins, which get their name from the Latin word “to select”, are a diverse class of proteins originally discovered in plants. The carbohydrate-binding properties of lectins are attributed to their

carbohydrate-recognition domain (CRD) which typically binds to terminal glycan structures present in a shallow binding-cleft [8]. On the other hand, positively charged amino acid residues on GAG-binding proteins enable the binding of negatively charged GAG repeats.

In the context of intracellular events, the lectins calnexin and calreticulin bind monoglucosylated high-mannose glycans that are found on newly synthesized glycoproteins. Upon binding, these lectins prevent the glycoproteins from leaving the endoplasmic reticulum (ER) until they are properly folded, thus serving a quality-control function. Similarly located in the ER, the M-type lectins participate in the process of ER-associated glycoprotein degradation (ERAD). Here, these lectins bind to misfolded glycoproteins bearing high-mannose glycans and target them for proteasomal degradation.

Probably the most studied function of GBP proteins is the involvement of lectins in both the innate and adaptive immune system through both intracellular and intercellular mechanisms. The galectin family of lectins can differentiate between self and non-self within innate and adaptive immunity[9]. More specifically, galecetin-3 can recognize and bind *Helicobacter pylori* and act as a bactericidal, possibly through membrane disruption [10]. Other lectins, for example the ficolins or mannose-binding lectins (MBL), can activate the complement pathway leading to complement-mediated cell death [11]. The lectin CD22, a member of the Siglec family of lectins, is present on B- lymphocytes and binds α2-6 linked sialic acid. Upon binding, this lectin receptor signals to induce an inhibitory response to prevent autoimmunity. Due to the widespread

involvement of lectins in many aspects of the immune response, it is not surprising that they are exceptional targets for novel therapeutics [12-14].

SOLUBLE LECTINS WITHIN THE IMMUNE SYSTEM

Chapter two of this thesis defines the glycan-binding properties of the human intelectins. As described below, these are secreted lectins with functional roles in the innate immune system. To highlight the diversity of secreted human lectins involved in the immune system, a review describing the different glycan-binding properties and functional roles of five different lectin families (C-type, ficolins, galectins, pentraxins, X-type) is presented.

As mentioned above, lectins are classified based on their primary amino acid sequence and structural homology [15]. The glycan-binding attributes of a CRD can be independent of the rest of the protein sequence. Such is the case of the prototypical galectin class (Gal-1, -2, -7, -10, -11, -13, -14, and -15), which are expressed as a single CRD without any accessory domains. Most lectins, however, will typically include additional domains such as a signal peptide for secretion, a cystine rich N-terminal region for oligomerization, or other domains that give enzymatic function and/or protein-protein interactions. To date, 16 different families of animal lectins have been discovered. Within a family, it is typical for the lectin to share a common ligand, due to the homology of the CRD domain. This aspect is best highlighted by the galectin and F-type family of lectins which preferentially bind to β-galactosides and fucose residues, respectively. In the case of galectins, a closer examination of the CRD reveals eight residues that are mostly identical throughout the 100+ galectin sequences known. In the following sections, a

summary of seven different lectins families (C-type, collectins, ficolins, galectins, pentraxins, reg lectins, and X-type lectins) will be presented (Table 1.1). These seven lectin families were selected of the sixteen total families due to their relevance to the topic of this dissertation: they are all secreted and soluble lectins with various roles within the immune system. A review of these lectins will put the work of this dissertation into context with other lectins that function in the immune system.

C-type lectins- C-type lectins (CTLs) are noted for their requirement for calcium for glycan binding as well as primary and secondary structural homology within their CRD. Interestingly, the CRD domain of CTLs do not always require the coordination of a calcium ion for glycan binding, nor does the presence of the CRD inherently impart carbohydrate-binding activity. Because of this peculiarity, the CRD of CTLs is commonly referred to as the CTL domain (CTLD). The CTL family consists of sixteen subdivisions based on domain structure and encompasses both soluble and membrane-bound lectins including collectins, selectins, the Reg proteins, and lecticans. This section will focus only on the collectins and Reg proteins as these are involved in innate immune functions.

The collectin family of C-type lectins includes 9 different members and exists as either in a membrane-bound or soluble form. These members include surfactant proteins A and -D (SP-A and SP-D), conglutinin, mannose-binding lectin (MBL), and collectins CL-43, CL-46, CL-P1, CL-L1, and CL-K1. These collectins contain a collagen-like domain and require a protein-bound calcium ion to bind to hydroxyl groups on the carbohydrate ligand [16]. Collectins are synthesized as a single polypeptide that can form oligomeric complexes ranging from 9-27 subunits [17]. In addition, collectins have been

reported to form complexes with other lectins. In one example, the lectin pentraxin 3 (PTX3) was shown to interact with collectins as well as ficolins to influence alphavirus disease progression [18]. The oligomeric status of collectins has an important effect on their function. By forming higher order structures, collectins recruit mannan-binding lectin serine protease 1 and 2 (MASP-1 and -2). MASP-1 then cleaves MASP-2 leading to the lectin pathway of complement [19]. This pathway contrasts with the classical pathway of complement in that the lectin pathway does not utilize antibodies to clear targeted antigens. Collectins function in the innate immune system as pattern-recognitions receptors by binding to different glycans expressed on the surface of pathogens such as viruses, bacteria, and fungi. These glycan ligands, termed pathogen-associated molecular patterns (PAMPs), are varied in their structure and consists of lipopolysaccharides (LPSs), lipoteichoic acids (LTAs), and β-glucans. Once bound to their specific PAMPs, collectins have the ability to activate phagocytosis or induce the complement pathway [20].

Structural analysis of human collectins show that these collectins share a conserved motif of glutamic acid-proline-asparagine. This motif allows for carbonyl side chains of the collectin to engage in calcium ion coordination to the hydroxyl groups on the glycan ligand [21, 22]. Collectins typically bind to terminating glycan residues; however, exceptions have been reported. In the case of collectin-K1 (CL-K1), the CTLD bound to a Man(α1-2)Man disaccharide while a terminal mannose residue engaged in hydrogen bonding with the protein [23].

Perhaps the most unique of the collectins is the surfactant protein-D (SP-D). Instead of forming a bouquet-like oligomer structure, it exhibits a cruciform structure via

disulfide bonds in its N-terminal domain. Each of the four ends of the cruciform is composed of a SP-D trimer with the CTLD domain pointed away. SP-D was found to be secreted by mucus-producing cells in the intestine as well as in the alveolar type II and Clara cells of the airway [24, 25]. SP-D has been shown to bind to both Gram-negative and Gram-positive bacteria such as *K. pneumoniae*, *Bacillus subtilis*, *S. aureus*, *P. aeruginosa*, and *E. coli.* [26-30]. Structural studies suggest the CTLD of SP-D interacts with this diverse set of bacteria via two different mechanisms [31]. In the first mechanisms, the protein-bound calcium ion coordinates to hydroxyl groups present on the carbohydrate ring ligand. In the second mechanism, the protein-bound calcium coordinates to exocyclic 1,2-terminal hydroxyl groups present on the carbohydrate ligand. This second mechanism is strikingly similar to the mechanism hIntL-1 utilizes (described below) to bind to its microbial glycans, suggesting a possible redundancy in the two lectins [32].

The Reg proteins are another lectin class within the CTL family with humans genome containing five different homologs RegIα, RegIβ, RegIIIα, RegIIIγ, and RegIV with murine RegIIIγ being the most well characterized. The RegIII proteins are secreted from intestinal epithelial cells and are also induced by inflammatory signals in the upper respiratory system of mice [33, 34]. The RegIII proteins are suggested to have their bactericidal properties by binding to the peptidoglycan of bacteria [35]. Because the peptidoglycan is not accessible in Gram-positive bacteria, this bactericidal function of RegIII is only relevant to Gram-negative bacteria [36]. The recognition of the peptidoglycan by REGIII is accomplished by a tripeptide repeat of glutamic acid-proline-asparagine, similar to the collectins, yet does not require the coordination of a calcium

ion [35]. Once bound to their bacterial targets, RegIII undergoes proteolysis of an inhibitory N-terminal propeptide and subsequently forms a hexameric membrane-permeabilizing oligomeric pore [37-39]. Pore formation depolarizes the bacteria ultimately leading to cell death. Interestingly, RegIII carbohydrate-binding is dependent on the chain length of the peptidoglycan: the shorter soluble peptidoglycan chains are unable to bind RegIII. It is possible that by only binding to the extended peptidoglycan chains, the RegIII proteins avoid competition from the free and soluble shorter chains that are shed by the bacteria [40].

Ficolins- Three human ficolins have been described, L-ficolin, H-ficolin, and M-ficolin, with a domain organization including a collagen-like, N-terminal domain, and fibrinogen-like domain (FBG) [41]. This class of lectin can bind both Gram-positive and Gram-negative bacteria as well as viruses. Both L-ficolin and H-ficolin are secreted from the liver, whereas M-ficolin is produced by monocytes [42-44]. Both ficolins and collectins are structurally similar with a N-terminal domain and a collagen-like domain. These similarities translate into functional similarities as both lectin types can activate the lectin pathway of complement [45]. However, the collectins bind to carbohydrate ligands using their CRD, whereas ficolins interact with carbohydrates with their FBG. This FBG domain is present on three other classes of proteins (fibrinogens, tenascins, and the microfibril-associated proteins) and is composed of 220-250 residues with 24 mostly hydrophobic residues, 4 cystine residues, and 40 highly conserved residues [46]. Despite these shared features, the ligands of the ficolins remain distinct. L-ficolin was originally reported to bind GlcNAc [46] and later showed to bind GlcNAc-containing structures such as lipoteichoic acid [47]. M-ficolin was reported to have a strong affinity for N-

acetylated and sialylated glycans [48, 49]. The ligands for H-ficolin are poorly understood. Interrogating the Consortium for Functional Glycomics (CFG) mammalian glycan array with H-ficolin did not show any significant binding, suggesting it may not interact with mammalian glycans [49]. Indeed, H-ficolin was later shown to interact with the O-specific polysaccharides of four specific strains of the commensal bacteria *H. alvei.*

The binding selectivities of the types of ficolins can be explained by the four different binding sites the ficolins employ. For example, H-ficolin was shown to bind galactose and fucose via the binding site 1 (S1). The S1 site is located proximal to the calcium-ion binding site, near the outer sider of the trimer and involves four aromatic residues. In contrast, L-ficolin uses three separate binding sites (S2-S4). This large binding surface is thought to give L-ficolin the ability to bind to the diverse set of glycans including phosphorylated, acetylated, sulfated molecules as well as neutral carbohydrates such as galactose and β-1,3-glucan, a component of yeast and fungal cell walls [50, 51].

Galectins- Galectins are a large family of lectins that are organized into three groups based on their primary sequence and domain organization [52]. The prototype/prototypical human galectins (Gal-1, -2, -7, -10, -13,) exist as a single CRD and can associate as homodimers with carbohydrate-binding sites on opposing sides. Chimera human galectins (Gal-3) have an N-terminal domain that allows for oligomerization and a single CRD domain. Lastly, tandem-repeat human galectins (Gal-4, -8, -9, -12) contain two separate CRD connected by a short peptide linker. Adding to the complexity, galectins transcripts have been shown to be differentially spliced leading to multiple isoforms [53, 54].

Galectins typically have some affinity towards β-galactosides; however, their true biological ligands are generally more structurally complex. These glycan ligands have been investigated using glycan arrays and show that the CRD between galectins families differ in carbohydrate-binding specificity. In one example, Gal-3 was shown to bind poly-*N*-acetyllactosamine (poly(LacNAc)) sequences (Galβ1-4GlcNAc), without a preference of a terminal β-Gal residue, whereas Gal-1 required terminal β-Gal for carbohydrate binding [55]. Interestingly, the chimera human Gal-8 contains two separate CRD domains and thus binds to two distinct carbohydrate ligands: the N-terminal CRD binds both sialylated and sulfated glycans while the C-terminal CRD binds polyLacNAc sequences [56, 57].

Galectins are reported to be involved in a variety of functions such as development, innate immune defense, microbial surveillance, cell death and signaling, with even intracellular functions being reported for Gal-1 [58-64]. In the case of innate immune function, galectins have been shown to interact with many different pathogens including bacteria, viruses, fungi, protozoans, as well as helminths [65]. As an antimicrobial agent, human Gal-4 and -8 directly binds and kills *E. coli* strains that express a human blood group B-like antigen [66]. Only the C-terminal CRD domain of Gal-4 and -8 mediates the bactericidal function of the lectin, and this function occurs independent of complement action. Other galectins have also showed anti-microbial properties with Gal-3 binding to the LPS of *H. pylori* and rapidly killing the microbe [10]. The molecular mechanisms of how galectins function as effector factors in microbial killing is not completely understood; however, studies involving Gal-8 suggest disruption of the integrity of the bacterial membrane is likely key [57].

Pentraxins- The pentraxins are found in both invertebrates and vertebrates and include both the short and long pentraxins. The C-reactive protein (CRP) and serum amyloid P component (SAP) are examples of the short pentraxins group with PTX3 belonging to the long pentraxin subset. Both the CRP and SAP, which are exclusively produced by hepatocytes, bind to microbial glycans via calcium coordination with CRP interacting with phosphocholine residues within the C-polysaccharide of *S. pneumoniae,* and SAP interacting with phosphoethanolamine and methyl 4,6-O-(1-carboxyethylidene)-β-D-galactopyranoside [67-69]. The long pentraxin PTX, expressed by vascular endothelial cells, recognizes protein A from the cell wall of *K. pneumoniae* as well as various other microbial and viral pathogens, although the binding mechanism is unknown [70].

The CRD of the pentraxins is unique in that it allows for the interaction with glycans, proteins, and lipids. This is accomplished by the eight amino acid motif (H-x-C-x-S/T-W-x-S), where x represents any amino acid). By engaging in protein-protein interactions, pentraxins modulate immune system function. For example, pentraxins have been shown to interact with C1q, the first component of the classical pathway of complement activation, to stimulate pathogen killing [71]. In addition, pentraxins bind ficolins and MBLs, other players in the lectin pathway of complement, to possibly modulate their function [18].

X-type lectins- Originally identified from *X. laevis* oocytes and embryos, the X-type lectin family has since expanded to include homologs present from the evolutionary ancient tunicates to humans [72, 73]. These lectins, termed intelectins, were first described to be calcium-dependent and galactose-binding.[74, 75]. The domain

organization of the intelectins includes a cystine-rich N-terminal region for oligomerization followed by a collagen-like and FBG domain, similar to the ficolins. However, whereas the ficolins can bind N-acetylated glycans via the FBG domains, intelectins are unable to do so [76, 77]. This is likely due to the lack of homology between FBG domains of the intelectins and ficolins.

The crystal structures of the human intelectin-1 (hIntL-1) with allyl-β-D-galactofuranose (β-Gal*f*) and *Xenopus* embryonic epidermal lectin (XEEL) with D-glycerol-1-phosphate (GroP) were recently solved [32, 78]. To date, these are the only structures of X-type lectins and show that the X-type lectins are structurally unique among the other classes of lectins. From the co-complex structure of hIntL-1, it was revealed that hIntL-1 binds to β-Gal*f* via coordination of a protein-bound calcium ion to the acyclic 1,2-terminal diol group present on β-Gal*f*. The monosaccharide β-Gal*f* is not found on mammalian cells, but is commonly found on surface of both pathogenic and commensal unicellular organisms such as *T. cruzi*, *S. pneumoniae*, and *B. fragilis* [79-82]. In addition to β-Gal*f*, hIntL-1 was also reported to bind to the structures GroP, heptoses, d-glycero-d-talo-oct-2-ulosonic acid (KO) and 3-deoxy-d-manno-oct-2-ulosonic acid (KDO) groups. Interestingly, each of these structures contain the terminal exocyclic 1,2-diol moiety; however, this moiety is not sufficient for binding. N-acetylneuraminic acid (Neu5Ac), a vertebrate sugar, also contains an acyclic 1,2-terminal diol, but due to possible steric hinderances and anion-repulsion between Neu5Ac and hIntL-1, no binding was observed [32].

The function of hIntL-1 as well as human intelectin-2 (hIntL-2) is not well understood. *In vitro*, hIntL-1 does not have bactericidal activity toward the serotype of *S.*

pneumoniae that it binds; however, it is possible that it may modulate other effector molecules via protein-protein interactions [32]. *In vivo* studies using mouse intelectin-1 and -2 (mIntL-1 and -2) have provided important clues as to the innate immune functions of the mammalian intelectins. mIntL-1 and -2 share high homology with the human intelectins (81% identical between mItnL-1 and hIntL-1; 77% identical between mIntL-1 and hIntL-1), and mIntL-1 is reported to also bind to galactofuranosyl moieties [32, 83]. Overexpression of either mIntL-1 or -2 did not affect the clearance of *N. brasiliensis* or *M. tuberculosis* compared to wild-type mice. Additionally, intelectin overexpression did not modulate the immune response during infection [84]. It was suggested, however, that an absence of intelectin expression may negatively affect pathogen clearance, and that expression above a certain threshold may not constitute any benefits in pathogen clearance [85]. Interestingly, mIntL-2 is naturally absent in the genome of C57BL/10 mice, and these mice have higher susceptibility to the parasite *T. spiralis* compared to mIntL-2-expressing BALB/c mice [86].

hIntL-1 and -2 differ in tissue and cellular expression. hIntL-1 is constitutively expressed in goblet cells of the colon, small intestine, as well as mesothelial cells surrounding the heart and adipose [75, 87]. hIntL-2, however, is only secreted by Paneth cells in the small intestine [72]. The goblet cells and Paneth cells are both specialized cell types with innate immune functions. Goblet cells are responsible for secreting the mucin proteins that densely coat the airway and intestinal epithelium. Serving as a protective barrier, these mucin proteins prevent direct microbial contact with the host [88]. Paneth cells, in contrast, are present only in the small intestine where they secrete antimicrobial factors such as defensins and lysozyme [89, 90].

THE CYTOKINE IL-13

In Chapter three of this dissertation, IL-13 is used to induce the expression and secretion of hIntL-1. A background of IL-13 including its effector functions and signaling mechanisms is essential to understand the significance of the work presented in chapter three. The following section examines these features.

IL-13 was first characterized as a small protein induced by activated mouse T_H2 cells and later as an effector molecule with anti-inflammatory and immune system modulating functions [91, 92]. The protein consists of 132 amino acids and belongs to the class of type II cytokines. The gene encoding IL-13 contains four exons and three introns and is located on 5q31 which also houses other cytokines including IL-3, IL-4, IL-5, IL-9, and GM-CSF [93, 94]. Key structural features of the IL-13 were identified by multidimensional NMR which include a lefthanded four-helical bundle, two disulfide bonds, and a 25-amino-acid hydrophobic core [95].

The pleiotropic effects of IL-13 is well documented with receptors being expressed on various cell types: human B cells, basophils, eosinophils, mast cells, endothelial cells, fibroblasts, monocytes, macrophages, respiratory epithelial cells, and smooth muscle cells [96]. The functional properties of IL-13 have strong overlap with IL-4, despite the two only sharing 25% homology [97]. It is possible that the similarities between the two cytokines are due to the highly conserved hydrophobic structural core: of these 25 amino acids, all are identical with the one exception being a conservative substitution [98]. Functionally, IL-13 has been shown to regulate a diverse set of

functions such as IgE class-switching, goblet cell hyperplasia, tumor growth, eosinophil and mast cell activation, and allergic inflammation [99].

Receptors for IL-13- Two receptors exist for IL-13: IL-13Rα1 and IL-13Rα2. These receptors share 37% homology and display different binding constants with IL-13. Whereas IL-13Rα1 binds with a low affinity (Kd =2-10 nmol/L), IL-13Rα2 shows a high-affinity (Kd = 250 pmol/L) interaction with IL-13 [100, 101]. The affinity for IL-13R α1 is significantly increased (Kd = 400 pmol/L) after forming a heterodimer complex with IL-4Rα. This complex is also the functional receptor for IL-4 signaling and both cytokines signal through the JAK/STAT pathway through this complex. The IL-13Rα2, unlike IL-13Rα1, contain neither a JAK/STAT binding sequence nor a long cytoplasmic tail and was thought to serve as a decoy receptor for IL-13 [101]. IL-13Rα2 typically resides intracellularly, but is mobilized to the cell surface upon treatment with IFN- γ [102]. Following upregulation of IL-13Rα2, IL-13 is sequestered resulting in a decrease of JAK/STAT signaling. Recently, IL-13Rα2 has been shown to modulate mitogen-activated protein kinase (MAPK) signaling through IL-13 binding [103-105].

JAK/STAT signaling - Upon binding to IL-13, IL-13Rα1 recruits IL-4Rα where a signal transduction pathway involving JAK/STAT is initiated. The JAKs are a group of tyrosine kinases that contain both a catalytic domain and a pseudokinase domain with four different JAKs characterized: JAK1, JAK2, JAK3, and Tyk2 [106]. Both IL-13Rα1 and IL-4Rα contain a proline-rich element termed Box 1 which has affinity for Tyk2 and JAK1, respectively [107]. By binding the heterodimer receptor, IL-13 activates both Tyk2 and JAK1 which subsequently leads to the phosphorylation of tyrosine residues present on IL-4Rα. STAT6 then utilizes its SH2 domain to bind to these phosphorylated

sites. JAK1 phosphorylates STAT6 on Y641 which allows for the dimerization of STAT6 [108]. Activated STAT6 dimers translocate to the nucleus and bind to and induce transcription of specific genes.

The mechanisms for propagating the IL-13 induced JAK/STAT signal cascade are well defined, but the inhibitory mechanisms, which include SH-2-containing phosphatases, suppressors of cytokine signaling (SOCS) and protein inhibitors of activated STAT of transcription (PIAS) are not well understood. The SH2 domain-containing tyrosine phosphatase 1 (SHP-1) downregulates not only IL-13 signaling but also IL-2, IL-3, CSF, and erythropoietin [96]. It is not clear if SHP-1 acts on STATs or JAKs as SHP-1 has been reported to bind to both JAK2 as well as IL-4 Rα [109, 110]. SOCS proteins negatively influence JAK/STAT signaling by binding to the activation loops of JAKs via the SH2 domain [111]. Once bound, SOCS modulate JAK activity through their kinase inhibitory regions which may act as a pseudosubstrate. In addition, SOCS and their ligands may be targeted for the proteasome for degradation [112]. This is further supported from the observation that a functional proteasomal degradation pathway is necessary for SOCS-mediated JAK/STAT inhibition [113]. Lastly, the PIAS bind to phosphorylated STAT dimers and prevent DNA binding [114]. It is hypothesized that STAT methylation plays a key role in PIAS binding. When STAT undergoes methylation at Arg31, the association between PAIS-STAT decreases significantly allowing for STAT to bind target DNA sequences [115].

MAPK signaling- Both the IL-4Rα and IL-13Rα2 receptor have been implicated in MAPK signaling upon activation. Upon IL-13 binding, IL-4Rα recruits insulin receptor substrate 1 and 2 (IRS-1 and -2) to the phosphorylated residue Tyr497 and are

subsequently activated [116]. Both IRS-1 and IRS-2 then signal through two separate pathways: phosphoinositol 3 (PI3) kinase and Ras/MAPK [117]. MAPK activation is also suggested to occur independently of IL-4R through the IL-13Rα2 receptor. As mentioned above, the receptor has a short cytoplasmic tail without any Box-1 or Box-2 signaling motifs. Despite this, reports suggest that this tail mediates signaling through an unknown mechanism [118]. In one study, deletion of this tail region prevented the IL-13 induction of TGF-β1 via an AP-1 variant [119]. Additionally, overexpression of IL-13Rα2 is common in various cancers and both human ovarian and pancreatic cancer cells showed increased MAPK activation following IL-13 stimulation [103-105].

Crosstalk between JAK/STAT and MAPK- Interactions between the JAK/STAT and MAPK signaling cascades occur on multiple levels with each influencing the signaling of the other [120]. Upon JAK activation, receptor tyrosine residues are phosphorylated which can lead to the recruitment of effector molecules with SH-2 domains. These proteins include SHP-1 and -2, as mentioned above, but also Src homology 2 domain-containing (Shc). From Shc, the GRB2 adaptor protein is recruited where it may stimulate the Ras cascade. Through indirect means, JAK/STAT signaling also regulates Ras activity by modulating RasGAP. Here, SOCS3 is induced by IL-6 where it binds RasGAP, an inhibitor of the Ras pathway, and negatively regulates its activity [121].

The MAPK pathway is also reported to influence JAK/STAT activity. In the first mechanism, activation of either the epidermal growth factor receptor (EGFR) or platelet-derived growth factor receptors (PDGFR) causes the phosphorylation of STATs independently of JAK activity [122]. STAT5a binds ERK1/2 *in vitro* through interactions

with the STAT5a COOH-terminal regions. This interaction may result in the phosphorylation of STAT5a at Ser780 by ERK1/2 [123]. Secondly, MAPK is activated downstream of the Ras pathway. Once activated, MAPK is known to phosphorylate a single serine residue on the C-terminal of STATs [124]. Although not required for STAT activity, this phosphorylation promotes STAT function.

IL-13 Function in Immunity- The expulsion of parasitic nematodes in mice is well documented to require IL-13 [125, 126]. In IL-4$^{-/-}$ mice, the parasitic nematode, *N. brasiliensis*, is expelled normally. Likewise, blocking IL-4 activity using an anti-IL-4 antibody resulted in mice that have the same susceptibility to infection as wildtype mice. In contrast, IL-13-deficient mice fail to expel these parasites. Mice that also lack the IL-13 signaling molecules IL-4Rα and STAT6 fail to expel *N. brasiliensis*. Interestingly, other parasites such as *T. muris* require both IL-13 and IL-4 for expulsion in mice. Mouse strains that express type 2 responses have a higher resistance to *T. muris* compared to mice that express type 1 responses [127]. Conversely, the expulsion of *T. spiralis* in IL-4 or IL-13 suppressed mice is not negatively affected. Only in the absence of both cytokines is worm expulsion suppressed [128]. Despite STAT6 being required for both *N. brasiliensis* and *T. spiralis* expulsion, the mechanisms of expulsions are different between the two parasites.

IL-13 also influences the hosts ability to combat intracellular parasitism. The course of disease progression with humans infected with *Leishmania major* is strongly dependent on the type of immune response generated against *L. major*. In lesions isolated from patients with localized cutaneous leishmaniasis, IL-13 was observed to be preferentially expressed over other type 2 cytokines, including IL-4 [129].

The following studies have reported differing roles for IL-13 on the progression of *L. major* infection. In IL-4Rα$^{-/-}$ mice, which should be defective in both IL-4 and IL-13 signaling, a decrease in the type 2 response was noted. Interestingly, these mice were resistant to the *L. major* substrain IR174, but not substrain LV39. IL-4$^{-/-}$ mice also showed a decrease in the type 2 response, as expected, but were equally susceptible to both substrains of *L. major*. The mechanism for the involvement of IL-13 as a disease-promoting factor was not investigated, but suggested inhibition of macrophage or monocytes may be at play [130]. In contrast to the previous report, IL-13 has also been suggested to function in a protective role against *L. major* disease progression. Both IL-4Rα$^{-/-}$ and IL-4$^{-/-}$ were studied in the context of chronic leishmaniasis. Whereas IL-4$^{-/-}$ mice were able to moderately control the infection, IL-4Rα$^{-/-}$ mice showed a more progressive nature of the disease manifesting in necrotic skin lesions and ultimately resulting in death [131].

Lastly, IL-13 may have a protective role in the intracellular murine pathogen *Listeria monocytogenes*. Recombinant IL-13 (rIL-13) was shown to increase IL-12, a type 1 cytokine critical in the protective immunity against intracellular pathogens [132]. Bone-marrow derived macrophages were responsible for this increase in IL-12 production. Mice treated with rIL-13 had reduced numbers of viable *L. monocytogenes* in both the spleen and liver. It was speculated that exogenous rIL-13 indirectly affected host macrophage via stimulation of IL-12 production and inhibition of IL-4 release.

The secretion of various mucins via goblet cells has been shown to be dependent on IL-13 stimulation. Goblet cells are responsible for producing different types of

mucins, depending on the tissue localization, which provide a protective barrier for the host against pathogen invasion. Co-culturing human bronchial epithelial cells with IL-13 resulted in an increase in goblet cell density (GCD) [133]. Curiously, IL-13 concentration was critical for goblet cell metaplasia: 1 ng/mL of IL-13 increased GCD whereas a concentration of 10 ng/mL resulted in a decrease of GCD. Pharmacological knockdown of the signaling molecules involved in the MAPK and PI3K pathways prevented the IL-13-induced goblet cell metaplasia. In both the airway and digestive systems, IL-13 acts as a potent mediator of Muc5ac and Muc2 expression, respectively [134, 135]. These mucins, as well as Muc5b and Muc6 are classified as 11p15 mucins, but only Muc2 and Muc5ac are reported to be under control of the EGFR signaling pathway [136].

In vivo studies utilizing knockout mice highlight the importance of the Muc2 in barrier function and gut homeostasis. $Muc2^{-/-}$ mice, when challenged with dextran sulfate, displayed severe symptoms of colitis compared to $Muc2^{+/-}$ and $Muc2^{+/+}$ mice. These Muc2-deficient developed an abnormal colonic morphology with a thickening of the gut mucosa [137]. Mice lacking in *Cosmc*, a chaperone protein required for O-glycosylation and thus functional Muc2, in intestinal epithelial cells resulted in a deficient mucus layer [88]. In addition to increasing susceptibility to induced colitis, these Muc2 deficient mice had direct bacterial-epithelial contact. Within wildtype mice, the inner Muc2 layer is proposed to be a sterile environment [138]. The disruption of the Muc2 layer in these mice may explain the altered gut microbiota observed.

THESIS OVERVIEW

Human intelectins were first identified from human expressed sequence tag databases searches using XCGL-1 cDNA and amino acid sequences [73]. Using different free saccharides, hIntL-1 was originally reported to be a calcium-dependent lectin with affinity toward D-pentoses and D-galactofuranosyl residues [76]. Surface plasmon resonance (SPR) and enzyme-linked immunosorbent like assay (ELISA) experiments later confirmed the interactions with galactofuranose, but yielded conflicted results on interactions with ribose and galactopyranose [32].

The first part of this thesis expands the knowledge of the glycan selectivity of three different X-type lectins: hIntL-1, -2 and XCGL-1. We interrogated both mammalian and microbial glycan microarrays with recombinantly expressed hIntL-1, -2 and XCGL-1. Briefly, array analysis showed that each lectin interacts with specific microbial glycans. hIntL-1 and -2 bound a set of microbial glycans from *Proteus vulgaris*, *Streptococcus pneumoniae*, *Proteus mirabilis*, and *Klebsiella pneumoniae*, while XCGL-1 bound the majority of *Streptococcus pneumoniae* serotypes glycans on the array. β-Gal*f* and GroP were revealed to be common structures within microbial glycans that bound hIntL-1 whereas hIntL-2 only showed binding to specific glycans with GroP. Molecular dynamics techniques were subsequently employed to validate and rationalize the array results of hIntL-1 and -2 with β-Gal*f*. hIntL-1 showed a stable interaction with β-Gal*f* whereas hIntL-2 did not. Amino acid substitution of key residues involved in ligand binding and calcium coordination were later shown to be substituted in the three lectins, possibly contributing to the observed differences in glycan binding.

The third chapter of this thesis investigates the mechanisms for IL-13 induced hIntL-1 activation. Here, IL-13 activates both the MAPK and JAK/STAT pathways in LS174T cells, a colon cancer adenocarcinoma cell line, resulting in secretion of hIntL-1. Knockdown of either MAPK or JAK/STAT prevented the IL-13-induction of hIntL-1, suggesting the lectin is co-dependent on both pathways. Previous reports show the promoter region of hIntL-1 contains a binding site for Sp1 [139]. This transcription factor is implicated in both the MAPK and JAK/STAT pathways, and could explain the co-dependence of hIntL-1 on two separate signaling pathways [140, 141]. By requiring two activated signaling pathways for induction, the regulation of hIntL-1 expression can be finely tuned.

REFERENCES

1. Rambourg, A. and C. Leblond, *Electron microscope observations on the carbohydrate-rich cell coat present at the surface of cells in the rat.* The Journal of cell biology, 1967. **32**(1): p. 27-53.
2. Herget, S., et al., *Statistical analysis of the Bacterial Carbohydrate Structure Data Base (BCSDB): characteristics and diversity of bacterial carbohydrates in comparison with mammalian glycans.* BMC Struct Biol, 2008. **8**: p. 35.
3. Adibekian, A., et al., *Comparative bioinformatics analysis of the mammalian and bacterial glycomes.* Chemical Science, 2011. **2**(2): p. 337-344.
4. Pless, D.D. and W.J. Lennarz, *Enzymatic conversion of proteins to glycoproteins.* Proc Natl Acad Sci U S A, 1977. **74**(1): p. 134-8.
5. Weinbaum, S., J.M. Tarbell, and E.R. Damiano, *The structure and function of the endothelial glycocalyx layer.* Annu Rev Biomed Eng, 2007. **9**: p. 121-67.
6. Schwartz, N.B. and M. Domowicz, *Chondrodysplasias due to proteoglycan defects.* Glycobiology, 2002. **12**(4): p. 57R-68R.
7. Stattin, E.L., et al., *A missense mutation in the aggrecan C-type lectin domain disrupts extracellular matrix interactions and causes dominant familial osteochondritis dissecans.* Am J Hum Genet, 2010. **86**(2): p. 126-37.
8. Weis, W.I. and K. Drickamer, *Structural basis of lectin-carbohydrate recognition.* Annu Rev Biochem, 1996. **65**: p. 441-73.
9. Vasta, G.R., et al., *Galectins as self/non-self recognition receptors in innate and adaptive immunity: an unresolved paradox.* Front Immunol, 2012. **3**: p. 199.
10. Park, A.M., et al., *Galectin-3 Plays an Important Role in Innate Immunity to Gastric Infection by Helicobacter pylori.* Infect Immun, 2016. **84**(4): p. 1184-93.
11. Fujita, T., M. Matsushita, and Y. Endo, *The lectin-complement pathway--its role in innate immunity and evolution.* Immunol Rev, 2004. **198**: p. 185-202.
12. Yamazaki, N., et al., *Endogenous lectins as targets for drug delivery.* Adv Drug Deliv Rev, 2000. **43**(2-3): p. 225-44.
13. O'Reilly, M.K. and J.C. Paulson, *Siglecs as targets for therapy in immune-cell-mediated disease.* Trends Pharmacol Sci, 2009. **30**(5): p. 240-8.
14. Yang, R.Y., G.A. Rabinovich, and F.T. Liu, *Galectins: structure, function and therapeutic potential.* Expert Rev Mol Med, 2008. **10**: p. e17.
15. Drickamer, K. and M.E. Taylor, *Biology of animal lectins.* Annu Rev Cell Biol, 1993. **9**: p. 237-64.
16. Drickamer, K., *C-type lectin-like domains.* Curr Opin Struct Biol, 1999. **9**(5): p. 585-90.
17. 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.
18. Foo, S.S., et al., *Pentraxins and Collectins: Friend or Foe during Pathogen Invasion?* Trends Microbiol, 2015. **23**(12): p. 799-811.
19. Bajic, G., et al., *Complement activation, regulation, and molecular basis for complement-related diseases.* EMBO J, 2015. **34**(22): p. 2735-57.
20. Hansen, S.W., et al., *The collectins CL-L1, CL-K1 and CL-P1, and their roles in complement and innate immunity.* Immunobiology, 2016. **221**(10): p. 1058-67.

21. Drickamer, K., *Engineering galactose-binding activity into a C-type mannose-binding protein.* Nature, 1992. **360**(6400): p. 183-6.
22. 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.
23. Venkatraman Girija, U., et al., *Molecular basis of sugar recognition by collectin-K1 and the effects of mutations associated with 3MC syndrome.* BMC Biol, 2015. **13**: p. 27.
24. Voorhout, W.F., et al., *Immunocytochemical localization of surfactant protein D (SP-D) in type II cells, Clara cells, and alveolar macrophages of rat lung.* J Histochem Cytochem, 1992. **40**(10): p. 1589-97.
25. Fisher, J.H. and R. Mason, *Expression of pulmonary surfactant protein D in rat gastric mucosa.* Am J Respir Cell Mol Biol, 1995. **12**(1): p. 13-8.
26. Pikaar, J.C., et al., *Opsonic activities of surfactant proteins A and D in phagocytosis of gram-negative bacteria by alveolar macrophages.* J Infect Dis, 1995. **172**(2): p. 481-9.
27. Van Iwaarden, J.F., et al., *Binding of surfactant protein A to the lipid A moiety of bacterial lipopolysaccharides.* Biochem J, 1994. **303 (Pt 2)**: p. 407-11.
28. Sahly, H., et al., *Surfactant protein D binds selectively to Klebsiella pneumoniae lipopolysaccharides containing mannose-rich O-antigens.* J Immunol, 2002. **169**(6): p. 3267-74.
29. van de Wetering, J.K., et al., *Characteristics of surfactant protein A and D binding to lipoteichoic acid and peptidoglycan, 2 major cell wall components of gram-positive bacteria.* J Infect Dis, 2001. **184**(9): p. 1143-51.
30. Kishore, U., et al., *The alpha-helical neck region of human lung surfactant protein D is essential for the binding of the carbohydrate recognition domains to lipopolysaccharides and phospholipids.* Biochem J, 1996. **318 (Pt 2)**: p. 505-11.
31. Wang, H., et al., *Recognition of heptoses and the inner core of bacterial lipopolysaccharides by surfactant protein d.* Biochemistry, 2008. **47**(2): p. 710-20.
32. Wesener, D.A., et al., *Recognition of microbial glycans by human intelectin-1.* Nat Struct Mol Biol, 2015. **22**(8): p. 603-10.
33. 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.
34. Choi, S.M., et al., *Innate Stat3-mediated induction of the antimicrobial protein Reg3gamma is required for host defense against MRSA pneumonia.* J Exp Med, 2013. **210**(3): p. 551-61.
35. Lehotzky, R.E., et al., *Molecular basis for peptidoglycan recognition by a bactericidal lectin.* Proc Natl Acad Sci U S A, 2010. **107**(17): p. 7722-7.
36. Cash, H.L., et al., *Symbiotic bacteria direct expression of an intestinal bactericidal lectin.* Science, 2006. **313**(5790): p. 1126-30.
37. Mukherjee, S., et al., *Antibacterial membrane attack by a pore-forming intestinal C-type lectin.* Nature, 2014. **505**(7481): p. 103-7.

38. Loonen, L.M., et al., *REG3gamma-deficient mice have altered mucus distribution and increased mucosal inflammatory responses to the microbiota and enteric pathogens in the ileum.* Mucosal Immunol, 2014. **7**(4): p. 939-47.
39. 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.
40. Mukherjee, S. and L.V. Hooper, *Antimicrobial defense of the intestine.* Immunity, 2015. **42**(1): p. 28-39.
41. Matsushita, M. and T. Fujita, *Ficolins and the lectin complement pathway.* Immunol Rev, 2001. **180**: p. 78-85.
42. Matsushita, M., et al., *A novel human serum lectin with collagen- and fibrinogen-like domains that functions as an opsonin.* J Biol Chem, 1996. **271**(5): p. 2448-54.
43. Akaiwa, M., et al., *Hakata antigen, a new member of the ficolin/opsonin p35 family, is a novel human lectin secreted into bronchus/alveolus and bile.* J Histochem Cytochem, 1999. **47**(6): p. 777-86.
44. Hashimoto, S., et al., *Serial analysis of gene expression in human monocyte-derived dendritic cells.* Blood, 1999. **94**(3): p. 845-52.
45. Holmskov, U., S. Thiel, and J.C. Jensenius, *Collections and ficolins: humoral lectins of the innate immune defense.* Annu Rev Immunol, 2003. **21**: p. 547-78.
46. Lu, J. and Y. Le, *Ficolins and the fibrinogen-like domain.* Immunobiology, 1998. **199**(2): p. 190-9.
47. Lynch, N.J., et al., *L-ficolin specifically binds to lipoteichoic acid, a cell wall constituent of Gram-positive bacteria, and activates the lectin pathway of complement.* J Immunol, 2004. **172**(2): p. 1198-202.
48. Frederiksen, P.D., et al., *M-ficolin, an innate immune defence molecule, binds patterns of acetyl groups and activates complement.* Scand J Immunol, 2005. **62**(5): p. 462-73.
49. Gout, E., et al., *Carbohydrate recognition properties of human ficolins: glycan array screening reveals the sialic acid binding specificity of M-ficolin.* J Biol Chem, 2010. **285**(9): p. 6612-22.
50. Laffly, E., et al., *Human ficolin-2 recognition versatility extended: an update on the binding of ficolin-2 to sulfated/phosphated carbohydrates.* FEBS Lett, 2014. **588**(24): p. 4694-700.
51. Ma, Y.G., et al., *Human mannose-binding lectin and L-ficolin function as specific pattern recognition proteins in the lectin activation pathway of complement.* J Biol Chem, 2004. **279**(24): p. 25307-12.
52. Thiemann, S. and L.G. Baum, *Galectins and Immune Responses-Just How Do They Do Those Things They Do?* Annu Rev Immunol, 2016. **34**: p. 243-64.
53. Wada, J. and Y.S. Kanwar, *Identification and characterization of galectin-9, a novel beta-galactoside-binding mammalian lectin.* J Biol Chem, 1997. **272**(9): p. 6078-86.
54. Bidon-Wagner, N. and J.P. Le Pennec, *Human galectin-8 isoforms and cancer.* Glycoconj J, 2002. **19**(7-9): p. 557-63.
55. Stowell, S.R., et al., *Galectin-1, -2, and -3 exhibit differential recognition of sialylated glycans and blood group antigens.* J Biol Chem, 2008. **283**(15): p. 10109-23.

56. Ideo, H., et al., *The N-terminal carbohydrate recognition domain of galectin-8 recognizes specific glycosphingolipids with high affinity.* Glycobiology, 2003. **13**(10): p. 713-23.
57. Stowell, S.R., et al., *Dimeric Galectin-8 induces phosphatidylserine exposure in leukocytes through polylactosamine recognition by the C-terminal domain.* J Biol Chem, 2008. **283**(29): p. 20547-59.
58. Liu, F.T., et al., *Galectins in regulation of apoptosis.* Adv Exp Med Biol, 2011. **705**: p. 431-42.
59. Compagno, D., et al., *Galectins: major signaling modulators inside and outside the cell.* Curr Mol Med, 2014. **14**(5): p. 630-51.
60. Liu, F.T., R.J. Patterson, and J.L. Wang, *Intracellular functions of galectins.* Biochim Biophys Acta, 2002. **1572**(2-3): p. 263-73.
61. Elola, M.T., et al., *Galectins: matricellular glycan-binding proteins linking cell adhesion, migration, and survival.* Cell Mol Life Sci, 2007. **64**(13): p. 1679-700.
62. Stowell, S.R., et al., *Microbial glycan microarrays define key features of host-microbial interactions.* Nat Chem Biol, 2014. **10**(6): p. 470-6.
63. Camby, I., et al., *Galectin-1: a small protein with major functions.* Glycobiology, 2006. **16**(11): p. 137R-157R.
64. Chen, H.Y., et al., *Galectins as bacterial sensors in the host innate response.* Curr Opin Microbiol, 2014. **17**: p. 75-81.
65. Vasta, G.R., *Roles of galectins in infection.* Nat Rev Microbiol, 2009. **7**(6): p. 424-38.
66. Stowell, S.R., et al., *Innate immune lectins kill bacteria expressing blood group antigen.* Nat Med, 2010. **16**(3): p. 295-301.
67. Black, S., I. Kushner, and D. Samols, *C-reactive Protein.* J Biol Chem, 2004. **279**(47): p. 48487-90.
68. Emsley, J., et al., *Structure of pentameric human serum amyloid P component.* Nature, 1994. **367**(6461): p. 338-45.
69. Hind, C.R., et al., *Binding specificity of serum amyloid P component for the pyruvate acetal of galactose.* J Exp Med, 1984. **159**(4): p. 1058-69.
70. Ma, Y.J., et al., *Ficolin-1-PTX3 complex formation promotes clearance of altered self-cells and modulates IL-8 production.* J Immunol, 2013. **191**(3): p. 1324-33.
71. Bottazzi, B., et al., *Multimer formation and ligand recognition by the long pentraxin PTX3. Similarities and differences with the short pentraxins C-reactive protein and serum amyloid P component.* J Biol Chem, 1997. **272**(52): p. 32817-23.
72. Lee, J.K., et al., *The X-lectins: a new family with homology to the Xenopus laevis oocyte lectin XL-35.* Glycoconj J, 2004. **21**(8-9): p. 443-50.
73. Lee, J.K., et al., *Human homologs of the Xenopus oocyte cortical granule lectin XL35.* Glycobiology, 2001. **11**(1): p. 65-73.
74. Roberson, M.M. and S.H. Barondes, *Lectin from embryos and oocytes of Xenopus laevis. Purification and properties.* J Biol Chem, 1982. **257**(13): p. 7520-4.
75. Lee, J.K., et al., *Cloning and expression of a Xenopus laevis oocyte lectin and characterization of its mRNA levels during early development.* Glycobiology, 1997. **7**(3): p. 367-72.

76. Tsuji, S., et al., *Human intelectin is a novel soluble lectin that recognizes galactofuranose in carbohydrate chains of bacterial cell wall.* J Biol Chem, 2001. **276**(26): p. 23456-63.
77. Krarup, A., D.A. Mitchell, and R.B. Sim, *Recognition of acetylated oligosaccharides by human L-ficolin.* Immunol Lett, 2008. **118**(2): p. 152-6.
78. Wangkanont, K., et al., *Structures of Xenopus Embryonic Epidermal Lectin Reveal a Conserved Mechanism of Microbial Glycan Recognition.* J Biol Chem, 2016. **291**(11): p. 5596-610.
79. Suzuki, E., et al., *A monoclonal antibody directed to terminal residue of beta-galactofuranose of a glycolipid antigen isolated from Paracoccidioides brasiliensis: cross-reactivity with Leishmania major and Trypanosoma cruzi.* Glycobiology, 1997. **7**(4): p. 463-8.
80. Richards, J.C., M.B. Perry, and D.J. Carlo, *The specific capsular polysaccharide of Streptococcus pneumoniae type 20.* Can J Biochem Cell Biol, 1983. **61**(4): p. 178-90.
81. Besra, G.S., et al., *A new interpretation of the structure of the mycolyl-arabinogalactan complex of Mycobacterium tuberculosis as revealed through characterization of oligoglycosylalditol fragments by fast-atom bombardment mass spectrometry and 1H nuclear magnetic resonance spectroscopy.* Biochemistry, 1995. **34**(13): p. 4257-66.
82. Baumann, H., et al., *Structural elucidation of two capsular polysaccharides from one strain of Bacteroides fragilis using high-resolution NMR spectroscopy.* Biochemistry, 1992. **31**(16): p. 4081-9.
83. Tsuji, S., et al., *Differential structure and activity between human and mouse intelectin-1: human intelectin-1 is a disulfide-linked trimer, whereas mouse homologue is a monomer.* Glycobiology, 2007. **17**(10): p. 1045-51.
84. Voehringer, D., et al., *Nippostrongylus brasiliensis: identification of intelectin-1 and -2 as Stat6-dependent genes expressed in lung and intestine during infection.* Exp Parasitol, 2007. **116**(4): p. 458-66.
85. Peebles, R.S., Jr., *The intelectins: a new link between the immune response to parasitic infections and allergic inflammation?* Am J Physiol Lung Cell Mol Physiol, 2010. **298**(3): p. L288-9.
86. Pemberton, A.D., et al., *Innate BALB/c enteric epithelial responses to Trichinella spiralis: inducible expression of a novel goblet cell lectin, intelectin-2, and its natural deletion in C57BL/10 mice.* J Immunol, 2004. **173**(3): p. 1894-901.
87. Washimi, K., et al., *Specific expression of human intelectin-1 in malignant pleural mesothelioma and gastrointestinal goblet cells.* PLoS One, 2012. **7**(7): p. e39889.
88. Kudelka, M.R., et al., *Cosmc is an X-linked inflammatory bowel disease risk gene that spatially regulates gut microbiota and contributes to sex-specific risk.* Proc Natl Acad Sci U S A, 2016. **113**(51): p. 14787-14792.
89. Deckx, R.J., G.R. Vantrappen, and M.M. Parein, *Localization of lysozyme activity in a Paneth cell granule fraction.* Biochim Biophys Acta, 1967. **139**(1): p. 204-7.
90. Ouellette, A.J. and M.E. Selsted, *Paneth cell defensins: endogenous peptide components of intestinal host defense.* FASEB J, 1996. **10**(11): p. 1280-9.
91. Minty, A., et al., *Interleukin-13 is a new human lymphokine regulating inflammatory and immune responses.* Nature, 1993. **362**(6417): p. 248-50.

92. Brown, K.D., et al., *A family of small inducible proteins secreted by leukocytes are members of a new superfamily that includes leukocyte and fibroblast-derived inflammatory agents, growth factors, and indicators of various activation processes.* J Immunol, 1989. **142**(2): p. 679-87.
93. McKenzie, A.N., et al., *Interleukin 13, a T-cell-derived cytokine that regulates human monocyte and B-cell function.* Proc Natl Acad Sci U S A, 1993. **90**(8): p. 3735-9.
94. Smirnov, D.V., et al., *Tandem arrangement of human genes for interleukin-4 and interleukin-13: resemblance in their organization.* Gene, 1995. **155**(2): p. 277-81.
95. Moy, F.J., et al., *Solution structure of human IL-13 and implication for receptor binding.* J Mol Biol, 2001. **310**(1): p. 219-30.
96. Hershey, G.K., *IL-13 receptors and signaling pathways: an evolving web.* J Allergy Clin Immunol, 2003. **111**(4): p. 677-90; quiz 691.
97. Chomarat, P. and J. Banchereau, *Interleukin-4 and interleukin-13: their similarities and discrepancies.* Int Rev Immunol, 1998. **17**(1-4): p. 1-52.
98. LaPorte, S.L., et al., *Molecular and structural basis of cytokine receptor pleiotropy in the interleukin-4/13 system.* Cell, 2008. **132**(2): p. 259-72.
99. Wynn, T.A., *IL-13 effector functions.* Annu Rev Immunol, 2003. **21**: p. 425-56.
100. Miloux, B., et al., *Cloning of the human IL-13R alpha1 chain and reconstitution with the IL4R alpha of a functional IL-4/IL-13 receptor complex.* FEBS Lett, 1997. **401**(2-3): p. 163-6.
101. Donaldson, D.D., et al., *The murine IL-13 receptor alpha 2: molecular cloning, characterization, and comparison with murine IL-13 receptor alpha 1.* J Immunol, 1998. **161**(5): p. 2317-24.
102. Zheng, T., et al., *Cytokine regulation of IL-13Ralpha2 and IL-13Ralpha1 in vivo and in vitro.* J Allergy Clin Immunol, 2003. **111**(4): p. 720-8.
103. Fujisawa, T., B.H. Joshi, and R.K. Puri, *IL-13 regulates cancer invasion and metastasis through IL-13Ralpha2 via ERK/AP-1 pathway in mouse model of human ovarian cancer.* Int J Cancer, 2012. **131**(2): p. 344-56.
104. Fujisawa, T., et al., *A novel role of interleukin-13 receptor alpha2 in pancreatic cancer invasion and metastasis.* Cancer Res, 2009. **69**(22): p. 8678-85.
105. Kioi, M., et al., *Interleukin-13 receptor alpha2 chain: a potential biomarker and molecular target for ovarian cancer therapy.* Cancer, 2006. **107**(6): p. 1407-18.
106. Darnell, J.E., Jr., I.M. Kerr, and G.R. Stark, *Jak-STAT pathways and transcriptional activation in response to IFNs and other extracellular signaling proteins.* Science, 1994. **264**(5164): p. 1415-21.
107. Welham, M.J., et al., *Interleukin-13 signal transduction in lymphohemopoietic cells. Similarities and differences in signal transduction with interleukin-4 and insulin.* J Biol Chem, 1995. **270**(20): p. 12286-96.
108. Mikita, T., et al., *Requirements for interleukin-4-induced gene expression and functional characterization of Stat6.* Mol Cell Biol, 1996. **16**(10): p. 5811-20.
109. Kashiwada, M., et al., *Immunoreceptor tyrosine-based inhibitory motif of the IL-4 receptor associates with SH2-containing phosphatases and regulates IL-4-induced proliferation.* J Immunol, 2001. **167**(11): p. 6382-7.

110. Klingmuller, U., et al., *Specific recruitment of SH-PTP1 to the erythropoietin receptor causes inactivation of JAK2 and termination of proliferative signals.* Cell, 1995. **80**(5): p. 729-38.
111. Yasukawa, H., et al., *The JAK-binding protein JAB inhibits Janus tyrosine kinase activity through binding in the activation loop.* EMBO J, 1999. **18**(5): p. 1309-20.
112. Zhang, J.G., et al., *The conserved SOCS box motif in suppressors of cytokine signaling binds to elongins B and C and may couple bound proteins to proteasomal degradation.* Proc Natl Acad Sci U S A, 1999. **96**(5): p. 2071-6.
113. Ram, P.A. and D.J. Waxman, *Role of the cytokine-inducible SH2 protein CIS in desensitization of STAT5b signaling by continuous growth hormone.* J Biol Chem, 2000. **275**(50): p. 39487-96.
114. Liao, J., Y. Fu, and K. Shuai, *Distinct roles of the NH2- and COOH-terminal domains of the protein inhibitor of activated signal transducer and activator of transcription (STAT) 1 (PIAS1) in cytokine-induced PIAS1-Stat1 interaction.* Proc Natl Acad Sci U S A, 2000. **97**(10): p. 5267-72.
115. Mowen, K.A., et al., *Arginine methylation of STAT1 modulates IFNalpha/beta-induced transcription.* Cell, 2001. **104**(5): p. 731-41.
116. Jiang, H., M.B. Harris, and P. Rothman, *IL-4/IL-13 signaling beyond JAK/STAT.* J Allergy Clin Immunol, 2000. **105**(6 Pt 1): p. 1063-70.
117. White, M.F. and L. Yenush, *The IRS-signaling system: a network of docking proteins that mediate insulin and cytokine action.* Curr Top Microbiol Immunol, 1998. **228**: p. 179-208.
118. Andrews, A.L., et al., *Cytoplasmic tail of IL-13Ralpha2 regulates IL-4 signal transduction.* Biochem Soc Trans, 2009. **37**(Pt 4): p. 873-6.
119. Fichtner-Feigl, S., et al., *IL-13 signaling through the IL-13alpha2 receptor is involved in induction of TGF-beta1 production and fibrosis.* Nat Med, 2006. **12**(1): p. 99-106.
120. Rawlings, J.S., K.M. Rosler, and D.A. Harrison, *The JAK/STAT signaling pathway.* J Cell Sci, 2004. **117**(Pt 8): p. 1281-3.
121. Luckett-Chastain, L.R., et al., *SOCS3 modulates interleukin-6R signaling preference in dermal fibroblasts.* J Interferon Cytokine Res, 2012. **32**(5): p. 207-15.
122. Vignais, M.L., et al., *Platelet-derived growth factor induces phosphorylation of multiple JAK family kinases and STAT proteins.* Mol Cell Biol, 1996. **16**(4): p. 1759-69.
123. Pircher, T.J., et al., *Extracellular signal-regulated kinase (ERK) interacts with signal transducer and activator of transcription (STAT) 5a.* Mol Endocrinol, 1999. **13**(4): p. 555-65.
124. Decker, T. and P. Kovarik, *Serine phosphorylation of STATs.* Oncogene, 2000. **19**(21): p. 2628-37.
125. Bancroft, A.J., et al., *Gastrointestinal nematode expulsion in IL-4 knockout mice is IL-13 dependent.* Eur J Immunol, 2000. **30**(7): p. 2083-91.
126. Urban, J.F., Jr., et al., *IL-13, IL-4Ralpha, and Stat6 are required for the expulsion of the gastrointestinal nematode parasite Nippostrongylus brasiliensis.* Immunity, 1998. **8**(2): p. 255-64.

127. Grencis, R.K., *Cytokine regulation of resistance and susceptibility to intestinal nematode infection - from host to parasite.* Vet Parasitol, 2001. **100**(1-2): p. 45-50.
128. Urban, J.F., Jr., et al., *Stat6 signaling promotes protective immunity against Trichinella spiralis through a mast cell- and T cell-dependent mechanism.* J Immunol, 2000. **164**(4): p. 2046-52.
129. Bourreau, E., et al., *Interleukin (IL)-13 is the predominant Th2 cytokine in localized cutaneous leishmaniasis lesions and renders specific CD4+ T cells unresponsive to IL-12.* J Infect Dis, 2001. **183**(6): p. 953-9.
130. Noben-Trauth, N., W.E. Paul, and D.L. Sacks, *IL-4- and IL-4 receptor-deficient BALB/c mice reveal differences in susceptibility to Leishmania major parasite substrains.* J Immunol, 1999. **162**(10): p. 6132-40.
131. Mohrs, M., et al., *Differences between IL-4- and IL-4 receptor alpha-deficient mice in chronic leishmaniasis reveal a protective role for IL-13 receptor signaling.* J Immunol, 1999. **162**(12): p. 7302-8.
132. Flesch, I.E., A. Wandersee, and S.H. Kaufmann, *Effects of IL-13 on murine listeriosis.* Int Immunol, 1997. **9**(4): p. 467-74.
133. Atherton, H.C., G. Jones, and H. Danahay, *IL-13-induced changes in the goblet cell density of human bronchial epithelial cell cultures: MAP kinase and phosphatidylinositol 3-kinase regulation.* Am J Physiol Lung Cell Mol Physiol, 2003. **285**(3): p. L730-9.
134. Zhen, G., et al., *IL-13 and epidermal growth factor receptor have critical but distinct roles in epithelial cell mucin production.* Am J Respir Cell Mol Biol, 2007. **36**(2): p. 244-53.
135. Iwashita, J., et al., *mRNA of MUC2 is stimulated by IL-4, IL-13 or TNF-alpha through a mitogen-activated protein kinase pathway in human colon cancer cells.* Immunol Cell Biol, 2003. **81**(4): p. 275-82.
136. Perrais, M., et al., *Induction of MUC2 and MUC5AC mucins by factors of the epidermal growth factor (EGF) family is mediated by EGF receptor/Ras/Raf/extracellular signal-regulated kinase cascade and Sp1.* J Biol Chem, 2002. **277**(35): p. 32258-67.
137. Van der Sluis, M., et al., *Muc2-deficient mice spontaneously develop colitis, indicating that MUC2 is critical for colonic protection.* Gastroenterology, 2006. **131**(1): p. 117-29.
138. Johansson, M.E., et al., *The inner of the two Muc2 mucin-dependent mucus layers in colon is devoid of bacteria.* Proc Natl Acad Sci U S A, 2008. **105**(39): p. 15064-9.
139. Schaffler, A., et al., *Genomic structure of human omentin, a new adipocytokine expressed in omental adipose tissue.* Biochim Biophys Acta, 2005. **1732**(1-3): p. 96-102.
140. Look, D.C., et al., *Stat1 depends on transcriptional synergy with Sp1.* J Biol Chem, 1995. **270**(51): p. 30264-7.
141. Milanini-Mongiat, J., J. Pouyssegur, and G. Pages, *Identification of two Sp1 phosphorylation sites for p42/p44 mitogen-activated protein kinases: their implication in vascular endothelial growth factor gene transcription.* J Biol Chem, 2002. **277**(23): p. 20631-9.

Table 1.1: SOLUBLE LECTINS INVOLVED IN THE IMMUNE SYSTEM

Secreted human lectins involved in innate immunity					
Lectin Type (CRD)	**Subtype**	**Examples**	**Glycan Specificity**	**Function**	**Pathogen Specificity**
C-Type	Collectin	Mannose-Binding Lectin	mannose	activates complement system	E. coli, Streptococci, Staphylococcus, fungal pathogens, trypanosomes
		Surfactant Protein A	LPS, mannose	stimulate phagocytosis, agglutination	Pseudomonas, Klebsiella, Haemophilus, Streptococcus, Staphylococcus, Rotavirus (only SP-D)
		Surfactant Protein D			
	Reg	RegIIIα	membrane phospholipids	bactercidal (pore forming)	Gram-positive bacteria
Galectins	Prototype	Gal-1	n-linked oligosaccharides	inhibits virus attachment	Nipah virus and Hendra virus
	Chimera	Gal-3	galactomannan, mannan, LPS	bactercidal (membrane disruption)	Fungal pathogens
	Tandem-repeat	Gal-4, Gal-8	Gal-alpha-(1,3) gal		E. coli
Pentraxins	Long	PTX3	Protein A	activate complement system(via ficolins)	Klebsiella, fungal, and viral pathogens
	Short	C-reactive Protein (CRP)	phosphocholine within the C-polysaccharide	activate complement system	Streptococcus, Leishmania
Ficolin	-	ficolin-1 (M-ficolin)	sialic acid and N-acetyl groups	activates complement system	bacteria, viral pathogens
		ficolin-3 (H-ficolin)	LPS		
X-Type	Intelectin	human intelectin-1 (hIntL-1)	Glycerol Phosphate, β-Gal*f*, KDO, KO, agarose	unknown	bacteria
		human intelectin-2 (hIntL-2)	Glycerol Phosphate, agarose	unknown	

CHAPTER 2

COMPARATIVE MICROBIAL GLYCAN BINDING SPECIFICITIES OF THREE INTELECTINS: HINTL-1, HINTL-2, AND XCGL-1

Jonathan Viola**,** Jin Kyu Lee, Amika Sood, Richard Cummings, James Paulson, Robert Woods, Michael Pierce. To be submitted to Journal of Biological Chemistry

ABSTRACT

Intelectins are a novel class of X-type lectins with roles in innate immunity. The first X-type lectin was identified from *Xenopus laevis* oocytes and embryos. This *X. laevis* cortical granule lectin (XCGL-1 or XL-35) crosslinks the jelly-coat protein surrounding the oocyte to prevent polyspermy. Following the discovery of XCGL-1, two human intelectin (hIntL-1 and -2) members were identified. Here, we interrogated glycan microarrays with purified recombinant hIntL-1 and -2 and native XCGL-1. We found that both hIntL-1 and -2 interacted with pathogenic microbial glycans including: *P. vulgaris*, *S. pneumoniae, P. mirabilis*, and *K. pneumoniae*. hIntL-1 showed binding to microbial glycans containing β-galactofuranose (β-Gal*f*) and glycerol phosphate (GroP) moieties. In contrast, hIntL-2 showed minimal binding to glycans with β-Gal*f* and only in specific glycans did it bind GroP. XCGL-1 interacted with the majority of the *S. pneumoniae* glycans on the array, but not with the affinity of hIntL-1 or -2. Molecular dynamic methods were employed to validate the interactions between Gal*f* and hIntL-1 and -2. A simulation of β-Gal*f* and hIntL-1 showed β-Gal*f* remained bound for the entire simulation, indicating a stable interaction. hIntL-2, however, was unable to bind β-Gal*f* for the duration of the simulation. Amino acid sequence alignment of the three lectins showed multiple ligand-binding residue substitutions, suggesting these residues are critical for glycan specificity and affinity. These findings examine the differential glycan binding specificities of intelectins, and further support their role in the immune system as microbial recognition molecules.

INTRODUCTION

Mammalian cells have been shown to be covered extensively in a layer of glycans, termed the glycocalyx [1]. Depending on the cell type, these extracellular glycans are implicated in numerous functional roles and can also act as an attachment site for pathogens. Bacteria also possess a glycocalyx that can differ by strain [2]. A class of proteins termed lectins are important mediators in immunity and in the host-pathogen recognition process. Currently, there are several different classes of mammalian lectins that function in an immunity-related biological role. These lectins can differ in tissue expression, structural features, and glycan binding specificity.

The X-type class of lectins, which includes the intelectin family members, was first identified and characterized from *Xenopus laevis* oocytes and embryos [3-5]. To date, there are eight X-type lectins identified from *X. laevis:* two *X. laevis* cortical granule lectins (XCGL-1 (also termed XL-35) and XCGL-2), three *X. laevis* serum lectins (XCL-1, XCL-2, XCL-3), two *X. laevis* intelectins (xIntL-3 and xIntL-4), and one *X. laevis* embryonic epidermal lectin (XEEL). Each of these lectins have been described to engage in various roles in *X. laevis* innate immunity: XCGL-1 crosslinks the jelly coat protein surrounding oocytes to prevent polyspermy [6], XCL-1 is induced by lipopolysaccharides and binds Gram positive and negative bacteria [7], xIntL-3 is induced by lipopolysaccharides and agglutinates *Escherichia coli* [8], and XEEL agglutinates specific strains of *Streptococcus pneumoniae* [9]. Interestingly, these X-type lectins require calcium for carbohydrate binding, but do not contain the C-type carbohydrate recognition domain (CRD). The carbohydrate binding specificities of three of these *X. laevis* X-type lectins have been investigated. XCGL-1 binding to the jelly coat

protein was shown to be inhibited with free galactose [6], and more recently, was shown to interact with glycans terminating in alpha-linked galactopyranose by using a combination of enzyme-linked lectin assays and surface plasmon resonance (SPR) experiments [10]. Indeed, the biological ligand for XCGl-1, the jelly-coat protein, is extensively *O*-glycosylated with many of the glycans containing a terminal alpha-linked galactose [11]. Analysis of competition assays with recombinant xIntl-3 preabsorbed to galactose-sepharose and free saccharides showed competitive elution with specific pentoses (xylose, ribose) [8]. Lastly, SPR and crystallographic analysis of XEEL suggested interactions with β-galactofuranose (β-Gal*f*) and D-glycerol-1-phosphate (GroP) via calcium coordination to an exocyclic terminal 1,2-diol [9].

In contrast to the numerous *X. laevis* intelectins, there are usually only two intelectin members in a mammalian species with sheep expressing three intelectin homologues [12]. The two human intelectins (hIntL-1 and hIntL-2) were first identified by searching human expressed sequence tag (EST) databases with XCGL-1 cDNA and amino acids sequences [13]. These two human homologues share a high sequence homology (85% identity at amino acid level), but differ greatly in tissue expression. hIntL-1 mRNA expression was localized to the heart, colon, small intestine, ovary and thymus whereas hIntL-2 mRNA expression was restricted to the small intestine [13]. A similar gene expression pattern was observed in mice: immunohistochemical staining indicated that mouse intelectin-2 (mIntL-2) was localized to the Paneth cells in the small intestine [14]. These specialized cell types sit at the base of the intestinal crypt where they secrete antimicrobial peptides (defensins) as part of the innate immune response [15]. hIntL-1 was reported to have specificity for pentoses as well as galactofuranose and

ribofuranose by competition assays using free saccharides [16]. Gal*f* is not synthesized in a mammalian cell, but is commonly found on the extracellular surface of certain microbes including *Trypanosoma cruzi* [17], *Streptococcus pneumoniae* [18], *Mycobacterium tuberculosis* [19], and *Bacteroides fragilis* [20]. Recombinant hIntL-1 was shown to not only bind to microbes containing Gal*f* residues [16, 21], but also microbes containing GroP, D-*glycero*-D-talo-oct-2-ulosonic acid (KO), 3-deoxy-D-*manno*-oct-2-ulosonic acid (KDO), and heptose moieties [22]. The crystal structure of hIntL-1 in complex with β-Gal*f* helped elucidate the mechanism for this unique specificity: hIntL-1 can interact with the wide range of glycans via calcium coordination to a terminal acyclic 1,2 diol that is shared between all the glycans.

In this report, we sought to determine the glycan binding and pathogen specificities of the intelectins. To accomplish this, we recombinantly expressed hIntL-1 and hIntL-2 in suspension human embryonic kidney 293T cells (HEK 293T). Purification of the lectins was achieved by using a galactose-sepharose column. The lectins were assessed for purity using a silver stain and then subsequently labeled with AlexaFluor 488. Labeled lectins were then analyzed on the Consortium for Functional Glycomics (CFG) microbial glycan array. hIntL-1 showed binding to glycans containing β-Gal*f* and GroP, but, in contrast to what has been previously reported [22], no binding to heptose, KO, and KDO groups was observed. Compared to hIntL-1, hIntL-2 showed minimal binding to β-Gal*f* but interacted with GroP moieties on specific glycans. Natively purified XCGL-1 interacted with the majority of *S. pneumoniae* glycans on the array. To understand the binding differences between hIntL-1 and hIntL-2, we generated a homology model of hIntL-2 using the co-crystal structure of hIntL-1 and β-Gal*f*.

Molecular dynamics (MD) simulations confirmed the stable interaction between hIntL-1 and β-Gal*f*, but due to key residue substitutions in hIntL-2, no interaction was observed between hIntL-2 and β-Gal*f*. Together, these findings further support the role of X-type lectins in the immune system as microbial recognition molecules

MATERIALS AND METHODS

Human intelectin-1 and -2 (hIntL-1 and -2) expression and purification

The expression constructs of hIntL-1 and hIntL-2 were synthesized by PCR reaction and then cloned into SpeI and XbaI sites of pTracer-EF vector. The expression plasmids were transfected into suspension culture HEK293 (FreeStyle 293-F cells, Invitrogen) with protein production continuing for 4-5 days at 37^{o}C. The conditioned culture media was then harvested and clarified by centrifugation and stored at -20^{o}C until purification. For purification, conditioned media was passed through a 5-µm filter (Millipore) and adjusted to 10 mM $CaCl_2$. Media was loaded onto a β-galactofuranose-sepharose column that was equilibrated with 20 mM Tris-HCL, pH 7.5, 150 mM NaCl, 0.1% Tween 20, (TBS-T) containing 10 mM $CaCL_2$. The resin was then washed with 3 volumes of TBS-T containing 10 mM $CaCl_2$ and bound intelectin was eluted using 20 mM Tris-HCL, pH 7.5, 150 mM NaCl (TBS) containing 10 mM EDTA. Purification of the intelectins could be accomplished also with either galactose-conjugated Sepharose (Vector Labs), or Sephrose-4B (Sigma). Elution of the bound intelectin was also possible using TBS containing 100 mM ribose. Elution fractions were analyzed on a Western blot using an anti-intelectin antibody (Millipore) and fractions containing the highest concentrations of intelectin were pooled together and concentrated using a 10,000 molecular-weight cutoff (MWCO) Amicon Ultra centrifugal filter. Purity of the intelectin was assessed by Silver stain and Western blot.

Purification of native xenopus cortical granule lectin (XCGL-1)

XCGL-1 was purified from *Xenopus* oocytes as previously described [10]. Briefly, *xenopus* ooyctes were homogenized in TCS buffer with the homogenate being precipitated with 9 volume of acetone. The pellet was resuspended in in 0.3 M galactose and centrifuged at 10,000 x *g* for 20 minutes. The supernatant was applied to a Sepharose-4B (Sigma) column equilibrated with TCS. The resin was washed with 3 volumes of TCS and then eluted using TBS containing 50 mM EDTA. Elution of XCGL-1 was also possible using TBS containing 100 mM ribose. Fractions were concentrated as described above and purity was assessed be Silver stain and Western blot.

Alexa Fluor 488 labeling of intelectins

Purified intelectins were buffered exchanged to PBS and concentrated to 1 mg/mL. Intelectins were labeled as instructed by using an Alex Fluor 488 labeling kit. Briefly, intelectins were incubated with the Alexa Fluor 488 TFP ester for 15 minutes. Unreactive dye was separated from the labeled proteins using spin filters containing Bio-Gel P-6. Labeled intelectins were stored at 4^{o}C and protected from light. To show labeled intelectins retained carbohydrate-binding activity, 1 µg of each lectin was combined with 50 µL of a suspension of Sepharose-4B. After a 1 hour incubation, the beads were washed 3x with TBS-T containing 10 mM $CaCl_2$ and eluted with 50 uL of TBS-T containing 10 mM EDTA. Input, unbound, and elution fractions were run on a Western blot to determine carbohydrate-binding activity.

Interrogation of CFG mammalian array with hIntL-1 and XCGL-1

70 uL of labeled intelectins were applied to the array slide (v2.1 for XCGL-1 and v5.0 for hIntL-1) at a concentration of 200 ug/mL and incubated with a coverslip in a humidified chamber for 1 hour. After washing and drying, the slide was detected for fluorescence with a microarray scanner. After removing the highest and lowest signals of each of the 6 replicate spots, the average intensity of the remaining 4 spots was calculated. Average relative fluorescence data was graphed against the unique glycan identification number. Graphs were generated using Microsoft Excel.

Interrogation of the Microbial Glycan Microarray (MGM) array with hIntL-1, -2, and XCGL-1

Labeled intelectins were analyzed on the MGMv2 as described above [23]. For a negative control, hIntL-1 and -2 were incubated on the array slide in the presence of 100 mM ribose and then quantified as described above.

SDS-PAGE and Western Blot of intelectins

Intelectin-containing samples were boiled in Laemmli buffer (Bio-Rad), with or without 2-mercaptoethanol, and then separated on a SDS-PAGE gel (Bio-Rad). Proteins were transferred to a polyvinylidene fluoride using the iBlot Dry Blotting system (ThermoFisher). Membranes were blocked overnight in 5% nonfat milk in TBS-T. Intelectins were detected using a polyclonal anti-Omentin antibody (1:1,000, Millipore). Secondary antibodies (1:10,000, Santa Cruz) conjugated to horseradish peroxidase were then used followed by detection using enhanced chemiluminescence (ECL) substrate.

Molecular Dynamics of hIntL-1 and -2 with Galf

Preparation of hIntL-1 and -2 structures in complex with Galf- A crystal structure of hIntL-1 in complex with Galf (PDB ID: 4WMY) reported at 1.6Å resolution was used for the analysis [24]. Monomer (Chain B) with lower B-factor values was selected. A homology model of hIntL-2 in complex with Galf was generated using the SWISS-MODEL homology modeling server with the crystal structure of hIntL-1 as a template (88% identity) [25]. All the crystallized water molecules and Calcium ions were retained. Appropriate hydrogen atoms were added using Reduce [26], a tool provided by AMBERTOOLS. The protonation states of Histidine residues were also determined using Reduce. The complexes were solvated in a truncated octahedral box of TIP5P water molecules with counter ions (Cl^-) added to neutralize the charge using the tLEAP module of AMBER. In the case of the hIntL-1 structure, 13288 waters were required, whereas hIntL-2 required 17006 water molecules. AMBER14SB [27] parameters were utilized for the amino acids. The parameters for the Gal*f* monosaccharide can be downloaded from www.glycam.org.

Energy Minimization- To remove bad contacts, the complexes were minimized in a two-step process. For the first step, solute atoms were restrained (500 kcal/mol-A**2) while the water and ions were minimized. Minimization of the entire system followed this. Each minimization step consisted of an initial phase of steepest descent method for 5,000 cycles and then by conjugate gradient for 25,000 cycles.

MD Simulation- The minimized complexes underwent a MD simulation performed with the pmemd.cuda version of AMBER12 [28]. MD simulations were performed with a

nonbonded cutoff distance of 8.0 Å and long-range electrostatic interactions were treated with the Particle-Mesh Ewald algorithm [29]. The SHAKE algorithm was used to constrain hydrogen-containing bonds, allowing an integration time step of 2fs [30]. The systems were heated from 100K to 300K under nVT conditions over a span of 100 ps by using the Berendsen thermostat and employing the Langevin thermostat using a collision frequency of 2.0 ps^{-1}. An MD simulation was performed for 1ns and discarded to equilibrate the system. The simulation was then continued for 100 ns under NPT conditions with weak restraints on the protein backbone (10 kcal/mol-A**2).

RESULTS

Expression and purification of human intelectins

cDNA of hIntL-1 and hIntL-2 were inserted separately into the pTracer vector. Each vector was then transiently transfected into a suspension culture of HEK293 as previously described [31]. The secreted intelectins were purified from the conditioned media using galactose-affinity chromatography in the presence of Ca^{2+} and subsequently eluted using 100 mM ribose or 5 mM EDTA. Both hIntL-1 and hIntL-2 share a trimeric structure of 120 kDa under nonreducing conditions and a 37 kDa size under reducing conditions (**Figure 2.1**). These data indicate hIntL-2 is a disulfide-linked homotrimer like hIntL-1 despite a previous report suggesting hIntL-2 is a monomer [32]. Only one of the cystine residues (Cys 48) [33] that have been suggested to be involved in intermolecular disulfide bond formation in hIntL-1 is conserved in hIntL-2; however, there are other

cystine residues in the N-terminal domain of hIntL-2 that could potentially serve as a substitute (**Figure 2.2).**

Analysis of intelectins on the CFG array

To determine the glycan binding specificity of the hIntL-1 and XCGL-1, we utilized previously developed glycan microarray technologies that allows for the probing of multiple unique glycan structures on a single glass slide [34]. We labeled purified hIntL-1 and XCGL-1 with AlexFluor488 and then assayed the Consortium for Functional Glycomics (CFG) mammalian glycan array v2.1. This array consists of 285 mammalian glycans containing an amino linker that is printed onto an N-hydroxysuccinimide (NHS)-activated glass microscope slide via a covalent amide linkage. No binding was observed on the mammalian glycan array for hIntL-1 (**Figure 2.3)**, suggesting it does not interact with mammalian glycans, in agreement with what has been reported [22]. Interestingly, the highest three glycan signals for XCGL-1 contain a terminal alpha-linked galactose to either a GalNAc or Gal (**Figure 2.4 and 2.5)**. These structures have been shown to occur on the O-glycans of the *X. laevis* jelly coat protein found surrounding the mature oocyte [35]**.**

Analysis of intelectins on the MGM array

To test the hypothesis that the intelectins can bind microbial glycans, we utilized the Microbial Glycan Microarray (MGM) developed by the CFG. This array contains over 300 microbial glycans isolated from both Gram-positive and Gram-negative bacteria [23]. Interrogating the MGM with either hIntL-1, -2, or XCGL-1 (**Figure 2.6**) showed strong binding to several microbial glycans (**Figure 2.7**). This binding was inhibited when done in the presence of free ribose (data not shown), a sugar previously shown to

elute hIntL-1 from galactose-sepharose [16]. The top microbial glycan binders for hIntL-1 included glycans from both Gram-positive and Gram-negative pathogens: *P. vulgaris*, *S. pneumoniae*, *P. mirabilis*, and *K. pneumoniae*. Similar to what has been previously reported [22], hIntL-1 showed binding to microbial glycans containing both GroP and β-Gal*f* moieties. hIntL-1 showed stronger binding intensities to microbial glycans containing GroP as opposed to β-Gal*f*, with the GroP-containing microbial glycan from *P. mirabilis* O54ab (OE) being the only exception (**Figure 2.6**). Despite the similarities between these results and those from [22], our results did not show any binding of hIntL-1 to microbial glycans from *Y. pestis*. These microbial glycans from *Y. pestis* do not contain GroP or β-Gal*f*. The *Y. pestis* microbial glycans are unique among the other glycan binders in that only Y. pestis contain heptose, KO, and KDO residues. The top microbial glycan binders from our results suggest that hIntL-1 only binds to microbial glycans containing GroP and β-Gal*f*. It is possible that differences in the hIntL-1 protein structure (Strap-tag II vs no tag) and array detection method (StrapMAB-Classic Chromeo vs Alexfluor-488) contributed to these array discrepancies. hIntL-2 also showed binding to microbial glycans containing GroP on the MGM, but unlike hIntL-1, limited binding to glycans containing β-Gal*f* was observed. Lastly, XCGL-1 showed the highest binding to the microbial glycan from *E. coli* O86. Interestingly, this is the only microbial glycan structure on the array that contains a terminal gal-α-gal, a structure that XCGL-1 bound on the synthetic glycan array. In addition, XCGL-1 showed binding (>20% of highest signal) to the majority of the *S. pneumoniae* glycans on the MGM.

Intelectins show different binding patterns to the different S. pneumoniae serotypes

There are 23 unique capsular polysaccharides of *S. pneumoniae* that are printed on the MGM, each representing a different serotype. Comparing the three different intelectin binding patterns to these 23 different serotypes showed striking differences (**Table II**). Both hIntL-1 and hIntL-2 showed strong binding (75%> of highest signal) to only a few serotypes, 4 out of 23 and 1 out of 23, respectively. In contrast, XCGL-1 showed a broader binding (20-60% of highest signal) to the majority of the serotypes (17 out of 23).

Human intelectins differ in key binding site amino acid residues

The co-crystal structure of hIntL-1 bound with β-Gal*f* yielded important information for understanding the specificity of hIntL-1 [22]. Located near the C-terminus of hIntL-1, there are three different groups of residues that contribute unique features for allowing hIntL-1 to bind to its ligand (**Figure 2.8**). The first group consists of 5 residues (N243, E244, N260, E262, E274) which allow for hIntL-1 to coordinate a calcium ion that then interacts with the ligand. A second group (W288 and Y293) forms a binding pocket for the ligand and may also increase ligand affinity due to CH-π interactions [36]. Lastly, only one residue (H263) directly binds to the O(6) hydroxyl of the ligand via hydrogen-bonding. Comparing the alignment of these residues between hIntL-1 and -2 reveals two differences. First, the corresponding residue for E274 in hIntL-1 is a glutamine. This substitution may potentially lessen the ability for hIntL-2 to chelate a calcium ion; however, four other residues are available to coordinate a calcium ion. The second substitution in hIntL-2 changes Y297 to a serine residue, thereby potentially altering the binding pocket formed by W288 and Y297. Also, by substituting a non-aromatic residue, any CH- π interactions that contribute to the affinity for hIntL-1

would not be present for hIntL-2. A comparison of the key residues between hIntL-1 and XCGL-1 show additional substitutions. First, the two residues of the binding pocket of hIntL-1 are substituted from a tryptophan to a phenylalanine and a tyrosine to an asparagine. Second, the residue N243 that is involved in calcium coordination of hIntL-1 is substituted to a threonine.

Molecular Dynamics of hIntl-1, -2 interactions with Galf

To validate the interactions of hIntL-1 and -2 between Gal*f* that were observed on the MGM, molecular dynamic simulations between hIntL-1 and -2 and Gal*f* were performed. A simulation from the co-crystal structure of hIntL-1 bound to Gal*f* (PDB: 42MY), was performed for 50 ns. Throughout the simulation, Gal*f* remained bound in the binding pocket of hIntL-1. Three residues on hIntL-1 (H263, E274 and Y288) formed hydrogen bonds with the Gal*f* ligand (**Table 2.1**). These hydrogen bonds represented a stable interaction due to their percent occupancy being at least 50%. In contrast to hIntL-1, hIntL-2 was unable to bind Gal*f* for the entire simulation. A r.m.s.d. plot was generated of the average atomic coordinate positions of the Gal*f* ligand over the 100 ns simulation between hIntL-1 and -2 and Gal*f* (**Figure 2.10**). From this, the interaction of hIntL-1 with Gal*f* was shown to be a stable interaction whereas the interaction of hIntL-2 was shown to be unstable. Throughout the simulation of hIntL-1, the atomic coordinates of Gal*f* were shown to be largely constant.

DISCUSSION

As a fascinating family of glycan binding proteins, the intelectins have an innate immune function as secreted proteins that can bind to at least some pathogenic microbes.

In the sea-dwelling deuterostomes such as the Sea Pineapple (*Halocynthia roretzi*), intelectin is secreted into the plasma by the hepatopancreas [37]. In higher animals, intelectins are secreted into both the digestive and respiratory systems by epithelia and, in particular, the specialized goblet cells that also secrete mucins [14, 38]. In humans, mice, and frogs, there are reports of intelectins in serum, and *Xenopus* contains an epidermal form [21, 39]. Frogs also have another form of intelectin (XCGL-1) which is packaged in specialized vesicles and found in oocytes. These specialized vesicles, termed cortical granules, release their contents extracellularly at fertilization and cross-link the mucin layer to prevent polyspermy. Binding data on the mammalian glycan array from the CFG show striking specificity and avidity for the Galα1,3GalNAc disaccharide shown to be present on in this cortical mucin layer. This barrier likely also functions to exclude microbes in the surrounding water, perhaps another innate immune function.

The two human intelectins, hIntL-1 and hIntL-2, which take their names from the mistaken conclusion that they are only found in the intestine, have quite distinct tissue and cellular localizations. hIntL-2, as mentioned above, is constitutively secreted from the small intestinal Paneth cells, implying a function in constant microbial surveillance. hIntL-1, by contrast, is found constitutively expressed in different tissues and cell types, including the goblet cells of the colon, small intestine, and to a lesser degree, the mesothelial cells surrounding the heart and adipose [40, 41]. Moreover, hIntL-1 transcripts are significantly upregulated by the T helper type 2 (T_h2) cytokine signaling pathway [42]. Perhaps these epithelia are the source of serum hIntL-1, which in healthy adults has a range of 100-500 ng/ml [21, 43]. These levels of lectin imply a function in vascular surveillance, as well as in the small intestine and airways.

The binding properties of hIntL-1 on the MGM and CFG synthetic glycan microarray have has been reported [22], and our results agree well with these, except we see no binding to the outer glycans of *Y. pestis* (**Appendix Table 1**). The differences could be due to different methods used to detect hIntL-1 binding, but this is uncertain. The results presented here show that there are clear differences, however, in the binding profiles of hIntL-1 and hIntL-2 to the MGM. While hIntL-1 binds to glycans expressing GroP and β-Gal*f*, hIntL-2 only binds to specific glycans containing GroP. Interestingly, GroP is not the minimal binding determinant for hIntL-2, since no hIntL-2 binding was seen to the GroP-containing microbial glycan *S. pneumoniae* type 56. Binding analysis of both human lectins to the complement of *S. pneumoniae* strains present on the array reveals hIntL-1 binding to four of 23 strains, while hIntL-2 binds only to one, type 43. Clearly the glycan binding site is different between these two lectins, and these results can likely be explained by changes in two residues in an amino acid sequence implicated in glycan interaction by examination of the co-crystal structure of hIntL-1 and β-Gal*f*.

On the other hand, XCGL-1 has a very specialized function, but still shows broad specificity in binding *S. pneumoniae* serotypes, though it does not appear to bind GroP or β-Gal*f* specifically. If we did not know of its specialized function, and looked only at its binding to the MGM, we would infer that XCGL-1 was binding to a spectrum of pathogens, in contrast to the MGM binding pattern of either hIntL-1 or hIntL-2. No binding of XSL-1, -2, or XEEL on the MGM has been reported, and therefore it is not possible to speculate on how specialized functions could have influenced the pattern of specificity. Also, it is not clear how such a restricted binding pattern (e.g., human intelectins binding a minority of the *S. pneumoniae* serotypes) would be beneficial to the

human innate immune function. Considering that the MGM represents a very small subset of the human microbiota diversity, it is certainly possible there are other bacteria, both pathogenic and commensal, which are present in the respiratory, digestive, and circulatory systems to which the intelectins broadly bind with equal or greater specificity. Since it is estimated that there are 500-1,000 different species gut microbiota, this is not really surprising [44]. Yet, hIntL-1 and hIntL-2 function in the same intestine with distinct binding specificities, with the latter secreted constitutively in the intestinal crypt and the former released after induction into the lumen of the gut by goblet cells and epithelia.

XCGL-1, after affinity purification, is observed predominantly in twelve-mer form. This lectin agglutinates rabbit erythrocytes, and even more so after trypsinization of the red cells, because of its binding to Forsmann antigen, a glycolipid with an Gal α1,3 GalNAc moiety. We observed agglutination of some *S. pneumoniae* serotypes *in vitro* using native XCGL-1, and this agglutination was limited to only bacteria that showed binding on the MGM (data not shown). By contrast, neither hIntL-1 nor hIntL-2, which each exist as trimers, failed to show agglutination of bacteria *in vitro* with those species that showed MGM binding (data not shown). Surprisingly, neither hIntL-1 nor hIntL-2 showed bacteriostatic or bactericidal activities *in vitro* to those species that they bound on the MGM. It is not clear, therefore, how hIntL-1 and hIntL-2 function to thwart successful pathogen infection. The binding clefts for their microbial glycans ligands are quite shallow, compared to glycan-binding pockets on other anti-microbial lectins, and most of their protein structure is not involved in glycan binding. This observation suggests that other domains of the intelectins (i.e., collagen-like and fibrinogen-like

domain) may have potential binding affinities for additional, yet-to-be-discovered, receptors that could mediate their innate immune or other function.

REFERENCES

1. Rambourg, A. and C.P. Leblond, *Electron microscope observations on the carbohydrate-rich cell coat present at the surface of cells in the rat.* J Cell Biol, 1967. **32**(1): p. 27-53.
2. Costerton, J.W., R.T. Irvin, and K.J. Cheng, *The Bacterial Glycocalyx in Nature and Disease.* Annual Review of Microbiology, 1981. **35**: p. 299-324.
3. Roberson, M.M. and S.H. Barondes, *Lectin from embryos and oocytes of Xenopus laevis. Purification and properties.* J Biol Chem, 1982. **257**(13): p. 7520-4.
4. Nishihara, T., et al., *Isolation and characterization of a lectin from the cortical granules of Xenopus laevis eggs.* Biochemistry, 1986. **25**(20): p. 6013-20.
5. Lee, J.K., et al., *Cloning and expression of a Xenopus laevis oocyte lectin and characterization of its mRNA levels during early development.* Glycobiology, 1997. **7**(3): p. 367-72.
6. Wyrick, R.E., T. Nishihara, and J.L. Hedrick, *Agglutination of jelly coat and cortical granule components and the block to polyspermy in the amphibian Xenopus laevis.* Proc Natl Acad Sci U S A, 1974. **71**(5): p. 2067-71.
7. Nagata, S., S. Nishiyama, and Y. Ikazaki, *Bacterial lipopolysaccharides stimulate production of XCL1, a calcium-dependent lipopolysaccharide-binding serum lectin, in Xenopus laevis.* Dev Comp Immunol, 2013. **40**(2): p. 94-102.
8. Nagata, S., *Identification and characterization of a novel intelectin in the digestive tract of Xenopus laevis.* Dev Comp Immunol, 2016. **59**: p. 229-39.
9. Wangkanont, K., et al., *Structures of Xenopus Embryonic Epidermal Lectin Reveal a Conserved Mechanism of Microbial Glycan Recognition.* J Biol Chem, 2016. **291**(11): p. 5596-610.
10. Arranz-Plaza, E., et al., *High-avidity, low-affinity multivalent interactions and the block to polyspermy in Xenopus laevis.* Journal of the American Chemical Society, 2002. **124**(44): p. 13035-13046.
11. Strecker, G., et al., *Primary structure of 12 neutral oligosaccharide-alditols released from the jelly coats of the anuran Xenopus laevis by reductive beta-elimination.* Glycobiology, 1995. **5**(1): p. 137-46.
12. French, A.T., et al., *Expression of three intelectins in sheep and response to a Th2 environment.* Vet Res, 2009. **40**(6): p. 53.
13. Lee, J.K., et al., *Human homologs of the Xenopus oocyte cortical granule lectin XL35.* Glycobiology, 2001. **11**(1): p. 65-73.
14. Pemberton, A.D., et al., *Innate BALB/c enteric epithelial responses to Trichinella spiralis: inducible expression of a novel goblet cell lectin, intelectin-2, and its natural deletion in C57BL/10 mice.* J Immunol, 2004. **173**(3): p. 1894-901.
15. Ayabe, T., et al., *Secretion of microbicidal alpha-defensins by intestinal Paneth cells in response to bacteria.* Nat Immunol, 2000. **1**(2): p. 113-8.
16. Tsuji, S., et al., *Human intelectin is a novel soluble lectin that recognizes galactofuranose in carbohydrate chains of bacterial cell wall.* The Journal of biological chemistry, 2001. **276**(26): p. 23456-23463.
17. Suzuki, E., et al., *A monoclonal antibody directed to terminal residue of beta-galactofuranose of a glycolipid antigen isolated from Paracoccidioides*

brasiliensis: cross-reactivity with Leishmania major and Trypanosoma cruzi. Glycobiology, 1997. **7**(4): p. 463-8.
18. Richards, J.C., M.B. Perry, and D.J. Carlo, *The specific capsular polysaccharide of Streptococcus pneumoniae type 20.* Can J Biochem Cell Biol, 1983. **61**(4): p. 178-90.
19. Besra, G.S., et al., *A new interpretation of the structure of the mycolyl-arabinogalactan complex of Mycobacterium tuberculosis as revealed through characterization of oligoglycosylalditol fragments by fast-atom bombardment mass spectrometry and 1H nuclear magnetic resonance spectroscopy.* Biochemistry, 1995. **34**(13): p. 4257-66.
20. Baumann, H., et al., *Structural elucidation of two capsular polysaccharides from one strain of Bacteroides fragilis using high-resolution NMR spectroscopy.* Biochemistry, 1992. **31**(16): p. 4081-9.
21. Tsuji, S., et al., *Capture of heat-killed Mycobacterium bovis bacillus Calmette-Guérin by intelectin-1 deposited on cell surfaces.* Glycobiology, 2009. **19**(5): p. 518-526.
22. Wesener, D.A., et al., *Recognition of microbial glycans by human intelectin-1.* Nat Struct Mol Biol, 2015. **22**(8): p. 603-10.
23. Stowell, S.R., et al., *Microbial glycan microarrays define key features of host-microbial interactions.* Nat Chem Biol, 2014. **10**(6): p. 470-6.
24. Wesener, D.A., et al., *Recognition of microbial glycans by human intelectin-1.* Nature Structural &Amp; Molecular Biology, 2015. **22**: p. 603.
25. Arnold, K., et al., *The SWISS-MODEL workspace: a web-based environment for protein structure homology modelling.* Bioinformatics, 2006. **22**(2): p. 195-201.
26. D.A. Case, D.S.C., T.E. Cheatham, III, T.A. Darden, R.E. Duke, T.J. Giese, H. Gohlke, A.W. Goetz, D. Greene, N. Homeyer, S. Izadi, A. Kovalenko, T.S. Lee, S. LeGrand, P. Li, C. Lin, J. Liu, T. Luchko, R. Luo, D. Mermelstein, K.M. Merz, G. Monard, H. Nguyen, I. Omelyan, A. Onufriev, F. Pan, R. Qi, D.R. Roe, A. Roitberg, C. Sagui, C.L. Simmerling, W.M. Botello-Smith, J. Swails, R.C. Walker, J. Wang, R.M. Wolf, X. Wu, L. Xiao, D.M. York and P.A. Kollman, *AMBER 2017.* University of California, San Francisco, 2017.
27. Maier, J.A., et al., *ff14SB: improving the accuracy of protein side chain and backbone parameters from ff99SB.* Journal of chemical theory and computation, 2015. **11**(8): p. 3696-3713.
28. Gotz, A.W., et al., *Routine Microsecond Molecular Dynamics Simulations with AMBER on GPUs. 1. Generalized Born.* J Chem Theory Comput, 2012. **8**(5): p. 1542-1555.
29. Darden, T., D. York, and L. Pedersen, *Particle Mesh Ewald - an N.Log(N) Method for Ewald Sums in Large Systems.* Journal of Chemical Physics, 1993. **98**(12): p. 10089-10092.
30. Ryckaert, J.-P., G. Ciccotti, and H.J. Berendsen, *Numerical integration of the cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes.* Journal of Computational Physics, 1977. **23**(3): p. 327-341.
31. Barb, A.W., et al., *NMR characterization of immunoglobulin G Fc glycan motion on enzymatic sialylation.* Biochemistry, 2012. **51**(22): p. 4618-26.

32. Tsuji, S., et al., *Secretion of intelectin-1 from malignant pleural mesothelioma into pleural effusion.* Br J Cancer, 2010. **103**(4): p. 517-23.
33. Tsuji, S., et al., *Differential structure and activity between human and mouse intelectin-1: human intelectin-1 is a disulfide-linked trimer, whereas mouse homologue is a monomer.* Glycobiology, 2007. **17**(10): p. 1045-51.
34. Blixt, O., et al., *Printed covalent glycan array for ligand profiling of diverse glycan binding proteins.* Proc Natl Acad Sci U S A, 2004. **101**(49): p. 17033-8.
35. Guerardel, Y., et al., *O-glycan variability of egg-jelly mucins from Xenopus laevis: characterization of four phenotypes that differ by the terminal glycosylation of their mucins.* Biochem J, 2000. **352 Pt 2**: p. 449-63.
36. Hudson, K.L., et al., *Carbohydrate-Aromatic Interactions in Proteins.* J Am Chem Soc, 2015. **137**(48): p. 15152-60.
37. Abe, Y., et al., *A unique primary structure, cDNA cloning and function of a galactose-specific lectin from ascidian plasma.* Eur J Biochem, 1999. **261**(1): p. 33-9.
38. French, A.T., et al., *The expression of intelectin in sheep goblet cells and upregulation by interleukin-4.* Vet Immunol Immunopathol, 2007. **120**(1-2): p. 41-6.
39. Nagata, S., et al., *Developmental expression of XEEL, a novel molecule of the Xenopus oocyte cortical granule lectin family.* Dev Genes Evol, 2003. **213**(7): p. 368-70.
40. Washimi, K., et al., *Specific expression of human intelectin-1 in malignant pleural mesothelioma and gastrointestinal goblet cells.* PLoS One, 2012. **7**(7): p. e39889.
41. Lee, J., et al., *Human homologs of the Xenopus oocyte cortical granule lectin XL35.* Glycobiology, 2001. **11**(1): p. 65-73.
42. Zhen, G., et al., *IL-13 and epidermal growth factor receptor have critical but distinct roles in epithelial cell mucin production.* Am J Respir Cell Mol Biol, 2007. **36**(2): p. 244-53.
43. Turan, H., et al., *Omentin serum levels and omentin gene Val109Asp polymorphism in patients with psoriasis.* Int J Dermatol, 2014. **53**(5): p. 601-5.
44. Savage, D.C., *Microbial ecology of the gastrointestinal tract.* Annu Rev Microbiol, 1977. **31**: p. 107-33.
45. Robert, X. and P. Gouet, *Deciphering key features in protein structures with the new ENDscript server.* Nucleic Acids Res, 2014. **42**(Web Server issue): p. W320-4.

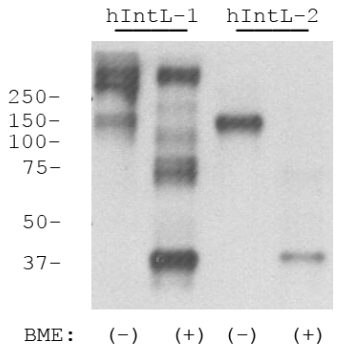


Figure 2.1 RECOMBINANTLY-EXPRESSED HUMAN INTELECTINS ARE BOTH A DISULFIDE-LINKED TRIMER. Western blot of hIntL-1 and hIntL-2 from purified supernatant of transfected HEK293 cells, +/- the addition of 2-mercaptoethanol (BME).

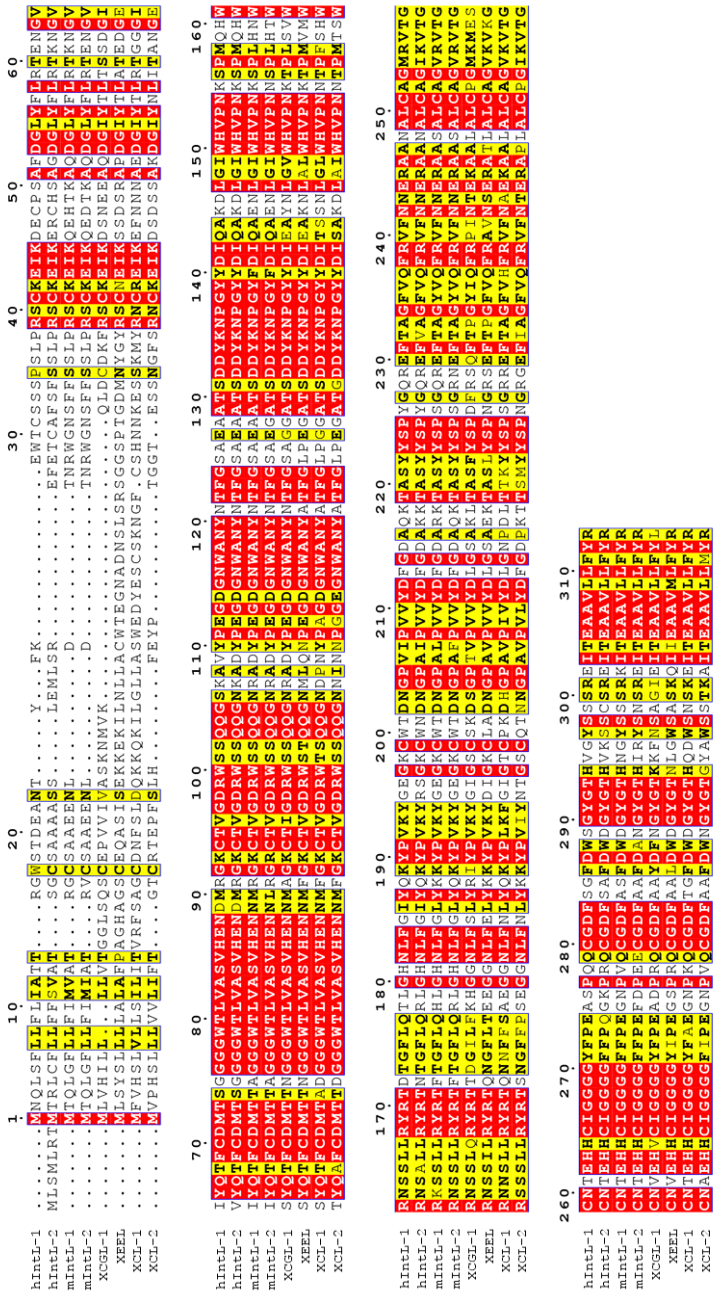


Figure 2.2 AMINO ACID SEQUENCE ALIGNMENT OF X-TYPE FAMILY MEMBERS. Sequence alignment of human (hIntL-1 and hIntL-2), mouse (mIntL-1 and mIntL-2) and xenopus (XCGL-1, XEEL, XCL-1 and XCL-2) intelectins using ESPript 3.0 [45]. Red boxes indicate identical residues and yellow boxes indicate similar residues based on physicochemical properties.

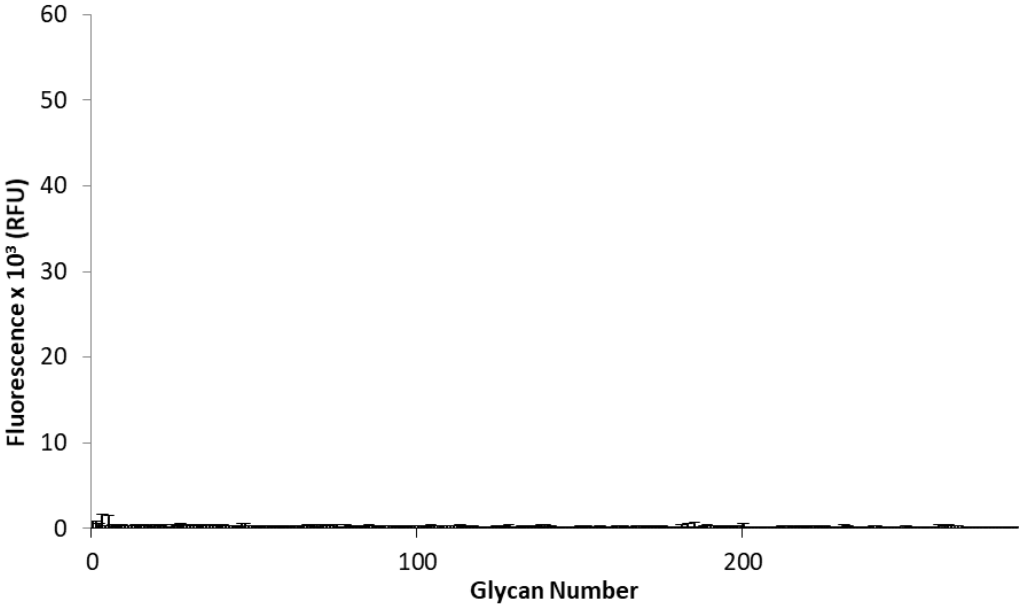


Figure 2.3 HINTL-1 DOES NOT BIND GLYCANS ON THE CFG MAMMALIAN GLYCAN ARRAY. CFG v2.1 array probed with hIntL-1

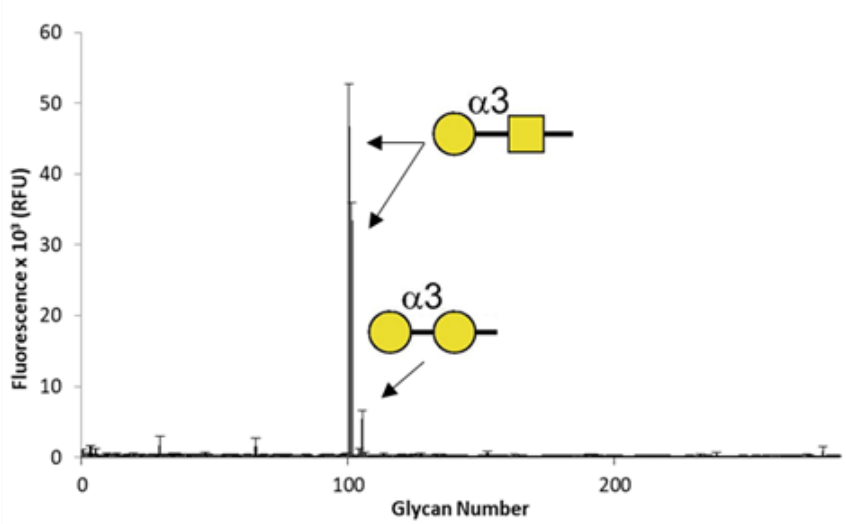


Figure 2.4 XCGL-1 INTERACTS WITH GLYCANS TERMINATING IN ALPHA-LINKED GALACTOSE**.** CFG v2.1 array proved with XCGL-1

Glycan	% of highest signal
Galα1-3GalNAcα-Sp8	100%
Galα1-3GalNAcβ–Sp8	71%
Galα1-3Galβ1-4Glcβ–Sp0	13%
Ceruloplasmine	3%
Galα1-3Galβ1-4GlcNAcβ–Sp8	2%
Transferrin	2%

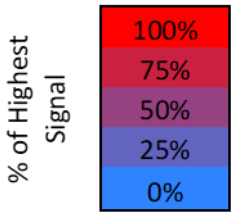
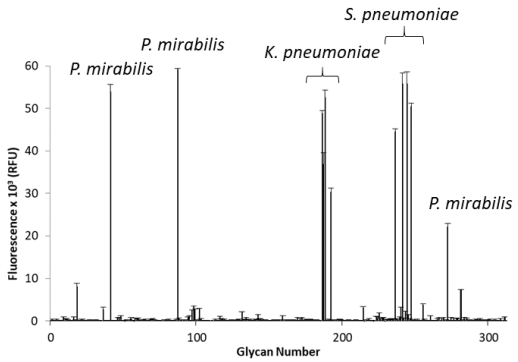
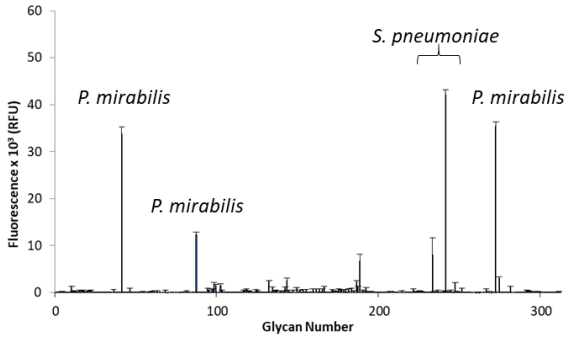


Figure 2.5 HEATMAP OF TOP GLYCAN HITS FOR XCGL-1 ON CFG. List of top glycans hits, sorted by percent of highest XCGL-1 binding. RFU, Relative Fluorescence Units. Error bars represent standard error of the mean

(A)



(B)



(C)

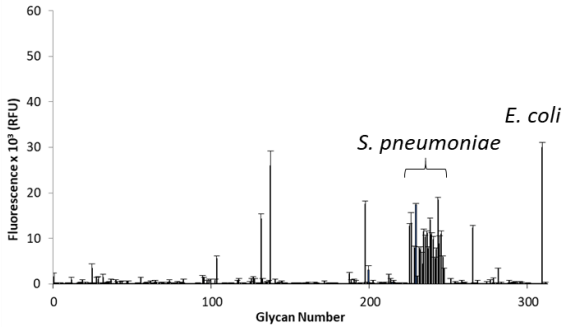


Figure 2.6 HINTL-1, HINTL-2, AND XCGL-1 INTERACT WITH SPECIFIC MICROBIAL GLYCANS ON THE MGM. (A) MGM v2 array probed with hIntL-1. (B) MGM v2 array probed with hIntL-2 (C) MGM v2 array probed with XCGL-1. RFU, Relative Fluorescence Units. Error bars represent standard error of the mean.

Proposed Ligand	Microbe	hIntL-1	hIntL-2	XCGL-1
Glycerol Phosphate	*P. vulgaris* O54a,54c (TG 103)	100%	29%	0%
	S. pneumoniae type 43	94%	100%	18%
	S. pneumoniae type 56	94%	1%	29%
	P. mirabilis O54a, 54b	91%	80%	0%
β-Gal*f*	*K. pneumoniae* O2ac OPS	89%	16%	3%
	S. pneumoniae type 70	85%	5%	11%
	K. pneumoniae O1 OPS	83%	5%	0%
	S. pneumoniae type 20	75%	1%	37%
	K. pneumoniae O2a OPS	62%	3%	7%
	K. pneumoniae O8 OPS	51%	3%	0%
Glycerol Phosphate	*P. mirabilis* O54ab (OE)	38%	84%	0%

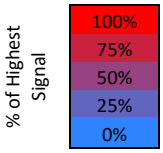


Figure 2.7 HEATMAP OF RELATIVE BINDING AFFINITIES OF INTELECTINS TO MICROBIAL GLYCANS. hIntL-1 interacts with glycans containing GroP and β-Gal*f*, hIntL-2 interacts with select glycans containing GroP, XCGL-1 interacts with select glycans containing GroP and β-Gal*f*, but at a lower affinity.

Microbe	hIntL-1	hIntL-2	XCGL-1
S. pneumoniae type 1	3%	0%	42%
S. pneumoniae type 2	1%	0%	45%
S. pneumoniae type 3	1%	0%	0%
S. pneumoniae type 4	1%	1%	26%
S. pneumoniae type 5	0%	0%	58%
S. pneumoniae type 8	0%	0%	5%
S. pneumoniae type 9	1%	1%	26%
S. pneumoniae type 12	0%	1%	25%
S. pneumoniae type 14	1%	0%	15%
S. pneumoniae type 17	1%	3%	38%
S. pneumoniae type 19	1%	0%	34%
S. pneumoniae type 20	75%	1%	37%
S. pneumoniae type 22	1%	1%	26%
S. pneumoniae type 23	0%	0%	47%
S. pneumoniae type 26	1%	0%	36%
S. pneumoniae type 34	5%	1%	33%
S. pneumoniae type 43	94%	100%	18%
S. pneumoniae type 51	1%	1%	25%
S. pneumoniae type 54	4%	0%	61%
S. pneumoniae type 56	94%	1%	29%
S. pneumoniae type 57	3%	0%	36%
S. pneumoniae type 68	1%	1%	19%
S. pneumoniae type 70	85%	5%	11%

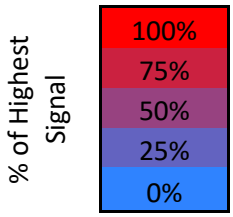


Figure 2.8: HEATMAP OF RELATIVE BINDING AFFINITIES OF INTELECTINS TO *S. PNEUMONIAE* MICROBIAL GLYCANS. List of 23 different *S. pneumoniae* microbial glycans printed on the MGM showing relative binding intensity of hIntl-1, hIntL-2 and XCGL-1.

Figure 2.9 ALIGNMENT OF KEY RESIDUES OF HINTL-1, -2, AND XCGL-1 INVOLVED IN LIGAND BINDING. Cyan boxes indicate residues involved in calcium coordination, red asterisks indicate residues involved in direct ligand binding, and yellow boxes indicate residues involved in forming the binding pocket for hIntL-1 as determined by [22].

Table 2.1: Average Hydrogen bond lengths and occupancies between hIntL-1 and Gal*f*

Donor		Acceptor			
Residue	Atom	Residue	Atom	Occupancy (%)	Distance (Å)
Galf	O6	His 263	NE2	99	2.90
Galf	O5	Glu 274	O5	77	2.72
Trp 288	NE1	Galf	O3	50	2.98

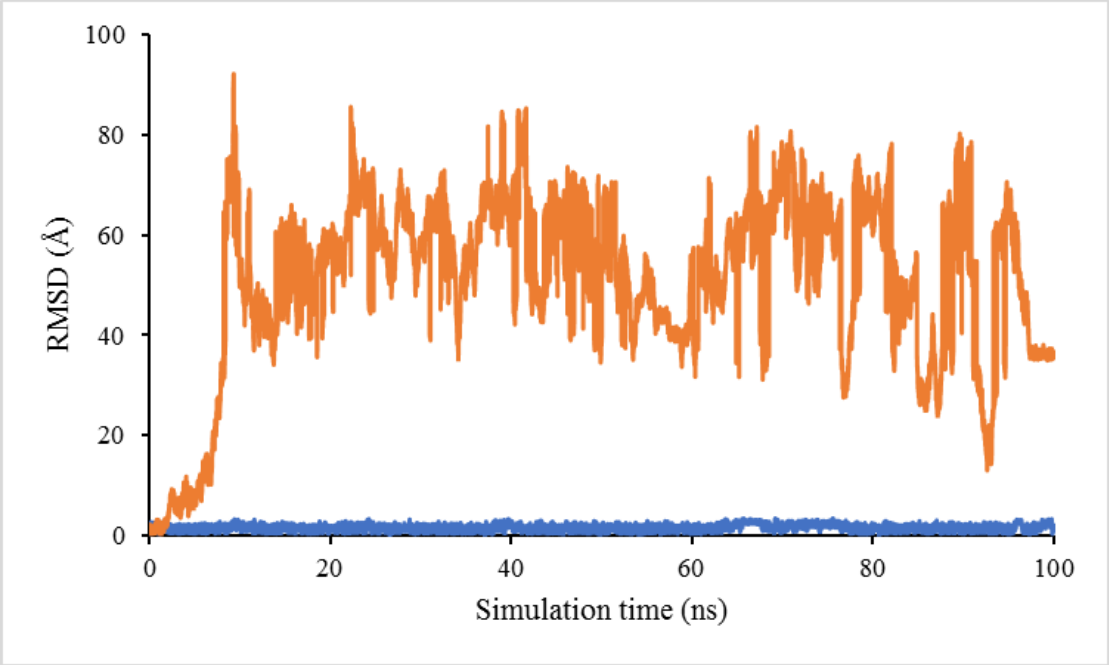


Figure 2.10 RMSD PLOT OF HINTL1- AND -2 INTERACTIONS WITH GALF. Time series of the RMSD values of the heavy atoms in Gal*f* interacting with hIntL-1 (blue) and hIntL-2 (orange), relative to the minimized conformation of the complex

CHAPTER 3

THE MAPK AND JAK/STAT SIGNALING PATHWAYS ARE REQUIRED FOR THE IL-13-INDUCTION OF HUMAN INTELECTIN-1 EXPRESSION

Jonathan Viola, Jin Kyu Lee, Michael Pierce. Submitted to FEBS Letters, 12/05/2017

ABSTRACT

The cytokine IL-13 is involved in numerous functions such as promoting cancer cell growth to inducing innate immunity with a complicated signaling network involving the MAPK and JAK/STAT pathways. In addition, IL-13 is responsible for the activation of many genes including Muc2 and human intelectin-1 (hIntL-1). In this report, the signaling mechanisms of IL-13-induced hIntL-1 activation using a colon cancer cell line are defined. IL-13 activates both the MAPK and JAK/STAT pathways to induce the expression and secretion of hIntL-1, an innate immune lectin. Pharmacologically inhibiting either the MAPK or JAK/STAT pathway prevents the IL-13-induced secretion of hIntL-1. These results suggest hIntL-1 activation is co-dependent on both the MAPK and JAK/STAT pathways, and highlight the complexities of the crosstalk between the MAPK and JAK/STAT.

INTRODUCTION

Following a pathogen challenge, it is common for the host to initiate a T helper type 2 (Th2) immune response. This Th2 response results in the recruitment and activation of an important group of cells involved in the innate immune response. These effector cells consist of basophils, mast cells, and eosinophils, which also have roles in allergic responses. The Th2 response causes the induction and secretion of several cytokines including IL-4, IL-5, IL-9, and IL-13, with the IL-4 cytokine initiating a forward feedback amplification loop for the Th2 response. IL-13 was originally thought to be a redundant cytokine due to it sharing the same receptor for IL-4; however, there are now several reports showing that IL-13 functions as more than an anti-inflammatory cytokine. These additional functions include nematode expulsion, allergic inflammation, goblet cell hyperplasia, and tumor cell growth [1].

IL-13 signaling can initiate two separate signal transduction pathways. The first pathway is through a heterodimer receptor consisting of IL-13Rα1 and IL-4Rα. IL-13 first binds to the receptor IL-13Rα1 which then recruits the receptor IL-4Rα to form a heterodimer receptor complex. This complex signals through a Janus kinase 1 (JAK1) which results in the activation of STAT6 and subsequently the activation of STAT6-dependent genes. Interestingly, IL-4 can bind the IL-4Rα receptor which recruits the IL-13Rα1 receptor to initiate the JAK/STAT signaling cascade [2]. IL-13 can also signal through a separate mechanism involving the receptor IL-13Rα2. This receptor has been reported to serve as a decoy receptor to sequester IL-13 with high affinity [3, 4] and its expression varies between different types of cancers: it is highly expressed in cancer cells such as head and neck cancer, pancreatic cancer, glioblastoma, and ovarian cancer [5-7],

but is undetectable via RT-PCR methods in colon colorectal epithelia [8]. More recently, IL-13Rα2 was shown to be required in the IL-13 activation of the MAPK pathway. Following IL-13 stimulation, human ovarian and pancreatic cancer cells expressing IL-13Rα2 exhibited MAPK activation with the downstream transcription factor API-1 also being activated [9] [7, 10]. Interestingly, IL-13Rα2 expression was positively correlated with the metastasis of these two cancer types, suggesting the potential of IL-13Rα2 as a biomarker. For cells that express low levels of IL-13Rα2, IL-13 stimulation was shown to have no effect on MAPK signaling [11]. Lastly, IL-13 may also initiate a forward feedback loop within this signaling cascade: human keratinocytes showed an increased expression of the IL-13Rα2 following IL-13 stimulation [12]. Together, these results illustrate the complex signaling mechanisms initiated by IL-13 on cancerous cell types to promote disease progression.

IL-13 has been shown to be an important effector molecule in both the respiratory and digestive systems. In the respiratory system, IL-13 stimulation of human bronchial cells will induce the expression of several genes including Muc5AC and human intelectin-1 (hIntL-1) [13] . A similar scenario is also reported in the digestive system: IL-13 stimulation of human colon cancer cells induces the expression of another mucin gene, MUC2 [14]. Both of the mucin proteins, Muc2 and Muc5AC, and the hIntL-1 are secreted proteins involved in innate immunity with an evolutionary history that can be traced back to early metazoans [15-17]. hIntL-1, a member of the X-type lectin family, binds specific pathogenic microbes including specific serotypes of *Streptococcus pneumoniae* and is reported to function as a microbial detector [18].

The signaling pathways required for the pathogen and cytokine induction of intelectin and Muc2 was previously investigated. The stimulation of Muc2 by IL-13 was shown to be MAPK-dependent in human colonic cell lines. With IL-13 and a MAPK inhibitor treatment, goblet-like colon cells were shown to revert to a basal level of Muc2 secretion [14]. Within a mouse system, intelectin induction via the helminth parasite *Nippostrongylus brasiliensis* or the pathogenic bacteria *Mycobacterium tuberculosis* was shown to be STAT6-dependent. Mice deficient in STAT6 were unable to induce either intelectin-1 or intelectin-2 expression in the lung. Interestingly, intelectin-1 expression in the small intestine was not affected in STAT6-deficient mince compared to wildtype, suggesting a complex regulatory mechanism between the two tissue types [19].

The secretion of intelectin in the respiratory system via the cytokines IL-13 and IL-4 has been studied in both humans and sheep [13, 20]; however, the induction of hIntL-1 within the digestive system is not well understood. We therefore sought to define the signaling mechanisms required for intelectin induction within the gut. Using a human goblet-like colon cancer cell line, we showed the induction and secretion of hIntL-1 via IL-13 signaling. Treatment with IL-13 activated both the STAT6 and MAPK pathways within these cells. By knocking down the MAPK or JAK/STAT pathway using pharmacological inhibitors, we showed a complete loss of IL-13-induced hIntL-1 secretion. These results indicate that hIntL-1 is co-dependent on the MAPK and JAK/STAT signaling pathways.

MATERIAL AND METHODS

Cell Culture

LS174T cells were obtained from American Type Culture Collection (ATCC). Cells were grown in Dulbecco's Modified Eagle Medium (DMEM), supplemented with 10% Fetal Bovine Serum (Atlanta Biologicals) at 37ºC in 5% CO_2 and 95% humidity.

IL-13 induction

LS174T cells were seeded onto a 6-well plate. 24 hours later, cells were washed with PBS and incubated for 72 hours in Opti-MEM (Life Technologies) with IL-13 (Peprotech) at a concentration of 10 ng/mL. For MAPK and STAT6 signaling analysis, cells were stimulated with IL-13 for 10 minutes and 30 minutes, respectively.

RT-PCR: +/- IL-13 on LS174T intelectin transcripts

LS174T cells were induced with IL-13 for 72 hours as described above. mRNA was isolated using TRIzol solution (ThermoFisher).

Intelectin from LS147T binding to agarose

LS174T cells were induced with IL-13 as described above. After 72 hours, the conditioned culutre media was collected and pelleted to remove cellular debris. Galactose-agarose (Vector Labs) was added to the media in the presence of 10 mM $CaCl_2$ and incubated for 2 hours at 4ºC. The resin was washed with tris-buffered saline and 0.05% Tween-20 (TBS-T) containing 10 mM $CaCl_2$. Bound hIntL-1 was eluted using TBS-T containing 5 mM EDTA and detected on a western blot.

MAPK and JAK/STAT inhibition

LS174T cells were seeded onto a 6-well plate. For MAPK inhibition, cells were washed with PBS and pretreated with 30 µM U0126 (Cell Signaling) or DMSO as a

negative control for 1 hour. For JAK/STAT knockdown, cells were washed with PBS and pretreated with 30 μM JAK1 inhibitor (Millipore) or DMSO as a negative control for 2 hours. Cells were than induced with IL-13 (10 ng/mL) for 10 minutes or 30 minutes for MAPK and STAT6 activation, respectively.

Western and Phospho-Western blot analysis:

Following IL-13 induction, conditioned media from LS174T was removed and the cell monolayer was washed 3x with PBS. Cells were lysed with a lysis buffer (50 mM Tris, pH 8.0, 150 mM NaCL, 1% Triton X-100) containing 100x protease inhibitor cocktail (Cell Signaling) and 100x phosphatase inhibitor cocktail (Cell Signaling). Cellular debris were pelleted at 4^{o}C at 18,000 x *g* for 10 min. Protein concentrations were determined using the bicinchoninic acid assay (BCA) (Pierce). Samples were boiled in Laemmli buffer (Bio-Rad) containing 5% 2-mercaptoethanol and then separated by sodium dodecyl sulphate (SDS)-polyacrylamide gels (10% acrylamide, Bio-Rad). Proteins were transferred to polyvinylidene fluoride using iBlot Dry Blotting system (ThermoFisher) and blocked overnight in 5% nonfat milk in tris-buffered saline and 0.05% Tween-20 (TBS-T). Intelectin was detected using anti-Omentin antibody (1:1,000, Milipore), Mucin 2 was detected using anti-Muc2 antibody (1:1,000, AbCam), β-Actin was detected using anti-β-Actin antibody (1:5,000, Sigma). Secondary antibodies (Santa Cruz) were conjugated to horseradish peroxidase and detected using enhanced chemiluminescence (ECL). For phopho-western blot analysis, membranes were blocked overnight in 3% bovine serum albumin (BSA) in TBS-T. Phopho-STAT6 was detected using anti-phospho-STAT6 (1:1,000 Cell Signaling), Phospho-MAPK was detected using anti-phospho-MAPK (1:500 Cell Signaling), STAT6 was detected using anti-STAT6

(1:1000, Cell Signaling) and MAPK was detected using anti-MAPK (1:1000, Cell Signaling).s Secondary antibodies (Santa Cruz) were conjugated to horseradish peroxidase and detected using enhanced chemiluminescence (ECL).

RESULTS:

IL-13 induces expression and secretion of human intelectin

To determine the signaling pathways required for human intelectin-1 (hIntL-1) induction, we utilized the LS174T cell line. This is a goblet-like colon adenocarcinoma and does not secrete detectable levels of hIntL-1 or -2 according to Western Blot analysis (**Figure 3.1**). However, upon stimulation with IL-13, LS174T cells upregulate hIntL-1 transcripts ($7x10^4$-fold) and protein levels (**Figure 3.1**). The secretion of Muc2, a gel-forming glycosylated protein, was also observed to be upregulated following IL-13 induction in accordance to other reports [14]. hIntl-2 transcripts were also upregulated by IL-13 ($2x10^2$-fold, data not shown), however, hIntL-2 was not detected by Western blotting (**Figure 3.1**).

Secreted hIntL-1 from LS174T is active

Previous studies have indicated that hIntL-1 binds specific moieties on microbial glycans such as glycerol phosphate and galactofuranose (Gal*f*) [18] as well as galactose-conjugated agarose [21, 22].To assess the carbohydrate-binding activity of the induced hIntL-1, we incubated conditioned culture media from IL-13 stimulated LS174T cells with galactose-agarose in the presence of calcium. hIntL-1 from the conditioned media bound to the galactose-agarose and was subsequently eluted off using EDTA, indicating the carbohydrate-binding properties of LS174T-deriverd hIntL-1 are active **(Figure 3.2)**.

IL-13 activates both MAPK and STAT6 pathways in LS174T cells

IL-13 stimulation has been shown to activate both the MAPK and STAT6 pathways in specific cell types [1]. To determine if the STAT6 and MAPK pathways in LS174T can also be activated in LS174T cells, we stimulated the cells with IL-13 and analyzed phosphoprotein levels of STAT6 and ERK1/2. The results reveal that IL-13 activates both the MAPK and STAT6 pathways in LS174T cells (**Figure 3.3A**). Using the MEK1/2 inhibitor U0126 and JAK1 inhibitor, a successful knockdown of the MAPK and JAK/STAT pathway was observed, respectively (**Figure 3.3**).

hIntL-1 induction via IL-13 is co-dependent on MAPK and JAK/STAT in LS174T. To determine the signaling pathway(s) required for IL-13 induction of hIntL-1, LS174T cells were inhibited with either U0126 or JAK1 inhibitor and then stimulated with IL-13. After knocking down the MAPK pathway in LS174T using U0126 and then stimulating with IL-13 for 72 hours, a complete loss of hIntL-1 secretion was observed (**Figure 3.4**). Muc2 secretion levels were also decreased following U0126 treatment. Pretreating LS174T cells with the JAK1 inhibitor and subsequent IL-13 stimulation also resulted in no secretion of hIntL-1 (**Figure 3.4**). Thus, hIntL-1 induction requires both activated MAPK and JAK/STAT pathways.

DISCUSSION

IL-13 has been shown to activate both the MAPK and JAK/STAT pathways in different cell types. We also report that IL-13 activates these pathways in the LS174T cell line. The LS174T cell line is a goblet-like adenocarcinoma that constitutively secretes Muc2. An innate immune protein, Muc2 forms a thick protective barrier, separating the

epithelial cells of the host from the microbiome within the lumen of the gut. The barrier formed by Muc2 can be classified into an inner and outer layer, with the inner layer suggested to be sterile [23]. The Muc2 protein can be upregulated in response to pathogen infection, cytokines, and EGFR ligands [24, 25]. Reports have shown that the induction of Muc2 via butyrate as well as the cytokines IL-4, IL-13 and TNF-α occurs through the MAPK pathway in human colon cancer cells [14, 26]. In agreement with this, we show that pharmacological inhibition of the MAPK pathway in LS174T blocks the induction and secretion of Muc2, further highlighting the requirement of MAPK signaling in Muc2 production.

The mouse homolog of hIntL-1 has been previously reported to be a STAT6-depenent gene [19]. In this report, we show for the first time that hIntL-1 induction via IL-13 is co-dependent on both the MAPK and STAT6 pathways. Pharmacologically inhibiting either the MAPK pathway or JAK/STAT pathway results in the loss of hIntL-1 secretion. Interestingly, knocking-down the MAPK pathway in LS174T negatively affects Muc2 secretion, whereas JAK/STAT inhibition has no effect on Muc2 secretion. It is possible that the codependence of hIntL-1 on the JAK/STAT and MAPK pathways is due to crosstalk between the two pathways. Such crosstalk has been documented in other cell types. Within Jurkat T cells, IL-4 was shown to activate the MAPK pathway, and that IL-4-induceded STAT6 activity was suppressed following Erk inhibition [27]. More recently, it was shown that both the MAPK and STAT3 signal transduction pathways are required for hepatic stellate cell (HSC) activation [28]. In this case, IL-6 was shown to induce the expression of both alpha-smooth muscle actin (αSMA) and collagen, type 1, alpha 1 (Col1a) in primary HSCs. Inhibition of either the MAPK or JAK/STAT pathway

resulted in the loss of IL-6-induced αSMA and Col1a mRNA upregulation. The mechanism for crosstalk between the MAPK and JAK/STAT pathways has been shown to occur on multiple transduction levels (e.g., phosphorylation cascade and transcriptional levels), creating a complex network of regulation [29]. Of relevance in the context of this report, both the MAPK and JAK/STAT signaling pathways were shown to be required for the induction of matrix metalloproteinase 1 (MMP-1) expression via the cytokine oncostatin M (OSM) in human astrocytes [30]. An OSM-responsive element (OMRE) was later shown to control the activation of the MMP-1 gene. Interestingly, OMRE contains an AP-1 (a transcription factor downstream of the MAPK pathway) and a STAT binding site, and requires both binding sites to be occupied for MMP-1 transcription.

It is possible that the promoter of hIntL-1 contains response elements specific to both the MAPK and JAK/STAT pathways. The genomic structure of hIntL-1 has been investigated, with multiple putative transcription factor binding sites being identified in the promoter region [31]. Of note is the Sp1 transcription factor binding site which has been implicated in both the MAPK and JAK/STAT pathways. Sp1 was shown to have a synergistic effect on the JAK/STAT pathway via direct binding to the STAT1 [32]. Sp1 is also implicated in the MAPK pathway in that p42/p44 MAPK can directly phosphorylate Sp1, causing its activation [33]. In conclusion, we report that both the MAPK and JAK/STAT signaling pathways are required for the IL-13 induced activation of hIntL-1 in LS174T cells. To our knowledge, this is the first example of an innate immune lectin whose expression is dependent on two separate signaling pathways.

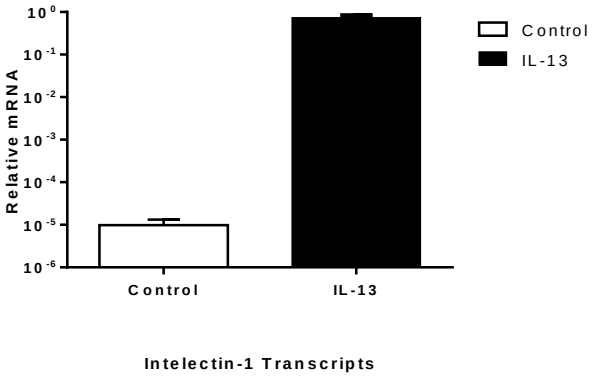
REFERENCES

1. Wynn, T.A., *IL-13 effector functions.* Annu Rev Immunol, 2003. **21**: p. 425-56.
2. Nelms, K., et al., *The IL-4 receptor: signaling mechanisms and biologic functions.* Annu Rev Immunol, 1999. **17**: p. 701-38.
3. Rahaman, S.O., et al., *IL-13R(alpha)2, a decoy receptor for IL-13 acts as an inhibitor of IL-4-dependent signal transduction in glioblastoma cells.* Cancer Res, 2002. **62**(4): p. 1103-9.
4. Chiaramonte, M.G., et al., *Regulation and function of the interleukin 13 receptor alpha 2 during a T helper cell type 2-dominant immune response.* J Exp Med, 2003. **197**(6): p. 687-701.
5. Kawakami, M., et al., *Interleukin-13 receptor alpha2 chain in human head and neck cancer serves as a unique diagnostic marker.* Clin Cancer Res, 2003. **9**(17): p. 6381-8.
6. Joshi, B.H., G.E. Plautz, and R.K. Puri, *Interleukin-13 receptor alpha chain: a novel tumor-associated transmembrane protein in primary explants of human malignant gliomas.* Cancer Res, 2000. **60**(5): p. 1168-72.
7. Fujisawa, T., et al., *A novel role of interleukin-13 receptor alpha2 in pancreatic cancer invasion and metastasis.* Cancer Res, 2009. **69**(22): p. 8678-85.
8. Heller, F., et al., *Interleukin-13 is the key effector Th2 cytokine in ulcerative colitis that affects epithelial tight junctions, apoptosis, and cell restitution.* Gastroenterology, 2005. **129**(2): p. 550-64.
9. Fujisawa, T., B.H. Joshi, and R.K. Puri, *IL-13 regulates cancer invasion and metastasis through IL-13Ralpha2 via ERK/AP-1 pathway in mouse model of human ovarian cancer.* Int J Cancer, 2012. **131**(2): p. 344-56.
10. Kioi, M., et al., *Interleukin-13 receptor alpha2 chain: a potential biomarker and molecular target for ovarian cancer therapy.* Cancer, 2006. **107**(6): p. 1407-18.
11. Cao, H., et al., *IL-13/STAT6 signaling plays a critical role in the epithelial-mesenchymal transition of colorectal cancer cells.* Oncotarget, 2016. **7**(38): p. 61183-61198.
12. David, M., et al., *Induction of the IL-13 receptor alpha2-chain by IL-4 and IL-13 in human keratinocytes: involvement of STAT6, ERK and p38 MAPK pathways.* Oncogene, 2001. **20**(46): p. 6660-8.
13. Zhen, G., et al., *IL-13 and epidermal growth factor receptor have critical but distinct roles in epithelial cell mucin production.* Am J Respir Cell Mol Biol, 2007. **36**(2): p. 244-53.
14. Iwashita, J., et al., *mRNA of MUC2 is stimulated by IL-4, IL-13 or TNF-alpha through a mitogen-activated protein kinase pathway in human colon cancer cells.* Immunol Cell Biol, 2003. **81**(4): p. 275-82.
15. Knowles, M.R. and R.C. Boucher, *Mucus clearance as a primary innate defense mechanism for mammalian airways.* J Clin Invest, 2002. **109**(5): p. 571-7.
16. Lang, T., G.C. Hansson, and T. Samuelsson, *Gel-forming mucins appeared early in metazoan evolution.* Proc Natl Acad Sci U S A, 2007. **104**(41): p. 16209-14.
17. Lee, J.K., et al., *The X-lectins: a new family with homology to the Xenopus laevis oocyte lectin XL-35.* Glycoconj J, 2004. **21**(8-9): p. 443-50.

18. Wesener, D.A., et al., *Recognition of microbial glycans by human intelectin-1.* Nat Struct Mol Biol, 2015. **22**(8): p. 603-10.
19. Voehringer, D., et al., *Nippostrongylus brasiliensis: identification of intelectin-1 and -2 as Stat6-dependent genes expressed in lung and intestine during infection.* Exp Parasitol, 2007. **116**(4): p. 458-66.
20. French, A.T., et al., *The expression of intelectin in sheep goblet cells and upregulation by interleukin-4.* Vet Immunol Immunopathol, 2007. **120**(1-2): p. 41-6.
21. Tsuji, S., et al., *Human intelectin is a novel soluble lectin that recognizes galactofuranose in carbohydrate chains of bacterial cell wall.* J Biol Chem, 2001. **276**(26): p. 23456-63.
22. Tsuji, S., et al., *Differential structure and activity between human and mouse intelectin-1: human intelectin-1 is a disulfide-linked trimer, whereas mouse homologue is a monomer.* Glycobiology, 2007. **17**(10): p. 1045-51.
23. Johansson, M.E., et al., *The inner of the two Muc2 mucin-dependent mucus layers in colon is devoid of bacteria.* Proc Natl Acad Sci U S A, 2008. **105**(39): p. 15064-9.
24. Perrais, M., et al., *Induction of MUC2 and MUC5AC mucins by factors of the epidermal growth factor (EGF) family is mediated by EGF receptor/Ras/Raf/extracellular signal-regulated kinase cascade and Sp1.* J Biol Chem, 2002. **277**(35): p. 32258-67.
25. Datta, R., et al., *Identification of novel genes in intestinal tissue that are regulated after infection with an intestinal nematode parasite.* Infect Immun, 2005. **73**(7): p. 4025-33.
26. Hatayama, H., et al., *The short chain fatty acid, butyrate, stimulates MUC2 mucin production in the human colon cancer cell line, LS174T.* Biochem Biophys Res Commun, 2007. **356**(3): p. 599-603.
27. So, E.Y., et al., *Ras/Erk pathway positively regulates Jak1/STAT6 activity and IL-4 gene expression in Jurkat T cells.* Mol Immunol, 2007. **44**(13): p. 3416-26.
28. Kagan, P., et al., *Both MAPK and STAT3 signal transduction pathways are necessary for IL-6-dependent hepatic stellate cells activation.* PLoS One, 2017. **12**(5): p. e0176173.
29. Heinrich, P.C., et al., *Principles of interleukin (IL)-6-type cytokine signalling and its regulation.* Biochem J, 2003. **374**(Pt 1): p. 1-20.
30. Korzus, E., et al., *The mitogen-activated protein kinase and JAK-STAT signaling pathways are required for an oncostatin M-responsive element-mediated activation of matrix metalloproteinase 1 gene expression.* J Biol Chem, 1997. **272**(2): p. 1188-96.
31. Schaffler, A., et al., *Genomic structure of human omentin, a new adipocytokine expressed in omental adipose tissue.* Biochim Biophys Acta, 2005. **1732**(1-3): p. 96-102.
32. Look, D.C., et al., *Stat1 depends on transcriptional synergy with Sp1.* J Biol Chem, 1995. **270**(51): p. 30264-7.
33. Milanini-Mongiat, J., J. Pouyssegur, and G. Pages, *Identification of two Sp1 phosphorylation sites for p42/p44 mitogen-activated protein kinases: their*

implication in vascular endothelial growth factor gene transcription. J Biol Chem, 2002. **277**(23): p. 20631-9.

(A)



(B)

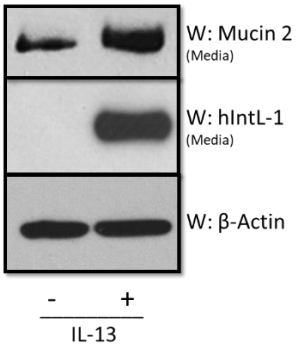


Figure 3.1 IL-13 INDUCES MUC2 AND HINTL-1 SECRETION IN LS174T CELLS. (A) Graph shows hIntL-1 transcripts in response to IL-13 induction in LS174T cells. (B) Western blot data shows hIntL-1 and Muc2 secretion levels from LS174T cells in response to 72 hours of IL-13 treatment.

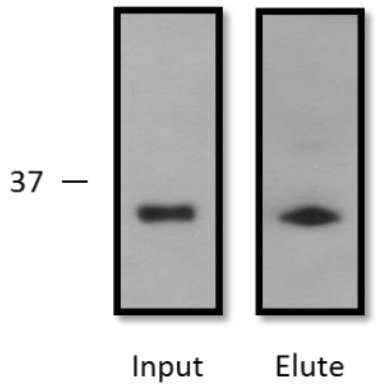
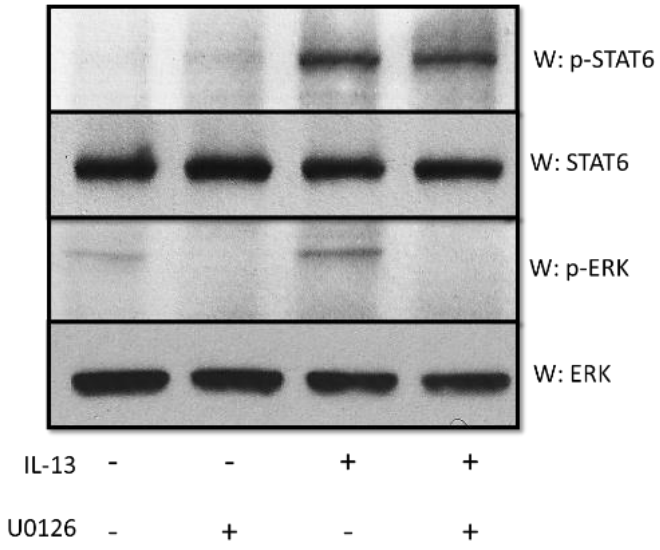


Figure 3.2 HINTL-1 FROM LS174T RETAINS CARBOHYDRATE-BINDING ACTIVITY. LS174T cells were stimulated with IL-13 for 72 hours. Conditioned media was collected and incubated with galactose-agarose in the presence of calcium. After washing, bound hIntL-1 eluted using EDTA.

(A)



(B)

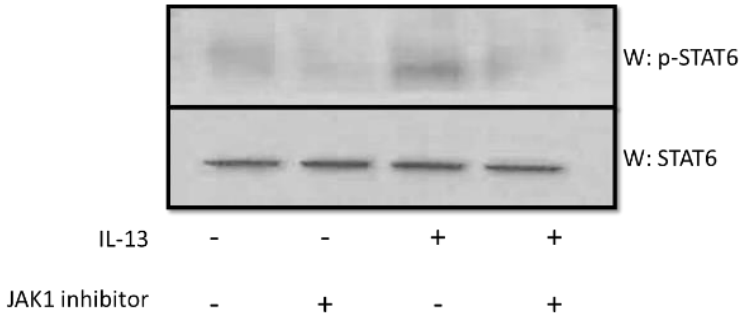
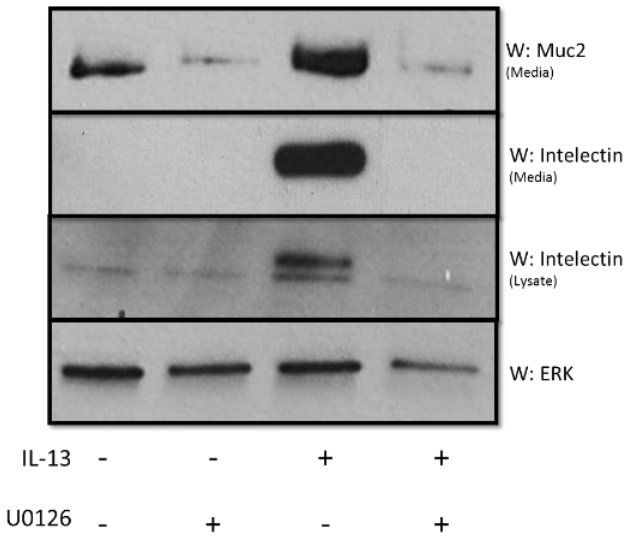


Figure 3.3 IL-13 ACTIVATES BOTH JAK/STAT AND MAPK PATHWAYS IN LS174T. (A) Western blot shows phosphoprotein levels of both STAT6 and ERK following IL-13 stimulation in LS174T cells. (B) Western blot showing inhibiton levels of p-STAT6 following JAK1 inhibition.

(A)



(B)

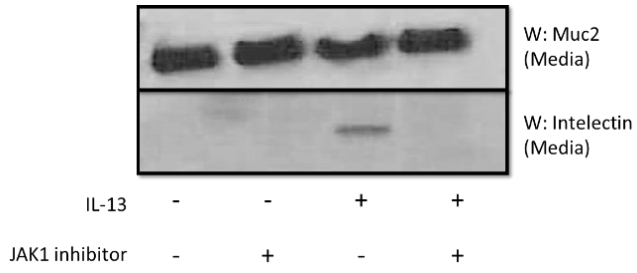


Figure 3.4: HINTL-1 INDUCTION IS CO-DEPENDENT ON BOTH JAK/STAT AND MAPK SIGNALING PATHWAYS. (A) Western blot of hIntL-1 and Muc2 levels following MAPK knockdown and IL-13 stimulation. Cells were pretreated with U0126 and then stimulated for 72 hours with IL-13. (B) Western blot of hIntL-1 following JAK/STAT knockdown and IL-13 stimulation. Cells were pretreated with JAK1 inhibitor and then stimulated for 72 hours with IL-13.

CHAPTER 4

DISCUSSION AND FUTURE PERSPECTIVES

The innate immune system is the first line of defense a host has against pathogen invasion and infection. This system acts on multiple levels, with functions ranging from a simplistic anatomical barrier such as the epidermis to more complex mechanisms that regulate immune cell recruitment and activation of both the complement cascade and adaptive immune system. Understanding the mechanisms within the innate immune system may provide valuable information that can be harnessed to develop novel strategies for combating infections. In this dissertation, the signaling pathways required for the activation of an innate immune lectin, hIntL-1, was determined. Second, X-type lectins from both human and *Xenopus laevis* were analyzed for their microbial glycan-binding properties. Molecular dynamics strategies were then used to validate and rationalize glycan-binding differences between the two human intelectins. In this concluding chapter, future goals stemming from the results of this dissertation as well as the subsequent challenges that will need to be addressed are discussed.

IL-13 stimulation causes cell-specific upregulation of its receptors. In normal human bronchial cells and human keratinocytes, IL-13 induces upregulation of IL-13Rα2, whereas IL-13 induces neither IL-13Rα1 nor IL-13Rα2 in A549 and BEAS-2B airway cell lines [1, 2]. Currently, it is not known if IL-13 affects the expression or externalization of IL-13Rα1 and/or IL-13Rα2 in goblet cells. To address the impact of

receptor regulation on IL-13 function, we can utilize RT-PCR and FACS analysis to quantify the upregulation and cell surface expression of these two receptors in response to stimulation. Moreover, separate knockdown of IL-13Rα1 and IL-13Rα2 using neutralizing antibodies can address what impact, if any, these receptors play in promoting a forward-feedback loop on their expression.

As is described in Chapter 3, hIntL-1 induction required both active MAPK and JAK/STAT signaling pathways. Pharmacologically inhibiting either pathway prevented the secretion of hIntL-1 in IL-13-stimulated LS174T cells. Analysis of the promoter and enhancer regions for hIntL-1 yielded multiple transcription factor binding sites, including Sp1, STAT1, and STAT3 [3]. To further probe the requirement of specific binding elements in hIntL-1 activation, we can develop a luciferase reporter gene. A construct of the 5'-adjacent region of hIntL-1 will be inserted into the luciferase reporter plasmid. Transfection of human goblet cells with the reporter gene and subsequent stimulation with IL-13 should result in high luciferase activity. Once confirmed, mutant constructs can be generated, each with point mutations in STAT or Sp1 response elements so that their respective transcription factors are unable to bind. Transfection with these mutant constructs will help define which binding elements are required for maximal hIntL-1 transcription.

Engineering lectins to exploit their biological functions is well described in the literature [4-6]. Gal-1 has immunomodulatory functions in both the adaptive and innate immune system and as such, is a target for cancer therapy [7]. Specifically, Gal-1 regulates the activation, cytokine secretion, and apoptosis of T-cells as well as enhances the migration of dendritic cells. Promising studies have used protein engineering to

prevent the inactivation of Gal-1 due to covalent dimerization of its cysteine residues. The cysteine residues of Gal-1 were substituted into serine residues and the mutated Gal-1 showed similar glycan-binding specifities as wildtype Gal-1. Importantly, this mutated Gal-1 showed full hemagglutination activity after 400 days of storage, in contrast to wildtype Gal-1 which lost activity after 10 days of storage [8]. The valency of Gal-1 can also be manipulated to increase therapeutic potential. Gal-1 contains a single CRD domains and exists as a monomer under low concentrations. Under high concentrations, Gal-1 forms noncovalent dimers with increased glycan-binding affinity but not affecting glycan selectivity. Stemming from this observation, leucine-zipper based stable Gal-1 homodimers were synthesized [9]. This recombinant Gal-1 homodimer induced T-cell apoptosis at a 20-fold reduced minimal effective concentration compared to wildtype Gal-1. Furthermore, recombinant Gal-1 induced IL-10 at a 100-fold lower minimum effective dose than wildtype Gal-1. These findings highlight the potential for engineering recombinant lectins for therapeutic purposes.

Similar to the work done with Gal-1, the intelectins can also be modified to exploit their glycan-binding properties. As discussed in Chapter 2, hIntL-1 failed to interact with a set of ~300 different glycans commonly found in mammals, suggesting it does not bind mammalian glycans. Interrogation of the MGM with hIntL-1 showed interactions with several different microbial glycans isolated from pathogenic bacteria. Using this information, one can generate a recombinant protein containing the CRD domain of hIntL-1 attached to cargo via a linker group. This cargo may be a small peptide that activates the complement pathway. A likely candidate would be the collagen-like domain present in both collectins and ficolins which lead to MASP binding and

subsequent complement activation. Additionally, other antibacterial drugs may be tethered to the CRD domain of hIntL-1 for efficient delivery. Administrating this recombinant protein in a clinical setting may help treat infections caused by the pathogens that hIntL-1 binds.

Although the amino acid sequences of the X-type lectin family are fairly conserved between species, key ligand-binding residues are substituted within their CRD domain (**Figure 2.9)**. These substitutions are likely to be involved in the differential glycan-binding properties observed on the MGM. Interrogation of the MGM with XCGL-1 resulted in interactions with the majority of different *S. pneumoniae* microbial glycans present on the array. In contrast to hIntL-1 and -2, XCGL-1 interacted with the glycans from *S. pneumoniae* at a much lower relative affinity (100% vs 30-60% highest binding). The structure of XCGL-1 is currently not known, but a homology model can be generated based on the previously solved structure of XEEL. Molecular dynamics and MM-GBSA analysis can be employed with XCGL-1 to calculate which residues of XCGL-1 contribute to the total binding energy of XCGL-1 ligands. With this information, one may work to generate a recombinant protein that combines the CRD domain elements of hIntL-1 that give high affinity with the CRD domain elements of XCGL-1 that give broad specificity for *S. pneumoniae* compounds.

Of the ~300 unique microbial glycans displayed on the MGM, 23 glycans represented are a subset of the different serotypes of *S. pneumoniae* [10]. A total of 93 different serotypes of *S. pneumoniae* have been described, with significant focus on each unique capsular polysaccharide as this is the most important virulence factor of the microbe [11]. A strong correlation exists between specific serotypes of *S. pneumoniae*

and infection rates which is why vaccines against the organism generally target a specific set of strains. For example, the 7-valent pneumococcal conjugate vaccine that targets pneumococcal serotypes 4, 6B, 9V, 14, 18C, 19F, and 23F has been successful in children with high-risk conditions [12]. These serotypes are suggested to be responsible for 80% of *S. pneumoniae* infections in children living in developed countries [13]. Interestingly, not all serotypes expressing the same capsular polysaccharide maintain the same virulence. This is likely due to differences in specific elements of the of *S. pneumoniae* virulome [14].

Typically, colonization of *S. pneumoniae* occurs by only one serotype at a given time, although cases do exists where multiple serotypes are present [15]. Considering that the pathogenicity of *S. pneumoniae* is tied to specific serotypes, a need for the identification of specific serotypes exists. Serotyping *S. pneumoniae* include the older methods involving the Quellung reaction and Neufeld test to the more contemporary methods involving dot blot assays, latex agglutination, and multiplex PCR [16-19]. Each method has their own inherent advantages and pitfalls which is reviewed by [20] and [21].

Because hIntL-1 and -2 selectively bind to *S. pneumoniae* serotypes containing β-Gal*f* and GroP (**Figure 2.7**), one could design an assay which utilizes the human intelectins to detect the serotypes of *S. pneumoniae* that it binds. The human intelectin can be generated recombinantly with a tag for downstream detection, or can be detected using monoclonal and polyclonal antibodies already developed. Diagnosing the serotype(s) present or absent in a clinical isolate may be useful in determining a specific treatment regimen.

hIntL-1 coordinates a protein-bound calcium ion to interact with an acyclic 1,2-terminal diol group present on microbial glycans. An acyclic 1,2-terminal diol group is not unique to bacterial glycans, as it is also present on the mammalian monosaccharide sialic acid. Computational modeling of α-Neu5Ac, a common form of sialic acid found in humans, in the binding site of hIntL-1 showed an unfavorable interaction due to two reasons [22]. First, an anion-anion repulsion was observed between the carboxylate group of α-Neu5Ac and Glu274 of hIntL-1. Second, steric hinderance between the *N*-acetyl group of α-Neu5Ac with the protein surface of hIntL-1 also destabilized binding. Sequence alignment of hIntL-1 with hIntL-2 (**Figure 2.9**) shows a substitution at position 274 from a glutamic acid to a glutamine. This substitution would potentially decrease the anion-anion repulsion reported in α-Neu5Ac with hIntL-1, and thereby possibly allow hIntL-2 to interact with α-Neu5Ac, and thus mammalian glycans.

We are currently determining the ability for hIntL-2 to bind α-Neu5Ac. We have sent the CFG purified recombinant hIntL-2 for analysis on the mammalian glycan array which contains this monosaccharide in various linkages. If binding to sialic acid is observed, we can further analyze what resides are critical for these glycan-binding properties by making point mutations in hIntL-1 and/or hIntL-2 based on **Figure 2.9**. In doing so, we may be able to interchange the glycan selectivity between hIntL-1 and -2. It is important to consider that by making point mutations within the CRD domain of the intelectins, a loss or change in glycan-binding may occur. Considering we purify the intelectins using galactose-affinity chromatography, it would be worthwhile to include a tag on the N-terminal region of the lectins for other methods of purification.

Intelectins from *xenopus* and humans have been purified using various types of affinity chromatography. We have purified XCGL-1 using melibiose affinity chromatography (a repeating disaccharide of α-linked galactose and glucose), and hIntL-1 and -2 using galactose-affinity chromatography [23] . Other groups successfully purified hIntL-1 and mIntL-1 using β-Gal*f* affinity chromatography [22]. Additionally, our group can purify *xenopus*, human, and rainbow trout intelectins in a calcium-dependent manner using underivatized agarose. Supporting this observation, XEEL was reported to also contain an affinity for agarose [24]. Agarose is linear polysaccharide composed of D-galactose and 3,6-anhydro-L-galactopyranose linked by α-(1,3) and β-(1-4). Intriguingly, agarose does not contain an acyclic 1,2-terminal diol, suggesting there are other binding mechanisms utilized by the intelectins. To address this, we can use computational docking techniques to answer the question of how the intelectins interact with agarose. It will be important to determine if intelectins require a specific length of the agarose polysaccharide to engage in binding, similar to how RegIII binding is dependent on the length of the peptidoglycan chain [25].

This work has determined the signaling pathways required for hIntL-1 secretion in goblet cells via IL-13 stimulation. Furthermore, we have analyzed the microbial-glycan binding properties of three intelectins from human and *xenopus*. Taken together, these results further implicate X-type lectin members in the immune response with important properties relevant to human pathogens.

REFERENCES

1. Zheng, T., et al., *Cytokine regulation of IL-13Ralpha2 and IL-13Ralpha1 in vivo and in vitro.* J Allergy Clin Immunol, 2003. **111**(4): p. 720-8.
2. Yasunaga, S., et al., *The negative-feedback regulation of the IL-13 signal by the IL-13 receptor alpha2 chain in bronchial epithelial cells.* Cytokine, 2003. **24**(6): p. 293-303.
3. Schaffler, A., et al., *Genomic structure of human omentin, a new adipocytokine expressed in omental adipose tissue.* Biochim Biophys Acta, 2005. **1732**(1-3): p. 96-102.
4. Hu, D., H. Tateno, and J. Hirabayashi, *Directed evolution of lectins by an improved error-prone PCR and ribosome display method.* Methods Mol Biol, 2014. **1200**: p. 527-38.
5. Hu, D., H. Tateno, and J. Hirabayashi, *Lectin engineering, a molecular evolutionary approach to expanding the lectin utilities.* Molecules, 2015. **20**(5): p. 7637-56.
6. Farhadi, S.A. and G.A. Hudalla, *Engineering galectin-glycan interactions for immunotherapy and immunomodulation.* Exp Biol Med (Maywood), 2016. **241**(10): p. 1074-83.
7. Rabinovich, G.A. and M.A. Toscano, *Turning 'sweet' on immunity: galectin-glycan interactions in immune tolerance and inflammation.* Nat Rev Immunol, 2009. **9**(5): p. 338-52.
8. Nishi, N., et al., *Functional and structural bases of a cysteine-less mutant as a long-lasting substitute for galectin-1.* Glycobiology, 2008. **18**(12): p. 1065-73.
9. van der Leij, J., et al., *Strongly enhanced IL-10 production using stable galectin-1 homodimers.* Mol Immunol, 2007. **44**(4): p. 506-13.
10. Stowell, S.R., et al., *Microbial glycan microarrays define key features of host-microbial interactions.* Nat Chem Biol, 2014. **10**(6): p. 470-6.
11. Kadioglu, A., et al., *The role of Streptococcus pneumoniae virulence factors in host respiratory colonization and disease.* Nat Rev Microbiol, 2008. **6**(4): p. 288-301.
12. Millar, E.V., et al., *Indirect effect of 7-valent pneumococcal conjugate vaccine on pneumococcal colonization among unvaccinated household members.* Clin Infect Dis, 2008. **47**(8): p. 989-96.
13. van der Poll, T. and S.M. Opal, *Pathogenesis, treatment, and prevention of pneumococcal pneumonia.* Lancet, 2009. **374**(9700): p. 1543-56.
14. Silva, N.A., et al., *Genomic diversity between strains of the same serotype and multilocus sequence type among pneumococcal clinical isolates.* Infect Immun, 2006. **74**(6): p. 3513-8.
15. Hare, K.M., et al., *Random colony selection versus colony morphology for detection of multiple pneumococcal serotypes in nasopharyngeal swabs.* Pediatr Infect Dis J, 2008. **27**(2): p. 178-80.
16. Austrian, R., *The quellung reaction, a neglected microbiologic technique.* Mt Sinai J Med, 1976. **43**(6): p. 699-709.
17. Fenoll, A., et al., *Dot blot assay for the serotyping of pneumococci.* J Clin Microbiol, 1997. **35**(3): p. 764-6.

18. Slotved, H.C., et al., *Simple, rapid latex agglutination test for serotyping of pneumococci (Pneumotest-Latex).* J Clin Microbiol, 2004. **42**(6): p. 2518-22.
19. Pai, R., R.E. Gertz, and B. Beall, *Sequential multiplex PCR approach for determining capsular serotypes of Streptococcus pneumoniae isolates.* J Clin Microbiol, 2006. **44**(1): p. 124-31.
20. Rayner, R.E., et al., *Genotyping Streptococcus pneumoniae.* Future Microbiol, 2015. **10**(4): p. 653-64.
21. Donkor, E.S., *Molecular typing of the pneumococcus and its application in epidemiology in sub-Saharan Africa.* Front Cell Infect Microbiol, 2013. **3**: p. 12.
22. Wesener, D.A., et al., *Recognition of microbial glycans by human intelectin-1.* Nat Struct Mol Biol, 2015. **22**(8): p. 603-10.
23. Lee, J.K., et al., *Cloning and expression of a Xenopus laevis oocyte lectin and characterization of its mRNA levels during early development.* Glycobiology, 1997. **7**(3): p. 367-72.
24. Wangkanont, K., et al., *Structures of Xenopus Embryonic Epidermal Lectin Reveal a Conserved Mechanism of Microbial Glycan Recognition.* J Biol Chem, 2016. **291**(11): p. 5596-610.
25. Mukherjee, S. and L.V. Hooper, *Antimicrobial defense of the intestine.* Immunity, 2015. **42**(1): p. 28-39.

APPENDIX: SUPPLEMENTAL TABLES

Appendix Table 1: FULL CFG ARRAY DATA FOR HINTL-1

Chart #	Glycan	Average	STDEV	SEM
3	Ceruloplasmine	1494	232	1296
5	Transferrin	1392	343	169
0	AGP	840	27	5952
1	AGP-A	702	308	2590
185	GlcAβ1-6Galβ-Sp8	615	283	60
46	[6OSO3]GlcNAcβ–Sp8	538	96	132
47	9-O-AcNeu5NAcα-Sp8	523	137	67
35	[3OSO3]Galβ1-4GlcNAcβ–Sp0	491	35	74
2	AGP-βl	489	94	1172
27	[3OSO3]Galβ1-4Glcβ-Sp8	489	146	119
28	[3OSO3]Galβ1-4(6OSO3)Glcβ–Sp0	469	43	88
20	β-GlcNAc–Sp0	458	68	37
19	β-GalNAc–Sp8	453	119	91
200	Neu5Acα2-3(Galβ1-3GalNAcβ1-4)Galβ1-4Glcβ-Sp0	452	178	33
78	GalNAcα1-3(Fucα1-2)Galβ1-3GlcNAcβ-Sp0	446	43	25
25	[3OSO3][6OSO3]Galβ1-4[6OSO3]GlcNAcβ-Sp0	437	139	73
6	α-D-Gal–Sp8	413	149	188
34	[3OSO3]Galβ1-4[6OSO3]GlcNAcβ-Sp8	411	164	40
29	[3OSO3]Galβ1-4(6OSO3)Glcβ–Sp8	400	125	78
75	Fucα1-3GlcNAcβ-Sp8	400	27	82
9	α-GalNAc–Sp8	395	80	701
183	GlcAβ-Sp8	395	308	35
16	β-D-Gal–Sp8	394	223	150
22	β-GlcN(Gc)-Sp8	388	122	257
32	[3OSO3]Galβ1-3GlcNAcβ–Sp8	388	77	91
13	α-Neu5Ac–Sp8	387	229	142
189	Manα1-2Manα1-3(Manα1-2Manα1-6)Manα-Sp9	385	111	38
40	6-H2PO3Manα–Sp8	384	97	46
71	Fucα1-2Galβ1-4GlcNAcβ–Sp8	380	99	96
104	Galα1-3Galβ1-4GlcNAcβ–Sp8	376	85	39
139	Galβ1-4[6OSO3]Glcβ–Sp8	373	160	81
36	[3OSO3]Galβ1-4GlcNAcβ-Sp8	371	146	62
38	[4OSO3][6OSO3]Galβ1-4GlcNAcβ-Sp0	368	193	50
66	Fucα1-2Galβ1-4(Fucα1-3)GlcNAcβ–Sp0	367	78	49
70	Fucα1-2Galβ1-4GlcNAcβ–Sp0	365	55	77
30	[3OSO3]Galβ1-3(Fucα1-4)GlcNAcβ–Sp8	362	42	113

232	Neu5Acα2-3Galβ1-4(Fucα1-3)GlcNAcβ1-3Galβ1-4GlcNAcβ-Sp8	355	45	31
181	Sorbitol-Sp8	348	183	86
221	Neu5Acα2-3Galβ-Sp8	345	30	62
128	Galβ1-3GalNAcβ1-4Galβ1-4Glcβ–Sp8	344	141	60
69	Fucα1-2Galβ1-4GlcNAcβ1-3Galβ1-4GlcNAcβ1-3Galβ1-4GlcNAcβ-Sp0	342	89	46
85	GalNAcα1-3Galβ–Sp8	340	127	30
65	Fucα1-2Galβ1-4(Fucα1-3)GlcNAcβ1-3Galβ1-4(Fucα1-3)GlcNAcβ1-3Galβ1-4(Fucα1-3)GlcNAcβ-Sp0	336	69	45
264	[3OSO3]Galβ1-4(Fucα1-3)(6OSO3)Glc-Sp0	335	60	60
113	Galβ1-2Galβ–Sp8	334	134	49
48	9-O-AcNeu5NAcα2-6Galβ1-4GlcNAcβ-Sp8	331	82	71
63	Fucα1-2Galβ1-3GlcNAcβ–Sp8	331	94	60
138	Galβ1-4[6OSO3]Glcβ–Sp0	328	193	66
260	Neu5Gcα2-3Galβ1-4Glcβ–Sp0	326	126	23
263	Neu5Gcα–Sp8	326	309	70
105	Galα1-3Galβ1-4Glcβ–Sp0	321	40	64
79	GalNAcα1-3(Fucα1-2)Galβ1-4(Fucα1-3)GlcNAcβ-Sp0	318	126	47
15	β-Neu5Ac-Sp8	317	69	55
17	β-D-Glc–Sp8	317	47	56
162	GlcNAcβ1-3Galβ1-3GalNAcα-Sp8	317	64	49
11	α-L-Fuc–Sp9	313	52	49
10	α-L-Fuc–Sp8	312	35	222
196	Manα1-6(Manα1-3)Manα1-6(Manα2Manα1-3)Manβ1-4GlcNAcβ1-4GlcNAcβ-N	312	126	53
73	Fucα1-2Galβ–Sp8	311	56	29
87	GalNAcβ1-3GalNAca–Sp8	305	83	89
67	Fucα1-2Galβ1-4(Fucα1-3)GlcNAcβ–Sp8	302	71	59
58	Fucα1-2Galβ1-3GalNAcβ1-4(Neu5Acα2-3)Galβ1-4Glcβ-Sp0	301	55	54
111	Galα1-4GlcNAcβ–Sp8	300	167	66
53	Neu5Acα2-6Galβ1-4GlcNAcβ1-2Manα1-3(Neu5Acα2-6Galβ1-4GlcNAcβ1-2Manα1-6)Manβ1-4GlcNAcβ1-4GlcNAcβ-Sp8	299	73	45
100	Galα1-3GalNAcα-Sp8	299	31	53
41	[6OSO3]Galβ1-4Glcβ–Sp0	297	94	63
33	[3OSO3]Galβ1-4(Fucα1-3)GlcNAcβ–Sp8	296	107	36
72	Fucα1-2Galβ1-4Glcβ–Sp0	295	145	33
26	[3OSO3][6OSO3]Galβ1-4GlcNAcβ-Sp0	293	164	53

127	Galβ1-3GalNAcβ1-4(Neu5Acα2-3)Galβ1-4Glcβ-Sp0	293	67	53
8	α-D-Man–Sp8	292	217	139
84	GalNAcα1-3GalNAcβ–Sp8	292	107	76
169	GlcNAcβ1-4Galβ1-4GlcNAcβ-Sp8	292	59	53
21	β-GlcNAc–Sp8	289	81	146
198	Man5_9mix N	289	68	45
92	GalNAcβ1-4GlcNAcβ–Sp8	286	109	73
217	NeuAcα2-3Galβ1-3(Fucα1-4)GlcNAcβ1-3Galβ1-4(Fucα1-3)GlcNAcβ Sp0	286	102	40
117	Galβ1-3(Fucα1-4)GlcNAc–Sp8	285	54	39
4	Fibrinogen	284	133	1147
43	[6OSO3]Galβ1-4GlcNAcβ–Sp8	284	187	80
74	Fucα1-3GlcNAcβ-Sp8	284	238	37
152	Galβ1-4GlcNAcβ–Sp8	283	153	67
54	Fucα1-2Galβ1-3GalNAcβ1-3Galα-Sp9	282	46	64
220	Neu5Acα2-3Galβ1-3(Neu5Acα2-6)GalNAcα–Sp8	282	136	25
42	[6OSO3]Galβ1-4Glcβ–Sp8	280	160	108
124	Galβ1-3GalNAcα-Sp8	278	115	43
108	Galα1-4Galβ1-4GlcNAcβ–Sp0	277	64	63
132	Galβ1-3GlcNAcβ–Sp0	276	87	56
250	Neu5Acα2-6Galβ–Sp8	276	111	41
192	Manα1-2Manα1-6(Manα1-3)Manα1-6(Manα2Manα2Manα1-3)Manβ1-4GlcNAcβ1-4GlcNAcβ-N	275	51	31
136	Galβ1-4(Fucα1-3)GlcNAcβ1-4Galβ1-4(Fucα1-3)GlcNAcβ-Sp0	274	23	66
133	Galβ1-3GlcNAcβ–Sp8	272	116	78
59	Fucα1-2Galβ1-3GalNAcβ1-4(Neu5Acα2-3)Galβ1-4Glcβ-Sp9	270	66	44
89	GalNAcβ1-3Galα1-4Galβ1-4GlcNAcβ-Sp0	270	103	84
191	Manα1-6(Manα1-2Manα1-3)Manα1-6(Manα2Manα1-3)Manβ1-4GlcNAcβ1-4GlcNAcβ-N	269	166	52
12	α-L-Rhα–Sp8	264	94	106
163	GlcNAcβ1-3Galβ1-4GlcNAcβ–Sp0	263	86	45
39	[4OSO3]Galβ1-4GlcNAcβ-Sp8	262	83	49
109	Galα1-4Galβ1-4GlcNAcβ–Sp8	260	80	36
44	[6OSO3]Galβ1-4[6OSO3]Glcβ-Sp8	259	95	73
214	Neu5Acα2-3GalNAcβ1-4GlcNAcβ-Sp0	259	90	30
14	α-Neu5Ac–Sp11	258	205	54
107	Galα1-4(Fucα1-2)Galβ1-4GlcNAcβ-Sp8	258	115	62
194	Manα1-3(Manα1-6)Manα–Sp9	256	67	59

64	Fucα1-2Galβ1-4(Fucα1-3)GlcNAcβ1-3Galβ1-4(Fucα1-3)GlcNAcβ-Sp0	254	55	49
37	[3OSO3]Galβ–Sp8	251	142	99
91	GalNAcβ1-4GlcNAcβ–Sp0	251	101	66
101	Galα1-3GalNAcβ–Sp8	251	108	18
241	Neu5Acα2-6GalNAcα–Sp8	251	48	18
231	Neu5Acα2-3Galβ1-4(Fucα1-3)GlcNAcβ1-3Galβ-Sp8	247	293	66
167	GlcNAcβ1-4MDPLys	245	50	74
240	Neu5Acα2-6(Galβ1-3)GalNAcα–Sp8	245	55	36
31	[3OSO3]Galβ1-3GalNAcα–Sp8	244	53	98
52	Neu5Acα2-6Galβ1-4GlcNAcβ1-2Manα1-3(Neu5Acα2-6Galβ1-4GlcNAcβ1-2Manα1-6)Manβ1-4GlcNAcβ1-4GlcNAcβ-Gly	244	71	77
225	Neu5Acα2-3Galβ1-3GlcNAcβ–Sp8	244	135	23
96	Galα1-3(Fucα1-2)Galβ1-4GlcNAc-Sp0	242	78	84
224	Neu5Acα2-3Galβ1-3GlcNAcβ–Sp0	241	55	50
50	GlcNAcβ1-2Manα1-3(GlcNAcβ1-2Manα1-6)Manβ1-4GlcNAcβ1-4GlcNAcβ-Gly	240	62	165
57	Fucα1-2Galβ1-3GalNAcα–Sp8	240	89	58
172	GlcNAcβ1-4GlcNAcβ1-4GlcNAcβ–Sp8	240	136	86
99	Galα1-3(Galα1-4)Galβ1-4GlcNAcβ-Sp8	239	109	84
95	Galα1-3(Fucα1-2)Galβ1-4(Fucα1-3)GlcNAcβ-Sp0	238	62	20
188	Manα1-2Manα1-2Manα1-3Manα-Sp9	236	84	39
174	GlcNAcβ1-6GalNAcα–Sp8	235	47	29
98	Galα1-3(Fucα1-2)Galβ–Sp8	233	58	75
141	Galβ1-4GalNAcβ1-3(Fucα1-2)Galβ1-4GlcNAcβ-Sp8	232	56	67
62	Fucα1-2Galβ1-3GlcNAcβ–Sp0	229	94	81
159	GlcNAcβ1-3(GlcNAcβ1-6)Galβ1-4GlcNAcβ–Sp8	227	9	60
115	Galβ1-3(Fucα1-4)GlcNAcβ1-3Galβ1-4GlcNAcβ-Sp0	226	63	71
150	Galβ1-4GlcNAcβ1-6GalNAcα–Sp8	226	132	98
161	GlcNAcβ1-3Galβ-Sp8	226	81	39
190	Manα1-2Manα1-3Manα-Sp9	224	71	35
156	GlcNAcα1-6Galβ1-4GlcNAcβ-Sp8	223	65	31
112	Galα1-6Glcβ-Sp8	221	62	71
129	Galβ1-3Galβ–Sp8	219	27	58
94	Galα1-3(Fucα1-2)Galβ1-3GlcNAcβ-Sp0	217	42	46
61	Fucα1-2Galβ1-3GlcNAcβ1-3Galβ1-4Glcβ–Sp8	216	62	91
106	Galα1-3Galβ–Sp8	216	84	40
266	[3OSO3]Galβ1-4[Fucα1-3][6OSO3]GlcNAc-Sp8	216	83	34
110	Galα1-4Galβ1-4Glcβ–Sp0	215	77	81
158	GlcNAcβ1-3(GlcNAcβ1-6)GalNAcα–Sp8	215	60	75

168	GlcNAcβ1-4(GlcNAcβ1-6)GalNAcα-Sp8	212	90	97
262	Neu5Gcα2-6Galβ1-4GlcNAcβ–Sp0	212	144	52
131	Galβ1-3GlcNAcβ1-3Galβ1-4Glcβ–Sp10	211	60	34
86	GalNAcα1-4(Fucα1-2)Galβ1-4GlcNAcβ-Sp8	210	99	82
102	Galα1-3Galβ1-4(Fucα1-3)GlcNAcβ–Sp8	210	91	77
160	GlcNAcβ1-3GalNAcα–Sp8	210	25	50
259	Neu5Gcα2-3Galβ1-4GlcNAcβ–Sp0	209	72	53
175	GlcNAcβ1-6Galβ1-4GlcNAcβ-Sp8	207	103	49
233	Neu5Acα2-3Galβ1-4GlcNAcβ1-3Galβ1-4(Fucα1-3)GlcNAc-Sp0	203	80	32
24	GlcNAcβ1-3(GlcNAcβ1-4)(GlcNAcβ1-6)GlcNAc-Sp8	201	59	65
51	Galβ1-4GlcNAcβ1-2Manα1-3(Galβ1-4GlcNAcβ1-2Manα1-6)Manβ1-4GlcNAcβ1-4GlcNAcβ-Gly	201	59	92
212	Neu5Acα2-3(Neu5Acα2-6)GalNAcα–Sp8	201	117	47
272	Fucα1-2-Galβ1-4[6OSO3]Glc-Sp0	198	66	31
7	α-D-Glc–Sp8	197	52	164
45	NeuAcα2-3[6OSO3]Galβ1-4GlcNAcβ–Sp8	197	63	85
82	GalNAcα1-3(Fucα1-2)Galβ1-4Glcβ-Sp0	197	111	55
135	Galβ1-4(Fucα1-3)GlcNAcβ–Sp8	197	18	55
210	Neu5Acα2-3(GalNAcβ1-4)Galβ1-4Glcβ–Sp0	196	72	45
81	GalNAcα1-3(Fucα1-2)Galβ1-4GlcNAcβ–Sp8	194	58	57
143	Galβ1-4GlcNAcβ1-3GalNAcα–Sp8	193	77	67
187	KDNα2-3Galβ1-4GlcNAcβ–Sp0	193	95	62
252	Neu5Acα2-8Neu5Acα2-3Galβ1-4Glcβ–Sp0	193	99	20
49	Manα1-3(Manα1-6)Manβ1-4GlcNAcβ1-4GlcNAcβ-Gly	192	134	100
118	Galβ1-3(Fucα1-4)GlcNAcβ–Sp8	190	39	33
146	Galβ1-4GlcNAcβ1-3Galβ1-4GlcNAcβ–Sp0	190	94	64
165	GlcNAcβ1-3Galβ1-4GlcNAcβ1-3Galβ1-4GlcNAcβ-Sp0	190	55	80
275	Galβ1-3(GlcNacβ1-6)GalNAc-T	190	63	57
201	Neu5Acα2-3Galβ1-3GalNAcα-Sp8	187	82	69
148	Galβ1-4GlcNAcβ1-3Galβ1-4Glcβ–Sp8	184	87	55
76	Fucα1-4GlcNAcβ–Sp8	182	59	74
193	Manα1-2Manα1-2Manα1-3(Manα1-2Manα1-3(Manα1-2Manα1-6)Manα1-6)Manβ1-4GlcNAcβ1-4GlcNAcβ-N	181	89	38
55	Fucα1-2Galβ1-3GalNAcβ1-3Galα1-4Galβ1-4Glcβ-Sp9	180	93	185
83	GalNAcα1-3(Fucα1-2)Galβ–Sp8	178	76	87
257	Neu5Gcα2-3Galβ1-3GlcNAcβ-Sp0	178	136	25
230	Neu5Acα2-3Galβ1-4(Fucα1-3)GlcNAcβ–Sp8	177	73	41

234	Neu5Acα2-3Galβ1-4GlcNAcβ1-3Galβ1-4GlcNAcβ1-3Galβ1-4GlcNAcβ–Sp0	177	61	35
121	Galβ1-3(Neu5Acα2-6)GalNAcα-Sp8	176	44	41
125	Galβ1-3GalNAcβ–Sp8	176	176	47
176	Glcα1-4Glcβ–Sp8	176	85	65
151	Galβ1-4GlcNAcβ–Sp0	174	75	29
216	Neu5Acα2-3Galβ1-3(Fucα1-4)GlcNAcβ–Sp8	174	67	51
155	GlcNAcα1-3Galβ1-4GlcNAcβ-Sp8	173	131	30
114	Galβ1-3(Fucα1-4)GlcNAcβ1-3Galβ1-4(Fucα1-3)GlcNAcβ-Sp0	172	84	48
179	Glcβ1-4Glcβ-Sp8	172	70	77
103	Galα1-3Galβ1-3GlcNAcβ-Sp0	171	91	82
93	Galα1-2Galβ–Sp8	170	75	86
243	Neu5Acα2-6Galβ1-4[6OSO3]GlcNAcβ-Sp8	170	110	48
80	GalNAcα1-3(Fucα1-2)Galβ1-4GlcNAcβ-Sp0	169	94	43
130	Galβ1-3GlcNAcβ1-3Galβ1-4GlcNAcβ-Sp0	169	80	23
254	Neu5Acβ2-6Galβ1-4GlcNAcβ-Sp8	167	100	25
223	NeuAcα2-3Galβ1-3GlcNAcβ1-3Galβ1-4GlcNAcβ-Sp0	166	63	72
197	Manα1-6(Manα1-3)Manα1-6(Manα1-3)Manβ1-4GlcNAcβ1-4 GlcNAcβ-N	165	34	30
202	NeuAcα2-8NeuAcα2-8NeuAcα2-8NeuAcα2-3(GalNAcβ1-4)Galβ1-4Glcβ-Sp0	165	42	33
273	Galβ1-3(Fucα1-4)GlcNAcβ1-3Galβ1-3(Fucα1-4)GlcNAcβ-Sp0	165	148	30
123	Galβ1-3(Neu5Acα2-6)GlcNAcβ1-4Galβ1-4Glcβ-Sp10	164	64	46
218	Neu5Acα2-3Galβ1-3(Neu5Acα2-3Galβ1-4)GlcNAcβ-Sp8	159	70	33
227	Neu5Acα2-3Galβ1-4(Fucα1-3)(6OSO3) GlcNAcβ –Sp8	157	83	37
208	Neu5Acα2-3(GalNAcβ1-4)Galβ1-4GlcNAcβ-Sp0	155	107	40
147	Galβ1-4GlcNAcβ1-3Galβ1-4Glcβ–Sp0	154	77	50
164	GlcNAcβ1-3Galβ1-4GlcNAcβ-Sp8	153	39	78
184	GlcAβ1-3Galβ-Sp8	153	48	38
178	Glcα1-6Glcα1-6Glcβ-Sp8	152	65	39
246	Neu5Acα2-6Galβ1-4GlcNAcβ1-3Galβ1-4(Fucα1-3)GlcNAcβ1-3Galβ1-4(Fucα1-3)GlcNAcβ-Sp0	152	24	38
142	Galβ1-4GlcNAcβ1-3(Galβ1-4GlcNAcβ1-6)GalNAcα-Sp8	149	80	36
149	Galβ1-4GlcNAcβ1-6(Galβ1-3)GalNAcα–Sp8	149	58	64
182	GlcAα-Sp8	149	93	44

211	NeuAcα2-3(NeuAcα2-3Galβ1-3GalNAcβ1-4)Galβ1-4Glcβ-Sp0	148	79	22
219	Neu5Acα2-3Galβ1-3[6OSO3]GalNAcα-Sp8	147	41	61
255	Neu5Acβ2-6(Galβ1-3)GalNAcα–Sp8	147	41	25
144	Galβ1-4GlcNAcβ1-3Galβ1-4(Fucα1-3)GlcNAcβ1-3Galβ1-4(Fucα1-3)GlcNAcβ-Sp0	146	29	48
229	Neu5Acα2-3Galβ1-4(Fucα1-3)GlcNAcβ–Sp0	146	43	46
77	Fucβ1-3GlcNAcβ-Sp8	145	117	83
235	Neu5Acα2-3Galβ1-4GlcNAcβ–Sp0	145	85	41
244	Neu5Acα2-6Galβ1-4GlcNAcβ–Sp0	144	60	27
195	Manα1-3(Manα1-2Manα1-2Manα1-6)Manα-Sp9	143	41	44
245	Neu5Acα2-6Galβ1-4GlcNAcβ–Sp8	142	30	28
270	Fucα1-2[6OSO3]Galβ1-4[6OSO3]Glc-Sp0	141	67	40
238	Neu5Acα2-3Galβ1-4Glcβ–Sp0	139	63	42
247	Neu5Acα2-6Galβ1-4GlcNAcβ1-3Galβ1-4GlcNAcβ-Sp0	139	97	101
203	Neu5Acα2-8Neu5Acα2-8Neu5Acα2-3(GalNAcβ1-4)Galβ1-4Glcβ-Sp0	137	79	29
122	Galβ1-3(Neu5Acβ2-6)GalNAcα-Sp8	136	142	79
88	GalNAcβ1-3(Fucα1-2)Galβ-Sp8	135	80	53
140	Galβ1-4GalNAcα1-3(Fucα1-2)Galβ1-4GlcNAcβ-Sp8	135	39	31
153	Galβ1-4Glcβ–Sp0	134	111	72
228	Neu5Acα2-3Galβ1-4(Fucα1-3)GlcNAcβ1-3Galβ1-4(Fucα1-3)GlcNAcβ1-3Galβ1-4(Fucα1-3) GlcNAcβ–Sp0	133	70	22
154	Galβ1-4Glcβ–Sp8	131	71	28
120	Galβ1-3(GlcNAcβ1-6)GalNAcα-Sp8	130	68	72
157	GlcNAcβ1-2Galβ1-3GalNAcα–Sp8	126	74	81
60	Fucα1-2Galβ1-3GlcNAcβ1-3Galβ1-4Glcβ–Sp10	124	56	61
269	Fucα1-2Galβ1-4[6OSO3]GlcNAc-Sp8	123	19	26
23	(Galβ1-4GlcNAcβ)2-3,6-GalNAcα–Sp8	121	48	70
166	GlcNAcβ1-3Galβ1-4Glcβ–Sp0	121	64	67
281	Galβ1-4(Fucα1-3)GlcNAcβ1-3Galβ1-3(Fucα1-4)GlcNAcβ-Sp0	121	92	17
237	Neu5Acα2-3Galβ1-4GlcNAcβ1-3Galβ1-4GlcNAcβ-Sp0	120	60	46
280	Galβ1-4[Fucα1-3][6OSO3]Glc-Sp0	120	41	36
134	Galβ1-4(Fucα1-3)GlcNAcβ–Sp0	117	52	82
205	Neu5Acα2-8Neu5Acα2-3(GalNAcβ1-4)Galβ1-4Glcβ–Sp0	116	85	66
215	Neu5Acα2-3Galβ1-3(6OSO3)GlcNAc-Sp8	116	57	25
68	Fucα1-2Galβ1-4GlcNAcβ1-3Galβ1-4GlcNAc–Sp0	115	92	71

256	Neu5Gcα2-3Galβ1-3(Fucα1-4)GlcNAcβ-Sp0	111	67	42
137	Galβ1-4(Fucα1-3)GlcNAcβ1-4Galβ1-4(Fucα1-3)GlcNAcβ1-4Galβ1-4(Fucα1-3)GlcNAcβ–Sp0	110	60	46
274	Galβ1-3-(Galβ1-4GlcNacβ1-6)GalNAc-T	110	61	33
282	Galβ1-4GlcNAcβ1-3Galβ1-3GlcNAcβ-Sp0	110	30	25
56	Fucα1-2Galβ1-3(Fucα1-4)GlcNAcβ–Sp8	108	55	63
204	Neu5Acα2-8Neu5Acα2-8Neu5Acα2-3Galβ1-4Glcβ–Sp0	108	48	42
236	Neu5Acα2-3Galβ1-4GlcNAcβ–Sp8	108	54	28
97	Galα1-3(Fucα1-2)Galβ1-4Glcβ-Sp0	104	64	53
116	Galβ1-3(Fucα1-4)GlcNAc–Sp0	104	103	46
206	Neu5Acα2-8Neu5Acα2-8Neu5Acα-Sp8	104	74	54
126	Galβ1-3GalNAcβ1-3Galα1-4Galβ1-4Glcβ-Sp0	103	46	47
207	Neu5Acα2-3(6-O-Su)Galβ1-4(Fucα1-3)GlcNAcβ–Sp8	100	43	46
18	β-D-Man–Sp8	97	46	52
173	GlcNAcβ1-6(Galβ1-3)GalNAcα–Sp8	96	68	44
253	Neu5Acβ2-6GalNAcα–Sp8	96	41	46
171	(GlcNAcβ1-4)5β-Sp8	95	48	71
248	Neu5Acα2-6Galβ1-4Glcβ–Sp0	95	14	63
177	Glcα1-4Glcα–Sp8	94	22	40
267	[3OSO3]Galβ1-4[Fucα1-3]GlcNAc-Sp0	94	42	42
186	KDNα2-3Galβ1-3GlcNAcβ–Sp0	93	42	51
284	Neu5Acα2-3Galβ1-4GlcNAcβ1-3Galβ1-3GlcNAcβ-Sp0	92	55	10
199	Manβ1-4GlcNAcβ-Sp0	90	6	56
283	Neu5Acα2-3Galβ1-3GlcNAcβ1-3Galβ1-3GlcNAcβ-Sp0	90	32	10
90	GalNAcβ1-4(Fucα1-3)GlcNAcβ-Sp0	89	37	22
222	NeuAcα2-3Galβ1-3GalNAcβ1-3Galα1-4Galβ1-4Glcβ-Sp0	88	56	57
278	Galβ1-3GlcNAcβ1-3Galβ1-3GlcNAcβ-Sp0	87	110	22
261	Neu5Gcα2-6GalNAcα–Sp0	79	21	40
276	Galβ1-3-(Neu5Aα2-3Galβ1-4GlcNacβ1-6)GalNAc-T	78	33	13
258	Neu5Gcα2-3Galβ1-4(Fucα1-3)GlcNAcβ-Sp0	77	34	33
268	Fucα1-2[6OSO3]Galβ1-4GlcNAc-Sp0	77	71	23
271	Fucα1-2-(6OSO3)-Galβ1-4Glc-Sp0	76	10	22
170	(GlcNAcβ1-4)6β-Sp8	74	29	66
239	Neu5Acα2-3Galβ1-4Glcβ–Sp8	74	35	55
249	Neu5Acα2-6Galβ1-4Glcβ–Sp8	74	45	30
251	Neu5Acα2-8Neu5Acα-Sp8	74	46	16

213	Neu5Acα2-3GalNAcα–Sp8	70	30	38
209	Neu5Acα2-3(GalNAcβ1-4)Galβ1-4GlcNAcβ-Sp8	68	71	57
119	Galβ1-3(Galβ1-4GlcNAcβ1-6)GalNAcα-Sp8	65	26	54
180	Glcβ1-6Glcβ-Sp8	65	12	76
145	Galβ1-4GlcNAcβ1-3Galβ1-4GlcNAcβ1-3Galβ1-4GlcNAcβ–Sp0	61	55	65
277	Galβ1-3GalNAc-T	61	15	27
279	Galβ1-4[Fucα1-3][6OSO3]GlcNAc-Sp0	57	20	23
265	[3OSO3]Galβ1-4(Fucα1-3)Glc-Sp0	55	3	33
242	Neu5Acα2-6GalNAcβ1-4GlcNAcβ-Sp0	52	23	34
226	Neu5Acα2-3Galβ1-4[6OSO3]GlcNAcβ-Sp8	43	15	41

Appendix Table 2: FULL CFG ARRAY DATA FOR XCGL-1

Chart #	Glycan	Average	STDEV	SEM
100	Gala1-3GlcNAca-Sp8	46749	14581	5952
101	Gala1-3GlcNAcb-Sp8	33429	6345	2590
105	Gala1-3Galb1-4Glcb-Sp0	5408	2872	1172
29	(3OSO3)Galb1-4(6OSO3)Glcb-Sp8	1671	3174	1296
65	Fuca1-2Galb1-4(Fuca1-3)GalNAcb1-3Galb1-4(Fuca1-3)GlcNAcb1-3Galb1-4(Fuca1-3)GalNAcb-Sp0	1537	2809	1147
3	Ceruloplasmin	1534	415	169
104	Gala1-3Galb1-4GlcNAcb-Sp8	1096	461	188
5	Transferrin	1047	401	164
0	Alpha1-acid glycoprotein (AGP)	912	339	139
278	Galb1-3GlcNAcb1-3Galb1-3GlcNAcb-Sp0	819	1716	701
152	Galb1-4GlcNAcb-Sp8	561	543	222
2	AGP-B (AGP ConA bound)	553	120	49
46	(6OSO3)GlcNAcb-Sp8	549	259	106
106	Gala1-3Galb-Sp8	532	348	142
1	AGP-A (AGP ConA flowthrough)	513	133	54
19	b-GalNAc-Sp8	505	134	55
113	Galb1-2Galb-Sp8	477	367	150
34	(3OSO3)Galb1-4(6OSO3)GlcNAcb-Sp8	468	137	56
35	(3OSO3)Galb1-4GlcNAcb-Sp0	467	128	52
99	Gala1-3(Gala1-4)Galb1-4GlcNAcb-Sp8	459	224	91
30	(3OSO3)Galb1-3(Fuca1-4)GlcNAcb-Sp8	424	90	37
13	a-NeuAc-Sp8	424	359	146
238	Neu5Aca2-3Galb1-4Glcb-Sp0	421	630	257
17	b-D-Glc-Sp8	417	171	70
47	9NAcNeu5Aca-Sp8	413	160	65
28	(3OSO3)Galb1-4(6OSO3)Glcb-Sp0	411	179	73
10	a-L-Fuc-Sp8	410	130	53
127	Galb1-3GalNAcb1-4(Neu5Aca2-3)Galb1-4Glcb-Sp0	408	292	119
37	(3OSO3)Galb-Sp8	407	217	88
22	b-GlcN(Gc)-Sp8	407	192	78
9	a-GalNAc-Sp8	405	276	113
77	Fucb1-3GlcNAcb-Sp8	402	241	98
21	b-GlcNAc-Sp8	393	222	91
54	Fuca1-2Galb1-3GalNAcb1-3Gala-Sp9	387	87	36
135	Galb1-4(Fuca1-3)GlcNAcb-Sp8	373	98	40

11	a-L-Fuc-Sp9	369	181	74
68	Fuca1-2Galb1-4GlcNAcb1-3Galb1-4GlcNAcb-Sp0	367	152	62
73	Fuca1-2Galb-Sp8	365	242	99
74	Fuca1-2GlcNAcb-Sp8	357	122	50
32	(3OSO3)Galb1-3GlcNAcb-Sp8	354	119	49
57	Fuca1-2Galb1-3GalNAca-Sp8	351	112	46
232	Neu5Aca2-3Galb1-4(Fuca1-3)GlcNAcb1-3Galb1-4GlcNAcb-Sp8	349	155	63
190	Mana1-2Mana1-3Mana-Sp9	345	265	108
133	Galb1-3GalNAcb-Sp8	344	196	80
62	Fuca1-2Galb1-3GalNAcb-Sp0	344	179	73
39	(4OSO3)Galb1-4GlcNAcb-Sp8	336	207	85
38	(4OSO3)(6OSO3)Galb1-4GlcNAcb-Sp0	334	323	132
26	(3OSO3)(6OSO3)Galb1-4GlcNAcb-Sp0	329	165	67
162	GlcNAcb1-3Galb1-3GalNAca-Sp8	326	174	71
12	a-L-Rha-Sp8	323	245	100
192	Mana1-2Mana1-6(Mana1-3)Mana1-6(Mana1-2Mana1-2Mana1-3)Manb1-4GlcNAcb1-4GlcNAcb-Asn	323	404	165
84	GalNAca1-3GalNAcb-Sp8	322	226	92
4	Fibrinogen	320	189	77
55	Fuca1-2Galb1-3GalNAcb1-3Gala1-4Galb1-4Glca-Sp9	315	111	45
93	Gala1-2Galb-Sp8	315	158	64
40	6-H2PO3Mana-Sp8	312	452	185
272	Fuca1-2Galb1-4(6OSO3)Glc-Sp0	312	155	63
42	(6OSO3)Galb1-4Glcb-Sp8	311	142	58
43	(6OSO3)Galb1-4GlcNAcb-Sp8	310	132	54
79	GalNAca1-3(Fuca1-2)Galb1-4(Fuca1-3)GlcNAcb-Sp0	307	107	44
269	Fuca1-2Galb1-4(6OSO3)GlcNAc-Sp8	305	148	61
91	GalNAcb1-4GlcNAcb-Sp0	303	222	91
120	Galb1-3(GlcNAcb1-6)GlcNAca-Sp8	302	199	81
97	Gala1-3(Fuca1-2)Galb1-4Glcb-Sp0	301	148	60
143	Galb1-4GlcNAcb1-3GalNAca-Sp8	299	120	49
78	GalNAca1-3(Fuca1-2)Galb1-3GlcNAcb-Sp0	298	111	45
210	Neu5Aca2-3(GalNAcb1-4)Galb1-4Glcb-Sp0	297	120	49
6	a-D-Gal-Sp8	297	145	59
33	(3OSO3)Galb1-4(Fuca1-3)GlcNAcb-Sp8	294	175	71
36	(3OSO3)Galb1-4GlcNAcb-Sp8	294	112	46
20	b-GlcNAc-Sp0	293	188	77
124	Galb1-3GalNAca-Sp8	292	236	96
194	Mana1-3(Mana1-6)Mana-Sp9	291	82	33
56	Fuca1-2Galb1-3(Fuca1-4)GlcNAcb-Sp8	289	70	29
198	Man5-9 mix -Asn	288	90	37

44	(6OSO3)Galb1-4(6OSO3)Glcb-Sp8	288	201	82
45	Neu5Aca2-3(6OSO3)Galb1-4GlcNAcb-Sp8	287	182	74
98	Gala1-3(Fuca1-2)Galb-Sp8	287	204	83
151	Galb1-4GlcNAcb-Sp0	285	62	25
228	Neu5Aca2-3Galb1-4(Fuca1-3)GlcNAcb1-3Galb1-4(Fuca1-3)GlcNAcb1-3Galb1-4(Fuca1-3)GlcNAcb-Sp0	283	114	47
69	Fuca1-2Galb1-4GlcNAcb1-3Galb1-4GlcNAcb1-3Galb1-4GlcNAcb-Sp0	283	106	43
139	Galb1-4(6OSO3)Glcb-Sp8	282	139	57
63	Fuca1-2Galb1-3GalNAcb-Sp8	279	135	55
75	Fuca1-3GlcNAcb-Sp8	278	213	87
16	b-D-Gal-Sp8	276	187	76
48	9NAcNeu5Aca2-6Galb1-4GlcNAcb-Sp8	276	74	30
136	Galb1-4(Fuca1-3)GlcNAcb1-4Galb1-4(Fuca1-3)GlcNAcb-Sp0	274	200	82
103	Gala1-3Galb1-3GlcNAcb-Sp0	271	217	89
70	Fuca1-2Galb1-4GlcNAcb-Sp0	270	131	53
27	(3OSO3)Galb1-4Glcb-Sp8	269	207	84
51	Galb1-4GlcNAcb1-2Mana1-3(Galb1-4GlcNAcb1-2Mana1-6)Manb1-4GlcNAcb1-4GlcNAcb-Gly	268	54	22
125	Galb1-3GalNAcb-Sp8	266	161	66
95	Gala1-3(Fuca1-2)Galb1-4(Fuca1-3)GlcNAcb-Sp0	265	178	73
71	Fuca1-2Galb1-4GlcNAcb-Sp8	261	211	86
217	Neu5Aca2-3Galb1-3(Fuca1-4)GlcNAcb1-3Galb1-4(Fuca1-3)GlcNAcb-Sp0	261	112	46
80	GalNAca1-3(Fuca1-2)Galb1-4GlcNAcb-Sp0	259	48	20
141	Galb1-4GalNAcb1-3(Fuca1-2)Galb1-4GlcNAcb-Sp8	258	206	84
186	KDNa2-3Galb1-3GlcNAcb-Sp0	258	129	53
183	GlcAb-Sp8	258	184	75
129	Galb1-3Galb-Sp8	257	206	84
126	Galb1-3GalNAcb1-3Gala1-4Galb1-4Glcb-Sp0	255	130	53
8	a-D-Man-Sp8	255	283	18
181	G-ol-amine	255	189	77
94	Gala1-3(Fuca1-2)Galb1-3GlcNAcb-Sp0	254	201	82
50	GlcNAcb1-2Mana1-3(GlcNAcb1-2Mana1-6)Manb1-4GlcNAcb1-4GlcNAcb-Gly	252	96	39
191	Mana1-6(Mana1-2Mana1-3)Mana1-6(Mana1-2Mana1-3)Manb1-4GlcNAcb1-4GlcNAcb-Asn	252	156	64
64	Fuca1-2Galb1-4(Fuca1-3)GalNAcb1-3Galb1-4(Fuca1-3)GlcNAcb-Sp0	251	98	40
174	GlcNAcb1-6GalNAca-Sp8	250	152	62
85	GalNAca1-3Galb-Sp8	249	155	63
115	Galb1-3(Fuca1-4)GlcNAcb1-3Galb1-4GlcNAcb-Sp0	249	89	36
15	b-NeuAc-Sp8	248	199	81

128	Galb1-3GalNAcb1-4Galb1-4Glcb-Sp8	247	162	66
92	GalNAcb1-4GlcNAcb-Sp8	247	175	71
83	GalNAca1-3(Fuca1-2)Galb-Sp8	245	121	49
121	Galb1-3(Neu5Aca2-6)GlcNAca-Sp8	244	118	48
188	Mana1-2Mana1-2Mana1-3Mana-Sp9	244	173	71
158	GlcNAcb1-3(GlcNAcb1-6)GlcNAca-Sp8	244	113	46
96	Gala1-3(Fuca1-2)Galb1-4GlcNAcb-Sp0	241	96	39
131	Galb1-3GalNAcb1-3Galb1-4Glcb-Sp10	240	80	33
283	Neu5Aca2-3Galb1-3GlcNAcb1-3Galb1-3GlcNAcb-Sp0	237	133	54
253	Neu5Aca1-6GalNAca-Sp8	236	175	72
165	GlcNAcb1-3Galb1-4GlcNAcb1-3Galb1-4GlcNAcb-Sp0	234	100	41
107	Gala1-4(Fuca1-2)Galb1-4GalNAcb-Sp8	233	193	79
24	GlcNAcb1-3(GlcNAcb1-4)(GlcNAcb1-6)GlcNAc-Sp8	233	112	46
59	Fuca1-2Galb1-3GalNAcb1-4(Neu5Aca2-3)Galb1-4Glcb-Sp9	232	106	43
261	Neu5Gca2-6GalNAca-Sp0	232	115	47
201	Neu5Aca2-3Galb1-3GalNAca-Sp8	232	114	47
132	Galb1-3GalNAcb-Sp0	231	129	53
49	Mana1-3(Mana1-6)Manb1-4GlcNAcb1-4GlcNAcb-Gly	229	146	60
25	(3OSO3)(6OSO3)Galb1-4(6OSO3)GlcNAcb-Sp0	227	142	58
118	Galb1-3(Fuca1-4)GlcNAcb-Sp8	225	57	23
52	Neu5Aca2-6Galb1-4GlcNAcb1-2Mana1-3(Neu5Aca2-6Galb1-4GlcNAcb1-2Mana1-6)Manb1-4GlcNAcb1-4GlcNAcb-Gly	225	84	34
212	Neu5Aca2-3(Neu5Aca2-6)GalNAca-Sp8	225	136	56
41	(6OSO3)Galb1-4Glcb-Sp0	224	191	78
154	Galb1-4Glcb-Sp8	224	200	82
180	Glcb1-6Glcb-Sp8	224	134	55
189	Mana1-2Mana1-3(Mana1-2Mana1-6)Mana-Sp9	223	162	66
109	Gala1-4Galb1-4GalNAcb-Sp8	223	114	46
138	Galb1-4(6OSO3)Glcb-Sp0	222	163	66
60	Fuca1-2Galb1-3GalNAcb1-3Galb1-4Glcb-Sp10	220	199	81
53	Neu5Aca2-6Galb1-4GlcNAcb1-2Mana1-3(Neu5Aca2-6Galb1-4GlcNAcb1-2Mana1-6)Manb1-4GlcNAcb1-4GlcNAcb-Sp8	217	76	31
82	GalNAca1-3(Fuca1-2)Galb1-4Glcb-Sp0	216	164	67
114	Galb1-3(Fuca1-4)GlcNAcb1-3Galb1-4(Fuca1-3)GlcNAcb-Sp0	214	88	36
31	(3OSO3)Galb1-3GalNAca-Sp8	212	164	67
116	Galb1-3(Fuca1-4)GlcNAc-Sp0	211	118	48
200	Neu5Aca2-3(Galb1-3GalNAcb1-4)Galb1-4Glcb-Sp0	211	159	65
90	GalNAcb1-4(Fuca1-3)Galb1-4GlcNAcb-Sp0	211	157	64
123	Galb1-3(Neu5Aca2-6)GlcNAc1-4Galb1Glcb-Sp10	210	123	50
117	Galb1-3(Fuca1-4)GlcNAc-Sp8	208	136	55
178	Glca1-6Glca1-6Glcb-Sp8	206	157	64

155	GlcNAcα1-3Galβ1-4GlcNAcβ-Sp8	206	240	98
264	(3OSO$_3$)Galβ1-4(Fucα1-3(6OSO$_3$)Glc-Sp0	205	72	29
258	Neu5Gcα2-3Galβ1-4(Fucα1-3)GlcNAcβ-Sp0	204	165	67
185	GlcAβ1-6Galβ-Sp8	202	176	72
122	Galβ1-3(Neu5Acβ2-6)GlcNAcα-Sp8	201	69	28
76	Fucα1-4GlcNAcβ-Sp8	201	72	30
89	GalNAcβ1-3Galα1-4Galβ1-4GlcNAcβ-Sp0	200	77	31
273	Galβ1-3(Fucα1-4)GlcNAcβ1-3Galβ1-3(Fucα1-4)GlcNAcβ-Sp0	200	197	81
176	Glcα1-4Glcβ-Sp8	200	183	75
108	Galα1-4Galβ1-4GalNAcβ-Sp0	200	147	60
23	(Galβ1-4GlcNAcβ)$_2$-3,6-GalNAcα-Sp8	198	121	50
187	KDNα2-3Galβ1-4GlcNAcβ-Sp0	198	95	39
219	Neu5Acα2-3Galβ1-3(6OSO$_3$)GalNAcβ-Sp8	197	121	49
81	GalNAcα1-3(Fucα1-2)Galβ1-4GlcNAcβ-Sp8	196	110	45
281	Galβ1-4(Fucα1-3)GlcNAcβ1-3Galβ1-3(Fucα1-4)GlcNAcβ-Sp0	196	191	78
215	Neu5Acα2-3Galβ1-3(6OSO$_3$)GlcNAcβ-Sp8	195	196	80
61	Fucα1-2Galβ1-3GalNAcβ1-3Galβ1-4Glcβ-Sp8	195	164	67
257	Neu5Gcα2-3Galβ1-3GlcNAcβ-Sp0	192	181	74
111	Galα1-4GlcNAcβ-Sp8	191	239	97
209	Neu5Acα2-3(GalNAcβ1-4)Galβ1-4GlcNAcβ-Sp8	191	131	53
66	Fucα1-2Galβ1-4(Fucα1-3)GlcNAcβ-Sp0	191	161	66
263	Neu5Gcα-Sp8	188	175	71
267	(3OSO$_3$)Galβ1-4(Fucα1-3)GlcNAc-Sp0	188	210	86
241	Neu5Acα2-6GalNAcα-Sp8	187	107	44
153	Galβ1-4Glcβ-Sp0	186	70	29
262	Neu5Gcα2-6Galβ1-4GlcNAcβ-Sp0	186	120	49
72	Fucα1-2Galβ1-4Glcβ-Sp0	184	160	65
196	Manα1-6(Manα1-3)Manα1-6(Manα1-2Manα1-3)Manβ1-4GlcNAcβ1-4GlcNAcβ-Asn	183	98	40
87	GalNAcα1-3GalNAcα-Sp8	182	96	39
260	Neu5Gcα2-3Galβ1-4Glcβ-Sp0	181	189	77
150	Galβ1-4GlcNAcβ1-6GalNAcα-Sp8	181	187	76
248	Neu5Acα2-6Galβ1-4Glcβ-Sp0	180	211	86
245	Neu5Acα2-6Galβ1-4GlcNAcβ-Sp8	179	108	44
254	Neu5Acα2-6Galβ1-4GlcNAcβ-Sp8	177	86	35
102	Galα1-3Galβ1-4(Fucα1-3)GalNAcβ-Sp8	177	93	38
259	Neu5Gcα2-3Galβ1-4GlcNAcβ-Sp0	177	148	60
207	Neu5Acα2-3(6-O-Su)Galβ1-4(Fucα1-3)GlcNAcβ-Sp8	177	124	51
235	Neu5Acα2-3Galβ1-4GlcNAcβ-Sp0	176	151	62
169	GlcNAcβ1-4Galβ1-4GlcNAcβ-Sp8	175	96	39
144	Galβ1-4GlcNAcβ1-3Galβ1-4(Fucα1-3)GlcNAcβ1-3Galβ1-4(Fucα1-3)GlcNAcβ-Sp0	175	94	38

282	Galb1-4GlcNAcb1-3Galb1-3GlcNAcb-Sp0	175	87	35
268	Fuca1-2(6OSO3)Galb1-4GlcNAc-Sp0	174	127	52
112	Gala1-6Glcb-Sp8	174	76	31
67	Fuca1-2Galb1-4(Fuca1-3)GlcNAcb-Sp8	173	93	38
195	Mana1-3(Mana1-2Mana1-2Mana1-6)Mana-Sp9	171	145	59
197	Mana1-6(Mana1-3)Mana1-6(Mana1-3)Manb1-4GlcNAcb1-4GlcNAcb-Asn	170	109	44
161	GlcNAcb1-3Galb-Sp8	168	130	53
157	GlcNAcb1-2Galb1-3GlcNAca-Sp8	167	74	30
229	Neu5Aca2-3Galb1-4(Fuca1-3)GlcNAcb-Sp0	167	109	45
243	Neu5Aca2-6Galb1-4(6OSO3)GlcNAcb-Sp8	166	138	56
204	Neu5Aca2-8Neu5Aca2-8Neu5Aca2-3Galb1-4Glcb-Sp0	165	82	33
256	Neu5Gca2-3Galb1-3(Fuca1-4)GlcNAcb-Sp0	165	169	69
206	Neu5Aca2-8Neu5Aca2-8Neu5Aca-Sp8	164	80	33
86	GalNAca1-4(Fuca1-2)Galb1-4GlcNAcb-Sp8	164	71	29
216	Neu5Aca2-3Galb1-3(Fuca1-4)GlcNAcb-Sp8	163	103	42
221	Neu5Aca2-3Galb-Sp8	163	161	66
266	(3OSO3)Galb1-4(Fuca1-3(6OSO3)GlcNAc-Sp8	163	132	54
14	Neu5Aca-Sp11	160	113	46
279	Galb1-4(Fuca1-3)(6OSO3)GlcNAc-Sp0	160	98	40
202	Neu5Aca2-8Neu5Aca2-8Neu5Aca2-8Neu5Aca2-3(GalNAcb1-4)Galb1-4Glcb-Sp0	160	140	57
222	Neu5Aca2-3Galb1-3GalNAcb1-3Gala1-4Galb1-4Glcb-Sp0	159	110	45
226	Neu5Aca2-3Galb1-4(6OSO3)GlcNAcb-Sp8	158	54	22
242	Neu5Aca2-6GalNAcb1-4GlcNAcb-Sp0	157	115	47
208	Neu5Aca2-3(GalNAcb1-4)Galb1-4GlcNAcb-Sp0	156	92	38
164	GlcNAcb1-3Galb1-4GlcNAcb-Sp8	155	73	30
137	Galb1-4(Fuca1-3)GlcNAcb1-4Galb1-4(Fuca1-3)GlcNAcb1-4Galb1-4(Fuca1-3)GlcNAcb-Sp0	155	61	25
142	Galb1-4GlcNAcb1-3(Galb1-4GlcNAcb1-6)GalNAca-Sp8	155	126	51
110	Gala1-4Galb1-4Galb-Sp0	155	97	40
280	Galb1-4(Fuca1-3)(6OSO3)Glc-Sp0	154	80	33
213	Neu5Aca2-3GalNAca-Sp8	153	149	61
274	Galb1-3(Galb1-4GlcNAcb1-6)GalNAc-T	153	60	25
140	Galb1-4GalNAca1-3(Fuca1-2)Galb1-4GlcNAcb-Sp8	152	151	62
58	Fuca1-2Galb1-3GalNAcb1-4(Neu5Aca2-3)Galb1-4Glcb-Sp0	152	141	57
250	Neu5Aca2-6Galb-Sp8	151	177	72
199	Manb1-4GlcNAcb-Sp0	151	123	50
218	Neu5Aca2-3Galb1-3(Neu5Aca2-3Galb1-4)GlcNAcb-Sp8	151	57	23
211	Neu5Aca2-3(Neu5Aca2-3Galb1-3GalNAcb1-4)Galb1-4Glcb-Sp0	146	101	41
119	Galb1-3(Galb1-4GlcNAcb1-6)GlcNAca-Sp8	146	91	37

246	Neu5Aca2-6Galb1-4GlcNAcb1-3Galb1-4(Fuca1-3)GlcNAcb1-3Galb1-4(Fuca1-3)GlcNAcb-Sp0	146	55	22
134	Galb1-4(Fuca1-3)GlcNAcb-Sp0	142	112	46
247	Neu5Aca2-6Galb1-4GlcNAcb1-3Galb1-4GlcNAcb-Sp0	142	100	41
240	Neu5Aca2-6(Galb1-3)GalNAca-Sp8	142	161	66
223	Neu5Aca2-3Galb1-3GlcNAcb1-3Galb1-4GlcNAcb-Sp0	141	76	31
88	GalNAca1-3(Fuca1-2)Galb-Sp8	140	79	32
170	(GlcNAcb1-4)$_6$b-Sp8	138	86	35
225	Neu5Aca2-3Galb1-3GlcNAcb-Sp8	138	101	41
233	Neu5Aca2-3Galb1-4GlcNAcb1-3Galb1-4(Fuca1-3)GlcNAcb-Sp0	137	69	28
160	GlcNAcb1-3GalNAca-Sp8	136	112	46
234	Neu5Aca2-3Galb1-4GlcNAcb1-3Galb1-4GlcNAcb1-3Galb1-4GlcNAcb-Sp0	136	103	42
214	Neu5Aca2-3GalNAcb1-4GlcNAcb-Sp0	132	134	55
167	GlcNAcb1-4MDPLys (bacterial cell wall)	132	87	36
237	Neu5Aca2-3Galb1-4GlcNAcb1-3Galb1-4GlcNAcb-Sp0	131	43	18
276	Galb1-3(Neu5Aca2-3Galb1-4GlcNAcb1-6)GalNAc-T	131	83	34
148	Galb1-4GlcNAcb1-3Galb1-4Glcb-Sp8	129	117	48
227	Neu5Aca2-3Galb1-4(Fuca1-3)(6OSO$_3$)GlcNAcb-Sp8	128	67	27
173	GlcNAcb1-6(Galb1-3)GalNAca-Sp8	128	70	28
193	Mana1-2Mana1-2Mana1-3[Mana1-2Mana1-3(Mana1-2Mana1-6)Mana1-6]Manb1-4GlcNAcb1-4GlcNAcb-Asn	128	93	38
182	GlcAa-Sp8	127	246	101
284	Neu5Aca2-3Galb1-4GlcNAcb1-3Galb1-3GlcNAcb-Sp0	127	155	63
146	Galb1-4GlcNAcb1-3Galb1-4GlcNAcb-Sp0	125	73	30
230	Neu5Aca2-3Galb1-4(Fuca1-3)GlcNAcb-Sp8	125	100	41
236	Neu5Aca2-3Galb1-4GlcNAcb-Sp8	124	40	16
224	Neu5Aca2-3Galb1-3GlcNAcb-Sp0	123	50	20
205	Neu5Aca2-8Neu5Aca2-3(GalNAcb1-4)Galb1-4Glcb-Sp0	123	113	46
145	Galb1-4GlcNAcb1-3Galb1-4GlcNAcb1-3Galb1-4GlcNAcb-Sp0	114	61	25
7	a-D-Glc-Sp8	111	62	25
244	Neu5Aca2-6Galb1-4GlcNAcb-Sp0	111	102	42
149	Galb1-4GlcNAcb1-6(Galb1-3)GalNAca-Sp8	107	61	25
166	GlcNAcb1-3Galb1-4Glcb-Sp0	107	81	33
203	Neu5Aca2-8Neu5Aca2-8Neu5Aca2-3(GalNAcb1-4)Galb1-4Glcb-Sp0	106	130	53
147	Galb1-4GlcNAcb1-3Galb1-4Glcb-Sp0	105	55	23
249	Neu5Aca2-6Galb1-4Glcb-Sp8	102	99	40
175	GlcNAcb1-6Galb1-4GlcNAcb-Sp8	101	126	52
159	GlcNAcb1-3(GlcNAcb1-6)Galb1-4GlcNAcb-Sp8	101	173	70

172	GlcNAcb1-4GlcNAcb1-4GlcNAcb-Sp8	100	146	60
277	Galb1-3GalNAc-T	97	81	33
275	Galb1-3(GlcNAcb1-6)GalNAc-T	95	82	34
220	Neu5Aca2-3Galb1-3(Neu5Aca2-6)GalNAcb-Sp8	95	102	42
265	(3OSO_3)Galb1-4(Fuca1-3)Glc-Sp0	87	55	23
231	Neu5Aca2-3Galb1-4(Fuca1-3)GlcNAcb1-3Galb-Sp8	84	64	26
163	GlcNAcb1-3Galb1-4GlcNAcb-Sp0	83	99	40
239	Neu5Aca2-3Galb1-4Glcb-Sp8	81	54	22
18	b-D-Man-Sp8	81	77	31
168	GlcNAcb1-4(GlcNAcb1-6)GalNAca-Sp8	77	75	30
179	Glcb1-4Glcb-Sp8	76	80	33
255	Neu5Aca2-6(Galb1-3)GalNAca-Sp8	75	139	57
171	(GlcNAcb1-4)$_5$b-Sp8	72	33	13
184	GlcAb1-3Galb-Sp8	68	66	27
271	Fuca1-2(6OSO_3)Galb1-4Glc-Sp0	65	55	22
177	Glca1-4Glca-Sp8	65	57	23
270	Fuca1-2(6OSO_3)Galb1-4(6OSO_3)Glc-Sp0	64	88	36
130	Galb1-3GalNAcb1-3Galb1-4GlcNAcb-Sp0	57	42	17
156	GlcNAca1-6Galb1-4GlcNAcb-Sp8	56	62	25
252	Neu5Aca2-8Neu5Aca2-3Galb1-4Glcb-Sp0	48	25	10
251	Neu5Aca2-8Neu5Aca-Sp8	45	24	10

Appendix Table 3: FULL MGM ARRAY DATA FOR HINTL-1

Chart #	Glycan	Average	STDEV	SEM
87	Proteus vulgaris O54a,54c (TG 103)	59156	486	198
241	Streptococcus pneumoniae type 43 (Danish type 11A)	55859	5974	2439
244	Streptococcus pneumoniae type 56 (Danish type 18C)	55786	6776	2766
41	Proteus mirabilis O54a, 54b (10704)	53927	4356	1778
188	Klebsiella O2ac OPS	52552	4507	1840
247	Streptococcus pneumoniae type 70 (Danish type 33F)	50444	1852	756
186	Klebsiella O1 OPS	48833	1653	675
236	Streptococcus pneumoniae type 20 (Danish type 20)	44539	1797	733
187	Klebsiella O2a OPS	36894	6729	2747
192	Klebsiella O8 OPS	30359	2367	966
272	Proteus mirabilis O54ab (OE)	22210	1783	728
18	Escherichia coli O128-B12 LPS	8028	2136	872
281	Shigella flexneri type 4a	7238	456	186
255	Yersinia pseudotuberculosis O:2a-dhmA	3793	524	214
98	Proteus mirabilis O27 (PrK 50/57)	3426	306	125
240	Streptococcus pneumoniae type 34 (Danish type 10A)	3100	274	112
214	Shigella dysenteriae type 2	3095	542	221
102	Proteus mirabilis O41 (PrK 67/57)	2822	288	118
36	Proteus mirabilis O3a, 3c (G1)	2730	1229	502
99	Proteus mirabilis O28 (PrK 51/57)	2711	298	122
97	Proteus mirabilis O26 (PrK 49/57)	2564	167	68
243	Streptococcus pneumoniae type 54 (Danish type 15B)	2201	120	49
131	Providencia alcalifaciens O5	2140	249	102
225	Streptococcus pneumoniae type 1 (Danish type 1)	1940	61	25
245	Streptococcus pneumoniae type 57 (Danish type 19A)	1513	67	28
142	Providencia alcalifaciens O36*	1476	98	40
260	Proteus mirabilis O3ab (S1959)	1209	168	68
159	Yersinia pestis, KM260(11)-ΔpmrF	1173	316	129
48	Proteus penneri O72a, 72b (4)	1156	86	35
116	Proteus penneri O67 (8)	1097	124	51

95	Proteus mirabilis O17 (PrK 32/57)	1093	196	80
223	Escherichia coli O55:B5 LPS- solution at 1 mg/mL	982	145	59
16	Serratia marcescens LPS	909	206	84
311	Yeast Mannan Sigma M-3640	905	15	6
172	Yersinia pestis, 11M-37	888	21	9
94	Proteus mirabilis O16 (4652)	888	67	27
239	Streptococcus pneumoniae type 26 (Danish type 6B)	851	436	178
9	Salmonella typhimurium SL 11881 (Re mut)	843	197	80
274	Proteus vulgaris O76 (HSC438)	840	152	62
266	Proteus mirabilis O20 (PrK 38/57)	808	51	21
46	Proteus penneri O69 (25)	806	30	12
134	Providencia alcalifaciens O19	796	70	28
269	Proteus mirabilis O43 (PrK 69/57)	789	116	47
283	Shigella flexneri type 5b	718	192	78
278	Shigella flexneri type 2b	713	61	25
169	Yersinia pestis, KIMD1-37	687	357	146
224	Escherichia coli O127:B8 LPS- solution at 1 mg/mL	674	745	304
60	Pseudomonas aerguinosa O6 Immuno 1	673	16	6
227	Streptococcus pneumoniae type 3 (Danish type 3)	669	57	23
285	Shigella flexneri type 6	643	246	100
228	Streptococcus pneumoniae type 4 (Danish type 4)	635	424	173
47	Proteus penneri O71 (42)	626	125	51
171	Yersinia pestis, 11M-25	603	35	14
242	Streptococcus pneumoniae type 51 (Danish type 7F)	580	290	118
268	Proteus mirabilis O33 (D52)	573	64	26
88	Proteus vulgaris O55 (TG 155)	573	234	96
58	Pseudomonas aerguinosa O6 6a	559	314	128
117	Proteus penneri O68 (63)	556	148	61
103	Proteus mirabilis O51 (19011)*	553	95	39
276	Shigella flexneri type 1b	553	224	91
233	Streptococcus pneumoniae type 14 (Danish type 14)	534	165	68
55	Pseudomonas aerguinosa O3 3a,3b,3c	531	781	319
294	Escherichia coli O85	531	55	22
308	Yeast Mannan	531	157	64
59	Pseudomonas aerguinosa O6 6a,6c	524	79	32
143	Providencia alcalifaciens O39	522	43	17
265	Proteus mirabilis O18 (PrK 34/57)	522	95	39
246	Streptococcus pneumoniae type 68 (Danish type 9V)	521	189	77
11	Salmonella typhimurium SL 684 (Rc mut)	517	54	22
70	Pseudomonas aerguinosa O15 15	491	41	17
279	Shigella flexneri type 3a	490	85	35

218	Shigella dysenteriae type 7	483	225	92
132	Providencia alcalifaciens O6*	479	39	16
118	Proteus penneri O70 (60)	477	48	20
190	Klebsiella O4 OPS	477	45	18
226	Streptococcus pneumoniae type 2 (Danish type 2)	476	425	173
21	Salmonella enteritidis LPS	456	55	23
237	Streptococcus pneumoniae type 22 (Danish type 22F)	445	209	85
22	Shigella bodyii type2	442	34	14
291	Shigella boydii type 13	440	26	10
174	Proteus vulgaris O24 (PrK 47/57)	434	129	53
175	Yersinia pestis KM260(11)-6C	433	112	46
65	Pseudomonas aerguinosa O10 10a,10c	423	16	7
10	Salmonella typhimurium TV 119 (Ra mut)	417	11	5
140	Providencia alcalifaciens O30	411	67	27
301	Salmonella enterica O55	407	83	34
69	Pseudomonas aerguinosa O14 14	406	99	40
198	Shigella boydii type 11	405	39	16
137	Providencia alcalifaciens O23	405	69	28
4	Pseudomonas aerguinosa O2 (Fisher immunotype 3)	403	11	4
144	Providencia rustigianii O14	399	62	25
235	Streptococcus pneumoniae type 19 (Danish type 19F)	396	181	74
1	Providencia stuartii O52	393	40	16
234	Streptococcus pneumoniae type 17 (Danish type 17F)	393	25	10
136	Providencia alcalifaciens O21	390	67	27
119	Proteus penneri O73a,b (103)	390	201	82
135	Providencia alcalifaciens O19	388	70	29
86	Proteus vulgaris O53 (TG 276-10)	387	128	52
251	Yersinia pestis KM260(11)-wabC/waaL	376	69	28
256	Yersinia pseudotuberculosis O:2c	375	59	24
178	Yersinia pestis 260(11)-37C-416	369	20	8
113	Proteus penneri O63 (22)	369	42	17
81	Proteus vulgaris O34 (4669)*	368	29	12
129	Providencia stuartii O49, Core 1	361	68	28
128	Providencia stuartii O47, Core 9	349	26	11
71	Proteus vulgaris O1 (18984)*	347	40	16
284	Shigella flexneri type 6a	346	18	7
176	Yersinia pestis 260(11)-37C-186	336	32	13
282	Shigella flexneri type 4b	334	39	16

78	Proteus vulgaris O21 (PrK 39/57)*	330	30	12
133	Providencia alcalifaciens O19	325	71	29
231	Streptococcus pneumoniae type 9 (Danish type 9N)	324	143	58
101	Proteus mirabilis O40 (10703)	320	79	32
124	Providencia stuartii O20*	313	40	16
298	Salmonella enterica O17	312	18	7
271	Proteus mirabilis O49 (PrK 75/57)	311	89	36
304	Escherichia coli K13	311	140	57
92	Proteus mirabilis O13 (PrK 26/57)	301	15	6
299	Salmonella enterica O28	300	27	11
17	Escherichia coli K235 LPS	300	12	5
232	Streptococcus pneumoniae type 12 (Danish type 12F)	294	9	3
68	Pseudomonas aerguinosa O13 13a,13c	291	309	126
12	Pseudomonas aerguinosa O10	289	13	5
216	Shigella dysenteriae type 5	287	153	62
193	Klebsiella O12 OPS	287	44	18
165	Yersinia pestis, KM260(11)-ΔwaaL	279	79	32
166	Yersinia pestis, KM260(11)-25	276	26	11
286	Shigella flexneri type X	274	63	26
296	Escherichia coli O145	273	7	3
141	Providencia alcalifaciens O32	269	79	32
238	Streptococcus pneumoniae type 23 (Danish type 23F)	269	20	8
54	Pseudomonas aerguinosa O3 3a,3b	268	110	45
61	Pseudomonas aerguinosa O7 7a,7b,7c	265	34	14
287	Shigella dysenteriae type 1	264	21	8
197	Shigella boydii type 9	262	44	18
153	Yersinia pestis, 1146-37	260	19	8
157	Yersinia pestis, KM218-25	259	41	17
180	Yersinia pestis P-1680-25C	258	42	17
199	Shigella boydii type 12	258	6	3
182	Yersinia pestis I-2377-25C	256	44	18
230	Streptococcus pneumoniae type 8 (Danish type 8)	253	34	14
161	Yersinia pestis, KM260(11)-Δ0186	252	7	3
194	Shigella boydii type 1	250	14	6
155	Yersinia pestis, 0KM218-37	249	36	15
292	Shigella boydii type 14	249	59	24
209	Escherichia coli O112ab	245	12	5
204	Escherichia coli O49	231	22	9
63	Pseudomonas aerguinosa O7 7a,7d	229	28	11

248	Yersinia pestis KM218-6C	224	42	17
52	Pseudomonas aerguinosa O2 2a,2d	223	68	28
31	Escherichia coli O106	222	190	77
162	Yersinia pestis, KM260(11)-Δ0186	217	127	52
191	Klebsiella O5 OPS	214	12	5
28	Shigella dysennteriae type 13	211	31	12
145	Providencia rustigianii O16	211	42	17
89	Proteus vulgaris O65 (TG 251)	211	99	40
258	Yersinia pseudotuberculosis O:4b	201	51	21
147	Yersinia pestis, KM260(11)-Δ0187	200	148	60
289	Shigella boydii type 7	198	63	26
249	Yersinia pestis KM260(11)-yjhW-6C	196	31	13
93	Proteus mirabilis O14a,14b (PrK 29/57)	194	308	126
149	Yersinia pestis, KM260(11)-Δrfe	191	26	11
183	Yersinia pestis I-2377-37C	191	64	26
206	Escherichia coli O58	190	38	16
27	Shigella dysennteriae type 11	188	268	110
96	Proteus mirabilis O23a,b,d (PrK 42/57)	185	16	7
297	Escherichia coli O107	183	39	16
290	Shigella boydii type 8	182	43	18
127	Providencia stuartii O47	176	13	5
2	Pseudomonas aerguinosa O4 (Habs serotype 4)	176	31	13
82	Proteus vulgaris O37a,b (PrK 63/57)	176	67	27
300	Salmonella enterica O47	176	29	12
254	Yersinia pseudotuberculosis O:2a	170	52	21
42	Proteus mirabilis O57 (TG319)	167	30	12
64	Pseudomonas aerguinosa O10 10a,10b	163	77	31
181	Yersinia pestis P-1680-37C	160	23	9
280	Shigella flexneri type 3b	160	18	7
177	Yersinia pestis 260(11)-37C-187	158	13	5
293	Escherichia coli O71	157	17	7
122	Providencia stuartii O4	157	21	9
38	Proteus mirabilis O10 (HJ4320)	153	24	10
189	Klebsiella O3 OPS	153	114	47
49	Pseudomonas aerguinosa O2 (2a),2d,2f	153	27	11
151	Yersinia pestis, 1146-25	153	35	14
126	Providencia stuartii O44	152	23	9
66	Pseudomonas aerguinosa O11 11a,11b	151	8	3
179	Yersinia pestis 260(11)-37C-417	150	37	15
62	Pseudomonas aerguinosa O7 7a,7b,7d	148	12	5
257	Yersinia pseudotuberculosis O:3	147	28	12

130	Providencia stuartii O57	146	5	2
201	Shigella boydii type 16	145	22	9
264	Proteus penneri O17 (16)	142	23	10
120	Proteus myxofaciens O60	138	31	13
24	Shigella bodyii type10	136	32	13
138	Providencia alcalifaciens O27	132	35	14
104	Proteus mirabilis O74 (10705, OF)	132	40	16
74	Proteus vulgaris O13 (8344)	130	31	13
163	Yersinia pestis, KM260(11)-ΔwaaQ	128	52	21
170	Yersinia pestis, KIMD1-25	127	24	10
106	Proteus mirabilis O77 (3 B-m)	125	44	18
13	Salmonella typhimurium dodeca saccharide	124	16	7
80	Proteus vulgaris O25 (PrK 48/57)	123	29	12
15	Salmonella typhimurium LPS	123	6	3
220	Shigella dysenteriae type 9	123	93	38
6	Pseudomonas aerguinosa O9 (9a, 9b, 9d)	121	9	4
213	Escherichia coli O168	119	59	24
302	Escherichia coli K92	119	10	4
50	Pseudomonas aerguinosa O2 2a,2b	118	25	10
67	Pseudomonas aerguinosa O12 12	117	63	26
84	Proteus vulgaris O44 (PrK 67/57)	115	20	8
203	Shigella boydii type 18	113	57	23
262	Proteus mirabilis O9 (PrK 18/57)	112	111	45
45	Proteus penneri O66 (2)	112	60	25
208	Escherichia coli O73	110	12	5
212	Escherichia coli O151	110	11	4
56	Pseudomonas aerguinosa O3 3a,3d	109	105	43
0	Providencia stuartii O49	108	32	13
288	Shigella boydii type 6	107	35	14
202	Shigella boydii type 17	107	50	21
168	Yersinia pestis, KM260(11)-37	106	47	19
7	Pseudomonas aerguinosa O6a (Habs serotype6, fraction IIa)	105	12	5
115	Proteus penneri O65 (34)	103	16	7
23	Shigella bodyii type4	102	8	3
32	Escherichia coli O130	101	20	8
152	Yersinia pestis 1146-25	101	23	9
154	Yersinia pestis, 1146-37	101	26	11
306	Davanat	101	210	86
39	Proteus mirabilis O29a, 29b (2002)	100	11	4
219	Shigella dysenteriae type 8 (Russian)	100	21	8

207	Escherichia coli O61	97	17	7
273	Proteus penneri O73ac (75)	96	17	7
19	Salmonella enterica abortus equi LPS	94	35	14
40	Proteus mirabilis O50 (TG332)	91	12	5
277	Shigella flexneri type 2a	91	21	9
195	Shigella boydii type 3	89	12	5
121	Proteus O56 (genomospecies 4)	89	26	10
173	Proteus mirabilis O23a, 23b, 23c (CCUG 10701)	88	33	13
309	Escherichia coli O86	87	18	7
30	Escherichia coli O40	86	8	3
253	Yersinia pseudotuberculosis 85pCad-20C	84	12	5
125	Providencia stuartii O43	79	23	9
263	Proteus mirabilis O11 (9B-m)	79	13	5
29	Escherichia coli O29	78	16	7
146	Providencia rustigianii O34	77	15	6
14	Salmonella enteritidis dodeca saccharide	77	10	4
205	Escherichia coli O52	75	22	9
267	Proteus penneri O31ab (28)	74	51	21
111	Proteus penneri O61 (21)	74	7	3
139	Providencia alcalifaciens O29	73	32	13
108	Proteus penneri O52 (15)	72	61	25
112	Proteus penneri O62 (41)	72	14	6
75	Proteus vulgaris O15 (PrK 30/57)	71	29	12
210	Escherichia coli O118	71	58	24
211	Escherichia coli O125	70	20	8
184	Francisella novicida OPS	70	14	6
8	Pseudomonas aerguinosa O6a (Habs serotype6, fraction IIb)	69	10	4
3	Pseudomonas aerguinosa O1 (Fisher immunotype 4)	67	9	4
196	Shigella boydii type 5	67	21	8
261	Proteus mirabilis O5 (PrK 12/57)	66	11	4
123	Providencia stuartii O18	65	22	9
164	Yersinia pestis, KM260(11)-ΔwaaQ	64	11	4
160	Yersinia pestis, KM260(11)-ΔpmrF	63	18	7
310	Galactomannan DAVANT (160102) Pro-Pharmacenti	61	68	28
312	1-2 Mannan Acetobacter methanolieus MB135	59	4	1
73	Proteus vulgaris O12 (PrK 25/57)	59	69	28
53	Pseudomonas aerguinosa O2 Immuno 7	58	21	9
200	Shigella boydii type 15	58	9	4
5	Pseudomonas aerguinosa O13 (Sandvik serotype II)	56	12	5

307	Laminarin	55	14	6
90	Proteus mirabilis O6 (PrK 14/57)	53	13	5
167	Yersinia pestis, KM260(11)-25	52	40	16
76	Proteus vulgaris O17 (PrK 33/57)	51	27	11
77	Proteus vulgaris O19a (PrK 37/57)	51	22	9
158	Yersinia pestis, KM218-25	51	24	10
305	Neisseria meningitidis Group C	50	24	10
110	Proteus penneri O59 (9)	50	101	41
37	Proteus mirabilis O8 (TG326)	49	10	4
35	Escherichia coli O180	48	33	14
34	Escherichia coli O150	45	12	5
43	Proteus penneri O8 (106)	44	25	10
85	Proteus vulgaris O45 (4680)	44	10	4
114	Proteus penneri O64a,b,c (27)	43	15	6
275	Shigella flexneri type 1a	43	15	6
250	Yersinia pestis KM260(11)-wabD/waaL	42	18	7
148	Yersinia pestis, KM260(11)-Δ0187	41	1	0
79	Proteus vulgaris O22 (PrK 40/57)	41	6	3
109	Proteus penneri O58 (12)	40	7	3
26	Shigella dysennteriae type 8 (batch 12)	40	85	35
100	Proteus mirabilis O29a (PrK 52/57)	38	3	1
185	Francisella tularensis OPS	35	15	6
83	Proteus vulgaris O37a,c (PrK 72/57)	35	7	3
259	Proteus vulgaris O2 (OX2)	34	3	1
105	Proteus mirabilis O75 (10702, OC)	34	9	4
295	Escherichia coli O99	33	37	15
150	Yersinia pestis, KM260(11)-Δrfe	33	5	2
217	Shigella dysenteriae type 6 SR-strain	31	5	2
270	Proteus vulgaris O47 (PrK 73/57)	31	6	3
156	Yersinia pestis, KM218-37	31	1	1
107	Proteus penneri O31a (26)	30	6	2
33	Escherichia coli O148	30	51	21
91	Proteus mirabilis O11 (PrK 24/57)	27	53	22
44	Proteus penneri O64a, 64b, 64d (39)	25	10	4
72	Proteus vulgaris O4 (PrK 9/57)	25	5	2
20	Salmonella typhosa LPS	19	105	43
303	Escherichia coli K5	19	2	1
25	Shigella dysennteriae type 3	16	40	16
222	Escherichia coli O26:B6 LPS- solution at 1 mg/mL	7	48	19
252	Yersinia pseudotuberculosis 85pCad-37C	7	7	3
215	Shigella dysenteriae type 4	-2	117	48

229	Streptococcus pneumoniae type 5 (Danish type 5)	-28	19	8
57	Pseudomonas aerguinosa O4 4a,4c	-37	38	15
221	Escherichia coli O111:B4 LPS- solution at 1 mg/mL	-37	71	29
51	Pseudomonas aerguinosa O2 2a,2b,2e	-52	33	14

Appendix Table 4: FULL MGM ARRAY DATA FOR HINTL-2

Chart #	Glycan	Average	STDEV	SEM
241	Streptococcus pneumoniae type 43 (Danish type 11A)	42097	2515	1027
272	Proteus mirabilis O54ab (OE)	35456	2294	937
41	Proteus mirabilis O54a, 54b (10704)	33813	3665	1496
87	Proteus vulgaris O54a,54c (TG 103)	12361	1181	482
233	Streptococcus pneumoniae type 14 (Danish type 14)	8049	8988	3670
188	Klebsiella O2ac OPS	6649	3626	1480
274	Proteus vulgaris O76 (HSC438)	3082	740	302
143	Providencia alcalifaciens O39	2601	1318	538
132	Providencia alcalifaciens O6*	2561	123	50
186	Klebsiella O1 OPS	2270	750	306
247	Streptococcus pneumoniae type 70 (Danish type 33F)	2058	90	37
98	Proteus mirabilis O27 (PrK 50/57)	1672	1119	457
99	Proteus mirabilis O28 (PrK 51/57)	1520	505	206
102	Proteus mirabilis O41 (PrK 67/57)	1469	1048	428
166	Yersinia pestis, KM260(11)-25	1266	72	30
10	Salmonella typhimurium TV 119 (Ra mut)	1252	126	51
187	Klebsiella O2a OPS	1252	200	82
281	Shigella flexneri type 4a	1187	306	125
142	Providencia alcalifaciens O36*	1140	166	68
192	Klebsiella O8 OPS	1120	55	23
149	Yersinia pestis, KM260(11)-Δrfe	1056	91	37
134	Providencia alcalifaciens O19	1049	430	176
251	Yersinia pestis KM260(11)-wabC/waaL	966	99	41
46	Proteus penneri O69 (25)	916	178	73
94	Proteus mirabilis O16 (4652)	903	81	33
183	Yersinia pestis I-2377-37C	881	70	29
221	Escherichia coli O111:B4 LPS- solution at 1 mg/mL	842	75	31
266	Proteus mirabilis O20 (PrK 38/57)	827	38	15
181	Yersinia pestis P-1680-37C	818	21	9
118	Proteus penneri O70 (60)	778	18	7
161	Yersinia pestis, KM260(11)-Δ0186	770	208	85
159	Yersinia pestis, KM260(11)-ΔpmrF	761	30	12
175	Yersinia pestis KM260(11)-6C	748	48	20
163	Yersinia pestis, KM260(11)-ΔwaaQ	718	135	55

177	Yersinia pestis 260(11)-37C-187	713	75	31
179	Yersinia pestis 260(11)-37C-417	703	101	41
176	Yersinia pestis 260(11)-37C-186	665	111	45
155	Yersinia pestis, KM218-37	658	47	19
171	Yersinia pestis, 11M-25	634	104	43
97	Proteus mirabilis O26 (PrK 49/57)	630	559	228
165	Yersinia pestis, KM260(11)-ΔwaaL	628	363	148
124	Providencia stuartii O20*	628	124	51
153	Yersinia pestis, 1146-37	620	157	64
103	Proteus mirabilis O51 (19011)*	575	50	20
117	Proteus penneri O68 (63)	554	52	21
291	Shigella boydii type 13	549	90	37
68	Pseudomonas aerguinosa O13 13a,13c	538	98	40
144	Providencia rustigianii O14	529	61	25
36	Proteus mirabilis O3a, 3c (G1)	512	259	106
178	Yersinia pestis 260(11)-37C-416	509	362	148
244	Streptococcus pneumoniae type 56 (Danish type 18C)	487	32	13
136	Providencia alcalifaciens O21	478	24	10
22	Shigella bodyii type2	477	20	8
236	Streptococcus pneumoniae type 20 (Danish type 20)	474	63	26
190	Klebsiella O4 OPS	473	256	105
116	Proteus penneri O67 (8)	470	127	52
17	Escherichia coli K235 LPS	461	122	50
95	Proteus mirabilis O17 (PrK 32/57)	453	70	29
214	Shigella dysenteriae type 2	436	52	21
225	Streptococcus pneumoniae type 1 (Danish type 1)	434	15	6
63	Pseudomonas aerguinosa O7 7a,7d	420	57	23
147	Yersinia pestis, KM260(11)-Δ0187	416	459	187
151	Yersinia pestis, 1146-25	408	271	111
61	Pseudomonas aerguinosa O7 7a,7b,7c	399	26	11
137	Providencia alcalifaciens O23	392	80	33
16	Serratia marcescens LPS	392	147	60
292	Shigella boydii type 14	372	50	20
240	Streptococcus pneumoniae type 34 (Danish type 10A)	370	47	19
125	Providencia stuartii O43	363	238	97
15	Salmonella typhimurium LPS	349	417	170
141	Providencia alcalifaciens O32	339	130	53
11	Salmonella typhimurium SL 684 (Rc mut)	337	84	34
21	Salmonella enteritidis LPS	333	126	51

311	Yeast Mannan Sigma M-3640	323	12	5
81	Proteus vulgaris O34 (4669)*	318	103	42
223	Escherichia coli O55:B5 LPS- solution at 1 mg/mL	316	14	6
172	Yersinia pestis, 11M-37	315	197	80
18	Escherichia coli O128-B12 LPS	313	20	8
139	Providencia alcalifaciens O29	306	236	96
312	1-2 Mannan Acetobacter methanolicus MB135	294	95	39
242	Streptococcus pneumoniae type 51 (Danish type 7F)	293	73	30
191	Klebsiella O5 OPS	293	79	32
120	Proteus myxofaciens O60	290	176	72
268	Proteus mirabilis O33 (D52)	283	16	7
294	Escherichia coli O85	282	23	9
194	Shigella boydii type 1	276	84	34
19	Salmonella enterica abortus equi LPS	273	120	49
20	Salmonella typhosa LPS	260	76	31
226	Streptococcus pneumoniae type 2 (Danish type 2)	250	107	44
184	Francisella novicida OPS	237	303	124
119	Proteus penneri O73a,b (103)	228	37	15
224	Escherichia coli O127:B8 LPS- solution at 1 mg/mL	223	123	50
135	Providencia alcalifaciens O19	220	19	8
174	Proteus vulgaris O24 (PrK 47/57)	220	158	65
296	Escherichia coli O145	215	41	17
248	Yersinia pestis KM218-6C	212	44	18
195	Shigella boydii type 3	209	227	93
207	Escherichia coli O61	207	15	6
12	Pseudomonas aerguinosa O10	201	213	87
298	Salmonella enterica O17	199	28	12
234	Streptococcus pneumoniae type 17 (Danish type 17F)	192	57	23
4	Pseudomonas aerguinosa O2 (Fisher immunotype 3)	185	166	68
228	Streptococcus pneumoniae type 4 (Danish type 4)	184	20	8
86	Proteus vulgaris O53 (TG 276-10)	183	40	16
230	Streptococcus pneumoniae type 8 (Danish type 8)	182	59	24
157	Yersinia pestis, KM218-25	182	67	27
88	Proteus vulgaris O55 (TG 155)	175	50	21
54	Pseudomonas aerguinosa O3 3a,3b	174	140	57
127	Providencia stuartii O47	168	54	22
152	Yersinia pestis 1146-25	166	28	11
59	Pseudomonas aerguinosa O6 6a,6c	165	145	59
47	Proteus penneri O71 (42)	160	19	8
255	Yersinia pseudotuberculosis O:2a-dhmA	159	13	5

278	Shigella flexneri type 2b	158	20	8
58	Pseudomonas aerguinosa O6 6a	151	15	6
13	Salmonella typhimurium dodeca saccharide	149	100	41
31	Escherichia coli O106	149	47	19
140	Providencia alcalifaciens O30	144	19	8
9	Salmonella typhimurium SL 11881 (Re mut)	143	67	27
222	Escherichia coli O26:B6 LPS- solution at 1 mg/mL	143	68	28
50	Pseudomonas aerguinosa O2 2a,2b	142	70	28
123	Providencia stuartii O18	140	52	21
231	Streptococcus pneumoniae type 9 (Danish type 9N)	140	9	4
40	Proteus mirabilis O50 (TG332)	139	147	60
104	Proteus mirabilis O74 (10705, OF)	139	90	37
239	Streptococcus pneumoniae type 26 (Danish type 6B)	136	15	6
238	Streptococcus pneumoniae type 23 (Danish type 23F)	136	58	24
48	Proteus penneri O72a, 72b (4)	134	61	25
78	Proteus vulgaris O21 (PrK 39/57)*	128	27	11
232	Streptococcus pneumoniae type 12 (Danish type 12F)	125	82	34
306	Davanat	125	30	12
44	Proteus penneri O64a, 64b, 64d (39)	122	180	74
145	Providencia rustigianii O16	122	74	30
113	Proteus penneri O63 (22)	121	122	50
246	Streptococcus pneumoniae type 68 (Danish type 9V)	121	51	21
276	Shigella flexneri type 1b	120	49	20
6	Pseudomonas aerguinosa O9 (9a, 9b, 9d)	119	84	34
122	Providencia stuartii O4	118	9	4
32	Escherichia coli O130	113	133	54
121	Proteus O56 (genomospecies 4)	113	48	20
169	Yersinia pestis, KIMD1-37	113	40	16
304	Escherichia coli K13	113	5	2
308	Yeast Mannan	111	5	2
307	Laminarin	109	8	3
245	Streptococcus pneumoniae type 57 (Danish type 19A)	109	42	17
42	Proteus mirabilis O57 (TG319)	109	101	41
289	Shigella boydii type 7	108	67	27
237	Streptococcus pneumoniae type 22 (Danish type 22F)	106	19	8
38	Proteus mirabilis O10 (HJ4320)	106	74	30
196	Shigella boydii type 5	105	36	15

271	Proteus mirabilis O49 (PrK 75/57)	103	14	6
51	Pseudomonas aerguinosa O2 2a,2b,2e	103	24	10
14	Salmonella enteritidis dodeca saccharide	103	20	8
235	Streptococcus pneumoniae type 19 (Danish type 19F)	102	14	6
92	Proteus mirabilis O13 (PrK 26/57)	101	25	10
25	Shigella dysennteriae type 3	100	49	20
112	Proteus penneri O62 (41)	95	29	12
205	Escherichia coli O52	94	17	7
193	Klebsiella O12 OPS	92	51	21
28	Shigella dysennteriae type 13	92	137	56
286	Shigella flexneri type X	91	22	9
1	Providencia stuartii O52	88	83	34
261	Proteus mirabilis O5 (PrK 12/57)	88	2	1
158	Yersinia pestis, KM218-25	87	57	23
280	Shigella flexneri type 3b	87	10	4
201	Shigella boydii type 16	87	28	12
167	Yersinia pestis, KM260(11)-25	86	48	20
115	Proteus penneri O65 (34)	86	22	9
243	Streptococcus pneumoniae type 54 (Danish type 15B)	86	7	3
227	Streptococcus pneumoniae type 3 (Danish type 3)	85	40	16
218	Shigella dysenteriae type 7	85	17	7
26	Shigella dysennteriae type 8 (batch 12)	83	32	13
197	Shigella boydii type 9	81	34	14
310	Galactomannan DAVANT (160102) Pro-Pharmacenti	79	5	2
210	Escherichia coli O118	79	5	2
279	Shigella flexneri type 3a	78	27	11
254	Yersinia pseudotuberculosis O:2a	78	9	4
106	Proteus mirabilis O77 (3 B-m)	77	13	5
154	Yersinia pestis, 1146-37	77	91	37
198	Shigella boydii type 11	77	16	7
126	Providencia stuartii O44	76	9	4
219	Shigella dysenteriae type 8 (Russian)	75	112	46
309	Escherichia coli O86	75	20	8
89	Proteus vulgaris O65 (TG 251)	75	12	5
110	Proteus penneri O59 (9)	75	8	3
80	Proteus vulgaris O25 (PrK 48/57)	74	14	6
293	Escherichia coli O71	74	10	4
131	Providencia alcalifaciens O5	73	29	12

3	Pseudomonas aerguinosa O1 (Fisher immunotype 4)	72	39	16
249	Yersinia pestis KM260(11)-yjhW-6C	72	10	4
96	Proteus mirabilis O23a,b,d (PrK 42/57)	72	27	11
305	Neisseria meningitidis Group C	69	12	5
62	Pseudomonas aerguinosa O7 7a,7b,7d	69	16	6
287	Shigella dysenteriae type 1	68	39	16
130	Providencia stuartii O57	68	29	12
203	Shigella boydii type 18	68	13	6
299	Salmonella enterica O28	68	24	10
273	Proteus penneri O73ac (75)	67	1	0
288	Shigella boydii type 6	66	11	5
49	Pseudomonas aerguinosa O2 (2a),2d,2f	66	6	3
170	Yersinia pestis, KIMD1-25	65	17	7
27	Shigella dysennteriae type 11	65	14	6
180	Yersinia pestis P-1680-25C	64	36	15
24	Shigella bodyii type10	63	3	1
182	Yersinia pestis I-2377-25C	63	22	9
303	Escherichia coli K5	60	12	5
199	Shigella boydii type 12	60	20	8
229	Streptococcus pneumoniae type 5 (Danish type 5)	59	10	4
164	Yersinia pestis, KM260(11)-ΔwaaQ	59	58	24
55	Pseudomonas aerguinosa O3 3a,3b,3c	58	8	3
290	Shigella boydii type 8	57	19	8
82	Proteus vulgaris O37a,b (PrK 63/57)	57	14	6
300	Salmonella enterica O47	56	12	5
5	Pseudomonas aerguinosa O13 (Sandvik serotype II)	55	69	28
264	Proteus penneri O17 (16)	55	11	5
105	Proteus mirabilis O75 (10702, OC)	55	30	12
138	Providencia alcalifaciens O27	55	6	2
64	Pseudomonas aerguinosa O10 10a,10b	54	15	6
128	Providencia stuartii O47, Core 9	54	14	6
53	Pseudomonas aerguinosa O2 Immuno 7	53	17	7
0	Providencia stuartii O49	53	38	15
206	Escherichia coli O58	53	8	3
90	Proteus mirabilis O6 (PrK 14/57)	52	4	1
212	Escherichia coli O151	51	13	5
216	Shigella dysenteriae type 5	51	10	4
156	Yersinia pestis, KM218-37	50	19	8
85	Proteus vulgaris O45 (4680)	50	10	4
208	Escherichia coli O73	50	4	2
162	Yersinia pestis, KM260(11)-Δ0186	50	16	7

220	Shigella dysenteriae type 9	49	11	5
57	Pseudomonas aerguinosa O4 4a,4c	49	16	7
202	Shigella boydii type 17	49	23	9
84	Proteus vulgaris O44 (PrK 67/57)	48	15	6
160	Yersinia pestis, KM260(11)-ΔpmrF	48	48	20
209	Escherichia coli O112ab	48	5	2
7	Pseudomonas aerguinosa O6a (Habs serotype6, fraction IIa)	47	19	8
2	Pseudomonas aerguinosa O4 (Habs serotype 4)	47	33	13
297	Escherichia coli O107	47	3	1
173	Proteus vulgaris O23a, 23b, 23c (CCUG 10701)	47	36	15
284	Shigella flexneri type 6a	47	9	4
204	Escherichia coli O49	47	13	5
146	Providencia rustigianii O34	47	14	6
211	Escherichia coli O125	47	4	2
168	Yersinia pestis, KM260(11)-37	47	8	3
213	Escherichia coli O168	46	12	5
257	Yersinia pseudotuberculosis O:3	45	6	2
189	Klebsiella O3 OPS	44	10	4
45	Proteus penneri O66 (2)	44	23	10
100	Proteus mirabilis O29a (PrK 52/57)	44	27	11
277	Shigella flexneri type 2a	42	4	2
282	Shigella flexneri type 4b	42	2	1
23	Shigella bodyii type4	41	4	2
108	Proteus penneri O52 (15)	41	8	3
111	Proteus penneri O61 (21)	41	8	3
83	Proteus vulgaris O37a,c (PrK 72/57)	40	9	4
270	Proteus vulgaris O47 (PrK 73/57)	39	14	6
109	Proteus penneri O58 (12)	39	18	7
259	Proteus vulgaris O2 (OX2)	38	7	3
52	Pseudomonas aerguinosa O2 2a,2d	37	52	21
8	Pseudomonas aerguinosa O6a (Habs serotype6, fraction IIb)	37	19	8
93	Proteus mirabilis O14a,14b (PrK 29/57)	37	7	3
253	Yersinia pseudotuberculosis 85pCad-20C	37	5	2
33	Escherichia coli O148	36	3	1
39	Proteus mirabilis O29a, 29b (2002)	35	16	7
133	Providencia alcalifaciens O19	34	30	12
101	Proteus mirabilis O40 (10703)	34	14	6
185	Francisella tularensis OPS	32	11	4
275	Shigella flexneri type 1a	32	0	0

114	Proteus penneri O64a,b,c (27)	31	9	4
301	Salmonella enterica O55	31	3	1
200	Shigella boydii type 15	30	9	4
70	Pseudomonas aerguinosa O15 15	29	27	11
56	Pseudomonas aerguinosa O3 3a,3d	29	15	6
77	Proteus vulgaris O19a (PrK 37/57)	28	21	9
217	Shigella dysenteriae type 6 SR-strain	28	45	19
150	Yersinia pestis, KM260(11)-Δrfe	27	20	8
283	Shigella flexneri type 5b	27	5	2
215	Shigella dysenteriae type 4	27	9	4
107	Proteus penneri O31a (26)	26	14	6
295	Escherichia coli O99	26	3	1
75	Proteus vulgaris O15 (PrK 30/57)	23	12	5
129	Providencia stuartii O49, Core 1	22	20	8
302	Escherichia coli K92	22	8	3
91	Proteus mirabilis O11 (PrK 24/57)	22	10	4
79	Proteus vulgaris O22 (PrK 40/57)	21	3	1
285	Shigella flexneri type 6	21	10	4
35	Escherichia coli O180	20	7	3
43	Proteus penneri O8 (106)	19	16	6
30	Escherichia coli O40	18	10	4
252	Yersinia pseudotuberculosis 85pCad-37C	16	7	3
250	Yersinia pestis KM260(11)-wabD/waaL	15	11	5
66	Pseudomonas aerguinosa O11 11a,11b	11	8	3
148	Yersinia pestis, KM260(11)-Δ0187	11	5	2
29	Escherichia coli O29	10	6	2
265	Proteus mirabilis O18 (PrK 34/57)	9	6	2
37	Proteus mirabilis O8 (TG326)	9	1	0
34	Escherichia coli O150	9	6	2
72	Proteus vulgaris O4 (PrK 9/57)	8	4	2
65	Pseudomonas aerguinosa O10 10a,10c	7	2	1
256	Yersinia pseudotuberculosis O:2c	6	6	2
267	Proteus penneri O31ab (28)	6	5	2
258	Yersinia pseudotuberculosis O:4b	4	2	1
263	Proteus mirabilis O11 (9B-m)	3	4	2
73	Proteus vulgaris O12 (PrK 25/57)	2	6	2
60	Pseudomonas aerguinosa O6 Immuno 1	1	2	1
76	Proteus vulgaris O17 (PrK 33/57)	1	3	1
71	Proteus vulgaris O1 (18984)*	0	6	2
74	Proteus vulgaris O13 (8344)	0	2	1
67	Pseudomonas aerguinosa O12 12	0	2	1

262	Proteus mirabilis O9 (PrK 18/57)	-1	3	1
269	Proteus mirabilis O43 (PrK 69/57)	-1	2	1
260	Proteus mirabilis O3ab (S1959)	-3	4	1
69	Pseudomonas aerguinosa O14 14	-3	4	2

Appendix Table 5: FULL MGM ARRAY DATA FOR XCGL-1

Chart #	Glycan	Average	STDEV	SEM
309	Escherichia coli O86	30064	2361	964
137	Providencia alcalifaciens O23	26047	7614	3109
243	Streptococcus pneumoniae type 54 (Danish type 15B)	18455	1239	506
197	Shigella boydii type 9	17537	1641	670
229	Streptococcus pneumoniae type 5 (Danish type 5)	17376	634	259
131	Providencia alcalifaciens O5	14378	2496	1019
238	Streptococcus pneumoniae type 23 (Danish type 23F)	13994	1213	495
226	Streptococcus pneumoniae type 2 (Danish type 2)	13443	5460	2229
225	Streptococcus pneumoniae type 1 (Danish type 1)	12722	1323	540
265	Proteus mirabilis O18 (PrK 34/57)	12341	1306	533
234	Streptococcus pneumoniae type 17 (Danish type 17F)	11517	1190	486
236	Streptococcus pneumoniae type 20 (Danish type 20)	11213	988	403
245	Streptococcus pneumoniae type 57 (Danish type 19A)	10972	1679	685
239	Streptococcus pneumoniae type 26 (Danish type 6B)	10823	1159	473
235	Streptococcus pneumoniae type 19 (Danish type 19F)	10292	901	368
240	Streptococcus pneumoniae type 34 (Danish type 10A)	9818	1356	553
244	Streptococcus pneumoniae type 56 (Danish type 18C)	8800	4344	1774
231	Streptococcus pneumoniae type 9 (Danish type 9N)	7935	404	165
228	Streptococcus pneumoniae type 4 (Danish type 4)	7866	1101	449
237	Streptococcus pneumoniae type 22 (Danish type 22F)	7761	903	369
242	Streptococcus pneumoniae type 51 (Danish type 7F)	7653	653	266
232	Streptococcus pneumoniae type 12 (Danish type 12F)	7473	1553	634
246	Streptococcus pneumoniae type 68 (Danish type 9V)	5835	737	301
103	Proteus mirabilis O51 (19011)*	5556	1527	624
241	Streptococcus pneumoniae type 43 (Danish type 11A)	5547	884	361
233	Streptococcus pneumoniae type 14 (Danish type 14)	4363	6247	2550
24	Shigella bodyii type10	3505	2149	877

281	Shigella flexneri type 4a	3348	364	148
247	Streptococcus pneumoniae type 70 (Danish type 33F)	3341	460	188
199	Shigella boydii type 12	3080	2359	963
187	Klebsiella O2a OPS	2252	701	286
212	Escherichia coli O151	2019	171	70
230	Streptococcus pneumoniae type 8 (Danish type 8)	1627	260	106
94	Proteus mirabilis O16 (4652)	1609	499	204
0	Providencia stuartii O49	1608	1885	770
126	Providencia stuartii O44	1526	112	46
31	Escherichia coli O106	1482	1677	685
27	Shigella dysennteriae type 11	1344	189	77
95	Proteus mirabilis O17 (PrK 32/57)	1327	199	81
55	Pseudomonas aerguinosa O3 3a,3b,3c	1321	309	126
117	Proteus penneri O68 (63)	1228	34	14
278	Shigella flexneri type 2b	1207	413	168
132	Providencia alcalifaciens O6*	1181	112	46
251	Yersinia pestis KM260(11)-wabC/waaL	1154	217	88
213	Escherichia coli O168	1127	299	122
127	Providencia stuartii O47	1124	122	50
125	Providencia stuartii O43	1117	146	60
140	Providencia alcalifaciens O30	1069	102	42
28	Shigella dysennteriae type 13	1039	1475	602
36	Proteus mirabilis O3a, 3c (G1)	1006	236	96
99	Proteus mirabilis O28 (PrK 51/57)	986	165	67
190	Klebsiella O4 OPS	963	287	117
188	Klebsiella O2ac OPS	937	77	32
82	Proteus vulgaris O37a,b (PrK 63/57)	909	203	83
11	Salmonella typhimurium SL 684 (Rc mut)	907	1327	542
39	Proteus mirabilis O29a, 29b (2002)	871	146	60
18	Escherichia coli O128-B12 LPS	869	50	21
276	Shigella flexneri type 1b	826	366	149
102	Proteus mirabilis O41 (PrK 67/57)	822	557	228
202	Shigella boydii type 17	787	24	10
136	Providencia alcalifaciens O21	739	49	20
191	Klebsiella O5 OPS	721	188	77
268	Proteus mirabilis O33 (D52)	700	77	32
98	Proteus mirabilis O27 (PrK 50/57)	684	817	334
73	Proteus vulgaris O12 (PrK 25/57)	674	568	232
47	Proteus penneri O71 (42)	674	22	9
46	Proteus penneri O69 (25)	669	79	32

144	Providencia rustigianii O14	654	107	44
215	Shigella dysenteriae type 4	654	309	126
171	Yersinia pestis, 11M-25	641	149	61
64	Pseudomonas aerguinosa O10 10a,10b	629	439	179
116	Proteus penneri O67 (8)	628	439	179
97	Proteus mirabilis O26 (PrK 49/57)	611	569	232
258	Yersinia pseudotuberculosis O:4b	593	312	127
288	Shigella boydii type 6	588	484	198
61	Pseudomonas aerguinosa O7 7a,7b,7c	581	13	5
295	Escherichia coli O99	513	27	11
124	Providencia stuartii O20*	511	59	24
43	Proteus penneri O8 (106)	476	562	229
293	Escherichia coli O71	460	87	35
42	Proteus mirabilis O57 (TG319)	459	155	63
254	Yersinia pseudotuberculosis O:2a	455	77	31
274	Proteus vulgaris O76 (HSC438)	449	63	26
78	Proteus vulgaris O21 (PrK 39/57)*	447	59	24
161	Yersinia pestis, KM260(11)-Δ0186	414	64	26
210	Escherichia coli O118	409	74	30
135	Providencia alcalifaciens O19	407	147	60
80	Proteus vulgaris O25 (PrK 48/57)	398	16	7
40	Proteus mirabilis O50 (TG332)	394	204	83
214	Shigella dysenteriae type 2	390	84	34
292	Shigella boydii type 14	388	189	77
34	Escherichia coli O150	384	80	33
198	Shigella boydii type 11	382	47	19
206	Escherichia coli O58	382	40	16
134	Providencia alcalifaciens O19	361	60	25
35	Escherichia coli O180	360	251	103
65	Pseudomonas aerguinosa O10 10a,10c	358	66	27
20	Salmonella typhosa LPS	356	134	55
72	Proteus vulgaris O4 (PrK 9/57)	344	229	93
283	Shigella flexneri type 5b	340	85	35
260	Proteus mirabilis O3ab (S1959)	331	64	26
69	Pseudomonas aerguinosa O14 14	330	156	64
296	Escherichia coli O145	326	46	19
17	Escherichia coli K235 LPS	320	81	33
291	Shigella boydii type 13	310	77	31
143	Providencia alcalifaciens O39	295	61	25
81	Proteus vulgaris O34 (4669)*	294	12	5
165	Yersinia pestis, KM260(11)-ΔwaaL	291	63	26

166	Yersinia pestis, KM260(11)-25	291	23	9
59	Pseudomonas aerguinosa O6 6a,6c	290	141	58
195	Shigella boydii type 3	281	60	25
153	Yersinia pestis, 1146-37	276	37	15
221	Escherichia coli O111:B4 LPS- solution at 1 mg/mL	271	17	7
311	Yeast Mannan Sigma M-3640	268	171	70
91	Proteus mirabilis O11 (PrK 24/57)	268	85	35
142	Providencia alcalifaciens O36*	268	19	8
208	Escherichia coli O73	267	55	22
266	Proteus mirabilis O20 (PrK 38/57)	256	18	7
129	Providencia stuartii O49, Core 1	254	88	36
300	Salmonella enterica O47	253	48	20
53	Pseudomonas aerguinosa O2 Immuno 7	247	64	26
259	Proteus vulgaris O2 (OX2)	243	38	15
108	Proteus penneri O52 (15)	242	7	3
118	Proteus penneri O70 (60)	242	93	38
16	Serratia marcescens LPS	239	22	9
289	Shigella boydii type 7	232	42	17
194	Shigella boydii type 1	227	79	32
62	Pseudomonas aerguinosa O7 7a,7b,7d	226	60	24
201	Shigella boydii type 16	225	39	16
2	Pseudomonas aerguinosa O4 (Habs serotype 4)	220	14	6
255	Yersinia pseudotuberculosis O:2a-dhmA	218	15	6
33	Escherichia coli O148	215	58	24
189	Klebsiella O3 OPS	215	13	5
23	Shigella bodyii type4	209	32	13
155	Yersinia pestis, KM218-37	208	51	21
4	Pseudomonas aerguinosa O2 (Fisher immunotype 3)	206	9	4
128	Providencia stuartii O47, Core 9	205	44	18
119	Proteus penneri O73a,b (103)	196	38	15
120	Proteus myxofaciens O60	193	46	19
218	Shigella dysenteriae type 7	193	13	5
9	Salmonella typhimurium SL 11881 (Re mut)	192	34	14
159	Yersinia pestis, KM260(11)-ΔpmrF	192	33	13
220	Shigella dysenteriae type 9	186	10	4
222	Escherichia coli O26:B6 LPS- solution at 1 mg/mL	186	12	5
76	Proteus vulgaris O17 (PrK 33/57)	183	66	27
145	Providencia rustigianii O16	183	25	10
10	Salmonella typhimurium TV 119 (Ra mut)	183	29	12
58	Pseudomonas aerguinosa O6 6a	182	36	15
175	Yersinia pestis KM260(11)-6C	180	9	3

176	Yersinia pestis 260(11)-37C-186	180	23	10
216	Shigella dysenteriae type 5	177	7	3
29	Escherichia coli O29	175	200	82
219	Shigella dysenteriae type 8 (Russian)	174	29	12
146	Providencia rustigianii O34	174	202	83
174	Proteus vulgaris O24 (PrK 47/57)	173	192	78
63	Pseudomonas aerguinosa O7 7a,7d	172	9	4
56	Pseudomonas aerguinosa O3 3a,3d	171	13	5
205	Escherichia coli O52	171	17	7
261	Proteus mirabilis O5 (PrK 12/57)	168	13	5
152	Yersinia pestis 1146-25	168	38	15
279	Shigella flexneri type 3a	167	43	18
107	Proteus penneri O31a (26)	162	23	10
6	Pseudomonas aerguinosa O9 (9a, 9b, 9d)	161	20	8
112	Proteus penneri O62 (41)	161	72	29
86	Proteus vulgaris O53 (TG 276-10)	160	12	5
183	Yersinia pestis I-2377-37C	160	26	11
71	Proteus vulgaris O1 (18984)*	156	25	10
262	Proteus mirabilis O9 (PrK 18/57)	153	53	22
26	Shigella dysennteriae type 8 (batch 12)	152	76	31
93	Proteus mirabilis O14a,14b (PrK 29/57)	149	8	3
133	Providencia alcalifaciens O19	149	9	4
249	Yersinia pestis KM260(11)-yjhW-6C	149	43	18
141	Providencia alcalifaciens O32	147	10	4
305	Neisseria meningitidis Group C	145	5	2
223	Escherichia coli O55:B5 LPS- solution at 1 mg/mL	145	24	10
204	Escherichia coli O49	144	44	18
177	Yersinia pestis 260(11)-37C-187	139	23	9
178	Yersinia pestis 260(11)-37C-416	139	102	42
269	Proteus mirabilis O43 (PrK 69/57)	139	82	34
88	Proteus vulgaris O55 (TG 155)	137	11	5
308	Yeast Mannan	137	28	11
68	Pseudomonas aerguinosa O13 13a,13c	135	158	65
181	Yersinia pestis P-1680-37C	132	12	5
211	Escherichia coli O125	130	25	10
285	Shigella flexneri type 6	127	15	6
272	Proteus mirabilis O54ab (OE)	126	34	14
157	Yersinia pestis, KM218-25	124	48	19
172	Yersinia pestis, 11M-37	122	90	37
25	Shigella dysennteriae type 3	122	86	35
45	Proteus penneri O66 (2)	121	76	31

149	Yersinia pestis, KM260(11)-Δrfe	121	17	7
277	Shigella flexneri type 2a	119	6	2
253	Yersinia pseudotuberculosis 85pCad-20C	119	11	4
54	Pseudomonas aerguinosa O3 3a,3b	118	14	6
271	Proteus mirabilis O49 (PrK 75/57)	117	12	5
248	Yersinia pestis KM218-6C	116	71	29
50	Pseudomonas aerguinosa O2 2a,2b	115	37	15
21	Salmonella enteritidis LPS	115	12	5
151	Yersinia pestis, 1146-25	114	88	36
267	Proteus penneri O31ab (28)	114	53	22
74	Proteus vulgaris O13 (8344)	112	191	78
22	Shigella bodyii type2	112	16	6
203	Shigella boydii type 18	110	16	7
32	Escherichia coli O130	109	14	6
130	Providencia stuartii O57	108	3	1
256	Yersinia pseudotuberculosis O:2c	107	16	6
224	Escherichia coli O127:B8 LPS- solution at 1 mg/mL	104	30	12
207	Escherichia coli O61	103	8	3
48	Proteus penneri O72a, 72b (4)	103	25	10
49	Pseudomonas aerguinosa O2 (2a),2d,2f	102	26	11
106	Proteus mirabilis O77 (3 B-m)	101	15	6
192	Klebsiella O8 OPS	100	11	5
286	Shigella flexneri type X	100	7	3
182	Yersinia pestis I-2377-25C	100	120	49
263	Proteus mirabilis O11 (9B-m)	100	29	12
109	Proteus penneri O58 (12)	98	10	4
52	Pseudomonas aerguinosa O2 2a,2d	97	185	76
121	Proteus O56 (genomospecies 4)	97	19	8
104	Proteus mirabilis O74 (10705, OF)	96	27	11
186	Klebsiella O1 OPS	95	18	8
227	Streptococcus pneumoniae type 3 (Danish type 3)	95	9	4
163	Yersinia pestis, KM260(11)-ΔwaaQ	94	65	26
37	Proteus mirabilis O8 (TG326)	94	9	4
15	Salmonella typhimurium LPS	91	60	24
264	Proteus penneri O17 (16)	91	18	7
154	Yersinia pestis, 1146-37	90	62	25
304	Escherichia coli K13	90	6	3
122	Providencia stuartii O4	87	10	4
90	Proteus mirabilis O6 (PrK 14/57)	87	16	7
139	Providencia alcalifaciens O29	85	7	3
200	Shigella boydii type 15	84	9	4

179	Yersinia pestis 260(11)-37C-417	82	13	5
41	Proteus mirabilis O54a, 54b (10704)	82	14	6
89	Proteus vulgaris O65 (TG 251)	78	11	4
92	Proteus mirabilis O13 (PrK 26/57)	76	15	6
1	Providencia stuartii O52	76	95	39
51	Pseudomonas aerguinosa O2 2a,2b,2e	74	32	13
169	Yersinia pestis, KIMD1-37	74	69	28
275	Shigella flexneri type 1a	74	5	2
290	Shigella boydii type 8	73	25	10
12	Pseudomonas aerguinosa O10	71	7	3
209	Escherichia coli O112ab	71	36	15
294	Escherichia coli O85	71	5	2
38	Proteus mirabilis O10 (HJ4320)	69	8	3
284	Shigella flexneri type 6a	68	11	4
138	Providencia alcalifaciens O27	67	15	6
87	Proteus vulgaris O54a,54c (TG 103)	66	4	2
111	Proteus penneri O61 (21)	66	9	4
79	Proteus vulgaris O22 (PrK 40/57)	65	7	3
8	Pseudomonas aerguinosa O6a (Habs serotype6, fraction IIb)	62	33	13
147	Yersinia pestis, KM260(11)-Δ0187	61	72	29
101	Proteus mirabilis O40 (10703)	61	16	7
193	Klebsiella O12 OPS	60	16	7
273	Proteus penneri O73ac (75)	59	6	2
19	Salmonella enterica abortus equi LPS	58	41	17
170	Yersinia pestis, KIMD1-25	58	41	17
7	Pseudomonas aerguinosa O6a (Habs serotype6, fraction IIa)	58	8	3
57	Pseudomonas aerguinosa O4 4a,4c	58	5	2
158	Yersinia pestis, KM218-25	58	37	15
70	Pseudomonas aerguinosa O15 15	56	41	17
180	Yersinia pestis P-1680-25C	55	75	30
252	Yersinia pseudotuberculosis 85pCad-37C	55	36	15
83	Proteus vulgaris O37a,c (PrK 72/57)	54	6	2
257	Yersinia pseudotuberculosis O:3	54	4	2
298	Salmonella enterica O17	53	6	2
217	Shigella dysenteriae type 6 SR-strain	52	15	6
150	Yersinia pestis, KM260(11)-Δrfe	51	57	23
96	Proteus mirabilis O23a,b,d (PrK 42/57)	50	9	4
115	Proteus penneri O65 (34)	50	8	3
14	Salmonella enteritidis dodeca saccharide	50	16	7

13	Salmonella typhimurium dodeca saccharide	50	18	7
184	Francisella novicida OPS	49	17	7
282	Shigella flexneri type 4b	48	20	8
168	Yersinia pestis, KM260(11)-37	47	31	12
299	Salmonella enterica O28	47	5	2
164	Yersinia pestis, KM260(11)-ΔwaaQ	46	51	21
160	Yersinia pestis, KM260(11)-ΔpmrF	44	45	19
303	Escherichia coli K5	44	4	2
156	Yersinia pestis, KM218-37	43	55	22
270	Proteus vulgaris O47 (PrK 73/57)	43	9	4
310	Galactomannan DAVANT (160102) Pro-Pharmacenti	42	9	4
113	Proteus penneri O63 (22)	41	56	23
306	Davanat	41	24	10
123	Providencia stuartii O18	40	21	9
280	Shigella flexneri type 3b	39	9	4
148	Yersinia pestis, KM260(11)-Δ0187	37	42	17
77	Proteus vulgaris O19a (PrK 37/57)	36	25	10
110	Proteus penneri O59 (9)	36	5	2
301	Salmonella enterica O55	36	5	2
312	1-2 Mannan Acetobacter methanolicus MB135	34	8	3
85	Proteus vulgaris O45 (4680)	34	5	2
30	Escherichia coli O40	34	3	1
196	Shigella boydii type 5	33	18	7
100	Proteus mirabilis O29a (PrK 52/57)	32	15	6
173	Proteus vulgaris O23a, 23b, 23c (CCUG 10701)	31	21	9
297	Escherichia coli O107	31	7	3
84	Proteus vulgaris O44 (PrK 67/57)	29	30	12
75	Proteus vulgaris O15 (PrK 30/57)	27	21	8
3	Pseudomonas aerguinosa O1 (Fisher immunotype 4)	27	3	1
167	Yersinia pestis, KM260(11)-25	27	29	12
105	Proteus mirabilis O75 (10702, OC)	25	18	7
162	Yersinia pestis, KM260(11)-Δ0186	24	27	11
114	Proteus penneri O64a,b,c (27)	24	1	1
5	Pseudomonas aerguinosa O13 (Sandvik serotype II)	23	4	2
67	Pseudomonas aerguinosa O12 12	21	36	15
185	Francisella tularensis OPS	21	3	1
307	Laminarin	15	8	3
287	Shigella dysenteriae type 1	15	18	7
302	Escherichia coli K92	15	4	2
250	Yersinia pestis KM260(11)-wabD/waaL	14	26	11

60	Pseudomonas aerguinosa O6 Immuno 1	9	17	7
66	Pseudomonas aerguinosa O11 11a,11b	1	2	1
44	Proteus penneri O64a, 64b, 64d (39)	-2	3	1