DEVELOPING STRATEGIES TO COMBAT *S. PNEUMONIAE*: TARGETING THE CAPSULE AND THE ADAPTIVE IMMUNE SYSTEM

by

PAETON LEIGH ANN WANTUCH

(Under the Direction of Fikri Avci)

ABSTRACT

Streptococcus pneumoniae (*Spn*) is a major human pathogen and a threat to public health with its ability to cause invasive pneumococcal diseases (IPD). Currently, there are over 100 different serotypes of *Spn* characterized by the distinct capsular polysaccharide (CPS) structures. These CPS are targeted for inclusion in glycoconjugate vaccines to combat infections. A [glyco]conjugate vaccine is comprised of a carrier protein covalently linked with the CPS of a bacterium. The current pneumococcal conjugate vaccine on the market, PCV13 is a multimeric vaccine utilizing CPSs isolated from 13 different highly invasive *Spn* serotyes and conjugated with a carrier protein (i.e., CRM197 or tetanus toxoid (TT)). PCV13 has seen great success in curving disease; however, limited effects have been observed in the elderly and immunocompromised populations as well as low or variable protection for the included serotypes. This demonstrates a need for continued research into conjugate vaccine design, mechanism of immune activation and novel methods for fighting IPD. In this dissertation, I provide evidence for adaptive immune activation by glycoconjugates utilizing the major histocompatibility class II (MHCII) pathway. Additionally, a new design for

glycoconjugates utilizing carrier peptides linked with defined lengths of oligosaccharides over the traditional full protein linked to full polysaccharide. I defined a set of human derived carrier peptides discovered from clinically relevant carrier proteins to be used in conjugate strategies. Finally, I elucidated the mechanism of action for the glycoside hydrolase Pn3Pase to aid in its use as a therapeutic against highly virulent type 3 *Spn* infections. These studies increase our knowledge and understanding of the mechanism of action of glycoconjugate vaccines to aid in new therapeutics and knowledge-based vaccine design to help battle IPD.

INDEX WORDS: *Streptococcus pneumoniae*, Capsular Polysaccharide, Carrier Protein, Adaptive Immunity, MHCII, T Cell Receptor, Glycoconjugate Vaccines, Glycoside Hydrolase

DEVELOPING STRATEGIES TO COMBAT *S. PNEUMONIAE*: TARGETING THE CAPSULE AND THE ADAPTIVE IMMUNE SYSTEM

by

PAETON LEIGH ANN WANTUCH

B.S., University of Wisconsin-Whitewater, 2014

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

DOCTOR OF PHILOSOPHY

ATHENS, GEORGIA

2020

DEVELOPING STRATEGIES TO COMBAT *S. PNEUMONIAE*: TARGETING THE CAPSULE AND THE ADAPTIVE IMMUNE SYSTEM

by

PAETON LEIGH ANN WANTUCH

Major Professor:	Fikri Avci
Committee:	Michael Tiemeyer
	Zachary Wood
	Robert Woods

Electronic Version Approved:

Ron Walcott
Interim Dean of the Graduate School
The University of Georgia
August 2020

DEDICATION

To my Dad for always believing me in and the endless love and support throughout my life.

To Capone for being my constant companion.

To my family at Trinity Baptist Church for being my home away from home.

To Elizabeth Malone for being a friend never forgotten.

To my strong foundation through faith; without this I wouldn't have been able to weather any storm. For the fruit of the Spirit is love, joy, peace, longsuffering, gentleness, goodness and faith. Finally, whatsoever things are true, whatsoever things are honest, whatsoever things are just, whatsoever things are pure, whatsoever things are lovely, whatsoever things are of good report; if there be any virtue think of these things.

Proverbs 3:5-6
Phillipians 4:13

ACKNOWLEDGEMENTS

I would like to first and foremost thank my advisor, Fikri Avci for his continued guidance and mentorship. Without his constant support and leadership, I would not be where I am today. Thank you for always pushing me towards greatness and believing in me. Your passion for science and dedication to the lab are inspiring.

I would also like to thank the Avci lab members: Dr. Lina Sun, Dr. Amy Paschall, Dr. Dustin Middleton, Ahmet Ozdilek, Jeremy Duke and Javid Aceil for their dedication to science and hard work in the lab. Thank you all for making my past six years in the Avci lab unforgettable and more importantly for your friendship.

I would like to give a special thanks to my family for their continued support, love and encouragement. Thank you for always believing in me.

I would like to thank my undergraduate mentors Dr. Robert Kuzoff and Dr. Christopher Veldkamp. Without you two I would have never dreamed of pursuing graduate school. Thank you for steering me towards research and always believing I was capable of more than I thought. Thank you for getting me started in my research path.

Finally, to my friends and colleagues in the CMM, CCRC, BCMB and my collaborators (particularly the Linhardt and Woods lab for their continued contributions to my research) and committee members: thank you for making graduate school a memorable experience and helping me grow and learn along the way.

TABLE OF CONTENTS

Page

APPENDICES

LIST OF TABLES

Page

LIST OF FIGURES

Page

CHAPTER 1

FORWARD AND LITERATURE REVIEW

The goal of this dissertation has been to elucidate the molecular mechanisms of glycoconjugate vaccine activation of the adaptive immune system. This information will be used to address the short comings in current glycoconjugate vaccine production strategies as well as work towards alternative prophylactic and therapeutic approaches to combat invasive pneumococcal diseases (IPD).

Chapter 1 discusses the antigen processing and presentation pathway leading to activation of the adaptive immune system. Further, this chapter discusses the interaction components MHCII and the T cell receptor (TCR) of CD4+ T cell in-depth and the roles they play in the immune response. Lastly, this chapter explores carbohydrates as T cell antigens and antigen presentation of glycoconjugates.

Chapter 2 examines the current state of glycoconjugate vaccines against IPD. It explores the underlying cause of disease as well as the various diseases attributed to pneumococcal bacteria. Additionally, this chapter covers the existing vaccines, successes, and potential shortcomings of the vaccines.

Chapter 3 follows closely with chapter 2 and discusses two serotypes of *Streptococcus pneumoniae* that continue to cause disease despite their inclusion in the current glycoconjugate vaccine on the market. Type 3 is of particular interest to my work and is continued to be explored in the subsequent chapters.

Chapter 4 uses a model glycoconjugate vaccine against type 3 *Streptococcus pneumoniae* prepared using oligosaccharides from the capsular polysaccharide of the bacteria to understand the molecular mechanisms of immune activation. I provide evidence that the glycoconjugates bind with purified MHCII molecules as well as the T cell receptor binding the glycan portion of the conjugate suggesting these glycoconjugates are following the glycan dependent model of T cell activation.

Chapter 5 investigates MHCII binding and T cell activating peptides from two proteins that are used in glycoconjugate vaccines on the market today. The goal of this work was to identify peptides binding human MHCII and activating human T cells to be applied to the strategy from chapter 4. I provide a list of human peptides that both bind a variety of human MHCII alleles and activate human T cells from healthy adult donors that express a set of MHCII alleles representative of a large portion of the population.

Chapter 6 explores the glycoside hydrolase Pn3Pase for its uses as a therapeutic against highly virulent type 3 pneumonia infection. Pn3Pase is produced by *Paenibacillus sp. 32352*, but exhibits depolymerization activity against the capsular polysaccharide of type 3 *Streptococcus pneumoniae* (Pn3P). I identified the active residues in this enzyme, domains important for function, determined the mechanism of action, and Pn3Pase affinity for Pn3P substrate.

Chapter 7 summarizes all the works and covers the implications of these discoveries. Additionally, it offers future directions for the research into glycoconjugate immune activation and implications. The co-author manuscript that is of direct relevance to my dissertation and that I contributed to is in appendix A. Co-author works that are not

directly relevant to my doctoral dissertation, but I worked on through collaborations are listed below.

1. Sun L, Paschall AV, Middleton DR, Ishihara M, Ozdilek A, **Wantuch PL**, Aceil J, Duke JA, LaBranche CC, Tiemeyer M, Avci FY. 2020. Glyopeptide epitope facilitates HIV-1 envelope specific humoral immune responses by eliciting T cell help. ***Nature Communications*.** 2020
2. Veldkamp CT, Kiermaier E, Gabel-Eissens SJ, Gillitzer ML, Lippner DR, DiSilvio FA, Mueller CJ, **Wantuch PL,** Chaffee GR, Famiglietti MW, Zgoba DM, Bailey AA, Bah Y, Engebretson SJ, Graupner DR, Lackber ER, LaRose VD, Medeiros T, Olson ML, Phillips AJ, Pyles H, Richard AM, Schoeller SJ, Touzeau B, Williams LG, Sixt M, Peterson FC. Solution structure of CCL19 and identifications of overlapping CCR7 and PSGL-1 binding sites. ***Biochemistry***. 2015, 54(27):4163-4166.
3. Middleton DR, Zhang X, **Wantuch PL,** Ozdilek A, Liu X, Lopilato R, Gangasani N, Bridger R, Wells L, Linhardt RJ, and Avci FY. Identification of the *Streptococcus pneumoniae* type 3 capsule-specific glycosyl hydrolase gene of *Paenibacillus sp.* 32352. ***Glycobiology***. 2018, 28(2):90-99.

My contribution to work (1) was developing the MHCII binding assay utilized and synthesizing/preparing all the peptides utilized in the study. My contribution to work (2) was protein synthesis and purification, 2D NMR, data analysis and manuscript writing. To work (3), I performed the kinetic analysis, protein purification and manuscript writing.

Literature Review

Adaptive Immune Responses

The host immune response to foreign invaders can simplistically be viewed as having two lines of defense, the innate response and adaptive immune response. Innate immunity is often referred to as the hosts' first line of defense (1,2). This defense mechanism is initiated immediately, or within hours, of exposure to a pathogen or antigen. Generally, the innate immune response comprises four types of barriers: anatomic, physiologic, endocytic and phagocytic (1). Since this immune response acts rapidly, it is antigen independent and unable to generate immune memory to the invading pathogen (1). Adaptive immunity, on the other hand, is highly antigen specific leading to long lasting immune memory to aid the host in continued defense against pathogens (2). While, these two immune responses are fundamentally different in mechanism of action, both work synergistically and are required for a fully functioning immune system.

The adaptive arm of the immune system utilizes antigen specific B and T lymphocytes to generate an immune response (2,3) **(Figure 1.1)**. Briefly, antigen specific receptors on the surface of B lymphocytes bind antigen and present it to T lymphocytes through the major histocompatibility complex class II (MHCII) protein (4). The T cell receptor (TCR) on the surface of the T lymphocyte binds the antigen which in turn stimulates the cell. This stimulation of the T lymphocyte provides cognate help to activate the B lymphocyte to produce antibodies and be further differentiated (3) **(Figure 1.1)**. The first step in generating this adaptive humoral (antibody dependent) immune response is referred to as antigen processing and presentation and will be covered in greater detail below. However, it is apparent that an effective immune response is

generated through this tri-molecular complex consisting of: antigen, MHCII, and TCR **(Figure 1.1)**. Further, these three components must function together as the antigen must bind MHCII and the TCR must be able to bind the antigen-MHCII complex. Thus, understanding the molecular mechanisms of interaction between these three components is crucial for our understanding of the immune response. These next sections cover MHCII proteins and TCRs in greater detail.

Major Histocompatibility Complex Class II

Major histocompatibility (MHC) proteins are a central part of generating an immune response. There are two classes of these proteins, class I and class II (MHCII). Both share similar functions which is to present antigens at the cell surface to T lymphocytes (5). However, the two proteins differ in the nature of the antigens they present. While class I proteins mostly present antigens derived from intracellular sources to CD8+ T cells, class II present exogenous antigens to CD 4+ T cells (5,6). These two proteins also differ in their binding grooves. Class I grooves are closed at both ends leading to the presentation of smaller peptides usually 8-10 residues in length (7). MHCII proteins, on the other hand, have binding grooves that are open at both ends accommodating longer peptides generally up to 25 amino acids (8-10) **(Figure 1.2)**. Additionally, while class I proteins are expressed on nearly every cell, class II proteins are limited to antigen presenting cells (APCs) such as dendritic cells, macrophages and B cells (5,11).

In humans there are three genetically distinct isotypes of MHCII proteins: HLA-DR, HLA-DQ and HLA-DP, while mice have two I-A and I-E (11). In humans these proteins are encoded by the highly polymorphic human leukocyte antigen (HLA) genes,

HLA-DR, *–DP* and *–DQ* (8). MHCII molecules are heterodimers consisting of two membrane inserted chains, α and β (12) **(Figure 1.2A)**. The peptide binding cleft of these proteins is formed by a β-sheet floor topped with two parallel α-helices, one from each subunit of the protein (13) **(Figure 1.2B)**. In some instances these α-helices may also run antiparallel (11). Peptide binding by MHCII molecules influences its conformation, stability, and surface appearance (10,14). Additionally, it has been noted that empty MHCII molecules are unstable and tend to aggregate (5,15).

These proteins possess the remarkable ability to bind a vast array of peptides to any given MHCII molecule forming a natural peptide library. While this topic has been the scope of many immunological research projects over the years, the mechanism by which MHCII proteins present antigens is still an active area of investigation (5). However, many works have been done to understand peptide binding with MHCII. Crystal structure analysis of peptide bound MHCII proteins reveals a mostly uniform mode of binding in which side chains of amino acids interact with pockets in the MHCII molecule (11,16-20). Additionally, it is thought peptides bind in a conserved N- to C-terminal orientation following two criteria: hydrogen bonds form between the MHCII protein and peptide backbone and four to five peptide side chains bind in the MHCII pockets (21,22).

It is generally accepted there are four major binding pockets: P1, P4, P6 and P9 with the numbering referring to the peptide side chain residing within the pocket (23) **(Figure 1.3)**. However, P7 is sometimes referenced as a major binding pocket as well (12). The peptide is bound in an extended conformation taking a fold similar to a polyproline II helix (24). During binding up to seventeen hydrogen bonds may form

between conserved residues on MHCII and the backbone of the peptide (19). Peptide side chains that are not residing within binding pockets are positioned outward where they can then interact with the T cell receptor (23) **(Figure 1.3)**. This hydrogen binding array allows MHCII to exhibit highly promiscuous binding (23). Given the high binding promiscuity of these proteins and the lack of a concrete consensus binding sequence, there has been evidence of peptides binding MHCII in different registers. That is, the same peptide binding MHCII only differing in the positions of the side chains within the MHC binding pockets (23,25).

Generally, there is little known about an exact binding motif or consensus sequence for peptides binding with MHCII. However, some work has been done to shed light on this matter. One study indicates that a stretch of nine amino acids may be critical for binding (12). Additionally, the side chain residue fitting into pocket P1 is almost always a hydrophobic anchor (12,26). However, beyond this binding requirements vary between alleles of MHCII. For example, HLA-DR1, -DR2, -DR3, and –DR4 prefers anchor spacing at P1, P4, P6, P7 and P9, while other alleles such as –DR5, DP4, and DQ7 bind with anchors at P3 and P5 (12). Another example of this difference at the residue level is HLA-DR1 prefers hydrophobic residues at P9, but HLA-DR4 prefers negatively charged residues (12). These allelic changes make it difficult to pin-point an exact binding preference for these molecules. Nevertheless, one study did look fourteen different peptides binding with the mouse allele I-A^d and noticed what appeared to be a common binding motif. Similar to above, the first position in P1 appeared to always be hydrophobic, next is typically a basic or polar residue, followed by two hydrophobics (P3 and P4), position five was variable, and position six was strictly alanine or serine (26).

These studies may point towards the anchor residues residing in the binding pockets to be hydrophobic. Lastly, MHCII molecules are able to bind antigens other than peptides. A number of studies have demonstrated MHCII unique ability to bind super antigens in a variety of different conformations (27-29).

In addition to the antigen presenting –DR, -DQ, -DP molecules there are two other MHCII proteins that aid in peptide binding and epitope selection. These proteins are HLA-DO and HLA-DM (30). To understand the roles of these two chaperones it is first important to understand the pathway of antigen processing and presentation.

Antigen Processing and Presentation

The central step in generating an adaptive immune response is T cell activation, which is induced by the T cell receptor recognizing antigen's epitopes bound with MHCII. How these epitopes are generated and antigen becomes associated with MHCII is through a process known as antigen processing and presentation. This process can be summarized simply into six steps: 1) Acquisition of the antigenic molecule by the APC; 2) targeting the antigenic molecule for destruction; 3) proteolysis of the antigenic molecule in the endosome; 4) delivery of the resulting antigenic peptides to the endolysome; 5) binding of the peptides with MHCII; and 6) present the peptide-MHCII complex on the surface of the APC (31) **(Figure 1.4)**.

Steps one through three above are all specific to the type of antigen presenting cell involved. Each type of cell: B cell, macrophage or dendritic cell (DC) acquires antigens and degrades them in its own process. B cells utilize an antigen specific B cell receptor (BCR) on the cell surface to capture antigen (Ag) (30,32). The BCR-Ag complex is then internalized into early endosomes (33). Macrophages, which serve as a

link between adaptive and innate immune response, are recruited to peripheral tissues during pathogen exposure and acquire antigens through phagocytosis (34) which are delivered to the phagosome (35). Dendritic cells circulate peripheral tissues for immune invaders and use endocytosis to acquire antigens derived from the pathogens (36) which are then delivered to endosomal/lysosomal multivesicular bodies (MVBs) (37). Both macrophages are dendritic cells require activation by mediators such as IFN-γ and GM-CSF (macrophages) or LPS (DCs) to induce higher expression levels of MHCII (38,39). Once the antigens are in their relevant compartments in each cell type, they are degraded in much the same way. Each APC type is equipped with both similar and distinct acidic proteases known as cathepsins (40,41) which degraded antigens into their respective T cell epitopes for presentation via MHCII. It is still largely unknown exactly which cathepsins degrade antigens into presentable T cell epitopes; however, it has been shown that Cat B, S and L are involved in processing internalized antigens (42,43), Cat D and E may play roles in certain cell types (44), and additional cathepsins such as Cat Z, F, K and H may be present (30).

Delivery of these newly degraded T cell epitopes for binding with MHCII, step four from above, is a complex process involving many steps and players. Therefore, it is first essential to look in the endoplasmic reticulum for the assembly of MHCII. Additionally, it is important to note that the synthesis and assembly of MHCII, HLA-DM and HLA-DO occurs concurrently or prior to antigen degradation. MHCII molecules are assembled in the endoplasmic reticulum and need to be associated with the polypeptide termed invariant chain (Ii) (45). Binding of the Ii not only helps form and stabilize the MHCII binding groove (46,47), but allows the release of MHCII from the endoplasmic

reticulum (ER) (5). Further, binding of the Ii blocks the peptide binding groove of MHCII, thus preventing other endogenous peptides in the ER from binding with the molecule (5). The invariant chain is additionally important as it's cytoplasmic tail contains the sorting motifs which direct the assembled MHCII molecules from the ER to late endosomal compartments, termed MIIC for MHC class II compartment, where the newly degraded antigens await (48).

While MHCII is assembled in the ER, the helper chaperones HLA-DM and HLA-DO are likewise assembled in the ER (49,50). While all APCs express HLA-DM, it was long accepted that only B cells were capable of expressing HLA-DO (51). However, recent studies have found evidence of –DO in thymic medullary cells (52) and dendritic cells (53-55). Both of these chaperones play a role in modulating peptide binding with MHCII. In the ER, HLA-DO must bind with HLA-DM to be transported from the ER into late endosomal compartments, i.e. MIIC (50). It is here, in the MIIC, where processed peptides, –DM, -DO, cathepsins and MHCII molecules combine and MHCII can acquire its antigenic peptide.

How the newly processed antigenic peptides were shuttled to the MIIC depends on in which APC cell type presentation is occurring. In macrophages and DCs were antigens are endocytosed into phagosomes, the phagosomes eventually fuse with late endosomes/lysosomes to form a phagolysosome exposing the peptides to proteases, HLA-DM and MHCII (30,31). In B cells, signal sequences located on the cytoplasmic tail of the BCR sort the processed peptides into MIIC (30). Before MHCII can bind with these antigenic peptides, it must first degrade the Ii currently bound within the binding groove. This is primarily accomplished by the cathepsins Cat S and L (56,57), which

degrade Ii to the minimum peptide class II-associated invariant chain (CLIP) (30). It is here that HLA-DM assists MHCII molecules in loading antigenic peptides. HLA-DM binds with class II molecules which induces the release of CLIP from the binding groove and stabilizes the molecules (58,59). Studies suggest that –DM binding with MHCII leads to a conformational change in MHCII structure which enables the dissociation of CLIP through the disruption of hydrogen bonds between the peptide backbone and MHCII (15,59). Further, HLA-DM also assists MHCII in binding of high affinity antigenic peptides (60-62). This peptide selection process if often referred to as "editing" in which –DM eliminates presentation of low affinity peptides while enhancing that of high affinity binders (59). Indeed, studies show that cells deficient in HLA-DM are defective in MHCII antigen presentation and that restoration of the molecule restores the cells ability to process and present antigens (63,64). The role of HLA-DO in antigen presentation is still a considerable question in the field; however, it is known that –DO functions as a modulator of –DM (49). HLA-DO binds HLA-DM in the same location that MHCII would, therefore acting as an inhibitor of –DM function with MHCII (49). This has led to thoughts that HLA-DO may play a key role in regulating autoimmunity (49). Additionally, it is thought that –DO may acts as a chaperone to –DM itself. HLA-DO binding with –DM may stabilize the molecule and prevent its degradation until the antigen containing endosomes fuse with lysosomes forming the MIIC and MHCII is ready to be loaded with antigenic peptides (65). Finally, -DO may have roles in promoting peptide presentation specifically in B cells (66). The sixth and final step in antigen processing and presentation is shuttling the peptide-MHCII complex from the MIIC to the plasma membrane to be presented to T cells. This is controlled by factors

such as cholesterol, pH, kinases and GTPases. Additionally, transport is driven by microtubule-based motors (dynein) and actin-based myosin motors, but the molecular basis for this processed is still largely undefined (5).

CD4+ T cells

CD4+ T cells play a major role in MHCII driven adaptive immune responses. T cells are developed in the thymus derived from lymphoid progenitor cells which come from either the bone marrow or fetal liver (67-69). T cell populations include CD4+ and CD8+ T cells, γδ TCR T cells, and natural killer T (NKT) cells (1,3). However, the most abundant T cell type in the human body is the CD4+ αβ TCR population (3). These cells most often serve in a helper function (enhance B and T cell responses) and are denoted as T_H cells (3). Upon activation T_H cells produce an array of cytokines leading them to be further classified into subcategories such as T_H1, T_H2, T_H17, T_{FH}, T_{reg}, or T_H9 based on the cytokines they produce (3). T cells become activated when their T cell receptor (TCR) binds an antigen-MHC complex. This is necessary not only to generate a cellular immune response, but also for the development and maintenance of the host T cell repertoire (70).

CD4+ αβ TCRs specifically interact with antigen-MHCII complexes to initiate T cell activation. Each TCR is made of two chains, an α and β, with each chain being composed of variable and constant regions (71). Additionally, TCRs contain an antigen binding site with three complementarity determining regions (CDRs) (70,71). The TCR binds with both the antigenic peptide (about 30% of the interface) (72) and the MHCII molecule (70) to form the immunological synapse (73). Peptide binding contacts are made through the CDR3 loop while binding with the MHCII helices are made through

CDR1 and CDR2 loops (74). TCR binding to their antigen-MHCII ligand is typically characterized by low affinity binding with slow kinetics, believed to be results of conformational changes during binding (75). The current accepted binding model of TCR with antigen-MHCII complexes is the TCR α and β chains are fixed over the α and β helices of MHCII that make up the peptide binding groove (76). However, a number of recent studies have demonstrated non-traditional docking topologies for TCRs with MHC or MHC-like molecules (76-78).

Glycoconjugates as T cell antigens (adapted from Sun, Middleton, Wantuch PL et al., *Glycobiology* 2016) (79)

The preceding sections have given in-depth explanation into the antigen processing and presentation pathway leading to an adaptive immune response, as well as exploring the tri-molecular complex responsible for generating the response: the antigen, MHCII, and TCR. This final section will review glycoconjugates as T cell antigens, specifically glycoconjugate vaccines, and how they elicit a T cell dependent adaptive immune response via the MHCII presentation pathway.

The introduction of glycoconjugate vaccines led to major breakthroughs in combating bacterial pathogens (80,81). Currently, there are a number of vaccines on the market against such pathogens as *S. pneumoniae*, *H. influenzae* type b and *N. meningitidis* (82,83). Glycoconjugate vaccines are composed of a carrier protein covalently conjugated to a bacterial carbohydrate pathogen. This is to induce a strong antibody response against the carbohydrate pathogen of the bacteria (84-86). The following sections will address glycoconjugate vaccines and their immune activation

broadly; for more in depth coverage of *S. pneumoniae* glycoconjugate vaccines specifically (the major focus of this thesis work) see Chapters 2 and 3.

Glycoconjugates generate T cell dependent immune protection

The capsular polysaccharide (CPS) of bacteria often contribute to their virulence mechanism and allow immune evasion from the host through increased bacterial attachment which prevents phagocytosis (87,88). In addition to being a major virulence factor, CPS are also surface localized with unique structures making them ideal vaccine candidates. These vaccines are designed to generate carbohydrate specific immune responses against the bacteria. However, polysaccharides alone are poorly immunogenic and cannot trigger T cell help to induce an adaptive immune response leading to memory B and T cells (89-91). There are a number of pure polysaccharide based vaccines used today; however, these vaccines while seeing some success in healthy adults fail to provide adequate protection in infants, elderly and immunocompromised individuals (92-94).

Glycoconjugate vaccines, on the other hand, are capable of eliciting a T cell dependent immune response against the CPS. This is accomplished by the coupling of the CPS to a carrier protein (95,96). These vaccines induce a CPS specific adaptive immune response which leads to IgM to IgG class switching, B and T memory and higher protection compared to the polysaccharide based vaccines (97,98). Indeed, immunization with current glycoconjugate vaccines on the market, such as Prevnar 13® against *S. pneumoniae*, have prevented infections by inducing long lasting protection against the bacteria (81,91,99-101). Despite the successes of glycoconjugate vaccines problem still persist. Namely in the vaccines failure to induce robust protective responses in the elderly

and immunocompromised individuals (81,93,102). These continued problems demonstrate the great need to enhance vaccine design and efficacy by developing structurally well-defined conjugates based on their mechanism of immune activation, i.e. presentation to T cells via MHCII.

Glycoconjugates elicit carbohydrate specific T cell responses via MHCII pathways

It has been a long standing hypothesis that carbohydrates are T cell independent antigens and therefore accepted that an adaptive immune response from glycoconjugate vaccine activation was due to the presentation of a peptide derived from the carrier protein (103) **(Figure 1.5)**. This hypothesis follows the same antigen processing and presentation pathway described in the above section with one difference; the BCR of the B cell binds the carbohydrate portion of the glycoconjugate (as opposed to carrier protein) which stimulates the production of CPS specific antibodies. The B cell would then process the carrier protein as previously explained and presents a peptide via MHCII to a CD4+ T cell. T cell activation induces the production of cytokine IL-4 which stimulates B cell maturation and subsequent CPS specific antibodies, IgG class switching and memory B and T cell production (94,98) **(Figure 1.5)**. At the time, considering that carbohydrates were T cell independent antigens and T cells could only bind protein antigens, this seemed the plausible hypothesis. However, this mechanism assumes one, the strong covalent linkage between the CPS and carrier protein is broken under endosomal conditions and two, fails to consider if T cells can bind an MHCII presented peptide linked carbohydrate (93,100,101,104-106).

Recently, the mechanism for glycoconjugate vaccine activation resulting in an adaptive immune response has been discovered. Importantly, this discovery shifted the

paradigm of peptide-centric T cell recognition (107). This work demonstrated that glycoconjugate vaccines are taken up by APCs and depolymerized into smaller molecular weight glycan-peptide conjugates showing the strong covalent linkage between the two is intact even under harsh endosomal conditions (107). After this processing by the APC, the smaller glycoconjugates follow the same antigen presentation pathway as above, in that they bind with MHCII and are presented at the APC surface **(Figure 1.6)**. Importantly, MHCII deficient APCs did not present the conjugates further demonstrating this presentation is MHCII dependent (107). The working model for this presentation shows that the peptide portion of the glycan-peptide binds with MHCII leaving the glycan exposed to be bound by the CD4+ TCR **(Figure 1.6)**. This work also demonstrated that immunizations with glycoconjugate vaccines induced a subset of CD4+ T cells, termed Tcarbs, which specifically recognize the carbohydrate portion of the conjugate, establishing that T cells are in fact able to recognize carbohydrates. The current model therefore is that it is not the T cell's inability to bind carbohydrate, but rather the carbohydrate's inability to bind MHCII without being linked with a carrier peptide.

Importantly, additional studies suggest the generalizability of Tcarb mediated immune response induced by glycoconjugate vaccines. In one such report, glycoconjugates for type 3 *S. pneumoniae* CPS provide evidence for CPS specific IgG responses in a Tcarb dependent manner (108). Additional works demonstrate evidence for glycoconjugate vaccine presentation at the APC surface (109) and Tcarb production from a meningococcal glycoconjugate vaccine (110). Further, a more recent report explored five conjugate vaccines in which four followed the Tcarb mediated mechanism,

while one followed the traditional peptide based mechanism described above (111). This is important as it demonstrates that while a Tcarb mediated response is the major activation pathway it is not the only pathway to generate carbohydrate specific adaptive immune responses. A further understanding of the immune mechanisms leading to glycoconjugate vaccine activation of adaptive immunity will provide a guide for the development of highly immunogenic vaccines to combat bacterial, parasitic and viral infections utilizing their surface glycans.

References

1. Marshall, J. S., Warrington, R., Watson, W., and Kim, H. L. (2018) An introduction to immunology and immunopathology. *Allergy Asthma Clin Immunol* **14**, 49
2. Chaplin, D. D. (2010) Overview of the immune response. *J Allergy Clin Immunol* **125**, S3-23
3. Bonilla, F. A., and Oettgen, H. C. (2010) Adaptive immunity. *J Allergy Clin Immunol* **125**, S33-40
4. Hunt, D. F., Michel, H., Dickinson, T. A., Shabanowitz, J., Cox, A. L., Sakaguchi, K., Appella, E., Grey, H. M., and Sette, A. (1992) Peptides presented to the immune system by the murine class II major histocompatibility complex molecule I-Ad. *Science* **256**, 1817-1820
5. Neefjes, J., Jongsma, M. L., Paul, P., and Bakke, O. (2011) Towards a systems understanding of MHC class I and MHC class II antigen presentation. *Nat Rev Immunol* **11**, 823-836
6. Tampé, R., and McConnell, H. M. (1991) Kinetics of antigenic peptide binding to the class II major histocompatibility molecule I-Ad. *Proc Natl Acad Sci U S A* **88**, 4661-4665
7. Matsumura, M., Fremont, D. H., Peterson, P. A., and Wilson, I. A. (1992) Emerging principles for the recognition of peptide antigens by MHC class I molecules. *Science* **257**, 927-934

8. Wieczorek, M., Abualrous, E. T., Sticht, J., Álvaro-Benito, M., Stolzenberg, S., Noé, F., and Freund, C. (2017) Major Histocompatibility Complex (MHC) Class I and MHC Class II Proteins: Conformational Plasticity in Antigen Presentation. *Front Immunol* **8**, 292

9. Chicz, R. M., Urban, R. G., Lane, W. S., Gorga, J. C., Stern, L. J., Vignali, D. A., and Strominger, J. L. (1992) Predominant naturally processed peptides bound to HLA-DR1 are derived from MHC-related molecules and are heterogeneous in size. *Nature* **358**, 764-768

10. Nalefski, E. A., Shaw, K. T., and Rao, A. (1995) An ion pair in class II major histocompatibility complex heterodimers critical for surface expression and peptide presentation. *J Biol Chem* **270**, 22351-22360

11. Scott, C. A., Peterson, P. A., Teyton, L., and Wilson, I. A. (1998) Crystal structures of two I-Ad-peptide complexes reveal that high affinity can be achieved without large anchor residues. *Immunity* **8**, 319-329

12. Rammensee, H. G., Friede, T., and Stevanoviíc, S. (1995) MHC ligands and peptide motifs: first listing. *Immunogenetics* **41**, 178-228

13. Witt, S. N., and McConnell, H. M. (1994) Formation and dissociation of short-lived class II MHC-peptide complexes. *Biochemistry* **33**, 1861-1868

14. Ferrante, A., Templeton, M., Hoffman, M., and Castellini, M. J. (2015) The Thermodynamic Mechanism of Peptide-MHC Class II Complex Formation Is a Determinant of Susceptibility to HLA-DM. *J Immunol* **195**, 1251-1261

15. Grotenbreg, G. M., Nicholson, M. J., Fowler, K. D., Wilbuer, K., Octavio, L., Yang, M., Chakraborty, A. K., Ploegh, H. L., and Wucherpfennig, K. W. (2007)

Empty class II major histocompatibility complex created by peptide photolysis establishes the role of DM in peptide association. *J Biol Chem* **282**, 21425-21436

16. Fremont, D. H., Matsumura, M., Stura, E. A., Peterson, P. A., and Wilson, I. A. (1992) Crystal structures of two viral peptides in complex with murine MHC class I H-2Kb. *Science* **257**, 919-927

17. Fremont, D. H., Hendrickson, W. A., Marrack, P., and Kappler, J. (1996) Structures of an MHC class II molecule with covalently bound single peptides. *Science* **272**, 1001-1004

18. Stern, L. J., Brown, J. H., Jardetzky, T. S., Gorga, J. C., Urban, R. G., Strominger, J. L., and Wiley, D. C. (1994) Crystal structure of the human class II MHC protein HLA-DR1 complexed with an influenza virus peptide. *Nature* **368**, 215-221

19. Fleckenstein, B., Jung, G., and Wiesmüller, K. H. (1999) Quantitative analysis of peptide-MHC class II interaction. *Semin Immunol* **11**, 405-416

20. Batalia, M. A., and Collins, E. J. (1997) Peptide binding by class I and class II MHC molecules. *Biopolymers* **43**, 281-302

21. Stern, L. J., and Wiley, D. C. (1994) Antigenic peptide binding by class I and class II histocompatibility proteins. *Behring Inst Mitt*, 1-10

22. Gunther, S., Schlundt, A., Sticht, J., Roske, Y., Heinemann, U., Wiesmuller, K. H., Jung, G., Falk, K., Rotzschke, O., and Freund, C. (2010) Bidirectional binding of invariant chain peptides to an MHC class II molecule. *Proc Natl Acad Sci U S A* **107**, 22219-22224

23. McFarland, B. J., Sant, A. J., Lybrand, T. P., and Beeson, C. (1999) Ovalbumin (323-339) peptide binds to the Major Histocompatibility Complex Class II I-A protein using two functionally distinct registers. *Biochemistry* **38**, 16663-16670

24. Jardetzky, T. S., Brown, J. H., Gorga, J. C., Stern, L. J., Urban, R. G., Strominger, J. L., and Wiley, D. C. (1996) Crystallographic analysis of endogenous peptides associated with HLA-DR1 suggests a common, polyproline II-like conformation for bound peptides. *Proc Natl Acad Sci U S A* **93**, 734-738

25. Wantuch, P. L., Sun, L., LoPilato, R. K., Mousa, J. J., Haltiwanger, R. S., and Avci, F. Y. (2020) Isolation and characterization of new human carrier peptides from two important vaccine immunogens. *Vaccine* **38**, 2315-2325

26. Sette, A., Buus, S., Appella, E., Adorini, L., and Grey, H. M. (1990) Structural requirements for the interaction between class II MHC molecules and peptide antigens. *Immunol Res* **9**, 2-7

27. Jablonsky, M. J., Subramaniam, P. S., Johnson, H. M., Russell, J. K., and Krishna, N. R. (1997) The solution structure of a class II major histocompatibility complex superantigen binding domain. *Biochem Biophys Res Commun* **234**, 660-665

28. Jardetzky, T. S., Brown, J. H., Gorga, J. C., Stern, L. J., Urban, R. G., Chi, Y. I., Staufferacher, C., Strominger, J. L., and Wiley, D. C. (1994) Three-dimensional structure of a human class II histocompatibility molecule complexed with superantigen. *Nature* **368**, 711-718

29. Kim, J., Urban, R. G., Strominger, J. L., and Wiley, D. C. (1994) Toxic shock syndrome toxin-1 complexed with a class II major histocompatibility molecule HLA-DR1. *Science* **266**, 1870-1874

30. Bryant, P., and Ploegh, H. (2004) Class II MHC peptide loading by the professionals. *Curr Opin Immunol* **16**, 96-102

31. Vyas, J. M., Van der Veen, A. G., and Ploegh, H. L. (2008) The known unknowns of antigen processing and presentation. *Nat Rev Immunol* **8**, 607-618

32. Siemasko, K., and Clark, M. R. (2001) The control and facilitation of MHC class II antigen processing by the BCR. *Curr Opin Immunol* **13**, 32-36

33. Drake, J. R., Lewis, T. A., Condon, K. B., Mitchell, R. N., and Webster, P. (1999) Involvement of MIIC-like late endosomes in B cell receptor-mediated antigen processing in murine B cells. *J Immunol* **162**, 1150-1155

34. Harding, C. V., Ramachandra, L., and Wick, M. J. (2003) Interaction of bacteria with antigen presenting cells: influences on antigen presentation and antibacterial immunity. *Curr Opin Immunol* **15**, 112-119

35. Duclos, S., Corsini, R., and Desjardins, M. (2003) Remodeling of endosomes during lysosome biogenesis involves 'kiss and run' fusion events regulated by rab5. *J Cell Sci* **116**, 907-918

36. Rescigno, M., and Borrow, P. (2001) The host-pathogen interaction: new themes from dendritic cell biology. *Cell* **106**, 267-270

37. Kleijmeer, M., Ramm, G., Schuurhuis, D., Griffith, J., Rescigno, M., Ricciardi-Castagnoli, P., Rudensky, A. Y., Ossendorp, F., Melief, C. J., Stoorvogel, W., and

Geuze, H. J. (2001) Reorganization of multivesicular bodies regulates MHC class II antigen presentation by dendritic cells. *J Cell Biol* **155**, 53-63

38. Laupeze, B., Fardel, O., Onno, M., Bertho, N., Drenou, B., Fauchet, R., and Amiot, L. (1999) Differential expression of major histocompatibility complex class Ia, Ib, and II molecules on monocytes-derived dendritic and macrophagic cells. *Hum Immunol* **60**, 591-597

39. Mellman, I., and Steinman, R. M. (2001) Dendritic cells: specialized and regulated antigen processing machines. *Cell* **106**, 255-258

40. Bryant, P. W., Lennon-Dumenil, A. M., Fiebiger, E., Lagaudriere-Gesbert, C., and Ploegh, H. L. (2002) Proteolysis and antigen presentation by MHC class II molecules. *Adv Immunol* **80**, 71-114

41. Chapman, H. A. (2006) Endosomal proteases in antigen presentation. *Curr Opin Immunol* **18**, 78-84

42. Hsieh, C. S., deRoos, P., Honey, K., Beers, C., and Rudensky, A. Y. (2002) A role for cathepsin L and cathepsin S in peptide generation for MHC class II presentation. *J Immunol* **168**, 2618-2625

43. Kakehashi, H., Nishioku, T., Tsukuba, T., Kadowaki, T., Nakamura, S., and Yamamoto, K. (2007) Differential regulation of the nature and functions of dendritic cells and macrophages by cathepsin E. *J Immunol* **179**, 5728-5737

44. Hewitt, E. W., Treumann, A., Morrice, N., Tatnell, P. J., Kay, J., and Watts, C. (1997) Natural processing sites for human cathepsin E and cathepsin D in tetanus toxin: implications for T cell epitope generation. *J Immunol* **159**, 4693-4699

45. Busch, R., Doebele, R. C., Patil, N. S., Pashine, A., and Mellins, E. D. (2000) Accessory molecules for MHC class II peptide loading. *Curr Opin Immunol* **12**, 99-106

46. Bikoff, E. K., Huang, L. Y., Episkopou, V., van Meerwijk, J., Germain, R. N., and Robertson, E. J. (1993) Defective major histocompatibility complex class II assembly, transport, peptide acquisition, and CD4+ T cell selection in mice lacking invariant chain expression. *J Exp Med* **177**, 1699-1712

47. Viville, S., Neefjes, J., Lotteau, V., Dierich, A., Lemeur, M., Ploegh, H., Benoist, C., and Mathis, D. (1993) Mice lacking the MHC class II-associated invariant chain. *Cell* **72**, 635-648

48. Landsverk, O. J., Bakke, O., and Gregers, T. F. (2009) MHC II and the endocytic pathway: regulation by invariant chain. *Scand J Immunol* **70**, 184-193

49. Mellins, E. D., and Stern, L. J. (2014) HLA-DM and HLA-DO, key regulators of MHC-II processing and presentation. *Curr Opin Immunol* **26**, 115-122

50. Poluektov, Y. O., Kim, A., and Sadegh-Nasseri, S. (2013) HLA-DO and Its Role in MHC Class II Antigen Presentation. *Front Immunol* **4**, 260

51. Karlsson, L., Surh, C. D., Sprent, J., and Peterson, P. A. (1991) A novel class II MHC molecule with unusual tissue distribution. *Nature* **351**, 485-488

52. Douek, D. C., and Altmann, D. M. (2000) T-cell apoptosis and differential human leucocyte antigen class II expression in human thymus. *Immunology* **99**, 249-256

53. Chen, X., Reed-Loisel, L. M., Karlsson, L., and Jensen, P. E. (2006) H2-O expression in primary dendritic cells. *J Immunol* **176**, 3548-3556

54. Fallas, J. L., Yi, W., Draghi, N. A., O'Rourke, H. M., and Denzin, L. K. (2007) Expression patterns of H2-O in mouse B cells and dendritic cells correlate with cell function. *J Immunol* **178**, 1488-1497

55. Hornell, T. M., Burster, T., Jahnsen, F. L., Pashine, A., Ochoa, M. T., Harding, J. J., Macaubas, C., Lee, A. W., Modlin, R. L., and Mellins, E. D. (2006) Human dendritic cell expression of HLA-DO is subset specific and regulated by maturation. *J Immunol* **176**, 3536-3547

56. Nakagawa, T., Roth, W., Wong, P., Nelson, A., Farr, A., Deussing, J., Villadangos, J. A., Ploegh, H., Peters, C., and Rudensky, A. Y. (1998) Cathepsin L: critical role in Ii degradation and CD4 T cell selection in the thymus. *Science* **280**, 450-453

57. Shi, G. P., Villadangos, J. A., Dranoff, G., Small, C., Gu, L., Haley, K. J., Riese, R., Ploegh, H. L., and Chapman, H. A. (1999) Cathepsin S required for normal MHC class II peptide loading and germinal center development. *Immunity* **10**, 197-206

58. Denzin, L. K., and Cresswell, P. (1995) HLA-DM induces CLIP dissociation from MHC class II alpha beta dimers and facilitates peptide loading. *Cell* **82**, 155-165

59. Pos, W., Sethi, D. K., Call, M. J., Schulze, M. S., Anders, A. K., Pyrdol, J., and Wucherpfennig, K. W. (2012) Crystal structure of the HLA-DM-HLA-DR1 complex defines mechanisms for rapid peptide selection. *Cell* **151**, 1557-1568

60. Katz, J. F., Stebbins, C., Appella, E., and Sant, A. J. (1996) Invariant chain and DM edit self-peptide presentation by major histocompatibility complex (MHC) class II molecules. *J Exp Med* **184**, 1747-1753

61. Weber, D. A., Dao, C. T., Jun, J., Wigal, J. L., and Jensen, P. E. (2001) Transmembrane domain-mediated colocalization of HLA-DM and HLA-DR is required for optimal HLA-DM catalytic activity. *J Immunol* **167**, 5167-5174

62. Sloan, V. S., Cameron, P., Porter, G., Gammon, M., Amaya, M., Mellins, E., and Zaller, D. M. (1995) Mediation by HLA-DM of dissociation of peptides from HLA-DR. *Nature* **375**, 802-806

63. Denzin, L. K., Robbins, N. F., Carboy-Newcomb, C., and Cresswell, P. (1994) Assembly and intracellular transport of HLA-DM and correction of the class II antigen-processing defect in T2 cells. *Immunity* **1**, 595-606

64. Morris, P., Shaman, J., Attaya, M., Amaya, M., Goodman, S., Bergman, C., Monaco, J. J., and Mellins, E. (1994) An essential role for HLA-DM in antigen presentation by class II major histocompatibility molecules. *Nature* **368**, 551-554

65. Adler, L. N., Jiang, W., Bhamidipati, K., Millican, M., Macaubas, C., Hung, S. C., and Mellins, E. D. (2017) The Other Function: Class II-Restricted Antigen Presentation by B Cells. *Front Immunol* **8**, 319

66. Alfonso, C., Williams, G. S., Han, J. O., Westberg, J. A., Winqvist, O., and Karlsson, L. (2003) Analysis of H2-O influence on antigen presentation by B cells. *J Immunol* **171**, 2331-2337

67. Hedrick, S. M. (2008) Thymus lineage commitment: a single switch. *Immunity* **28**, 297-299

68. Jenkinson, E. J., Jenkinson, W. E., Rossi, S. W., and Anderson, G. (2006) The thymus and T-cell commitment: the right niche for Notch? *Nat Rev Immunol* **6**, 551-555

69. Takahama, Y. (2006) Journey through the thymus: stromal guides for T-cell development and selection. *Nat Rev Immunol* **6**, 127-135

70. Piepenbrink, K. H., Gloor, B. E., Armstrong, K. M., and Baker, B. M. (2009) Methods for quantifying T cell receptor binding affinities and thermodynamics. *Methods Enzymol* **466**, 359-381

71. Grant, E. J., Josephs, T. M., Valkenburg, S. A., Wooldridge, L., Hellard, M., Rossjohn, J., Bharadwaj, M., Kedzierska, K., and Gras, S. (2016) Lack of Heterologous Cross-reactivity toward HLA-A*02:01 Restricted Viral Epitopes Is Underpinned by Distinct alphabetaT Cell Receptor Signatures. *J Biol Chem* **291**, 24335-24351

72. Rudolph, M. G., Stanfield, R. L., and Wilson, I. A. (2006) How TCRs bind MHCs, peptides, and coreceptors. *Annu Rev Immunol* **24**, 419-466

73. Dustin, M. L. (2009) The cellular context of T cell signaling. *Immunity* **30**, 482-492

74. Patarroyo, M. E., Bermudez, A., and Patarroyo, M. A. (2011) Structural and immunological principles leading to chemically synthesized, multiantigenic, multistage, minimal subunit-based vaccine development. *Chem Rev* **111**, 3459-3507

75. Willcox, B. E., Gao, G. F., Wyer, J. R., Ladbury, J. E., Bell, J. I., Jakobsen, B. K., and van der Merwe, P. A. (1999) TCR binding to peptide-MHC stabilizes a flexible recognition interface. *Immunity* **10**, 357-365

76. Beringer, D. X., Kleijwegt, F. S., Wiede, F., van der Slik, A. R., Loh, K. L., Petersen, J., Dudek, N. L., Duinkerken, G., Laban, S., Joosten, A., Vivian, J. P., Chen, Z., Uldrich, A. P., Godfrey, D. I., McCluskey, J., Price, D. A., Radford, K. J., Purcell, A. W., Nikolic, T., Reid, H. H., Tiganis, T., Roep, B. O., and Rossjohn, J. (2015) T cell receptor reversed polarity recognition of a self-antigen major histocompatibility complex. *Nat Immunol* **16**, 1153-1161

77. Gras, S., Chadderton, J., Del Campo, C. M., Farenc, C., Wiede, F., Josephs, T. M., Sng, X. Y. X., Mirams, M., Watson, K. A., Tiganis, T., Quinn, K. M., Rossjohn, J., and La Gruta, N. L. (2016) Reversed T Cell Receptor Docking on a Major Histocompatibility Class I Complex Limits Involvement in the Immune Response. *Immunity* **45**, 749-760

78. Le Nours, J., Gherardin, N. A., Ramarathinam, S. H., Awad, W., Wiede, F., Gully, B. S., Khandokar, Y., Praveena, T., Wubben, J. M., Sandow, J. J., Webb, A. I., von Borstel, A., Rice, M. T., Redmond, S. J., Seneviratna, R., Sandoval-Romero, M. L., Li, S., Souter, M. N. T., Eckle, S. B. G., Corbett, A. J., Reid, H. H., Liu, L., Fairlie, D. P., Giles, E. M., Westall, G. P., Tothill, R. W., Davey, M. S., Berry, R., Tiganis, T., McCluskey, J., Pellicci, D. G., Purcell, A. W., Uldrich, A. P., Godfrey, D. I., and Rossjohn, J. (2019) A class of gammadelta T cell receptors recognize the underside of the antigen-presenting molecule MR1. *Science* **366**, 1522-1527

79. Sun, L., Middleton, D. R., Wantuch, P. L., Ozdilek, A., and Avci, F. Y. (2016) Carbohydrates as T-cell antigens with implications in health and disease. *Glycobiology* **26**, 1029-1040

80. Goldblatt, D. (2000) Conjugate vaccines. *Clin Exp Immunol* **119**, 1-3

81. Wantuch, P. L., and Avci, F. Y. (2018) Current status and future directions of invasive pneumococcal diseases and prophylactic approaches to control them. *Hum Vaccin Immunother* **14**, 2303-2309

82. Pichichero, M. E. (2013) Protein carriers of conjugate vaccines: characteristics, development, and clinical trials. *Hum Vaccin Immunother* **9**, 2505-2523

83. Micoli, F., Adamo, R., and Costantino, P. (2018) Protein Carriers for Glycoconjugate Vaccines: History, Selection Criteria, Characterization and New Trends. *Molecules* **23**

84. Li, P., and Wang, F. (2015) Polysaccharides: Candidates of promising vaccine adjuvants. *Drug Discov Ther* **9**, 88-93

85. Vella, M., and Pace, D. (2015) Glycoconjugate vaccines: an update. *Expert Opin Biol Ther* **15**, 529-546

86. Zimmermann, S., and Lepenies, B. (2015) Glycans as Vaccine Antigens and Adjuvants: Immunological Considerations. *Methods Mol Biol* **1331**, 11-26

87. AlonsoDeVelasco, E., Verheul, A. F., Verhoef, J., and Snippe, H. (1995) Streptococcus pneumoniae: virulence factors, pathogenesis, and vaccines. *Microbiol Rev* **59**, 591-603

88. Casadevall, A., and Pirofski, L. A. (2009) Virulence factors and their mechanisms of action: the view from a damage-response framework. *J Water Health* **7 Suppl 1**, S2-S18

89. Coutinho, A., and Moller, G. (1973) B cell mitogenic properties of thymus-independent antigens. *Nature New Biol* **245**, 12-14

90. Barrett, D. J. (1985) Human immune responses to polysaccharide antigens: an analysis of bacterial polysaccharide vaccines in infants. Advances in Pediatrics

91. Weintraub, A. (2003) Immunology of bacterial polysaccharide antigens. *Carbohydr. Res.* **338**, 2539-2547

92. Jones, C. (2005) Vaccines based on the cell surface carbohydrates of pathogenic bacteria. *An Acad Bras Cienc* **77**, 293-324

93. Avci, F. (2013) Novel Strategies for Development of Next-Generation Glycoconjugate Vaccines. *Current Topics in Medicinal Chemistry* **13**, 2535-2540

94. Avci, F., Kasper, D., Paul, W., Littman, D., and Yokoyama, W. (2010) How Bacterial Carbohydrates Influence the Adaptive Immune System. *Annual Review of Immunology, Vol 28* **28**, 107-130

95. Jennings, H. J., and Lugowski, C. (1981) Immunochemistry of groups A, B, and C meningococcal polysaccharide-tetanus toxoid conjugates. *J Immunol* **127**, 1011-1018

96. Schneerson, R., Barrera, O., Sutton, A., and Robbins, J. B. (1980) Preparation, characterization,and immunogenicity of Haemophilus influenzae type b polysaccharide-protein conjugates. *J. Exp. Med.* **152**, 361-376

97. MR, W., LC, P., AK, R., F, M., J, D., HJ, J., and DL, K. (1993) Stimulation of protective antibodies against type Ia and Ib group B streptococci by a type Ia polysaccharide-tetanus toxoid conjugate vaccine. *Infect Immun* **61**, 4760-4766

98. Guttormsen, H., Wetzler, L., Finberg, R., and Kasper, D. (1998) Immunologic memory induced by a glycoconjugate vaccine in a murine adoptive lymphocyte transfer model. *Infection and Immunity* **66**, 2026-2032

99. Wantuch, P. L., and Avci, F. Y. (2019) Invasive pneumococcal disease in relation to vaccine type serotypes. *Hum Vaccin Immunother*, 1-2

100. Avci, F. Y., Li, X., Tsuji, M., and Kasper, D. L. (2013) Carbohydrates and T cells: a sweet twosome. *Semin Immunol* **25**, 146-151

101. Trotter, C., McVernon, J., Ramsay, M., Whitney, C., Mulholland, E., Goldblatt, D., Hombach, J., Kieny, M., and Subgrp, S. (2008) Optimising the use of conjugate vaccines to prevent disease caused by Haemophilus influenzae type b, Neisseria meningitidis and Streptococcus pneumoniae. *Vaccine* **26**, 4434-4445

102. Pirofski, L. (2001) Polysaccharides, mimotopes and vaccines for fungal and encapsulated pathogens. *Trends Microbiol* **9**, 445-451

103. Mitchison, N. (2004) Landmark - T-cell-B-cell cooperation. *Nature Reviews Immunology* **4**, 308-312

104. Kalka-Moll, W. M., Tzianabos, A. O., Bryant, P. W., Niemeyer, M., Ploegh, H. L., and Kasper, D. L. (2002) Zwitterionic polysaccharides stimulate T cells by MHC class II-dependent interactions. *J Immunol* **169**, 6149-6153

105. Cobb, B. A., and Kasper, D. L. (2008) Characteristics of carbohydrate antigen binding to the presentation protein HLA-DR. *Glycobiology* **18**, 707-718

106. Cobb, B. A., Wang, Q., Tzianabos, A. O., and Kasper, D. L. (2004) Polysaccharide processing and presentation by the MHCII pathway. *Cell* **117**, 677-687

107. Avci, F., Li, X., Tsuji, M., and Kasper, D. (2011) A mechanism for glycoconjugate vaccine activation of the adaptive immune system and its implications for vaccine design. *Nature Medicine* **17**, 1602-U1115

108. Middleton, D. R., Sun, L., Paschall, A. V., and Avci, F. Y. (2017) T Cell-Mediated Humoral Immune Responses to Type 3 Capsular Polysaccharide of. *J Immunol* **199**, 598-603

109. Lai, Z., and Schreiber, J. R. (2009) Antigen processing of glycoconjugate vaccines; the polysaccharide portion of the pneumococcal CRM(197) conjugate vaccine co-localizes with MHC II on the antigen processing cell surface. *Vaccine* **27**, 3137-3144

110. Muthukkumar, S., and Stein, K. E. (2004) Immunization with meningococcal polysaccharide-tetanus toxoid conjugate induces polysaccharide-reactive T cells in mice. *Vaccine* **22**, 1290-1299

111. Sun, X., Stefanetti, G., Berti, F., and Kasper, D. L. (2019) Polysaccharide structure dictates mechanism of adaptive immune response to glycoconjugate vaccines. *Proc Natl Acad Sci U S A* **116**, 193-198

Figure 1.1

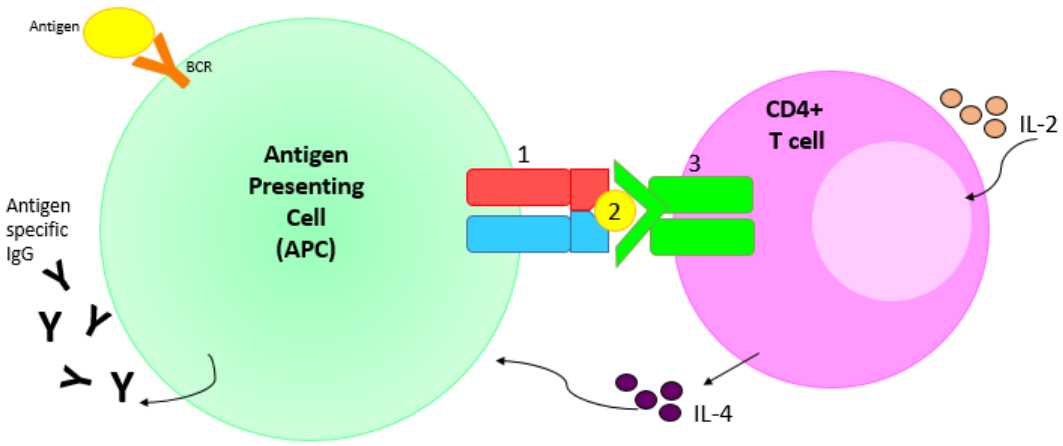


Figure 1.1: Mechanism of adaptive immune response and depiction of tri-molecular complex. Antigen enters the B cell through the B cell receptor (BCR) and is delivered to MHCII. The tri-molecular complex is labeled as 1) MHCII protein 2) Antigen and 3) CD4+ T cell receptor. The presentation of the antigen via MHCII to the TCR leads to T cell activation. CD4+ T cell activation leads to B cell induction to help produce antigen specific IgG antibodies.

Figure 1.2

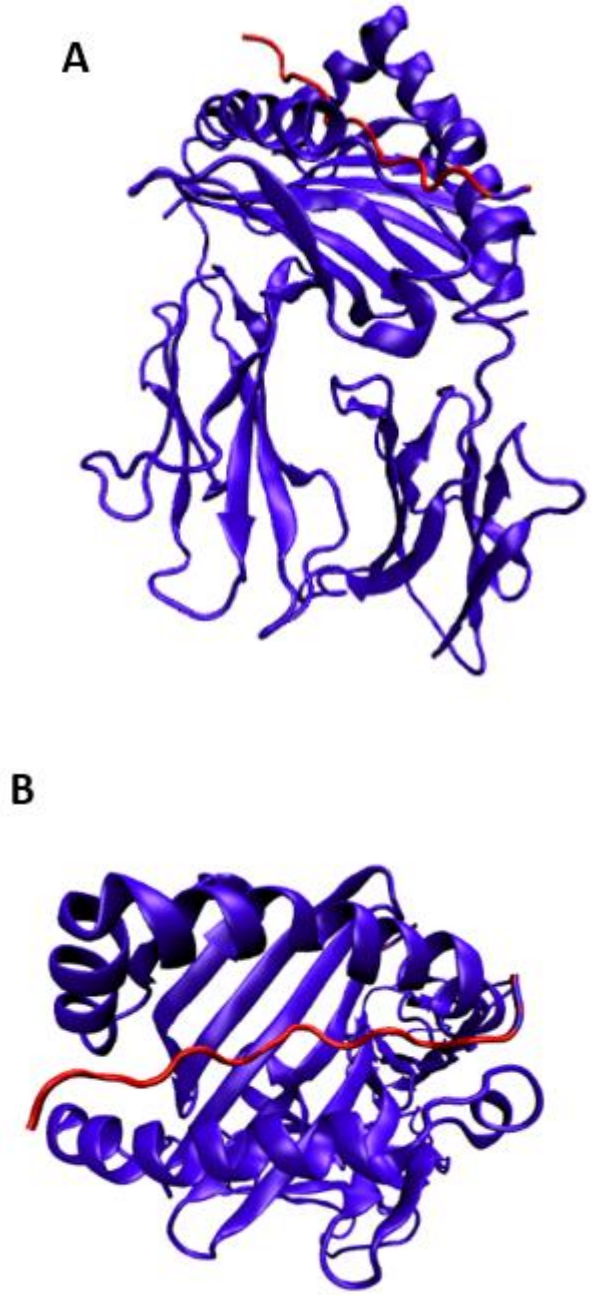


Figure 1.2: Structure of mouse MHCII protein, I-A^d. PDBID 1IAO was used to create images of protein, reference 11 **A)** Full view of the protein, chains A and B of MHCII are

colored in blue while the antigen OVAp is colored in red. **B)** View of the antigen binding groove. MHCII is colored in blue, antigen is colored in red.

Figure 1.3

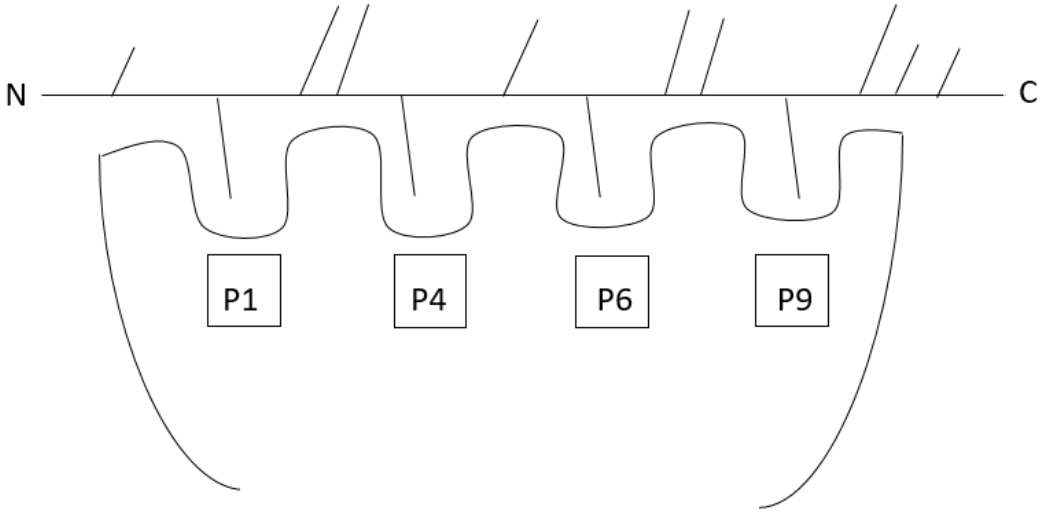


Figure 1.3: Cartoon representation of MHCII binding pockets. Top line depicts a peptide antigen binding with MHCII where some residues fit into designated pockets, while others point upwards. Residues pointing upwards may be bound by the T cell receptor (longer residues) or not (shorter residues). Bottom in the MHCII protein with major pockets (P1,4,6,9) labeled. Figure conceptualized from reference 23.

Figure 1.4

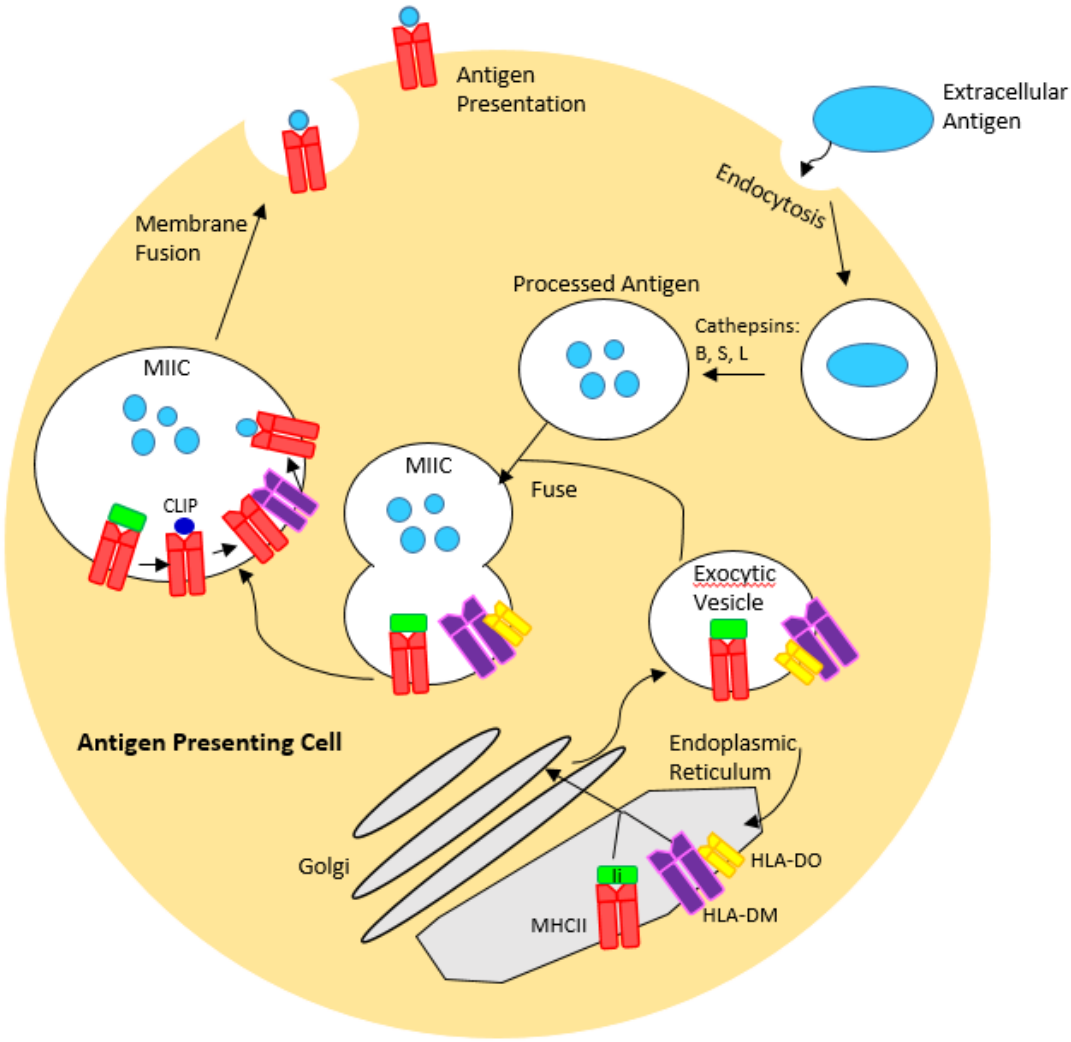


Figure 1.4: Antigen processing and presentation pathway. Extracellular antigen is endocytosed by the APC and processed into smaller antigen fragments by cathepsins. Meanwhile in the ER MHCII, HLA-DM and HLA-DO are assembled and transported to exocytic vesicles. The exocytic vesciles fuse with endosomes containing the processed antigens to form the MIIC. Here in the MIIC, Ii is cleaved into CLIP, with the help of HLA-DM CLIP is removed and swapped for the antigenic peptide. The loaded MHCII is transported to the surface of the APC for antigen presentation to T cells.

Figure 1.5

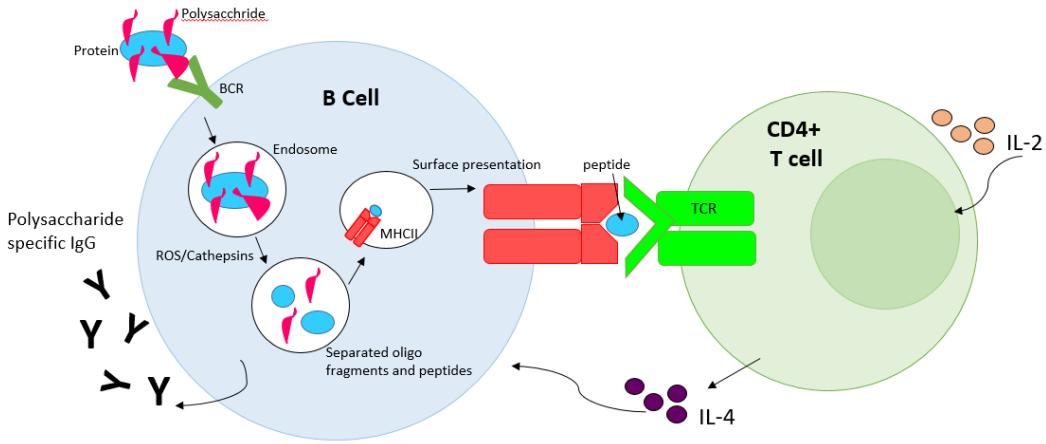


Figure 1.5: Antigen processing and presentation pathway of peptide based carbohydrate specific response. The B cell receptor (BCR) binds with the polysaccharide portion of the glycoconjugate. The glycoconjugate is endocytosed into the endosome and processed by reactive oxygen and nitrogen species (ROS/RNS) and cathepsins to yield separate oligosaccharide and peptide fragments. MHCII binds with a peptide fragment and presents the peptide on the cell surface where a T cell receptor (TCR) binds the complex. This activate the T cells to stimulate the B cell into producing antibodies against the polysaccharide in which it originally bound through the BCR.

Figure 1.6

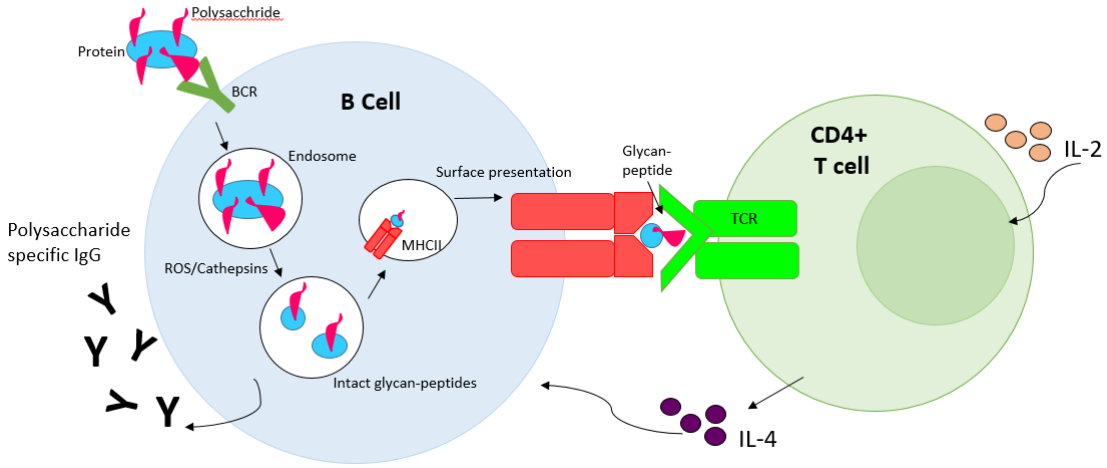


Figure 1.6: Antigen processing and presentation pathway of Tcarb based carbohydrate specific response. The B cell receptor (BCR) binds with the polysaccharide portion of the glycoconjugate. The glycoconjugate is endocytosed into the endosome and processed by reactive oxygen and nitrogen species (ROS/RNS) and cathepsins to yield covalently attached glycan-peptide fragments. MHCII binds with a peptide fragment of the glycan-peptide and presents the glycoconjugate on the cell surface. Here a T cell receptor (TCR) binds the glycan portion of the glycoconjugate. This activate the T cells to stimulate the B cell into producing antibodies against the polysaccharide.

CHAPTER 2

CURRENT STATUS AND FUTURE DIRECTIONS OF INVASIVE PNEUMOCOCCAL DISEASES AND PROPHYLACTIC APPROACHES TO CONTROL THEM

DOI: 10.1080/21645515.2018.1470726

Abstract

Streptococcus pneumoniae is a major human bacterial pathogen responsible for millions of deaths each year and significantly more illnesses worldwide. With over 90 different serotypes, providing effective vaccine programs has been a continuing challenge. Since 1983, the world has been introduced to three different pneumococcal vaccines (PPSV23, PCV7, and PCV13) each with their own complications and successes. Since vaccination programs began, a decrease in the overall rate of pneumococcal pneumonia and associated diseases has been observed, notably in higher risk populations. However, with a decrease in incidence of vaccine type pneumococcal serotypes, increases in non-vaccine serotypes of the bacteria have been observed along with serotype switching. Additionally, a rise in antibiotic resistant strains of *S. pneumoniae* is noted. Here we discuss both the positive and negative clinical manifestations of pneumonia vaccine programs and discuss the challenges in pneumococcal vaccine design.

Characteristics of the bacterium

Streptococcus pneumoniae is a gram-positive coccus bacterium (1). It most often colonizes in the nasopharynx of its host (2). Colonization with this bacterium in the nasopharynx is exceptionally common, with 40-95% of infants and 1-10% of adults being colonized at any time (3). There are over 90 distinctive serotypes of this bacterium, which vary in geographical prevalence (4). Different serotypes either cause mucosal colonization or invasive disease. For example, serotypes 1 and 3 are most often isolated from the lung, whereas serotypes 6,10, and 23 are regularly associated with meningitis

(5). Approximately 23 serotypes are responsible for 80-90% of invasive pneumococcal disease (IPD) (5). Each serotype is distinguished by its unique capsular polysaccharide (CPS), which decorates the outside surface of the bacterium. The CPS is a major virulence factor, significantly contributing to disease. Other virulence factors include the cell wall, pneumolysin, and pneumococcal surface protein A (6).

Virulence of the bacterium differs with the composition of the CPS, making it one of the primary virulence factors. The CPS allows the bacterium to evade phagocytosis by the host cells (7). Its antigenicity allows the host to make antibodies against specific serotypes.

Diseases associated with *S. pneumoniae*

S. pneumoniae is responsible for many invasive infections including meningitis, sepsis, and bacteremia, as well as more common bacterial infections such as community acquired pneumonia (CAP) and otitis media. On average, there are 4,100 cases of meningitis (8), 12,000 of bacteremia (9), 500,000 of pneumonia, and 7 million cases of otitis media annually in the United States (10,11). The bacterium most commonly affects pediatric, elderly, and immunocompromised hosts. *S. pneumoniae* is the leading cause of pneumonia in children worldwide and is responsible for 30% of adult pneumonia cases (12). This bacterium is particularly troublesome in at-risk populations, including diabetics, asthmatics, and the HIV positive. In fact, coinfection with influenza or HIV leads to increased *S. pneumoniae* carriage, which leads to a significantly higher risk of pneumococcal infection and mortality (3). *S. pneumoniae* and *H. influenzae* coinfection is the eighth leading cause of death in the United States, with a death toll of over 50,000 people in 2014 (13,14). An average of 1.2 million young children worldwide die due to

pneumococcal pneumonia and meningitis (10). In developed countries, the mortality rate for pneumococcal pneumonia can reach 11-40% (12).

Pneumococcal Vaccines

In 1983, Pneumovax® PPSV23, a vaccine against 23 prevalent serotypes of *S. pneumoniae* was introduced in the United States **(Table 2.1)** (5). This vaccine showed evidence of protection against invasive disease in adults (12,15). There is also some evidence to suggest the vaccine protects the elderly, which is a high-risk population for pneumococcal infections (16). However, the vaccine still has many shortcomings, including poor immunogenicity in children, especially those under the age of two (15). There is also no strong evidence showing protection in immunocompromised patients (5). The effectiveness of PPSV23 seems to decrease with age of the patient and time since administration (5). The majority of these factors can be attributed to the short-lived IgM antibodies elicited from the T-cell independent response, no affinity maturation, and lack of IgG class switching of antibodies (17). The solution for these issues came with the introduction of the T cell dependent conjugate vaccines in the 2000s.

In 2000, PCV7 was introduced in the United States. This vaccine contained the CPS of seven highly virulent serotypes of *S. pneumoniae* covalently conjugated to the carrier protein, non-toxic diphtheria toxin mutant CRM197 **(Table 2.1)** (18). While this conjugate vaccine showed evidence of greater protection in children and herd effect among the population (19), the effectiveness in adults was not well studied (5). In 2010, a new conjugate vaccine was released in the United States, Prevnar® PCV13. Much like PCV7, PCV13 contains the CPS of 13 virulent serotypes conjugated to CRM197 **(Table 2.1)** (15). PCV13 showed promising results for protecting against invasive pneumonia

infection as well as otitis media in young children. Additionally, a 2015 study on PCV13 found it to be protective against CAP in elderly patients (15).

Populations Affected

Pneumococcal pneumonia infections are prevalent among three population groups: children, elderly, and the immunocompromised. Children under the age of two are affected the most. This is most likely due to their immature immune systems, which makes administering vaccinations at this age critical. Children aged two to five are also considered at risk. It is estimated that worldwide pneumococcal disease causes 11% of all deaths in children less than five (19). A study looking at the effects of the PCV7 vaccine noted the risk in this group and the importance of implementing the PCV7 vaccine as PPSV23 was not considered effective in children (20).

Adults over the age of 65 are considered a major risk group. The high risk of pneumococcal infection mostly stems from altered immunity and increased carriage (3). While PPSV23 vaccine was not found effective in children, it did offer limited immunity in adults. The introduction of PCV7 saw a great reduction in pneumococcal pneumonia cases in children; however, its impact on adult disease was not well studied (5,15). With the introduction of PCV13, a greater impact of protection was observed in adults, specifically 65 and older (15).

Lastly, patients who are immunocompromised are at a greater risk of infection. This includes patients who are HIV positive and patients who suffer from chronic diseases (5). *S. pneumoniae* is the most common cause of pneumonia, sepsis, and meningitis among the immunocompromised (3). A more detailed look into how the vaccines have affected these populations is explored in the subsequent section.

Vaccine impact on disease

The overall impact of vaccines on pneumococcal disease has been a major success. The bulk of post vaccination data comes from PCV7 as it was released in 2000; data on PCV13 effectiveness is still being observed. After the introduction of PCV7 in the United States in 2000, a 77% reduction in pneumonia in children under the age of five was observed **(Figure 2.2A)** (21). Additionally, in 2001 there was a 69% decrease in IPD in young children compared to 1998-1999 **(Figures 2.1 and 2.2)** (20,22). The overall rate of vaccination in the United States was between 68-69% (21). A reduction in IPD among immunocompromised patients was also seen as an effect of PCV7 introduction (5).

There has also been reported data of effectiveness in Europe. Specifically, in the Netherlands there was a 35% reduction of disease in children under the age of two and vaccine-type caused IPD decreased by 67% after the introduction of PCV7 (21). However, this was lower than the results seen in the United States despite the Netherlands having an increased rate of vaccination among children (94.4%) (21). The Netherlands also used PCV10, a conjugate vaccine containing all the serotypes of PCV7 plus three more **(Table 2.1)** and observed a 34% reduction in the first episode of otitis media (5). Similar results were seen throughout Europe.

From a major study done on the impact of the PCV13 vaccine, we can see the first effects of this vaccine (15). In 2008 it was reported that 68.4% of invasive pneumonia incidents in adults 65 or older were caused by PCV13 serotypes and 49.7% by PCV7 serotypes (15). In 2013, after the introduction of PCV13, these numbers decreased to 42.3% by PCV13 serotypes, 6% by PCV7 serotypes **(Figure 2.2)** (15,22). Additionally, PCV13 has reduced the incidence of IPD in young children by 36% **(Figure 2.1)** (5,22).

However, with the great many positives observed from the introduction of pneumococcal vaccines there are still areas of concern. In 2006 it was reported that between 23-30% of children still carried pneumococci with no evidence of a decline in carriage prevalence (18,23). Additionally, while PCV7 reduced the incidence of IPD by 97% for serotypes included in the vaccine in children under the age of five, infection from non-vaccine type serotypes increased by 22% (5). This was also observed in the Netherlands, where IPD rates for non-vaccine serotypes increased by 13% (21). The rate of IPD eventually leveled off despite the huge reduction seen in PCV7 serotypes because an emergence in infections caused by serotypes 19A and 7F was observed **(Figures 2.1 and 2.2)** (21,22). It was also noted that the rate of IPD in adults 65 and older caused by 19A was 5.3% in 2008 but increased to 11.4% by 2013 (15). Large increases in non-vaccine serotypes were also observed for HIV/AIDs positive adults (24). Further insights into serotype replacement and the consequences of pneumonia vaccines is discussed in the following sections.

Challenges Associated with Pneumococcal Conjugate Vaccines

The negative consequences observed from pneumococcal pneumonia vaccines are not solely attributed to the vaccines themselves. While there are long-standing problems associated with conjugate vaccine production, such as cost effectiveness and empirical conjugation chemistries, more uncontainable factors are at play.

A major factor in variable responses from the vaccines is that serotypes are not the same geographically. For example, the serotypes contained in PCV7 covered 80% of the most prominent serotypes seen in the US, and 60% of those in Europe (21). The most common serotypes seen in Europe are 1, 3, 7F, 14, and 19A (5). However, even with this

general consensus, large differences are seen country to country. Serotypes 1, 5, 6A, 6B, 14, 19F, and 23F have been isolated in at least 50% of infected individuals worldwide, making them very common worldwide (5).

Additionally, serotype distribution changes between age groups and diseases. The most common serotypes seen in children are 6A, 14,19A, and 19F (25), while the most prevalent in adults are 3, 6A, 7F, and 19A (5). Among pneumococcal diseases, serotypes 1, 2, 4, 6A, 6B, 7F, 8, 9V, 14, 18C, 19F, and 23F are dominantly associated with IPD worldwide (5). However, serotypes 6, 10, and 23 are frequently linked with meningitis and serotypes 14 and 19F with otitis media (26).

One can see how producing a vaccine to capture these many serotype factors would be challenging. Furthermore, immunity is limited to the serotypes contained within the vaccines, with current conjugation chemistries controlling the number of capsules that can be incorporated. Another challenge is the increased frequency of host colonization with other respiratory pathogens such as *S. aureus* and *H. influenza* (16).

A major challenge with conjugate immunization is carrier-specific immune suppression. This is the idea that pre-existing immunity to the carrier protein due to vaccination suppresses the immune response to the carbohydrate linked to the same carrier protein in conjugate vaccines (27). A possible mechanism for this occurring within the host is currently existing antibodies to the carrier protein may prevent or hinder anti-carbohydrate B cells from interacting with their epitopes, thereby favoring anti-carrier protein B cell responses. There have been studies that found that pre-existing immunity to the carrier protein TT interferes with antibody responses to conjugate vaccines (27). The idea of carrier-specific immune suppression and possible mechanisms

by which it occurs is of particular importance to not only pneumococcal conjugate vaccines, but also all existing conjugate vaccines.

Lastly, an issue that is of rising concern is the increase in antibiotic-resistant strains of *S. pneumoniae*. After the introduction of the PCV7 vaccine there was a significant increase in strain 19A, a non-vaccine type strain (28). Importantly, the amount of antibiotic-resistance in this serotype also increased (29). Serotypes 6A, 6B, 9V, 14, 15A, 19F, and 23F have also shown the most antibiotic-resistance to penicillin and erythromycin, with many of these serotypes being included in the vaccines (25). A positive side to this is that many of these vaccine type serotypes, while showing antibiotic resistance, have decreased in frequency since the introduction of the conjugate vaccine. Importantly, the introduction of PCV13 has led to a 45% decrease in IPD caused by 19A in young children (5).

Serotype Replacement and Serotype Switching

A major element that contributes to the effectiveness of the vaccines is serotype replacement and serotype switching. Serotype replacement can be defined as the expanding of non-vaccine type serotypes within the population, while serotype (capsular) switching is a change in a serotype of a single clone by changing of its *cps* locus that synthesizes the CPS. These two events are not mutually exclusive, with capsular switch variants often expanding within the population (7).

The post PCV7 vaccination era saw an increase in multiple non-vaccine type serotypes. During pre-vaccination years 17% of pneumonia was caused by non-vaccine serotypes in children under five, this increased to 88% by 2004 (24). For adults 65 and older the same trend was witnessed. Between 1998 and 1999 the IPD rate by non-vaccine

serotypes was 44%, which increased to 78% by 2004 (24). There was a notable rise in strains 19A and 7F (18). The proportion of IPD caused by 19A increased from 0% in 1991-1994 to 18% in 2001-2003 (29). Additionally, strains 3, 15, 22F and 33F increased in children under the age of five, however, these strains were also seen among adults (24). Slight increases in strain 19F, 6A, 16F, 23A, and 35 were also observed in adults (24,25,29). The increase in strains 19A and 3 is particularly alarming since these two serotypes are commonly associated with IPD. The post PCV7 era also witnessed an increase in non-vaccine type IPD in the elderly, with this increase being more pronounced than the increase that was observed in children under five (29). There was also an increase in non-vaccine type IPD in HIV adults. However, there was a 33% reduction in overall *S. pneumoniae* caused meningitis in the elderly (29). While data is still being acquired for the post PCV13 era, there has been a noteworthy increase in serotype 35B since introduction of the vaccine (12,30).

Nonencapsulated *S. pneumoniae*

In addition to capsular switch variants and serotype replacement strains, it is important to consider nonencapsulated strains (NESp) of *S. pneumoniae*. As per their name, these strains do not produce a capsular polysaccharide, but NESp also include strains that express novel serotypes or capsules below detectable levels. These strains have been associated with causing conjunctivitis, IPD, otitis media, and carriage (31). Strains can be divided into two classes. Class I types contain the *cps* locus which normally codes for the CPS; however, mutations or disruptions prevent the bacteria from making a capsule. Class II types, on the other hand, completely lack the *cps* locus and genes required to produce a capsule, and instead have novel genes and virulence factors

(31-33). Typically, the CPS is the primary virulence factor and without it strains cannot readily colonize or cause disease. However, NESp have been shown to colonize the nasopharynx (32). This is probably due to increased exposure of host cell receptors to surface proteins due to lack of capsule. The surface proteins of *S. pneumoniae* are important because they adhere to the hosts' cells. Many group II NESp have a novel surface protein, PspK (33). Strains that produce PspK have been shown to have increased colonization over strains lacking PspK (33). Additionally, NESp may be able to cause disease because many contain DltA and EndA proteins which have been shown to inhibit neutrophil extracellular traps, thus allowing them to evade bacterial clearance (32). Lastly, NESp create larger biofilms than encapsulated strains. This is again most likely due to increased exposure of the surface proteins, which adhere to the host cells (32). Larger biofilm formation also reduces the strains sensitivity to antibiotics and host response.

Discussion

Despite the current challenges with pneumococcal vaccine programs, there is no denying the effectiveness of the vaccines. Post vaccine era statistics on the decreases seen in pneumococcal pneumonia infections attest to this **(Figure 2.1)**. However, one important aspect of these vaccines that needs to be researched and addressed further is their production. Up to this point, conjugate vaccines have been constructed using poorly controlled conjugation chemistries to link carbohydrate and protein (34). These designs do not consider the mechanism of action within the body once administered, with gaps in knowledge of the required immune response to induce protection being a major scientific challenge (35). This leads to poorly-characterized, heterogeneous, and variably

immunogenic conjugate vaccines (36,37). Recently, a mechanism for conjugate vaccine-triggered immune response has been proposed (38,39). In this model, uptake and processing of conjugate vaccines by an antigen-presenting cell yield the presentation of a carbohydrate epitope via the major histocompatibility complex class II. This in turn stimulates carbohydrate specific T-cells (Tcarbs) and leads to the production of high-affinity carbohydrate specific IgGs (38,40,41). Considering the mechanism of action in producing conjugate vaccines will help produce knowledge-based vaccines that are target specific, structurally better-defined, both immunogenic and protective, and produced at a much lower cost (39). This will allow for use on a more global scale to help control pneumococcal infections. Additionally, this will aid in designing vaccines targeted specifically for certain at-risk groups, as it is becoming increasingly clear that vaccination, particularly in adults, is not a one-size-fits-all concept. A recent study has found that adults aged 55-74, when previously vaccinated with PPSV23, showed reduced OPA response to PCV13 (42). This is known as immune hyporesponsiveness, in which the patient shows poor immune response to PCV13 as a result of prior polysaccharide antigen administration in PPSV23 (42). Similarly, a study found that children who had previously received doses of PCV7 when boosted with PPSV23 at 12 months exhibited immune hyporesponsiveness compared to children who did not receive the PPSV23 booster (43). Fortunately, by age 5-7 these children showed normal immune response to PCV13 and the hyporesponsiveness did not persist (43).

Furthermore, while it is true that we see a rise in non-vaccine type serotypes post introduction of the vaccines, there is some evidence that points to this occurring naturally. In a study by Wyres et al., it was noted that capsular switching most likely

arises naturally in pneumococcal strains as well as post vaccine introduction (7). While vaccination programs do induce a selective pressure on pneumococci, which contributes to the serotype epidemiology, these naturally occurring fluctuations among serotypes play a contributing role. An example of vaccine induced selective pressure and natural selection occurring simultaneously was observed in the post PCV7 era (18). Croucher et al., points out that a number of successful lineages that were vaccine-type prior to PCV7 introduction will likely persist through variants that have acquired non-vaccine type capsules through natural transformation from vaccine pressure, i.e., serotype switching (18). As was the circumstance after the introduction of PCV7, it is probable that non-vaccine type capsular switch variants will persist and emerge from the PCV13 vaccine program. Fortunately, the extent of serotype replacement and serotype switching on IPD and pneumonia rates has been minimal to this point (29).

Awareness of NESp is important for a multitude of reasons. Firstly, there has been an increase in NESp prevalence since the introduction of conjugate vaccines. It was observed that strains increased from 1.5% in 2001 to 5.1% in 2006 (31). However, at any point in time NESp frequency can be between 3-19% worldwide (32). This is important, as current conjugate vaccines do not protect against NESp. One detail to note, however, is that while this increase in NESp prevalence was observed from use of conjugate vaccines, it is also due to more stringent typing of pneumococcal samples. Additionally, as pointed out above, these strains are still able to cause diseases, albeit at a lower frequency than encapsulated strains (32). These strains are also found to be highly resistant to antibiotics. One study out of Portugal found 7% of carriage isolates to be NESp, with all NESp strains being resistant to one antibiotic and 89% of them being

resistant to multiple (32,44). Another possibility is genetic exchange between NESp and encapsulated strains. The *cps* locus is considered a hotspot due to abnormally high horizontal gene transfer (23). NESp have been shown to acquire mutants at a higher rate than encapsulated strains (32). With the increase in nasopharynx colonization by NESp and non-vaccine type serotypes the likelihood of transfer between the two is great (33). This could transfer novel genes, such as *pspK*, and antibiotic resistance to encapsulated strains. NESp also contain novel virulence factors, which could potentially be transferred. Lastly, as the use of conjugate vaccines becomes more widespread, the increased pressure on *S. pneumoniae* strains to adapt becomes greater. Conjugate vaccines affect the niche that non-vaccine type and NESp can now exploit to expand. This leaves room for increased risk of IPD caused by NESp. Pressure from vaccines also forces the vaccine type serotypes to adapt in order to persist, leaving room for genetic exchange between them and NESp. Over all, the frequency of NESp and non-vaccine type strains will likely continue to grow as conjugate vaccines continue to target specific serotypes.

Lastly, when thinking about the rise in antibiotic-resistant strains of pneumococci, it is possible that a subsidizing factor in this is the increased use of antibiotics and broad-spectrum antibiotics, not just vaccine programs. There is evidence that the increase in disease caused by serotype 19A is also universally related to the increase in antibiotic use and therefore the trend of antibiotic resistance observed in this particular serotype (25,29). Moreover, there is a rise in infection from antibiotic-resistant strains seen in countries without a national pediatric vaccination program, leaving room for uncertainty whether the spike in 19A, and antibiotic-resistance in this strain, is solely from the introduction of PCV7 (29). As noted in a 2013 study, rates of antibiotic-resistance tend to

vary from region to region and are influenced by such factors as serotype epidemiology of that region, vaccination programs, and antibiotic usage (25). Additionally, the use of conjugate vaccines against *S. pneumoniae* has led to a reduction in antibiotic use to treat these infections, which hopefully lessens the pressure for resistance to evolve (45). This further corroborates the belief that both antibiotic-resistance and serotype replacement/switching are not solely influenced by vaccination programs, but act via natural and outside factors. However, with an estimated 700,000 deaths annually from antibiotic-resistant infections, and that number only expected to rise, this is an important issue to consider (45).

In conclusion, numerous studies of the pneumococcal vaccines have all shown increases in non-vaccine type serotypes post vaccine era, the persistence of vaccine type serotypes, and the connection of antibiotic resistance with certain serotypes. However, the greater impact of these vaccines is undeniable. The overall percentage of pneumonia has decreased worldwide since the introduction of the conjugate vaccines. Multiple studies show evidence of reduction in carriage from vaccine type strains in those vaccinated and importantly even the unvaccinated are protected indirectly through herd immunity. This has been an added benefit of vaccination programs where vaccinated individuals are protected from disease and colonization, however, by preventing disease they also subsequently reduce transmission of the pathogen to individuals who are not/cannot be vaccinated (45). Reservations over serotype replacement, antibiotic resistance, and nonencapsulated strains should not dissuade the use of conjugate vaccines nor refute the fact that these vaccines have impeded serious illnesses among the population.

References

1. Watson, D. A., Musher, D. M., Jacobson, J. W., and Verhoef, J. (1993) A brief history of the pneumococcus in biomedical research: a panoply of scientific discovery. *Clin Infect Dis* **17**, 913-924
2. Griffith, F. (1928) The Significance of Pneumococcal Types. *J Hyg (Lond)* **27**, 113-159
3. Jochems, S. P., Weiser, J. N., Malley, R., and Ferreira, D. M. (2017) The immunological mechanisms that control pneumococcal carriage. *PLoS Pathog* **13**, e1006665
4. Geno, K. A., Gilbert, G. L., Song, J. Y., Skovsted, I. C., Klugman, K. P., Jones, C., Konradsen, H. B., and Nahm, M. H. (2015) Pneumococcal Capsules and Their Types: Past, Present, and Future. *Clin Microbiol Rev* **28**, 871-899
5. Aliberti, S., Mantero, M., Mirsaeidi, M., and Blasi, F. (2014) The role of vaccination in preventing pneumococcal disease in adults. *Clin Microbiol Infect* **20 Suppl 5**, 52-58
6. AlonsoDeVelasco, E., Verheul, A. F., Verhoef, J., and Snippe, H. (1995) Streptococcus pneumoniae: virulence factors, pathogenesis, and vaccines. *Microbiol Rev* **59**, 591-603
7. Wyres, K. L., Lambertsen, L. M., Croucher, N. J., McGee, L., von Gottberg, A., Linares, J., Jacobs, M. R., Kristinsson, K. G., Beall, B. W., Klugman, K. P., Parkhill, J., Hakenbeck, R., Bentley, S. D., and Brueggemann, A. B. (2013)

Pneumococcal capsular switching: a historical perspective. *J Infect Dis* **207**, 439-449

8. Thigpen, M. C., Whitney, C. G., Messonnier, N. E., Zell, E. R., Lynfield, R., Hadler, J. L., Harrison, L. H., Farley, M. M., Reingold, A., Bennett, N. M., Craig, A. S., Schaffner, W., Thomas, A., Lewis, M. M., Scallan, E., Schuchat, A., and Emerging Infections Programs, N. (2011) Bacterial Meningitis in the United States, 1998-2007. *New England Journal of Medicine* **364**, 2016-2025
9. Hamborsky J, K. A., Wolfe S. (2015) Pneumococcal Disease. in *Epidemiology and Prevention of Vaccine-Preventable Diseases* (Prevention, C. f. D. C. a. ed.), 13 Ed., Public Health Foundation, Washington D.C. pp
10. Hausdorff, W. P., Bryant, J., Kloek, C., Paradiso, P. R., and Siber, G. R. (2000) The contribution of specific pneumococcal serogroups to different disease manifestations: implications for conjugate vaccine formulation and use, part II. *Clin Infect Dis* **30**, 122-140
11. (1997) Prevention of pneumococcal disease: recommendations of the Advisory Committee on Immunization Practices (ACIP). *MMWR Recomm Rep* **46**, 1-24
12. Daniels, C. C., Rogers, P. D., and Shelton, C. M. (2016) A Review of Pneumococcal Vaccines: Current Polysaccharide Vaccine Recommendations and Future Protein Antigens. *J Pediatr Pharmacol Ther* **21**, 27-35
13. Morens, D. M., Taubenberger, J. K., and Fauci, A. S. (2008) Predominant role of bacterial pneumonia as a cause of death in pandemic influenza: implications for pandemic influenza preparedness. *J Infect Dis* **198**, 962-970

14. Kochanek, K. D., Murphy, S. L., Xu, J., and Tejada-Vera, B. (2016) Deaths: Final Data for 2014. *Natl Vital Stat Rep* **65**, 1-122

15. Bonten, M. J., Huijts, S. M., Bolkenbaas, M., Webber, C., Patterson, S., Gault, S., van Werkhoven, C. H., van Deursen, A. M., Sanders, E. A., Verheij, T. J., Patton, M., McDonough, A., Moradoghli-Haftvani, A., Smith, H., Mellelieu, T., Pride, M. W., Crowther, G., Schmoele-Thoma, B., Scott, D. A., Jansen, K. U., Lobatto, R., Oosterman, B., Visser, N., Caspers, E., Smorenburg, A., Emini, E. A., Gruber, W. C., and Grobbee, D. E. (2015) Polysaccharide conjugate vaccine against pneumococcal pneumonia in adults. *N Engl J Med* **372**, 1114-1125

16. Feldman, C., and Anderson, R. (2014) Review: current and new generation pneumococcal vaccines. *J Infect* **69**, 309-325

17. Sun, L., Middleton, D. R., Wantuch, P. L., Ozdilek, A., and Avci, F. Y. (2016) Carbohydrates as T-cell antigens with implications in health and disease. *Glycobiology* **26**, 1029-1040

18. Croucher, N. J., Finkelstein, J. A., Pelton, S. I., Mitchell, P. K., Lee, G. M., Parkhill, J., Bentley, S. D., Hanage, W. P., and Lipsitch, M. (2013) Population genomics of post-vaccine changes in pneumococcal epidemiology. *Nature genetics* **45**, 656-663

19. Hamaluba, M., Kandasamy, R., Ndimah, S., Morton, R., Caccamo, M., Robinson, H., Kelly, S., Field, A., Norman, L., Plested, E., Thompson, B. A., Zafar, A., Kerridge, S. A., Lazarus, R., John, T., Holmes, J., Fenlon, S. N., Gould, K. A., Waight, P., Hinds, J., Crook, D., Snape, M. D., and Pollard, A. J. (2015) A cross-sectional observational study of pneumococcal carriage in children, their parents,

and older adults following the introduction of the 7-valent pneumococcal conjugate vaccine. *Medicine (Baltimore)* **94**, e335

20. Whitney, C. G., Farley, M. M., Hadler, J., Harrison, L. H., Bennett, N. M., Lynfield, R., Reingold, A., Cieslak, P. R., Pilishvili, T., Jackson, D., Facklam, R. R., Jorgensen, J. H., Schuchat, A., and Active Bacterial Core Surveillance of the Emerging Infections Program, N. (2003) Decline in invasive pneumococcal disease after the introduction of protein-polysaccharide conjugate vaccine. *N Engl J Med* **348**, 1737-1746

21. Rodenburg, G. D., de Greeff, S. C., Jansen, A. G., de Melker, H. E., Schouls, L. M., Hak, E., Spanjaard, L., Sanders, E. A., and van der Ende, A. (2010) Effects of pneumococcal conjugate vaccine 2 years after its introduction, the Netherlands. *Emerg Infect Dis* **16**, 816-823

22. (2016) Trends by Serotype Group, 1998-2015. in *Active Bacterial Core surveillance* (National Center for Immunization and Respiratory Diseases, D. o. B. D. ed., Centers for Disease Control and Prevention

23. Croucher, N. J., Finkelstein, J. A., Pelton, S. I., Parkhill, J., Bentley, S. D., Lipsitch, M., and Hanage, W. P. (2015) Population genomic datasets describing the post-vaccine evolutionary epidemiology of Streptococcus pneumoniae. *Sci Data* **2**, 150058

24. Hicks, L. A., Harrison, L. H., Flannery, B., Hadler, J. L., Schaffner, W., Craig, A. S., Jackson, D., Thomas, A., Beall, B., Lynfield, R., Reingold, A., Farley, M. M., and Whitney, C. G. (2007) Incidence of pneumococcal disease due to non-

pneumococcal conjugate vaccine (PCV7) serotypes in the United States during the era of widespread PCV7 vaccination, 1998-2004. *J Infect Dis* **196**, 1346-1354

25. Hackel, M., Lascols, C., Bouchillon, S., Hilton, B., Morgenstern, D., and Purdy, J. (2013) Serotype prevalence and antibiotic resistance in Streptococcus pneumoniae clinical isolates among global populations. *Vaccine* **31**, 4881-4887

26. Rosenblut, A., Napolitano, C., Pereira, A., Moreno, C., Kolhe, D., Lepetic, A., and Ortega-Barria, E. (2017) Etiology of acute otitis media and serotype distribution of Streptococcus pneumoniae and Haemophilus influenzae in Chilean children <5 years of age. *Medicine (Baltimore)* **96**, e5974

27. Dagan, R., Poolman, J., and Siegrist, C.-A. (2010) Glycoconjugate vaccines and immune interference: A review. *Vaccine* **28**, 5513-5523

28. Croucher, N. J., Harris, S. R., Fraser, C., Quail, M. A., Burton, J., van der Linden, M., McGee, L., von Gottberg, A., Song, J. H., Ko, K. S., Pichon, B., Baker, S., Parry, C. M., Lambertsen, L. M., Shahinas, D., Pillai, D. R., Mitchell, T. J., Dougan, G., Tomasz, A., Klugman, K. P., Parkhill, J., Hanage, W. P., and Bentley, S. D. (2011) Rapid pneumococcal evolution in response to clinical interventions. *Science* **331**, 430-434

29. Dagan, R. (2009) Serotype replacement in perspective. *Vaccine* **27 Suppl 3**, C22-24

30. Richter, S. S., Heilmann, K. P., Dohrn, C. L., Riahi, F., Diekema, D. J., and Doern, G. V. (2013) Pneumococcal serotypes before and after introduction of conjugate vaccines, United States, 1999-2011(1.). *Emerg Infect Dis* **19**, 1074-1083

31. Park, I. H., Kim, K. H., Andrade, A. L., Briles, D. E., McDaniel, L. S., and Nahm, M. H. (2012) Nontypeable pneumococci can be divided into multiple cps types, including one type expressing the novel gene pspK. *MBio* **3**

32. Keller, L. E., Robinson, D. A., and McDaniel, L. S. (2016) Nonencapsulated Streptococcus pneumoniae: Emergence and Pathogenesis. *MBio* **7**, e01792

33. Pipkins, H. R., Bradshaw, J. L., Keller, L. E., and McDaniel, L. S. (2018) Virulence of an Encapsulated Streptococcus pneumoniae is Increased Upon Expression of Pneumococcal Surface Protein K. *J Infect Dis*

34. Avci, F., Kasper, D., Paul, W., Littman, D., and Yokoyama, W. (2010) How Bacterial Carbohydrates Influence the Adaptive Immune System. *Annual Review of Immunology, Vol 28* **28**, 107-130

35. Sheerin, D., Openshaw, P. J., and Pollard, A. J. (2017) Issues in vaccinology: Present challenges and future directions. *Eur J Immunol* **47**, 2017-2025

36. Avci, F. (2013) Novel Strategies for Development of Next-Generation Glycoconjugate Vaccines. *Current Topics in Medicinal Chemistry* **13**, 2535-2540

37. Costantino, P., Rappuoli, R., and Berti, F. (2011) The design of semi-synthetic and synthetic glycoconjugate vaccines. *Expert Opinion on Drug Discovery* **6**, 1045-1066

38. Avci, F., Li, X., Tsuji, M., and Kasper, D. (2011) A mechanism for glycoconjugate vaccine activation of the adaptive immune system and its implications for vaccine design. *Nature Medicine* **17**, 1602-U1115

39. Middleton, D. R., Sun, L., Paschall, A. V., and Avci, F. Y. (2017) T Cell-Mediated Humoral Immune Responses to Type 3 Capsular Polysaccharide of. *J Immunol* **199**, 598-603

40. Avci, F., Li, X., Tsuji, M., and Kasper, D. (2012) Isolation of carbohydrate-specific CD4(+) T cell clones from mice after stimulation by two model glycoconjugate vaccines. *Nature Protocols* **7**, 2180-2192

41. Avci, F. Y., Li, X., Tsuji, M., and Kasper, D. L. (2013) Carbohydrates and T cells: a sweet twosome. *Semin Immunol* **25**, 146-151

42. Jackson, L. A., El Sahly, H. M., George, S., Winokur, P., Edwards, K., Brady, R. C., Rouphael, N., Keitel, W. A., Mulligan, M. J., Burton, R. L., Nakamura, A., Ferreria, J., and Nahm, M. H. (2018) Randomized clinical trial of a single versus a double dose of 13-valent pneumococcal conjugate vaccine in adults 55 through 74 years of age previously vaccinated with 23-valent pneumococcal polysaccharide vaccine. *Vaccine* **36**, 606-614

43. Licciardi, P. V., Toh, Z. Q., Clutterbuck, E. A., Balloch, A., Marimla, R. A., Tikkanen, L., Lamb, K. E., Bright, K. J., Rabuatoka, U., Tikoduadua, L., Boelsen, L. K., Dunne, E. M., Satzke, C., Cheung, Y. B., Pollard, A. J., Russell, F. M., and Mulholland, E. K. (2016) No long-term evidence of hyporesponsiveness after use of pneumococcal conjugate vaccine in children previously immunized with pneumococcal polysaccharide vaccine. *J Allergy Clin Immunol* **137**, 1772-1779 e1711

44. Sa-Leao, R., Nunes, S., Brito-Avo, A., Alves, C. R., Carrico, J. A., Saldanha, J., Almeida, J. S., Santos-Sanches, I., and de Lencastre, H. (2008) High rates of

transmission of and colonization by Streptococcus pneumoniae and Haemophilus influenzae within a day care center revealed in a longitudinal study. *J Clin Microbiol* **46**, 225-234

45. Jansen, K. U., Knirsch, C., and Anderson, A. S. (2018) The role of vaccines in preventing bacterial antimicrobial resistance. *Nat Med* **24**, 10-19

Table 2.1

Table 2.1: Description of past and present pneumococcal vaccines. Table was adapted from references 4 and 9.

Vaccine	Year Introduced	Carrier Protein	Serotypes	Impacts
PSV23	1983	N/A	1,2,3,4,5,6B,7F,8,9N,9V,10A,11A,12F,14,15B 17F,18C,19A,19F,20,22F,23F,and33F	-Reduced invasive disease in adults -Not protective in children
PCV7	2000	CRM197	4,6B,9V,14,18C,19F, and 23F	-Reduced invasive disease -More protective in children -Increase 19A and 7F infections
PCV10	2011 (Not in US)	NTHi protein D; tetanus toxoid; diphtheria toxoid	1,4,5,6B,7F,9V,14,18C,19F, and 23F	-Decrease in otitis media infection
PCV13	2010	CRM197	1,3,4,5,6A,6B,7F,9V,14,18C,19A,19F, and 23F	-Reduced invasive disease -Protection in all age/risk groups -Increase in 35B infections

Figure 2.1

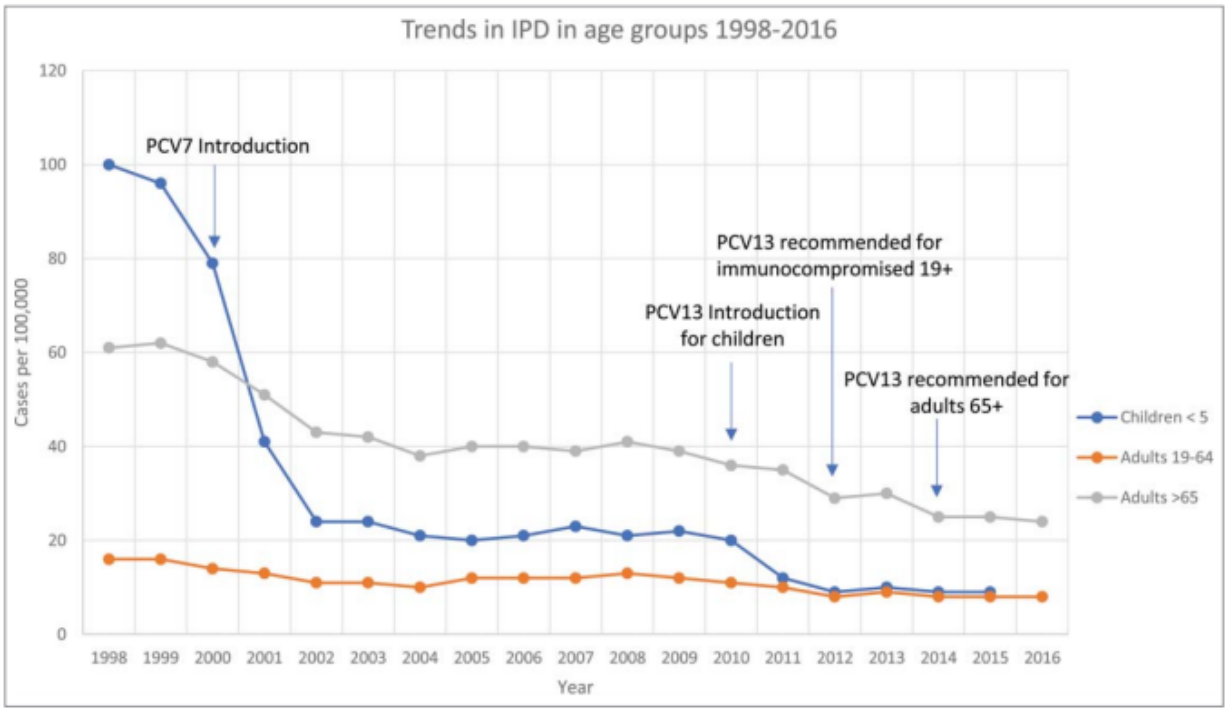
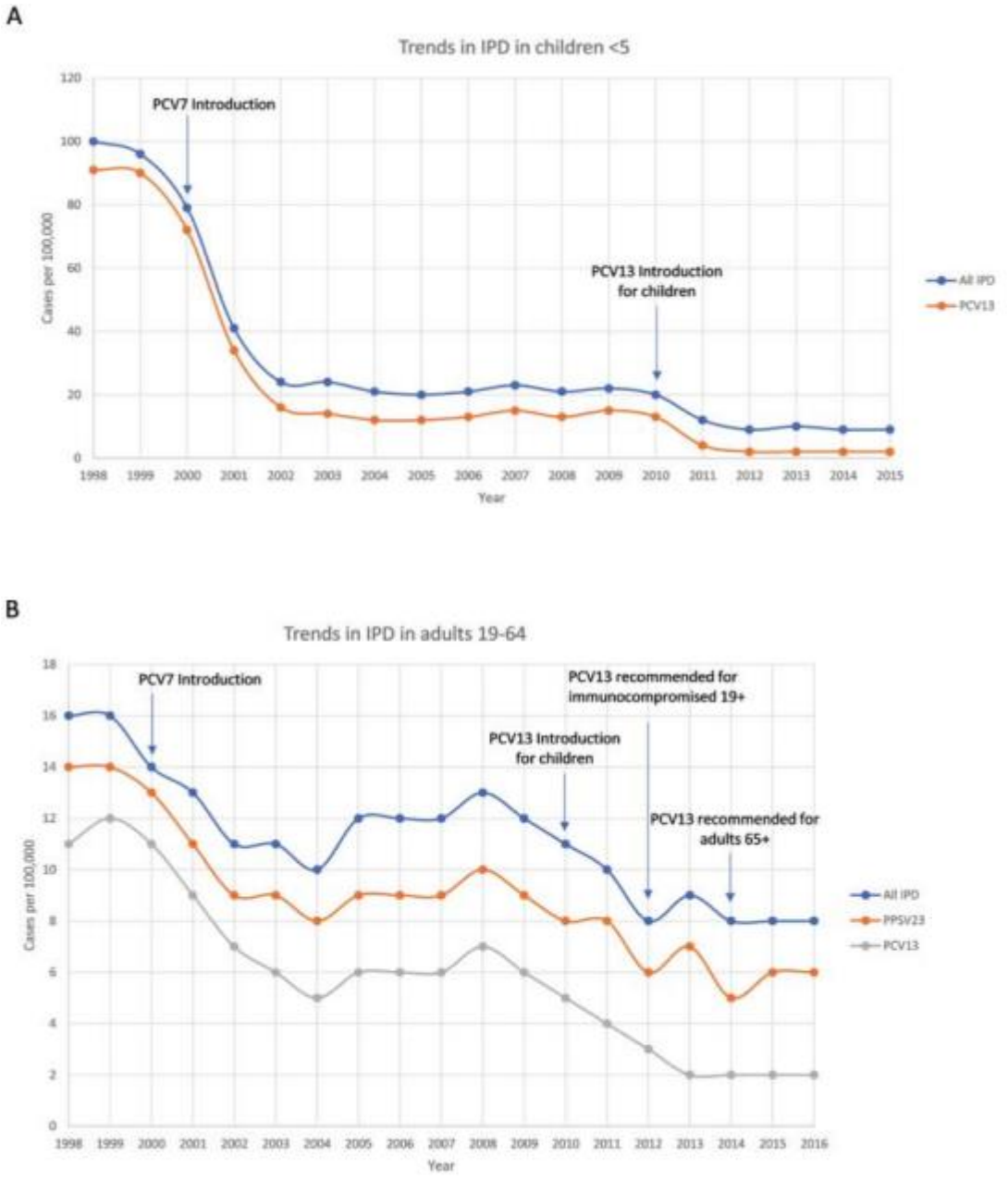


Figure 2.1: Trends in overall rate of IPD (invasive pneumococcal diseases) in different age groups between the years 1998-2016 in the United States. Blue line represents children under five, orange line adults ages 19-64, gray line adults over the age of 65. Important dates for vaccine introduction and recommendations are pinpointed. Data for children less than five was only available through 2015. Figure was constructed with data from the CDC ABCs bacterial surveillance programs, 2016 references 22.

Figure 2.2


A
Trends in IPD in children <5
PCV7 Introduction
PCV13 Introduction for children
Cases per 100,000
Year
All IPD
PCV13
B
Trends in IPD in adults 19-64
PCV7 Introduction
PCV13 recommended for immunocompromised 19+
PCV13 Introduction for children
PCV13 recommended for adults 65+
Cases per 100,000
Year
All IPD
PPSV23
PCV13

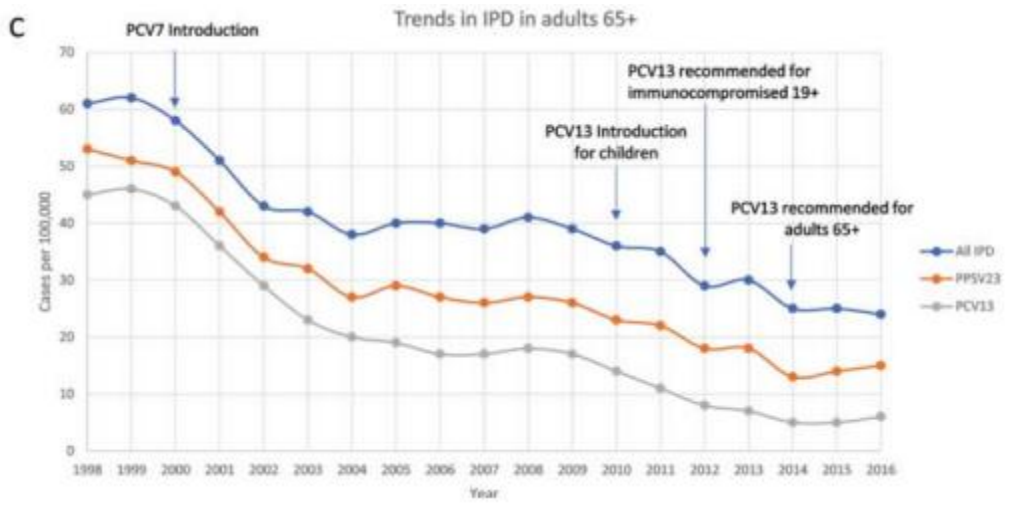


Figure 2.2: Trends in overall rate of IPD (invasive pneumococcal diseases), disease caused by PPSV23 included serotypes, and disease caused by PCV13 serotypes in different age groups between the year 1998-2016 in the United States. A) Children under 5, with only overall and PCV13 data available through 2015. B) Adults ages 19-64 C) Adults over 65. Figure was constructed with data from the CDC ABCs bacterial surveillance program, 2016 reference 22.

CHAPTER 3

INVASIVE PNEUMOCOCCAL DISEASE IN RELATION TO VACCINE TYPE SEROTYPES

Paeton L Wantuch and Fikri Y Avci. 2019. *Hum Vaccin Immunother.* 15(4), 874-875.

DOI: 10.1080/21645515.2018.1564444

Abstract

Since 1983 the world has been introduced to four vaccines combating disease caused by *Streptococcus pneumoniae* bacteria. However, despite vaccination programs disease caused by *S. pneumoniae* continues to lead to high morbidity and mortality worldwide. Surprisingly, instances of invasive pneumococcal disease (IPD) are still highly attributed to serotypes found in the current vaccine, such as serotypes 3 and 19A. Conversely, non-vaccine serotypes, such as 35B, are increasing and of rising interest. The persistence of vaccine type serotypes and the increase in non-vaccine type serotypes show the need for further research into conjugate vaccine design and the need for novel strategies to combat IPD.

Vaccine Type *S. pneumoniae* Strains

We would like to thank Ozkaya-Parlakay *et al.* for their response to our article and for further emphasizing a continuing cause for concern in invasive pneumococcal disease (IPD) caused by vaccine type *Streptococcus pneumoniae* strains. Despite the use of conjugate vaccines, we are seeing an increase in both non-vaccine type and vaccine type serotypes **(Figure 3.1)** (1), with the persistence or increase of vaccine types being particularly alarming. Of the vaccine type serotypes, types 3 and 19A are of great concern. The drastic increase seen in 19A **(Figure 3.1)**, is of interest as it included in the current marketed vaccine, PCV13, indicating a reduced efficacy of the vaccine to protect against this serotype. Additionally, this particular serotype has noted antibiotic resistance, more so than other serotypes (2-4). The rise in antibiotic resistant strains is becoming greater cause for concern and action as it is recently estimated that 30% of IPD cases are caused by *S. pneumoniae* resistant to one or more antibiotics (5).

The prevalence of disease caused by serotype 3, on the other hand, has not increased but rather stayed the same despite current vaccination programs **(Figure 3.1)**. This indicates reduced effectiveness of the vaccine against this serotype as well. Owing to its increased virulence and mortality rate, serotype 3 has raised global attention (6-9). In fact, serotype 3 is the second most common isolate of adult IPD and currently accounts for about 10% of all disease with that number ever increasing (6). In the recent study by Silva-Costa surveying pediatric pneumonia cases in Portugal between 2010-2015, they found that serotypes 3,1, and 19A accounted for 62% of all cases (10). With more and more data suggesting reduced vaccine effectiveness against type 3 it is imperative to seek new strategies to combat this serotype. Recently, work has been done to implement a capsule degrading enzyme against type 3 as a therapeutic agent (11,12).

It is our opinion that both vaccine type and non-vaccine type serotypes are emerging concerns in human health. This is due to increased antibiotic resistance seen in a growing number of serotypes (5), capsular switching and replacement, and the persistence of vaccine serotypes (4). This brings to light the need for novel strategies to fight IPD and the need for truly protective, new-generation conjugate vaccines that utilize effective immune mechanisms (13-15). However, it is imperative that these issues should not deter use of the current conjugate vaccines against *S. pneumoniae.* Numerous studies have shown the effectiveness of these vaccines (4), including recent work studying the administration of PCV13 in adults over the age of 65 and demonstrating its effectiveness against IPD (16). Remarkably, the prevalence of most vaccine-type serotypes has decreased since the introduction of conjugate vaccines **(Figure 3.1)**, which manifests the power of conjugate vaccines in combating infectious diseases.

References

1. Richter, S. S., Heilmann, K. P., Dohrn, C. L., Riahi, F., Diekema, D. J., and Doern, G. V. (2013) Pneumococcal serotypes before and after introduction of conjugate vaccines, United States, 1999-2011(1.). *Emerg Infect Dis* **19**, 1074-1083

2. Croucher, N. J., Harris, S. R., Fraser, C., Quail, M. A., Burton, J., van der Linden, M., McGee, L., von Gottberg, A., Song, J. H., Ko, K. S., Pichon, B., Baker, S., Parry, C. M., Lambertsen, L. M., Shahinas, D., Pillai, D. R., Mitchell, T. J., Dougan, G., Tomasz, A., Klugman, K. P., Parkhill, J., Hanage, W. P., and Bentley, S. D. (2011) Rapid pneumococcal evolution in response to clinical interventions. *Science* **331**, 430-434

3. Dagan, R. (2009) Serotype replacement in perspective. *Vaccine* **27 Suppl 3**, C22-24

4. Wantuch, P. L., and Avci, F. Y. (2018) Current status and future directions of invasive pneumococcal diseases and prophylactic approaches to control them. *Hum Vaccin Immunother* **14**, 2303-2309

5. Kim, L., McGee, L., Tomczyk, S., and Beall, B. (2016) Biological and Epidemiological Features of Antibiotic-Resistant Streptococcus pneumoniae in Pre- and Post-Conjugate Vaccine Eras: a United States Perspective. *Clin Microbiol Rev* **29**, 525-552

6. Sugimoto, N., Yamagishi, Y., Hirai, J., Sakanashi, D., Suematsu, H., Nishiyama, N., Koizumi, Y., and Mikamo, H. (2017) Invasive pneumococcal disease caused by mucoid serotype 3 Streptococcus pneumoniae: a case report and literature review. *BMC Res Notes* **10**, 21

7. Martens, P., Worm, S. W., Lundgren, B., Konradsen, H. B., and Benfield, T. (2004) Serotype-specific mortality from invasive Streptococcus pneumoniae disease revisited. *BMC Infect Dis* **4**, 21

8. Weinberger, D. M., Harboe, Z. B., Sanders, E. A., Ndiritu, M., Klugman, K. P., Rückinger, S., Dagan, R., Adegbola, R., Cutts, F., Johnson, H. L., O'Brien, K. L., Scott, J. A., and Lipsitch, M. (2010) Association of serotype with risk of death due to pneumococcal pneumonia: a meta-analysis. *Clin Infect Dis* **51**, 692-699

9. Briles, D. E., Crain, M. J., Gray, B. M., Forman, C., and Yother, J. (1992) Strong association between capsular type and virulence for mice among human isolates of Streptococcus pneumoniae. *Infect Immun* **60**, 111-116

10. Silva-Costa, C., Brito, M. J., Pinho, M. D., Friães, A., Aguiar, S. I., Ramirez, M., Melo-Cristino, J., Infections, P. G. f. t. S. o. S., and Society, P. S. G. o. I. P. D. o. t. P. I. D. (2018) Pediatric Complicated Pneumonia Caused by Streptococcus pneumoniae Serotype 3 in 13-Valent Pneumococcal Conjugate Vaccinees, Portugal, 2010-2015. *Emerg Infect Dis* **24**, 1307-1314

11. Middleton, D. R., Zhang, X., Wantuch, P. L., Ozdilek, A., Liu, X., LoPilato, R., Gangasani, N., Bridger, R., Wells, L., Linhardt, R. J., and Avci, F. Y. (2018) Identification and characterization of the Streptococcus pneumoniae type 3

capsule-specific glycoside hydrolase of Paenibacillus species 32352. *Glycobiology* **28**, 90-99

12. Middleton, D. R., Paschall, A. V., Duke, J. A., and Avci, F. Y. (2018) Enzymatic Hydrolysis of Pneumococcal Capsular Polysaccharide Renders the Bacterium Vulnerable to Host Defense. *Infect Immun*
13. Avci, F. Y., Li, X. M., Tsuji, M., and Kasper, D. L. (2011) A mechanism for glycoconjugate vaccine activation of the adaptive immune system and its implications for vaccine design. *Nature Medicine* **17**, 1602-1610
14. Sun, L., Middleton, D. R., Wantuch, P. L., Ozdilek, A., and Avci, F. Y. (2016) Carbohydrates as T-cell antigens with implications in health and disease. *Glycobiology* **26**, 1029-1040
15. Middleton, D. R., Sun, L., Paschall, A. V., and Avci, F. Y. (2017) T Cell-Mediated Humoral Immune Responses to Type 3 Capsular Polysaccharide of. *J Immunol* **199**, 598-603
16. Bonten, M. J., Huijts, S. M., Bolkenbaas, M., Webber, C., Patterson, S., Gault, S., van Werkhoven, C. H., van Deursen, A. M., Sanders, E. A., Verheij, T. J., Patton, M., McDonough, A., Moradoghli-Haftvani, A., Smith, H., Mellelieu, T., Pride, M. W., Crowther, G., Schmoele-Thoma, B., Scott, D. A., Jansen, K. U., Lobatto, R., Oosterman, B., Visser, N., Caspers, E., Smorenburg, A., Emini, E. A., Gruber, W. C., and Grobbee, D. E. (2015) Polysaccharide conjugate vaccine against pneumococcal pneumonia in adults. *N Engl J Med* **372**, 1114-1125

Figure 3.1

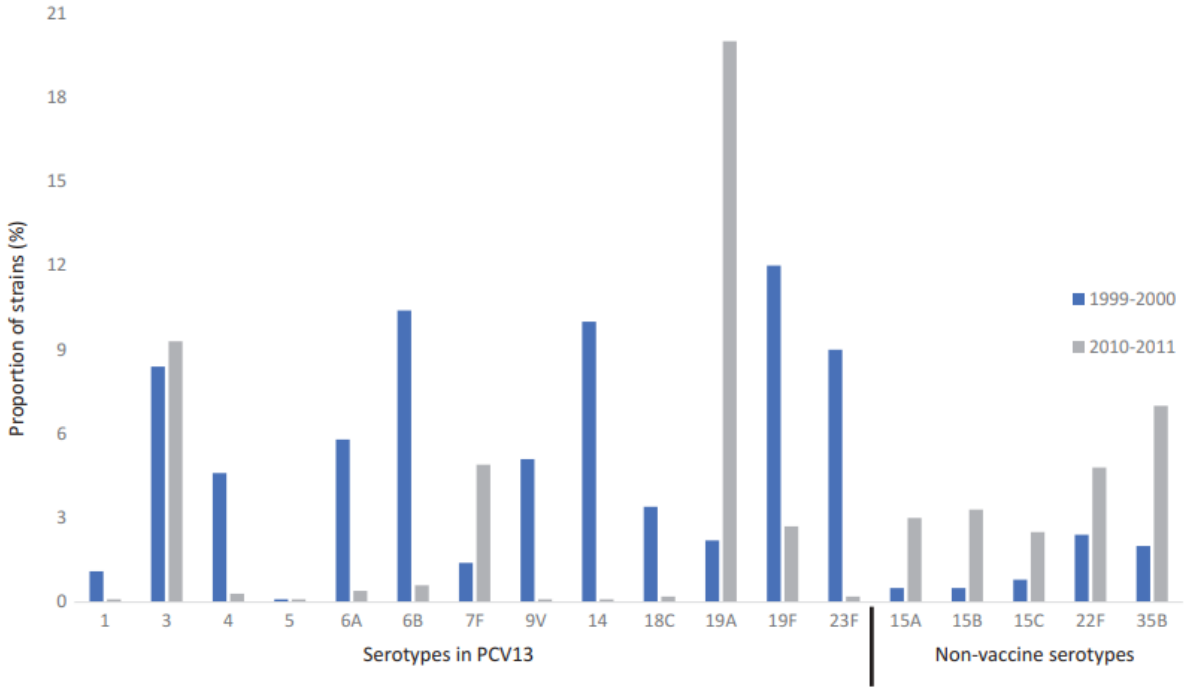


Figure 3.1: Serotype distribution before and after vaccine introduction, all ages. Distribution of vaccine type and non-conjugate vaccine type serotypes before (blue) and after (gray) implementation of conjugate vaccines. Figure was constructed with data from reference 1. Note: Serotypes 15B and 22F are included in the 23-valent polysaccharide vaccine, PPSV23. The term “non-vaccine” used throughout the articles refers to “non-conjugate vaccine” serotypes.

CHAPTER 4

MOLECULAR MECHANISMS FOR CARBOHYDRATE PRESENTATION TO CD4+ T CELLS VIA MHCII PATHWAY

Wantuch PL *et. al.*, unpublished. (In collaboration with Drs. Robert J Woods and Robert J Linhardt**)**

*****To be submitted to *Journal of Immunology*

Abstract

The introduction of glycoconjugate vaccines has led to great strides in combating bacterial infections. However, despite the successes of these vaccines infections still persist, especially in high risk populations. This is in part due to not fully understanding the immune mechanisms that govern glycoconjugate vaccine immune activation leading to variably constructed vaccines. Recently, a model for carbohydrate specific T cell immune activation has been elucidated; however, the molecular mechanisms of epitope presentation are still unclear. Using a library of covalently linked glycan-peptide conjugates, purified MHCII proteins and T cells this study works to shed light on these mechanisms. To initiate a T cell dependent immune response, glycan-peptides bind with MHCII leaving the glycan exposed for T cell recognition. The T cell receptor specifically binds the glycan to activate T cell help to produce carbohydrate specific immune response. The mechanism identified here will help create knowledge based glycoconjugates to better combat infections.

Introduction

Many bacterial pathogens express a large molecular weight capsular polysaccharide (CPS) that coat their external surface. Due to the their high antigenicity and accessibility CPSs have long been targeted as the main component of glycoconjugate vaccines (1). Glycoconjugate vaccines are composed of a bacterial pathogen, such as the CPS, covalently linked with a carrier protein. Over the last three decades glycoconjugate vaccines have been used to combat infectious disease in pathogens such as *S. pneumoniae*, *H. influenzae* and *N. meningitidis* (2,3). The covalently coupling of the CPS to a carrier protein leads to T cell help for B cells to produce a strong antibody response

against the bacterial CPSs (4-7). Immunizations with these vaccines induces IgM to IgG class switching, B and T cell memory and increased protection (8,9). While glycoconjugates have proven successful to help control bacterial disease, their immunogenicity has been variable especially in high risk populations (10). This has been attributed to construction of these vaccines, which is empirically linking these two molecules with minimum consideration for their mechanism of immune activation (7).

It has long been postulated that carbohydrates are T cell independent antigens and that glycoconjugate induced immune activation was due to a peptide, generated from the carrier protein, binding with major histocompatibility class II (MHCII) and being presented to CD4+ T cells (11). However, this hypothesis overlooks the strong covalent linkage between the two molecules, which is unlikely to be broken even under endolysosomal conditions (7,12). Recently, a model of glycoconjugate vaccine immune activation has been demonstrated which leads to carbohydrate specific CD4+ T cell activation via the presentation of a glycan-peptide by MHCII (13-15). This model reveals that under endosomal conditions the covalent bond between the two molecules is not broken; however, they are processed into smaller fragments, glycan-peptides. These glycan-peptides bind with MHCII and are presented at the cell surface where carbohydrate specific CD4+ T cells will bind the glycan activating the T cell to help the B cell produce glycan specific antibodies (13). Importantly, this model continues to be shown, demonstrating its wider applicability (16-18). However, the structural nature of the molecular mechanisms governing this model remain unclear.

In this current study, we utilized a library of glycan-peptides to mimic processed glycoconjugates that would be generated within the endosome to test the molecular

mechanisms underlying antigen processing and presentation of glycoconjugates. These glycoconjugates consist of defined lengths of oligosaccharides from the CPS of type 3 *S. pneumoniae* covalently coupled to the carrier peptide OVAp. This library was tested for the ability to bind with purified MHCII molecules and for T cell activation. Additionally, computational work went into elucidating the structure of MHCII binding with the glycan-peptides. A thorough understanding of the molecular mechanisms leading to glycoconjugate vaccine induced immune activation is of utmost importance to design next generation vaccines to combat infections not only against highly virulent type 3 *Spn* used in this study, but also other bacterial pathogens.

Material and Methods

Mice

We used primary cells from either wild type BALB/c mice or OVAp specific TCR transgenic mice [C-Cg-Tg(DO11.10)10D10/J] from Jackson Laboratories in all T cell assays. Mice were housed in Center for Molecular Medicine Rodent Vivarium at the University of Georgia. Mice were kept in microisolator cages and handled under a hood. All animal experiments were in compliance with the University of Georgia Institutional Animal Care and Use Committee under the approved animal use protocol # A2019 11-008-Y1-A0

Peptide Synthesis

We synthesized OVAp 323-339 and five derivatives modified with either C or K residues. Sequences of peptides can be found in **Table 4.1**. All couplings for peptides were carried out on an automated microwave-assisted solid-phase peptide synthesizer (CEM Corp. Liberty microwave synthesizer) using the standard protocols in the

instrument software. Peptides were synthesized on Rink amide resin (0.6 meq/g; Novabiochem) via N^R-N-(9-fluorenyl)methoxycarbonyl (Fmoc) approach in the primary solvent N,N-dimethylformamide (DMF). 20% 4-methylpiperidine in DMF was used for Fmoc removal. 2-(1*H*-Benzotriazole-1-yl)-oxy-1,1,3,3-tetramethyluronium hexafluorophosphate/1-hydroxybenzotriazole in the presence of *N*,*N*-diisopropylethylamine (DIPEA) were used as the coupling reagents. Peptides were cleaved from the resin through TFA/triisopropylsilane/H_2O (95:2.5:2.5) cocktail for ~2 hours. The cleavage cocktail was added dropwise through a filter to cold ether to precipitate the crude peptide and centrifuged to remove the ether supernatant. Purity was verified by analytical HPLC and MALDI-TOF MS.

Conjugations

Purified Pn3P (type 3 *Spn* CPS) powder was obtained from American Type Cell Collection (ATCC 172-X). Pn3P was reduced to oligosaccharides utilizing acid hydrolysis as previously described (17,19). Various lengths of oligosaccharides (tetra, hexa or octasaccharide) were conjugated with OVAp derivatives using reductive amination or chemoselective ligation as previously described (13,14,17).

T cell Activation

Splenic mononuclear cells were harvested fresh from DO11.10 transgenic or WT mice. The culture medium for the PBMCs was RPMI 1640 (Corning) supplemented with 2 g/L sodium bicarbonate, 50 µM 2-mercaptoethanol, 2 mM L-glutamine, 1 mM sodium pyruvate, 1% nonessential amino acids, 1% penicillin/streptomycin, and 10% heat inactivated FBS. Activation assays were performed with full splenocytes using $5x10^5$ cells/well. Cells were plated in duplicate per antigen in 200 µL supplemented RPMI in

96-well flat bottom plate. Cells were stimulated with each of the following antigens at 2 µg/ml: OVAp, scrambled OVAp, Nterm C/K OVAp, Cterm C/K OVAp, A326C/K, H328C/K, H331C/K, H331K-Octa, H328K-Tetra, H328K-Octa, A326K-Hexa, or Nterm-Tetra. Cells were harvested after 48 hours for activation assessment. The extent of activation was measured by percent CD69 positive among CD4+ T cells using anti-mouse CD4 and CD69 antibodies (Biolegend) in flow cytometry analysis (CytoFLEX, Beckman Coulter). Live cells were gated as CD69+ in CD4+ populations. Basal growth rate was determined from duplicate wells that contained full splenocytes without stimuli. All flow data was analyzed using FlowJo software (Tree Star).

IL-2 Cytokine ELISA

Cytokine production from T cell activation was determined by ELISA. 96-well plates (Costar) were coated overnight with anti-IL-2 antibody (1:1000 dilution; Biolegend) and blocked with 1% BSA in PBS. Plates were washed with 0.05% PBST, all subsequent washes were the same. After washing, wells were incubated with cell supernatants from T cell activation assays for 2 hours at room temperature. After washing, biotinylated anti-IL-2 (1:1000 dilution; Biolegend) was added for 1.5 hour at room temperature followed by HRP-Avidin (1:1000 dilution; Biolegend) for 1 hour at room temperature. Plates were developed using TMB substrate (Biolegend) and stopped with 2 N H_2SO_4. The optical densities were determined at 450 nm using a microplate reader (Synergy H1, Bio-Tek). Significance was determined using Student's *t* test with $p<0.05$.

MHCII Binding Assays

Purified MHCII monomers (mouse allele I-A^d) were graciously provided by the NIH tetramer facility. Monomers came peptide loaded with a 3C protease cleavage site. This

peptide was cleaved with 1µ of 3C protease (Pierce™ HRV 3C Protease Solution Kit) per 200µg monomer for 8 hours at room temperature. Cleaved monomer is then loaded with desired peptide or glycoconjugate through a peptide exchange reaction. Peptide exchange reaction use 200µg cleaved monomer with 270µM peptide (or glycan-peptide) sample at pH5 in a citrate buffer. Samples for this experiment were OVAp, scrambled OVAp, Nterm C/K OVAp, Cterm C/K OVAp, A326C/K, H328C/K, H331C/K, H331K-Octa, H331K-Hexa, H331K-Tetra, H328K-Hexa, A326K-Octa, or A326K-Hexa. Reactions were incubated for 5 days at room temperature. At the end of the incubation, reactions were neutralized with 1M sodium phosphate buffer pH 7.5 and spun down at max speed for 10 minutes to remove aggregates. Absorbance was measured at 280, all samples except the scrambled OVA peptide reaction gave significant concentration values. Binding was confirmed by isoelectric focusing (Novex™ pH 3-7 Protein IEF gels). Gel was visualized using silver stain (Pierce™ Silver Stain Kit).

MHCII Tetramer Preparation and Staining

MHCII monomers exchanged with OVAp were used to make tetramers. Starting with 20µL of monomer, 6µl of Streptavidin-APC antibody (Biolegend) was added every 10 minutes until 60µL total antibody had been added. Since monomers are biotinylated tetramers will form with the addition of streptavidin. To test if tetramers could bind/stain live CD4+ T cells, full splenocytes were collected from WT or DO11.10 mice. Cells were added as $5x10^5$ cells/tube in supplemented RPMI media (Corning). Cells were plated in duplicate and were incubated with no antigen or with tetramer for 3 hours at 37 degrees. After washing all cells were stained with anti-mouse CD4 antibody (Biolegend) and

analyzed in flow cytometry (CytoFLEX, Beckman Coulter). The extent of T cell bind by tetramers was measured as percent of APC+ populations in CD4+ populations.

MHCII Tetramer Activation

To test if tetramers could activate T cells *in vitro* T cell stimulation was used. 96 well plates (Thermo) were coated with no antigen or with tetramer at 2µg/mL. Full splenocytes were collected from WT or DO11.10 mice and CD4+ T cells were isolated using a negative selection CD4 enrichment kit (BD Biosciences). CD4+ T cells were plated as $5x10^4$ cells/well in supplemented RPMI (Corning) media and incubated for 48 hours at 37 degrees in 5% CO_2. Cells were collected, washed and stained with anti-mouse CD4 (Biolegend) and anti-mouse CD25 (Biolegend). The extent of activation was measured by percent CD25+ in CD4+ T cells in flow cytometry analysis (CytoFLEX, Beckman Coulter). Live cells were gated as CD69+ in CD4+ populations. Basal growth rate was determined from duplicate wells that contained full splenocytes without coat/stimuli.

Molecular Dynamics

Crystal structure of MHCII in complex with ovalbumin peptide 323-339 (PDBID 1IAO) was used as a starting structure for all models (20). The structure was modified by adding a neutral lysine residue at position indicated in Table 4.1 to match synthetic peptides. FF12SB and GLYCAM06 (J-1) force field were used to parameterize the systems (21,22). Varying lengths of oligosaccharides were attached to the structure. The systems were neutralized with counter ions and solvated with TIP3P in a rectangular box. Three energy minimization steps were performed each with 5,000 steps steepest descent followed by 20,000 steps of conjugate gradient methods. The three minimizations first

restrained all atoms, second restrained the protein backbone and carbohydrate ring atoms (if necessary) and the last left the system completely unrestrained. The system was heated from 5 to 300K over a span of 100 picoseconds using Langevin thermostat under NVT conditions. The MD simulations were performed under the same conditions as equilibrium for 20 nanoseconds.

Binding Affinity Calculations

Computational binding affinities were calculated using 20,000 snapshots extracted evenly from 20ns of MD simulations using a single trajectory method with the MMPBSA.py.MPI module of AMBER (23). The net binding energies were calculated as the difference between the energies of the complex minus the protein and ligand.

Results

OVAp and derivatives stimulated transgenic TCR

To elucidate the mechanism in which glycan-peptides bind with MHCII we started with a model peptide OVAp 323-339. This is a well-known MHCII binding peptide, with ample evidence showing it binding with mouse allele I-A^d (20,24-26) and evidence suggesting binding with human allele HLA-DR (13). The glycan-peptide conjugates which will be utilized in this study require either a cysteine or lysine amino acid residue to link with the oligosaccharide. To this end we modified the sequence of the OVAp either at the N-terminus, C-terminus or positions 326, 328, or 331. A scrambled version of the original OVAp was used as a negative control. The sequence of these peptide can be seen in **Table 4.1**. To assess whether the modifications made to the peptides alter their ability to stimulate T cells, we utilized the transgenic mouse line DO11.10 in which TCRs are specific for the OVAp. Using splenic mononuclear cells in

an *in vitro* T cell assay we tested T cell activation via flow cytometry and IL-2 cytokine production. As indicated by CD69+ cells in CD4+ populations, OVAp and derivatives all significantly activated CD4+ T cells, while the scrambled peptide and no antigen controls did not **(Figure 4.1A,B).** Similarly, all other peptides (OVAp and derivatives) produced significant levels of IL-2 cytokine while the scrambled version of OVAp gave no significant response comparable to the no antigen control **(Figure 4.1C)** indicating T cell activation. This data demonstrates that while the scrambled version of OVAp acts as a negative control all other OVAp derivatives have the ability to activate transgenic T cells as well as the unmodified OVAp, showing these peptide derivatives are comparable to OVAp as T cell epitopes.

Glycan-peptides inhibit T cell activation

With the production of a library of peptides we then attached different lengths of oligosaccharides through the Lys or Cys residues as previously described in collaboration with the Linhardt lab (17). This left us with an initial library of five peptides having either a tetra, hexa or octasaccharide attached at differing locations to simulate glycan-peptides which may be produced within the antigen presenting cell under endosomal conditions. The CPS of type 3 *Spn* consists of repeating disaccharides of glucose (Glc) and gluronic acid (GlcA). Therefore, the oligosaccharides used in this study are either two (tetra), three (hexa) or four (octasaccharide) repeating units. In a Tcarb (carbohydrate specific T cell) mediated immune response we hypothesize that the peptide of the glycan-peptides binds with MHCII leaving the glycan exposed to be bound by the TCR. Following this hypothesis, we used the transgenic mouse DO11.10 which should have an inhibited T cell response if the TCR is binding glycan and not peptide. T cell activation

was measured by CD69+ cells in CD4+ populations, all glycoconjugates had inhibited T cell activation compared with peptide alone **(Figure 4.2A,B)**. Similarly, as measured by IL-2 cytokine production the unmodified OVAp gave a robust T cell response as expected, but glycan-peptide conjugates gave no significant response **(Figure 4.2C)**. Taken together this suggests that the peptide specific transgenic TCR are binding with the exposed oligosaccharide thus inhibiting their T cell activation as our hypothesis states.

Peptide and glycoconjugates bind with MHCII

It is possible that the inhibited T cell response observed from the glycoconjugates could be due to their inability to bind with MHCII thus no antigen would be present for T cell binding. Our hypothesis states that the peptide portion of the conjugate binds with MHCII leaving the glycan exposed. Therefore, we first needed to test whether the modifications we made to the OVAp derivatives affect the peptides ability to bind MHCII. Using an isoelectric focusing gel we measured the peptides binding with MHCII. Compared to OVAp all peptide derivatives bind with MHCII as indicated by the presence of an intense band, although H328K peptide (lane 4) does give a lesser faint band **(Figure 4.3A).** However, due to data in **Figure 4.1** we know that H328K peptide stimulates T cells therefore must have the ability to bind MHCII. Contrary to this, the two negative controls scrambled peptide and hexasaccharide show no significant binding with MHCII as evidenced by the absence of a protein band **(Figure 4.3A)**. It is known that empty MHCII proteins are unstable and thus aggregate in solution (27), which explains why no protein is present when peptides do not bind.

Evidence of a glycan-peptide binding with MHCII is evident not only by the presence of a band, but additionally by a downward shift in protein location as the

slightly negative charge of the glycan would shift the isoelectric point of the protein. The glycoconjugates NtermC-tetrasaccharide **(Figure 4.3B)** and A326K-hexasacchardie **(Figure 4.3C)** give evidence of binding with MHCII as judged by these two properties. Importantly, both of these conjugates also inhibited DO11.10 T cell activation **(Figure 4.2).** Taken together these preliminary data suggest that a subset of the glycoconjugates bind with MHCII and inhibit transgenic T cell activation supporting the Tcarb mediated immune response hypothesis.

OVAp tetramers bind and activate T cells

MHCII tetramers are useful tools to bind and identify subsets of T cells. It is our hope to use these tetramers to isolate the population of T cells that are specific to the Pn3P carbohydrate. To optimize this protocol, we first generated tetramers that were bound with the OVAp **(Figure 4.4)**. To test the efficiency of the tetramers and confirm the OVAp is bound with the MHCII, we first performed an *in vitro* T cell assay using WT mice or the transgenic DO11.10 mice. Isolated CD4+ T cells from each group were incubated with either no antigen or the OVAp tetramers. Compared to the no antigen control, the tetramers were able to significantly activate DO11.10 T cells **(Figure 4.4A,B)**. Importantly, there was not a significant difference in activation between the no antigen or tetramer incubated WT as these were naïve mice never exposed to OVA antigen **(Figure 4.4A,B**). Additionally, the positive control CD3 incubated cells in both groups gave a robust positive response (**Figure 4.4A,B).** Next, we tested the tetramers ability to bind a population of CD4+ T cells. Total splenic mononuclear cells were harvested from either WT or DO11.10 mice and incubated with no antigen or OVAp tetramers. Using flow cytometry, the tetramer bound CD4+ T cells were separated

(tetramers are labeled with APC (what APC stands for)). This population would contain CD4+ T cells that tetramers bound to. Compared to no antigen control, the tetramers were able to bind a population of CD4+ T cells **(Figure 4.4C,D).** Importantly, the tetramers were not able to bind a significant population of cells in WT mice **(Figure 4.4C,D)**. At this time, work is still in progress to produce tetramers bound with glycan-peptide. However, this data suggests we are able to produce MHCII tetramers that both activate and bind CD4+ T cells of a specified population.

MHCII binding glycan-peptides

To visualize the structure of MHCII binding with the peptide and glycan-peptides we generated a model based on an existing crystal structure in collaboration with the Woods lab. We used PBDID 1IAO to model our OVAp and glycoconjugates into the binding groove**. Figure 4.5** depicts models of H331K-Tetra **(Figure 4.5A)** and Nterm-Tetra **(Figure 4.5B)** glycoconjugates. As the model shows, both conjugates bind in the MHCII binding groove and the peptide portion is what actually binds the MHCII groove leaving the glycan exposed. Additionally, the attached oligosaccharides are positioned in such a way they are exposed for potential TCR binding **(Figure 4.5)**. To help understand the binding requirements of MHCII with peptides and glycan-peptides we used MM-GBSA analysis **(Figure 4.6)**. This analysis is useful to estimate the free energy of binding of small ligands, such as peptides, to their receptor (28). **Figure 4.6A** gives the binding profile of OVAp with MHCII. Compared to that, there are slight differences in residues contributing to binding of glycan-peptides with MHCII **(Figure 4.6B)**. The more negative the binding value the color of the residue will become more red as evidenced by the scale below the figure. Additionally, the more negative value the more this amino

acid contributes to binding the ligand. It appears that certain residues of MHCII contribute more to binding the glycan-peptide than peptide alone. Indeed, there were five amino acids in particular that contributed more to binding in the presence of a glycan **(Figure 4.6C)**. We compared binding energies of OVAp alone, Nterm-Tetra, Cterm-Tetra, and H331K-Tetra glycoconjugates. I73, N63, and R77 contributed to binding in all simulations; however, contributed more when a glycan was present. Further, residues Y23 and L74 only contributed to binding when a glycan was present and not peptide alone **(Figure 4.6C)**. Taken together, these data suggest that there may be differences in binding requirements between peptide and glycoconjugates, but that both bind MHCII. Additionally, binding the glycoconjugates potentially positions the glycan in such a way that it in exposed for TCR binding.

Discussion

The insight into the immune mechanisms of glycoconjugate vaccine immune activation laid out in this chapter will aid in the design of novel knowledge based vaccines. Recent studies have begun to shed light on this subject (13,16,18); however, the structural requirements were still unclear. Herein we have reported evidence to strength the existing working model of Tcarb mediated immune response. This data suggests that the peptide portion of the glycoconjugate binds with MHCII protein leaving the attached glycan exposed for recognition by a carbohydrate specific T cell as previously proposed. While peptide alone is shown to bind MHCII, a hexasaccharide displayed no binding. However, a glycan-peptide consisting of the same peptide and hexasaccharide covalently linked together now binds MHCII. Further, the idea of T cells binding to carbohydrates was further strengthened through the use of peptide specific transgenic mice. While

peptide alone heavily activated these T cells as expected, glycoconjugates constructed with the same peptide inhibited T cell activation proposing the T cell binds to glycan, not peptide. With a greater understanding of the molecular mechanisms leading to antigen processing and presentation of glycoconjugates to stimulate a carbohydrate specific T cell response, we can begin applying this knowledge to vaccine design. A process which is often empirically linking the two molecules (protein and carbohydrate) without considering these mechanisms.

Further, the use of peptide based glycoconjugates in this study was not simply to mimic endosomal conditions. Rather, a strategy based on logic and understanding of the immune system to shift towards peptide based glycoconjugate vaccines. Indeed, several studies have explored this avenue in vaccine design (13,29-35). These studies have demonstrated that peptide vaccine candidates work as well as (36-38) or better than (13) full protein vaccines. The peptide based glycoconjugates accurately recapitulates what is processed in the endosome, therefore are the smallest units possible for MHCII binding. This means these vaccine candidates are not likely to be further degraded once administered, as traditional vaccines are, which could alter their immune activity. By supplying the immune system with the requirements for MHCII and T cell binding, we ensure that every epitope is utilized for a more robust T cell response.

For years, it has been thought that carbohydrates are T cell independent antigens. However, we postulate that it is not the T cells inability to bind carbohydrate, but rather the inability of MHCII to bind carbohydrates. Thus, MHCII is ill equipped to present such a carbohydrate epitope to T cells without the covalent linkage of a peptide. Besides the recent studies demonstrating a carbohydrate specific T cell response to challenge the

idea of glycans being T cell independent (13,16,18), there is increasing evidence of the immune system interacting with non-protein antigens such as carbohydrates. For example, zwitterionic polysaccharides have been shown to activate T cells (39,40). Glycolipids are presented to T cells by MHCI-like proteins CD1b and CD1d (41,42). Lastly, glycopeptides have been shown to be presented to CD4+ and CD8+ T cells (43-46). All these reports challenge the idea of carbohydrates being T cell independent antigens and strengthens the idea of Tcarb mediated immune response explored in this current study.

Despite the insights into molecular mechanism of glycoconjugate immune activation presented in this study, work remains to be done. To fully understand the Tcarb mediated immune response a thorough understanding of T cell receptor (TCR) binding is also necessary. It is important to understand what drives T cell binding to carbohydrates at the structural level of the TCR where there could be unique sequence changes allowing for such binding. Further, to test these glycoconjugates as possible vaccine candidates it is necessary to develop a larger library of possible epitopes. This requires moving away from the model OVAp used in this study and towards clinically relevant peptides. Indeed, this exact solution is explored in the next chapter. Once this library is constructed immunizations to test vaccine efficacy are also essential. However, the work shown here is a steppingstone towards future work into knowledge based glycoconjugate vaccines and is invaluable to our understanding of the immune mechanisms.

References

1. Avci, F. (2013) Novel Strategies for Development of Next-Generation Glycoconjugate Vaccines. *Current Topics in Medicinal Chemistry* **13**, 2535-2540
2. Trotter, C., McVernon, J., Ramsay, M., Whitney, C., Mulholland, E., Goldblatt, D., Hombach, J., Kieny, M., and Subgrp, S. (2008) Optimising the use of conjugate vaccines to prevent disease caused by Haemophilus influenzae type b, Neisseria meningitidis and Streptococcus pneumoniae. *Vaccine* **26**, 4434-4445
3. Weintraub, A. (2003) Immunology of bacterial polysaccharide antigens. *Carbohydr. Res.* **338**, 2539-2547
4. Li, P., and Wang, F. (2015) Polysaccharides: Candidates of promising vaccine adjuvants. *Drug Discov Ther* **9**, 88-93
5. Vella, M., and Pace, D. (2015) Glycoconjugate vaccines: an update. *Expert Opin Biol Ther* **15**, 529-546
6. Zimmermann, S., and Lepenies, B. (2015) Glycans as Vaccine Antigens and Adjuvants: Immunological Considerations. *Methods Mol Biol* **1331**, 11-26
7. Avci, F. Y., and Kasper, D. L. (2010) How Bacterial Carbohydrates Influence the Adaptive Immune System. *Annual Review of Immunology, Vol 28* **28**, 107-130
8. Guttormsen, H., Wetzler, L., Finberg, R., and Kasper, D. (1998) Immunologic memory induced by a glycoconjugate vaccine in a murine adoptive lymphocyte transfer model. *Infection and Immunity* **66**, 2026-2032

9. MR, W., LC, P., AK, R., F, M., J, D., HJ, J., and DL, K. (1993) Stimulation of protective antibodies against type Ia and Ib group B streptococci by a type Ia polysaccharide-tetanus toxoid conjugate vaccine. *Infect Immun* **61**, 4760-4766

10. Wantuch, P. L., and Avci, F. Y. (2018) Current status and future directions of invasive pneumococcal diseases and prophylactic approaches to control them. *Hum Vaccin Immunother* **14**, 2303-2309

11. Janeway, C. A., Travers, P., Walport, M., and Chlomchik, M. (2005) *Immunobiology, 6th edition., Garland Science Publishing:New York*,

12. Guttormsen, H.-K., Sharpe, A. H., Chandraker, A. K., Brigtsen, A. K., Sayegh, M. H., and Kasper, D. L. (1999) Cognate stimulatory B-Cell-T-Cell interactions are critical for T-cell help recruited by glycoconjugate vaccines. *Infect Immun* **67**, 6375-6384

13. Avci, F., Li, X., Tsuji, M., and Kasper, D. (2011) A mechanism for glycoconjugate vaccine activation of the adaptive immune system and its implications for vaccine design. *Nature Medicine* **17**, 1602-U1115

14. Avci, F., Li, X., Tsuji, M., and Kasper, D. (2012) Isolation of carbohydrate-specific CD4(+) T cell clones from mice after stimulation by two model glycoconjugate vaccines. *Nature Protocols* **7**, 2180-2192

15. Avci, F. Y., Li, X., Tsuji, M., and Kasper, D. L. (2013) Carbohydrates and T cells: a sweet twosome. *Semin Immunol* **25**, 146-151

16. Middleton, D. R., Sun, L., Paschall, A. V., and Avci, F. Y. (2017) T Cell-Mediated Humoral Immune Responses to Type 3 Capsular Polysaccharide of. *J Immunol* **199**, 598-603

17. Cheng, S., Wantuch, P. L., Kizer, M. E., Middleton, D. R., Wang, R., DiBello, M., Li, M., Wang, X., Li, X., Ramachandiran, V., Avci, F. Y., Zhang, F., Zhang, X., and Linhardt, R. J. (2019) Glycoconjugate synthesis using chemoselective ligation. *Org Biomol Chem* **17**, 2646-2650

18. Sun, X., Stefanetti, G., Berti, F., and Kasper, D. L. (2019) Polysaccharide structure dictates mechanism of adaptive immune response to glycoconjugate vaccines. *Proc Natl Acad Sci U S A* **116**, 193-198

19. Middleton, D. R., Zhang, X., Wantuch, P. L., Ozdilek, A., Liu, X., LoPilato, R., Gangasani, N., Bridger, R., Wells, L., Linhardt, R. J., and Avci, F. Y. (2018) Identification and characterization of the Streptococcus pneumoniae type 3 capsule-specific glycoside hydrolase of Paenibacillus species 32352. *Glycobiology* **28**, 90-99

20. Scott, C. A., Peterson, P. A., Teyton, L., and Wilson, I. A. (1998) Crystal structures of two I-Ad-peptide complexes reveal that high affinity can be achieved without large anchor residues. *Immunity* **8**, 319-329

21. Case, D. A., Cheatham, T. E., 3rd, Darden, T., Gohlke, H., Luo, R., Merz, K. M., Jr., Onufriev, A., Simmerling, C., Wang, B., and Woods, R. J. (2005) The Amber biomolecular simulation programs. *J Comput Chem* **26**, 1668-1688

22. Kirschner, K. N., Yongye, A. B., Tschampel, S. M., Gonzalez-Outeirino, J., Daniels, C. R., Foley, B. L., and Woods, R. J. (2008) GLYCAM06: a generalizable biomolecular force field. Carbohydrates. *J Comput Chem* **29**, 622-655

23. Miller, B. R., 3rd, McGee, T. D., Jr., Swails, J. M., Homeyer, N., Gohlke, H., and Roitberg, A. E. (2012) MMPBSA.py: An Efficient Program for End-State Free Energy Calculations. *J Chem Theory Comput* **8**, 3314-3321

24. Sette, A., Buus, S., Appella, E., Adorini, L., and Grey, H. M. (1990) Structural requirements for the interaction between class II MHC molecules and peptide antigens. *Immunol Res* **9**, 2-7

25. McFarland, B. J., Sant, A. J., Lybrand, T. P., and Beeson, C. (1999) Ovalbumin (323-339) peptide binds to the Major Histocompatibility Complex Class II I-A protein using two functionally distinct registers. *Biochemistry* **38**, 16663-16670

26. Tampé, R., and McConnell, H. M. (1991) Kinetics of antigenic peptide binding to the class II major histocompatibility molecule I-Ad. *Proc Natl Acad Sci U S A* **88**, 4661-4665

27. Grotenbreg, G. M., Nicholson, M. J., Fowler, K. D., Wilbuer, K., Octavio, L., Yang, M., Chakraborty, A. K., Ploegh, H. L., and Wucherpfennig, K. W. (2007) Empty class II major histocompatibility complex created by peptide photolysis establishes the role of DM in peptide association. *J Biol Chem* **282**, 21425-21436

28. Genheden, S., and Ryde, U. (2015) The MM/PBSA and MM/GBSA methods to estimate ligand-binding affinities. *Expert Opin Drug Discov* **10**, 449-461

29. Etlinger, H. M., Gillessen, D., Lahm, H. W., Matile, H., Schönfeld, H. J., and Trzeciak, A. (1990) Use of prior vaccinations for the development of new vaccines. *Science* **249**, 423-425

30. Cai, H., Chen, M. S., Sun, Z. Y., Zhao, Y. F., Kunz, H., and Li, Y. M. (2013) Self-adjuvanting synthetic antitumor vaccines from MUC1 glycopeptides

conjugated to T-cell epitopes from tetanus toxoid. *Angew Chem Int Ed Engl* **52**, 6106-6110

31. de Velasco, E. A., Merkus, D., Anderton, S., Verheul, A. F., Lizzio, E. F., Van der Zee, R., Van Eden, W., Hoffman, T., Verhoef, J., and Snippe, H. (1995) Synthetic peptides representing T-cell epitopes act as carriers in pneumococcal polysaccharide conjugate vaccines. *Infect Immun* **63**, 961-968

32. Jackson, D. C., Purcell, A. W., Fitzmaurice, C. J., Zeng, W., and Hart, D. N. (2002) The central role played by peptides in the immune response and the design of peptide-based vaccines against infectious diseases and cancer. *Curr Drug Targets* **3**, 175-196

33. Rodrigues-da-Silva, R. N., Correa-Moreira, D., Soares, I. F., de-Luca, P. M., Totino, P. R. R., Morgado, F. N., Oliveira Henriques, M. D. G., Peixoto Candea, A. L., Singh, B., Galinski, M. R., Moreno, A., Oliveira-Ferreira, J., and Lima-Junior, J. D. C. (2019) Immunogenicity of synthetic peptide constructs based on PvMSP9E795-A808, a linear B-cell epitope of the P. vivax Merozoite Surface Protein-9. *Vaccine* **37**, 306-313

34. Snook, A. E., Baybutt, T. R., Hyslop, T., and Waldman, S. A. (2016) Preclinical Evaluation of a Replication-Deficient Recombinant Adenovirus Serotype 5 Vaccine Expressing Guanylate Cyclase C and the PADRE T-helper Epitope. *Hum Gene Ther Methods* **27**, 238-250

35. Zeigler, D. F., Roque, R., and Clegg, C. H. (2019) Optimization of a multivalent peptide vaccine for nicotine addiction. *Vaccine* **37**, 1584-1590

36. Baraldo, K., Mori, E., Bartoloni, A., Norelli, F., Grandi, G., Rappuoli, R., Finco, O., and Del Giudice, G. (2005) Combined conjugate vaccines: enhanced immunogenicity with the N19 polyepitope as a carrier protein. *Infect Immun* **73**, 5835-5841

37. Bixler, G. S., Eby, R., Dermody, K. M., Woods, R. M., Seid, R. C., and Pillai, S. (1989) Synthetic peptide representing a T-cell epitope of CRM197 substitutes as carrier molecule in a Haemophilus influenzae type B (Hib) conjugate vaccine. *Adv Exp Med Biol* **251**, 175-180

38. Falugi, F., Petracca, R., Mariani, M., Luzzi, E., Mancianti, S., Carinci, V., Melli, M. L., Finco, O., Wack, A., Di Tommaso, A., De Magistris, M. T., Costantino, P., Del Giudice, G., Abrignani, S., Rappuoli, R., and Grandi, G. (2001) Rationally designed strings of promiscuous CD4+ T cell epitopes provide help to Haemophilus influenzae type b oligosaccharide: a model for new conjugate vaccines. *Eur. J. Immunol* **31**, 3816-3824

39. Avci, F., Kasper, D., Paul, W., Littman, D., and Yokoyama, W. (2010) How Bacterial Carbohydrates Influence the Adaptive Immune System. *Annual Review of Immunology, Vol 28* **28**, 107-130

40. Cobb, B. A., Wang, Q., Tzianabos, A. O., and Kasper, D. L. (2004) Polysaccharide processing and presentation by the MHCII pathway. *Cell* **117**, 677-687

41. Brigl, M., and Brenner, M. B. (2004) CD1: antigen presentation and T cell function. *Annu Rev Immunol* **22**, 817-890

42. Sieling, P. A., Chatterjee, D., Porcelli, S. A., Prigozy, T. I., Mazzaccaro, R. J., Soriano, T., Bloom, B. R., Brenner, M. B., Kronenberg, M., Brennan, P. J., and et al. (1995) CD1-restricted T cell recognition of microbial lipoglycan antigens. *Science* **269**, 227-230

43. Deck, B., Elofsson, M., Kihlberg, J., and Unanue, E. R. (1995) Specificity of Glycopeptide-Specific T Cells. *J. Immunol.* **155**, 1074-1078

44. Galli-Stampino, L., Meinjohanns, E., Frische, K., Meldal, M., Jensen, T., Werdelin, O., and Mouritsen, S. (1997) T-cell recognition of tumor-associated carbohydrates: the nature of the glycan moiety plays a decisive role in determining glycopeptide immunogenicity. *Cancer Res* **57**, 3214-3222

45. Haurum, J. S., Arsequell, G., Lellouch, A. C., Wong, S. Y., Dwek, R. A., McMichael, A. J., and Elliott, T. (1994) Recognition of carbohydrate by major histocompatibility complex class I-restricted, glycopeptide-specific cytotoxic T lymphocytes. *J Exp Med* **180**, 739-744

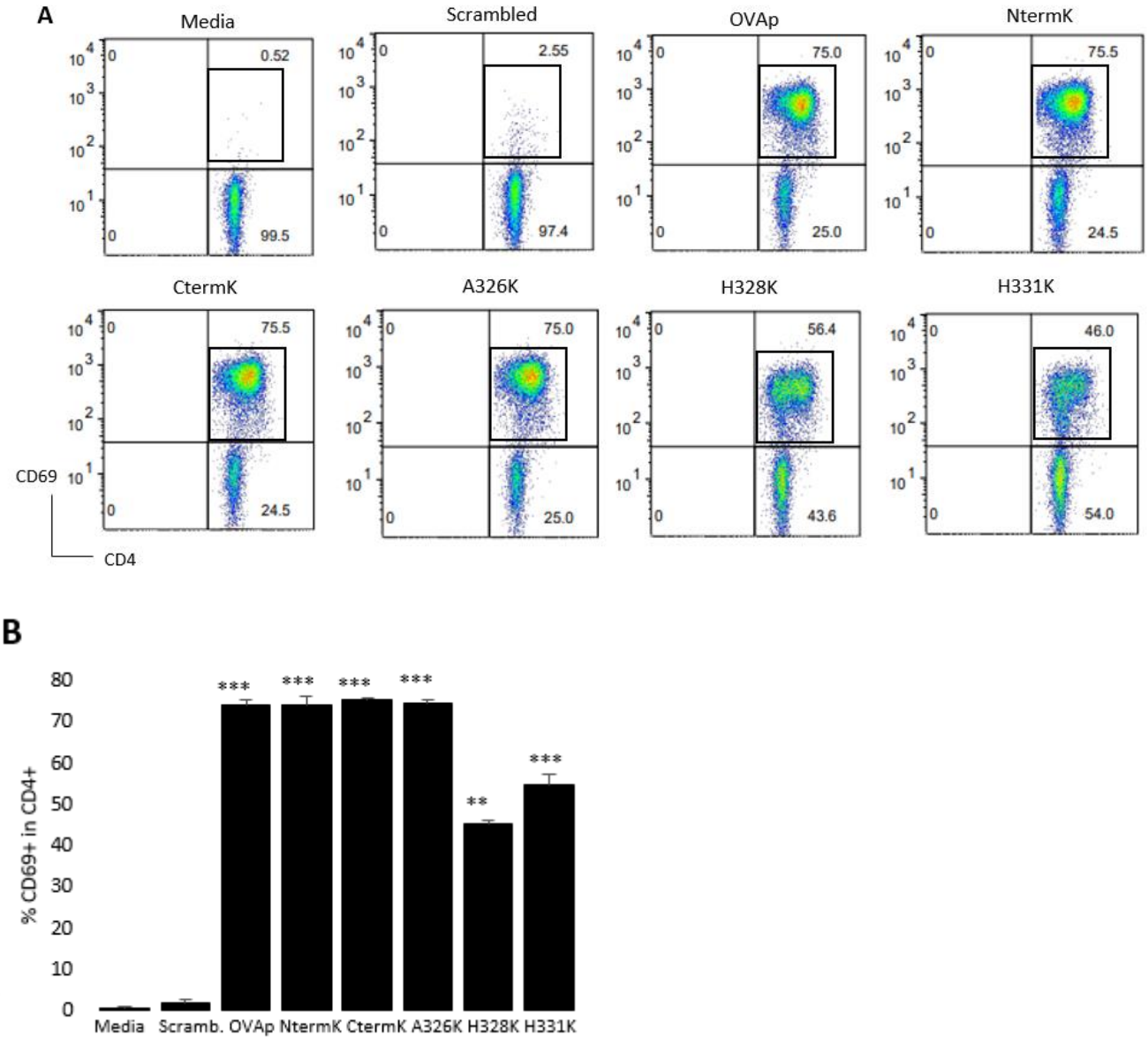
46. Mouritsen, S., Meldal, M., Christiansen-Brams, I., Elsner, H., and Werdelin, O. (1994) Attachment of oligosaccharides to peptide antigen profoundly affects binding to major histocompatibility complex class II molecules and peptide immunogenicity. *Eur J Immunol* **24**, 1066-1072

Table 4.1

Table 4.1: List of peptide derivatives use in the study. Modifications using either Cys (C) or Lys (K) are shown in locations and depicted in bold.

Peptide Derivatives	
OVAp	ISQAVHAAHAEINEAGR
Nterm C/K	**C/K**ISQAVHAAHAEINEAGR
Cterm C/K	ISQAVHAAHAEINEAGR**C/K**
A326C/K	ISQ**C/K**VHAAHAEINEAGR
H328C/K	ISQAV**C/K**AAHAEINEAGR
H331C/K	ISQAVHAA**C/K**AEINEAGR
Scrambled OVAp	AIGHASNEIQAVAEHAR

Figure 4.1


A
Media
Scrambled
OVAp
NtermK
CtermK
A326K
H328K
H331K
0.52
99.5
2.55
97.4
75.0
25.0
75.5
24.5
75.5
24.5
75.0
25.0
56.4
43.6
46.0
54.0
CD69
CD4
B
% CD69+ in CD4+
80
70
60
50
40
30
20
10
0
***
***
***
***
**
***
Media
Scramb.
OVAp
NtermK
CtermK
A326K
H328K
H331K

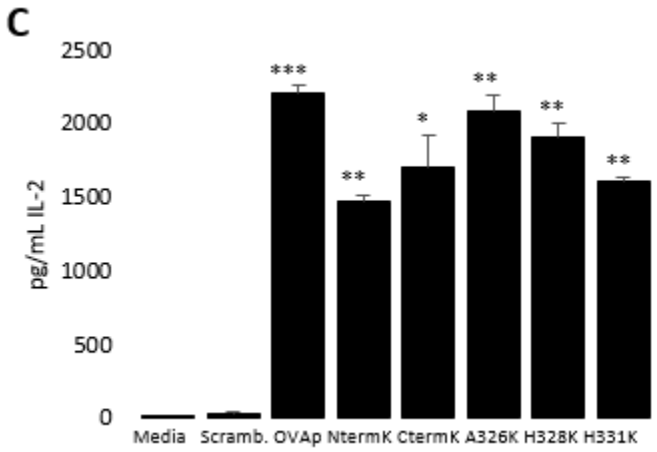
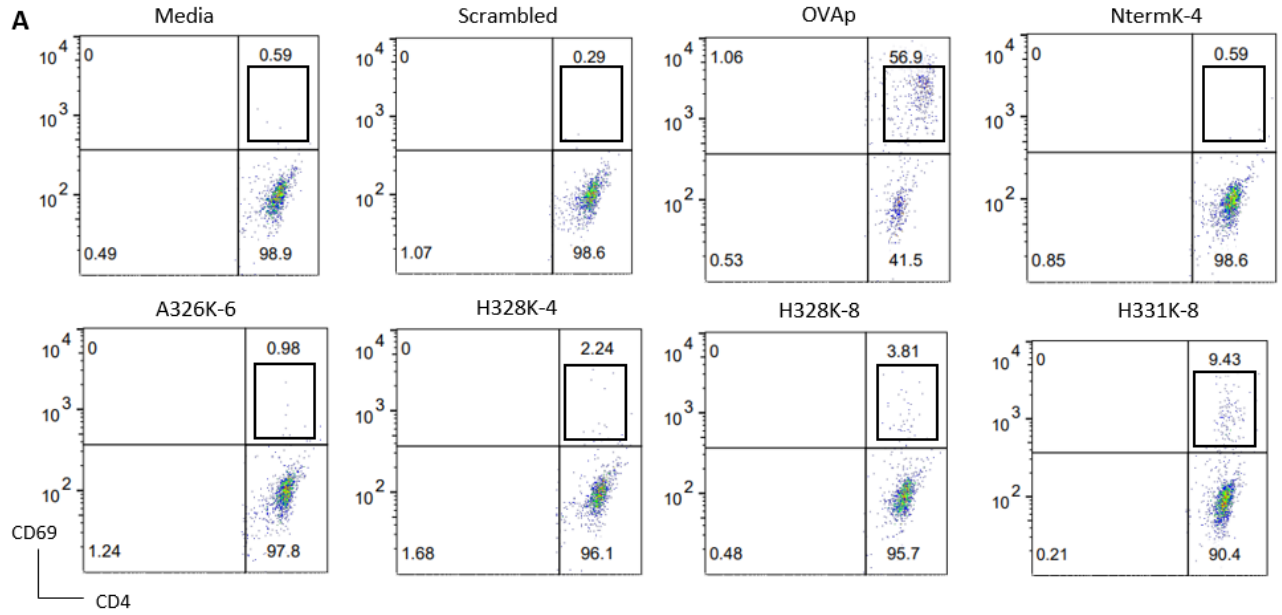
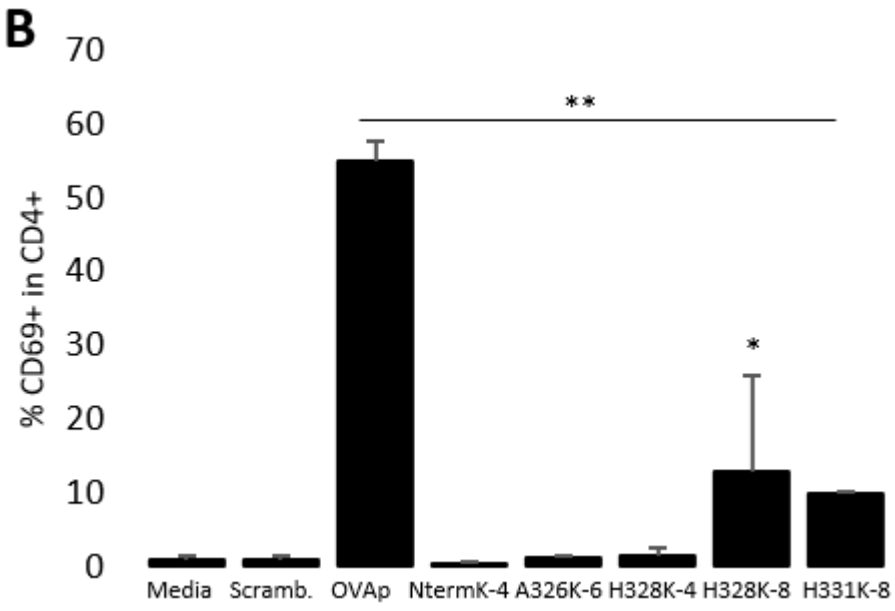


Figure 4.1: T cell activation of peptides A) T cell proliferation measured using flow cytometery gating percent CD69+ in CD4+ T cells. **B)** Quantification of flow cytometry gating using duplicates. **C)** T cell activation of transgenic mouse line DO11.10 using peptide derivatives as measured by IL-2 ELISA.

Figure 4.2


A
Media
Scrambled
OVAp
NtermK-4
A326K-6
H328K-4
H328K-8
H331K-8
0.59
0.49
98.9
0.29
1.07
98.6
1.06
56.9
0.53
41.5
0.59
0.85
98.6
0.98
1.24
97.8
2.24
1.68
96.1
3.81
0.48
95.7
9.43
0.21
90.4
CD69
CD4


B
% CD69+ in CD4+
70
60
50
40
30
20
10
0
**
*
Media
Scramb.
OVAp
NtermK-4
A326K-6
H328K-4
H328K-8
H331K-8

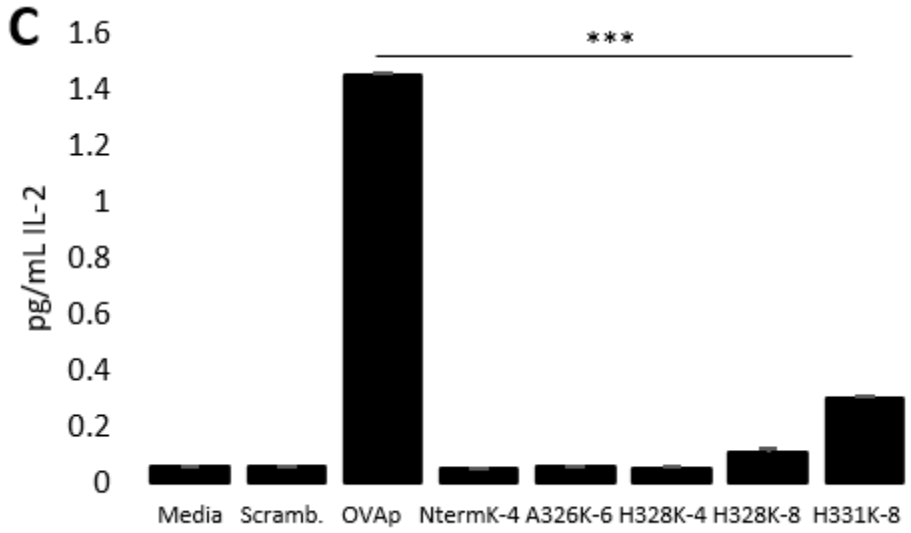


Figure 4.2: T cell activation of conjugates A) T cell proliferation measured using flow cytometery gating percent CD69+ cells in CD4+ T cells. **B)** Quantification of flow cytometry gating using duplicates **C)** T cell activation of transgenic mouse line DO11.10 using glycoconjugates as measured by IL-2 ELISA.

Figure 4.3

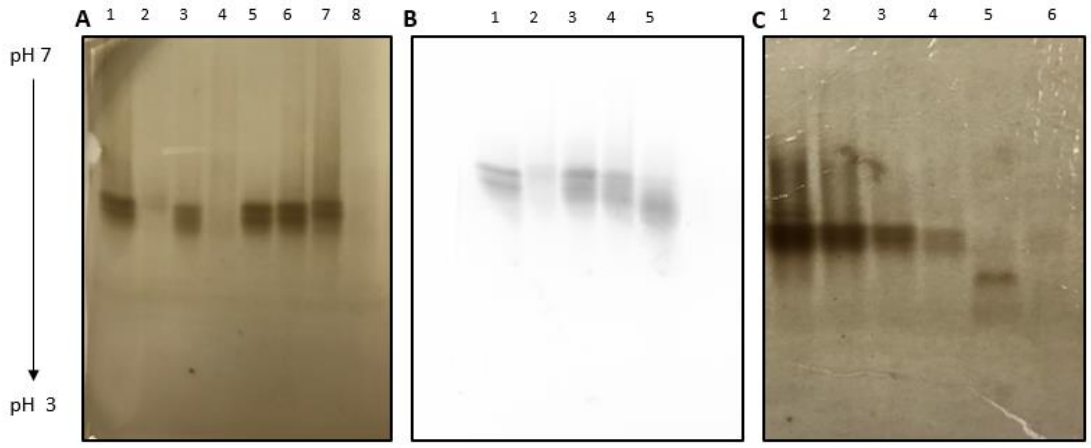
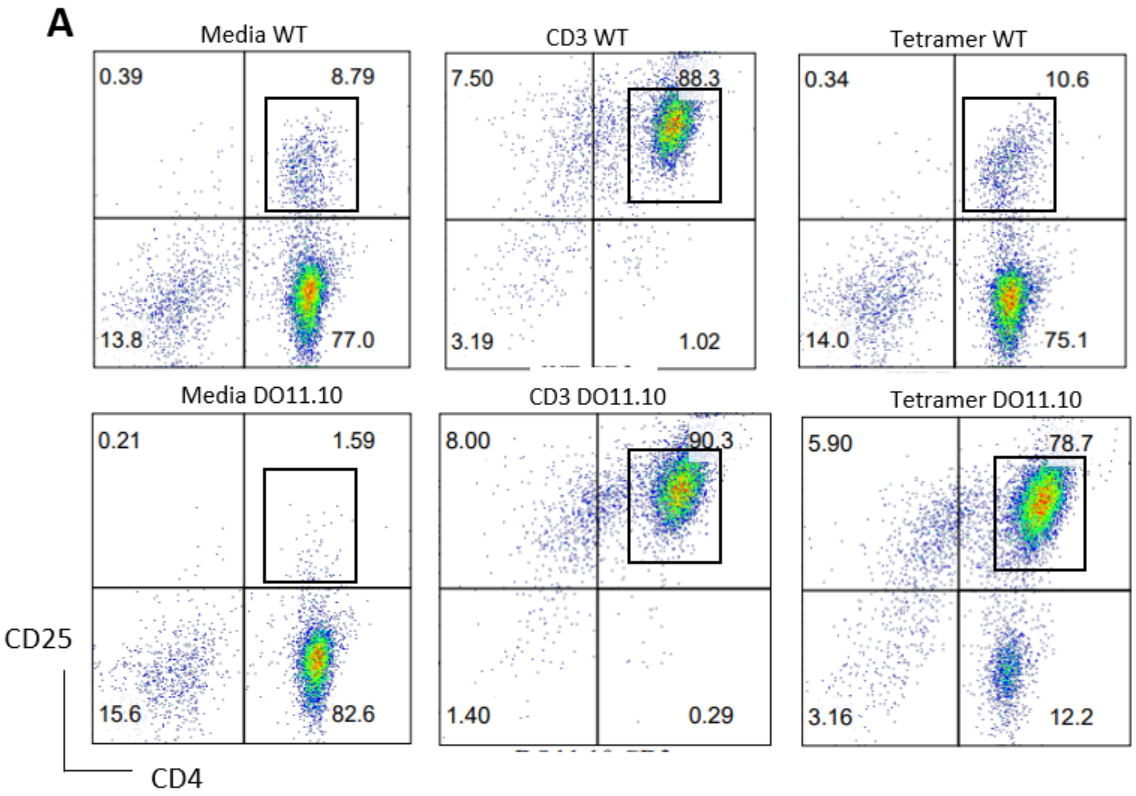
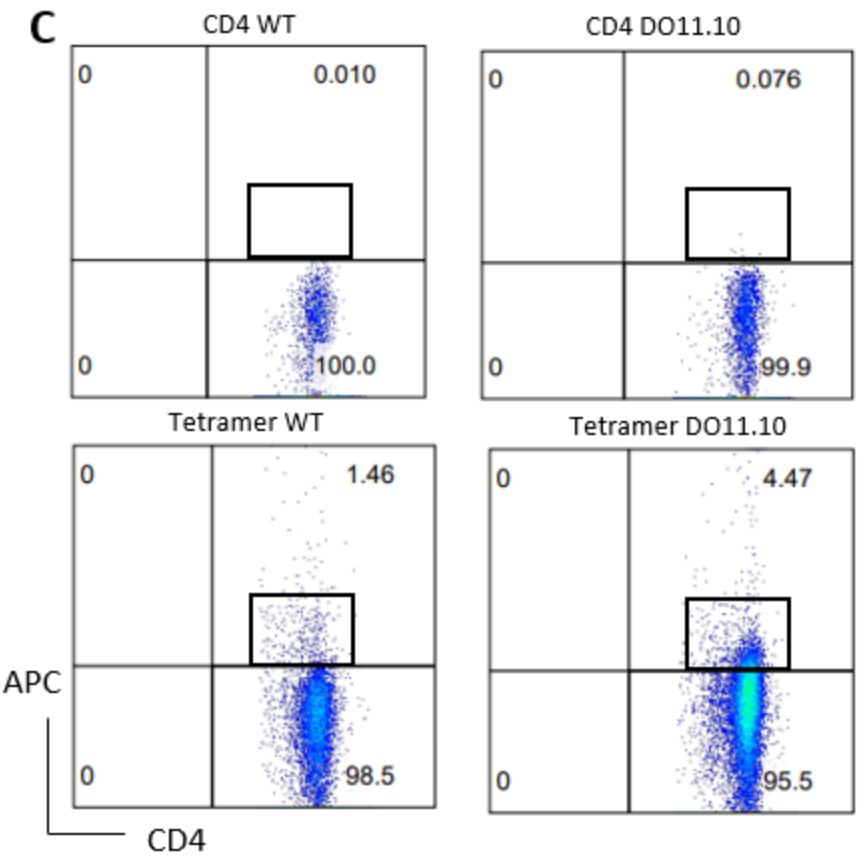


Figure 4.3: MHCII binding of peptides and conjugates. Binding of **A)** peptides lanes: 1) OVAp 2) scrambled 3) H331K 4) H328K 5) A326K 6) Nterm-K 7) Cterm-K 8) hexasaccharide; **B)** glycoconjugates lanes: 1) OVAp 2) scrambled 3) Nterm-C 4) Nterm-C-Di 5) Nterm-C-Tetra; **C)** glycoconjugates lanes: 1) H331K-Octa 2) A326K-Octa 4) H331K-Hexa 5) A326K-Hexa 6) H331K-Tetra to MHCII detected using IEF gel and silver stain.

Figure 4.4


A
Media WT
0.39
8.79
13.8
77.0
CD3 WT
7.50
88.3
3.19
1.02
Tetramer WT
0.34
10.6
14.0
75.1
Media DO11.10
0.21
1.59
15.6
82.6
CD3 DO11.10
8.00
90.3
1.40
0.29
Tetramer DO11.10
5.90
78.7
3.16
12.2
CD25
CD4
B
% CD25+ in CD4+
100
90
80
70
60
50
40
30
20
10
0
ns
***
***
Media WT
CD3 WT
Tetramer WT
Media DO11.10
CD3 DO11.10
Tetramer DO11.10


C
CD4 WT
0
0.010
0
100.0
CD4 DO11.10
0
0.076
0
99.9
Tetramer WT
0
1.46
0
98.5
Tetramer DO11.10
0
4.47
0
95.5
APC
CD4

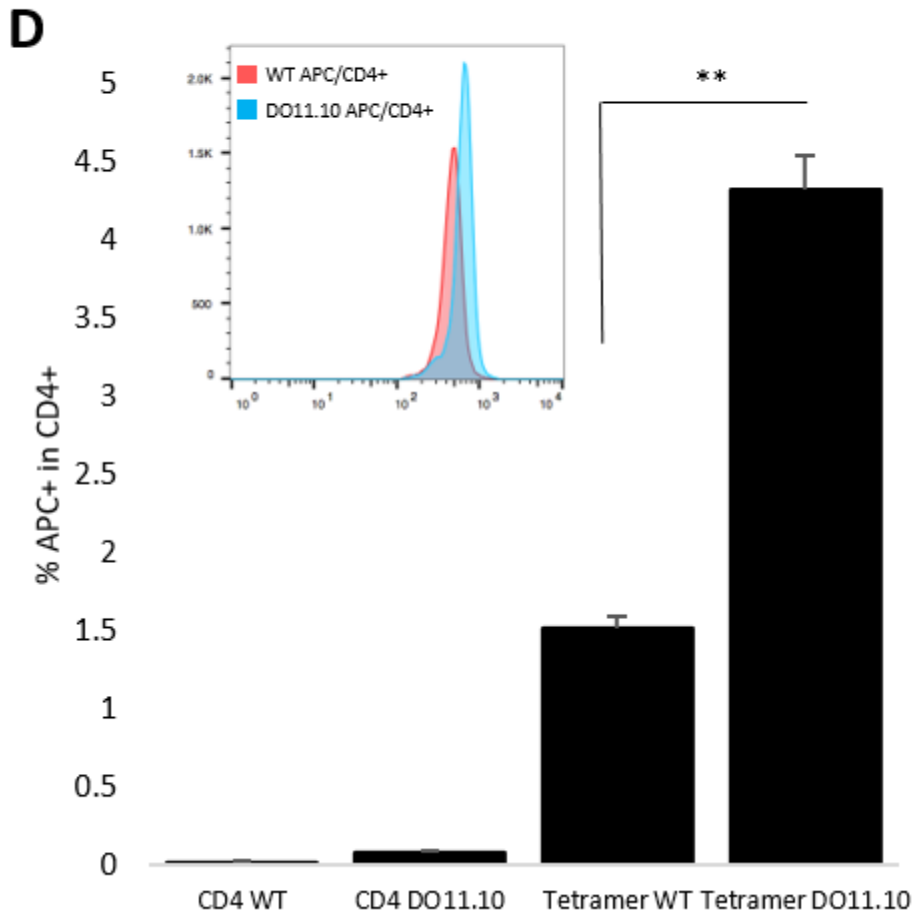
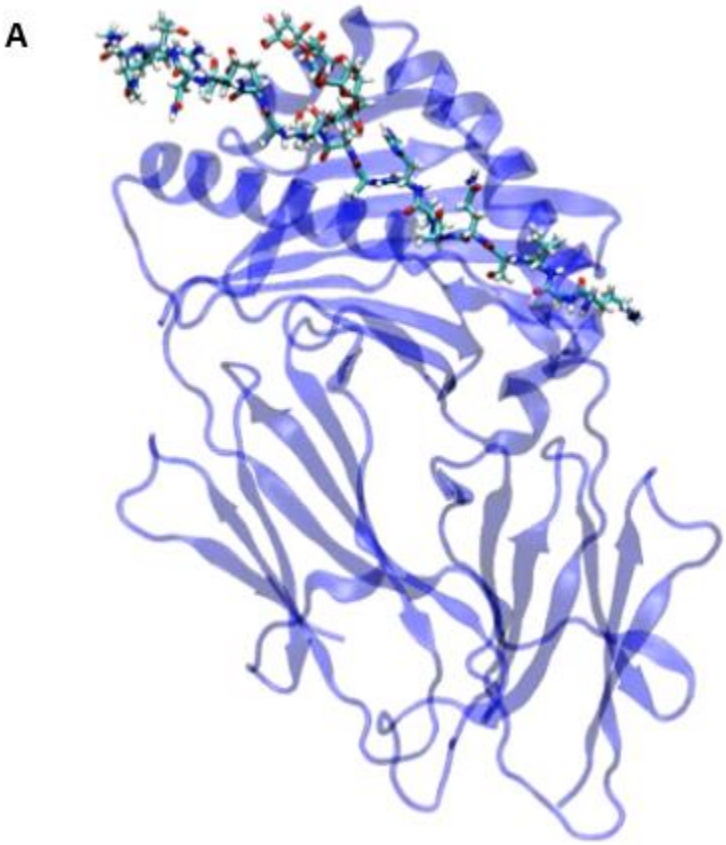


Figure 4.4: OVA peptide MHCII tetramer staining of T cells. A) T cell staining from T cell activation assay using WT or DO11.10 mice. Flow was gated as CD25+ in CD4+ cells. CD3 was used as a positive control. **B)** Quantification of flow cytometry gating using duplicates **C)** APC labeled OVAp MHCII tetramers binding T cells using WT or DO11.10 mice full splenocytes. Gating was percent APC+ (tetramers) in CD4+ positive populations. **D)** Quantification of flow cytometry gating using duplicates with MFI shift depicted.

Figure 4.5



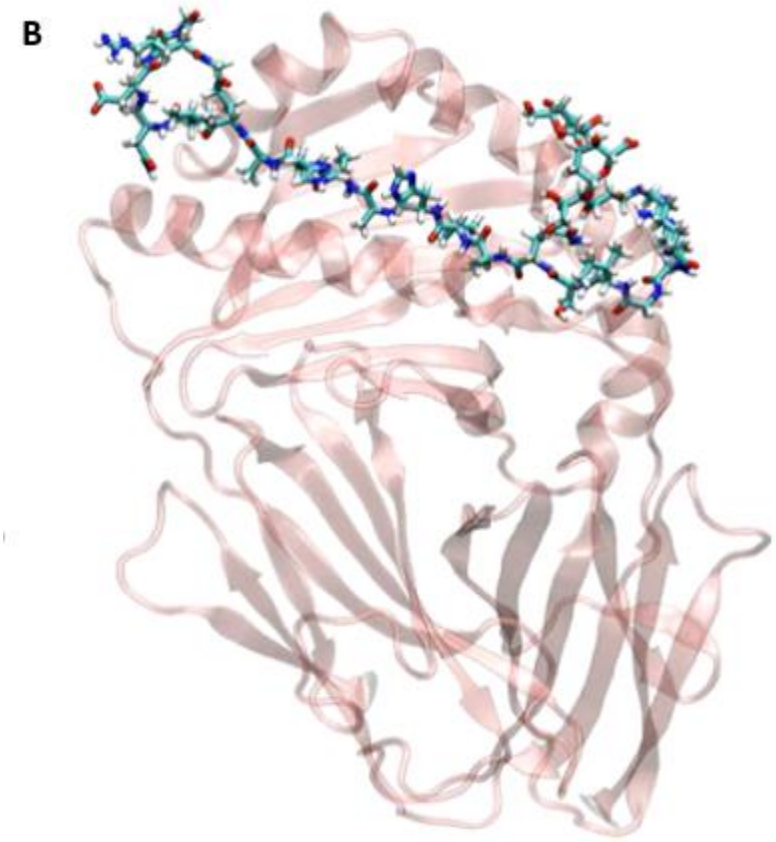


Figure 4.5: Structure of MHCII bound with antigen. MHCII molecule (PDBID: 1IAO) [reference 20] was modeled to bind **A)** H331K-Tetrasaccharide glycoconjugate **B)** Nterm-K-Tetrasaccharide glycoconjugate

Figure 4.6

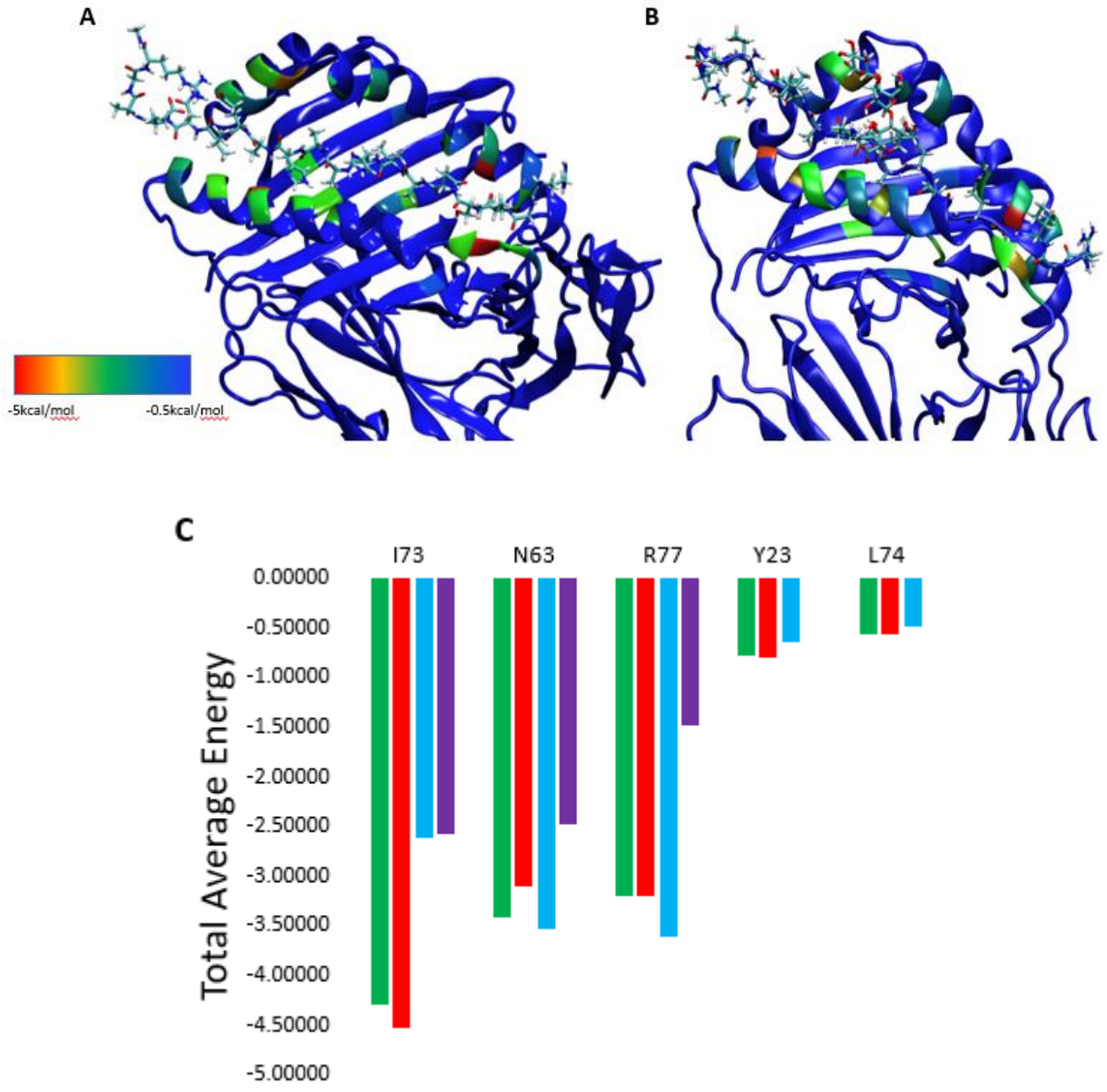


Figure 4.6: Binding requirements of peptide and conjugates. MHCII molecules (PDBID: 1IAO) [reference 20] were modeled with **A)** OVAp or **B)** H331K-Tetra glycoconjugate. MM-GBSA analysis was performed. MHCII residues contributing to binding of the antigen are colored by degree to which they contribute (red- highest binding, blue-lowest). **C)** Five residues were found to contribute more to binding when

glycans were present than peptide binding alone. Green bar: H331K-Tetra, Red bar: Nterm-K-Tetra, Blue bar: Cterm-K-Tetra, Purple bar: OVAp

CHAPTER 5

ISOLATION AND CHARACTERIZATION OF NEW HUMAN CARRIER PEPTIDES FROM TWO IMPORTANT VACCINE IMMUNOGENS

Paeton L Wantuch, Lina Sun, Rachel K LoPilato, Jarrod J Mousa, Robert S Haltiwanger and Fikri Y Avci. 2020. 38(10), 2315-2325. Reprinted here with permission of the publisher.

DOI: 10.1016/j.vaccine.2020.01.065

Abstract

In the preparation of glycoconjugate vaccines in clinical practice, two highly immunogenic carrier proteins, CRM_{197} and tetanus toxoid (TT), are predominantly used to glycoconjugate with the capsular polysaccharides (CPSs) of bacterial pathogens. In addition, TT has long been used as an effective vaccine to prevent tetanus. While these carrier proteins play an important role in immunogenicity and vaccine design alike, their defined human major histocompatibility complex class II (MHCII) T cell epitopes are inadequately characterized. In this current work, we use mass spectrometry to identify the peptides from these carrier proteins that are naturally processed and presented by human B cells via MHCII pathway. The MHCII-presented peptides are screened for their T cell stimulation using primary CD4+ T cells from four healthy adult donors. These combined methods reveal a subset of eleven CD4+ T cell epitopes that proliferate and stimulate human T cells with diverse MHCII allelic repertoire. Six of these peptides stand out as potential immunodominant epitopes by responding in three or more donors. Additionally, we provide evidence of these natural epitopes eliciting more significant T cell responses in donors than previously published TT peptides selected from T cell epitope screening. This study serves toward understanding carrier protein immune responses and thus enables the use of these peptides in developing novel knowledge-based vaccines to combat persisting problems in glycoconjugate vaccine design.

Introduction

The introduction of glycoconjugate vaccines to clinical practice in the late 80s has led to great strides in combating infections against bacterial pathogens (1-4). Glycoconjugate vaccines against a number of highly infective serotypes of *S.*

pneumoniae, *H. influenzae* type b, and *N. meningitidis* are currently available (3). These vaccines are composed of the bacterial capsular polysaccharide (CPS) covalently conjugated to a carrier protein. The most common carrier proteins include tetanus toxoid (TT) and a non-toxic mutant of diphtheria toxin, CRM_{197} (1,3). Surprisingly, while these carrier proteins play a prominent role in glycoconjugate vaccine design, the precise nature of their major histocompatibility complex class II (MHCII) epitopes has not been extensively studied.

Upon administration, vaccine components are endocytosed by antigen presenting cells (APCs) processed into smaller peptide fragments that bind with MHCII (5,6). This allows the processed peptide epitopes to be shuttled to the APCs surface to interact with T cell receptors (TCR) of CD4+ helper T cells. Since glycoconjugate vaccines primarily induce humoral immune responses, antigen presentation by B cells is essential for their subsequent activation and differentiation into plasma cells (6,7). There are two working models for how glycoconjugate vaccines induce CD4+ T cell stimulation. According to the first model, when the polysaccharide portion of the glycoconjugate vaccine is not fully degraded in the endolysosomes, a peptide-bound, processed carbohydrate T cell epitope is presented on the surface of APCs (8-10). In the second model, the polysaccharide portion of the glycoconjugate is fully degraded in the endolysosomes and the free peptide is presented on the APCs surface (10). Both models require processing of the carrier protein into MHCII-binding peptides with or without a covalently bound, processed glycan for T cell presentation (11-13). This in turn stimulates the T cells to help B cells produce high-affinity IgG antibodies against the CPS (8). Thus, it is evident that the immunogenicity of the given glycoconjugate vaccine is dependent upon the

endosomal processing of glycoconjugate vaccine to yield peptide-containing epitopes that bind to a wide variety of MHCII alleles for T cell stimulation (14). The trimolecular complex, consisting of TCR, MHCII and the T cell epitope are crucial for B cell activation and differentiation into plasma cells and therefore our understanding of the adaptive, humoral immune responses elicited by glycoconjugate vaccines.

The intimate knowledge of the nature of the trimolecular complex can then be exploited for the production of knowledge-based, new-generation glycoconjugate vaccines. Due to their empirical design and synthesis, current glycoconjugate vaccines have variable immunogenicity, especially in high-risk populations, such as elderly and immunocompromised individuals (2,15). Carrier-specific suppression wherein antibody response to the polysaccharide portion of the glycoconjugate vaccine can be inhibited due to pre-existing immune response to the carrier protein from prior immunizations is also a rising concern (16-19). Carrier-specific suppression may be contributing to lowered immunogenicity and efficacy of glycoconjugate vaccines as continued exposure to the same carrier proteins continues to rise with generation of new vaccines. Carrier-specific suppression stems, in part, from the presence of carrier-specific B cells and suppressor T cells (18). Therefore, one approach to eliminate this is the identification of immunogenic peptide epitopes with T cell helper function as has been the focus of multiple studies (16-18).

Importantly, the advantage of using immunogenic peptide epitopes as carriers for glycoconjugate vaccines over the traditional carrier proteins has been demonstrated in previous studies (8,19-28). For example, one study showed that the use of a peptide as a carrier led to higher IgG titers against the polysaccharide and greater protection in a

survival assay compared to a protein carrier (8). Another study performed the preclinical evaluation of a colorectal cancer vaccine target utilizing the established PAn DR Epitope (i.e., PADRE peptide) as the carrier (27). Most recently, a study showed that peptide as carrier leads to reduced anti-carrier antibody titers compared to protein carrier, indicating reduced carrier-specific suppression (28). These studies mark an important point in understanding immune activation and mechanism behind glycoconjugate vaccines. Only a very small subset of peptides generated from any protein carrier in the endolysosome will bind to MHC. Implementing the use of MHCII-binding peptides as carriers will lead to effective T cell mediated responses. To date, there have been a number of peptides derived from carrier protein TT for this very task. Some of the more common, often referred to as universal TT epitopes, are P2, P30, and P32 (14,18,29-34). Additionally, a subset of peptides has been identified for CRM_{197}, although research into T cell epitopes of this carrier protein is less extensive (19-21,35,36). While the insights gleaned from these works are significant and important, it remains that these peptides were identified through indirect approaches; namely, prediction softwares or T cell screening of overlapping synthetic peptides spanning the proteins.

In this study, we aimed to identify naturally processed and presented immunogenic peptide epitopes derived from carrier proteins TT and $CRM_{197,}$ through immunoprecipitation and mass spectrometry. The identified peptides were then probed for T cell proliferation in four healthy adult donors primed with TT and CRM_{197}. We determine peptide epitopes for both carrier proteins, in addition to variants of previously published peptides, which are capable of binding MHCII and stimulating human CD4+ T cells. Additionally, we present evidence that these defined peptide epitopes discovered

through natural MHCII presentation perform better in activating CD4+ T cells than previously reported TT epitopes. This information will be valuable for generating peptide-based carriers to be used in future vaccine design.

Material and Methods

Study Subjects

The studies described herein utilized human samples approved by the University of Georgia Institutional Review Board as STUDY00005127. Four adult male donors from Athens, GA USA who were vaccinated with PCV13 (Prevnar-13) less than one year prior to their blood collection, and with verbally confirmed Tdap vaccination within the past 10 years were recruited to the University of Georgia Clinical and Translational Research Unit. Participants provided written informed consent for participation in this study. Peripheral blood mononuclear cells (PBMCs) were purified by Ficoll-gradient density centrifugation and were used fresh in all assays.

Affinity Purification of MHCII Molecules

Approximately $5x10^7$ human B lymphoblasts (Raji ATCC CCL-86) were incubated with 1 mg of carrier protein (CRM_{197} [Fina Biosolutions, LLC], Tetanus toxoid [TT], or TT heavy chain [TT_{hc}][Fina Biosolutions, LLC]) in RPMI 1640 (Corning) medium containing 2 g/L sodium bicarbonate, 2 mM L-glutamine, 1 mM sodium pyruvate, 1% penicillin/streptomycin, and 10% heat inactivated FBS. Cells were incubated for 18 hrs at 37 degrees in 5% CO_2. After incubation MHCII molecules were obtained via immunoprecipitation after lysis of the cells in NP-40 buffer for 1 hour at room temperature. The lysate was cleared by centrifugation at 15,000xg for 15 minutes. The MHCII molecules were immunoprecipitated from the cleared lysate using 15 ng of each

anti-human HLA-DR antibody (L243 Biolegend), HLA-DP antibody (BRAFB6 Santa Cruz Biotechnology), or HLA-DQ antibody (B-K27 Santa Cruz Biotechnology) bound to Protein A agarose beads (Sigma-Aldrich). The affinity column was washed with PBS four times. The MHCII molecules were then eluted from the affinity column with 10% acetic acid at room temperature for 3 minutes with 4 elution fractions collected. Eluted MHCII proteins were evaluated for purity via mass spectrometry. Immunoprecipitation for mass spectrometry analysis was performed three independent times to select the peptide epitopes displayed here.

Separation of Peptides from MHCII Molecules

Eluted MHCII molecules were heated at 70 degrees for 10 minutes to denature MHCII and release bound peptides. Peptides were separated out from denatured HLA protein subunits by ultrafiltration using a 10 kDa cutoff membrane filter (Millipore) at 4000xg. The filter was washed two times using deionized water before loading samples. Recovered peptides in the filtrate were dried down and resuspended in 8M urea in 50 mM ammonium bicarbonate before sonication. Samples were desalted prior to mass spectrometry analysis with ZipTip C-18 columns per product protocol (Millipore). The eluted peptides were diluted to 10% acetonitrile with 0.1% formic acid and spun through a 0.2 μM nylon centrifugal filter (VWR) at 1000xg to remove any precipitants. The retentate fraction containing denatured HLA proteins was dried down and resuspended in 8M urea in 50 mM ammonium bicarbonate containing 10 mM TCEP. Samples were treated with 10 mM iodoacetamide and sonicated. Trypsin (250ng) was added and the samples were incubated overnight at 37 ^{o}C. The resulting peptides were desalted on ZipTips as described above.

Analysis by Mass Spectrometry

The samples were injected onto a PepMap RSLC C18 column (Thermo Scientific) with an Easy nano HPLC coupled to a Q-Exactive Plus mass spectrometry system (Thermo Scientific) at a flow rate of 300 nl/min with a 25 min 0-40% acetonitrile gradient in 0.1% formic acid followed by a 3 min gradient to 80% acetonitrile. Spectra were recorded with a resolution of 35,000 in the positive polarity mode over the range of m/z 350–2,000 and an automatic gain control target value was $1x10^6$. The 10 most prominent precursor ions in each full scan were isolated for higher energy collisional dissociation-tandem mass spectrometry (HCD-MS/MS) fragmentation with normalized collision energy of 35%, an automatic gain control target of $2x10^5$, an isolation window of m/z 3.0, dynamic exclusion enabled, and fragment resolution of 17,500.

Database Search

Targeted searches against tetanus toxin protein (Uniprot P04958) and diphtheria toxin protein (Uniprot Q6NK15) were performed by Byonic v (Protein Metrics) software. Non-specific cleavage was selected with a parent ion mass error of 10 ppm and MS2 ion mass error of 20 ppm. Peptides identified after 2% False Discovery Rate were manually evaluated. Byonic scores for positive identifications were greater than 50 but under 100, so spectra with discernable isotopic distributions and few MS2 contaminants were considered. Proteome Discoverer 2.1 was used for Sequest searches against the human proteome (UniProt ID UP000005640) to identify MHCII molecules in retentate fractions.

Synthesis of Discovered Peptides

We synthesized 11 peptides that were observed through mass spectrometry analysis. Peptides were derived from either CRM_{197} or tetanus toxoid (TT) proteins. Additionally,

we synthesized two known CD4+ cell epitopes of TT, P2 and P32. Sequences and protein positions of peptides can be found in Table 1. All couplings for peptides were carried out on an automated microwave-assisted solid-phase peptide synthesizer (CEM Corp. Liberty microwave synthesizer) using the standard protocols in the instrument software. Peptides were synthesized on Rink amide resin (0.6 meq/g; Novabiochem) via N^R-N-(9-fluorenyl)methoxycarbonyl (Fmoc) approach in the primary solvent N,N-dimethylformamide (DMF). 20% 4-methylpiperidine in DMF was used for Fmoc removal. 2-(1*H*-Benzotriazole-1-yl)-oxy-1,1,3,3-tetramethyluronium hexafluorophosphate/1-hydroxybenzotriazole in the presence of *N*,*N*-diisopropylethylamine (DIPEA) were used as the coupling reagents. Peptides were cleaved from the resin through TFA/triisopropylsilane/H_2O (95:2.5:2.5) cocktail for ~2 hours. The cleavage cocktail was added dropwise through a filter to cold ether to precipitate the crude peptide and centrifuged to remove the ether supernatant. Purity was verified by analytical HPLC and MALDI-TOF MS (data not shown).

ELISA of Donor Serum

Anti-carrier protein IgG titers were determined using enzyme-linked immunosorbent assay (ELISA). Briefly, 96-well plates (Costar) were coated in duplicate overnight with 2 μg/mL protein (CRM_{197}, TT_m, TT_{hc}, or BSA as negative control). Wells were blocked with 1% BSA in PBS and washed with 0.05% PBS-Tween (PBST) all subsequent washes were the same. Serial dilutions of donor serum starting at 1:200 was added to wells for 2 hours at room temperature and washed. Total IgG titers were detected by HRP conjugated anti-human IgG (Santa Cruz Biotechnology) (1:2000 dilution) added to wells for 2 hours at room temperature. After washing, plates were developed using 3,3',5,5'

tetramethyl benzidine (TMB) substrate (Biolegend) and stopped with 2 N H_2SO_4. The optical densities were determined at 450 nm using a microplate reader (Synergy H1, Bio-Tek). Serum titers were determined at OD 0.5 and significance determined using 2-tailed Student's *t* test with $p<0.05$.

T cell Proliferation

PBMCs were collected freshly from healthy donors and separated using Ficoll extraction. The culture medium for the PBMCs was RPMI 1640 (Corning) supplemented with 2 g/L sodium bicarbonate, 50 μM 2-mercaptoethanol, 2 mM L-glutamine, 1 mM sodium pyruvate, 1% nonessential amino acids, 1% penicillin/streptomycin, and 10% heat inactivated FBS. CD4+ T cells were separated out from PBMCs using a negative selection CD4 enrichment kit (BD Biosciences) and stained with 2 μM carboxyfluorescein diacetate succinimidyl ester (CFSE) (37). CD4- depleted PBMCs were treated with mitomycin-C at 25 μg/mL. Proliferation assays were performed with CFSE stained CD4+ T cells using 10^5 cells/well and mitomycin-C treated PBMCs ($2x10^5$ cells/well) as APCs. Cells were plated in quadruplicate per antigen in 200 μL supplemented RPMI in 96-well flat bottom plate. Cells were stimulated with 2.5 ng/mL IL-2 and each of the following antigens at 50 μg/ml: CRM_{197} protein (Fina Biosolutions, LLC), TT_{hc} protein (Fina Biosolutions, LLC), TT_m, P2, P32, TT_{94-107}, $TT_{660-667}$, $TT_{826-837}$, $TT_{1093-1102}$, $TT_{1169-1179}$, $TT_{1222-1236}$, $TT_{1228-1239}$, CRM_{26-39}, CRM_{87-97}, $CRM_{299-312}$, and $CRM_{425-440}$. After 72 hours cells were supplemented with 2.5 ng/mL IL-2 and 50 μg/ml antigen in 50 μL supplemented RPMI media. Cells were harvested after 6 days for proliferation assessment. The extent of proliferation was measured by CFSE depletion among CD4+ T cells using anti-human CD4 antibody (Biolegend) in flow cytometry

analysis (CytoFLEX, Beckman Coulter). Proliferating cells were gated as CFSE- in CD4+ populations. Basal growth rate was determined from quadruplicate wells that contained CD4+ cell enriched PBMCs without stimuli.

MHCII Binding Assays

MHCII binding was assessed using an ELISA based assay as previously described (38). Approximately $1x10^7$ Raji cells or MHCII-deficient Raji derived RJ2.2.5 cells were plated per well in 3mL supplemented RPMI medium in a 6-well plate. Cells were incubated with 100 μg of biotinylated peptides ($CRM_{299-312}$ or $TT_{1093-1102}$) in RPMI 1640 (Corning) medium containing 2 mM L-glutamine, 1% penicillin/streptomycin, and 10% heat inactivated FBS. After 18hrs incubation at 37 degrees, cells were lysed in NP-40 buffer for 1 hour at room temperature. The lysate was cleared by centrifugation at 15,000xg for 15 minutes. ELISA assay was performed to detect the presence of MHCII bound biotinylated peptides. Briefly, 96-well plates (Costar) were coated in duplicate overnight with 5 μg/mL L243 anti-HLA-DR (Biolegend). Wells were blocked with 1% BSA in PBS and washed with 0.05% PBS-Tween (PBST). Whole cell lysates were incubated for 1 hour at room temperature. Presence of MHCII bound biotinylated peptide was detected by adding HRP conjugated Avidin (Biolegend) (1:1000 dilution) for 1 hour at room temperature. After washing, plates were developed using TMB (Biolegend) and stopped with 2 N H_2SO_4. The optical densities were determined at 450 nm using a microplate reader (Synergy H1, Bio-Tek). Significance was determined using 2-tailed Student's *t* test with $p<0.05$ comparing no antigen negative control wells to experimental cell groups incubated with biotinylated peptides.

IFN- γ Cytokine ELISA

Cytokine production from T cell stimulation was determined by ELISA. 96-well plates (Costar) were coated overnight with anti-IFN-γ antibody (1:200 dilution; Biolegend) and blocked with 1% BSA in PBS. Plates were washed with 0.05% PBST, all subsequent washes were the same. After washing, wells were incubated with cell supernatants from T cell assays for 2 hours at room temperature. After washing, biotinylated anti-IFN-γ (1:200 dilution; Biolegend) was added for 1 hour at room temperature followed by HRP-Avidin (1:1000 dilution; Biolegend) for 30 minutes at room temperature. Plates were developed using TMB substrate (Biolegend) and stopped with 2 N H_2SO_4. The optical densities were determined at 450 nm using a microplate reader (Synergy H1, Bio-Tek). Significance was determined using Student's *t* test with $p<0.05$.

HLA Locus Genotyping of Donors

HLA typing of each donor was performed by CD Genomics. Alleles of DPA1, DPB1, DQA1, DQB1, DRB1, and DR345 locus were genotyped for each donor.

Results

Immunoprecipitation to pulldown peptide-loaded MHCII proteins

To identify peptides generated through processing of CRM_{197}, TT_m, or TT_{hc} by human APCs and presented by MHCII, we adapted a previously described method of immunoprecipitation (IP) and mass spectrometry (39) **(Figure 5.1)**. We utilized a human B cell lymphoblast line, Raji cells (ATCC CCL-86) as APCs and incubated with either CRM_{197}, TT_m, or TT_{hc}. The use of two TT proteins, TT_m (full protein:light chain and heavy chain) and TT heavy chain (TT_{hc}) was to assess if T cell epitopes existed in the light chain of the protein as the majority of reported T cell epitopes are in the heavy chain

(40). Raji cells are known to express multiple alleles of -DRB1 and B3, -DPB1, and -DQA1 and B1 (41). To observe a full spectrum of peptides presented via Raji MHCII proteins, we used antibodies against all three isotypes of MHCII, HLA-DR -DP and -DQ **(Figure 5.1)**. Mass spectrometry analysis of the retentate revealed that we were able to successfully pull down each isotype of MHCII protein selectively with little cross contamination between isotypes **(Figure 5.2)**. Elution fractions were heated to dissociate the MHCII protein into releasing the bound peptides, which were then separated using ultrafiltration. Importantly, mass spectrometry confirmed the presence of MHCII in the cutoff column retentate and absence in the filtrate, which contained the eluted peptides (data not shown). Taken together, these results indicate the IP protocol is efficiently pulling down each isotype of MHCII proteins with few contaminants or isotype crossover.

Mass spectrometry analysis reveals a subset of new MHCII binding peptides

To determine the identity of the eluted peptides from each MHCII IP, we used LC-MS/MS **(Figure 5.1)**. Mass spectral analysis of eluted peptide pools revealed a set of eleven MHCII binding peptides from either TT or CRM_{197} of varying lengths naturally presented via MHCII from human B cells **(Table 5.1)**. The majority of peptides were discovered through HLA-DR IP. Four peptides were isolated from HLA-DQ IP, and only one from HLA-DP IP. Three of the peptides were observed in IP with multiple alleles **(Table 5.1)**. We found the average length of peptide epitopes from TT processing and presentation to be 12 residues, with peptides ranging from 8-15 residues. Peptides found for CRM_{197}, on the other hand, had an average length of 14 residues, ranging from 11-16 amino acids **(Table 5.1)**. We found two peptides ($TT_{826-837}$ and $TT_{1169-1179}$) that shared

overlapping sequences with known TT peptides P2 (QYIKANSKFIGITEL) and P32 (LKFIIKRYTPNNEIDS) respectively (14). $TT_{826-837}$ shares 8 amino acid residues with P2 at the N-terminus of P2. Similarly, $TT_{1169-1179}$ shares 7 residues with P32 at the N-terminus of P32. Additionally, we observed one peptide from the TT light chain TT_{94-107}.

Human donors have IgG titers against CRM_{197} and TT

To determine if these identified peptides were able to stimulate antigen-specific T cell response in a physiological scenario, human PBMCs were utilized. We first screened human sera for reactivity to both carrier proteins used in this study to confirm that each donor had existing IgG titers, therefore responsive B and T cells. Donors were previously immunized with both PCV13 and Tdap. Serum titers against each carrier protein and a negative control of BSA were determined using ELISA **(Figure 5.3)**. Serum IgG titers against carrier proteins were significantly higher than negative control BSA for all donors **(Figure 5.3)**. The variability in serum titers, particularly for TT proteins compared to CRM_{197}, most likely results from timing of vaccination for donors. Donors received PCV13 in the preceding months before this study, but the time of vaccination for Tdap was as early as 10 years prior to the study. These results indicate that the selected donors had significant IgG titers against all three carrier proteins and would be sufficient to study T cell response against identified peptides.

Identified peptides are able to stimulate donor CD4+ T cell response

Next, we assessed the ability of the newly identified CRM_{197} and TT peptides to stimulate CD4+ T cells. For this purpose, we used CD4+ enriched donor PBMCs and monitored percent of CFSE depletion in CD4+ T cells after 6 days of incubation with each peptide. We observed that each donor responded to one or more of the identified

peptides as well as the full carrier proteins **(Table 5.2)**. Donor 1 displayed increased proliferation to peptides $TT_{1169-1179}$ (P32-like peptide), $TT_{1222-1236}$, CRM_{26-39}, and $CRM_{299-312}$ **(Figure 5.4b)**. Several other peptides showed slight increase compared to the basal growth rate but were not significant. Interestingly, Donor 1 did not show increased proliferation to previously described TT peptides P2 or P32 **(Figure 5.4b)**. Donor 2 had increased proliferation compared to basal growth for peptides P32, TT_{94-107}, $TT_{1169-1179}$, $TT_{1222-1236}$, $CRM_{299-312}$, and $CRM_{425-440}$ **(Figure 5.4c)**. Donor 2 did not show significant proliferation of the TT peptide P2. Additionally, slight proliferative increases were seen in $TT_{1093-1102}$ and in CRM_{87-97}, but they were not significant compared to basal growth rate **(Figure 5.4c)**. Donors 3 and 4 showed broad significant T cell proliferation compared to basal growth rate for all identified peptides except two TT peptides each **(Table 5.2, Figure 5.4a, d, e)**. Donor 3 had no significant response to peptides $TT_{826-837}$ and $TT_{1222-1236}$, while Donor 4 had no response to peptides $TT_{826-837}$ and $TT_{1169-1179}$. Notably, one peptide, $CRM_{299-312}$, gave a positive response in all four donors, while peptides $CRM_{425-440}$, CRM_{26-39}, TT_{94-107}, $TT_{1222-1236}$, and $TT_{1169-1179}$ (P32-like peptide) gave positive response in three out of the four donors. Taken together, these results suggest each identified peptide has the capability of proliferating donor CD4+ T cells and six of the peptides responded in three or more donors.

To examine the ability of the identified peptides to stimulate the production of cytokine IFN-γ by T cell, we tested cell supernatants via ELISA **(Figure 5.5)**. Culture supernatants from CD4+ T cell enriched donor PBMCs were stimulated with full carrier proteins or identified peptides and screened for IFN-γ production after 6 days. Each identified peptide, and carrier protein, was capable of stimulating IFN-γ production in

one or more donors **(Table 5.2)**. As expected, the results for IFN-γ screening closely matched results from the proliferation assay with few discrepancies. Donor 1 had significant levels of IFN-γ response to $TT_{660-667}$, $TT_{1093-1102}$, and CRM peptides 26-39, 87-97, and 299-312 **(Figure 5.5a)**. Donor 2 displayed significant production of IFN-γ for eight identified peptides, TT_{94-107}, $TT_{1093-1102}$, $TT_{1169-1179}$, $TT_{1222-1236}$, CRM_{26-39}, CRM_{87-97}, $CRM_{299-312}$, and $CRM_{425-440}$ **(Figure 5.5b)**. Donor 3 had significant IFN-γ production towards each peptide except two **(Figure 5.5c)**. Likewise, Donor 4 had significant IFN-γ production in response to every peptide except two **(Figure 5.5d)**. Overall, eight peptides were able to produce significant levels of IFN-γ in at least three donors: TT_{94-107}, $TT_{660-667}$, $TT_{1093-1102}$, $TT_{1169-1179}$ (P32-like peptide), CRM_{26-39}, CRM_{87-97}, $CRM_{299-312}$, and $CRM_{425-440}$.

Identified peptides bind to MHCII

To corroborate immunoprecipitation data on MHCII binding, we performed an ELISA-based *in vitro* binding assay **(Figure 5.6)** adopting a previously described method (38). One peptide from each TT and CRM subsets was selected and biotinylated for MHCII binding evaluation based on their significant T cell activity **(Table 5.2, Figures 5.4 and 5.5)**. Biotinylated peptides were incubated with Raji B cells or MHCII-deficient, Raji-derived RJ2.2.5 B cells and binding was assessed. MHCII-bound biotinylated peptides were pulled down together with MHCII molecules and detected by Avidin-HRP. Compared to no antigen control, both peptides showed a significant binding to MHCII **(Figure 5.6a)**. Importantly, RJ2.2.5 cells were used as a control of MHCII binding as these B cells lack MHCII expression **(Figure 5.6b)**. There was no detection of biotinylated peptides bound with MHCII compared to the no antigen negative control

when incubated with RJ2.2.5 cells, suggesting a lack of MHCII binding and presentation. Taken together this data further supports that selected peptides bind to MHCII and these peptides are processed and presented via MHCII in APCs to stimulate T cell response.

Donors have unique subset of class II alleles

We hypothesize that donors respond to different peptides due to their distinct HLA allele expression. To reveal the correlations, we assessed the class II genotype of each donor **(Table 5.3)**. In brief, individual donor DNA was isolated and HLA gene capture was performed. After library construction deep sequencing was completed to determine HLA alleles at loci DPA1, DPB1, DQA1, DQB1, DRB1, and DR345 for each donor. Each donor has two alleles per loci with resolution to six digits **(Table 5.3)**. However, Donors 1 and 3 only had a single allele for loci DR345 **(Table 5.3)**, as not every individual possess the DRB3 loci (7).

As expected, three donors were heterozygous for each loci alleles, with the exception of DPA1 **(Table 5.3)**. This is most likely due to the low polymorphism of the DPA1 loci (7). All four donors are homozygous and express the allele DPA1*01:03:01, which is the most dominant allele in the United States populations (42).

Next, we established which alleles were shared between donors and how this correlates with positively responding peptides **(Figure 5.7)**. Interestingly, all four donors express at least one allele of DPA1*01:03 and DPB1*04:01 **(Figure 5.7a, Table 5.3)**. $CRM_{299\text{-}312}$ was the only peptide pulled down in the HLA-DP specific immunoprecipitation and it gave a positive T cell proliferative response in all four donors as determined by CFSE depletion **(Figure 5.7a)**. In looking at the HLA-DQ alleles for each donor, there was no single allele shared between all; however, Donors 1 and 3 share

two alleles, while Donors 3 and 4 share one **(Figure 5.7b)**. There were four peptides pulled down in the -DQ IP, $TT_{1093-1102}$, $TT_{1228-1239}$, $CRM_{299-312}$, and $CRM_{425-440}$. Both TT peptides gave positive response in Donors 3 and 4, while $CRM_{450-465}$ gave response in Donors 2,3, and 4 and $CRM_{299-312}$ in all four donors **(Figure 5.7b)**. Unsurprisingly, there was little overlap between donors and alleles of DRB1 and DRB345. HLA-DR loci have one of the highest levels of polymorphism with hundreds of alleles present in the population (43). However, Donors 1 and 3 share allele DRB1*01:01 **(Figure 5.7c)**. There were nine identified peptides that gave positive response in two or more donors pulled down from the -DR IP ($TT_{826-837}$ was also identified through this IP but failed to give significant response in more than one donor). Peptides $TT_{660-667}$, $TT_{1093-1102}$, and CRM_{87-97} gave significant response in Donors 3 and 4 **(Figure 5.7c)**. Peptides, TT_{94-107}, $TT_{1169-1179}$, $TT_{1222-1236}$, CRM_{26-39}, and $CRM_{425-440}$ gave significant response in three donors each. Peptide $CRM_{299-312}$ was also identified through the -DR IP, much like -DP and -DQ, and gave a significant response in all four donors **(Figure 5.7c)**. Taken together, we have identified a subset of peptides that give significant T cell response in multiple donors that have different MHC class II genotypes, and that donors have different alleles from Raji B cells which were first used to identify peptides. This suggests possible immunodominant roles for these peptides in immune presentation and may have implications for rational vaccine design.

Discussion

Past reports have shed light on T cell epitopes of carrier proteins CRM_{197} and TT through bioinformatics or T cell screening of synthetic peptides spanning the proteins (14,18,21,29-36). This study sought to explore the epitopes that are naturally processed

by human B cells and presented via MHCII when exposed to the common glycoconjugate vaccine carrier proteins CRM_{197} and tetanus toxoid. Herein, we define a set of eleven CD4+ T cell epitopes for TT or CRM_{197} proteins utilizing mass spectrometry of MHCII presented peptides. To our best knowledge, none of these peptides have been reported previously with the specific amino acid sequences observed here. However, it is important to note that peptide epitopes described here share overlapping sequences in various degrees with epitopes reported on the Immune Epitope Database and Analysis Resource (IEDB). Each of the eleven peptides is capable of stimulating CD4+ T cells in at least one donor, with six peptides stimulating T cells in three or more donors, suggesting an immunodominant role for these epitopes. In support of this, all four donors have different class II genotypes from each other as well as from the Raji cell line originally used to identify the peptides demonstrating these peptides are associated with multiple alleles of MHCII. Based on a recently published study analyzing HLA-DR and DQ alleles in US population, one or more MHCII alleles in all four donors are identified as top 10 most common alleles in the United States population across ethnicities (44). Additionally, every allele (DRB1, DQB1 and DPB1) in all donors and Raji cell line are considered a common allele as described by the Common and Well-Documented (CWD) alleles catalogue (42,45,46). This allelic information paired with our donor T cell data suggests these identified peptides could bind MHCII in a large subset of the population, making them ideal vaccine components.

A subset of the observed peptides did not show significant response, either through proliferation or IFNγ secretion, in one or more donors. There could be a few explanations concerning this. First, it is possible that a negative response could be due to

a low frequency of responding cells. Since peptides were probed with total CD4+ T cells and not TT or CRM specific T cells alone, wells could have contained a lower frequency of carrier protein specific CD4+ T cells leading to a lower response. Second, the timing of Tdap for donors could play a role in T cell responses against TT. Verbal confirmation of Tdap was received from donors which was further tested by IgG titers against each carrier protein. While each donor displayed titers to both TT and CRM carrier proteins, it remains that the time since vaccination could cause variation in the observed results. For example, donor 1 displays relatively low levels of IgG against TT suggesting time of vaccination was years prior. Would a booster of the Tdap vaccine increase TT IgG titers in this donor and change the observed results? Lastly, there has been increased interest in the role that biological sex has on vaccine efficacy (47-49). Studies suggest that females tend to display greater vaccine efficacy and higher antibody titers post vaccination (47-49). In this current study all human donors were male which may affect the responses we observed here. However, it is encouraging that all peptides gave a positive response in one or more donors given that research suggest males may have lower vaccine efficacy. This further suggests these peptides would be ideal vaccine components. In any vaccination strategy or study the importance of biological sex, age, and vaccine components should be considered.

Interestingly, we observed no significant response against the previously published TT epitopes P2 and P32 in one or more donors. This is consistent with a recent study on T cell responses to TT (34) wherein researchers show no significant responses to TT epitope P2. Having this observation in multiple studies necessitates a direct

methodology that follows natural MHCII pathway to define T cell epitopes as laid out in this study.

During our MS analysis of the MHCII presented peptides, no proteases were utilized prior to MS to capture the full length and sequence of the MHCII bound peptides. Discovery of two peptides that overlap slightly with P2 and P32 with different N-terminal residues demonstrate that identification of naturally processed peptides can be preferable to alternative screening approaches. Thus, characterizing epitopes isolated from MHCII pathway could yield new immunodominant peptides. These findings may also suggest that MHCII proteins may prefer different binding registers than what has been previously described. Indeed, a number of reports have shown the ability of MHCII proteins to bind well-known MHCII epitope $OVA_{323-339}$ in distinct registers (50,51). These studies indicate that MHCII proteins are capable of binding the same peptide in different ways, suggesting why we observed similar, yet distinct peptides from what has been described. This also suggests that previously published peptides identified through use of overlapping synthetic peptides spanning the complete protein sequence or peptides found through prediction algorithms may not be: 1) what MHCII preferentially binds to and 2) the most effective approach for identifying naturally presented immunogenic peptides.

Recently, there has been a shift towards utilizing immunogenic peptides as carriers for vaccine candidates over the full-length carrier proteins. Several past studies have explored this idea (8,18,22-24,26-28) and shown that the candidates work as well (19-21), or better than traditional carrier protein vaccines (8). Importantly, this shift is not just a current trend, but rather based on logic and our knowledge of the immune system. After processing in the endolysosome of the APCs, proteins are broken into fragments

with only a small subset of these fragments capable of MHCII binding. By supplying the immune system with the exact epitopes necessary for presentation we can ensure every epitope is utilized to enrich for a more robust T cell mediated response. Additionally, the use of these peptide constructs for vaccine candidates accurately recapitulates what is occurring in the endolysosome. Furthermore, these peptides are the smallest units possible for MHC binding and are not likely degraded further, which may alter their immune activity.

Discovering a repertoire of MHCII-binding peptides derived from multiple carrier proteins is critical since most peptides by themselves are limited in their MHCII allelic coverage. Therefore, a number of strategies have been proposed and studied to overcome this limitation. Studies have been done on linking strings of MHCII promiscuous peptides together (19,20,25), utilizing liposomes (52), nanoparticles (53), and more. These studies show the feasibility and future direction of utilizing peptide epitopes for the generation of glycoconjugate vaccines over full-length carrier proteins. Moreover, shifting towards immunogenic peptide epitopes allows for more robust and cost-effective means of vaccine production as peptides can be produced on a larger scale, high yield, and at low cost. The knowledge gained from this work will aid in defining these MHCII binding peptides to be used in the production of knowledge-based, structurally-defined next generation vaccines.

References

1. Goldblatt, D. (2000) Conjugate vaccines. *Clin Exp Immunol* **119**, 1-3
2. Wantuch, P. L., and Avci, F. Y. (2018) Current status and future directions of invasive pneumococcal diseases and prophylactic approaches to control them. *Hum Vaccin Immunother* **14**, 2303-2309
3. Pichichero, M. E. (2013) Protein carriers of conjugate vaccines: characteristics, development, and clinical trials. *Hum Vaccin Immunother* **9**, 2505-2523
4. Micoli, F., Adamo, R., and Costantino, P. (2018) Protein Carriers for Glycoconjugate Vaccines: History, Selection Criteria, Characterization and New Trends. *Molecules* **23**
5. Hunt, D. F., Michel, H., Dickinson, T. A., Shabanowitz, J., Cox, A. L., Sakaguchi, K., Appella, E., Grey, H. M., and Sette, A. (1992) Peptides presented to the immune system by the murine class II major histocompatibility complex molecule I-Ad. *Science* **256**, 1817-1820
6. Avalos, A. M., and Ploegh, H. L. (2014) Early BCR Events and Antigen Capture, Processing, and Loading on MHC Class II on B Cells. *Front Immunol* **5**, 92
7. Murphy, K., and Weaver, C. (2017) *Janeway's Immunobiology*, 9 ed., Garland Science, Taylore & Francis Group LLC, New York, NY
8. Avci, F., Li, X., Tsuji, M., and Kasper, D. (2011) A mechanism for glycoconjugate vaccine activation of the adaptive immune system and its implications for vaccine design. *Nature Medicine* **17**, 1602-U1115

9. Middleton, D. R., Sun, L., Paschall, A. V., and Avci, F. Y. (2017) T Cell-Mediated Humoral Immune Responses to Type 3 Capsular Polysaccharide of. *J Immunol* **199**, 598-603

10. Sun, X., Stefanetti, G., Berti, F., and Kasper, D. L. (2019) Polysaccharide structure dictates mechanism of adaptive immune response to glycoconjugate vaccines. *Proc Natl Acad Sci U S A* **116**, 193-198

11. Sun, L., Middleton, D. R., Wantuch, P. L., Ozdilek, A., and Avci, F. Y. (2016) Carbohydrates as T-cell antigens with implications in health and disease. *Glycobiology* **26**, 1029-1040

12. Rappuoli, R., De Gregorio, E., and Costantino, P. (2019) On the mechanisms of conjugate vaccines. *Proc Natl Acad Sci U S A* **116**, 14-16

13. Avci, F. (2013) Novel Strategies for Development of Next-Generation Glycoconjugate Vaccines. *Current Topics in Medicinal Chemistry* **13**, 2535-2540

14. Panina-Bordignon, P., Tan, A., Termijtelen, A., Demotz, S., Corradin, G., and Lanzavecchia, A. (1989) Universally immunogenic T cell epitopes: promiscuous binding to human MHC class II and promiscuous recognition by T cells. *Eur J Immunol* **19**, 2237-2242

15. Wantuch, P. L., and Avci, F. Y. (2019) Invasive pneumococcal disease in relation to vaccine type serotypes. *Hum Vaccin Immunother*, 1-2

16. Schutze, M. P., Leclerc, C., Jolivet, M., Audibert, F., and Chedid, L. (1985) Carrier-induced epitopic suppression, a major issue for future synthetic vaccines. *J Immunol* **135**, 2319-2322

17. Schutze, M. P., Leclerc, C., Vogel, F. R., and Chedid, L. (1987) Epitopic suppression in synthetic vaccine models: analysis of the effector mechanisms. *Cell Immunol* **104**, 79-90

18. Etlinger, H. M., Gillessen, D., Lahm, H. W., Matile, H., Schönfeld, H. J., and Trzeciak, A. (1990) Use of prior vaccinations for the development of new vaccines. *Science* **249**, 423-425

19. Falugi, F., Petracca, R., Mariani, M., Luzzi, E., Mancianti, S., Carinci, V., Melli, M. L., Finco, O., Wack, A., Di Tommaso, A., De Magistris, M. T., Costantino, P., Del Giudice, G., Abrignani, S., Rappuoli, R., and Grandi, G. (2001) Rationally designed strings of promiscuous CD4+ T cell epitopes provide help to Haemophilus influenzae type b oligosaccharide: a model for new conjugate vaccines. *Eur. J. Immunol* **31**, 3816-3824

20. Baraldo, K., Mori, E., Bartoloni, A., Norelli, F., Grandi, G., Rappuoli, R., Finco, O., and Del Giudice, G. (2005) Combined conjugate vaccines: enhanced immunogenicity with the N19 polyepitope as a carrier protein. *Infect Immun* **73**, 5835-5841

21. Bixler, G. S., Eby, R., Dermody, K. M., Woods, R. M., Seid, R. C., and Pillai, S. (1989) Synthetic peptide representing a T-cell epitope of CRM197 substitutes as carrier molecule in a Haemophilus influenzae type B (Hib) conjugate vaccine. *Adv Exp Med Biol* **251**, 175-180

22. Jackson, D. C., Purcell, A. W., Fitzmaurice, C. J., Zeng, W., and Hart, D. N. (2002) The central role played by peptides in the immune response and the design

of peptide-based vaccines against infectious diseases and cancer. *Curr Drug Targets* **3**, 175-196

23. Cai, H., Chen, M. S., Sun, Z. Y., Zhao, Y. F., Kunz, H., and Li, Y. M. (2013) Self-adjuvanting synthetic antitumor vaccines from MUC1 glycopeptides conjugated to T-cell epitopes from tetanus toxoid. *Angew Chem Int Ed Engl* **52**, 6106-6110

24. de Velasco, E. A., Merkus, D., Anderton, S., Verheul, A. F., Lizzio, E. F., Van der Zee, R., Van Eden, W., Hoffman, T., Verhoef, J., and Snippe, H. (1995) Synthetic peptides representing T-cell epitopes act as carriers in pneumococcal polysaccharide conjugate vaccines. *Infect Immun* **63**, 961-968

25. Slingluff, C. L. (2011) The present and future of peptide vaccines for cancer: single or multiple, long or short, alone or in combination? *Cancer J* **17**, 343-350

26. Rodrigues-da-Silva, R. N., Correa-Moreira, D., Soares, I. F., de-Luca, P. M., Totino, P. R. R., Morgado, F. N., Oliveira Henriques, M. D. G., Peixoto Candea, A. L., Singh, B., Galinski, M. R., Moreno, A., Oliveira-Ferreira, J., and Lima-Junior, J. D. C. (2019) Immunogenicity of synthetic peptide constructs based on PvMSP9E795-A808, a linear B-cell epitope of the P. vivax Merozoite Surface Protein-9. *Vaccine* **37**, 306-313

27. Snook, A. E., Baybutt, T. R., Hyslop, T., and Waldman, S. A. (2016) Preclinical Evaluation of a Replication-Deficient Recombinant Adenovirus Serotype 5 Vaccine Expressing Guanylate Cyclase C and the PADRE T-helper Epitope. *Hum Gene Ther Methods* **27**, 238-250

28. Zeigler, D. F., Roque, R., and Clegg, C. H. (2019) Optimization of a multivalent peptide vaccine for nicotine addiction. *Vaccine* **37**, 1584-1590

29. Demotz, S., Lanzavecchia, A., Eisel, U., Niemann, H., Widmann, C., and Corradin, G. (1989) Delineation of several DR-restricted tetanus toxin T cell epitopes. *J Immunol* **142**, 394-402

30. Diethelm-Okita, B. M., Okita, D. K., Banaszak, L., and Conti-Fine, B. M. (2000) Universal epitopes for human CD4+ cells on tetanus and diphtheria toxins. *J Infect Dis* **181**, 1001-1009

31. Diethelm-Okita, B. M., Raju, R., Okita, D. K., and Conti-Fine, B. M. (1997) Epitope repertoire of human CD4+ T cells on tetanus toxin: identification of immunodominant sequence segments. *J Infect Dis* **175**, 382-391

32. Southwood, S., Sidney, J., Kondo, A., del Guercio, M. F., Appella, E., Hoffman, S., Kubo, R. T., Chesnut, R. W., Grey, H. M., and Sette, A. (1998) Several common HLA-DR types share largely overlapping peptide binding repertoires. *J Immunol* **160**, 3363-3373

33. Demotz, S., Barbey, C., Corradin, G., Amoroso, A., and Lanzavecchia, A. (1993) The set of naturally processed peptides displayed by DR molecules is tuned by polymorphism of residue 86. *Eur J Immunol* **23**, 425-432

34. da Silva Antunes, R., Paul, S., Sidney, J., Weiskopf, D., Dan, J. M., Phillips, E., Mallal, S., Crotty, S., Sette, A., and Lindestam Arlehamn, C. S. (2017) Definition of Human Epitopes Recognized in Tetanus Toxoid and Development of an Assay Strategy to Detect Ex Vivo Tetanus CD4+ T Cell Responses. *Plos One* **12**, e0169086

35. Leonard, E. G., Canaday, D. H., Harding, C. V., and Schreiber, J. R. (2003) Antigen processing of the heptavalent pneumococcal conjugate vaccine carrier protein CRM197 differs depending on the serotype of the attached polysaccharide. *Infect. Immun.* **71**, 4186-4189

36. Raju, R., Navaneetham, D., Okita, D., Diethelm-Okita, B., McCormick, D., and Conti-Fine, B. M. (1995) Epitopes for human CD4+ cells on diphtheria toxin: structural features of sequence segments forming epitopes recognized by most subjects. *Eur J Immunol* **25**, 3207-3214

37. Di Carluccio, A. R., Tresoldi, E., So, M., and Mannering, S. I. (2019) Quantification of Proliferating Human Antigen-specific CD4+ T Cells using Carboxyfluorescein Succinimidyl Ester. *J Vis Exp* **148**, e59545

38. Cobb, B. A., Wang, Q., Tzianabos, A. O., and Kasper, D. L. (2004) Polysaccharide processing and presentation by the MHCII pathway. *Cell* **117**, 677-687

39. Bozzacco, L., Yu, H., Zebroski, H. A., Dengjel, J., Deng, H., Mojsov, S., and Steinman, R. M. (2011) Mass spectrometry analysis and quantitation of peptides presented on the MHC II molecules of mouse spleen dendritic cells. *J Proteome Res* **10**, 5016-5030

40. Kerblat, I., Tongiani-Dahshan, S., Aude-Garcia, C., Villiers, M., Drouet, C., and Marche, P. N. (2000) Tetanus toxin L chain is processed by major histocompatibility complex class I and class II pathways and recognized by CD8+ or CD4+ T lymphocytes. *Immunology* **100**, 178-184

41. Bentley, G., Higuchi, R., Hoglund, B., Goodridge, D., Sayer, D., Trachtenberg, E. A., and Erlich, H. A. (2009) High-resolution, high-throughput HLA genotyping by next-generation sequencing. *Tissue Antigens* **74**, 393-403

42. Gonzalez-Galarza, F. F., Takeshita, L. Y., Santos, E. J., Kempson, F., Maia, M. H., da Silva, A. L., Teles e Silva, A. L., Ghattaoraya, G. S., Alfirevic, A., Jones, A. R., and Middleton, D. (2015) Allele frequency net 2015 update: new features for HLA epitopes, KIR and disease and HLA adverse drug reaction associations. *Nucleic Acids Res* **43**, D784-788

43. van Lith, M., McEwen-Smith, R. M., and Benham, A. M. (2010) HLA-DP, HLA-DQ, and HLA-DR have different requirements for invariant chain and HLA-DM. *J Biol Chem* **285**, 40800-40808

44. Maiers, M., Gragert, L., and Klitz, W. (2007) High-resolution HLA alleles and haplotypes in the United States population. *Hum Immunol* **68**, 779-788

45. Mack, S. J., Cano, P., Hollenbach, J. A., He, J., Hurley, C. K., Middleton, D., Moraes, M. E., Pereira, S. E., Kempenich, J. H., Reed, E. F., Setterholm, M., Smith, A. G., Tilanus, M. G., Torres, M., Varney, M. D., Voorter, C. E., Fischer, G. F., Fleischhauer, K., Goodridge, D., Klitz, W., Little, A. M., Maiers, M., Marsh, S. G., Müller, C. R., Noreen, H., Rozemuller, E. H., Sanchez-Mazas, A., Senitzer, D., Trachtenberg, E., and Fernandez-Vina, M. (2013) Common and well-documented HLA alleles: 2012 update to the CWD catalogue. *Tissue Antigens* **81**, 194-203

46. Castelli, F. A., Buhot, C., Sanson, A., Zarour, H., Pouvelle-Moratille, S., Nonn, C., Gahery-Ségard, H., Guillet, J. G., Ménez, A., Georges, B., and Maillère, B.

(2002) HLA-DP4, the most frequent HLA II molecule, defines a new supertype of peptide-binding specificity. *J Immunol* **169**, 6928-6934

47. Fink, A. L., and Klein, S. L. (2015) Sex and Gender Impact Immune Responses to Vaccines Among the Elderly. *Physiology (Bethesda)* **30**, 408-416

48. Fink, A. L., Engle, K., Ursin, R. L., Tang, W. Y., and Klein, S. L. (2018) Biological sex affects vaccine efficacy and protection against influenza in mice. *Proc Natl Acad Sci U S A* **115**, 12477-12482

49. Fischinger, S., Boudreau, C. M., Butler, A. L., Streeck, H., and Alter, G. (2019) Sex differences in vaccine-induced humoral immunity. *Semin Immunopathol* **41**, 239-249

50. Scott, C. A., Peterson, P. A., Teyton, L., and Wilson, I. A. (1998) Crystal structures of two I-Ad-peptide complexes reveal that high affinity can be achieved without large anchor residues. *Immunity* **8**, 319-329

51. McFarland, B. J., Sant, A. J., Lybrand, T. P., and Beeson, C. (1999) Ovalbumin (323-339) peptide binds to the Major Histocompatibility Complex Class II I-A protein using two functionally distinct registers. *Biochemistry* **38**, 16663-16670

52. Riaz, M. K., Riaz, M. A., Zhang, X., Lin, C., Wong, K. H., Chen, X., Zhang, G., Lu, A., and Yang, Z. (2018) Surface Functionalization and Targeting Strategies of Liposomes in Solid Tumor Therapy: A Review. *Int J Mol Sci* **19**

53. Fujita, Y., and Taguchi, H. (2011) Current status of multiple antigen-presenting peptide vaccine systems: Application of organic and inorganic nanoparticles. *Chem Cent J* **5**, 48

Figure 5.1

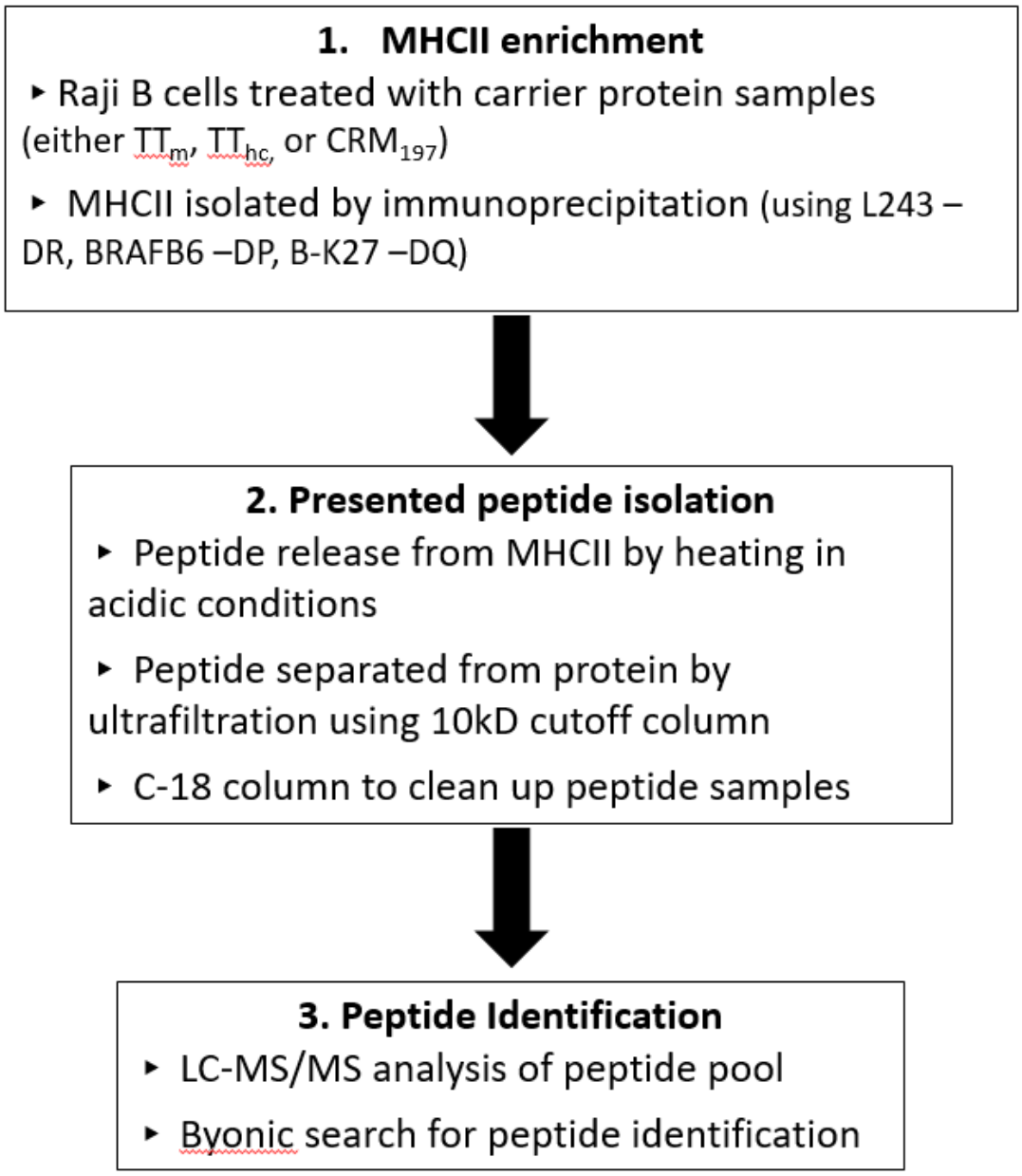


Figure 5.1: Schematic representation of experimental design to identify peptides presented on differing isotypes of MHCII molecules on human B cells after treatment with common carrier proteins used in vaccine design (CRM_{197} and tetanus toxoid [TT]). Diagram shows steps from cell treatment through peptide identification.

Figure 5.2

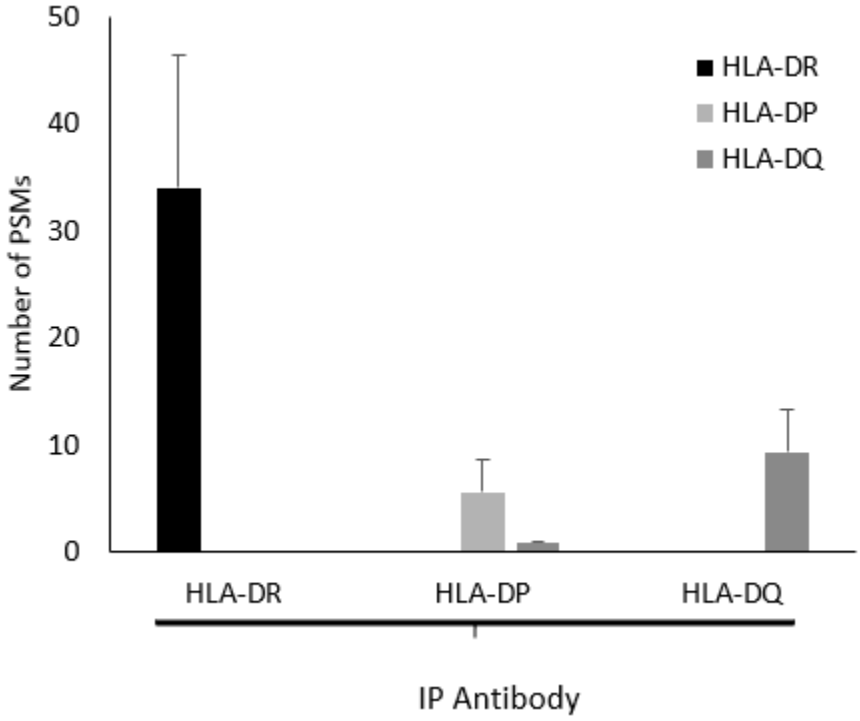


Figure 5.2: Mass spectral analysis of immunoprecipitation products. Identified MHCII proteins in bead bound product from each immunoprecipitated isotype compared via number of peptide spectral matches (PSMs). The data represents the average of number of PSMs from three biological replicates. The error bars are standard deviation.

Table 5.1

Table 5.1: List of peptides identified through LC-MS/MS analysis of MHCII immunoprecipitation. Table lists the peptide sequences identified from B cells treated with carrier protein and the amino acid residues at which they occur, the MHCII isotype used for their pulldown, and the calculated and observed parent ion masses (all masses converted to singly charged $M+H^+$ m/z) from mass spectra.

Peptide	Sequence	Observed in Isotypes	Parent Ion $[M+H]^+$	Predicted $[M+H]^+$
TT_{94-107}	LFNRIKNNVAGEAL	DR	1558.881	1558.870
$TT_{660-667}$	NFIGALET	DR	864.454	864.446
$TT_{826-837}$	NILMQYIKANSK	DR	1294.704	1294.682
$TT_{1093-1102}$	CKALNPKEIE	DR, DQ	1144.596	1144.603
$TT_{1169-1179}$	LYNGLKFIIKR	DR	1364.850	1364.841
$TT_{1222-1236}$	DRILRVGYNAPGIPL	DR	1653.943	1653.943
$TT_{1228-1239}$	GYNAPGIPLYKK	DQ	1320.733	1320.731
CRM_{26-39}	GYVDSIQKGIQKPK	DR	1560.873	1560.874
CRM_{87-97}	GLTKVLALKVD	DR	1156.723	1156.730
$CRM_{299-312}$	KTTAALSILPGIGS	DR, DP, DQ	1328.778	1328.787
$CRM_{425-440}$	TPLPIAGVLLPTIPGK	DR, DQ	1586.984	1586.988

Figure 5.3

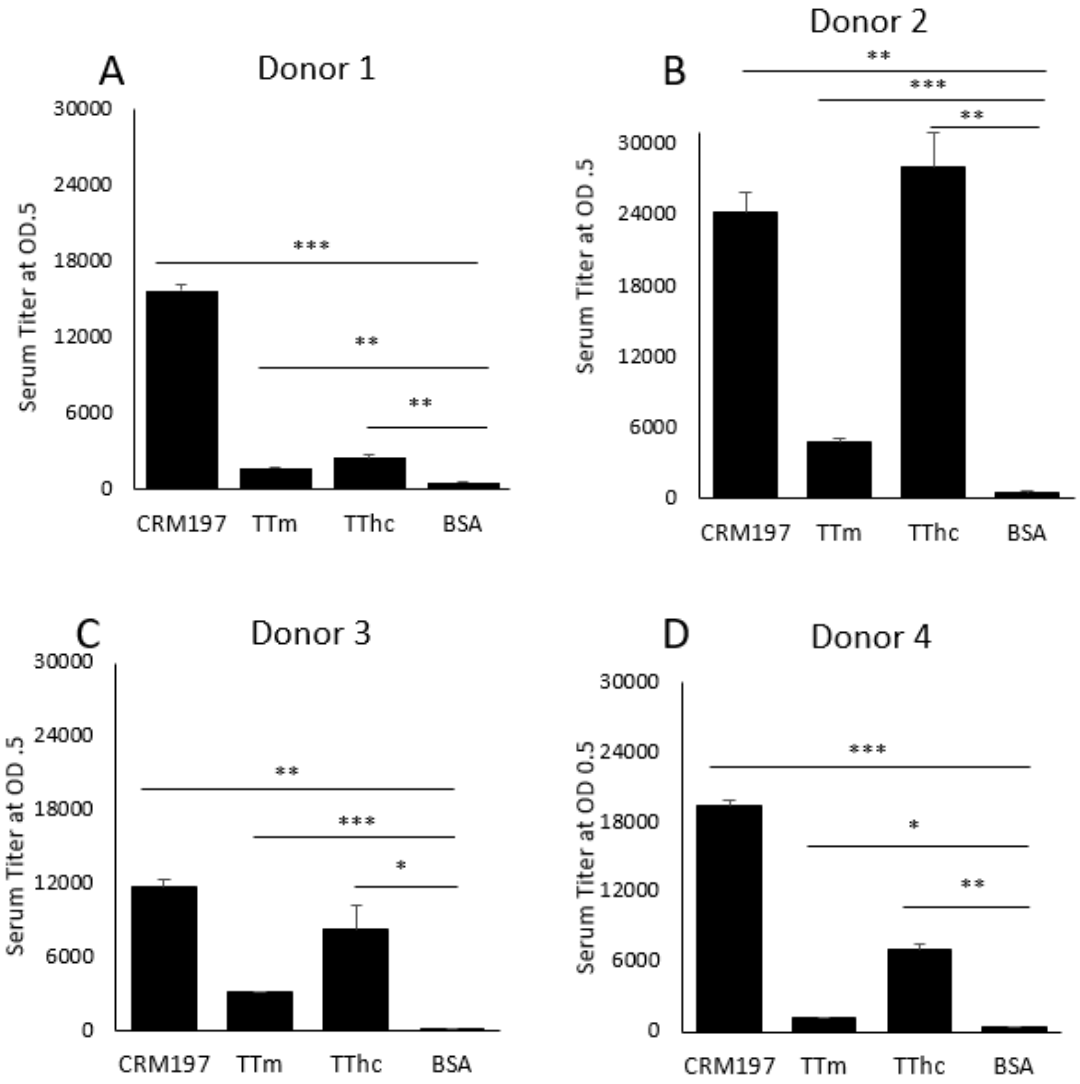


Figure 5.3: Serum IgG titers of four human donors. Serum titers are for three carrier proteins CRM_{197}, TT_m, and TT_{hc}. BSA was used as a negative control. **A)** Donor 1 **B)** Donor 2 **C)** Donor 3 **D)** Donor 4 serum titers were determined at OD 0.5. Significance was determined using Student's *t* test with $p<0.05$

Table 5.2

Table 5.2: List of peptides that gave a positive response in at least one donor. Results are shown for CFSE staining and IFN-γ ELISA per donor. Significance was determined using student's t test and is given as * p<0.05 **p<0.005 ***p<0.0005. Boxes are color coded based on significance value as pale green (*), light green (**), and dark green (***).

	Donor 1		Donor 2		Donor 3		Donor 4	
	CFSE	IFN-γ	CFSE	IFN-γ	CFSE	IFN-γ	CFSE	IGN-γ
P2					*		***	*
P32			***	**	**	**	***	*
TT_{94-107}			***	**	***	**	**	*
$TT_{660-667}$		*			**	*	***	*
$TT_{826-837}$						*		
$TT_{1093-1102}$		*		*	**	**	***	*
$TT_{1169-1179}$	*		***	*	*	*		*
$TT_{1222-1236}$	*		***	***			*	
$TT_{1228-1239}$					**	**	***	**
CRM_{26-39}	*	**		**	*	*	***	***
CRM_{87-97}		*		**	**	***	***	**
$CRM_{299-312}$	***	*	***	**	**	*	***	**
$CRM_{425-440}$			***	**	**	**	***	***

Figure 5.4

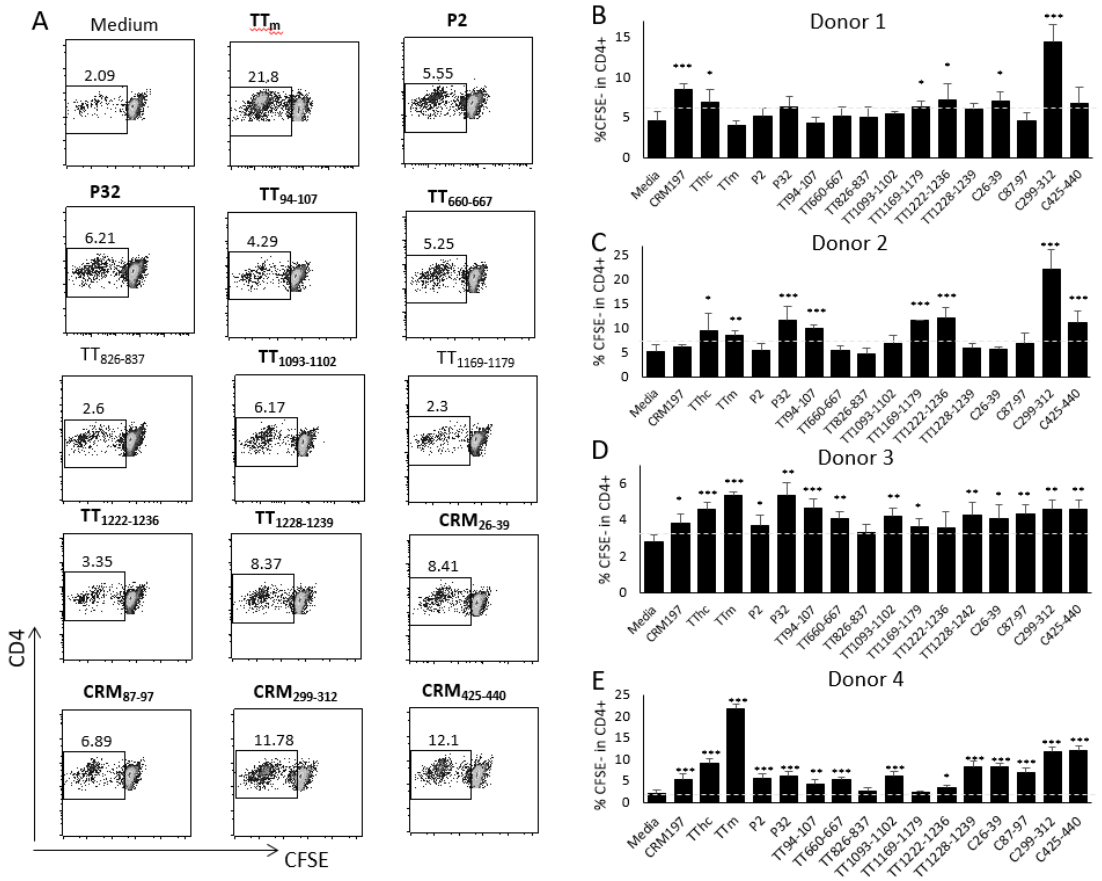
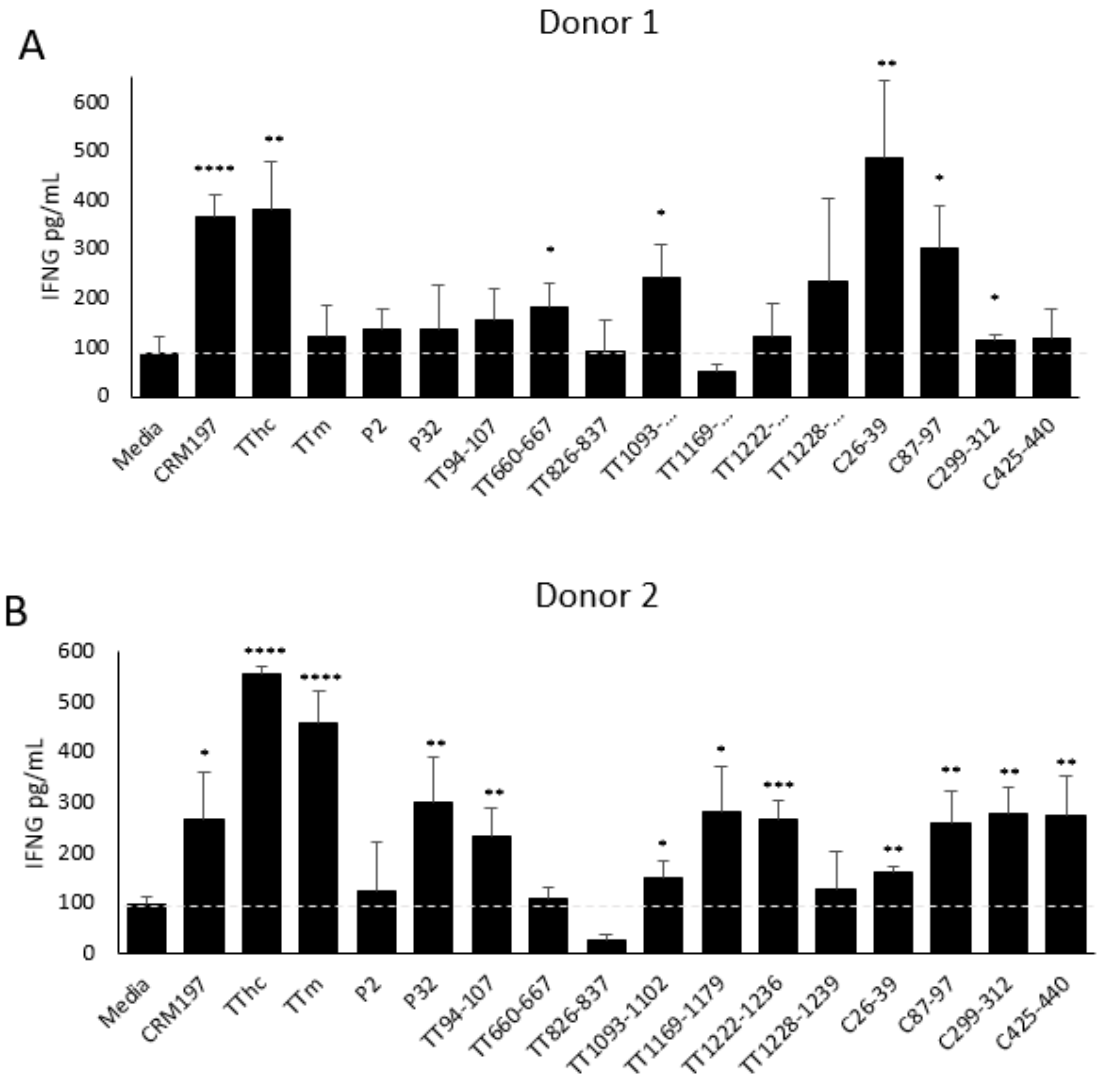


Figure 5.4: Proliferation of CD4+ enriched PBMCs from four donors. A) Donor 4 gating of CD4+ cells with CFSE- populations squared off. CFSE responsive peptide stimulations are bold. TT_m is shown as positive control and medium as negative. % CFSE- in CD4+ for **B)** Donor 1 **C)** Donor 2 **D)** Donor 3 and **E)** Donor 4. PBMCs were enriched for CD4+ T cells and APCs then incubated with respective antigen supplemented with IL-2. Proliferation was measured by gating CD4+ T cells and measuring percent of CFSE- in CD4+ populations. Student's *t* test was performed for statistical value with $p<0.05$.

Figure 5.5


A
Donor 1
IFNG pg/mL
600
500
400
300
200
100
0
Media
CRM197
TThc
TTm
P2
P32
TT94-107
TT660-667
TT826-837
TT1093-...
TT1169-...
TT1222-...
TT1228-...
C26-39
C87-97
C299-312
C425-440
B
Donor 2
IFNG pg/mL
600
500
400
300
200
100
0
Media
CRM197
TThc
TTm
P2
P32
TT94-107
TT660-667
TT826-837
TT1093-1102
TT1169-1179
TT1222-1236
TT1228-1239
C26-39
C87-97
C299-312
C425-440

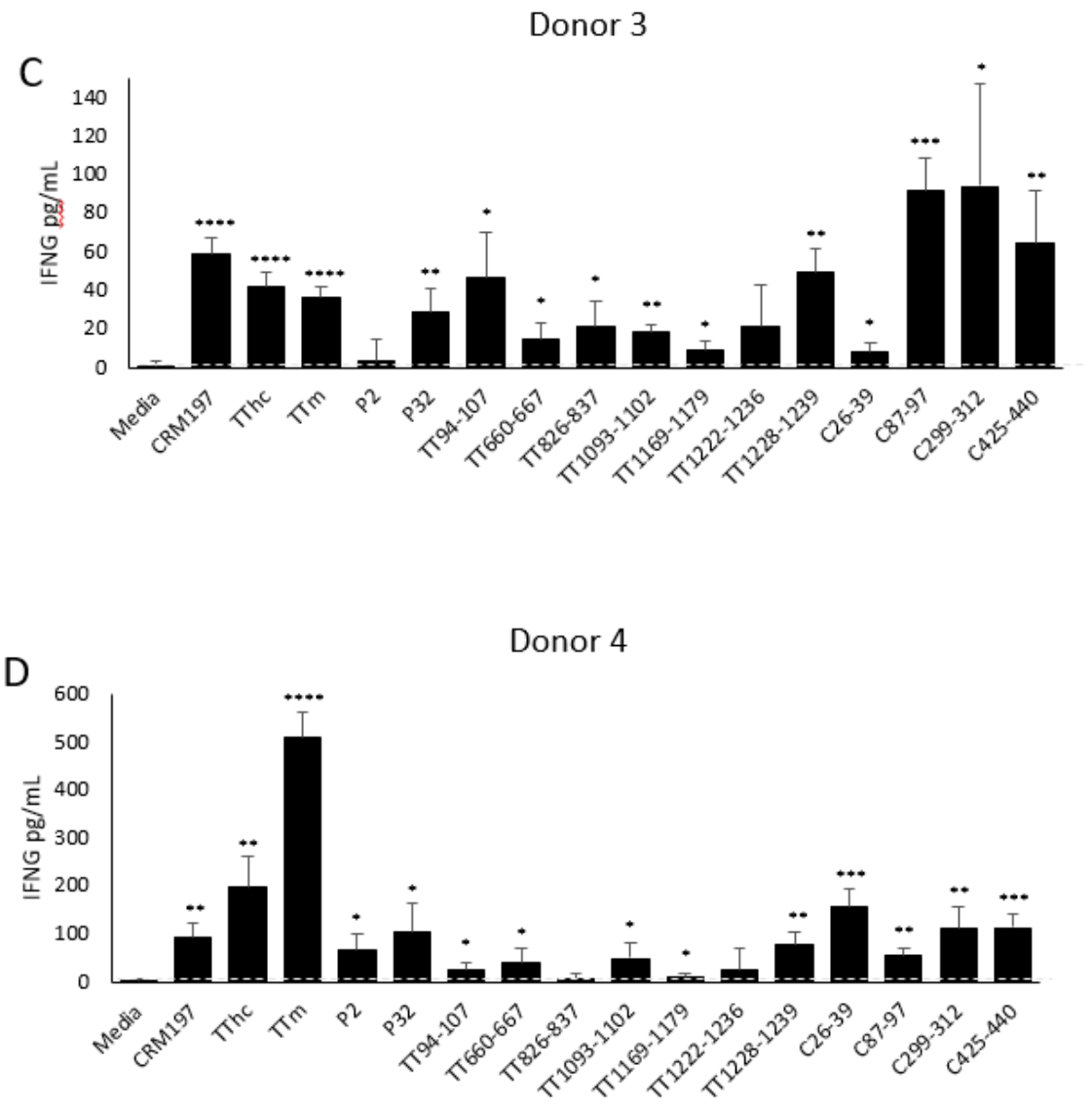


Figure 5.5: IFN-γ cytokine secretion of human PBMCs from four donors. **A)** Donor 1 **B)** Donor 2 **C)** Donor 3 **D)** Donor 4. PBMCs were enriched for CD4+ T cells and APCs then incubated with respective antigen supplemented with IL-2. Cytokine secretion was measured using ELISA assay and output was converted to product formed in pg/mL using IFN-γ human standards. Student's *t* test was performed for statistical value against media blank with $p<0.05$.

Figure 5.6

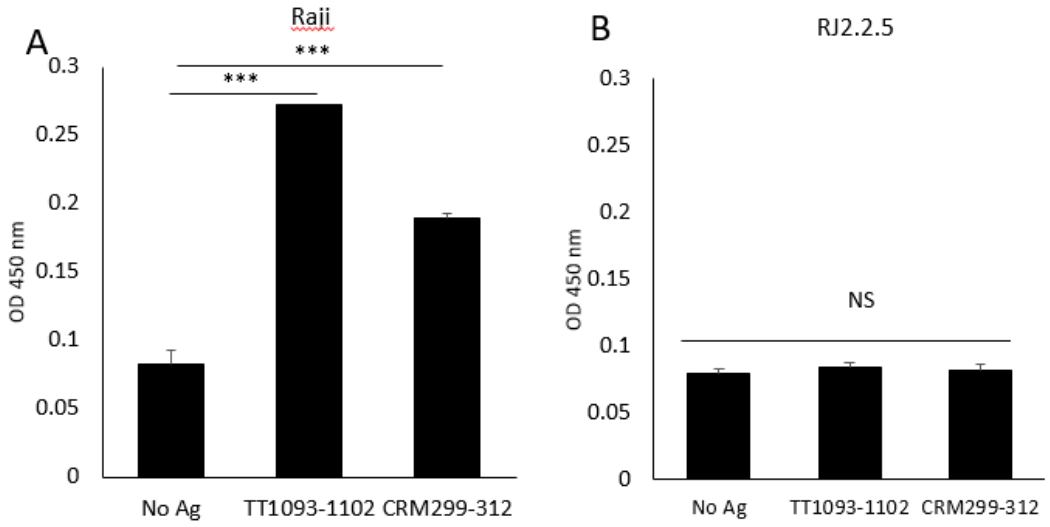


Figure 5.6: MHCII presentation of peptides for T cell recognition. A) Raji B or **B)** RJ2.2.5 cells were incubated with biotinylated peptides. Whole cell lysates were incubated on L243 anti-HLA-DR (Biolegend) coated ELISA plates and binding to MHCII was measured by Avidin-HRP.

Figure 5.7

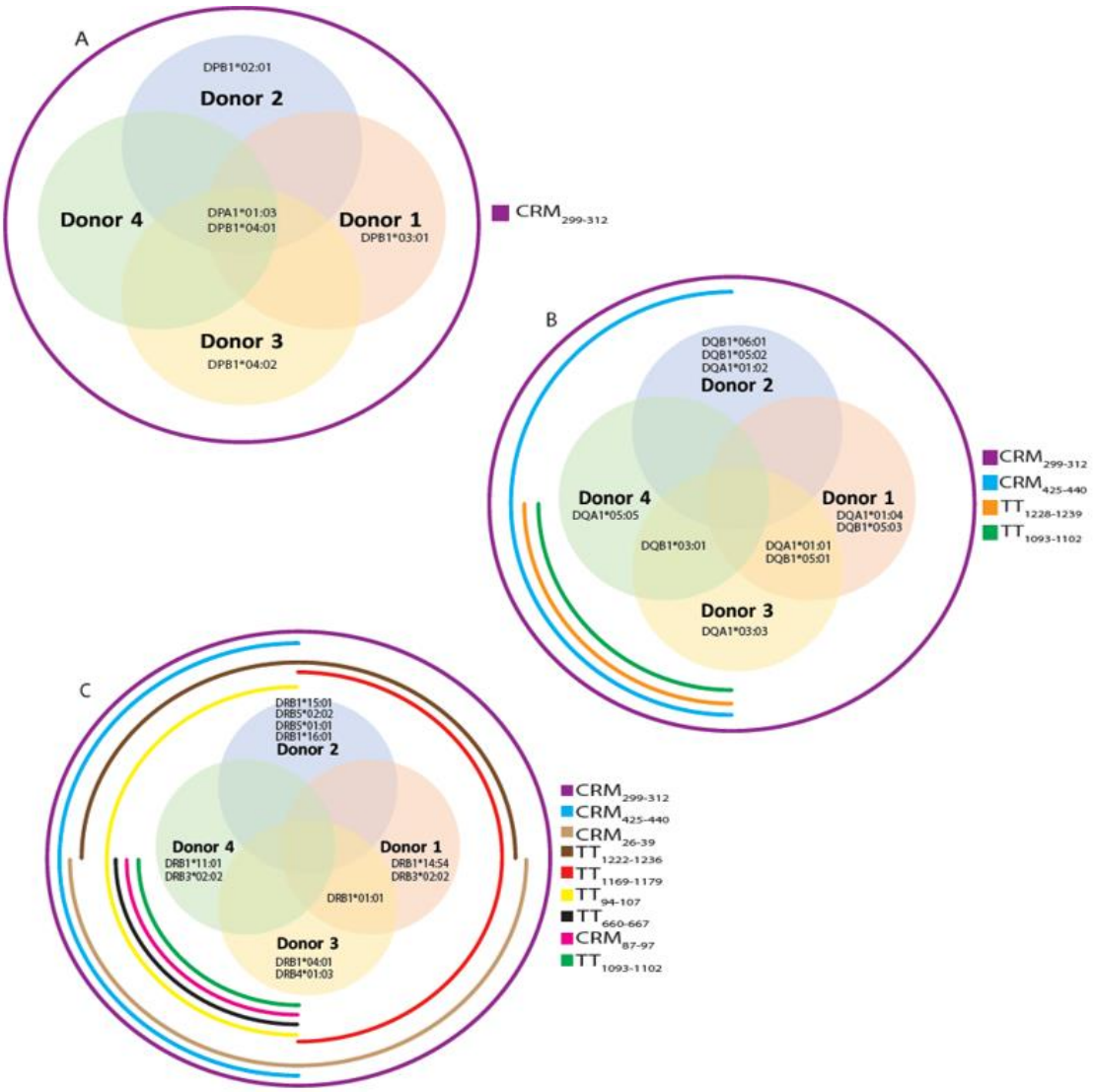


Figure 7: Three tier comparison of the class II alleles expressed in each donor for each isotype of MHCII and the peptides presented by that isotype. Venn diagram depicts comparison of class II alleles expressed in each donor for **A)** DP, **B)** DQ and **C)** DR isotypes of MHCII. The outer colored lines represent the peptides that were identified via mass spectrometry for that isotype of MHCII. These lines surround the donors which had a positive T cell proliferative response (as determined by CFSE staining).

CHAPTER 6

CHARACTERIZATION OF THE β-GLUCURONIDASE PN3PASE AS THE FOUNDING MEMBER OF GLYCOSIDE HYDROLASE FAMILY GHxxx

---

Paeton L Wantuch, Satya Jella, Jeremy A Duke, Jarrod J Mousa, Bernard Henrissat, John N Glushka and Fikri Y Avci. Accepted: *Glycobiology* 2020. Reprinted here with permission of publisher.

Abstract

Paenibacillus sp. 32352 is a soil-dwelling bacterium capable of producing an enzyme, Pn3Pase that degrades the capsular polysaccharide of *Streptococcus pneumoniae* serotype 3 (Pn3P). Recent reports on Pn3Pase have demonstrated its initial characterization and potential for protection against highly virulent *S. pneumoniae* serotype 3 infections. Initial experiments revealed this enzyme functions as an exo-β1,4-glucuronidase cleaving the β(1,4) linkage between glucuronic acid and glucose. However, the catalytic mechanism of this enzyme is still unknown. Here, we report the detailed biochemical analysis of Pn3Pase. Pn3Pase shows no significant sequence similarity to known glycoside hydrolase families, thus this novel enzyme establishes a new carbohydrate-active enzyme (CAZy) glycoside hydrolase family. Site directed mutagenesis studies revealed two catalytic residues along with truncation mutants defining essential domains for function. Pn3Pase and its mutants were screened for activity, substrate binding and kinetics. Additionally, NMR analysis revealed that Pn3Pase acts through a retaining mechanism. This study exhibits Pn3Pase activity at the structural and mechanistic level to establish the new CAZy glycoside hydrolase family GHxxx belonging to the large GH-A clan. (Final family number will be assigned at proof stage if manuscript is deemed acceptable for publication in *Glycobiology*) This study will also serve towards generating Pn3Pase derivatives with optimal activity and pharmacokinetics aiding in the use of Pn3Pase as a novel therapeutic approach against type 3 *S. pneumoniae* infections.

Introduction

A number of early studies demonstrated an enzyme found in *Bacillus circulans* was capable of degrading the capsular polysaccharide of highly virulent type 3 *Streptococcus pneumoniae* (Pn3P) (1-3). However, upon sequencing the genome of this bacterial strain "*Bacillus circulans* Jordan strain 32352" we discovered this bacterium rather belonged within the *Paenibacillus* genus and was reestablished as *Paenibacillus sp.* 32352 (4). As previously reported, this soil-dwelling *Paenibacillus* bacterium is capable of producing a glycoside hydrolase (GH), Pn3Pase, which degrades the linear repeating disaccharide polymer -3)βGlcA(1-4)βGlc(1- that makes up Pn3P (1-3). Recently, we reported further on Pn3Pase identifying and cloning the Pn3Pase gene from *Paenibacillus sp.* 32352, determining optimal activity parameters and characterizing oligosaccharide products from Pn3P degradation (5). Further, we have reported on Pn3Pase potential as a therapeutic agent against type 3 pneumococcal infections (6).

With the reclassification of the *Paenibacillus* bacteria and exploration into the Pn3Pase gene it appeared this glycoside hydrolase did not belong to any existing carbohydrate-active enzyme (CAZy) (7) (http://www.cazy.org) family. Sequence alignments revealed no significant homology to current CAZy GH families over the entire length of a known catalytic domain; however, InterPro analysis suggests local similarities to family GH39 (8). Additionally, Phyre2 analysis suggests sequence similarities to glucuronidases of family GH79 (9). Sequence alignments also suggest Pn3Pase may belong to the GH-A clan and thus may function with a retaining mechanism, as other GH-A clan members. Further, this enzyme has a primary sequence without homologs in the database and it is also functionally unique. To our knowledge,

this is the only enzyme that demonstrates Pn3P hydrolysis activity. Taken together, this information suggested Pn3Pase may establish and belong to a new GH family.

Given that Pn3Pase has potential as a therapeutic agent against the highly infectious serotype 3 *S. pneumoniae*, determining the structure function relationship of this enzyme is crucial to generating Pn3Pase derivatives with optimal activity and pharmacokinetics. Apart from the study exploring Pn3Pase as a therapeutic (6) and the work on elucidating the Pn3Pase gene for recombinant protein expression (5) little work has been done on this enzyme. We do not yet know enzymatic activity sites/domains, mechanism of action, substrate binding specificity, kinetic parameters, or structure.

In continuation of our work characterizing this important enzyme, here we present the detailed biochemical analysis of Pn3Pase. This work led to the establishment of a new GH family designated GHxxx of which Pn3Pase is its founding member. Site directed mutagenesis revealed the nucleophile and acid/base catalytic residues, while C-terminal truncations indicated domains crucial for activity. Binding and kinetic assays revealed this enzyme's activity towards Pn3P and its substrate specificity. NMR analysis demonstrated Pn3Pase utilizes a retaining mechanism similar to other GH-A clan members. Taken together, this work provides the biochemical basis for the Pn3P hydrolysis catalyzed by Pn3Pase. Additionally, this work will aid in developing Pn3Pase as a therapeutic agent against the highly pathogenic type 3 pneumococcus infection.

Material and Methods

Production of recombinant Pn3Pase and mutants

WT recombinant Pn3Pase (DDBJ/ENA/GenBank accession number: MZNT01000000) was produced as previously described (5). Briefly, the coding region of

Pbac_3331 (Pn3Pase) was amplified from *Paencibacillus sp. 32352* genomic DNA into pDONR221 using BP clonase reaction (Thermo Fisher Scientific). Primers for cloning WT and all mutants are listed in Table 2. After transformation into DH5α cells and DNA sequence confirmation, LR clonase reaction (Thermo Fisher Scientific) was performed to insert into pET-DEST42 destination vector for the expression of a carboxy-terminal His_6-tagged fusion protein in *E. coli* BL21(DE3) cells. C-terminal truncation mutants, TM1, TM2 and TM3 were produced by first amplifying DNA in PCR reaction containing 25µl 2x master mix, 2.5µl each of forward and reverse primers, 2.5µl template Pbac. genomic DNA, and brought up to 50µl with water. PCR reaction conditions were 98ºC 120 s, followed by 20 cycles of 98ºC for 15 s, 55ºC for 15s and 72ºC for 6 mins. PCR products were visualized on 1% agarose gel and excised. DNA was extracted from gel using EZNA gel extraction kit (Omega Bio-tek). Using extracted DNA, mutants were then produced following the same BP and LR clonase reactions as WT above. Single and double site mutant proteins were produced using the Quik Change XL Site Directed Mutagenesis Kit (Agilent). PCR reactions contained 5µl 10x buffer, 5µl Pbac_3551 DNA in pDONR221 vector, 2.5µl each of forward and reverse primers, 1µl dNTPs, 3µL QuikSolution, 1µl Pfu Ultra HF, and brought up to 50µl with water. PCR reaction conditions were 95ºC for 60 s followed by 18 cycles of 95ºC for 50 s, 60ºC for 50 s and 68ºC for 9 mins. Following PCR Dpn1 reaction from kit was performed followed by transformation into XL10 competent cells and DNA sequence confirmation. LR clonase reaction was done same as above.

Purification of enzymes

BL21 cells transformed with pET-DEST42-Pn3Pase or mutant plasmid were grown in LB medium supplemented with 100µg/mL ampicillin at 37^{o}C while cell density was monitored at absorbance 600nm. Once OD 600nm reached 1, cells were transferred to 20^{o}C, protein expression was induced by adding Isopropyl β-D-1-thiogalactopyranoside to a final concentration of 1mM and the cell culture was allowed to incubated for 24 hrs. Cells were harvested by centrifugation. Cells were resuspended in phosphate-buffered saline (PBS, pH 7.2) with 10µg/mL DNase and lysed using EmulsiFlex-C5 homogenizer (Avestin). Lysate was cleared by centrifugation at 18,000 xg for 45 mins at 4^{o}C and passed through a 0.45-um filter. Proteins were purified by Ni^{2+}-NTA resin at 4^{o}C and eluted with 300 mM imidazole. Elution was run through FPLC Superdex 200 sizing column (GE LifeSciences) with PBS as running buffer. Protein concentration was determined using NanoDrop (Thermo Fisher Scientific) using extinction coefficients for each protein determined using ExPASy ProtParam tool (10). Purity was assessed by visualizing proteins on stain free tris-glycine gel (BioRad) using gel doc EZ imager (BioRad).

Enzyme Activity Assays

Recombinant Pn3Pase and mutant hydrolysis activity was determined by measuring the increase in reducing ends using *p*-hydroxybenzoic acid hydrazide (PAHBAH) method (11). A reaction mixture (100µl) containing 10µg Pn3P (ATCC 172-X) and 30nM, 60nM or 120nM enzyme in PBS was incubated at 37^{o}C for 30, 60 or 120 minutes then heated at 100^{o}C for 5 minutes to stop reaction. Time 0, used as a negative control, was the same reaction mixtures containing 30nM, 60nM or 120nM of heat killed enzyme. A volume of

40µl of the reaction was mixed with 120µl of 1% (w/v) PAHBAH-HCl solution in duplicate, heated at 100^{o}C for 5 minutes. Absorbance at 410nm was measured on Biotek synergy H1 microplate reader in a clear flat bottom 96-well microplate. Statistical analysis was determined using Student's t test and compared all data points to their corresponding time zero data.

Enzyme Kinetics

The Michaelis-Menten constant (K_M) and the maximum velocity (V_{max}) of WT Pn3Pase was measured using Pn3P as substrate (average molecular weight: 400,000 Da). Pn3P was found to be 400 kDa as determined by size exclusion using the Sephacryl S-300 column (GE Healthcare) using the refractive index of defined mass dextran standards for comparison (12). The substrate was used at seven concentrations for WT (800, 400, 200, 100, 50, 25, and 0 nM) in PBS. WT Pn3Pase was added at 1µg/ml (5.9nM). Reactions were heated at 37^{o}C and aliquots were taken/stopped at 0, 4, 8, 12, 16, and 20 minutes (corresponding to approximately 10% of total hydrolysis yielding mostly tetrasaccharides) by heating reactions at 100^{o}C for 5 minutes. The amount of product formed was measured using PAHBAH-HCl assay as described above with tetrasaccharides, obtained from enzymatic degradation of Pn3P, used to generate a standard curve for data fitting. Initial velocity was determined using the amount of product formed in the linear region of absorbance. Initial velocities of each substrate concentration were curve fitted using non-linear regression with Michaelis-Menten model on GraphPad Prism to determine K_M and V_{max}. k_{cat} was likewise determined using GraphPad Prism as V_{max}/E_T with E_T set as total enzyme concentration 5.9nM. The catalytic efficiency k_{cat}/K_m was determined using these values in $min^{-1}M^{-1}$.

Tetrasaccharides used for standard curve were generated as previously described (5). Standard deviation of data was determined through independent experimental duplicates.

Binding Assay

Binding affinity of recombinant proteins to Pn3P substrate was determined using Biolayer Interferometry (BLI) (FortèBio OctetRED-384). All proteins and Pn3P were suspend in buffer (PBS, 0.5% bovine serum albumin [BSA], 0.05% Tween). Biotinylated Pn3P substrate (20µg/mL) was immobilized on streptavidin biosensor tips (FortèBio) for 150 s after an initial baseline in running buffer for 60 s. Baseline signal was measured again for 60 s before biosensor tips were immersed in wells containing protein (0nm, 100nM, 250nM, or 500nM) for 300 s. Dissociation was measured by returning biosensor tips to baseline for 300 s. Octet data analysis software was used to analyze data. Values of reference wells containing no protein (0nm) were subtracted from data and affinity values were calculated using local and partial fit curve function with a 1:1 binding model. k_{on}, k_{off} and K_D values were determined as the average of the three substrate concentration binding curves. Binding curves were graphed using GraphPad Prism.

NMR for stereochemical analysis

^{1}H NMR analysis was performed on reaction products released from Pn3Pase hydrolysis of Pn3P to determine reaction mechanism. Pn3P was resuspended as 1mg/mL (2.6 mM) in PBS and 10% D_2O and transferred to a 3mm NMR tube. After collecting an initial spectrum at 37^{o}C, the enzyme was added to the tube with a final concentration of 15µg/mL, mixed, and data collection then resumed after 2 minutes. Data were collected continually and four transients were summed every 8 seconds. NMR spectra were acquired on an Agilent 600 MHz DD2 spectrometer equipped with a 3mm cryoprobe, and

used the standard presaturation pulse sequence to reduce the water signal. Data was processed with Mnova software (Mestrelab, Inc.)

Generation of biotinylated Pn3P

Biotinylated Pn3P was prepared using hydrazide biotin. 2 mg of 25 kDa Pn3P was dissolved in 500 µL of 0.1 M sodium borate buffer (pH 5.4). EZ-Link hydrazide-biotin (Thermo Fisher Scientific) (12mg) was dissolved in 100 µL DMSO and added to Pn3P solution. EDC (1-ethyl-3-(3-dimethylaminopropyl) carboiimide HCL) (Thermo Scientific) (1.5mg) was added to the solution, vortexed and incubated before at 25˚C for 3 hours with agitation. Product was purified using Superdex 200 sizing column (GE LifeSciences).

Results

Pn3Pase domain analysis and production

We previously reported on the *Paenibacillus sp. 32352* gene, Pbac_3551, which encodes for the Pn3Pase protein (5). The translated protein sequence of Pn3Pase contains 1545 amino acids and yields a mature protein of 164.1 kDa. InterPro sequence analysis (8) of the full protein revealed putative domains (Figure 1A,B). There was homology between residues 180 and 353 to glycoside hydrolase superfamily, with suggested homology to GH family 39 at the N-terminal region. Other predicted regions of interests were homology to galactose-binding-like superfamily, which for the purpose of this study we took to be the potential carbohydrate binding module (CBM) (Figure 1B), a domain of unknown function 1080 (DUF1080), which has structural similarity to a β1,3-1,4 glucanase in GH family 16, and a concanavalin A-like lectin/glucanase superfamily (Figure 1B). Alignments with existing GHs in the CAZy database revealed short

segments of homology with carbohydrate-active enzymes; however, no significant homology across the full length of Pn3Pase exists. Pn3Pase was run against CAZy HMMs built for GH families (HMMer2,3) which yielded scores insufficient for inclusion in any existing CAZy GH family. The only two borderline hits, GH39 and GH79, showed local similarity that did not extend along the length of the catalytic domains. However, the alignments with GH39 and GH79 showed conservation of the subsequence surrounding the known acid/base residue of clan GH-A (E in GNEPN) and the known nucleophile (E in VSEYGW), with the later only visible in the alignment with GH39 (supplemental data 1) (13). Altogether these low score and incomplete alignments suggested Pn3Pase would establish a new GH family related to clan GH-A, have potential catalytic residues at amino acid sites 196 and 306, and suggested Pn3Pase works through a retaining mechanism. Enzymes belonging to clan GH-A function through retaining mechanisms with their catalytic domain having a $(\beta/\alpha)_8$ barrel fold (14) suggesting Pn3Pase may also have these features.

Moving forward with these preliminary evidences, we performed site directed mutagenesis and deletion studies to confirm the identity of catalytic residues and of domains important for function (Figure 1). Figures 1A and 1B describe the Pn3Pase derivatives and their predicted domains used in this study. WT Pn3Pase and mutant coding sequences were amplified and cloned into pET-DEST42 vector with a C-terminal His_6tag, and expressed in *E. coli*. Recombinant proteins were purified using affinity chromatography followed by size exclusion chromatography and purity was assessed using SDS-PAGE gel (Figure 1C). Bands at appropriate molecular weights (Figure 1A) for each protein were observed (Figure 1C). Oligonucleotides used to generate the

mutants are displayed in Table 1 and site directed mutants were confirmed through sequencing.

Biochemical characterization of Pn3Pase mutants

To assess domains important for function and predicted catalytic residues, we performed activity assays utilizing the *p*-hydroxybenzoic acid hydrazide (PAHBAH) method (11). The method labels the reducing ends of sugars and would reveal if Pn3P hydrolysis was occurring through each Pn3Pase derivative. Unmodified Pn3Pase and its derivatives were first assayed under previously established conditions utilizing 10µg/mL Pn3P substrate in PBS with 30nM enzyme for 30-120 minutes (Figure 2A). As expected WT reached Pn3P hydrolysis saturation within 30 minutes with labeled reducing ends not significantly increasing at 60 or 120 minutes (Figure 2A). Interestingly, TM1 and E196A mutant began displaying significant hydrolysis activity at 30 minutes which continued to increase over the time course. However, no other Pn3Pase mutant showed hydrolase activity even after 4 hours (Figure 2A). To determine if these derivatives truly exhibited no activity against Pn3P we increased the enzyme concentration to 2x and 4x that of the original concentration (Figure 2B,C). Again, no derivatives except TM1 and E196A displayed Pn3P hydrolase activity even at increased enzyme concentrations. TM1 displayed increased levels of activity compared to E196A, which was especially evident at 120nM where TM1 reached hydrolysis levels similar to WT after 120 minutes (Figure 3C).

To continue the biochemical assessment of Pn3Pase activity, we measured Michaelis-Menten kinetic parameters utilizing PAHBAH assay. Since WT was the only enzyme to display significant activity levels at early time points with low enzyme and

substrate levels we only carried on with this protein. Michaelis-Menten parameters of WT Pn3Pase were determined using different concentrations of Pn3P substrate ranging from 0-800 µM with 1µg/mL (5.9nM) enzyme. WT Pn3Pase displayed activity kinetics with a k_{cat} of 1483 min^{-1} and K_M of 0.32 $\pm$ 0.026 µM (Figure 3). The catalytic efficiency defined by the specificity constant k_{cat}/K_M was calculated as $4.6x10^9$ $min^{-1}M^{-1}$.

Taken together, these results suggest that the glutamic acid residues at 196 and 306 positions function as catalytic residues. Further, both E196 and E306 are conserved glutamic acids in Pn3Pase and its close homologs as is the case for clan GH-A glycoside hydrolases (15) making them the acid/base and the nucleophilic residues respectively. Additionally, amino acids and domains after residue 765 do not appear to be crucial for function, while the putative carbohydrate binding module is required for function. While TM1 and E196A are the only Pn3Pase mutants to display function, they hydrolyze at reduced efficiency compared to WT Pn3Pase.

Pn3Pase binding affinity to Pn3P

To determine Pn3Pase binding affinity for its substrate we utilized biolayer interferometry. Pn3P was biotinylated, bound to streptavidin biosensors and Pn3Pase proteins were used as ligand (Figure 4). We used only the non-catalytic Pn3Pase (low catalytic activity for E196A) mutants for this assay as WT and TM1 significantly hydrolyzed the substrate during the course of the experiment and therefore showed little to no binding (data not shown). The site mutants E196A, E306A, and E196/306A had high affinities with K_D (equilibrium dissociation constant) of 418 nM, 451 nM, and 357 nM respectively. TM2 binding curves at 100 nM and 250 nM are both present but overlapped and appeared as a single curve. TM2 had a K_D roughly 10-fold higher than

site mutants at 3.7μM which appears to be due to a slower k_{on} (second-order rate constant of the binding reaction) rate. This result is to be expected as TM2 mutant lacks the putative CBM which has been suggested to play a large role in substrate binding (16). Importantly, this previous study also suggests both the glycoside hydrolase and carbohydrate binding modules are involved in binding elucidating why we observed binding even in the absence of a CBM albeit at a lower affinity than non-catalytic full length mutants (16). In addition, we observed no apparent binding of TM3 to Pn3P substrate (Figure 4). As TM3 is the putative carbohydrate binding module mutant we expected to see binding; however, it has been suggested some CBMs display weak binding interactions as could be the case for our protein (17).

Pn3Pase reaction mechanism

Next, we determined whether Pn3Pase acts through a mechanism that yields either a retention or inversion of the stereochemistry at the GlcA anomeric reducing end (Figure 5). We used ^{1}H NMR spectroscopy to monitor the hydrolysis of the β-GlcA(1-4)β-Glc-glycosidic linkage in Pn3P by Pn3Pase over time. Chemical shifts of Pn3P were based on previously published data (5). Since the β-H1 signal of free reducing GlcA (4.60 ppm) is obscured by the overlapping water signal at 37^{o} C and further reduced by the water suppression NMR pulse program, the GlcA H5 signals were used to monitor the reaction (Figure 5). The Pn3P starting material shows the glycosidically linked β-GlcA H5 signal at 3.9 ppm (bottom trace, Figure 5). Immediately after the addition of enzyme a signal corresponding to the H5 of the free reducing β-anomer of GlcA (3.85 ppm) is produced and continues to grow over time (Figure 5). Only after around 10 minutes, does a small signal appear corresponding to the H5 of the free reducing α-GlcA (4.04 ppm) (Figure 5,

top trace), which is due to mutarotation. This experiment confirms that Pn3Pase functions through a retaining mechanism and supports an activity mechanism that functions through a nucleophile and an acid/base catalytic residues (15) as other clan GH-A members further suggesting Pn3Pase may belong to this GH clan. Taken together these results demonstrate that Pn3Pase acts via a retaining mechanism to hydrolyze Pn3P.

Discussion

The capsular polysaccharides (CPS) of many pathogenic bacteria are prominent features serving many functions such as assisting in adhesion and colonization and inhibition of opsonophagocytosis rendering them major virulence factors (18,19). Unencapsulated mutants often fail to colonize and rarely cause infections due to efficient opsonophagocytotic clearance by host cells (18,20). Among *S. pneumoniae* species, serotype 3 is one of the most virulent (21-23) and continues to be of concern despite its inclusion in the current conjugate vaccine against pneumococcal infections (24-27). Armed with this knowledge, the benefits of utilizing an alternative therapy against type 3 pneumococcal infection becomes apparent. Indeed, we have begun investigations into utilizing Pn3Pase, described in this study, as a therapeutic agent against highly virulent type 3 infections (6). Pn3Pase's unique ability to hydrolyze the CPS of type 3 *S. pneumoniae* makes it a potential therapeutic agent to reduce bacterial colonization and infection. Information from this current study will aid in further developing Pn3Pase as an alternative to conventional treatments such as antibiotics, establishing the biochemical characteristics necessary for optimum therapeutic potential. Additionally, recent work on a mucin-selective protease StcE (28) suggests a possible role for the non-catalytic Pn3Pase mutants in purifying capsule. Type 3 is among the pneumococcal serotypes that

has exhibited capsule shedding (29). Non-active mutants of Pn3Pase may be useful in isolating shed capsule *in vivo* to assess capsule shedding in bacterial virulence as well as for serotyping in clinics.

In this current work we observed unique binding characteristics in our truncation mutants compared to non-catalytic full length proteins. TM3 mutant did not exhibit any significant binding to Pn3P substrate as well as having abolished activity. TM3 being non-catalytic was expected (30). It was previously noted that substrate binding for CBMs varies in a wide range with some having relatively weak interactions (17) as could be the case in Pn3Pase's putative CBM. TM2 recorded binding to Pn3P with approximately 10-fold lower affinity than full length proteins. An important role of the CBM is to target the substrate and bring it into closer proximity with the enzyme (31) which increases hydrolysis of the substrate (32). TM2 lacks the putative CBM, which paired with the published work above may explain the lower binding affinity and no enzymatic activity. With no CBM, TM2 does not appear to bring Pn3P into close proximity with the hydrolase domain nor bind the substrate long enough for hydrolysis to occur. Several studies have suggested that removal of the CBM dramatically decreases enzymatic activity (32-34).

TM1 and E196A were the only mutants to display catalytic activity. We were unable to observe substrate binding with WT or TM1 due to degradation activity. Compared to WT, TM1 and E196A have significantly lower catalytic activities. In TM1 it is possible this could be due to structure and removal of the C terminal domains. InterPro analysis suggests full length Pn3Pase has two C-terminal domains that may possess carbohydrate-binding-like properties, the DUF1080 and a ConA-like/glucanase

superfamily domain. Reviews have suggested proteins that possess hydrolytic activity, i.e. glycoside hydrolases, can have one or more CBMs (32). If this is the case for Pn3Pase, TM1, being half the size of WT, could have lower catalytic activity due to removal of these other potential CBM-like domains. Additionally, the C terminal domains could be important for overall tertiary structure which could aid in activity. TM1 mutant does however demonstrate the minimum structure requirements for Pn3Pase activity, the glycoside hydrolase domain and putative CBM. All other C terminal domains do not appear to be essential for hydrolase activity. Our results also suggest the correct assignment of E196 as the acid/base catalytic residue and E306 as the nucleophile. This conclusion is apparent by E196A drastically decreased activity compared to WT and E306A complete loss of function. Prior work on GHs have shown that mutating these catalytic residues does not always lead to a complete ablation of activity (35). With catalysis taking place at the glucuronic acid residue bond, it is possible that the carboxylate of GlcA may be able to act as the acid/base residue in the E196A mutant thereby leading to partial activity rescue (Figure 6). Further, members of clan GH-A have conserved glutamic acid residues functioning as the acid/base and nucleophile catalytic residues as appears to be the case with Pn3Pase (13)

A Blast search using Pn3Pase sequence against the non-redundant protein sequence database of the NCBI shows that a small number of close homologs exist, all of which are essentially found in *Paenibacillus* sequences. Proteins with more distant similarity were also found (supplemental data 2), but their distance to the only characterized GHxxx member (Pn3Pase) is such that it is appropriate to place them into the non-classified CAZy section of the glycoside hydrolases until a sufficient taxonomical diversity is

captured in family GHxxx. The reason behind the small size of the family is unclear; perhaps it simply means that degradation of Pn3P is not widespread and that the more distant relatives of Pn3Pase (those outside the red box in the supplemental 2) may act on a different substrate.

Taken together, the work in this current study establishes the initial biochemical analysis of the glycoside hydrolase Pn3Pase which led to the establishment of GH family GHxxx. Despite our best efforts we have yet to obtain a crystal structure of this enzyme. Due to the large sequence divergence between Pn3Pase and other clan GH-A members we were also unable to build a homology model that had meaningful features. However, similarity with clan GH-A predicts Pn3Pase would have a β/α_8 barrel fold (TIM barrel), but the exact structural details that govern substrate recognition cannot be reliably predicted. Continuing work with Pn3Pase will involve a crystal structure to complement the work in this study and shed more light into the structure-function relationship of Pn3Pase with Pn3P. Additionally, further structural analysis will reveal if the correlations between activity and domains observed in this study are accurate and confirm the assignment of residues E196 and E306 as the acid/base and nucleophile. Further structural investigations will also aid in developing Pn3Pase as a therapeutic against type 3 pneumococcal infections.

References

1. Avery, O. T., and Dubos, R. (1930) THE SPECIFIC ACTION OF A BACTERIAL ENZYME ON PNEUMOCOCCI OF TYPE III. *Science* **72**, 151-152
2. Dubos, R., and Avery, O. T. (1931) DECOMPOSITION OF THE CAPSULAR POLYSACCHARIDE OF PNEUMOCOCCUS TYPE III BY A BACTERIAL ENZYME. *J Exp Med* **54**, 51-71
3. Sickles, G. M., and Shaw, M. (1934) A Systematic Study of Microörganisms Which Decompose the Specific Carbohydrates of the Pneumococcus. *J Bacteriol* **28**, 415-431
4. Middleton, D. R., Lorenz, W., and Avci, F. Y. (2017) Complete Genome Sequence of the Bacterium. *Genome Announc* **5**
5. Middleton, D. R., Zhang, X., Wantuch, P. L., Ozdilek, A., Liu, X., LoPilato, R., Gangasani, N., Bridger, R., Wells, L., Linhardt, R. J., and Avci, F. Y. (2018) Identification and characterization of the Streptococcus pneumoniae type 3 capsule-specific glycoside hydrolase of Paenibacillus species 32352. *Glycobiology* **28**, 90-99
6. Middleton, D. R., Paschall, A. V., Duke, J. A., and Avci, F. Y. (2018) Enzymatic Hydrolysis of Pneumococcal Capsular Polysaccharide Renders the Bacterium Vulnerable to Host Defense. *Infect Immun*

7. Lombard, V., Golaconda Ramulu, H., Drula, E., Coutinho, P. M., and Henrissat, B. (2014) The carbohydrate-active enzymes database (CAZy) in 2013. *Nucleic Acids Res* **42**, D490-495

8. Mitchell, A. L., Attwood, T. K., Babbitt, P. C., Blum, M., Bork, P., Bridge, A., Brown, S. D., Chang, H. Y., El-Gebali, S., Fraser, M. I., Gough, J., Haft, D. R., Huang, H., Letunic, I., Lopez, R., Luciani, A., Madeira, F., Marchler-Bauer, A., Mi, H., Natale, D. A., Necci, M., Nuka, G., Orengo, C., Pandurangan, A. P., Paysan-Lafosse, T., Pesseat, S., Potter, S. C., Qureshi, M. A., Rawlings, N. D., Redaschi, N., Richardson, L. J., Rivoire, C., Salazar, G. A., Sangrador-Vegas, A., Sigrist, C. J. A., Sillitoe, I., Sutton, G. G., Thanki, N., Thomas, P. D., Tosatto, S. C. E., Yong, S. Y., and Finn, R. D. (2019) InterPro in 2019: improving coverage, classification and access to protein sequence annotations. *Nucleic Acids Res* **47**, D351-D360

9. Kelley, L. A., Mezulis, S., Yates, C. M., Wass, M. N., and Sternberg, M. J. (2015) The Phyre2 web portal for protein modeling, prediction and analysis. *Nat Protoc* **10**, 845-858

10. Gasteiger, E., Gattiker, A., Hoogland, C., Ivanyi, I., Appel, R. D., and Bairoch, A. (2003) ExPASy: The proteomics server for in-depth protein knowledge and analysis. *Nucleic Acids Res* **31**, 3784-3788

11. Blakeney, A. B., and Mutton, L. L. (1980) A simple colorimetric method for the determination of sugars in fruit and vegetables. *J Sci Food Agric* **31**, 889-897

12. Li, G., Li, L., Xue, C., Middleton, D., Linhardt, R. J., and Avci, F. Y. (2015) Profiling pneumococcal type 3-derived oligosaccharides by high resolution liquid chromatography-tandem mass spectrometry. *J Chromatogr A* **1397**, 43-51

13. Henrissat, B., Callebaut, I., Fabrega, S., Lehn, P., Mornon, J. P., and Davies, G. (1996) Conserved catalytic machinery and the prediction of a common fold for several families of glycosyl hydrolases. *Proc Natl Acad Sci U S A* **93**, 5674

14. Henrissat, B., and Bairoch, A. (1996) Updating the sequence-based classification of glycosyl hydrolases. *Biochem J* **316 (Pt 2)**, 695-696

15. Davies, G., and Henrissat, B. (1995) Structures and mechanisms of glycosyl hydrolases. *Structure* **3**, 853-859

16. Klontz, E. H., Trastoy, B., Deredge, D., Fields, J. K., Li, C., Orwenyo, J., Marina, A., Beadenkopf, R., Günther, S., Flores, J., Wintrode, P. L., Wang, L. X., Guerin, M. E., and Sundberg, E. J. (2019) Molecular Basis of Broad Spectrum N-Glycan Specificity and Processing of Therapeutic IgG Monoclonal Antibodies by Endoglycosidase S2. *ACS Cent Sci* **5**, 524-538

17. Volkov, I. I. u., Lunina, N. A., and Velikodvorskaia, G. A. (2004) Prospects for practical application of substrate-binding modules of glycosyl hydrolases. *Prikl Biokhim Mikrobiol* **40**, 499-504

18. Magee, A. D., and Yother, J. (2001) Requirement for capsule in colonization by Streptococcus pneumoniae. *Infect Immun* **69**, 3755-3761

19. Moxon, E. R., and Kroll, J. S. (1990) The role of bacterial polysaccharide capsules as virulence factors. *Curr Top Microbiol Immunol* **150**, 65-85

20. Nelson, A. L., Roche, A. M., Gould, J. M., Chim, K., Ratner, A. J., and Weiser, J. N. (2007) Capsule enhances pneumococcal colonization by limiting mucus-mediated clearance. *Infect Immun* **75**, 83-90

21. Briles, D. E., Crain, M. J., Gray, B. M., Forman, C., and Yother, J. (1992) Strong association between capsular type and virulence for mice among human isolates of Streptococcus pneumoniae. *Infect Immun* **60**, 111-116

22. Weinberger, D. M., Harboe, Z. B., Sanders, E. A., Ndiritu, M., Klugman, K. P., Rückinger, S., Dagan, R., Adegbola, R., Cutts, F., Johnson, H. L., O'Brien, K. L., Scott, J. A., and Lipsitch, M. (2010) Association of serotype with risk of death due to pneumococcal pneumonia: a meta-analysis. *Clin Infect Dis* **51**, 692-699

23. Martens, P., Worm, S. W., Lundgren, B., Konradsen, H. B., and Benfield, T. (2004) Serotype-specific mortality from invasive Streptococcus pneumoniae disease revisited. *BMC Infect Dis* **4**, 21

24. Wantuch, P. L., and Avci, F. Y. (2018) Current status and future directions of invasive pneumococcal diseases and prophylactic approaches to control them. *Hum Vaccin Immunother* **14**, 2303-2309

25. Wantuch, P. L., and Avci, F. Y. (2019) Invasive pneumococcal disease in relation to vaccine type serotypes. *Hum Vaccin Immunother*, 1-2

26. Richter, S. S., Heilmann, K. P., Dohrn, C. L., Riahi, F., Diekema, D. J., and Doern, G. V. (2013) Pneumococcal serotypes before and after introduction of conjugate vaccines, United States, 1999-2011(1.). *Emerg Infect Dis* **19**, 1074-1083

27. Gruber, W. C., Scott, D. A., and Emini, E. A. (2012) Development and clinical evaluation of Prevnar 13, a 13-valent pneumococcal CRM197 conjugate vaccine. *Ann N Y Acad Sci* **1263**, 15-26

28. Malaker, S. A., Pedram, K., Ferracane, M. J., Bensing, B. A., Krishnan, V., Pett, C., Yu, J., Woods, E. C., Kramer, J. R., Westerlind, U., Dorigo, O., and Bertozzi, C. R. (2019) The mucin-selective protease StcE enables molecular and functional analysis of human cancer-associated mucins. *Proc Natl Acad Sci U S A* **116**, 7278-7287

29. Kietzman, C. C., Gao, G., Mann, B., Myers, L., and Tuomanen, E. I. (2016) Dynamic capsule restructuring by the main pneumococcal autolysin LytA in response to the epithelium. *Nat Commun* **7**, 10859

30. Boraston, A. B., Bolam, D. N., Gilbert, H. J., and Davies, G. J. (2004) Carbohydrate-binding modules: fine-tuning polysaccharide recognition. *Biochem J* **382**, 769-781

31. Hervé, C., Rogowski, A., Blake, A. W., Marcus, S. E., Gilbert, H. J., and Knox, J. P. (2010) Carbohydrate-binding modules promote the enzymatic deconstruction of intact plant cell walls by targeting and proximity effects. *Proc Natl Acad Sci U S A* **107**, 15293-15298

32. Shoseyov, O., Shani, Z., and Levy, I. (2006) Carbohydrate binding modules: biochemical properties and novel applications. *Microbiol Mol Biol Rev* **70**, 283-295

33. Carrard, G., and Linder, M. (1999) Widely different off rates of two closely related cellulose-binding domains from Trichoderma reesei. *Eur J Biochem* **262**, 637-643

34. Coutinho, J. B., Gilkes, N. R., Kilburn, D. G., Warren, R. A. J., and Miller Jr, R. C. (1993) The nature of the cellulose-binding domain effects the activities of a bacterial endoglucanase on different forms of cellulose. **113 (2)**, 211-217

35. Armstrong, Z., and Davies, G. J. (2020) Structure and function of Bs164 β-mannosidase from Bacteroides salyersiae the founding member of glycoside hydrolase family GH164. *J Biol Chem* **295**, 4316-4326

Figure 6.1

A

Mutant	Domain Boundary	MW	Ext. Coeff.
WT	Full length	168.24 kD	221750
TM1	1-765	85.56 kD	152070
TM2	1-500	56.76 kD	109670
E196A	Predicted acid catalytic residue E: GAG- A: GCG	168.18 kD	221750
E306A	Predicted nucleophilic catalytic residue E: GAA- A: GCA	168.18 kD	221750
E196/306A	Double site mutant of predicted active sites	168.13 kD	221750
TM3	562-765	22.27 kD	22920

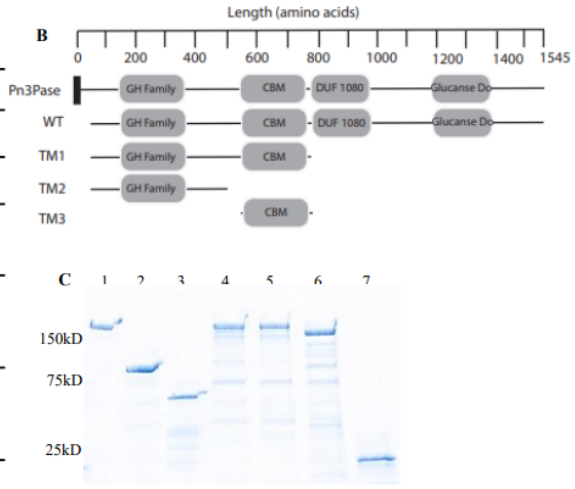


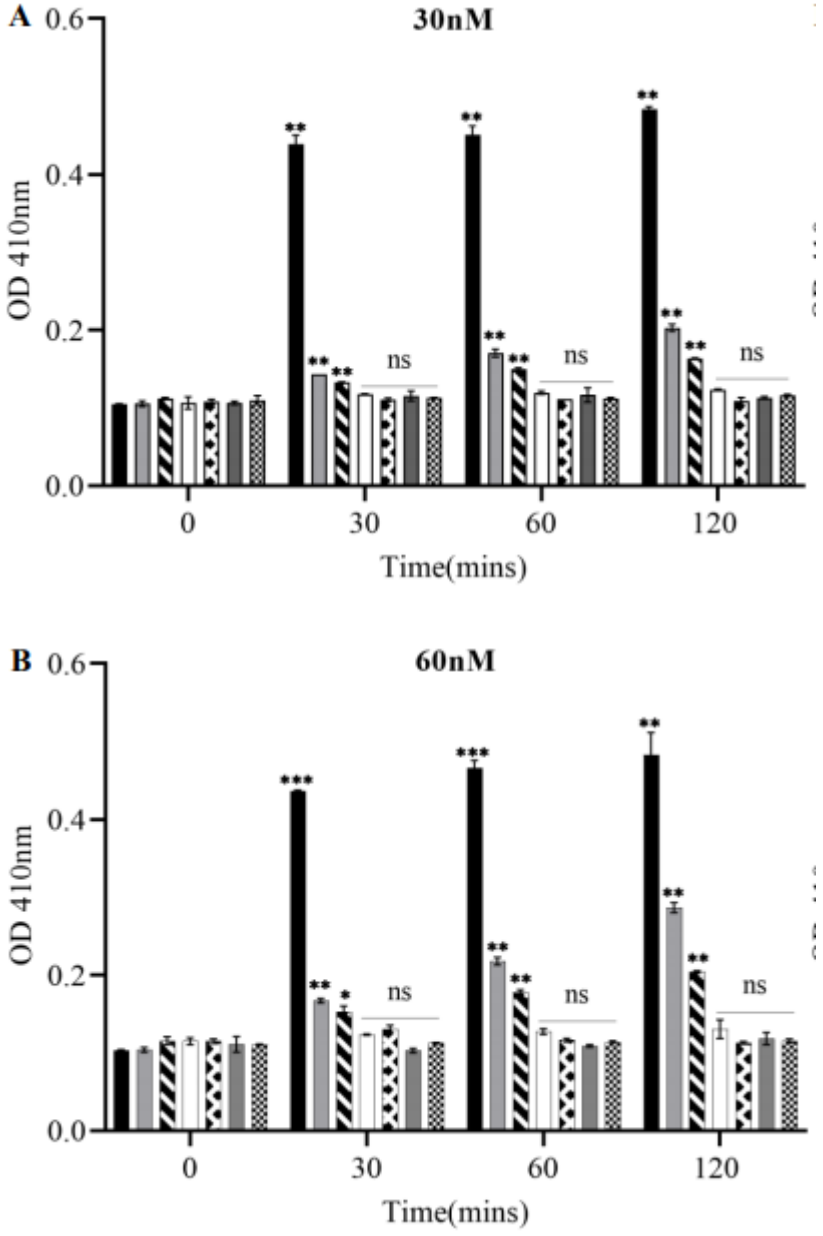
Figure 6.1: Characterization of Pn3Pase mutants. **A)** Description of Pn3Pase and derivatives used in this study. Extinction coefficients were determined using ExPASy ProtParam tool, reference 33. **B)** Schematic of predicted domains (InterPro). Pn3Pase is full length native protein, WT is recombinant Pn3Pase, TM1,2,3 are truncation mutants (TM). CBM, carrepresentationbohydrate binding module homologous to galactose binding domains. Signal peptide is represented by black rectangle. **C)** Stain free SDS-PAGE gel of Pn3Pase mutants: 1) WT 2) TM1 3) TM2 4) E196A 5) E306A 6) E196/306A 7) TM3

Table 6.1

Table 6.1: Oligonucleotides for each protein used in this study

Oligonucleotide	**Sequence 5'-3'**
WT_F	ggggacaagtttgtacaaaaaagcaggcttcgaaggagatagaaccatggcacccgtgaatctggaagc
WT_R	ggggaccactttgtacaagaaagctgggtgctccacgatcaccttattcgataacg
TM1_R	ggggaccactttgtacaagaaagctgggtgttcggcaaaaacctttacatcg
TM2_R	ggggaccactttgtacaagaaagctgggtgggtgtagaaagtgcggcc
E196A_F	gggcaacgcgccgaacc
E196A_R	ggttcggcgcgttgccc
E306A_F	tgttagcgcatacggctggaag
E306A_R	cttccagccgtatgcgctaaca
TM3_F	ggggacaagtttgtacaaaaaagcaggcttcgaaggagatagaaccatgaatatcgtgtctgccggc
TM3_R	ggggaccactttgtacaagaaagctgggtgttcggcaaaaacctttacatcg

Figure 6.2


A
30nM
OD 410nm
0.6
0.4
0.2
0.0
0
30
60
120
Time(mins)
ns
B
60nM
OD 410nm
0.6
0.4
0.2
0.0
0
30
60
120
Time(mins)
ns

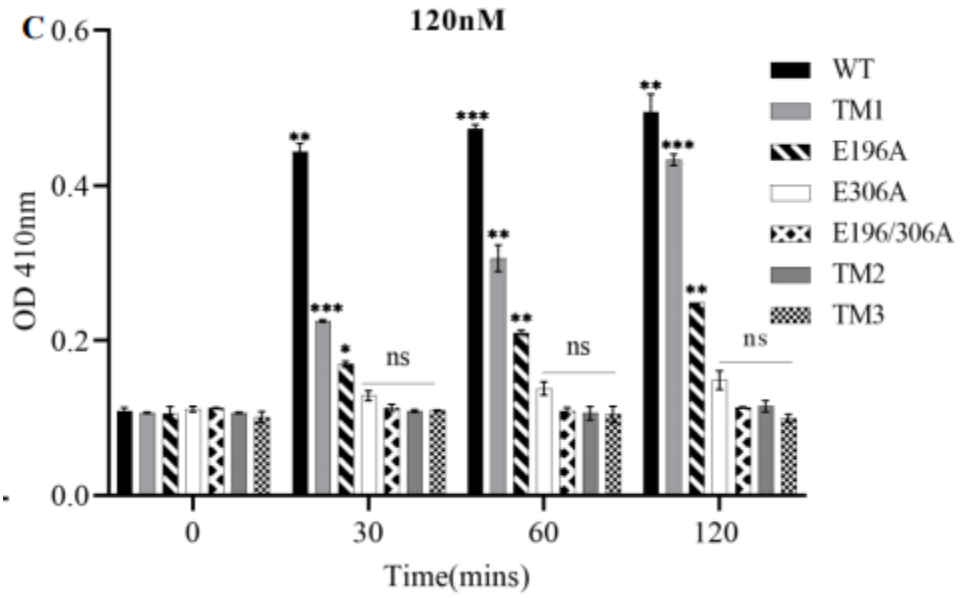


Figure 6.2: Activity characteristics of Pn3Pase mutants. PAHBAH assay was used to determine substrate degradation. All samples had .1mg/ml Pn3P substrate and A) 30nM B) 60nM C) 120nM enzyme incubated for 0, 30, 60 and 120 minutes. Legend for all graphs is the same and to the right of C. Significance for each data is determined using Student's t test comparing data to corresponding time zero data.

Figure 6.3

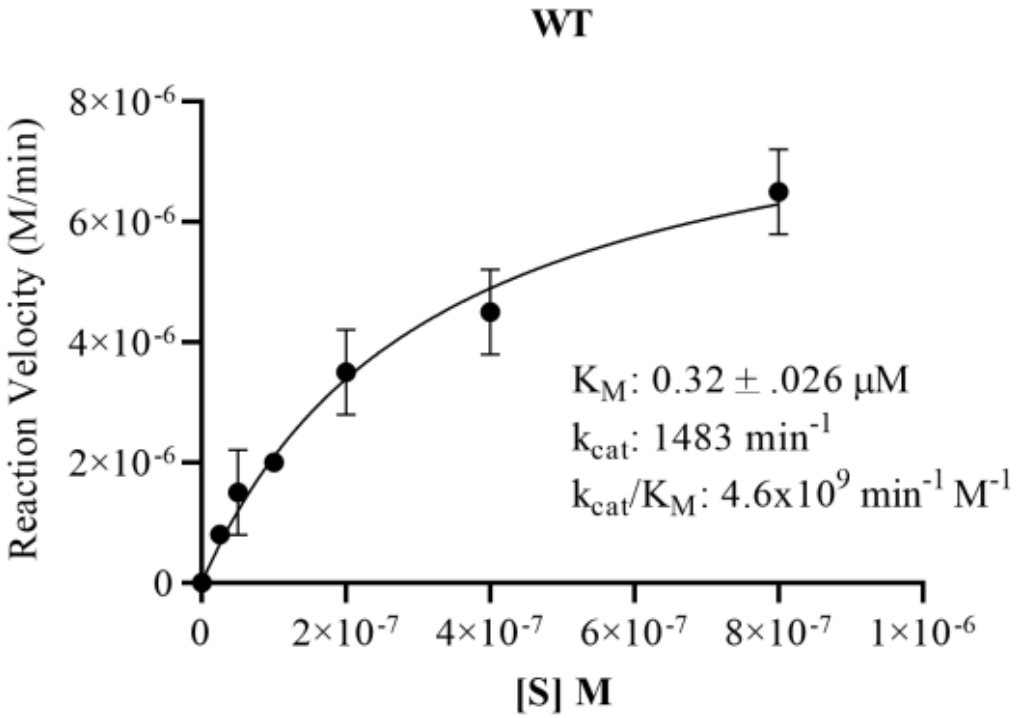


Figure 6.3: Substrate saturation curve for WT Pn3Pase. Michaelis-Menten kinetics were determined plotting substrate concentration against initial velocities for WT Pn3Pase. For kinetics experiments WT received 1µg/ml (5.97nM) enzyme respectively. Substrate concentrations were 800, 400, 200, 100, 50, 25 and 0 nM. Initial velocities were determined as the slope of the line in amount of product formed between 0 and 8 minutes. Amount product formed was determined using tetrasaccharide as standards. Kinetic parameters were determined using nonlinear regression (GraphPad Prism). Standard deviation of data was determined through independent duplicates.

Figure 6.4

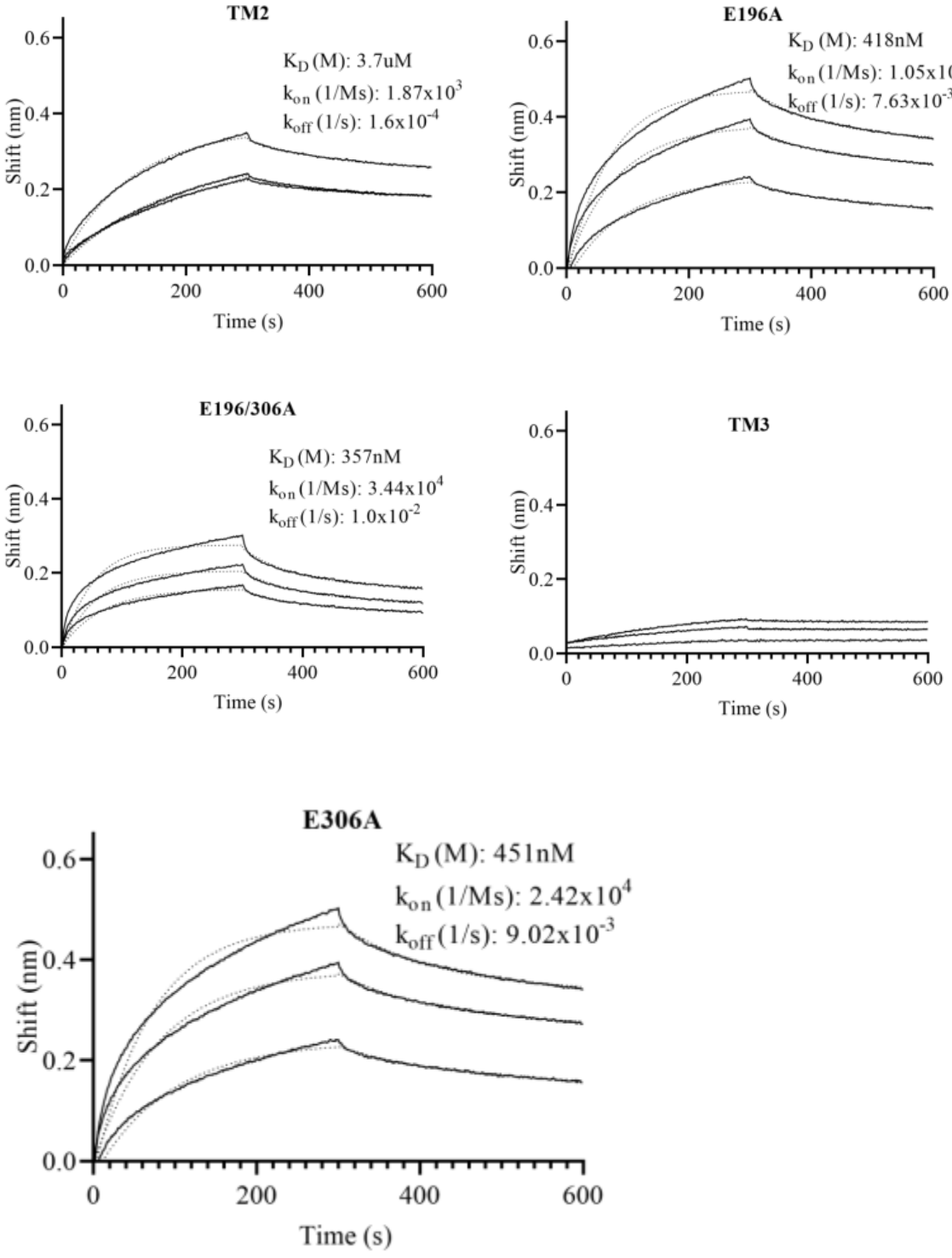

TM2
K_D (M): 3.7uM
k_{on} (1/Ms): 1.87x10^3
k_{off} (1/s): 1.6x10^{-4}
E196A
K_D (M): 418nM
k_{on} (1/Ms): 1.05x10^4
k_{off} (1/s): 7.63x10^{-3}
E196/306A
K_D (M): 357nM
k_{on} (1/Ms): 3.44x10^4
k_{off} (1/s): 1.0x10^{-2}
TM3
E306A
K_D (M): 451nM
k_{on} (1/Ms): 2.42x10^4
k_{off} (1/s): 9.02x10^{-3}
Shift (nm)
Time (s)
0.0
0.2
0.4
0.6
0
200
400
600

Figure 6.4: Binding characterization of Pn3Pase and mutants to Pn3P. Binding rates were determined using Biolayer interferometry (BLI). Biotinylated Pn3 polysaccharide (20µg/mL) was attached to streptavidin coated sensor for 150 seconds then returned to baseline buffer. Sensor was then exposed to Pn3Pase or mutants (100nm, 250nm and 500nm) for 300 seconds. Dissociation was measured over 300 seconds to determine K_D. K_D, k_{on} and k_{off} rates were determined as the average of the three binding curves. All three binding curves are present for TM2 mutant, 100nm and 250nm curves are overlapping. Fits are shown as dotted lines.

Figure 6.5

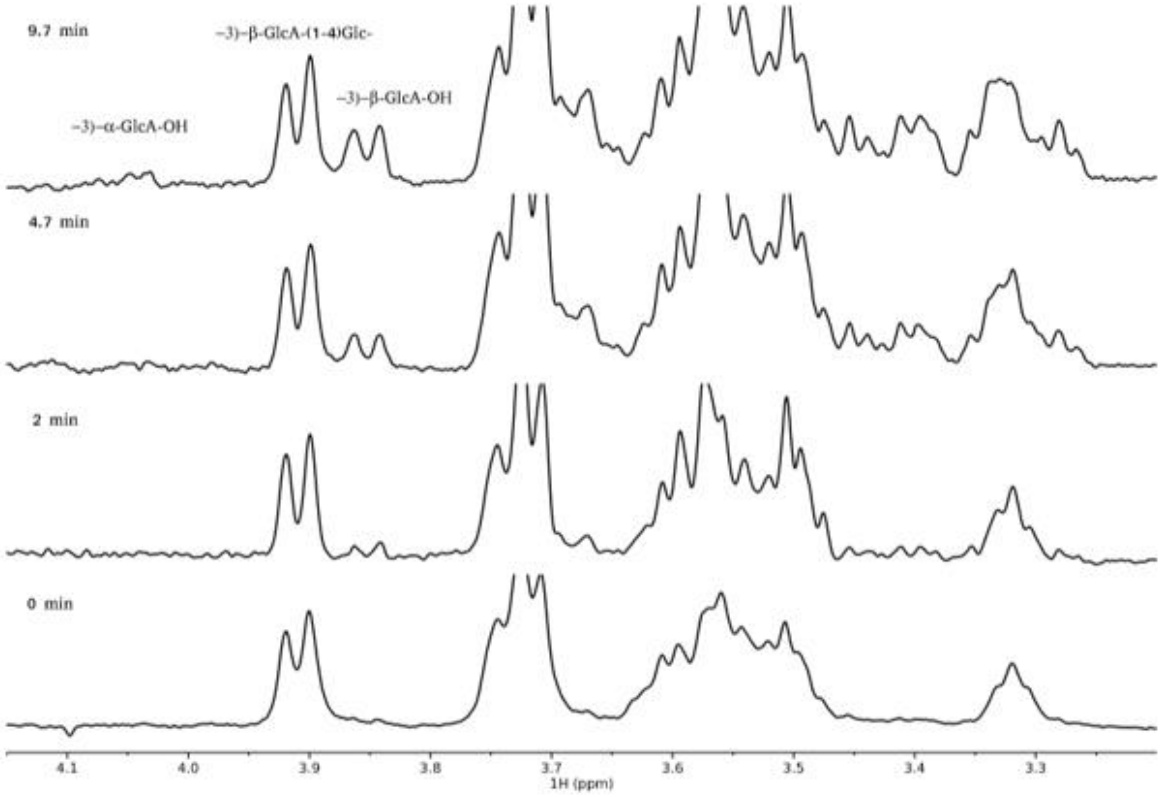


Figure 6.5: Reaction mechanism of Pn3Pase. Region of the 600 MHz proton spectrum of Pn3P, [-3)-β-D-GlcA(1-4)-β-D-Glc(1-]$_n$ in PBS buffer with 10% D2O at 37ºC identifying doublet signals from GlcA-H5. Bottom trace at time zero, additional spectra at times indicated.

Figure 6.6

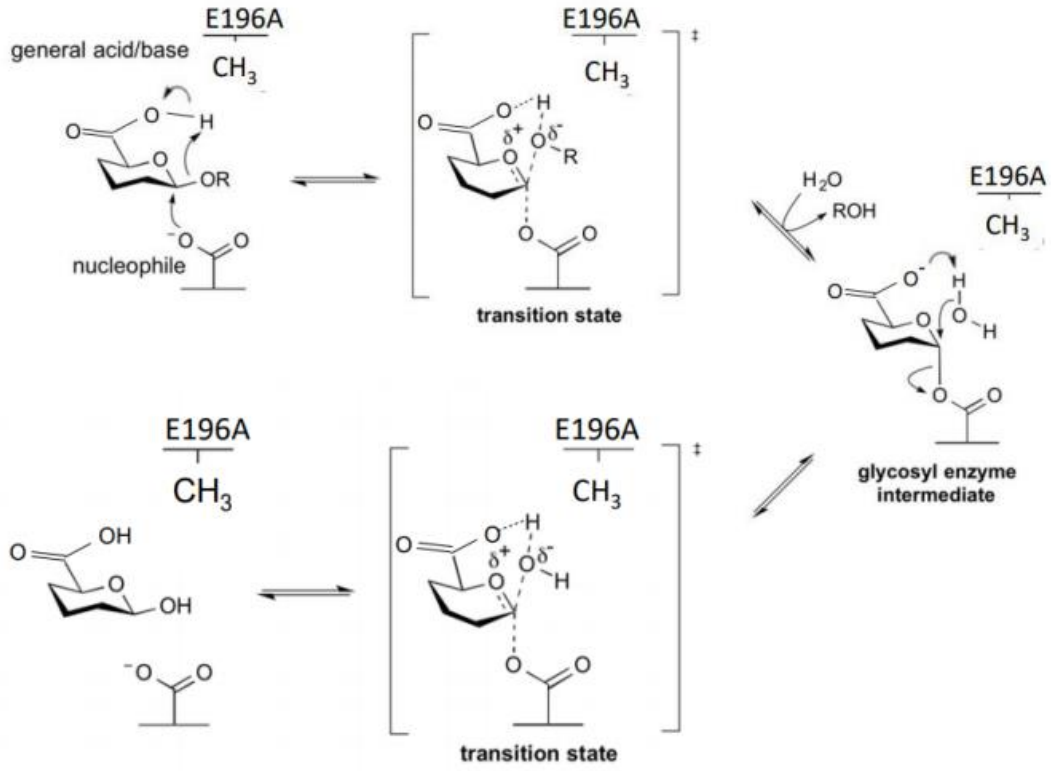
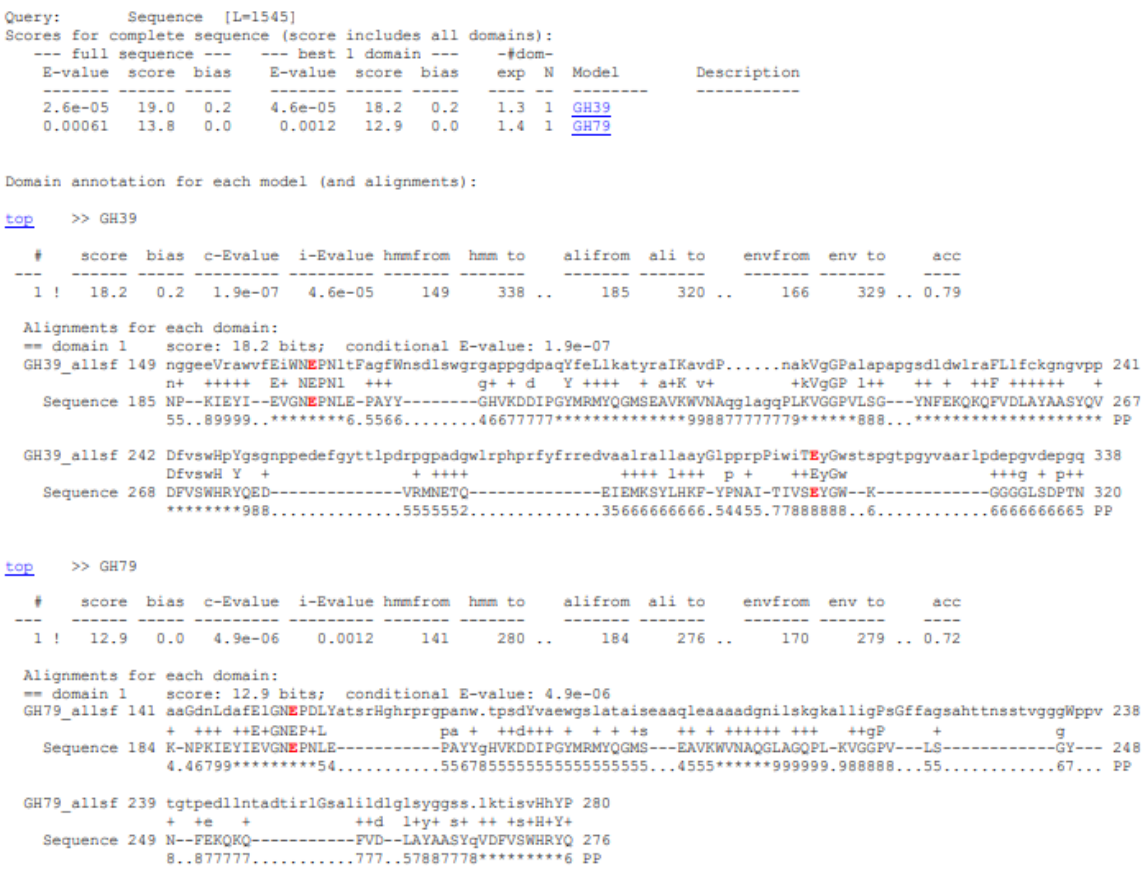


Figure 6.6: Proposed reaction mechanism for Pn3Pase and its E196A mutant (adapted from CAZypedia's reaction mechanism for a retaining b-glycoside hydrolase). Pn3Pase operates with a double displacement mechanism which involves a glucuronyl-enzyme intermediate. The second step is the nucleophile attack of the glucuronyl-enzyme intermediate by a water residue, which is normally activated by E196. In the E196A mutant the water is likely activated, albeit less efficiently, by the carboxylate of the GlcA accounting for the small retention of activity when the acid//base catalytic residue is mutated.

Figure 6.S1



```
Query:       Sequence  [L=1545]
Scores for complete sequence (score includes all domains):
   --- full sequence ---    --- best 1 domain ---     -#dom-
    E-value  score  bias     E-value  score  bias     exp  N  Model        Description
    -------  ------ -----    ------- ------ -----    ---- --  --------     -----------
    2.6e-05   19.0   0.2     4.6e-05   18.2   0.2     1.3  1  GH39
    0.00061   13.8   0.0      0.0012   12.9   0.0     1.4  1  GH79


Domain annotation for each model (and alignments):

top     >> GH39

   #    score  bias  c-Evalue  i-Evalue hmmfrom  hmm to    alifrom  ali to    envfrom  env to     acc
 ---   ------ ----- --------- --------- ------- -------    ------- -------    ------- -------    ----
   1 !   18.2   0.2   1.9e-07   4.6e-05     149     338 ..     185     320 ..     166     329 .. 0.79

  Alignments for each domain:
  == domain 1    score: 18.2 bits;  conditional E-value: 1.9e-07
  GH39_allsf 149 nggeeVrawvfEiWNEPNltFagfWnsdlswgrgappgdpaqYfeLlkatyraIKavdP......nakVgGPalapapgsdldwlraFLlfckgngvpp 241
                 n+  +++++  E+ NEPNl  +++          g+ + d   Y ++++  + a+K v+        +kVgGP l++   ++ +  ++F ++++++   +
    Sequence 185 NP--KIEYI--EVGNEPNLE-PAYY--------GHVKDDIPGYMRMYQGMSEAVKWVNAqglagqPLKVGGPVLSG---YNFEKQKQFVDLAYAASYQV 267
                 55..89999..********6.5566........46677777**************998877777779******888...********************** PP

  GH39_allsf 242 DfvswHpYgsgnppedefgyttlpdrpgpadgwlrphprfyfrredvaalrallaayGlpprpPiwiTEyGwstspgtpgyvaarlpdepgvdepgq 338
                 DfvswH Y  +               + ++++                ++++ l+++  p +    ++EyGw              +++g + p++
    Sequence 268 DFVSWHRYQED--------------VRMNETQ--------------EIEMKSYLHKF-YPNAI-TIVSEYGW--K------------GGGGLSDPTN 320
                 ********988..............5555552..............35666666666.54455.77888888..6............6666666665 PP

top     >> GH79

   #    score  bias  c-Evalue  i-Evalue hmmfrom  hmm to    alifrom  ali to    envfrom  env to     acc
 ---   ------ ----- --------- --------- ------- -------    ------- -------    ------- -------    ----
   1 !   12.9   0.0   4.9e-06    0.0012     141     280 ..     184     276 ..     170     279 .. 0.72

  Alignments for each domain:
  == domain 1    score: 12.9 bits;  conditional E-value: 4.9e-06
  GH79_allsf 141 aaGdnLdafElGNEPDLYatsrHghrprgpanw.tpsdYvaewgslataiseaaqleaaaadgnilskgkalligPsGffagsahttnsstvgggWppv 238
                 +  +++ ++E+GNEP+L            pa +  ++d+++ +  + + +s   ++ + ++++++ +++   ++gP     +            g
    Sequence 184 K-NPKIEYIEVGNEPNLE-----------PAYYgHVKDDIPGYMRMYQGMS---EAVKWVNAQGLAGQPL-KVGGPV---LS-----------GY--- 248
                 4.46799*********54...........5567855555555555555555...4555******999999.988888...55.....67... PP

  GH79_allsf 239 tgtpedllntadtirlGsalildlglsyggss.lktisvHhYP 280
                 +  +e   +           ++d  l+y+ s+ ++ +s+H+Y+
    Sequence 249 N--FEKQKQ-----------FVD--LAYAASYqVDFVSWHRYQ 276
                 8..877777...........777..57887778*********6 PP
```

Figure 6.S1: HMM search results against all HMMs for glycoside hydrolase families.

The catalytic residues are in red.

Figure 6.S2

Sequences producing significant alignments:

Select: All None Selected:0

Alignments Download GenPept Graphics Distance tree of results Multiple alignment

Description	Max Score	Total Score	Query Cover	E value	Per. Ident	Accession
DUF5011 domain-containing protein [Paenibacillus sp. 32352]	806	806	100%	0.0	100.00%	WP_079915027.1
discoidin domain-containing protein [Paenibacillus sp. URHA0014]	728	728	100%	0.0	88.95%	WP_028553222.1
discoidin domain-containing protein [Paenibacillus whitsoniae]	727	727	100%	0.0	88.68%	WP_126144750.1
discoidin domain-containing protein [Paenibacillus sp. Soil766]	726	726	100%	0.0	88.68%	WP_057315858.1
discoidin domain-containing protein [Paenibacillus sp. Soil750]	724	724	99%	0.0	88.65%	WP_056617367.1
discoidin domain-containing protein [Paenibacillus oryzisoli]	722	722	100%	0.0	88.16%	WP_068663733.1
hypothetical protein BK126_18045 [Paenibacillus sp. FSL H7-0326]	583	583	100%	0.0	71.65%	OMC67485.1
discoidin domain-containing protein [Paenibacillus sp. FSL H7-0326]	583	583	100%	0.0	71.65%	WP_083655951.1
DUF5011 domain-containing protein [Paenibacillus sp. YN15]	582	582	98%	0.0	73.70%	WP_113018908.1
DUF5011 domain-containing protein [Paenibacillus sp. YN15]	224	224	90%	2e-61	38.27%	WP_113018909.1
discoidin domain-containing protein [Vallitalea guaymasensis]	121	121	94%	5e-26	27.30%	WP_113675418.1
hypothetical protein [Streptomyces griseus]	74.7	74.7	85%	1e-10	25.62%	WP_037672359.1
hypothetical protein [Micromonospora echinospora]	73.6	73.6	48%	3e-10	30.41%	WP_145834095.1
hypothetical protein [Paenibacillus aceris]	73.6	73.6	82%	4e-10	25.66%	NHW38917.1
hypothetical protein [Micromonospora echinospora]	72.8	72.8	48%	5e-10	30.41%	WP_088980648.1
hypothetical protein C6376_36625 [Streptomyces sp. P3]	72.8	72.8	85%	6e-10	25.62%	AVV46075.1
hypothetical protein [Streptomyces sp. P3]	72.4	72.4	85%	6e-10	25.62%	WP_159083366.1
hypothetical protein [Streptomyces sp. SLBN-192]	70.5	70.5	82%	3e-09	25.35%	WP_142165674.1
hypothetical protein GA0074692_4185 [Micromonospora pallida]	70.1	70.1	48%	4e-09	29.35%	SCL35799.1
carbohydrate-binding protein [Paenibacillus sp. YN15]	65.9	65.9	50%	1e-07	28.83%	WP_113018907.1
hypothetical protein ACM01_07090 [Streptomyces viridochromogenes]	64.7	64.7	82%	2e-07	23.71%	KMS75983.1
hypothetical protein [Streptomyces viridochromogenes]	64.7	64.7	82%	2e-07	23.71%	WP_159029107.1
hypothetical protein [Nocardioides sp. SLBN-172]	64.3	64.3	84%	3e-07	22.75%	WP_141790907.1

Figure 6.S2: Blast analysis to see the extent of the new family (GHxxx). Until proof of activity has been found for the more remote members, it is safer to restrict family membership to those sequences in the red box. In the meantime, sequences outside of the red box will be placed in the non-classified section of CAZy. We will reexamine the possible membership of these sequences to family GHxxx when family GHxxx has sufficient diversity (too few members right now).

CHAPTER 7

CONCLUSIONS AND FUTURE DIRECTIONS

The primary goal of this dissertation was to explore the adaptive immune mechanisms of glycoconjugate vaccine activation using type 3 *S. pneumoniae* glycoconjugates as a model. We had previously shown that glycoconjugates against type 3 *Spn* followed a Tcarb (carbohydrate specific T cell) mediated adaptive immune response similar to prior work (1,2). Therefore, these glycoconjugates would serve as ideal models to uncover the molecular mechanisms that lead to this carbohydrate specific adaptive immune response. Our current hypothesis for this work postulates that T cells are not carbohydrate independent, rather MHCII lacks the ability to bind and present pure carbohydrates which contributed to this T cell notion. However, given the continued evidence for carbohydrate specific T cells (1-5), we hypothesize that intact glycan-peptides survive endosomal conditions and are presented via MHCII (6,7). It is the peptide portion of the glycan-peptide which binds MHCII and leaves the attached glycan exposed for T cell receptor binding. However, the molecular and structural nature of this hypothesis remains unclear.

In this dissertation I provide evidence of the above hypothesis. I demonstrate, using a library of glycan-peptide conjugates to mimic glycoconjugate processed in the endosome, that while peptide alone binds purified MHCII purified oligosaccharide does not. However, a glycoconjugate consisting of the same peptide and oligosaccharide

covalently attached together binds MHCII (Chapter 4). Further, using peptide specific transgenic mice, peptides activate the T cells while glycoconjugate inhibit the T cell response. Taken together, these provide evidence that the peptide portion anchors the glycoconjugate to MHCII and the T cell binds the glycan which would inhibit activation of peptide specific T cells. While this work helps identify molecular mechanisms of immune activation, contributes to the universalization of Tcarb mediated immune responses and aids in knowledge-based vaccine strategies, work to fully understand the immune response remains to be done. Future work will involve cloning carbohydrate specific T cell populations to understand the functional properties of the T cell and T cell receptor which leads to their engagement with carbohydrate antigens. Further, generation of peptide based glycoconjugates will continue to assist in producing knowledge-based vaccines, which will be structurally well defined and target specific, therefore highly immunogenic and protective across populations (8,9).

The idea to switch from full carrier protein vaccines to peptide based is not just a theory, but rather an idea based on logic and understanding of the immune system. Endosomal processing leaves only a small subset of peptides which will bind with MHCII at high affinity, leaving most of the generated peptide epitopes useless. Utilizing peptides already known to be high affinity MHCII binders will ensure a more robust T cell mediated response. Importantly, such advantages of peptide based vaccines has already been explored (1,10-17). Herein, I identified several MHCII binding peptides from clinically relevant carrier proteins currently used in marketed glycoconjugate vaccines (Chapter 5). Moreover, these peptides robustly stimulated CD4+ T cells from human donors expressing a wide variety of MHCII alleles, indicating these peptides may

be immunodominant. Future work will entail utilizing these identified peptides in the conjugation strategy above and testing their efficiency as glycoconjugate vaccines against type 3 *Spn.* Additionally, since peptides alone are limited in their immunogenicity we will explore strategies to combat this. Current strategies include linking these MHCII promiscuous peptides together (10,11,18) or including them as part of liposomes (19) or nanoparticles (20).

Despite a glycoconjugate vaccine against *Spn* being currently available, the morbidity and mortality from the bacterial pathogen remains high (21-24). This vaccine, while being ultimately successful, still elicits variable immune response in serotypes, especially highly virulent type 3 (9,25,26). This serotype alone accounts for over 10% of all IPD (invasive pneumococcal disease) (27). Additionally, widespread use of antibiotics to treat pneumococcal infections has led to a rise in antibiotic resistance (28). Therefore, while continuing to work on improving methods for glycoconjugate vaccine production, it is imperative to seek additional therapies to combat pneumococcal infections caused by type 3 *Spn.* To this end we have been investigating a novel glycoside hydrolase, Pn3Pase, which hydrolyzes the capsule (CPS) of type 3 *Spn* (29). We hypothesized that removal of the CPS through enzyme treatment would reduce the bacteria ability to cause infection and allow the host immune response to better clear infection (30).

While this previous work shed great insight into Pn3Pase role as a therapeutic, we still do not fully understand the mechanism of action. Pn3Pase is unique not only in its ability to acts a potential therapeutic, but also has unique structural properties. Pn3Pase does not fall within any existing Carbohydrate Active enZyme (CAZy) glycoside hydrolase (GH) family. In this dissertation, I established family GHxxx with Pn3Pase as

its founding member (Chapter 6). Additionally, I discovered that Pn3Pase acts through a retaining mechanism with unique kinetics, while identifying domains crucial for function and the active residues. This structural information on Pn3Pase will be crucial to further develop the enzyme with optimal pharmacokinetics. Future structural work will entail solving the structure of this enzyme as we have been unsuccessful despite many attempts. The work demonstrated here, along with previous work on Pn3Pase, provides promise that Pn3Pase may act as a valid alternative therapeutic to antibiotics. Importantly, further research into bacterial glycoside hydrolase could identify more enzymes with capsule degrading abilities that could function as therapies against more encapsulated highly virulent pathogens.

In summary, this dissertation sheds light on the molecular mechanisms of carbohydrate specific immune activation by glycoconjugate vaccines. The discoveries made through this thesis work will be valuable in generating novel glycoconjugate vaccines that function using a knowledge of the immune system. The findings of this work will be applied to strategies to combat invasive pneumococcal diseases.

References

1. Avci, F., Li, X., Tsuji, M., and Kasper, D. (2011) A mechanism for glycoconjugate vaccine activation of the adaptive immune system and its implications for vaccine design. *Nature Medicine* **17**, 1602-U1115
2. Middleton, D. R., Sun, L., Paschall, A. V., and Avci, F. Y. (2017) T Cell-Mediated Humoral Immune Responses to Type 3 Capsular Polysaccharide of. *J Immunol* **199**, 598-603
3. Sun, X., Stefanetti, G., Berti, F., and Kasper, D. L. (2019) Polysaccharide structure dictates mechanism of adaptive immune response to glycoconjugate vaccines. *Proc Natl Acad Sci U S A* **116**, 193-198
4. Avci, F., Li, X., Tsuji, M., and Kasper, D. (2012) Isolation of carbohydrate-specific CD4(+) T cell clones from mice after stimulation by two model glycoconjugate vaccines. *Nature Protocols* **7**, 2180-2192
5. Avci, F. Y., Li, X., Tsuji, M., and Kasper, D. L. (2013) Carbohydrates and T cells: a sweet twosome. *Semin Immunol* **25**, 146-151
6. Avci, F., Kasper, D., Paul, W., Littman, D., and Yokoyama, W. (2010) How Bacterial Carbohydrates Influence the Adaptive Immune System. *Annual Review of Immunology, Vol 28* **28**, 107-130
7. Guttormsen, H.-K., Sharpe, A. H., Chandraker, A. K., Brigtsen, A. K., Sayegh, M. H., and Kasper, D. L. (1999) Cognate stimulatory B-Cell-T-Cell interactions are

critical for T-cell help recruited by glycoconjugate vaccines. *Infect Immun* **67**, 6375-6384

8. Huang, S. S., Johnson, K. M., Ray, G. T., Wroe, P., Lieu, T. A., Moore, M. R., Zell, E. R., Linder, J. A., Grijalva, C. G., Metlay, J. P., and Finkelstein, J. A. (2011) Healthcare utilization and cost of pneumococcal disease in the United States. *Vaccine* **29**, 3398-3412
9. Wantuch, P. L., and Avci, F. Y. (2018) Current status and future directions of invasive pneumococcal diseases and prophylactic approaches to control them. *Hum Vaccin Immunother* **14**, 2303-2309
10. Falugi, F., Petracca, R., Mariani, M., Luzzi, E., Mancianti, S., Carinci, V., Melli, M. L., Finco, O., Wack, A., Di Tommaso, A., De Magistris, M. T., Costantino, P., Del Giudice, G., Abrignani, S., Rappuoli, R., and Grandi, G. (2001) Rationally designed strings of promiscuous CD4+ T cell epitopes provide help to Haemophilus influenzae type b oligosaccharide: a model for new conjugate vaccines. *Eur. J. Immunol* **31**, 3816-3824
11. Baraldo, K., Mori, E., Bartoloni, A., Norelli, F., Grandi, G., Rappuoli, R., Finco, O., and Del Giudice, G. (2005) Combined conjugate vaccines: enhanced immunogenicity with the N19 polyepitope as a carrier protein. *Infect Immun* **73**, 5835-5841
12. Bixler, G. S., Eby, R., Dermody, K. M., Woods, R. M., Seid, R. C., and Pillai, S. (1989) Synthetic peptide representing a T-cell epitope of CRM197 substitutes as carrier molecule in a Haemophilus influenzae type B (Hib) conjugate vaccine. *Adv Exp Med Biol* **251**, 175-180

13. Cai, H., Chen, M. S., Sun, Z. Y., Zhao, Y. F., Kunz, H., and Li, Y. M. (2013) Self-adjuvanting synthetic antitumor vaccines from MUC1 glycopeptides conjugated to T-cell epitopes from tetanus toxoid. *Angew Chem Int Ed Engl* **52**, 6106-6110

14. de Velasco, E. A., Merkus, D., Anderton, S., Verheul, A. F., Lizzio, E. F., Van der Zee, R., Van Eden, W., Hoffman, T., Verhoef, J., and Snippe, H. (1995) Synthetic peptides representing T-cell epitopes act as carriers in pneumococcal polysaccharide conjugate vaccines. *Infect Immun* **63**, 961-968

15. Jackson, D. C., Purcell, A. W., Fitzmaurice, C. J., Zeng, W., and Hart, D. N. (2002) The central role played by peptides in the immune response and the design of peptide-based vaccines against infectious diseases and cancer. *Curr Drug Targets* **3**, 175-196

16. Rodrigues-da-Silva, R. N., Correa-Moreira, D., Soares, I. F., de-Luca, P. M., Totino, P. R. R., Morgado, F. N., Oliveira Henriques, M. D. G., Peixoto Candea, A. L., Singh, B., Galinski, M. R., Moreno, A., Oliveira-Ferreira, J., and Lima-Junior, J. D. C. (2019) Immunogenicity of synthetic peptide constructs based on PvMSP9E795-A808, a linear B-cell epitope of the P. vivax Merozoite Surface Protein-9. *Vaccine* **37**, 306-313

17. Zeigler, D. F., Roque, R., and Clegg, C. H. (2019) Optimization of a multivalent peptide vaccine for nicotine addiction. *Vaccine* **37**, 1584-1590

18. Slingluff, C. L. (2011) The present and future of peptide vaccines for cancer: single or multiple, long or short, alone or in combination? *Cancer J* **17**, 343-350

19. Riaz, M. K., Riaz, M. A., Zhang, X., Lin, C., Wong, K. H., Chen, X., Zhang, G., Lu, A., and Yang, Z. (2018) Surface Functionalization and Targeting Strategies of Liposomes in Solid Tumor Therapy: A Review. *Int J Mol Sci* **19**

20. Fujita, Y., and Taguchi, H. (2011) Current status of multiple antigen-presenting peptide vaccine systems: Application of organic and inorganic nanoparticles. *Chem Cent J* **5**, 48

21. Richter, S. S., Heilmann, K. P., Dohrn, C. L., Riahi, F., Diekema, D. J., and Doern, G. V. (2013) Pneumococcal serotypes before and after introduction of conjugate vaccines, United States, 1999-2011(1.). *Emerg Infect Dis* **19**, 1074-1083

22. (1997) Prevention of pneumococcal disease: recommendations of the Advisory Committee on Immunization Practices (ACIP). *MMWR Recomm Rep* **46**, 1-24

23. Hamborsky J, K. A., Wolfe S. (2015) Pneumococcal Disease. in *Epidemiology and Prevention of Vaccine-Preventable Diseases* (Prevention, C. f. D. C. a. ed.), 13 Ed., Public Health Foundation, Washington D.C. pp

24. Hausdorff, W. P., Bryant, J., Kloek, C., Paradiso, P. R., and Siber, G. R. (2000) The contribution of specific pneumococcal serogroups to different disease manifestations: implications for conjugate vaccine formulation and use, part II. *Clin Infect Dis* **30**, 122-140

25. Wantuch, P. L., and Avci, F. Y. (2019) Invasive pneumococcal disease in relation to vaccine type serotypes. *Hum Vaccin Immunother*, 1-2

26. Dagan, R., Patterson, S., Juergens, C., Greenberg, D., Givon-Lavi, N., Porat, N., Gurtman, A., Gruber, W. C., and Scott, D. A. (2013) Comparative

immunogenicity and efficacy of 13-valent and 7-valent pneumococcal conjugate vaccines in reducing nasopharyngeal colonization: a randomized double-blind trial. *Clin Infect Dis* **57**, 952-962

27. Sugimoto, N., Yamagishi, Y., Hirai, J., Sakanashi, D., Suematsu, H., Nishiyama, N., Koizumi, Y., and Mikamo, H. (2017) Invasive pneumococcal disease caused by mucoid serotype 3 Streptococcus pneumoniae: a case report and literature review. *BMC Res Notes* **10**, 21

28. Kim, L., McGee, L., Tomczyk, S., and Beall, B. (2016) Biological and Epidemiological Features of Antibiotic-Resistant Streptococcus pneumoniae in Pre- and Post-Conjugate Vaccine Eras: a United States Perspective. *Clin Microbiol Rev* **29**, 525-552

29. Middleton, D. R., Zhang, X., Wantuch, P. L., Ozdilek, A., Liu, X., LoPilato, R., Gangasani, N., Bridger, R., Wells, L., Linhardt, R. J., and Avci, F. Y. (2018) Identification and characterization of the Streptococcus pneumoniae type 3 capsule-specific glycoside hydrolase of Paenibacillus species 32352. *Glycobiology* **28**, 90-99

30. Middleton, D. R., Paschall, A. V., Duke, J. A., and Avci, F. Y. (2018) Enzymatic Hydrolysis of Pneumococcal Capsular Polysaccharide Renders the Bacterium Vulnerable to Host Defense. *Infect Immun*

APPENDIX A

GLYCOCONJUGATE SYNTHESIS USING CHEMOSELECTIVE LIGATION

Shuihong Chen, **Paeton L Wantuch**, Megan E. Kizer, Dustin R. Middleton, Ruitong Wang, Mikaela DiBello, Mingli Li, Xing Wang, Xuebing Li, Vasanthi Ramachandiran, Fikri Y Avci, Fuming Zhang, Xing Zhang and Robert J Linhardt. 2019. 17(10), 2646-2650. Reprinted here with the permission of publisher.

*PLW contribution to this work was oligosaccharide purification, OVA peptide synthesis, MHCII binding assay and manuscript writing.

DOI: 10.1039/c9ob00270g

Abstract

Chemoselective ligation of carbohydrates and polypeptides was achieved using an adipic acid dihydrazide cross-linker. The reducing end of a carbohydrate is efficiently attached to peptides in two steps, constructing a glycoconjugate in high yield and with high regioselectivity, enabling the production of homogeneous glycoconjugates.

Introduction

Carbohydrates are often found linked to other biomolecules to form glycoconjugates. Glycoproteins and glycolipids, commonly found on the surface of cells, act as receptors for cell-cell recognition and are involved in cell growth, repair, adhesion and migration as well as tumor metastasis in cancer (1). Vaccination is a key strategy for the control of infectious diseases caused by viruses, bacteria and parasites (2). Oligosaccharides and polysaccharides can be antigenic (3), and capsular polysaccharides (CPSs) of bacterial pathogens, including *Streptococcus pneumoniae*, *Haemophilus influenzae type b* (Hib), and *Neisseria meningitidis,* are highly antigenic and recognizable by mammalian B cell receptors. CPSs are major targets for eliciting carbohydrate specific antibody responses to confer protection from these pathogens. However, CPSs are poorly immunogenic only eliciting a short-term immune response without immunological memory. The conjugation of the carbohydrates of these pathogens to protein carriers can induce long-lasting protection against encapsulated bacteria opening a path to the development of glycoconjugate vaccines (3-4). Glycoconjugate synthesis generally involves the random linking of carbohydrate and protein without regards to sites, leading to an incomplete understanding of mechanism of action (3). The cellular and molecular mechanisms for adaptive immune activation mediated by glycoconjugate vaccines have

been elucidated (5). Demystifying T cell activation mechanisms of glycoconjugate vaccines represents a key step towards designing a knowledge-based, structurally-defined, generation of new vaccines. This study offers an efficient conjugation strategy.

Different carrier proteins can remarkably impact immunogenicity and the efficacy of glycoconjugate vaccines. The ability to synthesize glycoconjugates of greater structural diversity should afford an improved understanding of vaccine mechanism and result in the development of more effective vaccines (6). In addition, the controlled glycosylation of peptide-based therapeutics can help protect against proteolytic degradation, denaturation and premature clearance, modulating their biophysical and physiological properties (7). Improved methods for glycoconjugate synthesis are urgently needed.

There are many strategies for glycopeptide synthesis relying on conventional chemical utilization of aldehydes, thiols, activated esters or hydrazides, carboxylic acids and amines, and even new bacterial protein glycan coupling technologies (PGCT) (2a,6b,6c,8). These approaches provide low yields and complex product mixtures, particularly when the reactants involve proteins and glycans with multiple reactive sites or with high levels of steric hindrance (2).

We describe a high-yield ligation chemistry affording homogenous glycoconjugates, in which the reducing sugar is first reacted with adipic acid dihydrazide (ADH) to form a carbohydrate bearing a linker at its reducing-end. The remaining acyl hydrazide is oxidatively converted to an acyl azide and then captured as a thioester and transesterified with the cysteine residue of a peptide to obtain a thioester-linked glycoconjugate. When this cysteine is at the N-terminus of the peptide chain, the thioester rapidly rearranges to form a stable amide linkage between the carbohydrate and peptide **(Figure A.1)**. The

installation of the ADH linker at the carbohydrate's reducing end and its selective reaction with the cysteine residue of the peptide affords homogeneous constructs with compositional control of the carbohydrate-protein conjugate and preserves the integrity of carbohydrate epitopes (9). In addition, ADH provides a 10-atom bridge between the carbohydrate and peptide after conjugation, potentially increasing detection limits of immunoassays by reducing steric hindrance (10).

Material and Methods

All chemicals unless noted otherwise were purchased from commercial sources and used as received. Yields are given after purification. NMR spectra were recorded on a Bruker 600 spectrometer with topspin 2.1.6 software at 298 K. Mass data were acquired by high-resolution ESI-MS (Thermo LTQ XL Orbitrap, Bremen, Germany).

I. General procedure for glycopeptides synthesis

a) Carbohydrate-ADH (adipic acid dihydrazide) synthesis

Carbohydrate (2.0 mg, 1.0 equiv) and ADH (50.0 equiv) were dissolved in 200 μL acetate buffer (pH 5.5, 100 mM, containing 1 mM PhNH2) at 50 oC and incubated for 24 h. The conjugation product was then recovered by Sephadex G-10 gel filtration **(for small size oligosaccharides-ADH, Table A.2, entries 1-7)** or size exclusion spin column (MWCO 3000 Da) **(for GAGs-ADH, Table A.2, entries 8-15)** and washed by distilled water and lyophilized.

b) Chemoselective ligation with N-terminal cysteine-peptides

Peptide[a] (2.0 mg, 1.0 equiv) and carbohydrate-ADH (2.0 equiv) were dissolved in 0.2 mL of ligation buffer (3.0 M Gn·HCl, 0.2 M Na2HPO4, pH 3.8), 20 μL of NaNO2 (200 mM) was added dropwise, and the reaction mixture was stirred for 20 min at -10 oC.

After that, 0.2 mL of freshly prepared MPAA (200 mM, dissolved in ligation buffer pH 7.0) was added, and the acidity of the mixed solution was adjusted to pH 7.0 with NaOH (2.0 M) slowly. The reaction mixture was stirred at room temperature for 2 hours. Before analysis, the reaction solution was reduced by 30 mM neutral tris(2-carboxyethyl)phosphine(TCEP). The reactions were monitored by HPLC and conjugation product were recovered by size exclusion spin column (MWCO 3000 Da) and Sephadex G-10 gel filtration washed by distilled water. The linkage of the corresponding conjugate is amido bond.

[a]N-terminal cysteine-peptides with no modification.

c) Chemoselective ligation with other peptides

Peptide[b] (2.0 mg, 1.0 equiv) and carbohydrate-ADH (2.0 equiv) were dissolved in 0.2 mL of ligation buffer (3.0 M Gn·HCl, 0.2 M Na2HPO4, pH 3.8), 20 μL of NaNO2 (200 mM) was added dropwise, and the reaction mixture was stirred for 20 min at -10 oC. After that, 0.2 mL of freshly prepared MPAA (200 mM, dissolved in ligation buffer pH 7.0) was added, and the acidity of the mixed solution was adjusted to pH 7.0 with NaOH (2.0 M) slowly. The reaction mixture was stirred at room temperature for 2 hours. The 5 reactions were monitored by HPLC and conjugation product were recovered by size exclusion spin column (MWCO 3000 Da) and Sephadex G-10 gel filtration washed by distilled water. The linkage of the corresponding conjugate is thioester bond.

[b]C-terminal, internal cysteine-peptides or acetyl protected N-terminal cysteine-peptides.

II. MHCII binding studies of glycopeptides dp2- and dp4-OVA

a) Disaccharide GlcA(1→4)Glc and tetrasaccharide

GlcA(1→4)Glc(1→3)GlcA(1→4)Glc preparation

Pure type 3 capsular polysaccharide (CPS, 40 mg) (ATCC 172-X) was added to 20 mL distilled H2O in a flask and heated on a hot plate. When most of the CPS is dissolved, 20 mL of 0.6 M TFA was added to the flask and continued to heat until fully dissolved and then incubated for exactly 2 hours at 100 oC. Immediately cool on ice and aliquot into tubes, speed vac to dryness. Re-suspend pellets in H2O and speed vac to dryness again. Re-suspended contents of all tubes in 2 mL H2O, centrifuged and the supernatant was subjected to Superdex 30 column to obtain the target compounds

b) Synthesis of glycopeptides dp2- and dp4-OVA

The procedures of dp2-ADH and dp4-ADH, and the corresponding conjugates with CysOVA (sequence: CISQAVHAAHAEINEAGR) are described in Section I.

c) MHCII binding test

Purified MHCII monomers (mouse allele I-Ad) were graciously provided by the NIH tetramer facility. Monomers came loaded with a placeholder peptide with a 3C protease cleavage site. Peptide is cleaved by treating monomer with 3C protease (PierceTM HRV 3C Protease Solution Kit) for 8 hours at room temperature. Cleaved monomer with empty binding groove is then loaded with desired peptide or glycan-peptide through an 7 exchange reaction. Exchange reactions use 200 μg cleaved monomer with 270 μM peptide (or glycan-peptide) sample at pH 5.0 in a citrate buffer. Samples for this experiment were OVA peptide 323-339, scrambled OVA peptide, OVA peptide 323-339 with an added Cys residue at the N terminus, disaccharide-peptide conjugate, and tetrasaccharide-peptide conjugate. Reactions were incubated for 5 days at room temperature. At the end of the incubation, reactions were neutralized with 1 M sodium phosphate buffer pH 7.5 and spun down at max speed for 10 minutes to remove

aggregates. Absorbance was measured at 280, all samples except the scrambled OVA peptide reaction gave significant concentration values. Binding was confirmed by isoelectric focusing (NovexTM pH 3-7 Protein IEF gels). Gel was visualized using silver stain (PierceTM Silver Stain Kit).

Results

Our study began by re-investigating the previously published reaction between heparin and adipic acid dihydrazide for the directional immobilization of heparin onto surfaces (11). A heparin dodecasaccharide, prepared through the controlled enzymatic depolymerization of heparin (12), and ADH were reacted to form a hydrazone bond **(Figure A.1 and Table A.1)**. This chemistry is particularly challenging with this glycan since the heparin dodecasaccharide is a polyanion (-48 charge) with a molecular weight 3990 Da. Its single reducing end is comprised of a relatively unreactive sugar with an anionic *N*-sulfo group adjacent to its anomeric center. Several common conjugation conditions including $NaCNBH_3$, $PhNH_2$/MeOH/H_2O (13) and $CeCl_3 \cdot 7H_2O$ (14) were initially tested and resulted in very low yields **(entries 1-3, Table A.1)**. Interestingly, conjugation in AcOH-DMSO (15) afforded a high yield of multiple ADH conjugated to the heparin dodecasaccharide **(entry 4, Table A.1)**. HCOONa buffer (16) afforded the single desired product at 75% yield **(entry 5, Table A.1)** and $PhNH_2$/acetate buffer (17) provided the same desired product in 80% yield **(entry 6, Table A.1)**. Based on these results, $PhNH_2$/acetate buffer was selected as it provided high yields in the shortest reaction time.

With the optimized conditions in hand, we next turned our attention to investigate the scope of this reaction by treating various reducing disaccharides, oligosaccharides and

polysaccharides, with adipic acid dihydrazide to prepare a small library of carbohydrate-ADH conjugates **(Table A.2)**. Coupling either simple neutral or acidic oligosaccharides with dihydrazide under these optimized conditions, such as lactose, maltotriose, maltotetraose, maltooligosaccharide and 6'-sialyllactose afforded the expected products in nearly quantitative yields, and the yields were greatly improved compared to those obtained using previous methods (16) **(entries 1-5, Table A.2)**. Two tetrasaccharides, Glc(1→3)GlcA(1→4)Glc(1→4)GlcA (18a) and GlcA(1→4)Glc(1→3)GlcA(1→4)Glc (18b), having glucuronic acid or glucose at the reducing end, prepared through controlled depolymerization of type-3 CPS of *Streptococcus pneumoniae* (Pn3P), also proved excellent substrates to conjugation with ADH, and are being investigated in vaccine development **(entry 6-7, Table A.2)**. Notably, more structurally complex glycosaminoglycans, having both short or long carbohydrate chains, and low or high levels of sulfation, also furnished good results suggesting the adaptability of this method to a variety substrate **(entries 8-15, Table A.2)**. Representative 1D ^{1}H NMR data for the analysis of heparin-ADH, depolymerized heparin dodecasaccharide-ADH and lactose-ADH, are presented in **Figure A.2A**. All the other compounds were also confirmed by NMR analyses.

We next investigated the conjugation of different carbohydrate-ADH derivatives with a variety of peptides containing N-terminal cysteine residues **(Table A.3)**. Initially, we utilized conventional carbodiimide chemistry, involving EDC/NHS, to react lactose-ADH with the Cys-Gly dipeptide. This resulted in the recovery of starting materials with only trace amounts of conjugate **(entry 1, Table A.3)**. 4-(4,6-Dimethoxy-1,3,5-triazin-2-yl)-4-methyl-morpholinium chloride (DMTMM) has been reported to be highly efficient for

the activation of hydrazide chemistry in protein-protein ligation (19). Thus, we undertook the reaction of lactose-ADH with the Cys-Gly dipeptide using DMTMM and obtained the desired glycopeptide conjugate in moderate 37% yield **(entry 2, Table A.3)**.

Liu and co-workers (20) recently developed a protein ligation method for the chemoselective reaction between a C-terminal peptide hydrazide and an N-terminal Cys-peptide affording a peptide bond. Inspired by this study, we investigated using sodium nitrite and thiol to form glycopeptides **(Figure A.1 and Table A.3)**. We were pleased that the reaction of lactose-ADH with the Cys-Gly dipeptide in the presence of $NaNO_2$/MPAA (4-mercaptophenylacetic acid) generated the glycopeptide in excellent 85% yield **(entry 3, Table A.3)**.

A representative characterization of lactose-dipeptide (Cys-Gly) by HRMS and NMR spectroscopy is presented in **Figure A.2B-E**. The signals corresponding to the anomeric protons of Gal and Glc residues are 4.36 and 4.04 ppm, respectively. The signals at 1.52, 2.15 and 2.28 ppm correspond to the adipic linker. The $^{3}J_{HH}$ coupling constant of the Glc anomeric proton was 9.0 Hz, suggesting a β-linkage with the adipic linker. Proton signal of α-position at 4.67 ppm was overlapped by HDO in ^{1}H NMR spectrum but could be detected (4.67, 52.2) in 2D HSQC spectrum. The H-H correlations between α and β position of cysteine residue were also observed. The structure was characterized by HRMS with $[M+H]^{+}$ 645.2285 **(Figure A.2C)**.

We next evaluated the substrate sensitivity of this reaction. Using the more complex cysteine-containing tridecapolypeptide CRKRLQVQLSIRT, corresponding to the syndecan-1 binding sequence within laminin (21), we coupled either simple lactose-ADH or the more complex heparin dp12-ADH **(entries 4-5, Table A.3)**.

Cysteine-hydrazide coupling chemistry involves three steps based on a mechanistic analysis. Carbohydrate-ADH is first oxidized to the acyl azide, which subsequently reacts with the thiol group-containing reagent 4-mercaptophenylacetic acid and is then transesterfied by the N-terminal cysteine residue generating an acetyl-thiol ester intermediate. Proceeding through a five-membered ring transition state, the vicinal amino group of this cysteine residue then attacks the acetyl-thiol ester forming a more stable amide linkage.

Ovalbumin (OVA) 323-339 is a major MHC class II-restricted T-cell epitope of OVA, which has been used for studies of immunological allergic responses. The N-cysteine-containing OVA 323-339, CISQAVHAAHAEINEAGR, was used in the preparation of glycopeptides with stable amide linkages to lactose and a disaccharide and tetrasaccharide derived from the type-3 CPS of *Streptococcus pneumoniae* derived **(entries 6-8, Table A.3)**. The binding of the major histocompatibility complex class II protein (MHCII) with the two synthetic glycopeptides **(entries 7-8, Table A.3)** was demonstrated **(Figure A.3)**. The OVA peptide (positive control, lane 1) and the two synthetic OVA-glycopeptides (lanes 4 and 5) bound to MHCII with comparable affinities. This is in contrast to the scrambled OVA peptide (negative control, lane 2), which only afforded a faint band due to aggregation of I-A^d protein in the absence of a binding peptide. MHCII molecules with an empty binding groove reportedly aggregate during prolonged incubation periods (22). Binding was observed through the downward shift in pI resulting from the introduction of the negative charge associated with the glycan that is especially noticeable in the tetrasaccharide-peptide conjugate. T-cells have long been considered to be non-responsive to carbohydrate antigens. Recently, however,

a model for T-cell recognition of a glycoconjugate vaccine has been advanced in which a carbohydrate epitope is presented on the surface of an antigen-presenting cell through MHCII. This presented epitope is then recognized by, and stimulates, carbohydrate specific CD4+ T cells (5a-5c). Carbohydrate does not interfere with peptide binding MHCII **(Figure A.3)**. This suggests that in the case of glycoconjugate vaccines, the carrier peptide binds MHCII exposing the covalently linked carbohydrate for T-cell recognition.

Finally, we examined if coupling could occur with thiol-functionalized biomolecules in place of ones with an unprotected N-terminal cysteine. Enfuvirtide (ENF), a 36-residue synthetic peptide drug, is a potent inhibitor of HIV-1 infection used in the treatment of AIDS. Sialic acid conjugated ENF has a dramatically extended circulating half-life (23). AcHNCys-T20, Ac-CYTSLIHSLIEESQNQQEKNEQELLELDKWASLWNWF-NH_2, a modified version of ENF with an acetyl protected cysteine at its N-terminus, was conjugated with 6'-sialyllactose-ADH and obtained the thioester linked glycocojugate in good yield **(entry 9, Table A.3)**. Additional glycoconjugates, to N-terminal protected, C-terminal, or internal cysteine residues also afford thioester-linked products.

Discussion

In summary, we have successfully implemented a cross-linking strategy targeting glycoconjugate preparation. This chemistry involves ADH as cross-linking reagent combining the reducing end of carbohydrates with the thiol groups of target peptides. Based on Liu and coworkers (20), who used a synthetic mono-hydrazide modified peptide as precursor in polypeptide synthesis, we introduced a

bifunctional dihydrazide linker that could be directly introduced in high yield into the reducing end of the carbohydrate. The second hydrazide group was then oxidatively converted to an acyl azide, captured as a thioester and then chemoselectively transesterified with a Cys-peptide to afford the target glycopeptide. The strength of this approach lies in its ability to produce structurally defined, homogeneous glycan-peptide conjugates. Based on our mechanistic understanding of the activation of the immune response to glycoconjugates, this approach holds promise for the synthesis of glycopeptide vaccine (5). This approach also facilitates the synthesis of glycosylated peptide-based therapeutics, protecting polypeptides against proteolysis, denaturation and improving their pharmacokinetics by modulating the biophysical and physiological properties. Finally, this chemistry has been extended to the chemoselective ligation of carbohydrate-hydrazide to biomolecules containing a single thiol group further broadening the scope of this reaction.

References

1. In *Essentials of Glycobiology* (Eds.: A. Varki, R. D. Cummings, J. D. Esko, P. Stanley, G. W. Hart, M. Aebi, A. G. Darvill, T. Kinoshita, N. H. Packer, J. H. Prestegard, R. L. Schnaar and P. H. Seeberger), Cold Spring Harbor Press, NY, 2017.
2. B. Kuberan and R. J. Linhardt, *Curr. Org. Chem.* 2000, *4*, 653-677; b) J. Hütter, B. Lepenies in *Carbohydrate-Based Vaccines: Methods and Protocols* (Ed.: B. Lepenies), Springer New York, New York, NY, 2015, pp. 1-10.
3. F. Y. Avci and D. L. Kasper, *Annu. Rev. Immunol.* 2010, *28*, 107-130.
4. O. T. Avery and W. F. Goebel, *J. Exp. Med.* 1931, *54*, 437-447.
5. F. Y. Avci, X. M. Li, M. Tsuji and D. L. Kasper, *Nat. Med.* 2011, *17*, 1602-1610; b) F. Y. Avci, X. M. Li, M. Tsuji and D. L. Kasper, *Nat. Protoc.* 2012, *7*, 2180-2192; c) D. R. Middleton, L. Sun, A. V. Paschall and F. Y. Avci, *J. Immunol.* 2017, *199*, 598-603.
6. R. Dagan, J. Poolman and C.-A. Siegrist, *Vaccine* 2010, *28*, 5513-5523; b) K. M. Koeller and C.-H. Wong, *Chem. Rev.* 2000, *100*, 4465-4494; c) B. G. Davis, *J. Chem. Soc., Perkin Trans. 1* 1999, 3215-3237.
7. R. J. Solá and K. Griebenow, *BioDrugs* 2010, *24*, 9-21.
8. V. S. Terra, D. C. Mills, L. E. Yates, S. Abouelhadid, J. Cuccui and B. W. Wren, *J. Med. Microbiol*. 2012, *61*, 919-926; b) P.-A. Ashford and S. P. Bew, *Chem. Soc. Rev.* 2012, *41*, 957-978; c) H. Hojo and Y. Nakahara, *Pept. Sci.* 2007, *88*,

308-324; d) H. Herzner, T. Reipen, M. Schultz and H. Kunz, *Chem. Rev.* 2000, *100*, 4495-4538.

9. R. Adamo, *Acc. Chem. Res*. 2017, *50*, 1270-1279.
    a. Basu, T. G. Shrivastav and K. P. Kariya, *Clin. Chem*. 2003, *49*, 1410-1412.
10. V. D. Nadkarni, A. Pervin and R. J. Linhardt, *Anal. Biochem*. 1994, *222*, 59-67.
    a. Pervin, C. Gallo, K. A. Jandik, X.-J. Han and R. J. Linhardt, *Glycobiology* 1995, *5*, 83-95.
11. T. P. Coxon, T. W. Fallows, J. E. Gough and S. J. Webb, *Org. Biomol. Chem.* 2015, *13*, 10751-10761.
12. J. M. dos Santos, *Eur. J. Org. Chem*. 2014, 6411-6417.
13. G. Li, L. Li, C. Xue, D. Middleton, R. J. Linhardt and F. Y. Avci, *J. Chromatogr. A* 2015, *1397*, 43-51.
14. N. S. Flinn, M. Quibell, T. P. Monk, M. K. Ramjee and C. J. Urch, *Bioconjugate Chem.* 2005, *16*, 722-728.
15. K. Godula and C. R. Bertozzi, *J. Am. Chem. Soc.* 2010, *132*, 9963-9965.
16. D. R. Middleton, X. Zhang, P. L. Wantuch, A. Ozdilek, X. Liu, R. Lopilato, N. Gangasani, R. Bridger, L. Wells, R. J. Linhardt and F. Y. Avci, *Glycobiology*, 2018, *28*, 90-99; b) D. J. Lefeber, R. Gutiérrez Gallego, C. H. Grün, D. Proietti, S. D'Ascenzi, P. Costantino, J. P. Kamerling and J. F. G. Vliegenthart, *Carbohydr. Res.* 2002, *337*, 819-825.

a. Leitner, L. A. Joachimiak, P. Unverdorben, T. Walzthoeni, J. Frydman, F. Forster and R. Aebersold, *Proc. Natl. Acad. Sci. USA* 2014, *111*, 9455-9460.

17. G.-M. Fang, Y.-M. Li, F. Shen, Y.-C. Huang, J.-B. Li, Y. Lin, H.-K. Cui and L. Liu, *Angew. Chem. Int. Ed.* 2011, *50*, 7645-7649.
18. L. R. Balaoing, A. D. Post, A. Y. Lin, H. Tseng, J. L. Moake and K. J. Grande-Allen, *PLoS One* 2015, *10*, e0130749.
19. G. M. Grotenbreg, M. J. Nicholson, K. D. Fowler, K. Wilbuer, L. Octavio, M. Yang, A. K. Chakraborty, H. L. Ploegh and K. W. Wucherpfennig, *J. Biol. Chem.* 2007, *282*, 21425-21436.
20. S. Cheng, X. Chang, Y. Wang, G. F. Gao, Y. Shao, L. Ma and X. Li, *J. Med. Chem.* 2015, *58*, 1372-1379.

Figure A.1

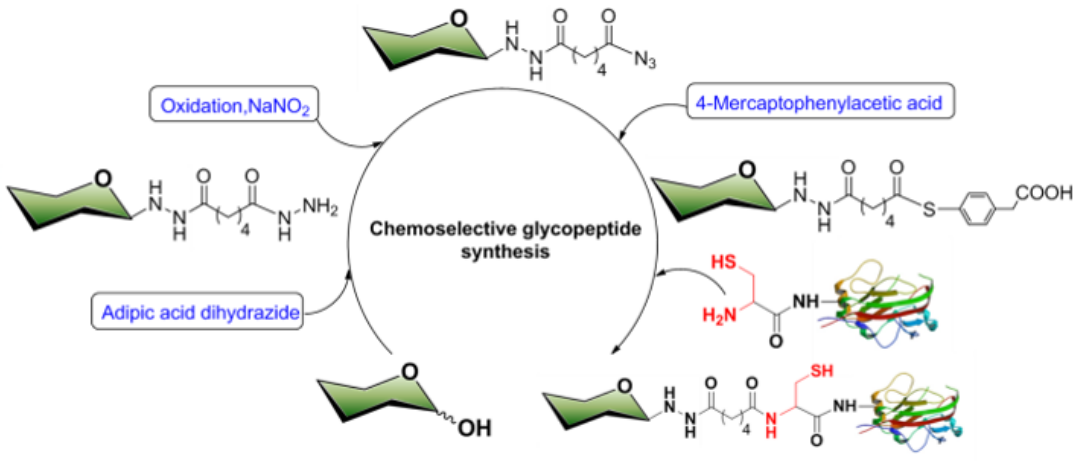


Figure A.1: **Glycopeptide preparation.** Peptides where cysteine is not at the N-terminus could also be synthesized but instead result in a thioester linkage

Table A.1

Table A.1: Optimization for conjugation of heparin dodecasaccharide with adipic acid dihydrazide

Entry	Reagents/solvents	Temp(°C)	Time(h)	Yield
1	$NaCNBH_3$, H_2O	70	48	10%
2	$PhNH_2$, MeOH/H_2O	75	24	trace
3	$CeCl_3 \cdot 7H_2O$, EtOH/H_2O	50	1	trace
4	AcOH/DMSO	45	48	Multiple conjugates
5	HCOONa buffer pH 4.75	37	72	75%
6	$PhNH_2$, acetate buffer pH 5.5	50	24	80%

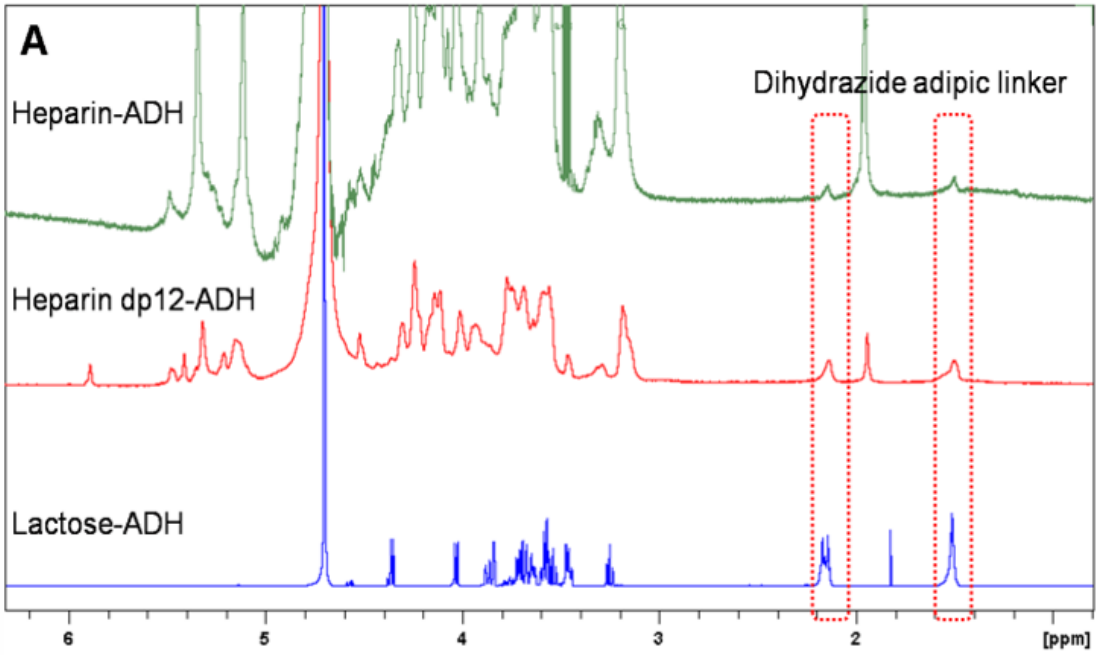
Table A.2

Table A.2: Carbohydrate substrate scope in dihydrazide reaction*

Entry	Carbohydrate	Yield (%)
1	Lactose	99
2	Maltotriose	95
3	Maltotetraose	91
4	Maltooligosaccharide	87
5	6'-Sialyllactose	97
6	Glc-GlcA-Glc-GlcA	85
7	GlcA-Glc-GlcA-Glc	89
8	Heparin dodecasaccharide	80
9	Heparan sulfate	80[b]
10	Heparin	80[b]
11	LMW-heparin	80[b]
12	*N*-desulfated heparin	90[b]
13	Chondroitin sulfate A	80[b]
14	Chondroitin sulfate E	80[b]
15	Dermatan sulfate	50[b]

*The ADH reaction with carbohydrate was conducted at PhNH2/acetate buffer (pH 5.5) conditions described in entry 6, Table 1. [b]Yields of ADH reaction at the reducing end of a heterogeneous, polydisperse polysaccharides were estimated by 1H-NMR.

Figure A.2


A
Heparin-ADH
Dihydrazide adipic linker
Heparin dp12-ADH
Lactose-ADH
6
5
4
3
2
[ppm]

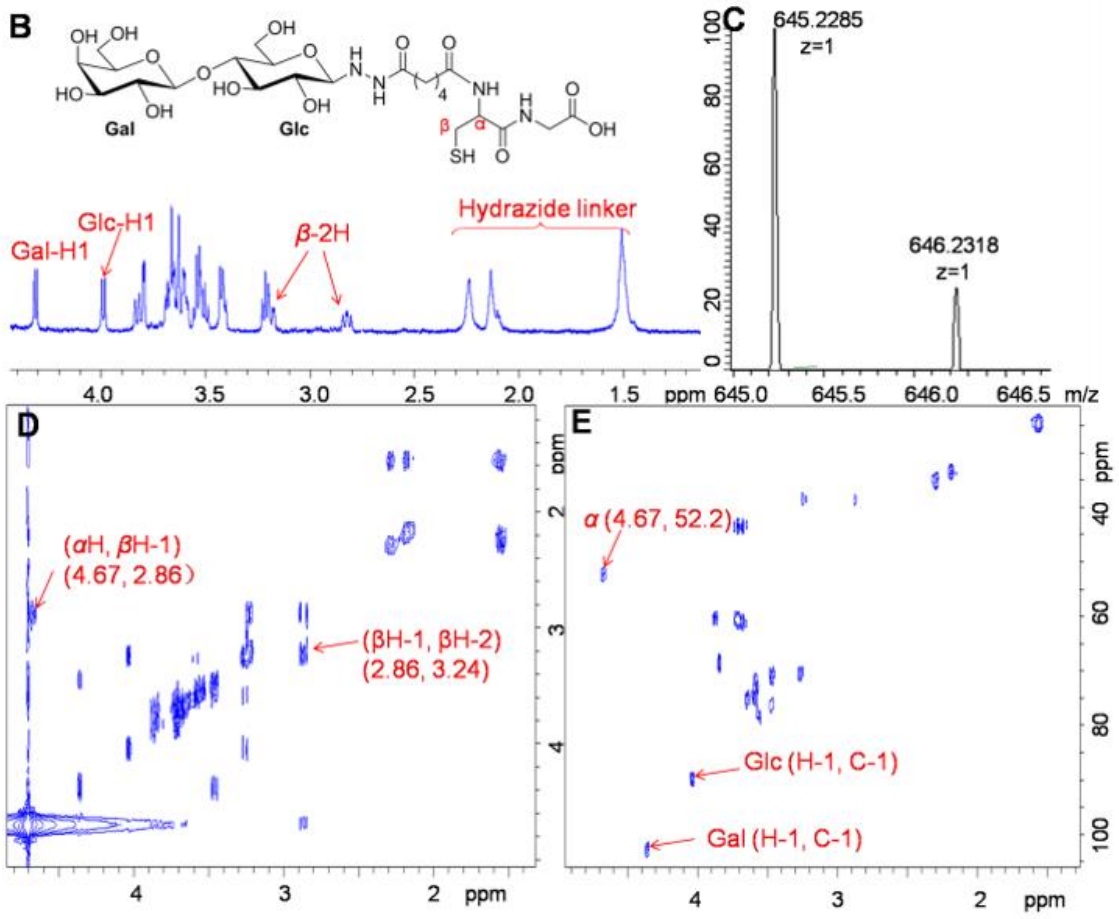


Figure A.2: NMR and HRMS analysis. Panel (A) shows 1D ^{1}H NMR spectra of heparin-, heparin dodecasaccharide- and lactose-ADH. Two sets of peaks at 1.51 and 2.16 ppm in the ^{1}H NMR correspond to dihydrazide adipic linker. Panel (B), (D), (E) and (C) show the 1D ^{1}H, 2D ^{1}H-^{1}H COSY, ^{1}H-^{13}C HSQC NMR and HRMS (positive-mode) spectra of lactose-dipeptide (Cys-Gly) conjugate.

Table A.3

Table A.3: Conjugation of carbohydrate-adipic acid dihydrazide (ADH) with peptides

Entry	Carbohydrate-ADH	Aglycones	Conditions	Yield (%)
1	Lactose	Cys-Gly dipeptide	EDC/NHS	Trace
2	Lactose	Cys-Gly dipeptide	DMTMM	37
3	Lactose	Cys-Gly dipeptide	$NaNO_2$, MPAA	85
4	Lactose	Cys-Laminin peptide	$NaNO_2$, MPAA	75
5	Heparin dp12	Cys-Laminin peptide	$NaNO_2$, MPAA	57
6	Lactose	Cys-OVA-323-339	$NaNO_2$, MPAA	68
7	GlA-Glc	Cys-OVA-323-339	$NaNO_2$, MPAA	67
8	$(GlcA\text{-}Glc)_2$	Cys-OVA-323-339	$NaNO_2$, MPAA	64
9	6'-Sialyllactose	AcHNCys-T20	$NaNO_2$, MPAA	62

Figure A.3

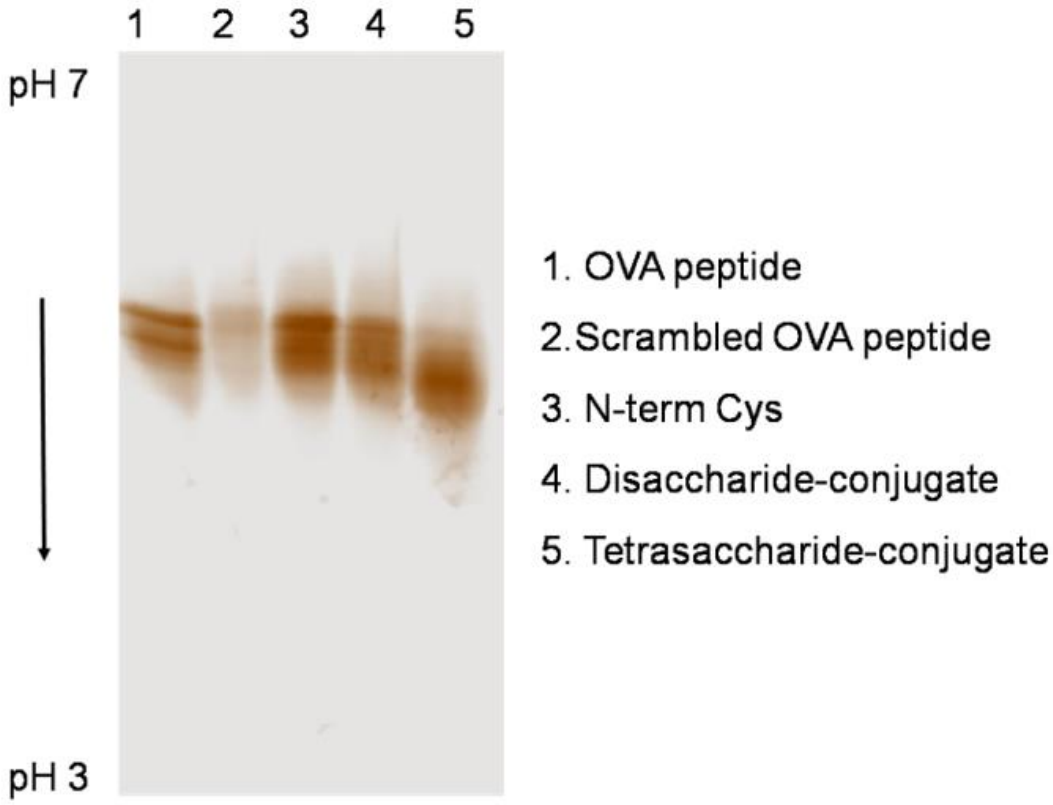


Figure A.3: NovexTM pH 3-7 protein IEF gel analysis of MHCII binding with peptide and glycopeptides visualized using silver stain.