BIOPHYSICAL CHARACTERIZATION OF CLASS II MAJOR HISTOCOMPATIBILITY COMPLEX (MHCII) MOLECULES

By

Jaspreet Kaur Osan, M.S.

A Dissertation Submitted in Partial Fulfillment of the Requirements

for the Degree of

Doctor of Philosophy

in

Biochemistry

University of Alaska Fairbanks

May 2020

APPROVED:

Dr. Andrea Ferrante, Co-Chair
Dr. Thomas Kuhn, Co-Chair
Dr. Andrej Podlutsky, Committee Member
Dr. Jack Chen, Committee Member
Dr. Thomas Green, Chair
Department of Chemistry & Biochemistry
Dr. Kinchel C. Doerner, Dean
College of Natural Science & Mathematics
Dr. Michael Castellini,
Dean of the Graduate School

ABSTRACT

Class II Major Histocompatibility Complex (MHCII) molecules are transmembrane glycoproteins expressed on the surface of antigen-presenting cells (APCs). APCs engulf pathogens and digest pathogenic proteins into peptides, which are loaded onto MHCII in the MHCII compartment (MIIC) to form peptide-MHCII complexes (pMHCII). These pMHCII are then presented to CD4+ T cells on the surface of APCs to trigger an antigen-specific immune response against the pathogens. HLA-DM (DM), a non-classical MHCII molecule, plays an essential role in generating kinetically stable pMHCII complexes which are presented to CD4+ T cells. When a few peptides among the pool of the peptide repertoire can generate the efficient CD4+ T cell response, such peptides are known as immunodominant. The selection of immunodominant epitopes is essential to generate effective vaccines against pathogens. The mechanism behind immunodominant epitope selection is not clearly understood. My work is focused on investigating various factors that help in the selection of immunodominant epitopes. For this purpose, peptides derived from H1N1 influenza hemagglutinin protein with known CD4+ T cell responses have been used. We investigated the role of DM-associated binding affinity in the selection of immunodominant epitopes. Our analysis showed that the presence of DM significantly reduces the binding affinity of the peptides with low CD4+ T cell response and inclusion of DM-associated IC_{50} in training MHCII algorithms may improve the binding prediction. Previous studies have shown that there is an alternate antigen presentation depending on antigen protein properties. Here, we showed that the immunodominant epitope presentation is dependent on the pH and length of the peptides. To study the MHCII in its native form, we assembled full-length MHCII in a known synthetic membrane model known as nanodiscs. We noted that, based on the lipid composition, assembly of the MHCII differs. Preliminary binding

studies with this tool showed that there might be a difference in the binding based on the type of the nanodisc. Collectively, our results showed that the immunodominant epitope selection is a complex process that is driven by various biochemical features.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ABBREVIATIONS

MHCII	Class II Major Histocompatibility Complex
MIIC	Class Major histocompatibility complex
APCs	Antigen Presenting Cells
DM	HLA-DM
pMHCII	peptide-MHCII complexes
CLIP	Class II-associated invariant chain peptide
CD4+ T cells	Helper T cells
ER	Endoplasmic reticulum
DR1	HLA-DR1 (One of the most common MHCII allele)
HA	Hémagglutinine protein
IC_{50}	Inhibition Constant
K_d	Equilibrium dissociation constant
bioHA	Biotinylated HA
IEDB	Immune Epitope Database
AUC	Area Under the Curve
ROC	Receiver Operating Characteristics
MWCO	Molecular weight cut off
HEK	Human Embryonic Kidney
SDS	Sodium Dodecyl Sulphate
BME	Beta mercaptoethanol
PBS	Phosphate Buffer Saline

FPLC	Fast Protein Liquid Chromatography
mDR1	Full length membrane DR1
MSP1E3D1	Membrane Scaffold Protein 1
POPC	1-palmitoyl-2-oleoyl-glycero-3-phosphocholine
PSM	Sphingomyelin
PAGE	Polyacrylamide Gel Electrophoresis
NBT	Nitro-blue tetrazolium chloride
BCIP	5-bromo-4-chloro-3'-indolyphosphate p-toluidine salt
ELISA	Enzyme-linked Immunosorbent Assay
L243	Antibody against HLA-DR1 allele

ACKNOWLEDGMENT

Education is an important part of a person life. It shapes one's personality, intellectual and prepare them for their future. I was never certain towards my career goals as growing up, but it was in my high school that a science fair at national level sealed my love and interest in research. I had never thought to move out of my country and come to United States to do my PhD. Although US was a new environment and a completely new world University of Alaska Fairbanks soon became my home away from home. I am thankful to Dr. Andrea Ferrante to give me an opportunity to work in his lab. I could not ask for any better mentor than him. I am extremely grateful for his guidance even after his move to Eli Lily. His constant encouragement was helpful to me specially to apply for fellowship and help me through writing my grants. I want to thank him to introducing me to the world of molecular immunology.

When Dr. Ferrante moved to Eli Lily in 2017, I at first thought that I would have to start from scratch but thanks to Dr. Thomas Kuhn who stepped in to fill in Dr. Ferrante shoes. I am grateful to Dr. Kuhn to not only guiding me as a Committee co-chair and mentor but also helping me when I needed a guidance in the absence of Dr. Ferrante. His expertise in protein biochemistry helped a lot in understanding my work. Dr. Kuhn and Dr. Ferrante have been an amazing mentor throughout my PhD. They have taught me how to become a scientist and how to become a mentor who support their mentee.

Research is never achieved by one person; it is a collaborative work and research is not possible without the guidance of expertise and support of lab members. I am thankful to my committee member Dr. Jack Chen and Dr. Andrej Podlutsky for their continuous support and guidance. They have provided me valuable feedback during committee meeting which shaped my PhD research. I also like to thank my collaborator at University of Montana Missoula Dr. Alexander

Ross who helped me in learning Fluorescence anisotropy. I am incredibly grateful to Dr. Harman Steele and Zifan Wang in his lab who helped me in processing my sample and for teaching me fluorescence anisotropy and analysis.

I am also thankful to Tynan Becker, my other half as she always says. We both started our PhD together. She has not only helped me during my research time in the lab, but she has also helped me personally to feel like Alaska as my second home.

I also like to thank Dr. Yong li, Melanie Roed and Margaret Castellini for their support as lab managers. They helped run lab smoothly and because of them we never have to worry about anything.

I am thankful to my undergrad Kristian Rivera whom I met in Fall 2018. She is a quick learner and very responsible. She helped me doing so many experiments when I was not able to do them. Research is so vast that sometimes you want to explore other things. During my PhD I met Dr. Stephanie Kennedy who was an environmental chemistry PhD student working on effect of mercury on Steller sea lion immune system. She wanted to check if mercury can bind to immune proteins. I helped her and her undergrad Roger Vang to process their sample on Isothermal titration calorimeter. I really enjoyed that little project. I am thankful to them to let me be a part of their project.

I like to thank all the past and current lab members of Kuhn lab who helped me during my PhD by one way or another. I am glad to have crossed path with you.

I also like to thank INBRE and Department of Biochemistry for providing me funding during my PhD. I am thankful to INBRE program as they have provided me an opportunity to learn how to write a competitive grant. I am thankful to Department of Chemistry for giving me a teaching assistant position during my first year. I really enjoyed teaching undergrad chemistry lab.

I would also like to thank my friends Dr. Saurabh Bhowmik, Neeraj Kulkarni, Priyam Sharma who provided me an emotional support during my journey.

At the end I like to thank the people without whom I never have decided to come to University of Alaska Fairbanks and pursue my PhD. First, I like to thank my parents Jaspal Singh Osan and Rajinder Kaur Osan and my brother Dilpreet Singh Osan for their unconditional love and encouragement to pursue my dreams. I am also thankful to my in-laws Jagdish Shetty and Bhavna Shetty for always supporting me. I also like to thank husband's uncle late Madhav Shah who introduced us to the education system in USA. My special thanks to my loving husband Madhur Shetty without his support I would have given up on my PhD at an early stage. He supported me on every step not only as a life partner but also as a mentor. Thank you for being there for me, I could not have done this without you. Special mention to my little angel, Trisha Osan Shetty, for bringing a new dimension to my life and a motivation towards finishing my PhD sooner than expected.

Chapter 1: Introduction

An introduction to the immune system

The Immune system is a complex system comprised of different branches with the primary function of protecting us from foreign pathogens and the potentially harmful effect of host invasion. When an organism encounters a pathogen, the immune system is activated and generates a response that tailors to the nature of the antigen. The Immune system could be divided broadly into two parts: Innate immunity and adaptive immunity. Innate immunity generates an immune response targeting primary infection and inflammation and acts as the first line of defense in the immune system (Bonilla & Oettgen, 2010). Studies have also shown that innate immunity is altered during sepsis, which makes hosts susceptible to secondary infections (Delano et al., 2011). Innate immunity does not have a memory, and with repeated encounters to the same pathogen, it generates comparable responses, and the same processes are triggered as soon as the infection occurs. Instead, adaptive immunity is developed over time, is specific for the antigen and modified upon repeated encounters, and finally establishes a long-lasting memory. Adaptive immunity involves crosstalk between antigen-presenting cells (APCs), T-cells, and B-cells.

Cells and molecules of the immune system

The immune system is made up of many cells and proteins, which helps in generating an immune response. All cell types originate from stem cells and classified into two main classes: Lymphoid stem cells (Lymphocytes) and myeloid stem cells (Granulocytes). Lymphocytes are further classified into three types: B-cells, T-cells, and Natural killer cells. Granulocytes are classified into macrophages, dendritic cells, neutrophil, eosinophil, basophil, and mast cells. Granulocytes cells, specialized in phagocytosis, engulf the whole pathogen and destroy them using digestive enzymes.

On the other hand, lymphocytes have more specialized roles. Natural killer cells induce apoptosis of infected cells. B-cells originate and mature in the bone marrow and express antibodies against specific antigens as well as present antigens to T-cells. T-cells are generated in bone marrow yet mature in the thymus. T-cells express unique receptors (T-cell receptors or TCR) on their surface along with CD (cluster of differentiation) glycoproteins. TCR is specific for the recognition of MHC bound antigen. T-cells are classified into two types based on CD glycoproteins: CD4+ T cells (helper T cells) – which recognize MHCII proteins and CD8+ T cells (Cytotoxic T cells) – which recognize MHCI proteins (Meuer et al., 1983). T cells generate an immune response by secreting cytokines, which are responsible for the differentiation of monocytes, macrophages. They are also responsible for the maturation of B-cells, antibody production, and activation of phagocytic cells.

Apart from immune cells, different types of immune proteins play an essential role in the regulation of immune response. The primary immune proteins are antibodies, cytokines, complement proteins, CD receptors, and MHC proteins. Antibodies are present on B-cells, which help in defense against bacteria and viruses, also develop memory against the pathogens. Cytokines are secreting proteins that act as a messenger to stimulate different cells and keep a check on the immune system. Complement proteins are a set of proteins that help in destroying pathogens by activating a set of cascade reactions. There are many CD molecules present in the immune system with a variety of functions. One of the functions of CD receptors are to act as cell surface receptors and help in antigen recognition. The list of the proteins mentioned here is by far not comprehensive considering the plethora of the proteins required for proper immune function. This thesis revolves around Major histocompatibility complex (MHC), a family of transmembrane glycoproteins present on the surface of APCs composed of MHCI, MHCII, and MHCIII. These MHC molecules

play a crucial role in the antigen presentation pathway. This thesis is focused on MHCII.

Overview of antigen presentation pathways

Antigen presentation is a process where the antigen is presented by antigen-presenting cells (APCs) to T-cells. There are two main antigen-presenting pathways – class I MHC and Class II MHC pathway (Neefjes, Jongsma, Paul, & Bakke, 2011). Class I MHC pathway is involved in presenting endogenous proteins. All nucleated cells express MHCI molecules and present proteins to CD8+ T cells. The proteins synthesized by a healthy cell are tolerated by CD8+ T cells. On the other hand, when a cell gets infected or undergoes mutations, the peptidome generated by them allows CD8+ T cells to detect them and help the immune system to destroy these infected or abnormal cells.

Class II MHC pathway present exogenous proteins (Figure 1.1). APCs express Class II MHC molecules (Hume, 1985). They take up the foreign pathogen, leading to the degradation of the pathogen proteins by the endosome. These protein fragments get presented to MHCII for binding. The selected peptide bound to MHCII gets transported to the surface of APCs, and it is here that these proteins are presented to CD4+ T cells. CD4+ T cells generate an immune response against this pathogen after the presentation of peptides generated from the pathogen.

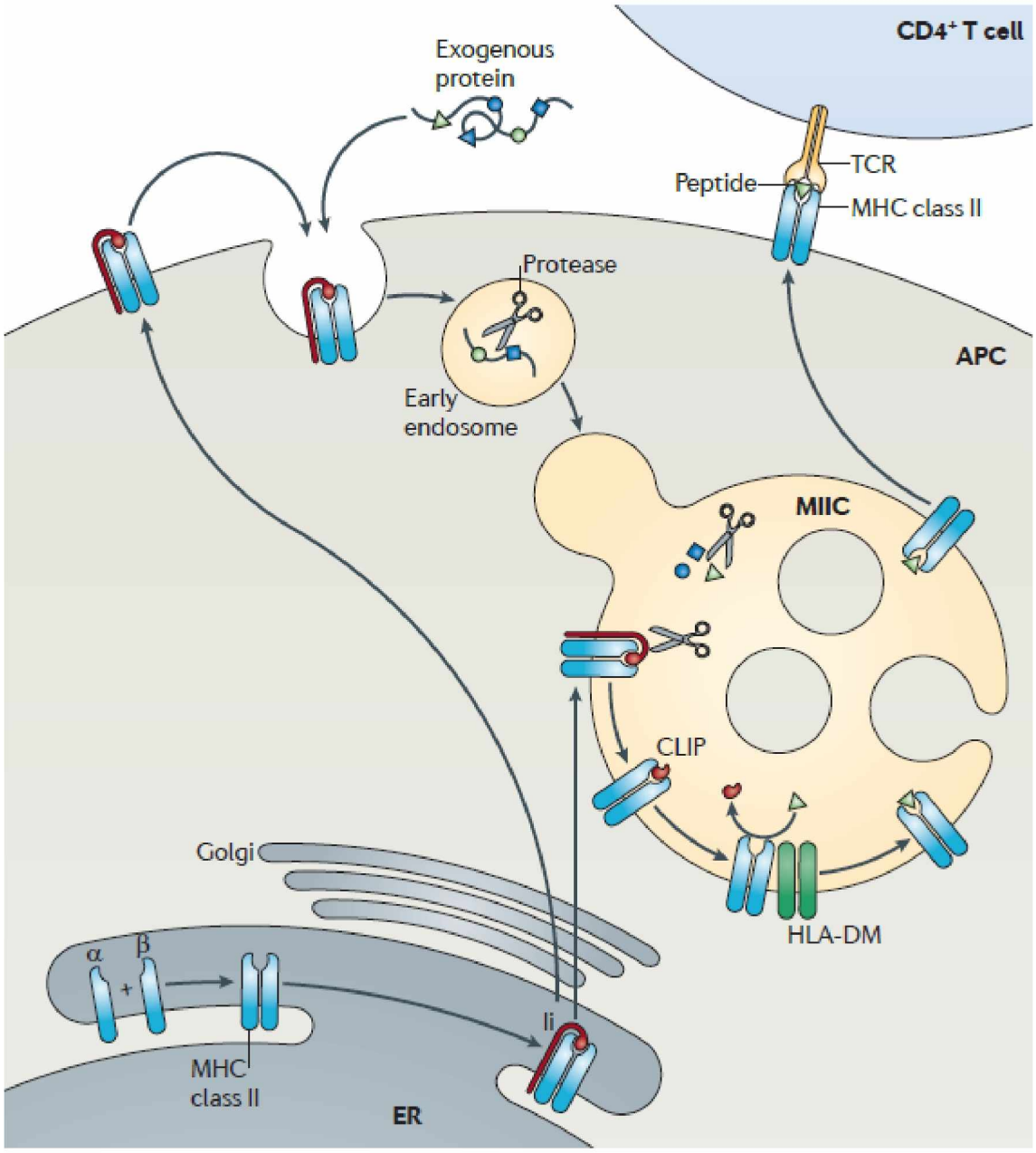


Figure 1.1: Class II MHC pathway: Newly synthesized MHCII molecules are transported via invariant chain to MIIC, directly or via cell membrane, where cathepsin cleaves the invariant chain to a shorter peptide CLIP, which is removed by non-classical MHCII molecule DM. The empty peptide binding site exchanges various peptides derived from antigen protein in the presence of DM. The selected pMHCII complex gets transported to the cell surface for the activation of CD4+ T cells (Neefjes et al., 2011).

MHC (Major histocompatibility class) gene

MHC gene is present on chromosome 6 in humans and represented as HLA (Human leukocyte antigen) (Stimpfling, 1971). There are three MHC genes: MHCI, MHCII, and MHCIII. For MHCI, there are three genes: HLA-A, HLA-B, and HLA-C. There are six genes for MHCII: HLA-DPA1, HLA-DPB1, HLA-DQA1, HLA-DQB1, HLA-DRA, and HLA-DRB1, while not much is known for the MHCIII gene other than it is involved in inflammation and other immune functions.

Class I MHC or endogenous pathway

Regularly, healthy proteins are degraded by the proteasome in the cells which get cleared from the system without generating an immune response since CD8+ T cells have tolerated such proteins. T-cells learn to differentiate between self and non-self-proteins so that the immune system doesn't produce an immune response against its protein. However, in the case of a productive mutation, or if the cell becomes infected, the protein fragments generated by the proteasome is recognized by CD8+ T cells. Cytoplasmic protein-derived peptides produced by the proteasome are further trimmed or degraded by cytosolic peptidase. Some fragments escape to the endoplasmic reticulum (ER) via a transporter embedded in the ER membrane known as transporter associated with antigen processing (TAP) (Solheim, Carreno, & Hansen, 1997). In the ER, the peptide loading complex (PLC) selects these fragments and facilitates it's complexing to MHCI, for presentation to CD8+ T cells. The PLC is comprised of MHCI, ERp57, Calreticulin and tapasin. As the peptide-binding site of MHCI can accommodate peptides 8-10 amino acid long, the peptide transported via TAP might require further trimming by the ER aminopeptidase associated with antigen (ERAAP). In the ER, MHCI is partially folded and stabilized by calreticulin and ERp57 (Farmery, Allen, Allen, & Bulleid, 2000).

Tapasin acts as a peptide editor and helps in accelerating the exchange of the peptides generated by ERAAP and transported by TAP. Once the PLC produces an efficient peptide-MHCI complex, this is carried to the cell surface to interact with CD8+ T cells (Solheim, 1999).

Class II MHC or exogenous pathway

MHCII molecules present peptides that derive from extracellular proteins or self-proteins generated by the endosomal pathway, as shown in Figure 1.1. MHCII unit assembles in the ER along with the invariant chain, which shields the peptide binding site from peptides present in the ER and directs the trafficking of MHCII molecules to the MHCII compartment (MIIC). In MIIC, the invariant chain is cleaved, leaving only a smaller peptide fragment known as class II-associated invariant chain peptide (CLIP) bound to the groove. HLA-DM (DM), a non-classical MHCII molecule, removes the CLIP peptide and makes the peptide binding site accessible to peptides derived from extracellular fragments. DM also helps in accelerating the peptide exchange and in the selection of kinetically stable peptide-MHCII (pMHCII) complexes (Kropshofer et al., 1996). pMHCII complexes then are transported to the cell surface where they interact with CD4+ T cells.

The research presented here focuses on understanding various factors that are required for MHC II-restricted peptide selection.

Immunodominance

The phenomenon of immunodominance is associated with the T-cell response. For a protein antigen, T-cell response is specific to only a few of the panoply of peptides generated by endosomal digestion. These peptides responsible for driving the antigen-specific T-cell responses are known as immunodominant. Immunodominance is found in class I and class II MHC, but it is mostly studied in MHCII. The need for immunodominant epitope is due to the limited space of

memory cells in lymph nodes (Berkower, Kawamura, Matis, & Berzofsky, 1985). So, to reduce the number of memory cells for the same antigen, only a few epitopes from that antigen generate a T-cell specific response.

Immunodominant epitope selection is a complex process that depends on various factors, such as peptide affinity for MHCII, susceptibility to specific proteases, structural features of antigens, and T cell receptor affinity for pMHCII complexes. There are currently two main hypotheses for immunodominant epitope selection: 1. epitope accessibility, and 2. kinetic stability (Kim & Sadegh-Nasseri, 2015).

Epitope accessibility presumes that immunodominant epitopes must be accessible to MHC binding sites or cathepsins cleavage. The support for this theory comes from the studies which have shown that the position of many known immunodominant epitope is near C- or N- termini of the protein or its flexible strands (Dai, Steede, & Landry, 2001). Instead, the kinetic stability model defines immunodominant epitopes based on their ability to form stable complexes with MHCII. Support for this model comes from the studies which show that high-affinity complexes show DM resistance (Lazarski et al., 2005).

Structural characteristics of MHC

MHCI and MHCII proteins are quite similar in structure. They have a membrane-proximal immunoglobulin-like domain and a membrane distal peptide binding site which comprises of eight-stranded beta-sheet and two-alpha helical region. The polymorphism mostly resides in the binding site. The significant difference in the structure is at the peptide-binding region, in that MHCI has a closed-groove peptide binding site that can accommodate a peptide residue of 8-10 amino acid long (Wilson & Fremont, 1993), while MHCII has an open-groove binding site due to which the peptide length is not restricted. The primary reason for differences in the epitope

mapping between MHCI and MHCII is due to their peptide-binding site. As MHCI has a limitation on amino-acid residues binding to the pockets lining the groove, it is relatively easier to map binding-motifs specific to different alleles while this is more difficult in the case of MHCII, as the alignment of amino-acid residues for longer peptides creates more probability of side-chain positioning to each binding pocket (Chaves, Lee, Nayak, Richards, & Sant, 2012). The first crystal structure of MHCII (HLA-DR1 allele) with $HA_{306-318}$ shows that peptide binds in a polyproline helix conformation. The peptide binds as an extended straight strand where N and C termini projected out of the binding pocket. In the MHCII binding site, various pockets are numbered as P1 to P10. These pockets refer to the position of the amino-acid chain of the peptide which will reside in that specific pocket. The lining of these pockets is highly polymorphic, causing the allele dependent peptide specificity. P1, P4, P6, P7, and P9 accommodate the side chain of the peptide which are crucial in determining the interaction between MHCII and peptide, as shown in Figure 1.2. The P1 pocket is the deepest pocket that accommodates mainly hydrophobic side chains near the peptide N terminus. Peptide binding to MHCII relies primarily on the H-bond network and hydrophobic interactions. H-bond network is established between the main chain of the peptide to the central and conserved side chain residues of MHCII, as shown in Figure 1.3. The alpha-chain residues which are involved in H-bond is α-53, α-9, α-62, α-69 and α-76. The beta-chain residues are β-81, β-82, β71, β-62, and β-57 (Painter & Stern, 2012).

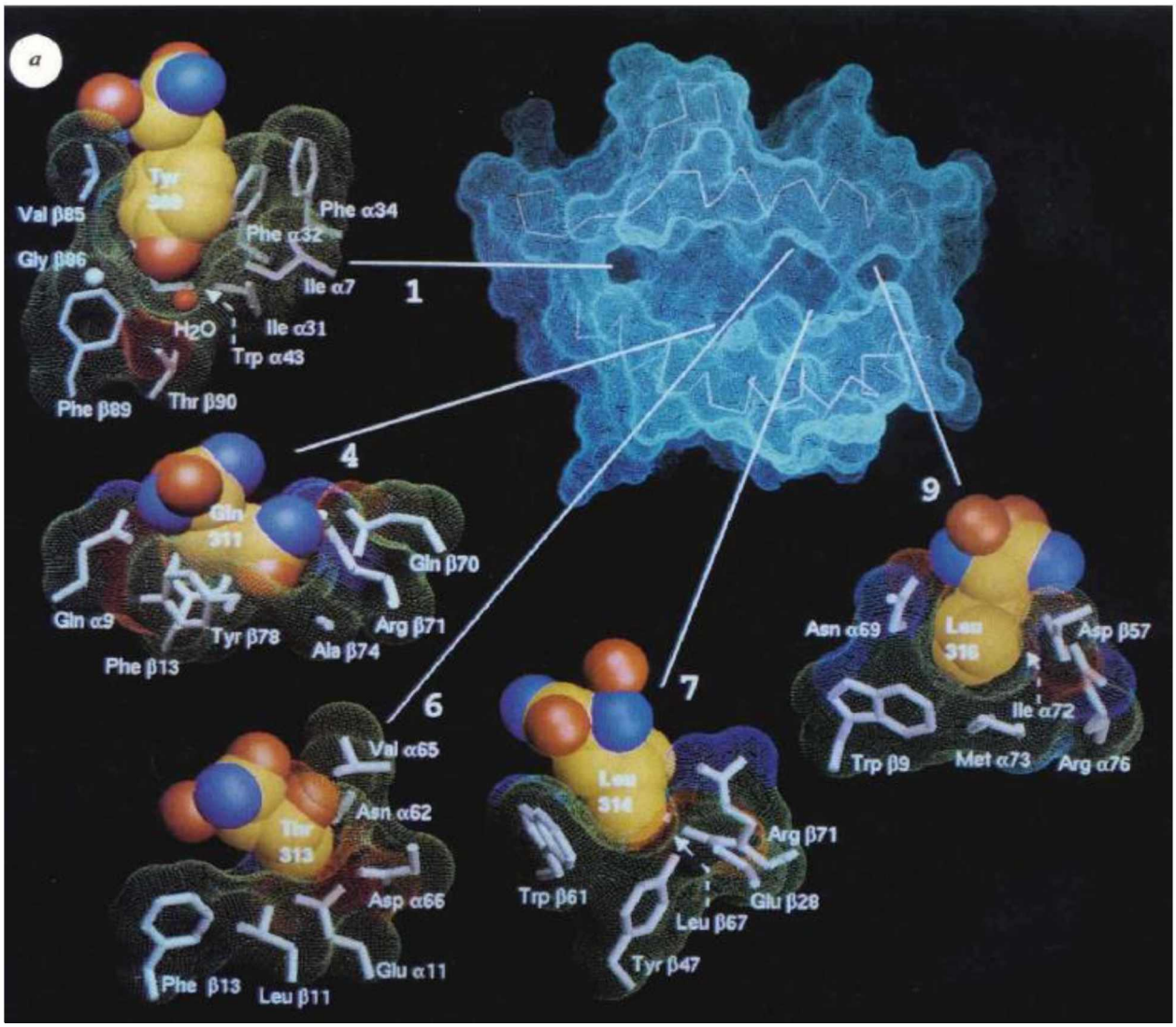


Figure 1.2: The major pocket P1, P4, P6, P7 and P9 in HLA-DR1, which plays crucial role in peptide interaction with MHCII (Stern et al., 1994).

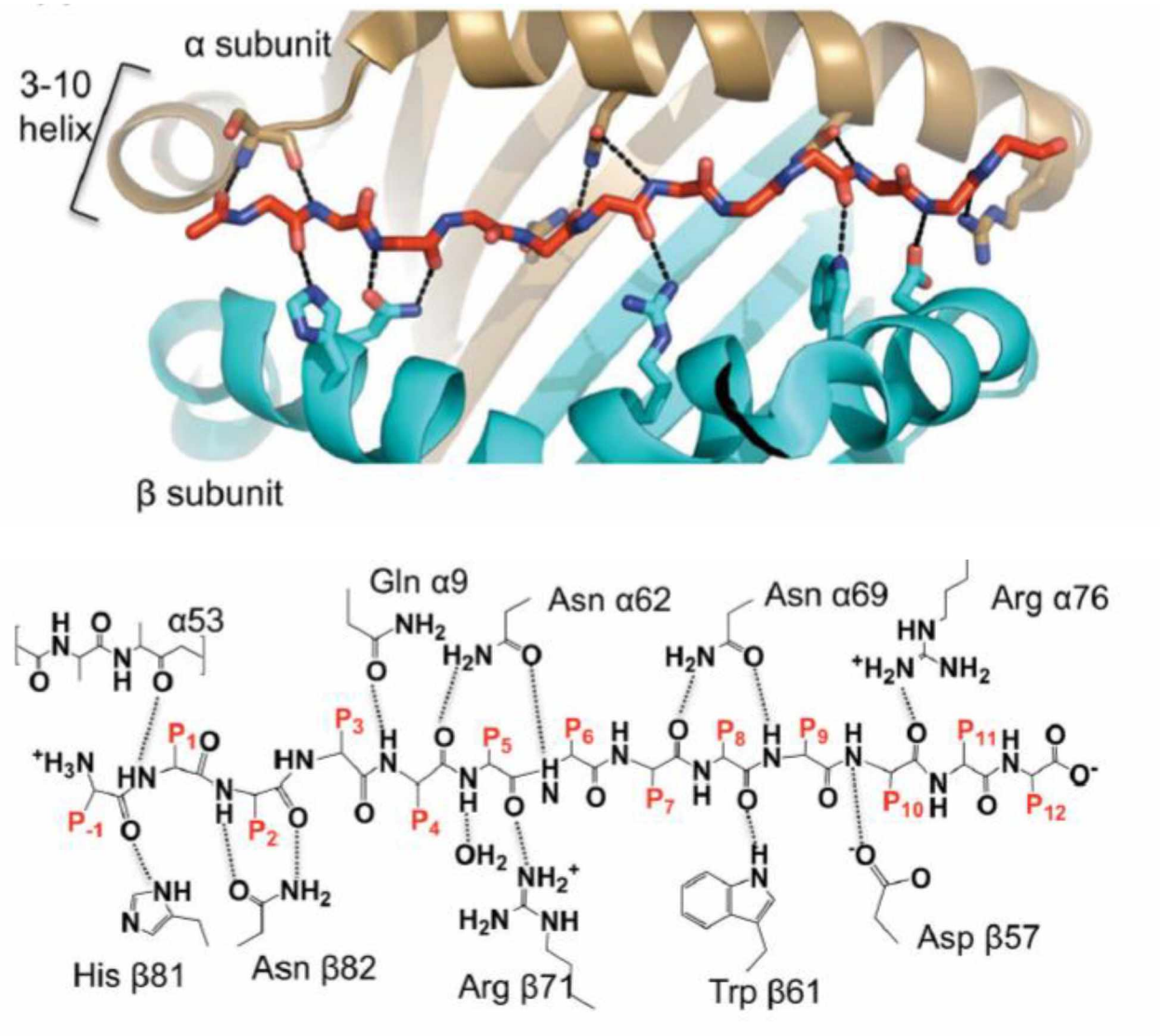


Figure 1.3: Hydrogen bond network between peptide $HA_{306-318}$ and MHCII binding site. Peptide is shown as stick representation. Bottom figure shows chemical diagram, with conserved residue of MHCII which contribute to H-bonding, with position of peptide side chain mentioned in red (Painter & Stern, 2012).

HLA-DM : non-classical MHCII molecule

DM is a non-classical MHCII molecule which itself does not present epitopes but contributes to the selection of the peptides, which presents to CD4+ T cells. A newly synthesized MHCII molecule binding site is covered by the invariant chain, which helps in the transportation of MHCII to MIIC and shields the peptide binding site from the endogenous peptides. In MIIC, invariant chain is cleaved to a shorted peptide known as CLIP. DM removes the CLIP peptide

from the binding site and enables the binding of the antigenic peptide to MHCII (Kropshofer et al., 1996; Stebbins, Peterson, Suh, & Sant, 1996). Studies have shown that DM also accelerates the exchange of peptides and helps in the selection of kinetically stable pMHCII complexes. The mechanism behind DM action is still poorly understood. Various studies are proposing different mechanisms via which DM helps in the selection of immunodominant epitopes. Studies have shown that the rate of peptide exchange is directly proportional to intrinsic dissociation rate of pMHCII complexes (Weber, Evavold, & Jensen, 1996). These studies supported the kinetic stability model of immunodominant selection. DM role is not only limited to accelerating the exchange of the peptides, but it has also shown that DM can act as MHCII-chaperone and stabilize the empty MHCII at low pH (Denzin, Hammond, & Cresswell, 1996; Kropshofer, Arndt, Moldenhauer, Hammerling, & Vogt, 1997). To further understand the DM mechanism of action, the crystal structure of DM along with DR1-HA complex molecule was solved. Pos et al. manipulated the N terminal site near P1 pocket to achieve this complex, as DM does not bind to DR1 when binding site occupies by covalently linked peptide (Pos et al., 2012). As P1 pocket plays an important for HA binding to DR. The structure showed that αW43 residue of the P1 pocket of DR1 plays a vital role in DM binding. This residue stabilizes the P1 pocket in the DR1-HA complex, but in DM-DR1-HA complex it is rotated away from P1 pocket. The indole ring of αW43 forms an H-bond with DM αN125. Mutation of αN125 to alanine completely abolishes the DM activity. So, based on this observation, the model they proposed for DM activity is that occlusion of P1 pocket by high-affinity peptides makes them resistant to DM activity (Pos et al., 2012).

Another model that is widely accepted is that DM activity is dependent on the conformation of the MHCII. It has shown that low-affinity peptides form a less rigid complex with MHCII, which

makes them DM susceptible. There is a specific region that is involved in MHCII conformational change. One study showed that substitution of αF54C in the DR1 molecule makes it sensitive to DM activity as compared to wild-type DR1, even when a high-affinity peptide is bound. The structure revealed a reorientation in the α45-50 region and changes in the flanking extended region α39-44 and α51-54. These regions are involved in DM interaction (Painter et al., 2011). DM activity is also dependent on MHCII polymorphism. It has been shown that there are alleles that feature DM-independent antigen presentation. One such type allele is DQ, and it can easily attain SDS stable conformation with Ii alone while DR alleles stability depends on Ii as well as DM.

Apart from dependence on MHCII allele specificity and structural conformation, another essential factor that drives DM activity is pH. Indeed, DM activity is optimum at pH 4.5 – 5.5 and reduces at neutral pH. The effect of pH on DM activity is related to the protonation of the proteins. The non-polar region of DR and DM plays a vital role in their interaction. In acidic conditions, the non-polar patch on both the proteins are usually exposed as compared to neutral pH. CD and fluorescence spectra have also shown that DM structure is different at pH 5 and 7. The work presented here investigates biochemical and biophysical factors that can impact DM function.

Forces responsible for peptide binding to MHCII

Significant efforts have been put forth to identify the biochemical and biophysical features of a pMHCII complex that makes it immunodominant. Historically, the two main biochemical characteristics that have been studied are binding affinity and kinetic stability. One report has extensively studied all the biochemical features associated with pMHCII for epitope selection. They studied IC_{50} (binding affinity), intrinsic dissociation half-life, and DM-mediated

dissociation half-life for peptides derived from the vaccinia virus with known DR1 epitopes. They showed that DM-associated half-life is an essential factor in selecting immunogenic epitope (Yin, Calvo-Calle, Dominguez-Amorocho, & Stern, 2012). There are other reports which showed a direct correlation between peptide intrinsic half-life and immunogenicity, stating immunodominant peptide having half-lives >150 hrs and cryptic peptide with less than 10 hrs (Lazarski et al., 2005). Another important feature that contributes to epitope selection is peptide binding to MHCII. Peptide binding to MHCII fundamentally involves encapsulation of the hydrophobic side chain in polymorphic pockets lining the binding site, and the establishment of an extensive H-bond network between side chains of non-polymorphic MHC residues and the backbone of the peptide. There are nine major positions in the binding site of MHCII, indicated with P1 through P9. The major pockets that are involved in the peptide interaction are P1, P4, P6, P7, and P9 (Stern et al., 1994). Particularly for DR alleles, and for other class II human and murine MHCII, the binding specificity of P1 to P9 positions have been extensively studied. The P1 pocket as mentioned above plays a vital role in stable peptide binding as well as in DM activity. This is the most hydrophobic pocket and prefers accepting large hydrophobic side chains (Trp, Tyr, Phe, Leu, and Ile). P4, P6, P7 are shallower pockets. P4 binds to large aliphatic side chains, P6 preferred smaller residues like Threonine. P9 is the deep pocket on the C-term side of the complex and prefers hydrophobic residues. One report has shown that the P10 positions can significantly contribute to binding on a residue-by-residue and peptide-by-peptide basis (Zavala-Ruiz, Strug, Anderson, Gorski, & Stern, 2004). Another rarely studied biochemical feature of pMHCII complex is its thermodynamic feature. Ferrante et al showed a correlation between structural and entropic component of pMHCII complexes. They have shown that flexible complexes with greater entropic penalty shows susceptibility to DM action (Ferrante,

Templeton, Hoffman, & Castellini, 2015). Thus, biochemical and biophysical characteristics of pMHCII complexes play a significant role in the selection of immunodominant epitope as determined by their DM susceptibility.

Immunodominant epitope prediction

MHCII presentation pathway generates a peptide repertoire for a pathogen, among which only a few determine the CD4+ T cell response against that pathogen. Those peptides are immunodominant epitopes. Engineering peptide-based vaccines require the knowledge of the pathogen-specific immunodominant epitopes, which can be used in lieu of the full pathogen or relevant recombinant antigens to induce a cellular and humoral memory. Indeed, the standard approach for epitope discovery involves isolation of pathogenic protein, constructing its protein and its fragments, measuring its binding affinity, or testing with CD4+ T cells to regulate the immune response. This approach is evidently time- and labor-intensive. To overcome the limitations of the classical epitope discovery approach, various epitope predictive algorithms have been generated (Bian, 2003; Nielsen, Lund, Buus, & Lundegaard, 2010; Nielsen et al., 2008; Nielsen, Lundegaard, & Lund, 2007). These algorithms are trained based on a database collecting experimentally validated epitopes. They used different computational training methods such as neural networks, multivariate statistical analysis, and consensus (Bisset & Fierz, 1993; Burden & Winkler, 2005). The prediction considers peptide binding affinity fundamentally to MHCII as a proxy for immunogenicity. The crystal structure of pMHCII has shown that the primary force of binding is peptide side chain binding to major pockets of MHCII; the algorithms investigate the amino-acid interaction which can fit into these pockets. The training dataset helps in refining this prediction, and the algorithm predicts the 9-core amino-acid which will fit in the binding site. This approach is useful for class I, but for class II this approach has

various limitations. One major limitation is due to the difference in the binding groove of these two molecules. Class I has a close groove while class II has an open groove; thus, it can bind peptides varying in length from 9-25 amino acids. For longer peptides, identifying the amino acids driving MHCII binding and those causing T-cell recognition is not known. Studies have also shown that the flanking residues also play an essential role in CD4+ T cell interaction (Holland, Cole, & Godkin, 2013). Hence, the epitope for class II alone is not enough to train the algorithm. The training dataset is generally only focused on the binding dataset or limited available CD4+ T cell data. The presentation of the epitope is such a complex process that involves several factors such as DM activity, cathepsin activity, CD4+ T cell interaction points, and allele dependence. Thus, training algorithms based only on binding datasets and limited epitope data is not enough. It is essential to consider all factors to generate efficient algorithms.

Significance of MHCII assembly in the membrane

MHCII are transmembrane glycoproteins which are assembled as a heterodimer in the membrane. MHCII molecule is comprised of α and β chain with membrane proximal, distal, and transmembrane portion (Stern et al., 1994). In the ER, a newly synthesized molecule transported as a nonameric complex along with an invariant chain as a chaperone to MIIC. MHCII molecule has also shown the tendency of dimerization of heterodimer (superdimers) in the soluble form. Its crystal structure has shown the dimer of αβ heterodimers arranged in parallel fashion, suggesting its tendency to dimerize associated with CD4+ T cell recognition (Brown et al., 2015; Cochran & Stern, 2000; Schafer, Malapati, Hanfelt, & Pierce, 1998). Studies supporting this hypothesis have shown that monomers are not stimulatory for T cells. Trimeric or tetrameric agonist ligand shows better TCR stimulation (Boniface et al., 1998). One study demonstrated that this dimer of dimers could stimulate CD4+ T cells for low-affinity antigens much better than

high-affinity antigens (Schafer & Pierce, 1994). The single-particle imaging study has also shown the presence of dimer of dimers in living cells, which are more prevalent at 22°C as compare to 37°C, showing its dependence on temperature and lipid environment (Cherry et al., 1998). The empty MHCII has the inherent property to aggregate in the absence of peptide and binding of the peptide stabilize it against aggregation (Stern & Wiley, 1992). The arrangement of MHCII in the membrane and tendency of purified MHCII to aggregate as dimer of dimers suggest that the arrangement of MHCII plays an essential role in the stabilization of MHCII and its function. In the early and late '90s, these superdimers were studied extensively regarding their role in CD4+ T cell activation. How they will impact the binding of the antigen has not been well-studied. We attempted to study the assembly of MHCII in model membranes of various lipid compositions and assess whether these model membranes are suitable for peptide binding studies.

The big picture

MHCII antigen presentation required various factors to select immunodominant epitope for presentation to CD4+ T cells, which generate an efficient response against a specific pathogen. Our work focuses on investigating different factors that are involved in immunodominant epitope selection. DM role in the selection of immunodominant epitope is extensively studied from the kinetic stability of the pMHCII perspective, but we investigated its role from binding affinity point of view. The epitope predictive algorithms are focused on binding affinity data without DM and epitope identification through T cell analysis. We have shown that DM-associated binding affinity is a proxy for immunodominance and has better predictive property. Next, to understand the DM mechanism of action, we have studied the structural features of the pMHCII complex. We have also considered the impact of pH on the kinetic stability of pMHCII and its

association with DM activity. In the end, we incorporate a full-length MHCII protein in the synthetic membrane to study the impact of lipid composition on its assembly. Studying these various biochemical features provide us significant insight into the role of DM and assembly of MHCII in the selection of immunodominant epitope.

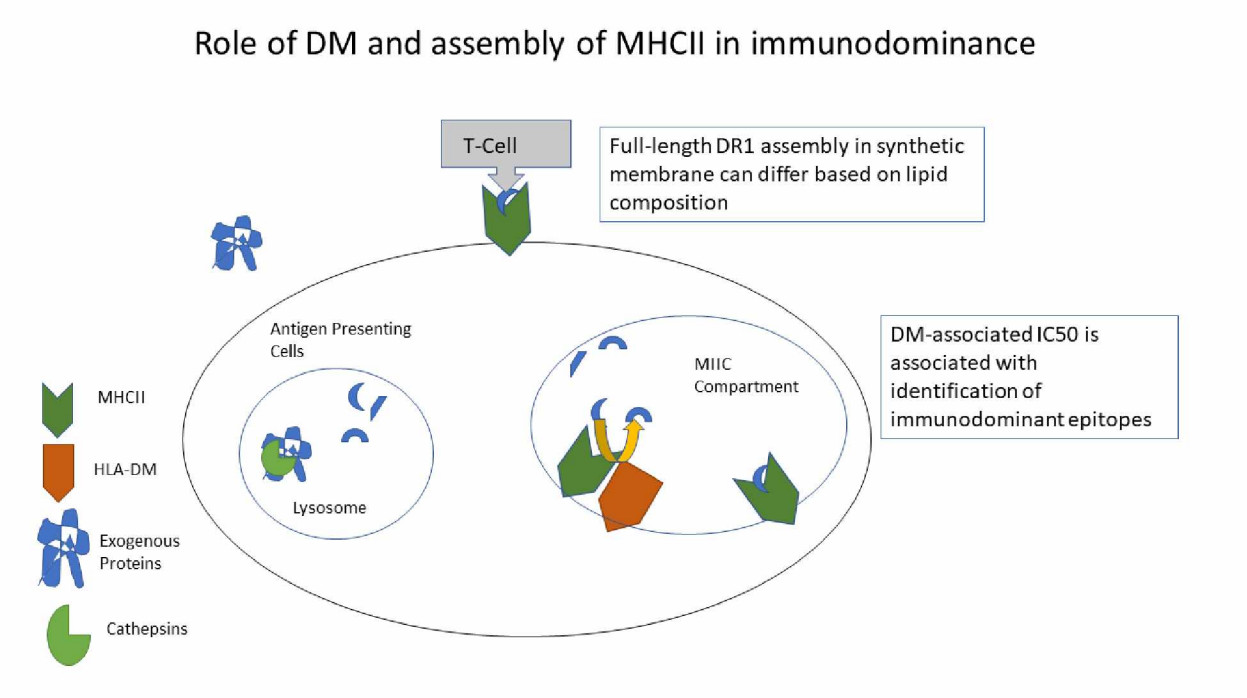


Figure 1.4: The overview of the research work presented in this thesis

References

Berkower, I., Kawamura, H., Matis, L. A., & Berzofsky, J. A. (1985). T cell clones to two major T cell epitopes of myoglobin: effect of I-A/I-E restriction on epitope dominance. *J Immunol, 135*(4), 2628-2634.

Bian, H. (2003). The use of bioinformatics for identifying class II-restricted T-cell epitopes. *Methods, 29*(3), 299-309. doi:10.1016/s1046-2023(02)00352-3

Bisset, L. R., & Fierz, W. (1993). Using a neural network to identify potential HLA-DR1 binding sites within proteins. *J Mol Recognit, 6*(1), 41-48. doi:10.1002/jmr.300060105

Boniface, J. J., Rabinowitz, J. D., Wulfing, C., Hampl, J., Reich, Z., Altman, J. D., . . . Davis, M. M. (1998). Initiation of signal transduction through the T cell receptor requires the multivalent engagement of peptide/MHC ligands [corrected]. *Immunity, 9*(4), 459-466.

Bonilla, F. A., & Oettgen, H. C. (2010). Adaptive immunity. *J Allergy Clin Immunol, 125*(2 Suppl 2), S33-40. doi:10.1016/j.jaci.2009.09.017

Brown, J. H., Jardetzky, T. S., Gorga, J. C., Stern, L. J., Urban, R. G., Strominger, J. L., & Wiley, D. C. (2015). Pillars article: three-dimensional structure of the human class II histocompatibility antigen HLA-DR1. Nature. 1993. 364: 33-39. *J Immunol, 194*(1), 5-11.

Burden, F. R., & Winkler, D. A. (2005). Predictive Bayesian neural network models of MHC class II peptide binding. *J Mol Graph Model, 23*(6), 481-489. doi:10.1016/j.jmgm.2005.03.001

Chaves, F. A., Lee, A. H., Nayak, J. L., Richards, K. A., & Sant, A. J. (2012). The utility and limitations of current Web-available algorithms to predict peptides recognized by CD4 T cells in response to pathogen infection. *J Immunol, 188*(9), 4235-4248. doi:10.4049/jimmunol.1103640

Cherry, R. J., Wilson, K. M., Triantafilou, K., O'Toole, P., Morrison, I. E., Smith, P. R., & Fernandez, N. (1998). Detection of dimers of dimers of human leukocyte antigen (HLA)-DR on the surface of living cells by single-particle fluorescence imaging. *J Cell Biol, 140*(1), 71-79. doi:10.1083/jcb.140.1.71

Cochran, J. R., & Stern, L. J. (2000). A diverse set of oligomeric class II MHC-peptide complexes for probing T-cell receptor interactions. *Chem Biol, 7*(9), 683-696.

Dai, G., Steede, N. K., & Landry, S. J. (2001). Allocation of helper T-cell epitope immunodominance according to three-dimensional structure in the human immunodeficiency virus type I envelope glycoprotein gp120. *J Biol Chem, 276*(45), 41913-41920. doi:10.1074/jbc.M106018200

Delano, M. J., Thayer, T., Gabrilovich, S., Kelly-Scumpia, K. M., Winfield, R. D., Scumpia, P. O., . . . Moldawer, L. L. (2011). Sepsis induces early alterations in innate immunity that impact mortality to secondary infection. *J Immunol, 186*(1), 195-202. doi:10.4049/jimmunol.1002104

Denzin, L. K., Hammond, C., & Cresswell, P. (1996). HLA-DM interactions with intermediates in HLA-DR maturation and a role for HLA-DM in stabilizing empty HLA-DR molecules. *J Exp Med, 184*(6), 2153-2165. doi:10.1084/jem.184.6.2153

Farmery, M. R., Allen, S., Allen, A. J., & Bulleid, N. J. (2000). The role of ERp57 in disulfide bond formation during the assembly of major histocompatibility complex class I in a synchronized semipermeabilized cell translation system. *J Biol Chem, 275*(20), 14933-14938. doi:10.1074/jbc.275.20.14933

Ferrante, A., Templeton, M., Hoffman, M., & 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*(3), 1251-1261. doi:10.4049/jimmunol.1402367

Holland, C. J., Cole, D. K., & Godkin, A. (2013). Re-Directing CD4(+) T Cell Responses with the Flanking Residues of MHC Class II-Bound Peptides: The Core is Not Enough. *Front Immunol, 4*, 172. doi:10.3389/fimmu.2013.00172

Hume, D. A. (1985). Immunohistochemical analysis of murine mononuclear phagocytes that express class II major histocompatibility antigens. *Immunobiology, 170*(5), 381-389. doi:10.1016/S0171-2985(85)80062-0

Kim, A., & Sadegh-Nasseri, S. (2015). Determinants of immunodominance for CD4 T cells. *Curr Opin Immunol, 34*, 9-15. doi:10.1016/j.coi.2014.12.005

Kropshofer, H., Arndt, S. O., Moldenhauer, G., Hammerling, G. J., & Vogt, A. B. (1997). HLA-DM acts as a molecular chaperone and rescues empty HLA-DR molecules at lysosomal pH. *Immunity, 6*(3), 293-302. doi:10.1016/s1074-7613(00)80332-5

Kropshofer, H., Vogt, A. B., Moldenhauer, G., Hammer, J., Blum, J. S., & Hammerling, G. J. (1996). Editing of the HLA-DR-peptide repertoire by HLA-DM. *EMBO J, 15*(22), 6144-6154.

Lazarski, C. A., Chaves, F. A., Jenks, S. A., Wu, S., Richards, K. A., Weaver, J. M., & Sant, A. J. (2005). The kinetic stability of MHC class II:peptide complexes is a key parameter that dictates immunodominance. *Immunity, 23*(1), 29-40. doi:10.1016/j.immuni.2005.05.009

Meuer, S. C., Hodgdon, J. C., Cooper, D. A., Hussey, R. E., Fitzgerald, K. A., Schlossman, S. F., & Reinherz, E. L. (1983). Human cytotoxic T cell clones directed at autologous virus-transformed targets: further evidence for linkage of genetic restriction to T4 and T8 surface glycoproteins. *J Immunol, 131*(1), 186-190.

Neefjes, J., Jongsma, M. L., Paul, P., & Bakke, O. (2011). Towards a systems understanding of MHC class I and MHC class II antigen presentation. *Nat Rev Immunol, 11*(12), 823-836. doi:10.1038/nri3084

Nielsen, M., Lund, O., Buus, S., & Lundegaard, C. (2010). MHC class II epitope predictive algorithms. *Immunology, 130*(3), 319-328. doi:10.1111/j.1365-2567.2010.03268.x

Nielsen, M., Lundegaard, C., Blicher, T., Peters, B., Sette, A., Justesen, S., . . . Lund, O. (2008). Quantitative predictions of peptide binding to any HLA-DR molecule of known sequence: NetMHCIIpan. *PLoS Comput Biol, 4*(7), e1000107. doi:10.1371/journal.pcbi.1000107

Nielsen, M., Lundegaard, C., & Lund, O. (2007). Prediction of MHC class II binding affinity using SMM-align, a novel stabilization matrix alignment method. *BMC Bioinformatics, 8*, 238. doi:10.1186/1471-2105-8-238

Painter, C. A., Negroni, M. P., Kellersberger, K. A., Zavala-Ruiz, Z., Evans, J. E., & Stern, L. J. (2011). Conformational lability in the class II MHC 310 helix and adjacent extended strand dictate HLA-DM susceptibility and peptide exchange. *Proc Natl Acad Sci U S A, 108*(48), 19329-19334. doi:10.1073/pnas.1108074108

Painter, C. A., & Stern, L. J. (2012). Conformational variation in structures of classical and non-classical MHCII proteins and functional implications. *Immunol Rev, 250*(1), 144-157. doi:10.1111/imr.12003

Pos, W., Sethi, D. K., Call, M. J., Schulze, M. S., Anders, A. K., Pyrdol, J., & Wucherpfennig, K. W. (2012). Crystal structure of the HLA-DM-HLA-DR1 complex defines mechanisms for rapid peptide selection. *Cell, 151*(7), 1557-1568. doi:10.1016/j.cell.2012.11.025

Schafer, P. H., Malapati, S., Hanfelt, K. K., & Pierce, S. K. (1998). The assembly and stability of MHC class II-(alpha beta)2 superdimers. *J Immunol, 161*(5), 2307-2316.

Schafer, P. H., & Pierce, S. K. (1994). Evidence for dimers of MHC class II molecules in B lymphocytes and their role in low affinity T cell responses. *Immunity, 1*(8), 699-707.

Solheim, J. C. (1999). Class I MHC molecules: assembly and antigen presentation. *Immunol Rev, 172*, 11-19. doi:10.1111/j.1600-065x.1999.tb01352.x

Solheim, J. C., Carreno, B. M., & Hansen, T. H. (1997). Are transporter associated with antigen processing (TAP) and tapasin class I MHC chaperones? *J Immunol, 158*(2), 541-543.

Stebbins, C. C., Peterson, M. E., Suh, W. M., & Sant, A. J. (1996). DM-mediated release of a naturally occurring invariant chain degradation intermediate from MHC class II molecules. *J Immunol, 157*(11), 4892-4898.

Stern, L. J., Brown, J. H., Jardetzky, T. S., Gorga, J. C., Urban, R. G., Strominger, J. L., & Wiley, D. C. (1994). Crystal structure of the human class II MHC protein HLA-DR1 complexed with an influenza virus peptide. *Nature, 368*(6468), 215-221. doi:10.1038/368215a0

Stern, L. J., & Wiley, D. C. (1992). The human class II MHC protein HLA-DR1 assembles as empty alpha beta heterodimers in the absence of antigenic peptide. *Cell, 68*(3), 465-477. doi:10.1016/0092-8674(92)90184-e

Stimpfling, J. H. (1971). Recombination within a histocompatibility locus. *Annu Rev Genet, 5*, 121-142. doi:10.1146/annurev.ge.05.120171.001005

Weber, D. A., Evavold, B. D., & Jensen, P. E. (1996). Enhanced dissociation of HLA-DR-bound peptides in the presence of HLA-DM. *Science, 274*(5287), 618-620. doi:10.1126/science.274.5287.618

Wilson, I. A., & Fremont, D. H. (1993). Structural analysis of MHC class I molecules with bound peptide antigens. *Semin Immunol, 5*(2), 75-80. doi:10.1006/smim.1993.1011

Yin, L., Calvo-Calle, J. M., Dominguez-Amorocho, O., & Stern, L. J. (2012). HLA-DM constrains epitope selection in the human CD4 T cell response to vaccinia virus by favoring the presentation of peptides with longer HLA-DM-mediated half-lives. *J Immunol, 189*(8), 3983-3994. doi:10.4049/jimmunol.1200626

Zavala-Ruiz, Z., Strug, I., Anderson, M. W., Gorski, J., & Stern, L. J. (2004). A polymorphic pocket at the P10 position contributes to peptide binding specificity in class II MHC proteins. *Chem Biol, 11*(10), 1395-1402. doi:10.1016/j.chembiol.2004.08.007

Chapter 2: HLA-DM associated binding affinity as a proxy of immunodominance[1]

Abstract

Binding of antigenic peptides to class II MHC molecules (MHCII) and the activity of the "editing" molecule HLA-DM (DM) on the resulting peptide-MHCII complexes are critical factors in the antigen presentation pathway. During processing, a panoply of antigen-derived peptides generated, of which only a few known as "immunodominant" are the focus of CD4+ T cell recognition. Different algorithms are available to predict the peptide-MHC binding and to identify CD4+ T cell epitopes within pathogenic proteins. Here we examined whether the inclusion of DM during peptide-MHCII binding events would facilitate identifying immunodominant epitopes. We used a competition binding assay to quantify the binding affinity of each peptide from a library of sequences covering the entire hemagglutinin protein from H1N1 influenza virus for which an HLA-DR1 restricted immunodominance hierarchy had been already established. Our data showed that the presence of DM significantly lowered the binding affinity of weaker epitopes but not that of immunodominant epitopes. Statistical analysis showed that DM-associated binding affinity of immunodominant was significantly different than weaker epitopes ($P = 0.0028$). Upon comparing binding affinity predicted by algorithms to the measured binding affinity, we observed that DM-associated binding affinity yields a better prediction. These findings indicate that DM-associated binding is a viable proxy of immunodominance and should be considered when designing algorithms for MHCII.

[1] Osan, J.K, Kuhn T. B, Ferrante, A. (In Prep) HLA-DM associated binding affinity as a proxy of immunodominance. Journal of Immunology

Introduction

Class II major histocompatibility complex (MHCII) molecules are glycoproteins expressed on the surface of antigen-presenting cells (APCs) and present peptides derived from foreign pathogens to CD4+ T cells, leading to a T-dependent immune response (Neefjes, Jongsma, Paul, & Bakke, 2011). APCs take up pathogens, or fragments thereof, through different mechanisms, and process them within endo-lysosomal compartments to be enzymatically cleaved into peptides. In one of these compartments named MHCII compartment (MIIC), peptides compete for binding to MHCII molecules. Peptide-MHCII complexes (pMHCII) are then shuttled to the plasma membrane for recognition by CD4+ T cells. The peptides, among the repertoire of peptides generated by endosomal digestion, which produce the most potent CD4+ T cell response, are known as immunodominant (Kim & Sadegh-Nasseri, 2015). Apart from the binding of MHCII to the peptide, another factor with a critical role in the selection of immunodominant peptides is HLA-DM (DM), a non-classical MHCII molecule which itself does not bind to peptides but helps in the selection of pMHCII for presentation to CD4+ T cells (Ferrante, Anderson, Klug, & Gorski, 2008). When a newly synthesized MHCII molecule is generated in the endoplasmic reticulum (ER), its peptide-binding site is occupied with the invariant chain, preventing the binding of other peptides present in the ER. When the MHCII molecule reaches the MIIC, cathepsins (such as cathepsin S) cleave the invariant chain leaving a smaller peptide known as Class II-associated invariant chain (CLIP) bound. DM first removes CLIP from the binding site of newly synthesized MHCII, enabling the binding of antigenic peptides to MHCII. Studies have shown that DM also accelerates the exchange of peptides and generates stable pMHCII, which presented to CD4+ T cells (Ferrante et al., 2008).

The mechanism behind immunodominance and epitope selection via MHCII is not clearly understood. Past studies have suggested that many factors such as length of peptides, cathepsin sensitivity, available T cell repertoire, pMHCII affinity to T cell receptor, and DM susceptibility can influence epitope selection or immunodominance (Kim & Sadegh-Nasseri, 2015). Altogether these factors make epitope selection a complex process.

Currently, two models can explain how immunodominant epitopes might be selected. The "epitope accessibility" model assumes that immunodominant epitopes are easily accessible to MHCII binding groove and cathepsins, which process these epitopes (Dai, Steede, & Landry, 2001). This model is supported by the evidence that many known immunodominant epitopes are located in the region of the protein included or adjacent to the more solvent-exposed C – or N – termini or flexible portion of the protein (Dai et al., 2001; Guillet, Lai, Briner, Smith, & Gefter, 1986; Nepom et al., 2001; Thomas, Hsieh, Schauster, Mudd, & Wilner, 1980). The "kinetic stability" model assumes that immunodominant epitopes form highly stable complexes with MHCII (Lazarski et al., 2005). An important factor that can help explain stable peptide binding to MHCII is the peptide dissociation in the presence of DM. Using a broad set of peptides derived from the vaccinia virus, Yin et al. investigated factors like MHCII binding affinity (IC_{50}), intrinsic half-life, and DM-mediated half-life to understand the factors contributing to peptide immunogenicity. The authors found that DM mediated half-life was an independent factor that identified MHCII epitopes (Yin, Calvo-Calle, Dominguez-Amorocho, & Stern, 2012). Based on these models, we can say that not one factor alone is responsible for selection of immunodominant epitope, and multiple factors drive the selection of immunodominant epitopes. Immunodominant epitope discovery remains the focus of many efforts to understand the immune response and for vaccine developments against various pathogens. One such effort is dedicated

to the construction of prediction algorithms. The advantage of in-silico predictions is that they can generate large binding datasets in a shorter time frame than the typical in-vitro assays, and they require fewer resources. Many algorithms such as NetMHCII 2.2, NetMHCIIpan 3.1, SMM-align, NN-align, and SYFPEITHI are currently available on-line (Karosiene et al., 2013; Nielsen & Lund, 2009; Nielsen, Lundegaard, & Lund, 2007). The Immune Epitope Database and Analysis Resource (IEDB) and SYFPEITHI are two major binding database sources for MHCI and MHCII. As of August 2012, the SYFPEITHI database contains 7000 peptide sequences that bind to MHCI and MHCII while the updated IEDB site shows that it contains 44,541 MHCII binding affinity data covering 26 allelic variants (IEDB and SYFPEITHI reference). Although predictive algorithms are used for both MHCI and MHCII, predictions for the former are relatively superior to the latter (Bisset & Fierz, 1993; Jensen et al., 2018). The reason for the difference in accuracy between MHCI and MHCII prediction models is not fully understood. For MHCII, large binding datasets are available for well-studied alleles, like DRB1*0101, 0301, 0701, 1501, and a limited number of DP and DQ alleles, but they are scarce for more rare alleles, making it difficult to provide reliable predictions. A major difference between MHCI and MHCII lies in their binding grooves. While MHCI has a closed binding groove allowing for the binding of only 9-10-amino acid long peptides, MHCII has an open binding groove, which enables binding of peptides of various length (between 8-9 and 25 residues or more), longer peptides can slide across the open-ended binding site and potentially interact with the MHCII in different "registers" with each allele or across alleles. This phenomenon adds a layer of complication to the prediction for class II (Nielsen, Lund, Buus, & Lundegaard, 2010). MHCII peptide interaction is comprised of three main interactions: hydrophobic interaction between peptide and deep pockets at either end of the groove, H-bond network across the length of the groove which

involves binding to the peptide backbone, and t ionic interactions due to charged side chains of shallower pockets in the center of the groove (Nelson & Fremont, 1999). The ionic interaction may vary between different alleles. The peptide side-chain residues labeled as P1 through P9 define the so-called "binding core". Of these 9 positions, P1, P4, P6, and P9 correspond to major pockets in the MHCII where peptide side chains are encapsulated (Sato et al., 2000). Many algorithms focus on these 4-5 anchoring residues for their predictions (McSparron, Blythe, Zygouri, Doytchinova, & Flower, 2003; Sathiamurthy et al., 2003). Studies have shown that the MHCII can bind to peptides ranging in length from 9-25 amino-acid, and such longer peptides is also presented to CD4+ T cells (Chicz et al., 1992; Srinivasan, Domanico, Kaumaya, & Pierce, 1993). These longer peptides have different binding cores which make it difficult to predict the anchor residues which will interact with the binding site and generate a better CD4+ T cell response, thus making MHCII epitope prediction difficult.

Studies have shown that algorithms indeed predict epitopes, but they also have a high degree of false positives (FP) and false negatives (FN) (Chaves, Lee, Nayak, Richards, & Sant, 2012), which is one of the reasons for the limited accuracy. The high degree of false positives in algorithms suggested that maybe there are epitopes that can bind to MHCII well but lack the features required for activation of CD4+ T cells. Consequently, one could speculate that improved predictions would depend less on either change made in training method or increased data sets and more on including those factors and parameters involved in the presentation of antigen but currently disregarded. Apart from binding to MHCII molecule, another factor which plays a pivotal role in immunodominant selection is the cleavage pattern of the protein and stability of pMHCII complexes in the presence of DM, but the algorithms that are currently available do not include DM-associated binding affinity and kinetics data in their training datasets.

As mentioned above, studies have shown that DM-mediated dissociation is one of the determinants which separate epitopes from non-epitopes. Investigating the IC_{50}s has led to the finding that IC_{50} correlation with CD4+ T cell response is not as strong as DM mediated dissociation (Yin et al., 2012). This finding may not be 100% accurate as the study did not measure the IC_{50}s in the presence of DM.

Here we used a set of peptides with known CD4+ T cell responses to study the DM activity and its relationship with the binding affinity of the peptides. To understand how DM affects the binding affinity of different epitopes, we measured the binding affinity of peptides derived from H1N1 hemagglutinin protein in the absence and presence of DM. For these peptides, a CD4+ T cell response had been already assessed by ELISPOT assay (Richards et al., 2007). We observed that in the presence of DM the binding affinity of epitopes with weaker CD4+ T cell response always decreases, whereas the binding affinity of epitopes with stronger CD4+ T cell response could either increase or decrease. We also evaluated the prediction efficiency of four known algorithms (NetMHCII, NetMHCIIpan, SMM-align, and NN-align) in comparison to DM-associated binding affinity. A comparison between measured and predicted binding affinity showed that the inclusion of DM leads to a better prediction as compared to the one from algorithms. Based on these observations we propose that DM-associated IC_{50} is a better predictor of immunodominant epitopes, and inclusion of peptide binding affinity data measured in the presence of DM would lead to more accurate algorithms.

Materials and methods

Expression and purification of recombinant soluble protein DR1 and DM: Recombinant soluble empty (peptide free) HLA-DR1 was produced in stably transfected Drosophila S2 cell line. Supernatant from the culture collected after inducing with copper sulfate. DR1 was purified from the supernatant by using L243-immunoaffinity chromatography as previously published (Stern et al., 1994). FLAG-tagged HLA-DM was stably transfected in Drosophila S2 cell lines and was purified by using anti-FLAG crosslinked M2 beads as described (Richards et al., 2007).

Peptide synthesis: Seventy-four 18-mer peptides offset by 11 amino acids spanning the entire length of H1N1 influenza virus hemagglutinin protein were synthesized, for which a DR1-restricted CD4+ T cell response in a mouse model had been already measured (Richards et al., 2007). These peptides were classified into four categories based upon CD4+ T cell response normalized to the response against peptide HA435-452 taken as a 100% reference: peptides with less than 5% response considered negative, those with 6 to 24% response considered weak and those with 25-60% response was considered subdominant. Any peptide showing at least 61% of the response was considered immunodominant. Five additional 15-mer peptides consistently reported within the top ten hits across algorithms were selected and synthesized to assess the impact of length on the binding. Peptide IDs and corresponding sequences are listed in Table I. Biotinylated-$HA_{306-318}$ (GPKYVKQNTLKLAT) from influenza A virus H3 subtype was used as a benchmark peptide in competition binding assays. Five 15-mer peptides representing shorter variations of the 18-mer peptides HA204-221 and HA435-452 were also synthesized. Peptides were synthesized by Anaspec Inc and ABI Scientific using Fmoc chemistry and fully automated multiple peptide synthesizer. Peptides had greater than 90% purity and verified by reverse-phase

HPLC and mass spectrometry. Peptides were dissolved in 10-100% DMSO based on their solubility and stored at -20°C.

Competition Binding Assay: Inhibition constant (IC_{50}) for each peptide was measured as described (Ferrante & Gorski, 2012). Briefly, 40 nM DR1 was incubated with 40 nM biotinylated HA (bioHA) peptide in citrate-phosphate buffer pH 5.4 (0.1% BSA, 0.01% Tween 20, 0.1 mM iodoacetamide, 5 mM EDTA, 0.02% NaN_3) in the presence of varying amounts of inhibitor peptides (2.5 mM to pM) for 3 days at 37°C. The incubation time ensured that over >65% of the DR1 protein participates in the peptide-binding reaction to reach equilibrium. To measure the IC_{50} values in the presence of DM, 240 nM DM added along with DR1 during reaction setup. The bound biotinylated peptide was detected using a solid-phase immunoassay, and Eu^{2+} labeled streptavidin. Plates were read using a Wallac VICTOR counter (PerkinElmer Wallac). Data was fit to logistic equation $y= a/ [1+ (x/x_0) b]$ by using Systat SigmaPlot. IC_{50} values obtained from the curve fit of the binding data. Each experiment was performed in quadruplicate and three individual experiments performed for each peptide. The mean of the three experiments was plotted to calculate the IC_{50} of each peptide. IC_{50} values can be used to calculate the equilibrium dissociation constant (K_d). K_d values for each peptide can be calculated by using Cheng-Prusoff equation $K_d = (IC_{50})/(1+[bioHA]/K_d\, bioHA)$ (Yung-ChiCheng, 1973). As K_d usually used as the measure of the binding affinity, here we have used IC_{50} as a proxy for binding affinity as IEDB uses IC_{50} values as a measure for epitope binding to MHCII.

IC_{50} predictions using algorithms: Four algorithms: NetMHCII 2.3, NetMHCIIpan 3.2, NN-align, and SMM-align (Karosiene et al., 2013; Nielsen & Lund, 2009; Nielsen et al., 2007) were used for predicting IC_{50} of the peptides under scrutiny. NetMHCII 2.3 and NetMHCIIpan 3.2 were the most updated version as of August 2018. The entire sequence of the HA protein from the

human influenza A/New Caledonia/20/99 (H1N1; Uniprot ID: QWG600) was inserted into the website to predict the IC_{50} of each peptide. NetMHCII 2.3 and NetMHCIIpan 3.2 predicted the IC_{50} for 18-mer peptides, and SMM-align and NN-align predicted IC_{50} for 15-mer peptides. SMM-align and NN-align predictions were made on 01/04/2018 using the IEDB analysis resource SMM-align and NN-align tool (Nielsen & Lund, 2009; Nielsen et al., 2007).

Unpaired two-tailed t-test and one-way Analysis of Variance (ANOVA): Unpaired two-tailed t-test and one-way ANOVA was performed by using Graph prism5. These tests were performed to determine if the IC_{50}s measured by competition binding assay without and with DM are significantly different between the four categories of peptides distinguished by Richards et al. 2007. We employed unpaired two-tailed t-test to compared immunodominant peptide with other categories of the peptides (subdominant, weak, and negative) while one-way ANOVA was used to compared immunodominant, subdominant, weak, and negative peptides.

Comparative analysis of predicting IC_{50} and measured IC_{50} by using receiver operating characteristic (ROC) curves and area under the curve (AUC) scores: Each binding prediction method compared with the epitopes identified by ELISPOT assay. Peptides with greater than 5% proliferation of CD4+ T cell response as compared to the HA435-452 (the reference peptide) were considered as an epitope for ELISPOT analysis. Prediction accuracy was measured as described (Chaves et al., 2012). False-positive (FP), false negative (FN), true positive (TP), and true negative (TN) rates were calculated for each method. For ELISPOT analysis, peptides which were binding and eliciting a T cell response were considered TP, peptides which were binding but not eliciting a T cell response were considered FP, peptides which did not bind nor elicited a T cell response were considered TN and peptides which did not bind but elicited a T cell response were considered as FN. Since algorithms are designed to predict binding, we set IC_{50} threshold to 5000 nM as per

IEDB recommendation, such that any peptide with IC_{50} greater than 5000 nM does not activate T cell.

True positive rate (TPR) = (TP/TP+FN) and False positive rate (FPR) = (FP/FP+ TN) were calculated. ROC curve was generated by plotting FPR against TPR using Graph Prism5. AUC values from the curves were used to calculate the prediction accuracy of each method. The prediction accuracy of algorithms was also calculated for a binding method without and with DM, as described above. To calculate the TP, TN, FP, and FN we differentiate between the epitope and non-epitope. In our binding assay, some peptides did not show binding at all; such peptides considered non-epitope while remaining peptide, which shows binding, were considered epitope.

Results

Binding affinity of weaker epitopes is reduced by the presence of DM more consistently than for dominant epitopes: To understand whether the addition of DM has an impact on the affinity of peptides classified according to their CD4+ T cell response, we measured IC_{50} of a set of peptides. These peptides were derived from hemagglutinin protein of the H1N1 influenza virus for which a DR1-restricted CD4+ T cell immunodominance hierarchy had been already established (Richards et al., 2007). In the original report, transgenic DR1 mice intranasally were infected with A/New Caledonia/20/99 at 2-4 months of age. These mice were sacrificed, and spleen was used as a source of CD4+ T cells for ELISPOT assay. The entire sequence of the hemagglutinin protein was covered by eighty 18-mer peptides overlapped by 11 amino acids, were synthesized, and the peptide-dependent CD4+ T cell response was measured. These peptides were categorized into four categories based on their CD4+ T cell response, as shown in Table 2: Immunodominant (6 peptides), subdominant (6 peptides), weak (29 peptides) and negative (39 peptides). We measured the IC_{50} of all 74 peptides both in the absence and presence of DM using a competition binding assays (CBA) with biotinylated $HA_{306-318}$ as our reference sequence (Table 3). Various concentrations of unlabeled test peptides could compete against biotinylated $HA_{306-318}$ for binding to DR1 in the absence of DM or along with 3-fold excess of DM. DR1 bound to biotinylated $HA_{306-318}$ was measured by using Eu^{+2} labeled streptavidin in a solid-phase assay.

To test whether there is any difference between IC_{50} values measured by our system based on the CD4+ T cell response (as determined by Richards et al. 2007), we plotted $\log IC_{50}$ values of peptides from individual categories (immunodominant, subdominant, weak and negative) and performed one-way ANOVA. We tested these for IC_{50} values in the absence and presence of DM. As shown in Figure 2.1A, there was a significant difference between immunodominant and weak

(p<0.05), immunodominant, and negative (p<0.05) and subdominant and weak (p<0.05). The absence of DM was not revealing a comparable difference between the individual peptide categories (data not shown). This indicates that the inclusion of DM can help distinguish between peptides eliciting different CD4+ T cell responses.

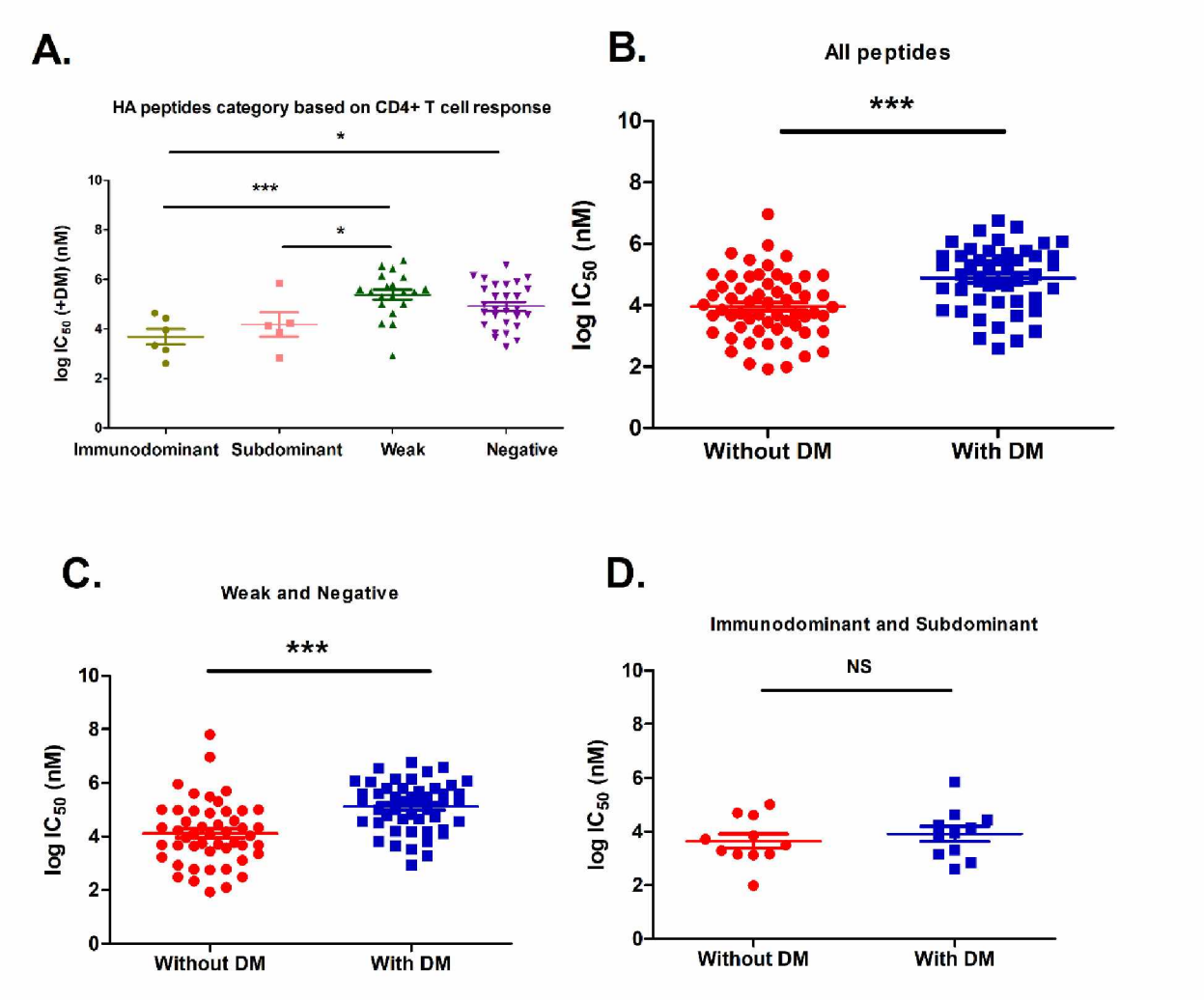


Figure 2.1: Inclusion of DM impact the binding affinity of weaker epitopes: (A) One-way ANOVA test performed for all four categories of peptides (immunodominant, subdominant, weak and negative). Two-tailed unpaired t-test between binding experiments performed in the absence and in the presence of DM was performed for (B) all peptides, (C) weak and negative peptides, (D) immunodominant and subdominant peptides.

To understand how DM presence impacts the IC_{50} of the tested peptides, we ran a set of two-tailed unpaired t-test on IC_{50} values of the peptides measured in the absence and presence of DM. Upon

comparing the log IC_{50} of peptides with and without DM, irrespective of their CD4+ T cell response, we found the IC_{50} values to be significantly different ($P<0.0001$), as shown in Figure 2.1B. The inclusion of DM significantly reduced the IC_{50} values of the peptides. Next, to check if this drop in IC_{50} values is across all the peptides or restricted to only a few peptides, we analyzed the logIC_{50} of peptides based on their CD4+ T cell response categories, without and with DM. We found that weak and negative peptides had an IC_{50} significantly different ($P<0.0001$), as shown in Figure 2.1C, due to the presence of DM while immunodominant and subdominant peptides did not show a change in IC_{50} values, as shown in Figure 2.1D. These indicate that DM-associated binding change directly correlates to CD4+ T cell response.

DM-associated IC_{50} is significantly different for immunodominant peptides when compared to other peptides: To see if the inclusion of DM in measuring peptide binding affinity can help distinguish between immunodominant peptides and other peptides categories we plotted logIC_{50} values of immunodominant peptides against the remaining peptide categories (subdominant, weak and negative). As shown in Figure 2.2A, we found that the IC_{50} values of immunodominant peptides without DM is not significantly different from the remaining peptides ($n = 3$, $p = 0.18$, unpaired t-test). On the other hand, as shown in Figure 2.2B, the presence of DM led to the IC_{50} values of immunodominant peptides to be significantly different ($n = 3$, $p = 0.0028$, unpaired t-test). We also ran the unpaired t-test where we check subdominant, weak, and negative individually against remaining peptides (data not shown) to test if there IC_{50} is different. Only weak peptides IC_{50} values were significantly different from remaining peptides IC_{50} values. The IC_{50} values of weak peptides were different from other categories irrespective of DM presence. Based on this

observation, we conclude that DM-associated IC_{50} is a significant determinant of immunodominant epitopes but not of others.

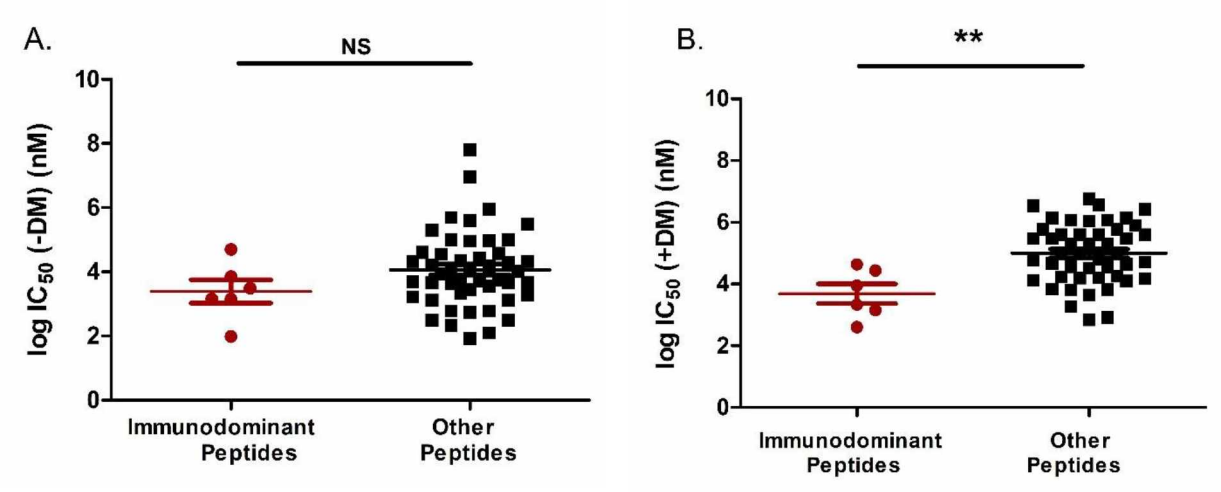


Figure 2.2: DM-associated IC50 is significantly different for immunodominant peptide: Two-tailed unpaired t-test between immunodominant peptides and other peptides (subdominant, weak, and negative) performed for (A) IC50 in the absence of DM, (B) IC50 in the presence of DM.

Predictive IC_{50} values differ from measured IC_{50}: In the last two decades, various online algorithms have been developed to predict the MHC epitopes (Bisset & Fierz, 1993). They predicted MHC-peptide affinity and used them as a proxy for the T-cell epitope. However, various studies have shown that these algorithms' efficiency for MHCII is lower than MHCI due to the high number of false-positive epitopes (Chaves et al., 2012). We assessed the accuracy of IC_{50} values predicted by different algorithms as opposed to experimentally determined IC_{50} values. We used different on-line available algorithms to predict the IC_{50} values of the peptides tested in competition binding assay and only considered IC_{50} values in the absence of DM since all current algorithms trained on uncatalyzed peptide-binding datasets. While NetMHCIIpan 3.2 and NetMHCII 2.3 could predict IC_{50} values for 18 amino acid peptides, SMM-align and NN-align predicted IC_{50} values for 15 amino acid peptides. As shown in Table II, predicted IC_{50} values were within a range of 6 nM to 24000 nM, while measured IC_{50} values were above 40 nM. Upon plotting

a heat map to display the difference between the measured and predicted IC_{50} of each peptide, as shown in Figure 2.3, we concluded that measured IC_{50} values were mostly in the μM range, while predicted IC_{50} values were mostly in the nM range. Comparison of the IC_{50} generated in vitro and in silico using a one-way ANOVA revealed a significant difference between measured and predicted IC_{50} (data not shown, $p<0.05$).

On ranking these peptides based on their binding affinity (data not shown), NN- and SMM- align showed the weak and negative peptides HA526-543 and HA533-550, respectively as rank 1, and both NetMHCII algorithms showed the negative peptide HA267-284 as rank 1. The immunodominant peptide HA440-455 always ranked in the top ten as per algorithm predictions. Our binding assay identified the negative peptide HA267-284 to be ranked one, followed by immunodominant peptide HA162-179 in the absence of DM, while immunodominant peptide HA204-221 was ranked 1 in the presence of DM. These findings indicate that the competition binding assay with DM was more accurate in identifying the immunodominant epitope.

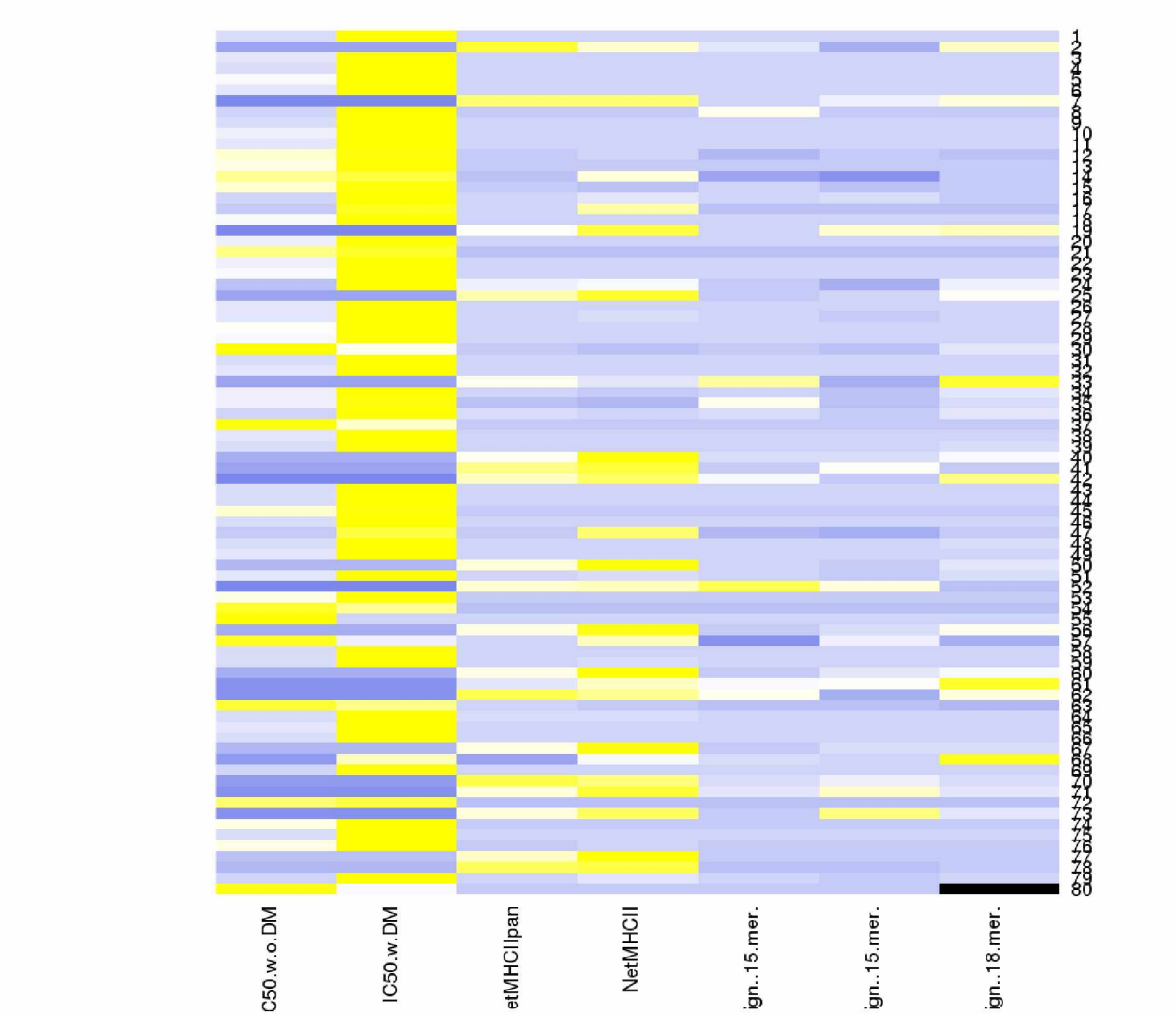


Figure 2.3: Heatmap of binding affinity for peptides derived from hemagglutinin protein for which CD4+ T cell restricted immunodominance hierarchy is already established (Richards et al. 2007). IC_{50} for each peptide is measured by competition binding assay in the presence and absence of DM. IC_{50} for each peptide also predicted by algorithms NetMHCIIpan 3.2, NetMHCII2.3, NN-align, and SMM-align. A comparison between predicted and measured IC_{50} showed that predicted and measured IC_{50} has a significant difference in their values. For algorithms mostly, a peptide with weaker immune response showed better binding affinity while in measured IC_{50} peptide with stronger immune response as well as weaker response showed comparable binding affinity. For with DM, the top peptide was the one with a stronger immune response, and in the presence of DM affinity for a weaker response, peptide always decreased.

Accounting for peptide length did not improve the predictions: SMM-align and NN-align do not provide the option for selecting the length of the peptide, so we used the result of the 15-mer prediction for the comparison with the experimental IC_{50} of the 18-mer peptide. To see if the length

of the peptide is a factor for the difference between predicted and measured IC_{50} values we measure the binding affinity of selected 15-mer peptides derived from immunodominant peptide HA204-221 and HA435-452. To obtain the 15-mer peptides we fed the sequence of the peptide in the IEDB database and selected the top 10 percentile peptides for our binding assay. As shown in Table II, there is a significant difference between measured and predicted IC_{50} values. These results point to the likelihood that peptide length is not a factor responsible for the difference between experimental and predicted binding affinity values as we noted for SMM and NN-align.

DM-mediated binding affinity improves the epitope prediction: Next, we wanted to test which method is better in predicting immunogenicity. We tested the efficiency of CBA without DM, CBA with DM, and web-based available algorithms. IEDB recommendations were used to classify the peptides as binder v/s non-binders. IEDB proposes the following recommendations for classifying affinity based on the IC_{50} values: high (IC_{50} values <50 nM), intermediate (IC_{50} values >50 and <500 nM), and low (IC_{50} values >500 and <5000 nM). Peptides with IC_{50} values >5000 nM are not known to be viable T cell epitopes, so we classified such peptides as non-binders (These peptides are different from the peptides for which we have no binding in our binding assay). Based on this classification, 25 out of the 74 peptides had an IC_{50} <5000 nM. We found that 35 peptides had an IC_{50} >5000 nM, and 14 peptides were not competitive at all in our binding assay for DR1 in the absence of DM. On the other hand, eight peptides showed IC_{50} <5000 nM, 51 peptides showed IC_{50} >5000 nM, and 15 peptides showed no binding at all for DR1 in the presence of DM. Out of the six immunodominant peptides, 4 featured an IC_{50} <5000 nM and 2 featured IC_{50} >5000 nM for DR1 in the absence of DM. When DM added in the binding assay, 50% of the immunodominant peptides showed an IC_{50} value greater than 5000 nM.

To check which method is best for prediction, we plotted the ROC curve for each in vitro and in silico approach, using the published ELISPOT assay data to define positive and negative epitopes. TP, FP, TN, and FN values were calculated for each threshold along with TPR and FPR, as described in methods. AUC values of the ROC curves calculated where 1 represents a perfect predictor, and 0.5 represents a random algorithm. We plotted three ROC curves for the overall comparison between different binding methods. First, we compare all the binding methods against the ELISPOT assay data. To differentiate between epitope and non-epitope we set the cut-off at 5%. Any peptide with CD4+ T cell response >5% was considered an epitope. Upon plotting the TPR and FPR values for peptides which have IC_{50} <5000 nM (IEDB recommendations that no known T cell epitope has IC_{50} >5000 nM) CBA with DM had the highest AUC value of 0.84 (Table 1) as compared to other methods which were very close to 0.5 values (Figure 2.4A). We hypothesize that DM is better for predictions when epitopes IC_{50} value cutoff set to <5000 nM. This is because we cover fewer FP epitopes. This becomes clear when we plot a ROC curve for CBA with and without DM (data not shown) up to 50000 nM which were the maximum IC_{50} value we calculated by using our binding assay. We observed that CBA with DM has its AUC value decreased to 0.68, and for CBA without DM has its AUC values decreased from 0.68 to 0.63.

NetMHCII provided the most accurate predictions compared to other algorithms: Next, we wanted to analyze which algorithm is better in predicting binding affinity upon comparing CBA as a method to select epitope. TP, FP, TN, and FN were measured as follows: peptides which predict binding, and for which we also observed binding in our method, were considered TP, peptides which predicted binding but did not show binding in our method considered as FP, peptides which predicted no binding (IC_{50} >5000 nM) and did not show binding in our method were considered as TN, and finally the peptides which were predicted no binding but shows

binding in our method were considered FN. In the past, it had been demonstrated that NetMHCII 2.2 performed better compared to NetMHCIIPan 3.1 when a large number of peptides were tested (Chaves et al., 2012). When testing with CBA without DM, we found that NetMHCII 2.3 (updated version) and NN align better in prediction with an AUC value of 0.84, as shown in Figure 2.4B while for CBA with DM, NetMHCII 2.3 was better in prediction with an AUC value of 0.86, as shown in Figure 2.4C. Even for comparison with the ELISPOT assay, the highest AUC value was NetMHCII 2.3 with AUC value of 0.67. Thus, our findings were consistent with previous literature indicating that NetMHCII latest version is better in prediction among the chosen four algorithms.

Table 2.1: AUC value for ROC curves plotted to calculate the prediction efficiency of various methods

Figure 2.4A		
Method	AUC	P-value
CBA w/o DM	0.68	0.34
CBA w DM	0.84	0.07
NetMHCIIpan 3.2	0.56	0.67
NetMHCII 2.3	0.67	0.24
NN-align	0.53	0.83
SMM-align	0.57	0.59
Figure 2.4B		
Method	AUC	P-value
NetMHCIIpan 3.2	0.81	0.011
NetMHCII 2.3	0.84	0.006
NN-align	0.84	0.006
SMM-align	0.73	0.06
Figure 2.4C		
Method	AUC	P-value
NetMHCIIpan 3.2	0.80	0.013
NetMHCII 2.3	0.86	0.003
NN-align	0.85	0.004
SMM-align	0.73	0.06

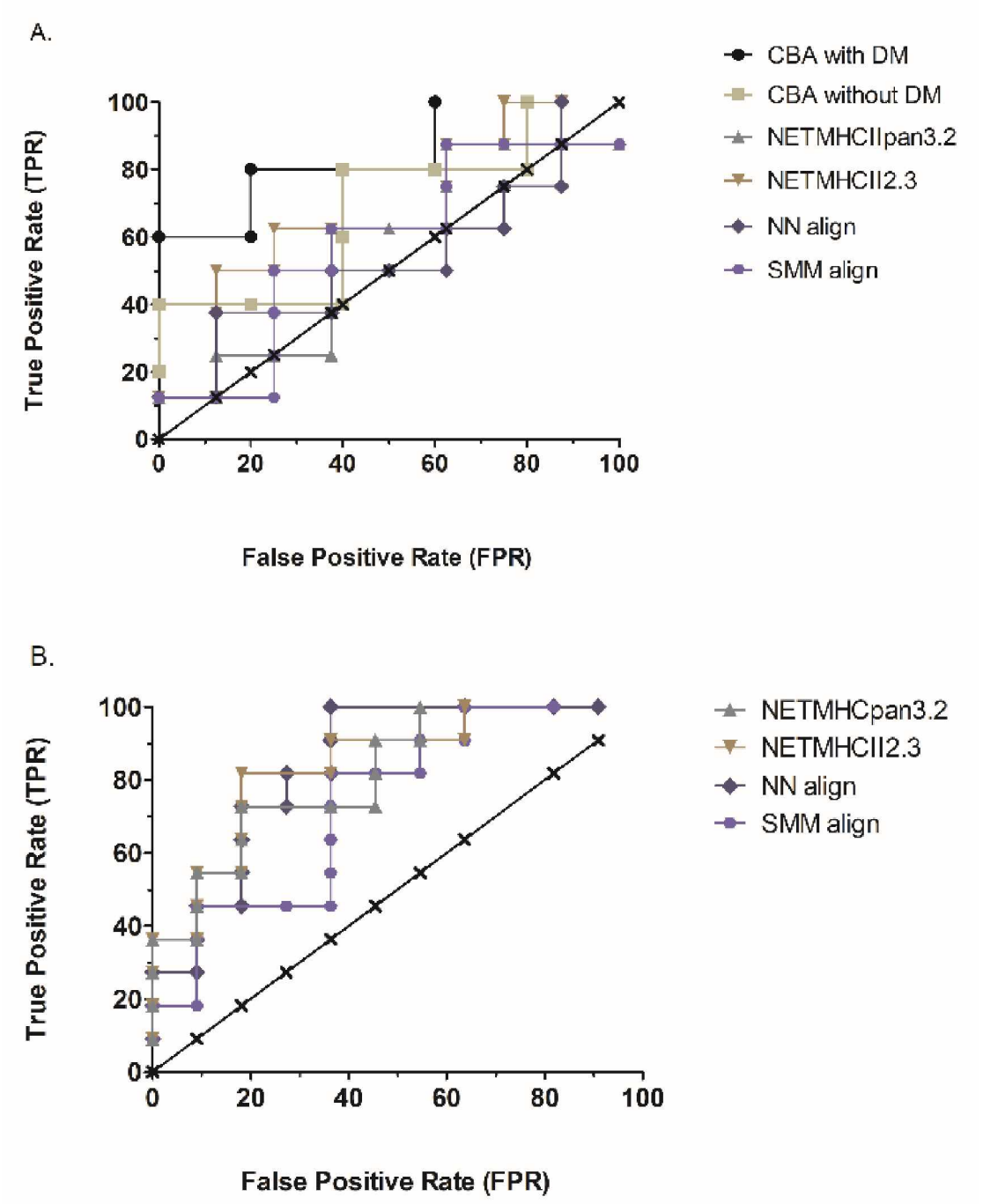
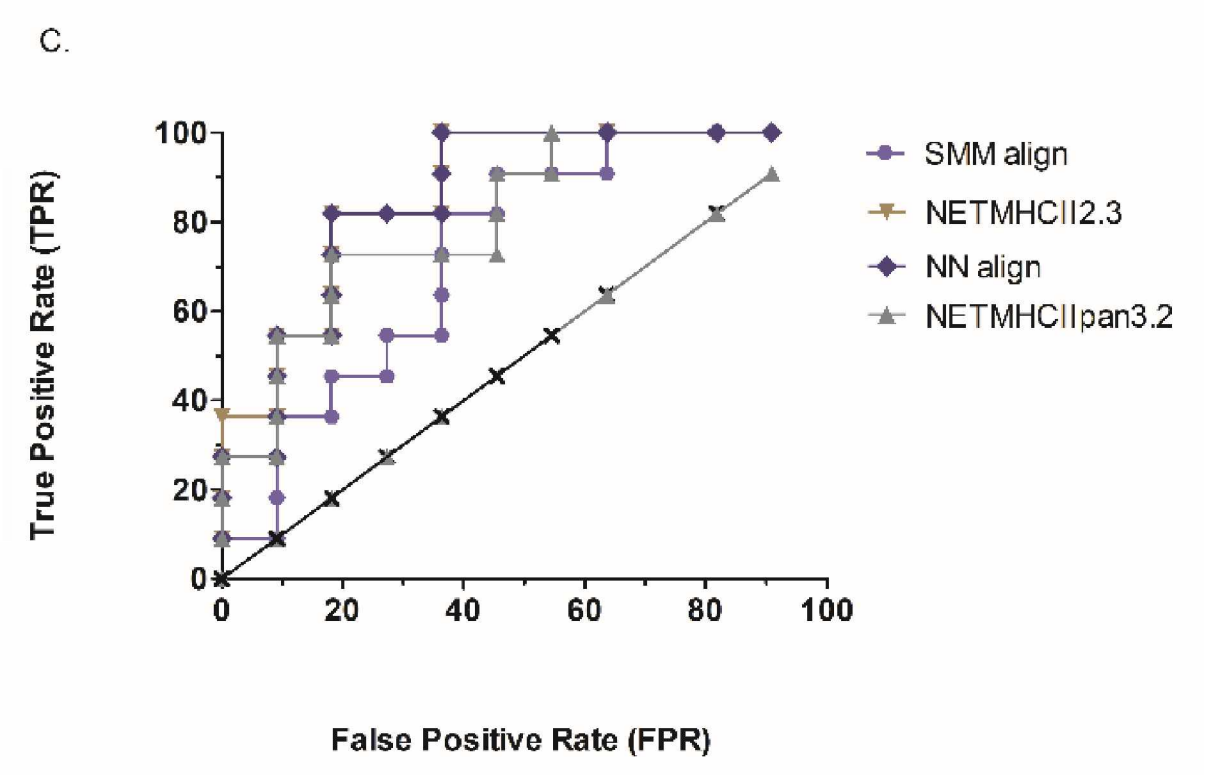


Figure 2.4: ROC curve and area under the curve (AUC) values for different binding methods. CBA with DM is better in prediction. A. ROC curve for comparison between measured IC_{50} and predicted IC_{50} to determine the method best for prediction of the epitope. NetMHCII 2.3 is a better predictor of binding when compared with other algorithms. B. ROC curve for comparison between different algorithms with competition binding assay without DM. C. ROC curve for comparison between different algorithms with competition binding assay with DM.

Figure 2.4: Continued



Discussion

Immunodominant epitope discovery remains the focus of many studies along with online-available prediction tools. The available online algorithms are not as efficient for MHCII epitope prediction as they are for MHCI. The reason for the lower accuracy rate for MHCII is mainly due to a high rate of false-positive epitopes. The high rate of false positives can be partially explained to omitting various factors that are known to be involved in epitope selection. One of the crucial factors which are not considered in the training dataset is DM activity. Previously published studies have investigated the role of DM in shaping the peptide repertoire by using DM-mediated half-life and DM-susceptibility (Yin et al., 2012). It is well established that DM accelerates the exchange of peptides and helps in the selection of kinetically stable MHCII-peptide (Yin, Maben, Becerra, & Stern, 2015). It is not very well studied how binding affinity (IC_{50}) of peptides would change in the presence of DM and if this factor can be used as a correlate of immunodominance.

To understand whether the DM-associated binding affinity can help in identifying immunodominant epitopes and whether DM can improve epitope prediction, we used peptides from hemagglutinin protein of H1N1 influenza virus, for which a DR1-restricted CD4+ T cell immunodominance hierarchy had been already established. The peptides classified into four categories based on their CD4+ T cell response: immunodominant, subdominant, weak, and negative, where immunodominant peptides showed the strongest T cell response and negative peptides showed the weakest T cell response. We measured the IC_{50} values for each of these peptides in the absence and presence of DM. Upon comparing the IC_{50} values of immunodominant peptides against the pool of the other three categories of peptides, we found that IC_{50} measured in the presence of DM was significantly different. We also studied the impact

of DM on the IC_{50} values of the peptides irrespective of their categories (Figure 1A) and found that there was a significant change in the IC_{50} values due to the presence of DM. When the individual category of the peptides was tested using ANOVA, we observed that for weak and negative peptides there was a significant difference between without and with DM IC_{50} values but not so for immunodominant and subdominant peptides. Taken all together these results showed that DM has significantly altered the binding affinity of the peptides for MHCII, and the impact of DM action is more significant on peptides with weaker CD4+ T cell response.

Another critical observation seen in our analysis is that certain peptides, like HA267-284 or HA323-340, which are associated with a weaker CD4+ T cell response, featured a binding affinity greater than peptides eliciting a stronger immune response. These results illustrate that a strong association with MHCII alone cannot determine the strong CD4+ T cell response. The binding affinity for these peptides reduced 22-fold and 6-fold in the presence of DM. These demonstrates that DM might be one of the factors which can explain the rationale behind how peptides with a greater binding affinity generate a weaker immune response.

Currently, available epitope prediction algorithms focus on binding affinity as a major factor because it is an essential step during antigen processing, and binding interaction between MHCII and antigen will also determine its capability to activate CD4+ T cells to generate an immune response. To ascertain the accuracy of such algorithms for the experimental epitopes adopted in our study, we tested the difference between measured and predicted binding affinity. We observed that the gap between these two values is significant. When we assessed the accuracy of the algorithm and competition binding assay, we found that competition binding assay with DM shows the best prediction with an AUC value of 0.84. We also compared different algorithms with each

other and found that NetMHCII's latest version is better in prediction compared to all the tested algorithms. This result is consistent with previously published studies (Chaves et al., 2012). Upon comparing various updated versions of these algorithms, there has been not much difference in terms of predictions. These results are due to not considering all the factors which are involved in MHCII antigen processing. Studies are started to investigate cathepsin cleavage (Schneidman-Duhovny et al., 2018) and role of flanking region in the activation of CD4+ T cells after binding affinity, but DM still neglected.

Current literature has shown that in the case of MHCII, the flanking residues on both N-termini and C-termini of an antigen can play a role in the activation of CD4+ T cells. Length of the peptide is an essential factor that can alter the antigen immune response. More recent studies have begun to focus on this element and started including the length of peptides in MHCII antigen prediction. Recently IEDB has introduced a new binding prediction method, which predicts the cleavage pattern as well as the binding affinity of an antigen; the performance of this new prediction method appeared to be better than the algorithms which only predict binding affinity.

DM is a non-classical molecule that removes CLIP from a newly synthesized MHCII so that the antigen can bind to MHCII. Studies have shown that DM catalyzes the exchange of peptides and selects for kinetically stable peptide-MHCII complex (Ferrante, 2013). Immunodominant peptides generally show DM resistance, while peptides with weaker immune response appear to be DM-sensitive. Upon comparison between DM associated binding affinity with CD4+ T cell response, we observed that the peptide with the weaker immune response always loses binding capability in presence of DM. This can be reconducted to DM ability to exchange peptide where it can accelerate the off rate, thus possibly decreasing the binding affinity of such antigens. For

immunodominant peptide we observed that the binding affinity could either increase or decrease. The possible explanation for immunodominant peptides not showing a one-directional shift in binding affinity could be the length of the peptide, which can affect the DM action on the peptide exchange. The example of such a mechanism is seen in case of few peptides, for instance, HA442-459, which is an 18-mer with decreased affinity in the presence of DM, but HA440-455, which is the same sequence but 16-mer with

better affinity in the presence of DM. Thus, a difference of 2amino acids can drastically change DM action, with the consequence that an antigen cleavage pattern can dictate the outcome of DM action. Indeed, an additional question our group is addressing concerns the possible correlation between CD4+ T cell response and cathepsin resistance and between cathepsin and DM activity. MHCII antigen processing is a complex process, and binding affinity alone cannot predict immunodominance accurately. To generate a better performing algorithm, integration of the various factors affecting antigen selection is needed, starting from DM-associated IC_{50} measurements, with evident impact on our capacity to efficiently and effectively assess immunodominance in a high-throughput, low-cost fashion.

Table 2.2: Peptides derived from the hemagglutinin protein of the H1N1 influenza virus used in this study. Based on the ELISPOT assay result (Richards et al. 2007) peptides were categorized into immunodominant, subdominant, weak and negative. The percentage of CD4+ T cell response measured against response generated by immunodominant peptide HA435-452. Peptides with less than 5% response were considered negative, those with 6 to 24% response considered weak, and those with a 25-60% response were considered subdominant. Any peptide with > 61% of the response considered immunodominant. 15-mer modified peptides were selected from the results of in-silico predictions in that they ranked among the highest ten scoring peptides across algorithms, and they contained binding motifs of experimental immunodominant peptides.

Peptide label	**Category**	**Peptide Sequence**
HA1-18	Negative	MKAKLLVLLCTFTATYAD
HA8-25	Negative	LLCTFTATYADTICIGYH
HA15-32	Weak	TYADTICIGYHANNSTDT
HA22-39	Negative	IGYHANNSTDTVDTVLEK
HA29-46	Weak	STDTVDTVLEKNVTVTHS
HA36-53	Weak	VLEKNVTVTHSVNLLEDS
HA43-60	Weak	VTHSVNLLEDSHNGKLCL
HA50-67	Weak	LEDSHNGKLCLLKGIAPL
HA57-74	Weak	KLCLLKGIAPLQLGNCSV
HA64-81	Weak	IAPLQLGNCSVAGWILGN
HA71-88	Weak	NCSVAGWILGNPECELLI
HA78-95	Negative	ILGNPECELLISKESWSY
HA85-102	Weak	ELLISKESWSYIVETPNP
HA92-109	Subdominant	SWSYIVETPNPENGTCYP
HA99-116	Weak	TPNPENGTCYPGYFADYE
HA106-123	Weak	TCYPGYFADYEELREQLS
HA113-130	Negative	ADYEELREQLSSVSSFER
HA120-137	Weak	EQLSSVSSFERFEIFPKE
HA127-144	Subdominant	SFERFEIFPKESSWPNHT
HA134-151	Weak	FPKESSWPNHTVTGVSAS
HA141-158	Negative	PNHTVTGVSASCSHNGKS
HA148-165	Negative	VSASCSHNGKSSFYRNLL
HA155-172	Immunodominant	NGKSSFYRNLLWLTGKNG
HA162-179	Immunodominant	RNLLWLTGKNGLYPNLSK
HA169-186	Negative	GKNGLYPNLSKSYVNNKE
HA176-193	Negative	NLSKSYVNNKEKEVLVLW
HA183-200	Negative	NNKEKEVLVLWGVHHPPN
HA190-207	Weak	LVLWGVHHPPNIGNQRAL
HA197-214	Subdominant	HPPNIGNQRALYHTENAY
HA204-221	Immunodominant	QRALYHTENAYVSVVSSH
HA211-228	Negative	ENAYVSVVSSHYSRRFTP

Table 2.2: (continued)

HA218-235	Weak	VSSHYSRRFTPEIAKRPK
HA225-242	Weak	RFTPEIAKRPKVRDQEGR
HA232-249	Negative	KRPKVRDQEGRINYYWTL
HA239-256	Negative	QEGRINYYWTLLEPGDTI
HA246-263	Negative	YWTLLEPGDTIIFEANGN
HA253-270	Subdominant	GDTIIFEANGNLIAPWYA
HA260-277	Subdominant	ANGNLIAPWYAFALSRGF
HA267-284	Negative	PWYAFALSRGFGSGIITS
HA274-291	Weak	SRGFGSGIITSNAPMDEC
HA281-298	Negative	IITSNAPMDECDAKCQTP
HA288-305	Weak	MDECDAKCQTPQGAINSS
HA295-312	Weak	CQTPQGAINSSLPFQNVH
HA302-319	Weak	INSSLPFQNVHPVTIGEC
HA309-326	Weak	QNVHPVTIGECPKYVRSA
HA316-333	Negative	IGECPKYVRSAKLRMVTG
HA323-340	Weak	VRSAKLRMVTGLRNIPSI
HA330-347	Negative	MVTGLRNIPSIQSRGLFG
HA337-354	Negative	IPSIQSRGLFGAIAGFIE
HA344-361	Negative	GLFGAIAGFIEGGWTGMV
HA351-368	Negative	GFIEGGWTGMVDGWYGYH
HA358-375	Weak	TGMVDGWYGYHHQNEQGS
HA365-382	Negative	YGYHHQNEQGSGYAADQK
HA372-389	Weak	EQGSGYAADQKSTQNAIN
HA379-396	Weak	ADQKSTQNAINGITNKVN
HA386-403	Negative	NAINGITNKVNSVIEKMN
HA393-410	Subdominant	NKVNSVIEKMNTQFTAVG
HA400-417	Negative	EKMNTQFTAVGKEFNKLE
HA407-424	Negative	TAVGKEFNKLERRMENLN
HA414-431	Negative	NKLERRMENLNKKVDDGF
HA421-438	Negative	ENLNKKVDDGFLDIWTYN
HA428-445	Negative	DDGFLDIWTYNAELLVLL
HA435-452	Immunodominant	WTYNAELLVLLENERTLD
HA442-459	Immunodominant	LVLLENERTLDFHDSNVK
HA449-466	Negative	RTLDFHDSNVKNLYEKVK
HA456-473	Negative	SNVKNLYEKVKSQLKNNA
HA463-480	Negative	EKVKSQLKNNAKEIGNGC
HA470-487	Negative	KNNAKEIGNGCFEFYHKC
HA477-494	Negative	GNGCFEFYHKCNNECMES
HA484-501	Weak	YHKCNNECMESVKNGTYD

Table 2.2: (continued)

HA491-508	Negative	CMESVKNGTYDYPKYSEE
HA498-515	Negative	GTYDYPKYSEESKLNREK
HA505-522	Weak	YSEESKLNREKIDGVKLE
HA512-529	Negative	NREKIDGVKLESMGVYQI
HA519-536	Negative	VKLESMGVYQILAIYSTV
HA526-543	Weak	VYQILAIYSTVASSLVLL
HA533-550	Negative	YSTVASSLVLLVSLGAIS
HA540-557	Weak	LVLLVSLGAISFWMCSNG
HA547-565	Negative	GAISFWMCSNGSLQCRIC
HA440-455	Immunodominant	ELLVLLENERTLDFHD
HA204-218	Modified 15-mer	QRALYHTENAYVSVV
HA205-219	Modified 15-mer	RALYHTENAYVSVVS
HA206-220	Modified 15-mer	ALYHTENAYVSVVSS
HA437-451	Modified 15-mer	YNAELLVLLENERTL
HA438-452	Modified 15-mer	NAELLVLLENERTLD

Table 2.3: IC50 for HA peptides measured by using competition binding assay in the absence and presence of DM and as predicted by NetMHCIIpan 3.2, NetMHCII 2.3, SMM-align (15 aa and 18 aa and NN-align (15 aa). NS: not synthesized due to difficulty in the synthesis. NB: no binding experimentally observed. LB: low binding; the peptide shows minimal competition capability at concentration >mM against the biotinylated benchmark.

Peptide label	IC50 (-DM) (nM)	IC50 (+DM) (nM)	NetMHCIIpan 3.2	NetMHCII 2.3	SMM align (15 aa)	NN align (15 aa)	SMM align (18 aa)
HA1-18	8800	200000	462.09	731.9	37	40.1	92
HA8-25	NS	NS	445.66	219.3	111	19.8	234
HA15-32	20000	300000	1540.01	2189.4	1477	280.4	467
HA22-39	22000	1100000	3993.87	9533.8	164	884.4	443
HA29-46	96000	600000	1565.12	2490.1	444	481.4	1010
HA36-53	89700	1200000	1342.56	6854.4	449	191.1	1050
HA43-60	NB	NB	2890.61	2880.1	825	1230.9	1702
HA50-67	607	15700	318.02	444.3	3888	419.7	266
HA57-74	561	44200	77.97	39.9	21	12.9	56
HA64-81	26700	300000	641.19	283.9	265	105	614
HA71-88	21500	400000	949.07	1312.4	835	2248.3	1149
HA78-95	5415	12700	1439.62	1811.5	662	1179.8	1087
HA85-102	100000	400000	627.21	1219.2	829	425.9	1388
HA92-109	5200	6920	1153.63	3390.8	573	93.3	1496
HA99-116	74000	200000	9230.11	5823.4	16120	4508.8	10684
HA106-123	2800	100000	2803.89	7804.1	2360	4042.8	257
HA113-130	306	6660	618.35	3613.3	49	65.9	141
HA120-137	200000	1400000	1125.28	3162	461	1480.5	1300

Table 2.3: (continued)

Peptide label	IC50 (-DM) (nM)	IC50 (+DM) (nM)	NetMHCIIpan 3.2	NetMHCII 2.3	SMM align (15 aa)	NN align (15 aa)	SMM align (18 aa)
HA127-144	NB	NB	1063.23	2641.1	683	1416.1	1595
HA134-151	36000	300000	4101.07	5262.6	2344	1528.1	1221
HA141-158	400000	600000	1890.97	1760	127	120.3	342
HA148-165	21300	200000	3410.16	3729	3071	2357	4045
HA155-172	7050	43300	245.27	106.3	94	36.2	186
HA162-179	98	1420	307.35	373.8	139	29.6	295
HA169-186	NB	NB	642.02	1122.5	155	186.4	373
HA176-193	4700	60600	1625.37	1231.9	522	1319.2	1250
HA183-200	5000	51600	962.01	3081.9	556	135	1308
HA190-207	86000	500000	1454.13	1420.5	870	302.6	1910
HA197-214	100000	700000	1705.25	2262.6	1177	1592.9	2603
HA204-221	1420	401	61.64	30.7	69	10.1	188
HA211-228	1290	33300	37.14	129.2	44	6.8	129
HA218-235	5660	100000	316.5	154.6	721	563.2	997
HA225-242	NB	NB	805.77	499.1	1395	95.9	2080
HA232-249	12200	65300	4027.15	3650.3	4290	1461.1	9371
HA239-256	841	3410	322	229.6	1143	335.1	688
HA246-263	306	6400	468.86	242.4	503	77.1	627
HA253-270	40000	13300	89.62	43.5	97	39.9	249
HA260-277	1290	17360	161.44	307.2	220	129.6	223

Table 2.3: (continued)

Peptide label	IC50 (-DM) (nM)	IC50 (+DM) (nM)	NetMHCIIpan 3.2	NetMHCII 2.3	SMM align (15 aa)	NN align (15 aa)	SMM align (18 aa)
HA267-284	84	1860	29.89	13.2	43	9.8	111
HA274-291	NB	NB	987.33	2985.5	443	456.8	734
HA281-298	NB	NB	8002.97	10748.3	1585	4012.4	1806
HA288-305	NB	NB	6466.75	9424.7	4379	2563.6	8246
HA295-312	90900	3500000	806.38	1964.2	1370	2217.2	1258
HA302-319	4600	200000	326.41	497	48	33.7	135
HA309-326	900000	2700000	3675.85	5164.1	3376	2754.7	8339
HA316-333	15200	400000	205.67	2706.7	59	84.6	142
HA323-340	124	834	126.32	692.8	56	18.4	117
HA330-347	213	4510	128.87	75.3	67	16.2	137
HA337-354	2210	35500	324.05	748.2	140	25.7	156
HA344-361	NS	NS	559.14	1503.3	133	99.5	268
HA351-368	4760	36700	1682.28	2621.3	1453	1049.5	2687
HA358-375	NB	NB	5201.05	5799.9	8909	5000.4	1866
HA365-382	LB	LB	5169.8	1685.7	498	1282.4	1445
HA372-389	LB	LB	4544.15	8468.3	2384	4087	5257
HA379-396	LB	NB	2001.28	6348.3	2448	3549.8	1442
HA386-403	NS	NS	1208.28	3326.2	373	586.6	1054
HA393-410	1940	690	525.39	1103.6	114	683.3	278
HA400-417	29880	400000	990.99	2774.5	417	1230.4	1138

Table 2.3: (continued)

Peptide label	IC50 (-DM) (nM)	IC50 (+DM) (nM)	NetMHCIIpan 3.2	NetMHCII 2.3	SMM align (15 aa)	NN align (15 aa)	SMM align (18 aa)
HA407-424	10500	200000	1865.85	5187.7	866	2398.1	2219
HA414-431	NB	NB	4479.86	13326.4	1147	2637.6	3337
HA421-438	NB	NB	5564.57	10845.2	7049	7780.4	20369
HA428-445	NS	NS	317.09	252.1	149	28.5	173
HA435-452	3080	2130	333.09	225.1	93	140.9	62
HA442-459	1410	27300	1171.85	1119.3	84	25.7	87
HA449-466	100000	1400000	2044.42	951.8	891	932	2047
HA456-473	38000	800000	714.04	3136.6	36	66.1	99
HA463-480	NB	NB	3244.52	10681.5	421	1375.1	1239
HA470-487	4270	15080	5160.02	10426.2	8363	8154.4	24048
HA477-494	1670	3700000	2189.42	1488.8	4764	4998.8	1711
HA484-501	NB	NB	4932.28	3919.1	1302	1598.4	1313
HA491-508	NB	NB	6879.47	14497.9	4358	8144.7	4325
HA498-515	500000	60000	5118.15	10735.6	3716	4044.3	8834
HA505-522	NB	NB	2677.36	4668.2	1165	4129.8	1551
HA512-529	13300	46500	328.36	567.2	158	402.9	238
HA519-536	3660	96300	274.85	1231.7	148	65.4	47
HA526-543	4580	16100	72.87	322.6	19	6.8	10
HA533-550	NS	NS	253.24	621.1	11	9.1	10
HA540-557	NS	NS	308.88	338.1	15	7.2	25

Table 2.3: (continued)

Peptide label	**IC50 (-DM) (nM)**	**IC50 (+DM) (nM)**	**NetMHCIIpan 3.2**	**NetMHCII 2.3**	**SMM align (15 aa)**	**NN align (15 aa)**	**SMM align (18 aa)**
HA547-565	603	17900	515.88	1624.4	131	23	360
HA440-455	49800	9020	131.23	50.1	27	8.7	80
HA204-218	701	441	24.4	8.4	69	10.10	69
HA205-219	334	433	26.4	8.8	67	8.8	67
HA206-220	1187	1236	43.3	15.2	73	12.9	73
HA437-451	51658	12250	67.7	18.9	24	9	24
HA438-452	5173	714	73.3	15.7	27	9.10	27

References

Bisset, L. R., & Fierz, W. (1993). Using a neural network to identify potential HLA-DR1 binding sites within proteins. *J Mol Recognit, 6*(1), 41-48. doi:10.1002/jmr.300060105

Chaves, F. A., Lee, A. H., Nayak, J. L., Richards, K. A., & Sant, A. J. (2012). The utility and limitations of current Web-available algorithms to predict peptides recognized by CD4 T cells in response to pathogen infection. *J Immunol, 188*(9), 4235-4248. doi:10.4049/jimmunol.1103640

Chicz, R. M., Urban, R. G., Lane, W. S., Gorga, J. C., Stern, L. J., Vignali, D. A., & 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*(6389), 764-768. doi:10.1038/358764a0

Dai, G., Steede, N. K., & Landry, S. J. (2001). Allocation of helper T-cell epitope immunodominance according to three-dimensional structure in the human immunodeficiency virus type I envelope glycoprotein gp120. *J Biol Chem, 276*(45), 41913-41920. doi:10.1074/jbc.M106018200

Ferrante, A. (2013). HLA-DM: arbiter conformationis. *Immunology, 138*(2), 85-92. doi:10.1111/imm.12030

Ferrante, A., Anderson, M. W., Klug, C. S., & Gorski, J. (2008). HLA-DM mediates epitope selection by a "compare-exchange" mechanism when a potential peptide pool is available. *PLoS One, 3*(11), e3722. doi:10.1371/journal.pone.0003722

Ferrante, A., & Gorski, J. (2012). A Peptide/MHCII conformer generated in the presence of exchange peptide is substrate for HLA-DM editing. *Sci Rep, 2*, 386. doi:10.1038/srep00386

Guillet, J. G., Lai, M. Z., Briner, T. J., Smith, J. A., & Gefter, M. L. (1986). Interaction of peptide antigens and class II major histocompatibility complex antigens. *Nature, 324*(6094), 260-262. doi:10.1038/324260a0

Jensen, K. K., Andreatta, M., Marcatili, P., Buus, S., Greenbaum, J. A., Yan, Z., . . . Nielsen, M. (2018). Improved methods for predicting peptide binding affinity to MHC class II molecules. *Immunology, 154*(3), 394-406. doi:10.1111/imm.12889

Karosiene, E., Rasmussen, M., Blicher, T., Lund, O., Buus, S., & Nielsen, M. (2013). NetMHCIIpan-3.0, a common pan-specific MHC class II prediction method including all three human MHC class II isotypes, HLA-DR, HLA-DP and HLA-DQ. *Immunogenetics, 65*(10), 711-724. doi:10.1007/s00251-013-0720-y

Kim, A., & Sadegh-Nasseri, S. (2015). Determinants of immunodominance for CD4 T cells. *Curr Opin Immunol, 34*, 9-15. doi:10.1016/j.coi.2014.12.005

Lazarski, C. A., Chaves, F. A., Jenks, S. A., Wu, S., Richards, K. A., Weaver, J. M., & Sant, A. J. (2005). The kinetic stability of MHC class II:peptide complexes is a key parameter that dictates immunodominance. *Immunity, 23*(1), 29-40. doi:10.1016/j.immuni.2005.05.009

McSparron, H., Blythe, M. J., Zygouri, C., Doytchinova, I. A., & Flower, D. R. (2003). JenPep: a novel computational information resource for immunobiology and vaccinology. *J Chem Inf Comput Sci, 43*(4), 1276-1287. doi:10.1021/ci030461e

Neefjes, J., Jongsma, M. L., Paul, P., & Bakke, O. (2011). Towards a systems understanding of MHC class I and MHC class II antigen presentation. *Nat Rev Immunol, 11*(12), 823-836. doi:10.1038/nri3084

Nelson, C. A., & Fremont, D. H. (1999). Structural principles of MHC class II antigen presentation. *Rev Immunogenet, 1*(1), 47-59.

Nepom, G. T., Lippolis, J. D., White, F. M., Masewicz, S., Marto, J. A., Herman, A., . . . Nepom, B. S. (2001). Identification and modulation of a naturally processed T cell epitope from the diabetes-associated autoantigen human glutamic acid decarboxylase 65 (hGAD65). *Proc Natl Acad Sci U S A, 98*(4), 1763-1768. doi:10.1073/pnas.98.4.1763

Nielsen, M., & Lund, O. (2009). NN-align. An artificial neural network-based alignment algorithm for MHC class II peptide binding prediction. *BMC Bioinformatics, 10*, 296. doi:10.1186/1471-2105-10-296

Nielsen, M., Lund, O., Buus, S., & Lundegaard, C. (2010). MHC class II epitope predictive algorithms. *Immunology, 130*(3), 319-328. doi:10.1111/j.1365-2567.2010.03268.x

Nielsen, M., Lundegaard, C., & Lund, O. (2007). Prediction of MHC class II binding affinity using SMM-align, a novel stabilization matrix alignment method. *BMC Bioinformatics, 8*, 238. doi:10.1186/1471-2105-8-238

Richards, K. A., Chaves, F. A., Krafcik, F. R., Topham, D. J., Lazarski, C. A., & Sant, A. J. (2007). Direct ex vivo analyses of HLA-DR1 transgenic mice reveal an exceptionally broad pattern of immunodominance in the primary HLA-DR1-restricted CD4 T-cell response to influenza virus hemagglutinin. *J Virol, 81*(14), 7608-7619. doi:10.1128/JVI.02834-06

Sathiamurthy, M., Hickman, H. D., Cavett, J. W., Zahoor, A., Prilliman, K., Metcalf, S., . . . Hildebrand, W. H. (2003). Population of the HLA ligand database. *Tissue Antigens, 61*(1), 12-19.

Sato, A. K., Zarutskie, J. A., Rushe, M. M., Lomakin, A., Natarajan, S. K., Sadegh-Nasseri, S., . . . Stern, L. J. (2000). Determinants of the peptide-induced conformational change in the human class II major histocompatibility complex protein HLA-DR1. *J Biol Chem, 275*(3), 2165-2173. doi:10.1074/jbc.275.3.2165

Schneidman-Duhovny, D., Khuri, N., Dong, G. Q., Winter, M. B., Shifrut, E., Friedman, N., . . . Sali, A. (2018). Predicting CD4 T-cell epitopes based on antigen cleavage, MHCII presentation, and TCR recognition. *PLoS One, 13*(11), e0206654. doi:10.1371/journal.pone.0206654

Srinivasan, M., Domanico, S. Z., Kaumaya, P. T., & Pierce, S. K. (1993). Peptides of 23 residues or greater are required to stimulate a high affinity class II-restricted T cell response. *Eur J Immunol, 23*(5), 1011-1016. doi:10.1002/eji.1830230504

Stern, L. J., Brown, J. H., Jardetzky, T. S., Gorga, J. C., Urban, R. G., Strominger, J. L., & Wiley, D. C. (1994). Crystal structure of the human class II MHC protein HLA-DR1 complexed with an influenza virus peptide. *Nature, 368*(6468), 215-221. doi:10.1038/368215a0

Thomas, D. W., Hsieh, K. H., Schauster, J. L., Mudd, M. S., & Wilner, G. D. (1980). Nature of T lymphocyte recognition of macrophage-associated antigens. V. Contribution of individual peptide residues of human fibrinopeptide B to T lymphocyte responses. *J Exp Med, 152*(3), 620-632. doi:10.1084/jem.152.3.620

Yin, L., Calvo-Calle, J. M., Dominguez-Amorocho, O., & Stern, L. J. (2012). HLA-DM constrains epitope selection in the human CD4 T cell response to vaccinia virus by favoring the presentation of peptides with longer HLA-DM-mediated half-lives. *J Immunol, 189*(8), 3983-3994. doi:10.4049/jimmunol.1200626

Yin, L., Maben, Z. J., Becerra, A., & Stern, L. J. (2015). Evaluating the Role of HLA-DM in MHC Class II-Peptide Association Reactions. *J Immunol, 195*(2), 706-716. doi:10.4049/jimmunol.1403190

Yung-ChiCheng, W. H. P. (1973). Relationship between the inhibition constant (KI) and the concentration of inhibitor which causes 50 per cent inhibition (I50) of an enzymatic reaction. *Biochemical Pharmacology, Volume 22, , 1 December 1973,* (Issue 23), Pages 3099-3108. doi:10.1016/0006-2952(73)90196-2

Chapter 3: Alternate antigen presentation pathway for peptides generated from Hemagglutinin protein[2]

Abstract

Classical antigen presentation pathway starts with antigen uptake at the surface of antigen-presenting cells (APCs), which undergoes enzymatic cleavage through various endosomal compartments and binds to Class II major histocompatibility complex (MHCII) in MHCII compartments (MIIC). In the MIIC, antigenic peptides bind to MHCII in the presence of the non-classical MHCII molecule DM. Studies have shown the presentation of proteins through routes different from the classical pathway of antigen presentation and presented in the absence of the peptide editing factor HLA-DM (DM). In this study, we tested DM-susceptible immunodominant peptides under the hypothesis that they were selected through an alternate antigen presentation route by assessing their stability at acidic and neutral pH. Our results show that these immunodominant peptides showed greater stability at neutral pH as compared to acidic pH. We also tested length of the peptides based on the assumption that cleavage patterns can influence the presentation of the antigen. We showed that 15-mer peptides have greater stability at acidic pH as compared to 18-mer peptides. The 15-mer peptides also showed better stability with DM. Taken together these results indicated that DM-sensitive peptides of a certain length could nevertheless be presented and elicit a significant T cell response through favorable alternate pathways characterized by a less acidic environment.

[2] Osan, J.K, Ferrante, A. Kuhn T. B (In Prep) Alternate antigen presentation pathway for peptides generated from Hemagglutinin protein.

Introduction

Class II major histocompatibility complex (MHCII) are glycoproteins that present antigenic peptides at the surface of antigen-presenting cells (APCs) to CD4+ T cells for generating an immune response (Rothbard & Gefter, 1991). A pool of antigenic peptides is generated from enzymatic cleavage, but only a few of the peptides bind to MHCII and present to CD4+ T cells, and out of those presented peptides only a few of them generate efficient CD4+ T cell responses. These peptides which generate the main CD4+ T cell response are known as immunodominant peptides (Kim et al., 2014). A newly synthesized MHCII molecule consisting of α and β chain is assembled in the endoplasmic reticulum (ER) and transported through Golgi most often in the form of a nonameric complex to MHC class II compartment (MIIC). This nonameric complex comprises of three (αβ) dimers along with three invariant chains (Ii), which shield each peptide binding site and act as a chaperone for efficient transport of the MHCII multimer. In the MIIC, the invariant chain is cleaved into a shorter peptide named Class II-associated Ii peptide (CLIP). CLIP was removed from the binding site by HLA-DM (DM), a non-classical MHCII molecule, which enables the binding of antigenic peptides to MHCII. DM role is not only limited to removing CLIP, as a seminal work has shown inefficient antigen presentation in DM-deficient cells (Martin et al., 1996). Kinetic stability studies have shown that DM plays an essential role in generating stable peptide-MHCII (pMHCII) complexes by enhancing the release of less stable peptides. Additional observations indicate that DM-associated half-life is an independent factor for epitope selection. However, the mechanism of DM action is not clearly understood.

MHCII antigen presentation pathway starts when a pathogen is engulfed through endocytosis by APCs and transported through various endosomal vesicles to MIIC. Apart from deeper endo-lysosomal compartments, peptides potentially can bind in the early endosomal vesicles and

possibly on the surface of APCs (Qiu, Xu, Wandinger-Ness, Dalke, & Pierce, 1994; Santambrogio et al., 1999; West, Lucocq, & Watts, 1994). This other antigen presentation pathway is independent of classical factors like an invariant chain and DM, uses recycled MHCII, and it has been suggested that the peptide loading could be happening in a specialized compartment (Lindner & Unanue, 1996; Pinet, Malnati, & Long, 1994; Pinet & Long, 1998). This specialized compartment is an early endosome, where matured or recycled MHCII is present, but in the absence of DM, or a less active form thereof. These MHCII are not complexed to CLIP, and they are internalized from the surface of APCs (Pathak & Blum, 2000; Robinson & Delvig, 2002; ten Broeke, Wubbolts, & Stoorvogel, 2013). One of the studies showed the alternate presentation of HEL 48-61 immunodominant peptide, where the whole protein was presented in the absence of a newly synthesized MHCII molecule, which is a DM-independent pathway as the protein was bound to mature MHCII. This study also showed a difference in the SDS-stability for the HEL protein depending upon where it binds to MHCII. They showed HEL protein formed a SDS-unstable complex at the surface of APCs but a SDS-stable complex when bound to mature MHCII in endosomal compartment (Lindner & Unanue, 1996). MHCII in the presence of SDS buffer separates into α and β chains, but studies have shown that when bound to certain peptides after incubation at neutral pH, MHCII does not separate into its individual chains even in the presence of SDS buffer (Sadegh-Nasseri & Germain, 1991; Springer, Kaufman, Siddoway, Mann, & Strominger, 1977). pH is an essential factor that regulates the biological function of the endosomal compartment and the many molecules acting therein. DM activity is more efficient at late-endosomal pH (4.5 -6.0). The binding and kinetics of pMHCII impacted by the pH of the compartment in which the interaction takes place. The alternate antigen presentation pathway is also modulated by the pH of the compartment. Here, we present

a study in which we investigate the effect of pH on the behavior of DM-susceptible immunodominant peptides and correlate pH-sensitivity to SDS stability. For our research, we selected peptides derived from the hemagglutinin protein of the H1N1 influenza virus for which an HLA-DR1 (DR1)-restricted CD4+ T cell response already established. We measured the half-life of the immunodominant peptides and showed their DM-susceptibility. Past studies have shown that hemagglutinin protein could be presented through a pathway that is independent of newly synthesized MHCII and invariant chain (Pinet et al., 1994). We used peptides derived from hemagglutinin protein with known CD4+ T cell response and measured their kinetics at acidic and neutral pH. We observed that immunodominant peptides showed slower dissociation at neutral pH as compared to acidic pH. These immunodominant peptides were DM-susceptible. We also measured their SDS-stability, and we observed that the majority of immunodominant peptides showed SDS-stability with few exceptions. We propose that these immunodominant peptides are displayed through an alternate antigen presentation pathway, which is DM-independent.

Materials and Methods:

Purification of HLA-DR1 and HLA-DM: HLA-DR1 (DR1) was co-transfected with alpha and beta chain in S2 Drosophila melanogaster cell line. DR1 was purified by using L243 cross-linked protein A sepharose beads (Stern et al., 1994). HLA-DM (DM) was also expressed in the S2 Drosophila melanogaster cell line. M2-anti flag affinity beads were used to purified DM (Hartman et al., 2010). Purified proteins were stored in 50% PBS and 50% glycerol for long term storage.

Peptide synthesis: We selected FAM-labeled 18-mer peptides derived from the hemagglutinin protein of the H1N1 influenza virus for which CD4+ T cell response already measured in ELISPOT assay (Richards et al., 2007). The sequence $HA_{306-318}$ (GPKYVKQNTLKLAT) from influenza A virus H3 subtype was adopted as benchmark or competitor peptide as needed. Anaspec Inc and ABI scientific synthesized peptides. Peptides were synthesized using Fmoc chemistry and fully automated multiple peptide synthesizer. Peptides have greater than 90% purity and verified by reverse-phase HPLC and mass spectrometry. Peptides were dissolved in 10-100% DMSO based on their solubility.

Fluorescence polarization assay: DR1 (5 μM) was incubated with a 4-fold excess of FAM-labeled peptides in citrate phosphate buffer (pH 5.4) for 18-24 hours at 37°C to form peptide/DR1 complexes (pDR1). pDR1 was then purified from unbound peptide by buffer exchange into PBS with Centricon-30 spin filter that had been pre-incubated with 25mM MES (pH 6.5). Purified DR1/peptide complexes were then quantified by reading the UV absorbance @ 280nm. Purified DR1/peptide complexes (100 nM) were then incubated with ten μM unlabelled HA peptide and 300 nM of DM when needed at 37°C in a citrate-phosphate buffer (pH 5.4). The reaction was carried out in a black polystyrene 96-well plate and was covered with

mineral oil to prevent evaporation. Measurements were performed using Wallac VICTOR counter (Perkin Elmer Wallac) with the excitation wavelength = 485 nm and emission wavelength = 535 nm. Specific control groups included peptide only and buffer only and used for background correction. FP values transformed to fraction of bound peptide with the equation: P bound = $(FP_x - FP_{free})/(FP_{bound} - FP_{free})$, where FPx indicates the value of FP measured at t=x minutes, FP_{free} suggests the value of FP relative to free peptide, and FP_{bound} indicates the value of fluorescence polarization of the complex (Ferrante, Templeton, Hoffman, & Castellini, 2015). The fraction of bound peptide was then plotted against time and fit one- or a two-phase exponential function for half-life calculation.

SDS stability assay: 5 μM DR1 and a 20-fold excess of FAM-labelled peptides were incubated in phosphate buffer saline (PBS pH 7.4) for 16-24 hours. The unbound peptide was washed away using Centricon-30 spin filters. Complexes were incubated with laemmli buffer without β-mercaptoethanol and incubated for 15 minutes at room temperature. Samples were loaded on 4-15 pre-cast polyacrylamide gel and ran at 130 V for an hour. The gels were scanned using a Typhoon scanner.

Results

Immunodominant peptides showed DM susceptibility: DM-associated half-life is historically adopted as a measure of immunodominance/immunogenicity. It is well established that DM generates kinetically and energetically stable pMHCII complexes, which are presented to CD4+ T cells. Here, we adopted a set of peptides with known CD4+ T cell responses to ascertain whether peptides with greater T cell responses are resistant to DM-mediated release. We incubated purified soluble DR1 with excess fluorescein-labelled peptides overnight at 37ºC. The excess peptide was removed by washing in centrifugal filtering microdevices with MWCO of 30 kDa. 100 nM of complexes were loaded in black, low-retention plates without or with 300 nM of DM. The release of the peptide was measured overnight. FP values were converted into a fraction of bound peptide and plotted against the time to calculate the half-life of each pDR1 The comparison between the half-life in the absence and presence of DM (as reported in Table 4) shows that the majority of peptides released faster in the presence of DM irrespective of their T-cell response. Past studies have shown that CD4+ T cell epitopes feature a DM-associated half-life > 6h, which is equivalent to the time the MHCIIs spend in the peptide loading compartment (Yin, Calvo-Calle, Dominguez-Amorocho, & Stern, 2012). We observed that none of the immunodominant peptides under scrutiny showed a half-life of 6 hours or greater. This result indicates that these peptides are susceptible to DM action and therefore do not follow the canonical requirement for MHCII selection and presentation.

Table 3.1: Kinetic stability of peptides derived from hemagglutinin protein. The half-life of the peptides measured at pH 5.4 in the absence and presence of DM.

Peptide Category	Peptide Name	T1/2 (minutes) (-DM)	T1/2 (minutes) (+DM)
Immunodominant	H1A23	253.12	257.81
	H1A24	4906.25	105.46
	H1A30	857.81	213.28
	H1A63	2162.1	328.12
	H1A64	3886.71	89.84
	H1A440	5027.34	239.06
Subdominant	H1A14	2625	960.93
	H1A38	918.75	2187.5
Weak	H1A08	839.06	42
	H1A09	6289.06	109.37
	H1A47	796.87	49
Negative	H1A17	10828.125	437.5
	H1A36	17648.43	546.87
	H1A35	1804.68	176.56
	H1A39	1914.06	847.65
	H1A48	857.81	49
	H1A79	5437.75	1570.31

SDS-stability for immunodominant peptide abolishes at acidic pH: Lindner et al. have shown that HEK protein SDS-stability changes after being processed by APCs (Lindner & Unanue, 1996). The partially folded HEK protein, when bound to recycled MHCII on the surface of the APCs, form unstable SDS-complex, but when they internalized, the complex dissociated. When the HEK internalized and bound to recycled MHCII in the endosome, it generates an SDS-stable complex while when HEK is internalized and processed through newly synthesized MHCII they generate SDS-stable HEK peptides. This study shows that there is a distinct pathway for a protein and its derived peptides based on the compartment where it internalizes and what kind of antigen processing machinery it exposes to. They showed the presentation of an antigen in DM-independent pathway where antigen bound to recycled MHCII. In our study we see that immunodominant peptides showed DM-susceptibility, which suggested that presentation of these

peptides might be DM-independent. Past studies have shown that SDS-stability is one of the properties, which is a sign of MHCII maturity, and immunodominant peptides showed SDS stability as one of the characteristics (Nelson, Petzold, & Unanue, 1993; Nelson, Roof, McCourt, & Unanue, 1992; Verreck et al., 1996). This suggests that SDS-stability may be required for efficient antigen presentation. To test if our peptides showed SDS-stability irrespective of DM-susceptibility, we tested SDS-stability for our peptides (mentioned in Table 4) which are derived from hemagglutinin protein of H1N1 influenza virus. Past studies have examined the SDS-stability at pH 7.4, but here we also tested the peptides SDS-stability when they formed at pH 5.4. We incubated the fluorescent-labeled peptides with DR1 overnight at 37°C. The unbound labeled peptide was removed from the bound complex by extensive buffer wash. Complexes were incubated for 15 minutes at room temperature with a laemmli buffer without BME. Samples ran on SDS-gel electrophoresis for an hour. Fluorescence was detected using a Typhoon scanner. The immunodominant peptides H1A440, H1A63, and H1A30 showed SDS stability along with the negative peptides H1A17 and H1A36, as shown in Figure 3.1A and B. We then tested SDS-stability for these peptides at pH 5.4, as shown in Figure 3.1C. We incubated the complexes at pH 5.4, and unbound peptide washed with PBS. Complexes were analyzed by SDS-PAGE. SDS stability was lost for H1A30 and negative peptide H1A36 when complexes were formed at pH 5.4 but not for the immunodominant peptide H1A440, H1A64 and negative peptide H1A17.

Kinetic and SDS-stability of immunodominant peptides dependent on the length of the peptides: Next, we wanted to test if the length of the peptides can alter the kinetics and SDS-stability of the above-tested peptides. Past studies have shown that peptide length influences the binding of the peptides to MHCII (O'Brien, Flower, & Feighery, 2008). We tested SDS-stability

and kinetic stability of new 15-mer peptides derived from immunodominant H1A30 and H1A63 by shortening the flanking ends. We also tested longer immunodominant peptides consisting of the overlapping H1A23, H1A24 (H1A2324), and H1A63, H1A64 (H1A6364). H16364 showed SDS-stability at pH 5.4 and 7.4 (Figure 3.1B and C). 15-mer derivatives of peptides H1A30 and H1A63 showed SDS-stability at pH 7.4 and 5.4 (Figure 3.1B and D). These observations suggest that based on peptide lengths the antigen could be presented through different presentation pathways.

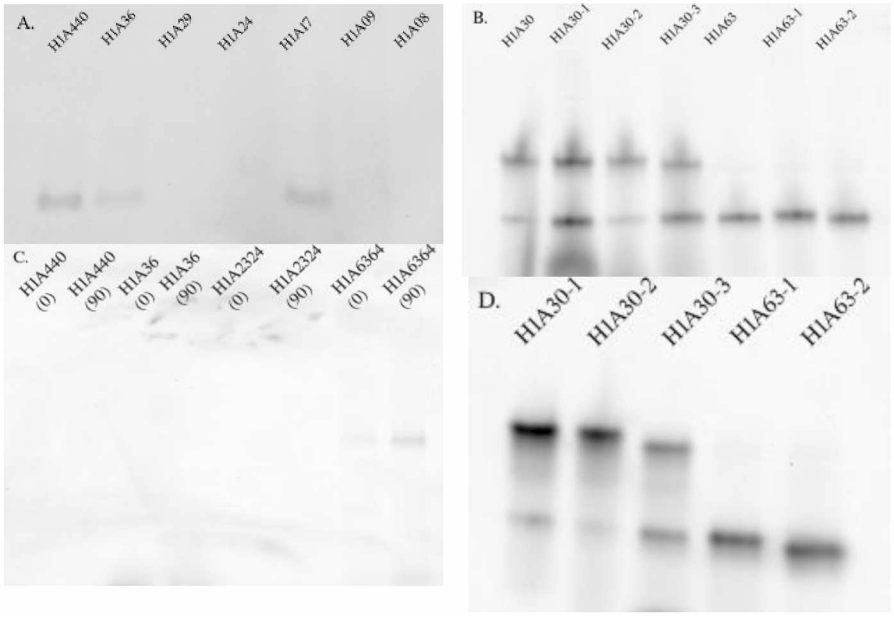


Figure 3.1: SDS-stability of the peptides differs based on the pH of the reaction: 5 μM of DR1 incubated with a 20-fold excess of fluorescence-labeled peptides overnight in PBS (A). SDS-stability of H1A440, H1A36, H1A29, H1A24, H1A17, H1A09 and H1A08 (B). SDS-stability of H1A30, H1A30-1, H1A30-2, H1A30-3, H1A63, H1A63-1 and H1A63-2. SDS-stability of peptides incubated in citrate-phosphate buffer (pH 5.4). The sample was also incubated with DM at time 0 and time 90 min. (C). SDS-stability of H1A440, H1A36, H1A2324 and H1A6364 (D). SDS-stability of H1A30-1, H1A30-2, H1A30-3, H1A63-1 and H1A63-2

We measured the half-life of 15-mer derivatives of H1A30 and H1A63 (Table 5). Two of the 15-mer derivatives of H1A30 showed approximately 3-to-4-fold increase in half-life in the presence of DM. In the case of H1A63 derivatives, the half-life increase showed only a 2-fold increase for

one of the derivatives. These results suggest that in the case of H1A30 and H1A63 the length of the peptides might be an essential factor in dictating DM activity.

Table 3.2: Kinetic stability of 15-mer derived from 30 and 63. The half-life of the peptides measured at pH 5.4 in the absence and presence of DM.

Peptide Category	Peptide Name	T1/2 (minutes) (-DM)	T1/2 (minutes) (+DM)
H1A30 Derivative	30-1	11473	866.25
	30-2	11503	770
	30-3	6930	346.5
H1A63 Derivative	63-1	6930	223.5
	63-2	11629	533.07

Immunodominant peptide H1A440, H1A30, and H1A23 are more stable at neutral pH: To test the hypothesis that immunodominant peptides H1A440 and H1A30 presented by binding at the surface of APCs, we incubated those peptides with DR1 at pH 7.4 and measured their release at pH 7.4 without DM. We observed that H1A440 half-life increased almost 7-fold, H1A23 up to 4-fold, and H1A30 increased up to 2-fold (Table 6). H1A63 half-life did not increase significantly.

Table 3.3: Kinetic stability of selected immunodominant peptides measured at pH 7.4. Both incubation and release measured at pH 7.4.

Peptide Name	T1/2 (minutes) (-DM)
H1A23	832.03
H1A30	2128.90
H1A63	1992.18
H1A440	34179.68

Discussion

The classical pathway of MHCII antigen presentation involves newly synthesized MHCII molecule, DM removal of CLIP from MHCII binding site, and exchange of antigen-derived peptides in the MIIC. Past studies have shown the existence of an alternate antigen processing pathway in which the antigen is presented through recycled MHCII at the surface of APCs (Pinet et al., 1994; Pinet, Vergelli, Martin, Bakke, & Long, 1995; Pinet & Long, 1998). This alternative pathway was analyzed for the processing and presentation of peptides generated from hemagglutinin protein, myelin basic protein and HEK protein (Lindner & Unanue, 1996; Pinet et al., 1994; Vergelli et al., 1997). Here, we tested a set of peptides with known CD4+ T cell responses derived from hemagglutinin protein for DM activity. Past studies have shown that DM-associated half-life is an independent factor of epitope prediction and in their studies, epitopes showed at least 6 hours of half-life in the presence of DM (Yin et al., 2012). The set of immunodominant peptides here analyzed showed a half-life of fewer than 6 hours in the presence of DM. We propose that these specific HA-derived immunodominant peptides is presented through an alternate, DM-independent pathway.

In our study, we selected peptides derived from H1N1 hemagglutinin protein, for which a DR1-restricted immunodominance hierarchy is already established. We measured the half-life of these peptides in the absence and presence of DM at pH 5.4. Unexpectedly, immunodominant peptides, along with weaker peptides, showed DM susceptibility. We also tested the SDS-stability of these peptides and observed that immunodominant sequences H1A30, H1A63, and H1A440 showed SDS-stability along with negative peptides H1A17, H1A36. The effect of pH tested along with SDS-stability showed that the SDS-stability of immunodominant peptide H1A30 abolishes at acidic pH. In contrast, the 15-mer derivatives tested within the

immunodominant peptides H1A30 showed SDS-stability even at acidic pH. We also measured the kinetic stability of these 15-mers at pH 5.4. These peptides showed greater kinetic stability in the presence of DM than the peptides from which they were originated. These results suggest that the observed DM susceptibility of these peptides could be due to the length of these peptides and can be presented through alternate pathways.

To test our hypothesis that these immunodominant peptides could be binding at the surface of APCs, we measured the half-life of these peptides at pH 7.4. H1A440 and H1A30 showed slower release at pH 7.4. DM is more active at acidic pH as these peptides are more stable at pH 7.4; these results suggest that these peptides presentation will be DM independent.

Our data showed that the antigen presentation of the hemagglutinin protein could be through various pathways. Currently, for immunodominant epitope selection, there are two models: kinetic stability and other one epitope accessibility. Kinetic stability model proposed that immunodominant pMHCII complexes have intrinsic stability, and DM helps in the selection of kinetically stable complexes (Yin et al., 2012; Yin, Maben, Becerra, & Stern, 2015). Epitope accessibility model suggests that immunodominant epitope selection is dictated by structural features of protein and accessibility to MHCII binding groove. This model indicates two paths for epitope accessibility: bind first, cut later and cut first, bind later (Kim et al., 2014; Kim & Sadegh-Nasseri, 2015). The classical pathway of MHCII presentation describes that antigens are first cleaved by cathepsins and later bind to MHCII in MIIC. Our study has two significant findings: 1. Kinetic stability of immunodominant peptides is pH-dependent which could rescue them from DM susceptibility 2. Immunodominant peptides stability is dependent on peptide length. Based on these findings we suggest that these immunodominant peptides are presented through alternate pathways where they would escape DM activity either by binding at the surface

of APCs or by further trimming via cathepsins. Our results support a model in which multiple cross-over pathways coexist within the APC and the outcome of the process is a consequence of the extent to which they can modulate peptide-intrinsic properties such as binding to MHCII, cleavability and DM-susceptibility.

References

Ferrante, A., Templeton, M., Hoffman, M., & 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*(3), 1251-1261. doi:10.4049/jimmunol.1402367

Hartman, I. Z., Kim, A., Cotter, R. J., Walter, K., Dalai, S. K., Boronina, T., . . . Sadegh-Nasseri, S. (2010). A reductionist cell-free major histocompatibility complex class II antigen processing system identifies immunodominant epitopes. *Nat Med, 16*(11), 1333-1340. doi:10.1038/nm.2248

Kim, A., Hartman, I. Z., Poore, B., Boronina, T., Cole, R. N., Song, N., . . . Sadegh-Nasseri, S. (2014). Divergent paths for the selection of immunodominant epitopes from distinct antigenic sources. *Nat Commun, 5*, 5369. doi:10.1038/ncomms6369

Kim, A., & Sadegh-Nasseri, S. (2015). Determinants of immunodominance for CD4 T cells. *Curr Opin Immunol, 34*, 9-15. doi:10.1016/j.coi.2014.12.005

Lindner, R., & Unanue, E. R. (1996). Distinct antigen MHC class II complexes generated by separate processing pathways. *EMBO J, 15*(24), 6910-6920.

Martin, W. D., Hicks, G. G., Mendiratta, S. K., Leva, H. I., Ruley, H. E., & Van Kaer, L. (1996). H2-M mutant mice are defective in the peptide loading of class II molecules, antigen presentation, and T cell repertoire selection. *Cell, 84*(4), 543-550. doi:10.1016/s0092-8674(00)81030-2

Nelson, C. A., Petzold, S. J., & Unanue, E. R. (1993). Identification of two distinct properties of class II major histocompatibility complex-associated peptides. *Proc Natl Acad Sci U S A, 90*(4), 1227-1231. doi:10.1073/pnas.90.4.1227

Nelson, C. A., Roof, R. W., McCourt, D. W., & Unanue, E. R. (1992). Identification of the naturally processed form of hen egg white lysozyme bound to the murine major histocompatibility complex class II molecule I-Ak. *Proc Natl Acad Sci U S A, 89*(16), 7380-7383. doi:10.1073/pnas.89.16.7380

O'Brien, C., Flower, D. R., & Feighery, C. (2008). Peptide length significantly influences in vitro affinity for MHC class II molecules. *Immunome Res, 4*, 6. doi:10.1186/1745-7580-4-6

Pathak, S. S., & Blum, J. S. (2000). Endocytic recycling is required for the presentation of an exogenous peptide via MHC class II molecules. *Traffic, 1*(7), 561-569.

Pinet, V., Malnati, M. S., & Long, E. O. (1994). Two processing pathways for the MHC class II-restricted presentation of exogenous influenza virus antigen. *J Immunol, 152*(10), 4852-4860.

Pinet, V., Vergelli, M., Martin, R., Bakke, O., & Long, E. O. (1995). Antigen presentation mediated by recycling of surface HLA-DR molecules. *Nature, 375*(6532), 603-606. doi:10.1038/375603a0

Pinet, V. M., & Long, E. O. (1998). Peptide loading onto recycling HLA-DR molecules occurs in early endosomes. *Eur J Immunol, 28*(3), 799-804. doi:10.1002/(SICI)1521-4141(199803)28:03<799::AID-IMMU799>3.0.CO;2-5

Qiu, Y., Xu, X., Wandinger-Ness, A., Dalke, D. P., & Pierce, S. K. (1994). Separation of subcellular compartments containing distinct functional forms of MHC class II. *J Cell Biol, 125*(3), 595-605. doi:10.1083/jcb.125.3.595

Richards, K. A., Chaves, F. A., Krafcik, F. R., Topham, D. J., Lazarski, C. A., & Sant, A. J. (2007). Direct ex vivo analyses of HLA-DR1 transgenic mice reveal an exceptionally broad pattern of immunodominance in the primary HLA-DR1-restricted CD4 T-cell response to influenza virus hemagglutinin. *J Virol, 81*(14), 7608-7619. doi:10.1128/JVI.02834-06

Robinson, J. H., & Delvig, A. A. (2002). Diversity in MHC class II antigen presentation. *Immunology, 105*(3), 252-262. doi:10.1046/j.0019-2805.2001.01358.x

Rothbard, J. B., & Gefter, M. L. (1991). Interactions between immunogenic peptides and MHC proteins. *Annu Rev Immunol, 9*, 527-565. doi:10.1146/annurev.iy.09.040191.002523

Sadegh-Nasseri, S., & Germain, R. N. (1991). A role for peptide in determining MHC class II structure. *Nature, 353*(6340), 167-170. doi:10.1038/353167a0

Santambrogio, L., Sato, A. K., Carven, G. J., Belyanskaya, S. L., Strominger, J. L., & Stern, L. J. (1999). Extracellular antigen processing and presentation by immature dendritic cells. *Proc Natl Acad Sci U S A, 96*(26), 15056-15061. doi:10.1073/pnas.96.26.15056

Springer, T. A., Kaufman, J. F., Siddoway, L. A., Mann, D. L., & Strominger, J. L. (1977). Purification of HLA-linked B lymphocyte alloantigens in immunologically active form by preparative sodium dodecyl sulfate-gel electrophoresis and studies on their subunit association. *J Biol Chem, 252*(17), 6201-6207.

Stern, L. J., Brown, J. H., Jardetzky, T. S., Gorga, J. C., Urban, R. G., Strominger, J. L., & Wiley, D. C. (1994). Crystal structure of the human class II MHC protein HLA-DR1 complexed with an influenza virus peptide. *Nature, 368*(6468), 215-221. doi:10.1038/368215a0

ten Broeke, T., Wubbolts, R., & Stoorvogel, W. (2013). MHC class II antigen presentation by dendritic cells regulated through endosomal sorting. *Cold Spring Harb Perspect Biol, 5*(12), a016873. doi:10.1101/cshperspect.a016873

Vergelli, M., Pinet, V., Vogt, A. B., Kalbus, M., Malnati, M., Riccio, P., . . . Martin, R. (1997). HLA-DR-restricted presentation of purified myelin basic protein is independent of intracellular processing. *Eur J Immunol, 27*(4), 941-951. doi:10.1002/eji.1830270421

Verreck, F. A., Vermeulen, C., Poel, A. V., Jorritsma, P., Amons, R., Coligan, J. E., . . . Koning, F. (1996). The generation of SDS-stable HLA DR dimers is independent of efficient peptide binding. *Int Immunol, 8*(3), 397-404. doi:10.1093/intimm/8.3.397

West, M. A., Lucocq, J. M., & Watts, C. (1994). Antigen processing and class II MHC peptide-loading compartments in human B-lymphoblastoid cells. *Nature, 369*(6476), 147-151. doi:10.1038/369147a0

Yin, L., Calvo-Calle, J. M., Dominguez-Amorocho, O., & Stern, L. J. (2012). HLA-DM constrains epitope selection in the human CD4 T cell response to vaccinia virus by favoring the presentation of peptides with longer HLA-DM-mediated half-lives. *J Immunol, 189*(8), 3983-3994. doi:10.4049/jimmunol.1200626

Yin, L., Maben, Z. J., Becerra, A., & Stern, L. J. (2015). Evaluating the Role of HLA-DM in MHC Class II-Peptide Association Reactions. *J Immunol, 195*(2), 706-716. doi:10.4049/jimmunol.1403190

Chapter 4: Distinct Assembly of full-length HLA-DR1 into nanodisc depends on the lipid composition[3]

Abstract

Class II major histocompatibility complex molecules (MHCII) are transmembrane glycoproteins found on the surface of antigen-presenting cells. MHCII expression is not limited to the cell surface, but it found in endosomal compartments, which have their signature membrane characteristics. They display to $CD4^+$ T cells peptides that have been generated and selected by intracellular antigen processing and presentation mechanisms, thus initiating an adaptive immune response. The traditional strategy to investigate peptide binding to MHCII has relied on the expression and purification of soluble MHCII, in which the transmembrane portion of the protein is removed. One question address whether membrane-embedding of native MHCII impacts its function and interactions with peptides. To this end, full-length human MHCII allele HLA-DR1 (DR1) was isolated from B-lymphoblastoid cell lines via immunoaffinity chromatography and subsequently incorporated into nanodiscs, synthetic model membrane device, to evaluate the potential effects of membrane lipid composition on MHCII assembly. Three types of nanodisc were generated: simple, fluid disordered, and rigid ordered, each of them characterized by unique lipid composition. Nanodiscs were separated using fast protein liquid chromatography (FPLC), indicating apparent differences between nanodisc types. Whereas fluid disordered nanodisc suggested DR1 assembly as a cluster (one major FPLC-generated peak), both rigid and straightforward nanodiscs revealed multiple peaks. As DR1 tends to form aggregates, we inferred that DR1 formed tetramers and dimers in simple and rigid nanodiscs. Our results indicate that membrane lipid composition has a substantial impact on native MHCII assembly

[3] Osan J.K., Rivera K., Ferrante A., Kuhn T. B. (In prep) Distinct assembly of full-length HLA-DR1 into nanodisc depends on the lipid composition.

and possibly peptide interaction. We propose that MHCII conformation and activity are a function of the cell compartment where they reside at any point in time.

Introduction

Class II major histocompatibility complex (MHCII) molecules are transmembrane glycoproteins expressed on the membrane of antigen-presenting cells (APCs) and endosomal compartment (Guillet, Lai, Briner, Smith, & Gefter, 1986). Upon infection, APCs uptake the pathogen and via the endo-lysosomal system, process the pathogenic proteins into smaller peptides, which can then attempt to bind to MHCII in the MIIC compartment. Selected peptide: MHCII complexes are presented to CD4+ T cells on the surface of APC, for engagement of the adaptive cellular immunity (Thomas, Hsieh, Schauster, Mudd, & Wilner, 1980).

MHCII is a heterodimeric membrane protein made up of α and β chains. These chains are glycosylated and have an extracellular domain, which contains the peptide-binding site, a transmembrane, and a cytoplasmic region (Gorga, Horejsi, Johnson, Raghupathy, & Strominger, 1987). The cytoplasmic domain of MHCII plays an essential role in signaling to B cells (Harton & Bishop, 1993; Wade, Ward, Rosloniec, Barisas, & Freed, 1994). Cytoplasmic domain and the transmembrane portion of MHCII play a vital role in the efficient expression of MHCII on the plasma membrane (Wade et al., 1994). MHCII molecules incorporate into two types of microdomain: (i). Cholesterol and glycosphingolipid-enriched domains denoted lipid rafts (ii). Microdomains made up of tetraspan proteins, which enrich MHCII molecule loaded with peptides(Vogt, Spindeldreher, & Kropshofer, 2002). Studies have shown that the peptide-MHCII complexes are presented on dendritic cell surfaces in cholesterol-rich micro clusters and these clusters play an essential role in activating CD4 T cells (Bosch, Heipertz, Drake, & Roche, 2013).

MHCII molecules tend to aggregate in the absence of peptide in their binding groove. Newly synthesized MHCII are rescued from such aggregation by forming a complex with invariant

chain, often as a nonamer ($(\alpha\beta)_3Ii_3$). Previous crystal structures have shown that MHCII can exist as a dimer of heterodimer (superdimers) (Brown et al., 2015). Studies have shown the presence of superdimers on the surface of mouse B cells and these superdimers show thermal and pH stability like MHCII dimer (Schafer, Malapati, Hanfelt, & Pierce, 1998). Superdimers were shown to be involved in T cell response to low-affinity antigens (Schafer & Pierce, 1994).

To study the binding of MHCII and antigenic peptides, the extracellular domain of MHCII is usually expressed in insect cells and purified as a soluble protein. Studies have shown that the empty human MHCII, HLA-DR1 (DR1) expressed by insect cell line tends to aggregate, although it can be rescued by incubating with antigenic peptide (Stern & Wiley, 1992; Yin, Maben, Becerra, & Stern, 2015). Our group has observed that purified DR1 shows a band at 120 kDa, and one at ~250 kDa, along with the usual dimer band at 56 kDa (data not shown here). It is possible that these aggregate or superdimers are DR1 bound to antibody (Hitzel, Gruneberg, van Ham, Trowsdale, & Koch, 1999); however, the DR1 fractions above 100 kDa are capable of binding the antigenic peptide $HA_{306-318}$ (data not published), suggesting that these heavier fractions are DR1 superdimers. To understand how DR1 assembles in the membrane and whether superdimers also assemble in the membrane, we adopted nanodiscs as a surrogate of cell membranes.

Nanodiscs are a synthetic membrane model currently widely used to study the membrane proteins in their soluble form. They are composed of phospholipids encircled by amphipathic membrane scaffold protein (MSP). Nanodiscs are a self-assembly system where detergent solubilizes component of nanodisc assembled when detergent is slowly removed (Denisov & Sligar, 2017). Studies have shown that the assembly of the membrane proteins is dependent on the lipid environment (Amin & Hazelbauer, 2012). Phosphatidylcholine is the most used

synthetic lipids in nanodisc assembly. Cholesterol, in combination with phosphatidylcholine, has been used previously for nanodisc preparation. The fluidity of the membrane depends upon the type and composition of lipids. Lipid rafts are sphingomyelin- and cholesterol-rich domains (Koukalova et al., 2017). By manipulating lipid ratios, fluid or rigid and ordered or disordered phases can be achieved in the nanodisc.

Here, we report the generation of a technology whereby full-length MHCII is used in a soluble form for application in biochemical assays. To this aim, we have taken advantage of an already established nanodisc assembly system. Full-length DR1 molecules from human B-cell lymphoblastoid cell lines were purified, which were subsequently embedded in nanodiscs during assembly. We also analyzed the effect of the lipid environment on the assembly of DR1 and its superdimers by changing the extent of fluidity and order of the nanodiscs. We observed a significant difference in DR1 assembly when different compositions of lipids used as well as the successful assembly of DR1 superdimers. These results suggest that the DR1 assembly within membranes is dependent on the lipid environment, and MHCII organization might be different in different compartments as a function of lipid composition.

Materials and methods

Proteins and Lipids: Full-length HLA-DR1 (mDR1) was purified from B-cell lymphoblastoid cell lines (B-LCL), which were acquired from Fred Hutch research cell bank.

MSP1E3D1 protein (M7074-5MG) was ordered from Sigma-Aldrich. The protein comes in a lyophilized form, which was stored in -20°C. The protein was dissolved in sterile water when ready for use.

1-palmitoyl-2-oleoyl-glycero-3-phosphocholine (POPC), cholesterol and sphingomyelin lipids were used for the preparation of nanodisc. Lipids were purchased from Avanti Polar Lipids. They were dissolved in chloroform and evaporated using dry N_2. The dried lipids were further dissolved using cholate buffer (100 mM sodium cholate, 20 mM Tris-CL, 100 mM NaCl buffer).

Purification of full-length DR1 from B-LCL: B-LCL was cultured in the laboratory, and mDR1 was purified from the cell culture. The cell suspension was spun down at 4°C, and intact cells were dissolved in 0.3 M sucrose lysis buffer. Pellet was homogenized by using a Dounce homogenizer and further spun down to collect clear supernatant. Pellet was also suspended using lysis buffer without sucrose, and the clear supernatant was collected. This clear supernatant is used to purify mDR1 by using immunoaffinity chromatography. The protein was stored in the 10 mM Tris CL with 0.1% deoxycholate pH 8.0.

Nanodisc Preparation: Three types of nanodiscs as defined by their lipid compositions prepared: simple, fluid disordered, and rigid ordered. Ratios of lipids, MSPE3D1, and mDR1 used in the preparation of nanodisc reported in Table 7.

Table 4.1: Description of lipids, MSPE3D1 and mDR1 ratios used in the preparation of different nanodiscs

Nanodisc	Lipids	MSPE3D1	mDR1
Simple	100 (POPC)	5	1
Simple	1000 (POPC)	10	1
Fluid-disordered	1000 (POPC: PSM: Cholesterol: 60: 1: 1)	10	1
Rigid-ordered	1000 (POPC: PSM: Cholesterol: 1: 1: 1)	10	1

To prepare nanodiscs, we made a reaction mix with the appropriate ratios of lipids, MSPE3D1, and mDR1, depending on the type of nanodisc. The reaction mixture was prepared in 100 mM sodium cholate, 20 mM Tris-CL, 100 mM NaCl buffer, and stored at 4^{o}C for simple and fluid disordered nanodisc and @ RT for rigid disordered nanodisc for one hour. The detergent from the reaction mix was separated by incubating it with Biorad SM2 beads for 3 – 4 hours. Beads were removed from the mix, which finally contained nanodisc. Nanodiscs were stored at 4^{o}C.

Separation of nanodisc by using fast protein liquid chromatography (FPLC): To purify the nanodisc from any unbound protein and lipids, we used FPLC with an SEC 650 column. The separation was performed in 20 mM Tris-CL, 100 mM NaCl, pH 7.4 at a flow rate of 0.5 ml/min, and each aliquot of nanodisc were collected and further concentrated using Millipore centricon filter (MWCO 30 kDa).

Identification of nanodisc using gel electrophoresis: The FPLC-purified fractions were visualized in native polyacrylamide gel electrophoresis (PAGE) and SDS-PAGE to confirm the presence of nanodiscs. For both runs, we used mDR1 and MSPE3D1 alone as a control. Gels were stained with silver stain kit.

Identification of nanodisc using immunoblotting: Samples were run on 4-15% Tris-Glycine pre-cast gel purchased from Bio-Rad. 10 ug of total protein was loaded in each well. Proteins were resolved by SDS-electrophoresis by running at 120 Volts for 1 hour. Proteins were

transferred to a nitrocellulose membrane by running at 25 Volts for 1 hour 45 minutes in 1X transfer buffer (25 mM Tris, 192 mM Glycine, 20% methanol). Membranes were blocked for 2 hours at room temperature using 5% BSA in PBS. Blocked membranes were probed with primary antibodies against full-length DR1 (MEM-267 – monoclonal antibody raised in mouse-specific to empty form of HLA-DR1) and MSP (anti-his monoclonal antibody raised in rabbit specific against 6*-His tag) by incubating overnight at 4^{o} C. Primary antibodies were detected by using secondary antibodies (goat-anti-mouse IgG Alkaline phosphatase for anti-DR1 and goat-anti-rabbit IgG Alkaline phosphatase for anti-his). Blot was developed by adding 5 ml of NBT/BCIP substrate solution.

Direct binding assay: 10 nM of nanodisc and control proteins was incubated with various concentration of biotinylated HA (bioHA) peptide in phosphate buffer saline (PBS) pH 7.4 (0.1% BSA, 0.01% Tween 20, 0.1 mM iodoacetamide, 5 mM EDTA, 0.02% NaN_3) for 5 days at 37^{o}C. For controls, soluble DR1 used as a positive control, and an empty nanodisc was used as a negative control. We tested fluid nanodisc and full-length DR1 alone for binding to bioHA. BioHA was serially diluted from 2.98 pM to 200 μM concentration. 20 nM of protein and nanodisc stock were prepared in PBS. In the reaction plate, 100 μl of protein solution and 100 μl of bioHA were added and incubated for 5 days at 37ºC. L243 antibody was used to capture DR1, and an anti-his antibody was used to capture nanodisc. The incubation time ensured binding reaction to reach equilibrium. The bound biotinylated peptide was detected using a solid-phase immunoassay, and Eu^{2+} labeled streptavidin. Plates were read using a Wallac VICTOR counter (PerkinElmer Wallac). Each experiment performed in quadruplicate.

Binding with labeled peptides: 5 μM protein or nanodisc were incubated with 100 μM of a fluorescein-labeled peptide ($HA_{306-318}$) in 20 mM Tris-CL and 100 mM sodium chloride buffer

(pH 7.4) overnight at 37°C. The unbound peptide was removed from the bound peptide by washing the complex with buffer ten times using an ultra-centrifugal filtering device with 30 kDa cutoff. 10 µl of the complex was incubated with 10 µl of native loading buffer and loaded on precast 4-15% Tris-Glycine gel. The gel was run using a native running buffer for 130 Volts for an hour. The labeled peptide was detected using a UV-transilluminator. Proteins were stained by using a silver stain kit. Silver stain and fluorescent bands positions were compared to identify the binding. Fluorescent bands were quantified by using ImageJ.

Results

Assembly of full-length DR1 using one lipid shows differences based on the amount of lipid used in the nanodisc assembly mix: Nanodisc is a synthetic membrane model adopted to investigate the biophysics and biochemistry of membrane proteins. In this study, we used the nanodisc as a surrogate of the cell membrane to embed full-length MHCII proteins and investigate whether their conformations differ as a function of the nanodisc lipidic composition and ratio. As the MHCII allele, we used the commonly studied and widely available DR1. We purified the full-length DR1 from the B-lymphoblastoid cell line using immunoaffinity chromatography. First, we prepared simple nanodisc with one lipid (POPC) in the reaction mix to ascertain whether DR1 assemble in the nanodisc. The details of assembly mix are shown in Table 7. When we used 100: 5: 1 ratio (POPC: MSP1E3D1: mDR1), we observed on the FPLC chromatogram (Figure 4.1A) one major peak and two smaller peaks before the peak expected to contain the nanodisc assembled with DR1. We ran the FPLC fractions on native (Figure 4.1B) and SDS (Figure 4.1C) gel to identify the fraction-containing nanodisc. Nanodisc formation is confirmed by comparing nanodisc fractions with MSP and DR1 alone control lane on SDS and native PAGE. Nanodisc fractions run differently on native PAGE as compared to controls. In SDS PAGE we observed both MSP and DR1 band in the nanodisc fractions. For earlier small peak we found that tetramer and dimers of DR1 first assembled, and in the later fractions, all the forms of DR1 assembled. The most significant peak fractions show assembly of monomer DR1 in majority amount. These observations indicate that different nanodiscs favor the incorporation of different conformations of DR1. Next, we prepared the nanodisc by using a different ratio of assembly mix 1000: 10: 1 ratio (POPC: MSP1E3D1: mDR1). As shown in Figure 4.1D, the FPLC chromatogram shows one major peak with a broader small peak before the major peak.

When analyzed in SDS-PAGE (Figure 4.1E) we observed that the significant peak fraction contains all conformation of DR1 along with MSP. Taken together, these results indicate that the composition of the nanodisc assembly mix determines the conformation of DR1 embedded in the assembly.

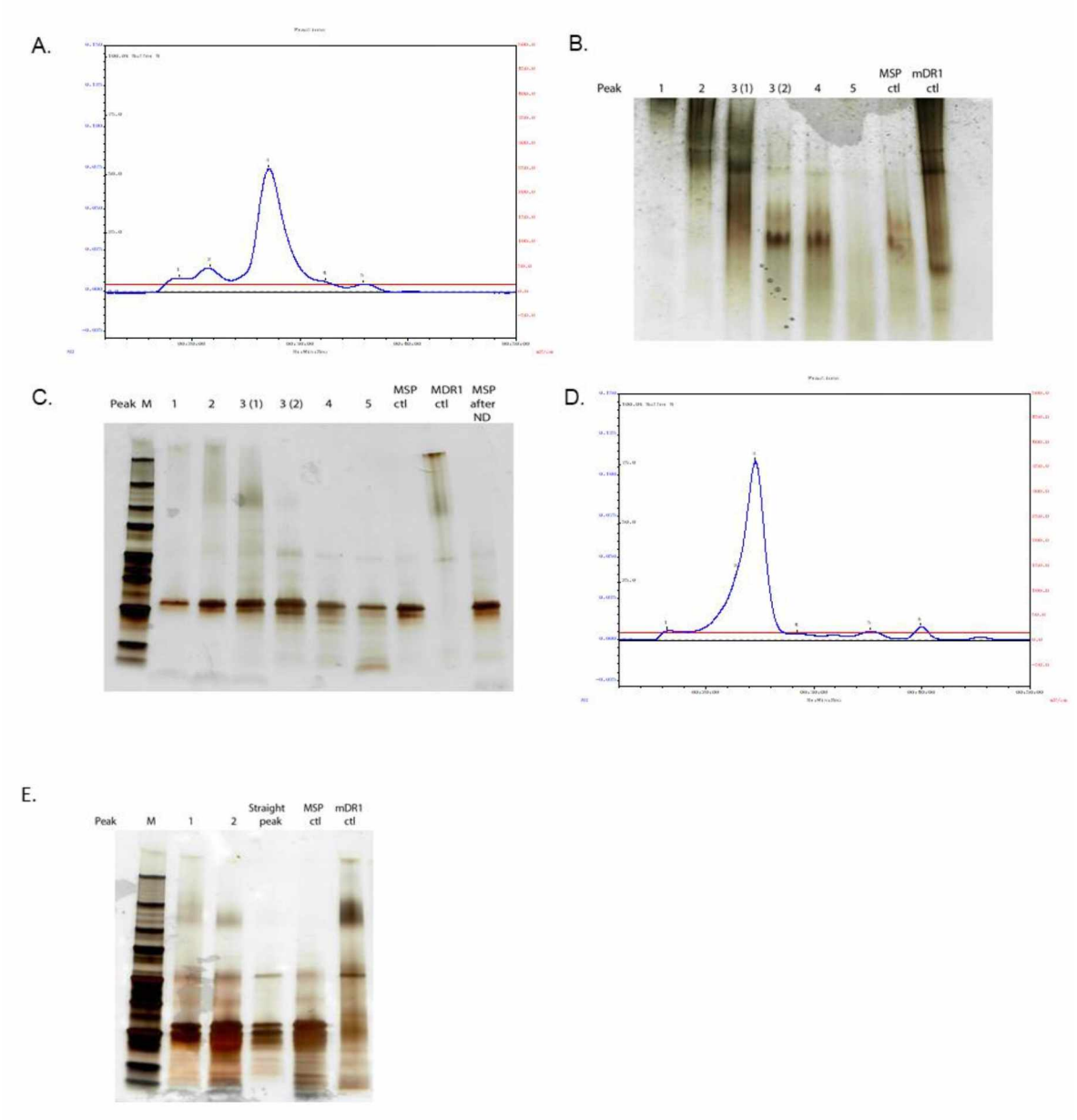


Figure 4.1: Assembly of full-length DR1 in the simple nanodisc. Assembly of DR1 differs as a function of lipid ratio in the nanodisc mix. Ratio of POPC: MSPE3D1: mDR1 (100: 5: 1) (A). FPLC chromatogram. (B). Native gel (C). SDS-PAGE gel Ratio of POPC: MSPE3D1: mDR1 (1000:10:1) (D). FPLC Chromatogram (E). SDS-PAGE gel

Assembly of full-length of DR1 in fluid-disordered nanodisc differs from rigid-ordered nanodisc: Lipid rafts are sphingomyelin- and cholesterol-enriched platforms in the plasma membrane (Koukalova et al., 2017). Based on the composition of the lipid ratio, the system can be rigid-ordered or fluid-disordered. Studies have shown that MHCII associates with lipid rafts membrane microdomains on the surface of APCs, and this association is essential for their ability to stimulate CD4+ T cells (Bosch et al., 2013). To mimic the MHCII in lipid rafts and study the effect of membrane composition on assembly of MHCII, we prepared two types of nanodiscs: fluid-disordered and rigid-ordered. Composition of fluid and rigid nanodiscs are shown in Table 1. As shown in Figure4.2A, for fluid-disordered nanodiscs, we observed one single peak on the FPLC. The collected fractions run on native (Figure 4.2B) and SDS (Figure4.2C) gel revealing that all the fractions of the major peak have monomer and superdimers DR1 conformation in one form of nanodisc. Next, we prepared rigid-ordered nanodiscs with DR1 and analyzed with FPLC for separation and SDS gel for identification. We observed a major peak on the chromatogram (Figure 4.2D), but the peak was separated at the top making it look like a double peak. On the SDS gel (Figure 4.2E) we observed that the first half-peak has more amount of superdimers DR1 while the second peak contains mostly monomer in the nanodisc. Overall, the FPLC chromatogram showed distinct differences between fluid and rigid nanodisc suggesting that the full-length DR1 assembly differs based on the composition of lipids in the membrane. All the nanodiscs were also analyzed by immunoblotting to identify the proteins. Nanodisc formation was confirmed by immunoblotting, and both MSP and full-length DR1 bands were observed in immunoblotting (Figure4.2F).

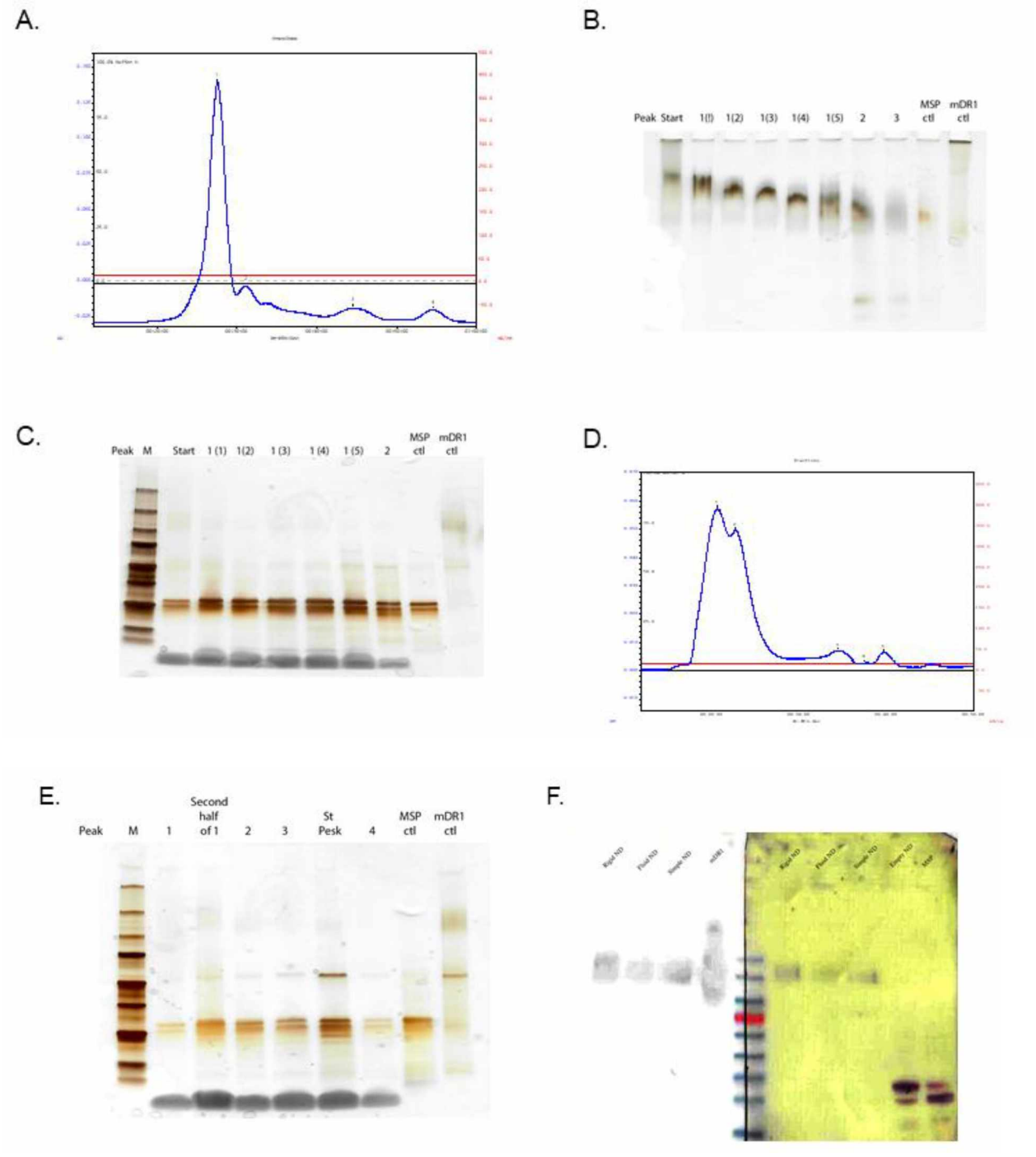


Figure 4.2: Assembly of full-length DR1 differs based on the composition of lipids. Fluid disordered nanodisc, ratio of lipids (POPC: PSM: Cholesterol (60:1:1)): MSPE3D1: mDR1 (1000: 10: 1) (A) FPLC chromatogram (B). Native gel (C). SDS-PAGE gel. Rigid ordered nanodisc, ratio of lipids (POPC: PSM: Cholesterol (1:1:1)): MSPE3D1: mDR1 (1000: 10: 1) (D). FPLC Chromatogram (E). SDS-PAGE gel (F). Immunoblot of various nanodiscs

Binding of the peptide to MHCII is interrupted due to the presence of transmembrane region or due to the presence of lipids: For almost four decades, soluble-DR1 has been used to measure the binding affinity of antigenic peptides. Here we present a tool where full-length DR1 is embedded in the synthetic membrane and available in the soluble form. We wanted to test if this assembly can be used for binding studies and if there is any difference between the binding affinity of soluble DR1 and full-length DR1 embedded in the nanodisc. For this purpose, we used an ELISA-based direct binding assay to measure binding affinity of biotinylated $HA_{306-318}$ (bioHA). Different concentrations of bioHA were incubated with soluble DR1, full-length DR1, and nanodisc to measure the binding affinity of bioHA. As shown in Figure 4.3A, 4.3B, 4.3C, and 4.3D, the fluorescent count for full-length DR1 and nanodisc with DR1 were less than the soluble DR1.

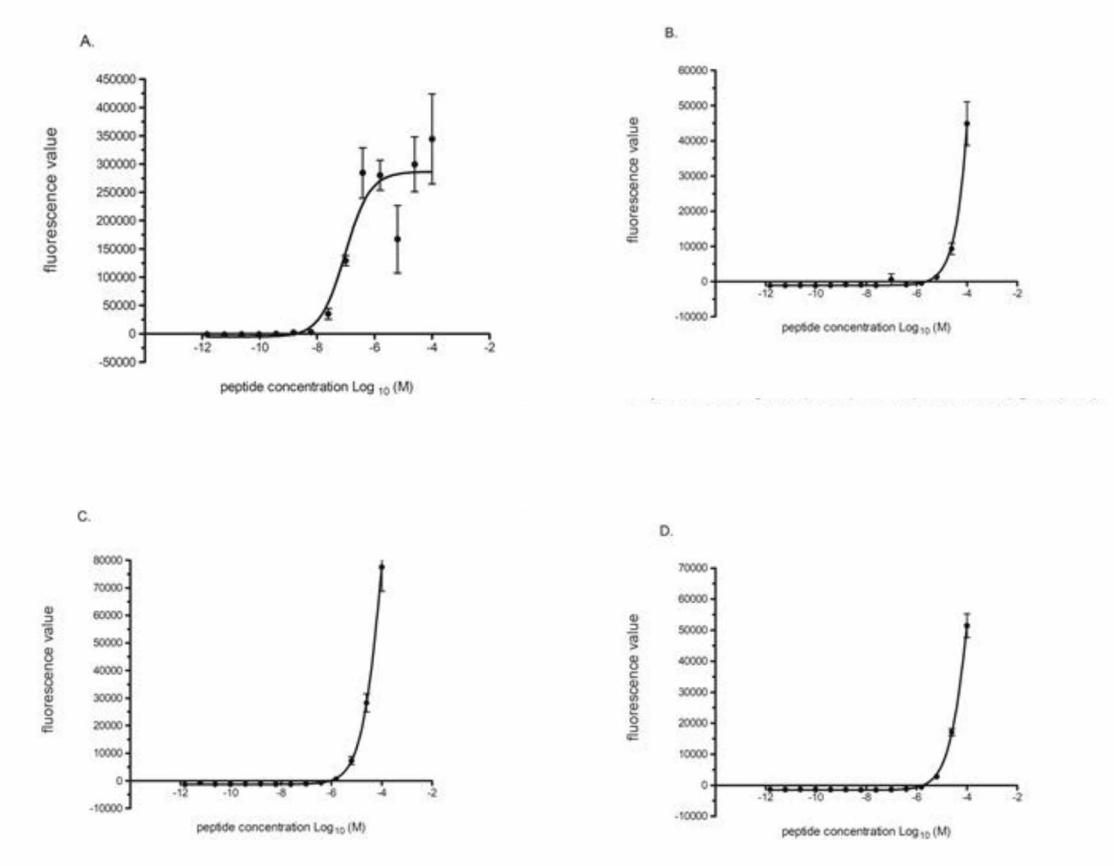


Figure 4.3: Direct binding assay to measure the binding affinity of biotinylated HA (bioHA): (A). Soluble-DR1 (B). mDR1 (C). Empty-nanodisc (D). Fluid-nanodisc.

The difference in binding based on the lipid composition: Our ELISA-based direct binding assay showed low fluorescence count for the nanodisc, and we believe that it is due to inefficient capturing of nanodisc via antibody on the surface of the binding plate. To overcome this issue, we used native-gel electrophoresis to visualize the binding. To test if the peptide binds to full-length DR1 embedded in the nanodisc, we used fluorescein-labeled $HA_{306-318}$ peptide. We incubated the proteins with a 20-fold excess of peptide overnight at 37°C and loaded the protein on the native-gel. We observed that empty nanodisc also showed a fluorescent band at lower intensity. The presence of fluorescent band with empty nanodisc could be due to non-specific binding of the peptide to lipids. We measured the band intensity using ImageJ and subtracted the empty nanodisc intensity from the samples band. We observed that simple nanodisc (1000: 10:

1) and fluid nanodisc showed binding which is comparable to soluble DR1 (Figure 4.4A and 4.4B). Rigid nanodisc showed no binding with HA peptide. The non-specific binding observed with empty nanodisc can be explained with the buffer composition (Bockmann, Hac, Heimburg, & Grubmuller, 2003). We also observed a significant difference between in-band intensity when we used two different buffers but with the same pH (pH 7.4). We suggest that the difference in binding is due to variation in ionic composition of the buffer.

Taken together these results suggest that the full-length DR1 embedded in the nanodisc can be used as a tool to study the binding of antigenic peptides. Our preliminary results indicate that there is a difference in the binding based on the type of nanodisc and the buffers used during the binding reaction. Our initial binding assays also suggest non-specific binding with empty nanodisc which shows there is a need of different binding assays to study the nanodisc. In the future, nanodisc embedded full-length DR1 showed the potential to replace the soluble DR1 protein.

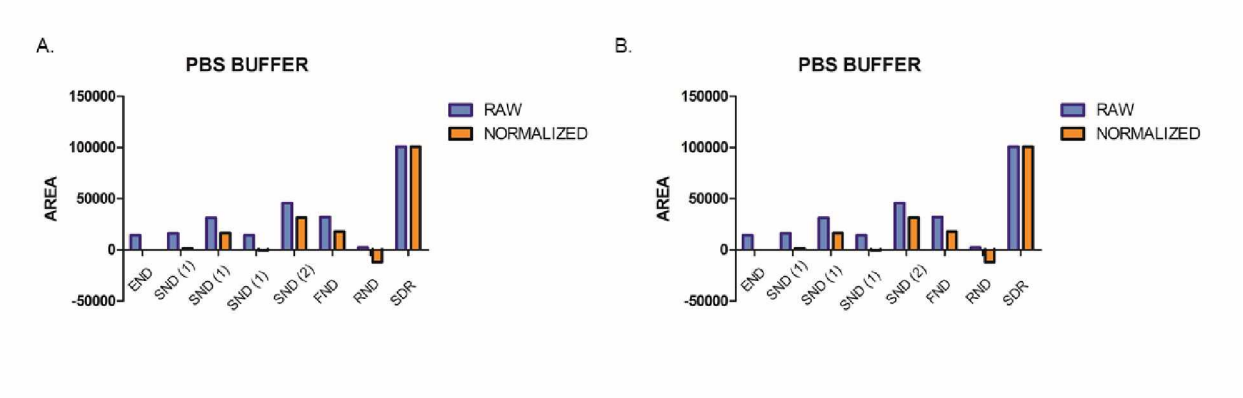


Figure 4.4: Difference in binding due to lipid composition: (A). 5 µM protein incubated with 100 µM of fluorescein-labeled HA peptide in PBS buffer (B). 5 µM protein incubated with 100 µM of fluorescein-labeled HA peptide 20 mM Tris-Cl 100 mM NaCl

Discussion

MHCII is extensively studied in its soluble form by using only the extracellular region. Studies have shown that the transmembrane and cytoplasmic domains of the protein play an essential role in the expression of MHCII as well as in B cell signaling (Harton, Jin, Hahn, & Drake, 2016; Harton & Bishop, 1993). However, the role of transmembrane and cytoplasmic region on the binding capacity of the MHCII is still unknown. There is currently no *in-vitro* system available, which can mimic the MHCII embedded in the cell membrane. The assembly of MHCII in the membrane also plays a vital role in the CD4+ T cell stimulation. Studies have suggested that MHCII can assemble in superdimers capable of triggering CD4+ T cell activation for low-affinity antigens (Schafer & Pierce, 1994). These superdimers show thermal and pH stability like heterodimer, as SDS is denatured both at a pH below 5 and temperature above 50ºC (Schafer et al., 1998). The goal of the work reported here was to investigate how a full-length MHCII would incorporate in membranes and if the membrane lipid composition would affect the assembly of MHCII. To achieve this, we used a synthetic model membrane known as nanodiscs widely adopted to study membrane proteins.

First, we prepared a simple nanodisc containing only phosphatidylcholine (POPC) in the assembly mix. We used two different ratios of assembly mix in the preparation. We observed that based on the ratio of lipid, the assembly of DR1 within the nanodisc changed. For 100: 5: 1 we found that the first fractions have nanodisc which has just superdimers while the later one shows dimers and superdimers both for 1000: 10: 1 assembly mix we observed superdimers and dimers assembled in nanodisc in first fractions and following fractions have dimers alone in the nanodisc. These results suggest that the percentage of lipids in the composition can alter the assembly of DR1 in the membrane.

Next, we tested the membrane fluidity and phase impact on the assembly of DR1. For this, we prepared fluid-disordered and rigid-ordered nanodiscs by using cholesterol and sphingomyelin along with POPC in the assembly mix. For fluid-disordered nanodiscs, we observed one major peak in FPLC, and when the fractions were analyzed on native and SDS gel we found that the nanodisc was composed of dimers and superdimers altogether. While for rigid-ordered nanodisc on FPLC we observed one major peak which bifurcated at the top of the peak. The fractions on SDS gel show that nanodisc in the first-half peak contains superdimers with dimers while the later fractions have mostly dimer of DR1. These observations suggest that the fluidity and phase of the membrane can determine how DR1 assembles in the membrane.

Studies have shown that cholesterol-rich microdomains are associated with MHCII-peptide complexes at the plasma membrane of dendritic cells, and the abolishment of these cholesterol rich-domains reduces the activation of CD4+ T cells (Bosch et al., 2013). The clustering of MHCII plays an essential role in enhancing TCR signaling. In our nanodisc, we observed that DR1 dimers and superdimers were both present. How they arranged in the nanodisc is not known. But based on how clustering of the MHCII complexes is essential for T-cell activation we think DR1 might arrange as a cluster in the nanodiscs.

Having shown the assembly of full-length DR1 in the nanodisc, full exploitation of this tool requires the assessment of its peptide binding capacity. For preliminary binding studies we used ELISA based assay and Fam-labeled peptides. In both cases, we observed non-specific interactions that masked the actual binding. Through an ELISA-based assay, we noted that full-length DR1, empty-nanodisc, and fluid-nanodisc showed comparable and lower fluorescence count as compared to soluble DR1. We believe the reason for the low count is maybe due to the inefficient capturing of the peptide-MHCII complex via L243 on the binding plates. For full-

length DR1 we think this is possibly due to free moving transmembrane region, while for DR1 embedded in nanodisc this is likely due to the bulky size of the nanodisc. Future approaches for testing the biochemical behavior of membrane embedded MHCII will include SPR and ITC. In our opinion, this system has the potential of mimicking the MHCII behavior as found on the membrane of antigen-presenting cells, and possibly revolutionizing the research field of antigen presentation and T cell activation by providing a substitute technology to cell culture-based assays.

References

Amin, D. N., & Hazelbauer, G. L. (2012). Influence of membrane lipid composition on a transmembrane bacterial chemoreceptor. *J Biol Chem, 287*(50), 41697-41705. doi:10.1074/jbc.M112.415588

Bockmann, R. A., Hac, A., Heimburg, T., & Grubmuller, H. (2003). Effect of sodium chloride on a lipid bilayer. *Biophys J, 85*(3), 1647-1655. doi:10.1016/S0006-3495(03)74594-9

Bosch, B., Heipertz, E. L., Drake, J. R., & Roche, P. A. (2013). Major histocompatibility complex (MHC) class II-peptide complexes arrive at the plasma membrane in cholesterol-rich microclusters. *J Biol Chem, 288*(19), 13236-13242. doi:10.1074/jbc.M112.442640

Brown, J. H., Jardetzky, T. S., Gorga, J. C., Stern, L. J., Urban, R. G., Strominger, J. L., & Wiley, D. C. (2015). Pillars article: three-dimensional structure of the human class II histocompatibility antigen HLA-DR1. Nature. 1993. 364: 33-39. *J Immunol, 194*(1), 5-11.

Denisov, I. G., & Sligar, S. G. (2017). Nanodiscs in Membrane Biochemistry and Biophysics. *Chem Rev, 117*(6), 4669-4713. doi:10.1021/acs.chemrev.6b00690

Gorga, J. C., Horejsi, V., Johnson, D. R., Raghupathy, R., & Strominger, J. L. (1987). Purification and characterization of class II histocompatibility antigens from a homozygous human B cell line. *J Biol Chem, 262*(33), 16087-16094.

Guillet, J. G., Lai, M. Z., Briner, T. J., Smith, J. A., & Gefter, M. L. (1986). Interaction of peptide antigens and class II major histocompatibility complex antigens. *Nature, 324*(6094), 260-262. doi:10.1038/324260a0

Harton, J., Jin, L., Hahn, A., & Drake, J. (2016). Immunological Functions of the Membrane Proximal Region of MHC Class II Molecules. *F1000Res, 5*. doi:10.12688/f1000research.7610.1

Harton, J. A., & Bishop, G. A. (1993). Length and sequence requirements of the cytoplasmic domain of the A beta molecule for class II-mediated B cell signaling. *J Immunol, 151*(10), 5282-5289.

Hitzel, C., Gruneberg, U., van Ham, M., Trowsdale, J., & Koch, N. (1999). Sodium dodecyl sulfate-resistant HLA-DR "superdimer" bands are in some cases class II heterodimers bound to antibody. *J Immunol, 162*(8), 4671-4676.

Koukalova, A., Amaro, M., Aydogan, G., Grobner, G., Williamson, P. T. F., Mikhalyov, I., . . . Sachl, R. (2017). Lipid Driven Nanodomains in Giant Lipid Vesicles are Fluid and Disordered. *Sci Rep, 7*(1), 5460. doi:10.1038/s41598-017-05539-y

Schafer, P. H., Malapati, S., Hanfelt, K. K., & Pierce, S. K. (1998). The assembly and stability of MHC class II-(alpha beta)2 superdimers. *J Immunol, 161*(5), 2307-2316.

Schafer, P. H., & Pierce, S. K. (1994). Evidence for dimers of MHC class II molecules in B lymphocytes and their role in low affinity T cell responses. *Immunity, 1*(8), 699-707.

Stern, L. J., & Wiley, D. C. (1992). The human class II MHC protein HLA-DR1 assembles as empty alpha beta heterodimers in the absence of antigenic peptide. *Cell, 68*(3), 465-477. doi:10.1016/0092-8674(92)90184-e

Thomas, D. W., Hsieh, K. H., Schauster, J. L., Mudd, M. S., & Wilner, G. D. (1980). Nature of T lymphocyte recognition of macrophage-associated antigens. V. Contribution of individual peptide residues of human fibrinopeptide B to T lymphocyte responses. *J Exp Med, 152*(3), 620-632. doi:10.1084/jem.152.3.620

Vogt, A. B., Spindeldreher, S., & Kropshofer, H. (2002). Clustering of MHC-peptide complexes prior to their engagement in the immunological synapse: lipid raft and tetraspan microdomains. *Immunol Rev, 189*, 136-151.

Wade, W. F., Ward, E. D., Rosloniec, E. F., Barisas, B. G., & Freed, J. H. (1994). Truncation of the A alpha chain of MHC class II molecules results in inefficient antigen presentation to antigen-specific T cells. *Int Immunol, 6*(10), 1457-1465. doi:10.1093/intimm/6.10.1457

Yin, L., Maben, Z. J., Becerra, A., & Stern, L. J. (2015). Evaluating the Role of HLA-DM in MHC Class II-Peptide Association Reactions. *J Immunol, 195*(2), 706-716. doi:10.4049/jimmunol.1403190

Chapter 5: Conclusion

The class II MHC processing and presentation pathway is a complex process, which involves multiple factors influencing the selection of the antigens that are ultimately displayed to CD4+ T cells (Neefjes, Jongsma, Paul, & Bakke, 2011). Exogenous proteins are presented via MHCII. The recognition of an exogenous protein is started with the uptake by antigen-presenting cells (APCs) through endocytosis. The protein moves through various lysosomal compartments with different pH and cathepsin enzymes, which generate a pool of peptides competing for binding to MHCII in the MHC class II compartment (MIIC) (Geuze, 1998). In MIIC, a newly synthesized MHCII molecule, or one recycle from the membrane, along with non-classical MHCII molecule HLA-DM, selects the peptides for presentation to CD4+ T cells (Nanda & Bikoff, 2005). There is more than one peptide that can bind to MHCII and presented to CD4+ T cells; among such peptides there are only a few that can generate significant CD4+ T cell response. Peptides that are the focus of T cell response are known as immunodominant. Immunodominant epitope discovery is the focus of substantial efforts due to the impact of this information on vaccine engineering and the understanding of the pathology of emerging pathogens or autoimmunity. Two models can explain how immunodominant epitopes selected via MHCII. The first model is a kinetic stability model, which postulates that the immunodominant epitope has intrinsic kinetic stability. Studies have shown that immunodominant peptides have kinetic stability > 100 hours, and weaker epitopes have kinetic stability < 6 hours (Sant et al., 2005). These epitopes are resistant to DM activity and weaker epitopes are also susceptible to DM activity. The second model is known as the "epitope accessibility" model, which proposes that immunodominant epitopes are easily accessible to MHCII binding sites due to their position in the protein structure, and they are also resistant to cathepsin activity (Kim et al., 2014; Kim & Sadegh-

Nasseri, 2015). Thus, two mechanisms can potentially determine the accessibility of the epitope to MHCII: "cut first and bind later" whereby a protein is first trimmed by cathepsin and then binds to MHCII, or "bind first and cut later" whereby a protein binds first to the MHCII and trimmed later at the flaming sides, whereas the MHCII itself shields intra-peptide cleavage sisters.

Literature shows that immunodominant epitope selection is not determined by one single factor, but it is dependent on multiple intertwined components. Here, we have investigated a few such mechanisms, namely the role of MHCII molecule biochemical behavior, the activity of DM, and the impact of MHCII assembly in the membrane on the selection of epitopes.

In most of our studies, we have used a set of peptides derived from H1N1 influenza hemagglutinin, for which a DR1-restricted immunodominance hierarchy was already established (Richards et al., 2007). Our first aim was to see if DM-associated IC_{50} can be considered a viable proxy of immunodominance. We measured the binding affinity of the various peptides in the absence and presence of DM. We observed a direct correlation between the presence of DM and the reduction in affinity to MHCII of those epitopes for which a weaker CD4+ T cell response recorded. We also tested the epitope prediction accuracy of DM-associated IC50 by plotting the Receiver operating characteristics (ROC) curve. When compared to different online prediction tools, the experimental measurement of DM-associated affinity showed better accuracy in predicting epitope by eliminating false-positive epitopes. The correlation between DM-associated IC_{50} and immunodominant epitope selection is novel and of certain impact in the field. Indeed, past and current studies focus on DM-associated half-lives instead of determining the role of DM during peptide binding to MHCII. We used competition binding assay to study the affinity of each peptide in a reaction that somewhat mimics the MIIC system, where peptides

of various affinity for the MHCII compete for binding in the presence of DM. We conclude that the DM-associated IC50 is a better correlate of immunodominance, and inclusion of this factor in epitope prediction tools can improve their accuracy.

Our second aim was to test the impact of pH and the length of the peptides in the MHCII presentation pathway. We started by looking into the role of pH in the presentation when we observed that immunodominant epitopes of hemagglutinin protein showed DM-susceptibility. Past studies have shown that immunodominant epitope kinetics is DM-resistant (Sant et al., 2005). To reconcile the apparent paradox of experimental immunodominant peptides showing DM susceptibility, we started looking into the effect of pH and the length of the peptides. We measured the off rate of immunodominant peptides at pH 7.4 and observed greater stability in comparison to pH 5.4. We also tested the SDS-stability of these peptides at pH 5.4 and 7.4 and observed that the pH drop abolishes SDS stability of immunodominant peptides. Next, we used 15-mer derivatives of selected immunodominant peptides and measured their kinetics at pH 5.4 in the absence and presence of DM. These derivatives showed better stability in the absence and presence of DM as compared to their original longer peptides. These findings indicate that pH and length of the peptides can impact their kinetic stability and DM activity therein. We propose that these immunodominant peptides could be presented through an alternate pathway in which peptides bind to empty MHCII at the surface of the APCs and are recycled in shallow compartments. Past studies have suggested such routes of processing and presentation for hemagglutinin, myelin basic protein and HEK protein (Lindner & Unanue, 1996; Pinet, Vergelli, Martin, Bakke, & Long, 1995; Pinet & Long, 1998; Vergelli et al., 1997).

The MHCII research of the past two decades was performed by using soluble MHCII, where the transmembrane region of the protein is not expressed (Frayser, Sato, Xu, & Stern, 1999; Gorga,

Horejsi, Johnson, Raghupathy, & Strominger, 1987). Soluble MHCII is a natural choice to study the biochemistry and biophysics of the protein interacting with peptides. Studies have shown that the MHCII organization in the plasma membrane is an essential factor to trigger T-cell activation (Fooksman, 2014). It is not known if MHCII assembly is directly related to lipid composition. Thus, our third objective was to study the impact of lipid composition in the MHCII assembly. To this aim, we used a synthetic membrane model known as nanodisc in which the lipid composition is easily manipulated. We prepared different types of nanodisc with different lipids composition and incorporated full-length MHCII in it. We showed that, based on lipid composition, various oligomeric forms of MHCII assembled in the nanodiscs. In the end, we wanted to use the nanodisc embedded full-length MHCII as a tool to study the binding affinity of antigenic peptides. We have shown that this tool is not fully developed for binding study adopting solid-phase immunoassay steps. We believe that this difficulty is due to the orientation of the MHCII in the nanodisc which makes it inaccessible by the capture antibody. We also tested peptide-binding qualitatively using the native-gel electrophoresis, and we did observe a difference in binding capacity between different nanodiscs. This observation is quite interesting because it suggests that the orientation of the MHCII in the nanodisc dictates how the antigen can interact and ultimately bind. This characteristic could be reflection of what happens in the cells. MHCII is found on the surface of the cells, MIIC, and different endosomal compartments. It might be possible that the orientation of the MHCII dictates that it will bind to an antigen or not in that specific compartment.

Our finding addresses essential factors that can dictate the selection of immunodominant epitopes. Here, we showed that the inclusion of DM activity could be a crucial addition to any strategy aiming at predicting immunodominant epitopes. pH and cleavage susceptibility,

however, may impact the generation of a specific epitope and need to be considered as well. Finally, we show the organization of MHCII in a synthetic membrane model and the implications of these findings in the biophysical characterization of MHCII. This work has therefore led the foundation for future research aimed at deepening our understanding and predicting the selection of immunodominant epitopes by MHCII.

References

Fooksman, D. R. (2014). Organizing MHC Class II Presentation. *Front Immunol, 5*, 158. doi:10.3389/fimmu.2014.00158

Frayser, M., Sato, A. K., Xu, L., & Stern, L. J. (1999). Empty and peptide-loaded class II major histocompatibility complex proteins produced by expression in Escherichia coli and folding in vitro. *Protein Expr Purif, 15*(1), 105-114. doi:10.1006/prep.1998.0987

Geuze, H. J. (1998). The role of endosomes and lysosomes in MHC class II functioning. *Immunol Today, 19*(6), 282-287.

Gorga, J. C., Horejsi, V., Johnson, D. R., Raghupathy, R., & Strominger, J. L. (1987). Purification and characterization of class II histocompatibility antigens from a homozygous human B cell line. *J Biol Chem, 262*(33), 16087-16094.

Kim, A., Hartman, I. Z., Poore, B., Boronina, T., Cole, R. N., Song, N., . . . Sadegh-Nasseri, S. (2014). Divergent paths for the selection of immunodominant epitopes from distinct antigenic sources. *Nat Commun, 5*, 5369. doi:10.1038/ncomms6369

Kim, A., & Sadegh-Nasseri, S. (2015). Determinants of immunodominance for CD4 T cells. *Curr Opin Immunol, 34*, 9-15. doi:10.1016/j.coi.2014.12.005

Lindner, R., & Unanue, E. R. (1996). Distinct antigen MHC class II complexes generated by separate processing pathways. *EMBO J, 15*(24), 6910-6920.

Nanda, N. K., & Bikoff, E. K. (2005). DM peptide-editing function leads to immunodominance in CD4 T cell responses in vivo. *J Immunol, 175*(10), 6473-6480. doi:10.4049/jimmunol.175.10.6473

Neefjes, J., Jongsma, M. L., Paul, P., & Bakke, O. (2011). Towards a systems understanding of MHC class I and MHC class II antigen presentation. *Nat Rev Immunol, 11*(12), 823-836. doi:10.1038/nri3084

Pinet, V., Vergelli, M., Martin, R., Bakke, O., & Long, E. O. (1995). Antigen presentation mediated by recycling of surface HLA-DR molecules. *Nature, 375*(6532), 603-606. doi:10.1038/375603a0

Pinet, V. M., & Long, E. O. (1998). Peptide loading onto recycling HLA-DR molecules occurs in early endosomes. *Eur J Immunol, 28*(3), 799-804. doi:10.1002/(SICI)1521-4141(199803)28:03<799::AID-IMMU799>3.0.CO;2-5

Richards, K. A., Chaves, F. A., Krafcik, F. R., Topham, D. J., Lazarski, C. A., & Sant, A. J. (2007). Direct ex vivo analyses of HLA-DR1 transgenic mice reveal an exceptionally broad pattern of immunodominance in the primary HLA-DR1-restricted CD4 T-cell response to influenza virus hemagglutinin. *J Virol, 81*(14), 7608-7619. doi:10.1128/JVI.02834-06

Sant, A. J., Chaves, F. A., Jenks, S. A., Richards, K. A., Menges, P., Weaver, J. M., & Lazarski, C. A. (2005). The relationship between immunodominance, DM editing, and the kinetic stability of MHC class II:peptide complexes. *Immunol Rev, 207*, 261-278. doi:10.1111/j.0105-2896.2005.00307.x

Vergelli, M., Pinet, V., Vogt, A. B., Kalbus, M., Malnati, M., Riccio, P., . . . Martin, R. (1997). HLA-DR-restricted presentation of purified myelin basic protein is independent of intracellular processing. *Eur J Immunol, 27*(4), 941-951. doi:10.1002/eji.1830270421

Appendix: Fluorescence anisotropy – based analysis of the conformational modifications in the peptide-MHCII complex structure[4]

Introduction

Class II major histocompatibility complex (MHCII) are glycoproteins expressed on the surface of antigen-presenting cells (APCs) and display antigenic peptides to CD4+ T cells (Hume, 1985). A MHCII molecule is synthesized in the endoplasmic reticulum (ER) usually as a nonameric complex, which comprises of a trimer of trimers, formed by an α and a β-chain and an invariant chain. This complex is transported via Golgi to MIIC-compartment where the invariant chain is shortened to a peptide named CLIP (Class II-associated invariant chain peptide), then removed by a non-classical MHCII molecule HLA-DM (DM) (Ferrante, 2013). Upon CLIP release, the pool of antigenic peptides that are generated via cathepsin cleavage can bind to MHCII molecules to form MHCII-peptide complexes (pMHCII) (Neefjes, Jongsma, Paul, & Bakke, 2011). The selected pMHCII is transported to the cell surface where they are presented to CD4+ T cells, which generate the immune response against the specific peptide. Among the peptide repertoire generated, only few of the peptides can produce a strong CD4+ T cell response, and these peptides are known as immunodominant peptides (Kim et al., 2014; Kim & Sadegh-Nasseri, 2015). Many areas of inquiry would benefit from an understanding of the factors determining the presentation of immunodominant peptides, such as vaccinology or immunopathology. Several aspects of the presentation process are usually considered to explain the selection of immunodominant peptides, the most relevant being pMHCII binding affinity and

[4] Osan J.K, Steele H., Wang Z., Ross Alexander, Ferrante A., Kuhn T.B. Fluorescence anisotropy-based analysis of the conformational modifications in the peptide-MHCII structure.
This study is a preliminary study for my project. It is not completed and will be carried out in my lab in the future.

kinetics, structural characteristics of the complex and role of DM therein (Dai, Steede, & Landry, 2001).

The crystal structure of several pMHCII's has shown that the peptide binds to MHCII by hydrophobic, ionic interaction and H-bond. There are various pockets in the MHCII binding site, which play an essential role in the interaction with the peptide. The major pockets are P1, P4, P6/P7, and P9. The numbering denotes the side chain of the peptides, which interacts with the binding site of MHCII. P1 is the deepest pocket at the N-terminal side of the pMHCII that accommodates the hydrophobic and sometimes aromatic amino acid side chain depending on the allele (Stern et al., 1994). Studies have shown that the region of the MHCII binding site including and surrounding the P1 pocket is the one undergoing the most dramatic changes upon complexation, in particular, the α-subunit 3_{10} helical region and the adjacent extended strand; in addition, the β2 Ig-like domain, and the pronounced kink in the β-subunit helical region (β62-71) appear to be structurally labile as well (Pos et al., 2012; Rupp et al., 2011; Schulze, Anders, Sethi, & Call, 2013).

Another important factor in the selection of immunodominant epitopes is the activity of DM. The role of DM is not only limited to the removal of CLIP, but it also plays an essential role in the selection of kinetically stable pMHCII. Studies have shown that DM accelerates the exchange of the peptides and generates the pMHCII, which are kinetically stable. Recent reports have shown that two main factors play an essential role in determining DM susceptibility. One factor is occupancy of the P1 pocket, and the second factor is the overall conformational modification in the MHCII as it interacts with the peptide (Pos et al., 2012; Yin & Stern, 2013).

In this study, we aimed to analyze changes in MHCII structural conformation upon the binding of the peptide in the absence and presence of DM. For this purpose, we generated two separate

single-cysteine mutations in the α-subunit 3_{10} helical region and one in the β-subunit helical region. We used fluorescence anisotropy to measure the anisotropy of the pMHCII complex in the absence and presence of the DM. We used $HA_{306-318}$ from H3 subtype of influenza and HA440-455 and HA323-340 from H1N1 influenza to form pMHCII complexes. Our preliminary data showed that the anisotropy change is dependent on the function of the bound peptide. We also observed that for peptide with stronger CD4+ T cell response there was less change in anisotropy when incubated with DM as compare to peptide with lower CD4+ T cell response. Initially we observed that there was inconsistency between three repeats. Upon performing three more repeats with our samples, we found similar inconsistency. This result suggested that it is possible that the fluorescein dye is binding non-specifically to our protein. Further mass spectrometry analysis of our labeled protein confirmed that the dye is binding non-specifically to our protein. This makes it very difficult to analyze our data as it is difficult to tell how much anisotropy is contributed by non-specific labeling v/s specific labeling. Our preliminary results suggested that anisotropy studies can reveal the mechanism behind DM activity. Future work should focus on labeling of the protein and repeating anisotropy experiments to provide valuable information needed to understand how DM screen different peptides for antigen presentation.

Materials and methods:

DR1 mutant proteins for fluorescence anisotropy: For fluorescent labeling of DR1, cysteine mutations were introduced at specific positions of the α and the β-chains thus producing three cysteine mutants: α45-DR1, α52-DR1, and β65-DR1. Here the number denotes the position of the amino acid which is substituted to cysteine, and the symbol indicates the chain of DR1 protein. Genscript Inc prepared plasmid for these mutants. S2 Drosophila melanogaster cell line was transfected using calcium phosphate transfection kit, and hygromycin was used as a selection agent. The cells were induced using copper sulfate at a concentration of 500 μM. Soluble protein were purified by using protein A sepharose beads, which cross-linked with the L243 antibody. The mutant protein was confirmed by mass spectrometry analysis performed at University of Montana, Bozeman. Wild-type DR1 (WT DR1) was expressed in our lab in S2 Drosophila melanogaster cell line and purified by using L243 affinity column chromatography (Stern et al., 1994).

Peptides synthesis: Four peptides were selected for the fluorescence anisotropy measurement. Peptide $HA_{306-318}$ (GPKYVKQNTLKLAT) from influenza A virus H3 subtype was selected as a benchmark peptide due to past structural studies with this peptide. We also selected two peptides derived from the hemagglutinin protein of the H1N1 influenza virus for which CD4+ T cell response was already known. We used HA440-455 (Peptide 440) and HA323-340 (Peptide 47) with two different types of CD4+ T cell responses, where HA440-455 showed strong CD4+ T cell response and HA323-340 low CD4+ T cell response in the ELISPOT assay performed by Sant et al (Richards et al., 2007). These peptides were synthesized by ABI scientific using Fmoc chemistry and fully automated multiple peptide synthesizer.

Fluorescence anisotropy: DR1 (WT and mutants) were labeled with fluorescein-5-maleimide at room temperature for three hours. Free dye was removed by washing with 20 mM PBS buffer ten times using Centricon-30 spin filters. 5 µM labeled proteins incubated with a 10-fold excess of unlabeled peptide overnight at 37C to generate the peptide-MHCII (pMHCII) complexes. The complexes were analyzed on the fluorimeter available in Dr. Alexander Ross lab at the University of Montana, Missoula, to measure the anisotropy of each complex. The overview of the fluorimeter shown in figure 1.

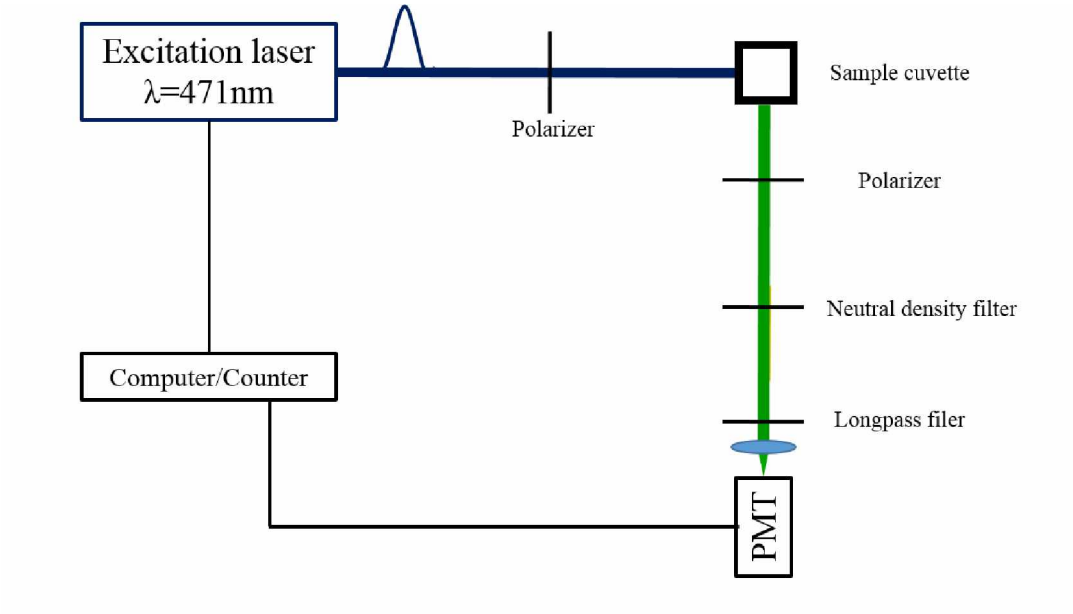


Figure 1. Scheme of Fluorimeter setup for the dynamic anisotropy measurement

Mass spectrometry analysis: Specific and non-specific labeling were analyzed by mass spectrometry. Unlabeled WT and mutant DR1 along with fluorescent labeled DR1 were sent to Colorado State University. Samples were analyzed on MALDI-TOF to calculate the mass to charge ratio for each sample.

Results:

Anisotropy measurement for pMHCII complexes: In this study, we wanted to understand which residues of the MHCII molecules are involved during the binding of the peptides and if there is a change in the molecular environment when DM is around. Thus, we designed three cysteine DR1 mutants for this purpose. Each mutant has a single cysteine mutation at a specific residue which was neighbor to the α-subunit 3_{10} helical region and the β-subunit helical region (β62-71). Previous studies have shown that these regions upon binding to the peptides showed changes in confirmation (Painter, Cruz, Lopez, Stern, & Zavala-Ruiz, 2008; Pos et al., 2012). We created mutations at positions 45 and 52 of α-chain and position 65 of β-chain. After mutation was created, the samples were sent to University of Montana, Bozeman for mass spectrometry analysis to confirm cysteine mutation. Three peptides $HA_{306-318}$ (H3 influenza subtype), HA440-455 and HA323-340 (H1N1 influenza) were selected for our anisotropy experiment. The fluorescence anisotropy measurement was performed at University of Montana, Missoula in Dr. Alexander Ross lab. First, we labeled WT and mutant DR1 with fluorescein-5-maleimide. WT DR1 was used as a control since it has a free cysteine which can be labeled with fluorescein. After labeling with fluorescein dye, 5 μM labeled DR1 proteins were incubated with 10-fold excess of peptide overnight. The pMHCII complexes were analyzed on the fluorometer. The anisotropy was measured for complexes in the absence and presence of DM. We calculated the mean of the steady-state anisotropy (Table 1). When we calculated the change in the anisotropy after addition of DM, we observed that for peptide 47 the anisotropy decrease was more as compared to other peptides (as shown in Figure 1). These show that the MHCII-peptide 47 complex might be more flexible in the presence of DM as compared to the other peptides.

Table 1: Steady-state anisotropy measurement for pMHCII complexes in the absence and presence of DM: Anisotropy was measured for complexes made with α45 DR1, α52 DR1, and β65 DR1. For each constant value mean of repeats were calculated.

Protein name	Peptide used for complex	DM added	Time Point (hour)	Mean of Steady-state Measurement	SD of Steady-state measurement
DR1	None	No		0.073	0.005
α-45 DR1	None	No		0.172	0.030
α-52 DR1	None	No		0.180	0.027
β-65 DR1	None	No		0.161	0.041
DR1	$HA_{306-318}$	No		0.166	0.037
	$HA_{306-318}$	Yes	0	0.175	0.043
	$HA_{306-318}$	Yes	24	0.163	0.059
α-45 DR1	$HA_{306-318}$	No		0.194	0.020
	$HA_{306-318}$	Yes	0	0.205	0.030
	$HA_{306-318}$	Yes	24	0.185	0.023
α-52 DR1	$HA_{306-318}$	No		0.154	0.044
	$HA_{306-318}$	Yes	0	0.168	0.040
	$HA_{306-318}$	Yes	24	0.190	0.015
β-65 DR1	$HA_{306-318}$	No		0.203	0.012
	$HA_{306-318}$	Yes	0	0.190	0.016
	$HA_{306-318}$	Yes	24	0.183	0.011
α-45 DR1	440	No		0.183	0.12
	440	Yes	0	0.198	0.012
	440	Yes	24	0.183	0.022
α-52 DR1	440	No		0.191	0.015
	440	Yes	0	0.191	0.005
	440	Yes	24	0.172	0.041
β-65 DR1	440	No		0.207	0.016
	440	Yes	0	0.192	0.016
	440	Yes	24	0.184	0.016
α-45 DR1	47	No		0.201	0.021
	47	Yes	0	0.200	0.007
	47	Yes	24	0.166	0.039
α-52 DR1	47	No		0.204	0.011
	47	Yes	0	0.196	0.009
	47	Yes	24	0.152	0.030
β-65 DR1	47	No		0.201	0.012
	47	Yes	0	0.175	0.050
	47	Yes	24	0.184	0.011

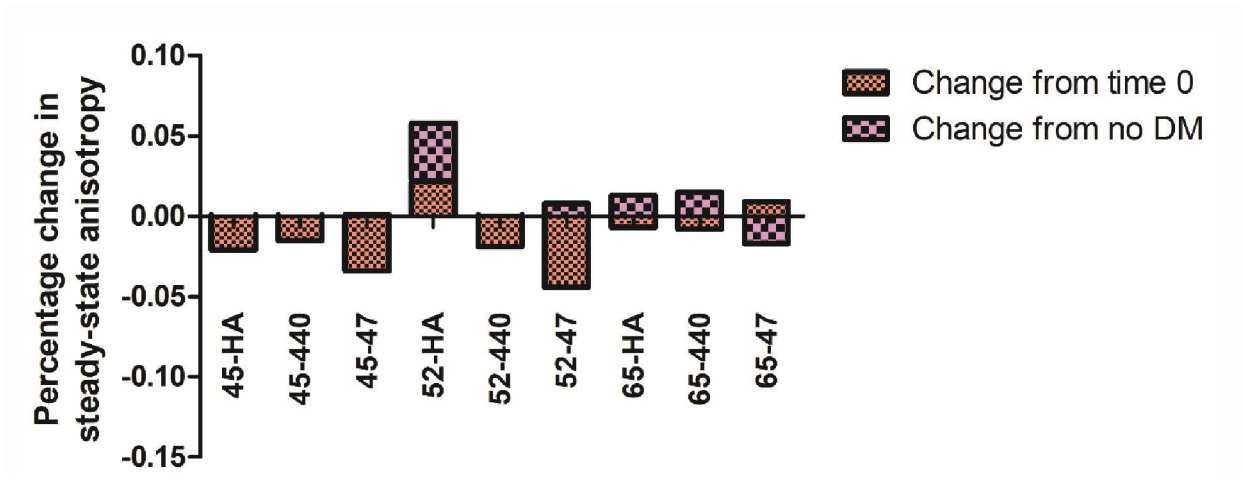


Figure 2: Change in the steady-state anisotropy after addition of DM

Mass spectrometry analysis to quantify the amount of fluorescein bound to MHCII protein specifically or non-specifically: Our preliminary analysis of fluorescence anisotropy measurement showed a discrepancy between the repeats. Upon increasing our repeat runs from three to six we still observed the same problem. We thought that the difference between repeats might be due to unspecific binding of the fluorescein dye to our protein. To verify this, we sent the unlabeled proteins along with labeled proteins to the mass spectrometry facility of Colorado State University. The proteins were analyzed using MALDI-TOF. The results showed that there was no difference in the mass to charge ratio between unlabeled and labeled protein (data not shown). Based on this result we concluded that there is a possibility that there is no covalently attached dye portion of the protein flew during the measurement because the amount of labeled protein is too low as compare to unlabeled protein. These results support our theory that the inconsistency between the repeats might be due to the unspecific labeling of the fluorescein dye to the MHCII protein. Due to the amount of unknown unspecific label dye it is difficult to know how much anisotropy contribution is from the unspecific labeling and how much is from labeled protein.

Conclusion and future directions:

MHCII molecule presents antigenic peptides to CD4+ T cells to generate the immune response against specific pathogens. It is shown that the non-classical MHCII molecule DM plays an essential role in generating kinetically stable pMHCII complexes which are presented to CD4+ T cells. Here, we wanted to understand the structural component of MHCII which is involved with DM to select the peptides which are presented to CD4+ T cells. To understand how the molecular environment of pMHCII complex changed in the presence of DM, we generated single cysteine DR1 mutant. The cysteine mutation is in that region of α and β-chain of MHCII, which are known to change during the binding of a peptide to MHCII molecule. We labeled the cysteine molecule with fluorescein-5-maleimide and incubated the labeled protein with different peptides for which the CD4+ T cell response is already known. We measured the steady-state anisotropy for each pMHCII complexes in the absence and presence of DM by using fluorimeter. Our preliminary data showed that the change due to the presence of DM was more imminent in the peptide 47 which has a lower CD4+ T cell response as compared to other peptides. Increased experimental replicates lead to inconsistency between the repeats. Mass spectrometry analysis showed that there is no difference in the mass between unlabeled and labeled proteins. This can be due to the lack of efficient labelling of dye to our protein and there being more amount of dye sticking non-specifically to our protein. These results raise a reasonable doubt on our collected data as it makes it difficult to separate the anisotropy associated with specific and non-specific labeling. Since our preliminary studies indicate that there might be a difference in anisotropy based on the flexibility of the complex, and DM susceptibility may be dependent on the structural change of pMHCII complexes, we believe a detail structural studies of pMHCII

complexes will provide useful information to understand how DM helps in the selection of the immunodominant epitopes.

References

Dai, G., Steede, N. K., & Landry, S. J. (2001). Allocation of helper T-cell epitope immunodominance according to three-dimensional structure in the human immunodeficiency virus type I envelope glycoprotein gp120. *J Biol Chem, 276*(45), 41913-41920. doi:10.1074/jbc.M106018200

Ferrante, A. (2013). HLA-DM: arbiter conformationis. *Immunology, 138*(2), 85-92. doi:10.1111/imm.12030

Hume, D. A. (1985). Immunohistochemical analysis of murine mononuclear phagocytes that express class II major histocompatibility antigens. *Immunobiology, 170*(5), 381-389. doi:10.1016/S0171-2985(85)80062-0

Kim, A., Hartman, I. Z., Poore, B., Boronina, T., Cole, R. N., Song, N., . . . Sadegh-Nasseri, S. (2014). Divergent paths for the selection of immunodominant epitopes from distinct antigenic sources. *Nat Commun, 5*, 5369. doi:10.1038/ncomms6369

Kim, A., & Sadegh-Nasseri, S. (2015). Determinants of immunodominance for CD4 T cells. *Curr Opin Immunol, 34*, 9-15. doi:10.1016/j.coi.2014.12.005

Neefjes, J., Jongsma, M. L., Paul, P., & Bakke, O. (2011). Towards a systems understanding of MHC class I and MHC class II antigen presentation. *Nat Rev Immunol, 11*(12), 823-836. doi:10.1038/nri3084

Painter, C. A., Cruz, A., Lopez, G. E., Stern, L. J., & Zavala-Ruiz, Z. (2008). Model for the peptide-free conformation of class II MHC proteins. *PLoS One, 3*(6), e2403. doi:10.1371/journal.pone.0002403

Pos, W., Sethi, D. K., Call, M. J., Schulze, M. S., Anders, A. K., Pyrdol, J., & Wucherpfennig, K. W. (2012). Crystal structure of the HLA-DM-HLA-DR1 complex defines mechanisms for rapid peptide selection. *Cell, 151*(7), 1557-1568. doi:10.1016/j.cell.2012.11.025

Richards, K. A., Chaves, F. A., Krafcik, F. R., Topham, D. J., Lazarski, C. A., & Sant, A. J. (2007). Direct ex vivo analyses of HLA-DR1 transgenic mice reveal an exceptionally broad pattern of immunodominance in the primary HLA-DR1-restricted CD4 T-cell response to influenza virus hemagglutinin. *J Virol, 81*(14), 7608-7619. doi:10.1128/JVI.02834-06

Rupp, B., Gunther, S., Makhmoor, T., Schlundt, A., Dickhaut, K., Gupta, S., . . . Kuhne, R. (2011). Characterization of structural features controlling the receptiveness of empty class II MHC molecules. *PLoS One, 6*(4), e18662. doi:10.1371/journal.pone.0018662

Schulze, M. S., Anders, A. K., Sethi, D. K., & Call, M. J. (2013). Disruption of hydrogen bonds between major histocompatibility complex class II and the peptide N-terminus is not sufficient to form a human leukocyte antigen-DM receptive state of major histocompatibility complex class II. *PLoS One, 8*(7), e69228. doi:10.1371/journal.pone.0069228

Stern, L. J., Brown, J. H., Jardetzky, T. S., Gorga, J. C., Urban, R. G., Strominger, J. L., & Wiley, D. C. (1994). Crystal structure of the human class II MHC protein HLA-DR1 complexed with an influenza virus peptide. *Nature, 368*(6468), 215-221. doi:10.1038/368215a0

Yin, L., & Stern, L. J. (2013). HLA-DM Focuses on Conformational Flexibility Around P1 Pocket to Catalyze Peptide Exchange. *Front Immunol, 4*, 336. doi:10.3389/fimmu.2013.00336