Characterization of MHC Class I and Class II-associated Peptides

Qianqian Shao

Alexandria, Virginia

Bachelor of Science in Chemistry with specialization in Biochemistry

University of Virginia, 2019

A Dissertation Presented to the Graduate Faculty of the University of Virginia

in Candidacy for the Degree of Doctor of Philosophy

Department of Chemistry

University of Virginia

Nov 2024

Acknowledgement

I would first like to express my deep gratitude to my advisor, Dr. Donald F. Hunt, for giving me the opportunity to conduct research in mass spectrometry and proteomics. I am also immensely thankful to Dr. Jeffrey Shabanowitz for his invaluable guidance and support throughout my graduate studies. My sincere thanks go to Dr. Mark M. Ross and Dr. Dina L. Bai for their constructive feedback on my presentations and writing. I am deeply appreciative of everyone who made my graduate school experience far better than I could have ever expected.

I am also grateful to the current and former members of the Hunt lab for sharing their knowledge and support during my time in both academia and industry. Special thanks to Dr. Keira Mahoney, Dr. Xi Peng, Dr. Maria Panepinto, and Dr. Negin Ghafourian for mentoring me on various lab techniques and sharing your insightful expertise. I extend my heartfelt thanks to Dr. Feng Yang and Dr. Erin D. Jeffery for their support during my internship at Genentech. I am also grateful to Dr. Seamus Kelley, Larry Reser, and Melissa Leyden for the pleasure of working together. Larry, thank you for stopping by my bench occasionally—I'll certainly miss our conversations—and for taking care of my plants while I was away on my internship.

I also want to thank Dr. Miaomiao Chen and Wenxin (Summer) Xu for the incredible friendship we built, supporting each other throughout our Ph.D. journeys. You are role

models of strong women. I believe the strength we gained from facing challenges together will empower us in the future.

I am incredibly grateful to my collaborators, Dr. John Sidney, Dr. Alessandro Sette, Paul Rubiro, Dr. Marcos Pires, Joey Kelly, Dr. James Heather, and Dr. Sara Pai, for the wonderful collaborations we shared.

I would like to thank Dr. David Cafiso, Dr. Jill Venton, Dr. Linda Columbus, and Dr. David Kittlesen for serving on my committee, and I extend additional thanks to Dr. David Cafiso, Dr. Jill Venton, and Dr. Ken Hsu for previously serving on my candidacy committee.

A very special thanks to my high school teacher and best friend Mr. John Hawes. Over the past ten years, our discussions on politics, history, philosophy, sociology, literature, and countless other topics have broadened my worldview and inspired my efforts in diversity, equity, and inclusion (DEI) during my Ph.D. studies. Thank you for your wisdom and friendship.

Lastly, I would like to express my heartfelt thanks to my parents Ying Zhang and Zong Shao for their immense courage in bringing our family to the United States, making all these meaningful encounters and opportunities possible. I want to thank my uncles Yinrong Shao and Yu Shao, and my aunt Emily Chen for their support of my family. My grandparents Chenlan Lin and Kangdi Shao, Yuanlan Qu and Jiayun Zhang for their unconditional love.

Table of Contents

List of Figures and Tables

Chapter 1: Introduction

Chapter 2: Characterization of MHC Class I-Associated Peptide in Mouse Cell Line

Chapter 3: Characterization of Hamster MHC for the Study of COVID

Chapter 4: Characterization of MHC Class I-Associated Phosphopeptides

List of Abbreviations

℃	degrees Celsius
β2m	beta-2 microglobulin
AC	analytical column
ACN	acetonitrile
AGC	automatic gain control
Angio	angiotensin I (DRYIHPFHL)
Ala,	A alanine
Arg,	R arginine
Asn,	N asparagine
Asp,	D aspartic acid
C18	octadecyl
CID	collision-induced dissociation
CHAPS	3- [(3-cholamidopropyl) dimethylammonio]-1-propane sulfonate
CIP2A	cancerous inhibitor of PP2A

Cys, C	cysteine
CLIP	Class II-associated invariant chain peptide
Da	Dalton, 1 amu
EDTA	ethylenediaminetetraacetic acid
ELISpot	enzyme-linked immunospot assay
ER	endoplasmic reticulum
ESI	electrospray ionization
ETD	electron transfer dissociation
EIC	extracted ion chromatogram
Gln, Q	glutamine
Glu, E	glutamic acid
Gly, G	glycine
HCD	higher-energy collisional dissociation
HCl	hydrochloric acid
HeLa	cervical adenocarcinoma cell line
HILIC	hydrophilic interaction liquid chromatography

His, H	histidine
HIV	human immunodeficiency virus
HLA	human leukocyte antigen
HPLC	high performance/pressure liquid chromatography
HPV	human papillomavirus
I2PP2A	inhibitor 2 of PP2A (protein SET)
i.d.	inner diameter
IDA	iminodiacetic acid
IL	interleukin
Ile, I	isoleucine
IMAC	immobilized metal affinity chromatography
IP	immunoprecipitation
IFN-γ	interferon gamma
IRM	ion routing multipole
KaSil	potassium silicate solution
LC	liquid chromatography/ chromatograph

Leu, L	leucine
Lys, K	lysine
MeOH	methanol
Met, M	methionine
Mg2+	magnesium (II) cation
MgCl2	magnesium chloride
MHC	major histocompatibility complex
MS	mass spectrometry
MS1	full mass spectrum
MS2	tandem mass spectrum
MW	molecular weight
NTA	nitrilotriacetic acid
N-term	amino-terminus of peptide
o.d.	outer diameter
p53	tumor protein 53
PBS	phosphate-buffered saline

PC	pre-column
PEG	polyethylene glycol
Phe, F	phenylalanine
PHEA	polyhydroxy ethyl aspartamide
pKa	acid dissociation constant
PP1	protein phosphatase 1
PP2A	protein phosphatase 2A
ppm	parts per million
pRb	retinoblastoma protein
Pro, P	proline
PTM	post-translational modification
RB	retinoblastoma
RF	radio frequency
RNA	ribonucleic acid
RP	reverse phase
RT	room temperature, retention time

S/N	signal to noise
Ser, S	serine
SET	protein SET
STAGE	stop and go extraction
TAP	transporter associated with antigen processing
TCR	T-cell receptor
TFA	trifluoroacetic acid
Thr, T	threonine
Treg	regulatory T-cells
Tris	tris(hydroxymethyl)aminomethane
Trp, W	tryptophan
Tyr, Y	tyrosine
Val, V	valine
Vaso/VIP	vasoactive intestinal peptide (HSDAVFTDNYTR)

Dissertation Abstract

Major Histocompatibility Complex (MHC) molecules play a crucial role in the immune system by presenting peptides to T-cells, allowing for the recognition of infected or abnormal cells. This dissertation presents three projects focused on the identification of peptides presented through the MHC class I and/or class II pathways, contributing to a better understanding of immune recognition and potential therapeutic applications.

The first project focuses on the verification of peptide presentation using a combination of chromatographic and tandem mass spectrometry (MS2) data. This project involved ensuring that the correct synthetic peptides—either the wild-type ovalbumin peptide (OvaWT) or its trimethylated lysine variant (OvaK7m3)—were bound to MHC molecules. By applying rigorous analytical techniques, including the RMA-S stabilization assay, this study confirmed that the expected peptides were presented on MHC molecules. This validation is critical for ensuring the reliability of MHC-peptide interaction studies, paving the way for further exploration of post-translational modifications and their effects on immune responses.

The second project aimed to characterize the MHC class I peptide binding motifs of Golden Syrian hamsters, an emerging model organism for COVID-19 research. This study was essential for identifying viral epitopes, particularly from the spike protein of SARS-CoV-2, that could be presented by hamster MHC molecules. By elucidating these binding motifs, the project provides a foundation for developing predictive algorithms to identify

immunogenic peptides from viral proteins in this species. This work enhances the use of hamsters in immunological studies, especially for infectious disease research and vaccine development.

The third project investigates the presentation of phosphopeptides on MHC class I molecules in head and neck squamous cell carcinoma (HNSCC) cells. Phosphorylation is a key regulator of cellular processes, and in cancerous cells, phosphopeptides can be presented on MHC molecules due to disruptions in normal phosphatase activity, particularly by SET and CIP2A proteins. Using advanced sample preparation techniques, including hydrophilic interaction chromatography (HILIC) and immobilized metal affinity chromatography (IMAC), along with LC-MS/MS analysis, 27 phosphopeptides were identified in HNSCC cells. Many of these phosphopeptides were also found in other cancerous and infected cells, suggesting their potential as universal vaccine candidates for both cancer and infectious diseases.

Together, these three projects provide significant insights into MHC peptide presentation, with applications ranging from the study of post-translational modifications in model systems to the identification of viral epitopes and cancer-specific antigens. These findings have broad implications for the development of immunotherapies and vaccines.

Chapter 1: Introduction

1.1 Introduction to Major Histocompatibility Complex (MHC)

1.1.1 Major Histocompatibility Complex (MHC) Antigen Processing and Presentation Pathway

The MHC peptide presentation pathway is a critical component of the immune system, responsible for the processing and presentation of foreign and self-antigens to T cells. This process enables the immune system to distinguish between normal, healthy cells and those that are infected or abnormal, such as cancerous cells. [1, 2]

There are two primary classes of MHC molecules: MHC class I and MHC class II.

MHC Class I Pathway is primarily involved in presenting intracellular peptide antigens, such as those derived from a virus or other intracellular pathogen proteins. The proteins in the cell are degraded into small peptides by the proteasome, and these peptides are then transported into the endoplasmic reticulum (ER) via the transporter associated with antigen processing (TAP). Within the ER, the peptides bind to MHC class I molecules, which are then transported to the cell surface, where they are presented to cytotoxic CD8+ T cells, which can recognize and kill infected or abnormal cells through an interaction between the T-cell receptor and the MHC-peptide complex.

In contrast, the MHC class II pathway is involved in presenting extracellular antigens. These antigens are typically derived from pathogens that have been endocytosed by antigen-presenting cells (APCs) such as macrophages, dendritic cells, or B cells. The endocytosed proteins are processed in endosomes and lysosomes, where they are broken down into peptides. These peptides are then loaded onto MHC class II molecules within the endosomal compartments. The MHC class II-peptide complexes are subsequently transported to the cell surface, where they are presented to helper CD4+ T cells, which play a crucial role in activating B cells to produce antibodies. [3,4]

Both pathways are essential for the immune system's ability to monitor and respond to a wide range of pathogens, ensuring that the body can effectively target and eliminate infected or abnormal cells while maintaining tolerance to self-antigens (proteins/peptides).

In the context of peptide presentation, MHC class I and class II molecules function differently to achieve this immune surveillance. Unlike MHC class I molecules, which typically bind shorter peptides with defined termini, MHC class II molecules have an open binding groove that can accommodate peptides of varying lengths, with a central core region anchoring the peptide to the MHC (**Figure 1-1**). This structural feature leads to the formation of overlapping sequences, where different peptides can share a common core but vary at the end. [5]

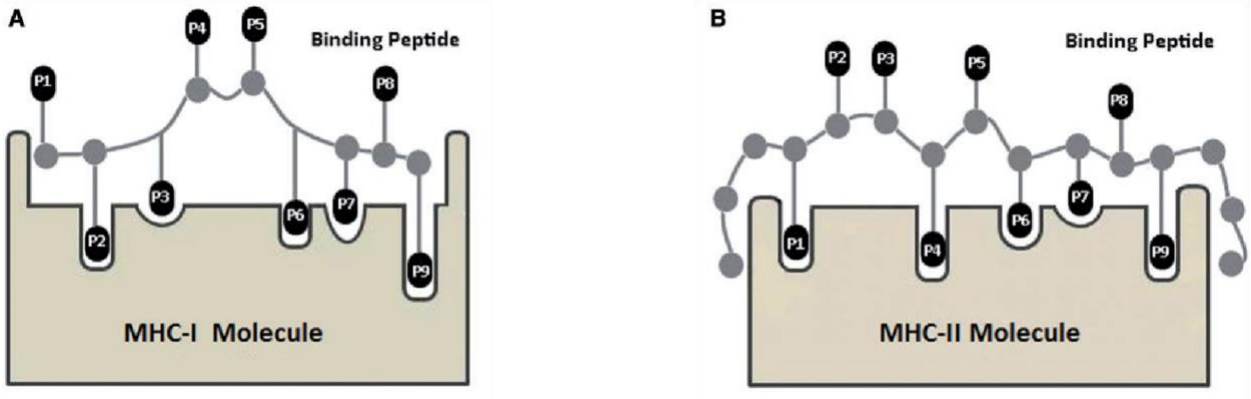


Figure 1- 1: *MHC Class I and Class II Binding Groove*[5]

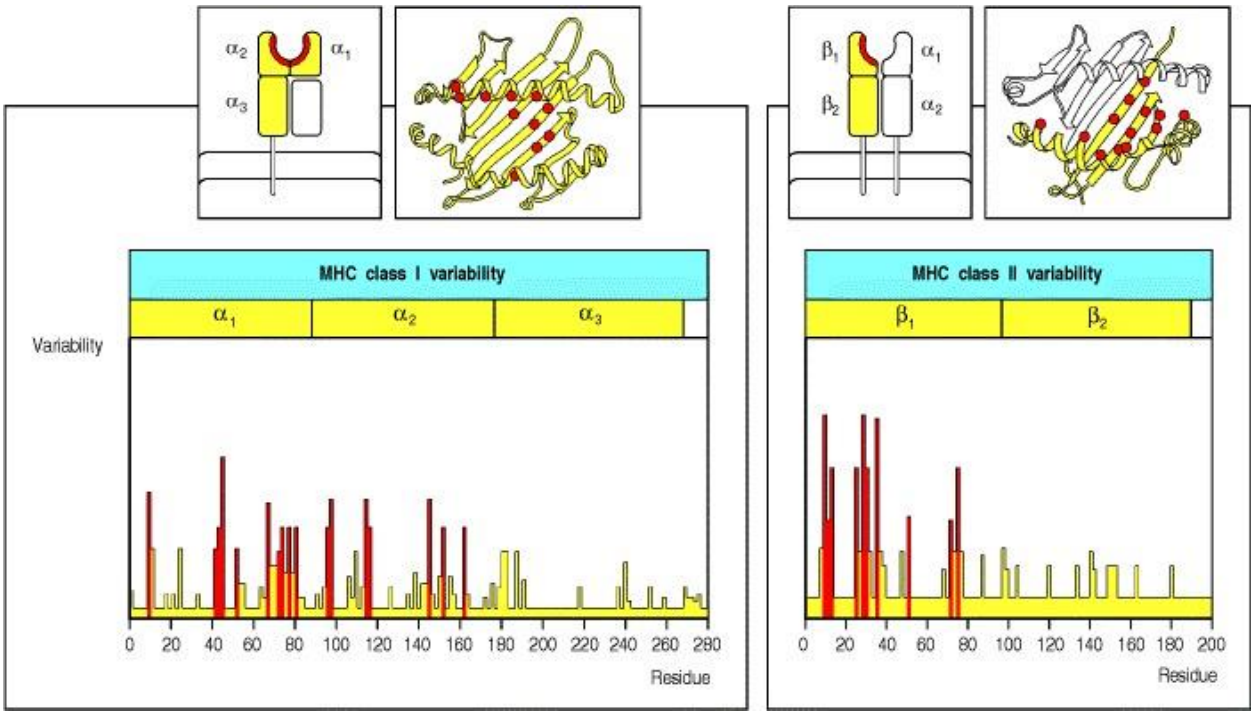
1.1.2 MHC Polymorphism

MHC molecules are highly polymorphic, making it difficult to generate pan-antibodies that could enrich for all MHC molecules, and make it hard to distinguish different alleles for each MHC gene.

For MHC class I molecules, HLA-A, HLA-B, and HLA-C are composed of an α chain and a beta-2 microglobulin (β2m). Since β2m is shared across different HLA types, the antibody W6/32 is highly effective as a pan-MHC antibody for enriching all HLA class I molecules.

However, MHC class II molecules do not have β2m, making it much more challenging to generate a pan-MHC antibody for class II molecules. Among the individual MHC class II proteins, HLA-DR has an α chain that is nearly invariant, with significant polymorphism occurring only in the β chain (shown in **Figure 1-2**). In contrast, both the α and β chains of HLA-DQ and HLA-DP exhibit significant polymorphism. [6,7]

a.



b.

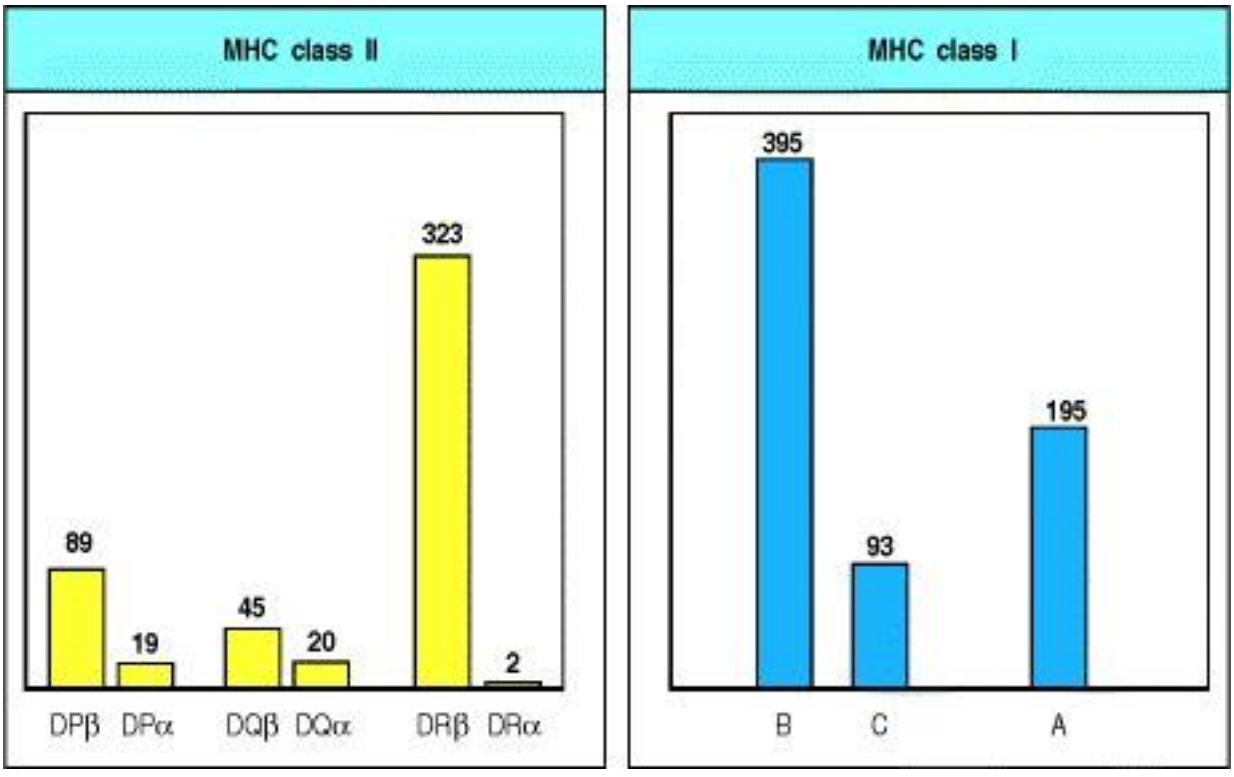


Figure 1- 2: *Polymorphism of MHC class I and class II molecules*[6]

a. The variability plots of the amino acid sequences of MHC molecules are shown. For the MHC class II molecule, the variability of the HLA-DR alleles is shown. b. The height of the bars in this figure represents the number of different alleles. These numbers correspond to the HLA alleles officially designated by the WHO Nomenclature Committee for Factors of the HLA System as of August 2000.

1.2 MHC-peptide Enrichment by Immunoprecipitation

Immunoprecipitation (IP) is a widely used laboratory technique that enables isolation of MHC molecules from cell lysate. The principle behind immunoprecipitation is based on the specific interaction between an MHC-specific antibody and its targeted MHC molecules.

An antibody specific to the MHC molecules of interest is introduced into a sample containing a complex mixture of cell lysate. The antibody binds to MHC molecules, forming an antibody-antigen complex.

To isolate this complex, the antibody-antigen pairs are then captured using a solid support, such as agarose or magnetic beads that are coated with protein A, protein G, or an antibody that recognizes the primary antibody. In this research, agarose beads were used. The beads allow for the antibody-protein complexes to be separated from the cell lysate protein mixture by centrifugation or a magnetic field.

The bead-bound complexes are washed multiple times to remove non-specifically bound proteins and other contaminants, ensuring that only the MHC-peptide complexes remain attached to the beads.

The MHC associated peptides are then eluted from the beads, usually by acid elution that disrupts the antibody-antigen interaction. The isolated MHC-associated peptides then can be analyzed using mass spectrometry to identify the peptides, including their modifications and abundances. [8,9,10]

1.3 Reverse Phase Liquid Chromatography (RPLC)

Reverse-phase liquid chromatography (RPLC) is a widely used technique in proteomics, particularly when coupled with mass spectrometry (MS) for the analysis of complex protein mixtures. In this technique, the stationary phase is nonpolar, while the mobile phase is more polar. [11,12]

One of the most common stationary phases used in RPLC is C18 media, which consists of octadecylsilane (ODS) bonded silica particles with 18-carbon alkyl chains, making it highly hydrophobic.

The mobile phase solvents typically consist of a combination of water (with 0.1% acetic acid for better ionization, separation and sensitivity) and an organic solvent such as acetonitrile (ACN). Other acids, such as TFA, yield improved separation but are not compatible with ESI-MS.

When a peptide sample is introduced into the column, the peptides interact with the hydrophobic C18 chains.

A gradient is employed where the proportion of the organic solvent is gradually increased. Hydrophilic peptides elute earlier when the mobile phase is aqueous, and hydrophobic peptides elute later as the mobile phase organic content increases.

Peptides are separated based on their differential partitioning between the mobile and stationary phases, which is influenced by their hydrophobicity. As peptides pass through the C18 column, peptides with greater hydrophobicity are retained longer, resulting in their separation from less hydrophobic peptides.

This separation is crucial in proteomics because it enables the mass spectrometer analysis of a smaller number of peptides per elution time, rather than as a complex mixture, leading to higher sensitivity and resolution, hence more accurate peptide identification and quantification.

1.4 Introduction to Mass Spectrometry

1.4.1 Overview of Mass Spectrometry

Mass spectrometry is an analytical technique used to analyze a wide range of molecules, including small molecules, peptides, proteins, inorganic ions, organic compounds, biomolecules, and large macromolecules like nucleic acids and polymers. The basic components of a mass spectrometer include the ion source, mass analyzer, and detector. The ion source ionizes analytes to make them charged, and a common peptide/protein ionization technique is electrospray ionization (ESI). The mass analyzer sorts ions based on their mass-to-charge ratio (m/z) and can induce fragmentation in

tandem mass spectrometry instrument configurations. Quadrupoles, Orbitraps, and quadrupole ion traps are examples of mass analyzers.

Ions are separated in the mass analyzer according to their m/z ratios, following distinct trajectories within the electric field. They are then detected either with an electron multiplier (quadrupole ion trap) or through ion-induced image current (Orbitrap), producing a mass spectrum that displays ion abundance and m/z ratios.

Additional components like lenses and quadrupoles focus and transport ions, often using voltage ramps.

Quadrupoles can also induce fragmentation via Collision-Induced Dissociation (CID) or High-Energy Collision Dissociation (HCD). In Orbitrap mass spectrometry, a C-trap accumulates ions and injects them into the Orbitrap for high-resolution analysis. An ion routing multiple is used as an HCD collision cell for HCD fragmentation, and it can be used to transport ions for CID fragmentation in the ion trap. For tandem MS, ions can be fragmented in an ion trap and then analyzed by the Orbitrap to achieve detailed spectral data. [13,14]

1.4.2 Significance of m/z

Mass spectrometry (MS) is used to measure the masses of molecules ranging from low (<100 Da) to very high (mega Da) molecular weights, requiring only minimal quantities of analyte. To be analyzed, the molecules must be ionized in the gas phase, as MS relies on electric fields to manipulate, focus, and separate ions. The masses are measured relative

to their charge, represented by the mass-to-charge ratio (m/z). The presence of charge is essential, as only charged species can be manipulated and detected by the mass spectrometer, while neutral species do not generate a signal.

The m/z value is significant because it allows for the differentiation of ions. Different ions have unique compositions and isotopic patterns, leading to distinct m/z values in the mass spectrum. By analyzing these m/z values, one can distinguish between different ions and, in some cases, infer structural or compositional details.

Additionally, the charge state of an ion can be determined by examining the spacing between peaks corresponding to different isotopes of the same ion. This information, combined with m/z, is critical for identifying and characterizing the analyte, providing insights into its molecular weight and structure.

1.4.3 Electrospray Ionization

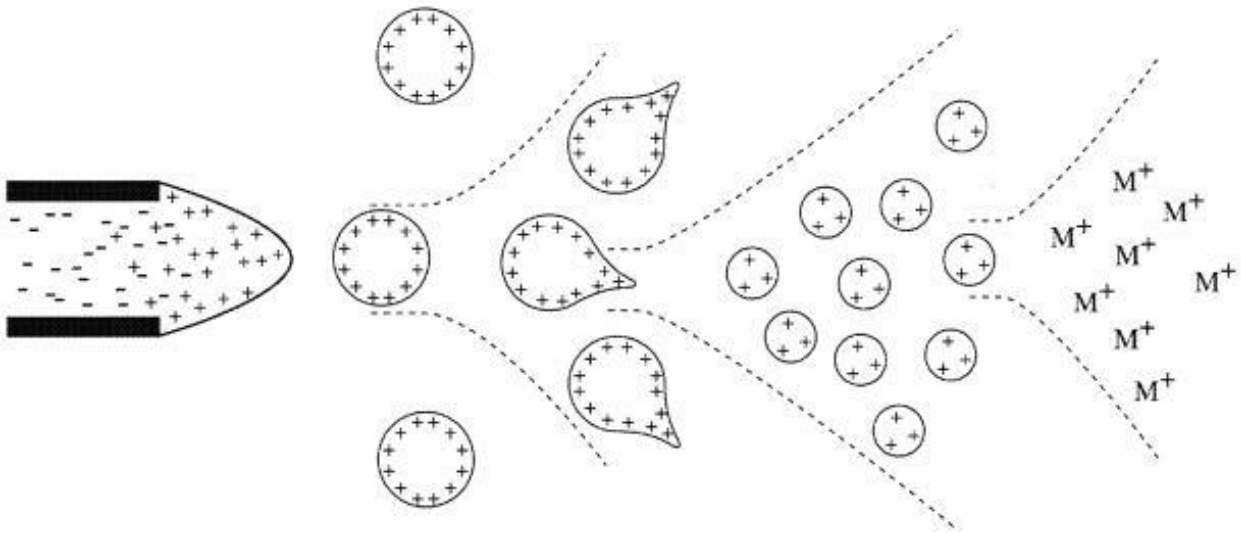


Figure 1- 3: *Mechanism of Electrospray Ionization* [15]

Electrospray Ionization (ESI) is a widely used ionization technique in mass spectrometry, particularly effective for analyzing large biomolecules like proteins, peptides, and nucleic acids, as well as smaller organic compounds. ESI is known for its ability to gently ionize complex molecules without causing significant fragmentation, making it ideal for studying proteomics.

ESI is considered a "soft" ionization technique because it ionizes molecules without causing extensive fragmentation, preserving the integrity of large and complex biomolecules by imparting low internal energy.[16]

The sample is dissolved in a liquid solvent and is introduced into the mass spectrometer by liquid chromatography (LC) through a narrow, fused silica capillary.

A high voltage is applied to the capillary to create a strong electric field between the capillary tip and the MS ion source. As the liquid exits the capillary, it forms a fine spray of highly charged droplets. This process can be aided by a counterflow of gas, which helps to desolvate the droplets.

As the solvent evaporates from the droplets, they become smaller and more concentrated in charge (**Figure 1-3**). Eventually, the Coulombic repulsion between like charges within a droplet causes it to undergo fission, producing even smaller droplets. This process continues until individual analyte molecules are released from the droplets as multiply charged ions. The resulting ions are then directed (focused/transmitted) into the mass analyzer for measurement. [15,16,17,18]

In our lab, instead of a positively charged ESI probe needle, a wire is connected to the ion source and the waste line. Voltage is applied to the wire, creating a circuit with a positive end at the waste line pulling negatively charged contaminants into the waste, and a negative end at the ion source facilitating the ejection of positively charged analytes.

1.4.4 Fragmentation Techniques

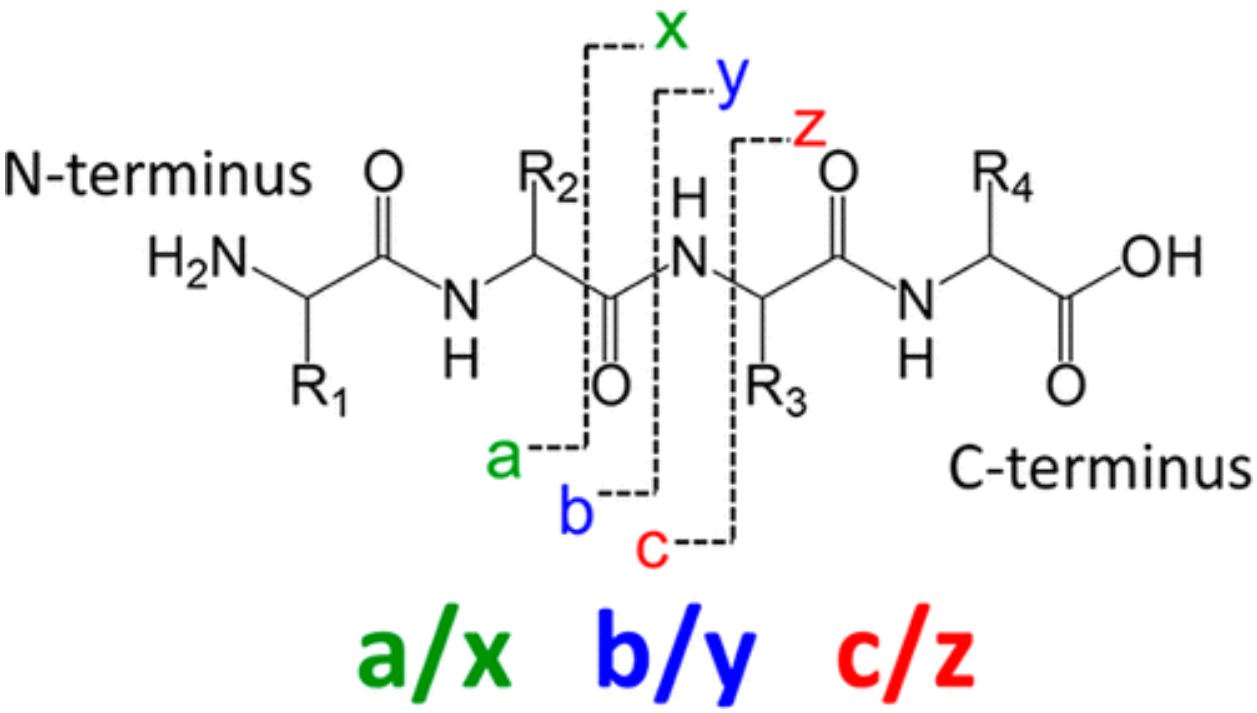


Figure 1- 4: *Nomenclature of Peptide Fragmentation* [19]

Fragmentation, which reveals the composition of the peptide amino acid sequence, is typically accomplished in a dedicated collision cell or in certain types of mass spectrometers such as ion traps or quadrupoles. These instruments can function as collision cells in techniques like Collision-Induced Dissociation (CID) and Higher-Energy

Collisional Dissociation (HCD), where gasses such as helium and nitrogen, respectively, are used to generate collision energy that breaks chemical bonds within the analytes, generating b and y ions (**Figure 1-4**).

An alternative fragmentation method is Electron Transfer Dissociation (ETD), in which electrons are transferred from ETD reagent ions to the analyte ions, inducing bond cleavage while preserving certain structural features, like post-translational modifications, making it a valuable tool for proteomic analysis.

In ETD, an electron is transferred from a radical anion to the peptide or protein ion, inducing fragmentation of the peptide backbone without disrupting fragile PTMs like phosphorylation, glycosylation, or ubiquitination. This results in the production of c- and z-type fragment ions (**Figure 1-4**), which can be used for sequence determination and localization of modifications.

ETD is especially valuable for analyzing large, multiply charged ions, such as those found in proteomics, as it complements other fragmentation methods like collision-induced dissociation (CID), which often leads to loss of labile modifications. By preserving these modifications while cleaving the peptide backbone, ETD provides more detailed structural information, making it a powerful tool for studying complex biomolecules. [19,20]

1.4.5 Mass Analyzer: Ion Trap

Mass analyzers are used to detect ions after fragmentation has occurred. Ion Trap Mass Spectrometry is a versatile and powerful analytical technique widely used for the

analysis of ions based on their mass-to-charge ratio (m/z). The ion trap functions both as a mass analyzer and as a tool for ion isolation and fragmentation.

An 3-dimensional (3D) ion trap, also known as a quadrupole ion trap, consists of three main components: a ring electrode and two end cap electrodes, which together form a three-dimensional quadrupole electric field. This setup creates a potential well where ions can be trapped and manipulated.

Ions are introduced into the trap and confined within the potential well created by applying RF voltage to the ring electrode. The applied RF voltage generates a dynamic electric field that traps ions in stable orbits within the trap. The ions oscillate in a specific pattern based on their m/z ratio, allowing for the selective retention or ejection of ions.

In the ion trap, ions are analyzed by destabilizing their orbits, causing them to be ejected sequentially based on their m/z ratio. This is done by gradually varying the RF voltage, which causes ions with different m/z values to resonate and be ejected from the trap at different times. The ejected ions are then detected, and their m/z ratios are determined. [21,22,23]

An 2-dimensional (2D) linear ion trap consists of three quadrupole mass filters, DC voltage is applied to each segment, creating a radially confining force to trap the ion inside. Because there is no force applied to the axial dimension, it allows the ions to spread out and can store more charges than 3D ion trap. [23]

Both 3D and 2D ion traps can perform MS/MS experiments. After isolating a specific ion (precursor ion), the ion trap can induce fragmentation, through collision-induced dissociation (CID) or electron transfer dissociation (ETD). The resulting fragment ions (product ions) are then analyzed, providing detailed structural information about the precursor ion.

The ion detection in a quadrupole ion trap is accomplished with an electron multiplier, where ions are ejected from the trap and directly detected by amplifying the impact of ions. [21,22]

1.4.6 Mass Analyzer: Orbitrap

In addition to quadrupole ion trap instruments, an Orbitrap (**Figure 1-2**) provides ion detection and m/z measurement with higher resolution for MS1 and MS2 analyses. The ion detection in Orbitrap is accomplished with image current detection.

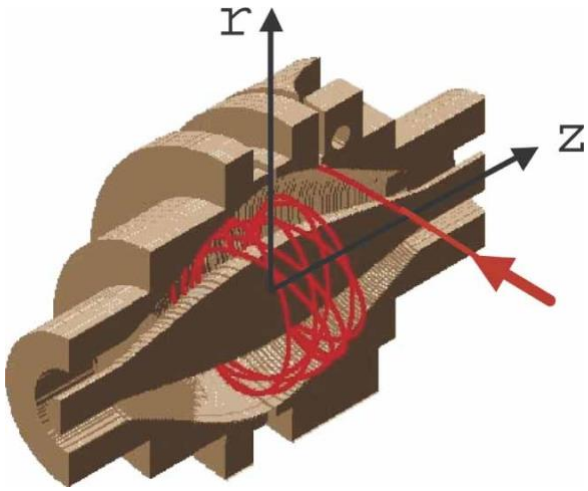


Figure 1- 5: *Orbitrap Cutaway View* [24]

Ion detection in a quadrupole ion trap is achieved using an electron multiplier, where ions are ejected from the trap and directly detected by amplifying the impact of ions. In contrast, an Orbitrap offers significantly higher resolution for both MS1 and MS2 analyses. Instead of direct ion counting, the Orbitrap detects ions using image current detection, where the oscillations of trapped ions generate a measurable current without destroying the ions. This method allows for more precise mass-to-charge (m/z) measurements compared to the quadrupole ion trap.

Mass analyzers are used to detect ions after fragmentation has occurred. High-resolution mass analyzers such as Orbitrap and Fourier Transform Ion Cyclotron Resonance (FTICR) provide highly accurate detection of ions, including those that have been fragmented by other parts of the instrument. [21, 23]

Orbitrap differs from other techniques in (A) how ions are transferred to Orbitrap, (B) how ions are trapped, and (C) how ions are detected.

(A) Ions are first produced in the ion source. These ions are focused and guided into a quadrupole which selects ions based on their mass-to-charge (m/z) ratio. The selected ions are then directed to the C-trap. The C-trap accumulates ions before they are injected into the Orbitrap. When ions are ejected from the C-trap, they are introduced into the Orbitrap in a way that aligns them with the central axis of the Orbitrap's electric field. This careful alignment helps ensure that the ions enter the Orbitrap in a controlled manner, optimizing trapping efficiency and minimizing deviations in their harmonic oscillation. [14]

(B) The Orbitrap consists of a central spindle-like electrode and a coaxial outer barrel electrode. Ions are trapped in the Orbitrap by the electrostatic attraction to the central spindle electrode, which is held at a high DC potential, typically negative for trapping positive ions. The outer electrode confines the ions in a stable orbit around the central electrode. The ions oscillate harmonically along the axis of the Orbitrap due to the specific shape of the electric field generated by the electrodes.

(C) The detection in an Orbitrap is based on the frequency of harmonic oscillations of ions along the central axis. These oscillations generate an image current in the outer barrel electrode, which is then detected. This signal is processed using Fourier transform to convert the time-domain oscillation data into a frequency-domain mass spectrum. The frequency of oscillation is inversely proportional to the square root of the m/z ratio, meaning ions with higher m/z values oscillate at lower frequencies, while those with lower m/z values oscillate at higher frequencies. This predictable oscillation behavior is key to the Orbitrap's ability to achieve high-resolution mass spectrometry. The oscillation frequency directly reflects the m/z. [24]

There is a trade-off between mass resolution and sensitivity in the Orbitrap. The high-resolution capability of the Orbitrap stems from its precise measurement of ion oscillation frequencies, which correspond to the m/z-dependent motion of ions within the trap. This intrinsic accuracy in detecting the frequencies of ion image currents allows for superior resolution in m/z measurements. However, the instrument's sensitivity to

harmonic oscillations also makes it more susceptible to space-charge effects, where an excessive number of ions can distort the electric field and reduce overall resolution.

To mitigate this, the number of ions entering the Orbitrap is controlled, which can limit sensitivity. While longer analyzing times and a larger number of oscillations can improve resolution, this also means fewer ions are detected, impacting sensitivity.

While the Orbitrap offers high resolution, it is also sensitive to space-charge effects like other ion traps, necessitating a balance between resolution and sensitivity, achieved through careful control of ion density and analysis conditions.

1.5 Data Processing

1.5.1 De Novo Sequencing

After a MS2 mass spectrum of a peptide is generated, the peptide sequence can be determined by analyzing the mass differences between consecutive peaks.

These mass differences correspond to the masses of individual amino acids. Starting from either the N- or C-terminus, the peptide sequence is built by matching these mass differences to known amino acid masses. The sequence is then validated by ensuring that the total mass matches the precursor ion and that all fragment ions (both the b- and y-ion for CID/HCD, and c- and z- ions for ETD) are consistent.

This process can be done manually, known as de novo sequencing, or using algorithms that compare the mass spectrum to known protein sequence databases, which is faster and more automated.

1.5.2 Byonic Protein Metrics

Byonic is a powerful software search program designed for mass spectrometry-based proteomics analysis. It is used to identify proteins, peptides, and post-translational modifications (PTMs) from complex MS data.

The engine supports high-resolution mass spectrometry data and allows users to perform flexible searches, accommodating missed cleavages, mass shifts, and unexpected modifications. Byonic uses advanced scoring algorithms to match experimental MS/MS spectra to theoretical peptide sequences, ensuring high sensitivity and accuracy in identification. Its strength lies in handling complex biological samples, such as those with non-tryptic peptides, labile modifications, or glycopeptides, making it a favored tool in advanced proteomics research. [25]

However, like any other search engine, Byonic is also prone to false-positive identifications. In this research, manual validation was performed to ensure the accuracy of the results.

1.5.3 Data Visualization for MHC Peptide Binding Motif Determination

Data visualization is aided by R programming. In this research, I wrote an R program to visualize and present the results of MS data analysis based on sequence alignment, phosphopeptide site localization, and peptide length count.

I also wrote R programs to convert the Byonic results into an appropriate format for further analysis using other programs such as GibbsCluster 2.0 and NetMHCIIpan 4.2 for MHC class I and II peptides, respectively.

1.5.4 GibbsCluster 2.0

For analysis of MHC class I peptides, GibbsCluster 2.0 was used to find peptide binding motifs. GibbsCluster 2.0 is a computational tool designed for clustering peptide sequences, typically derived from mass spectrometry experiments. It is particularly useful for analyzing large sets of peptides, such as those associated with MHC binding or post-translational modifications and identifying shared sequence motifs. The tool uses a probabilistic clustering approach based on Gibbs sampling, which allows it to group peptides into clusters with similar sequence patterns and motifs. [26]

GibbsCluster 2.0 ranked the frequency of amino acids at each position of the MHC class I binding cleft and generated a sequence logo, as shown below. In the logo, the increasing font size of an amino acid indicates its increasing frequency at a given position. GibbsCluster is very useful for revealing binding motifs for MHC class I molecules. As shown in **Figure 1-6**, analysis of a set of MHC class I-associated peptides identified by MS

yielded a high frequency of glycine at position 2, high frequency of proline and alanine at position 3, and high frequency of leucine, isoleucine, phenylalanine and valine at position 9.

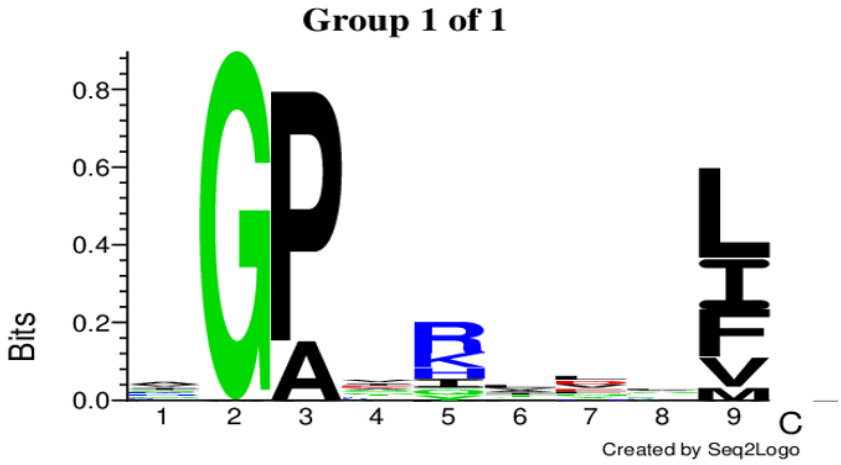


Figure 1- 6*: GibbsCluster Sequence Logos*

1.5.5 NetMHCIIpan 4.2

For analysis of MHC class II peptides, NetMHCIIpan 4.2 was used to predict the binding affinity of MHC class II peptides identified in the sample. NetMHCIIpan 4.2 is a computational prediction tool used to identify peptide binding to MHC class II molecules. It is part of the NetMHC family of tools and is designed to predict peptide binding to MHC class II alleles across multiple species, including humans. Unlike traditional methods that are limited to specific alleles, NetMHCIIpan uses pan-specific models, allowing it to predict binding for any MHC class II allele, even those for which experimental data are limited or unavailable. [27,28]

The NetMHCIIpan 4.2 uses the NNAlign_MA machine-learning framework [29] and was trained on combined data from MHC binding affinity (BA) assay, MHC II immunopeptidome eluted ligand (EL) assay obtained by LC-MS/MS, and patterns of proteolytic digestion.

1.6 References

1. Neefjes, J.; Jongsma, M. L. M.; Paul, P.; Bakke, O., Towards a systems understanding of MHC class I and MHC class II antigen presentation. *Nature Reviews Immunology* **2011**, 11 (12), 823-836.

2. Yang, K., Halima, A. & Chan, T.A. Antigen presentation in cancer — mechanisms and clinical implications for immunotherapy. *Nat Rev Clin Oncol* 20, 604–623 (**2023**). https://doi.org/10.1038/s41571-023-00789-4

3. Parham, P. The Immune System, 4th ed.; Garland Publishing: New York, NY, **2014**.

4. Murphy, K.; Weaver, C. Janeway's Immunobiology, 9th ed.; CRC Press: Boca Raton, FL, **2016**.

5. Zhang, L.; Udaka, K.; Mamitsuka, H.; Zhu, S. Toward More Accurate Pan-Specific MHC-Peptide Binding Prediction: A Review of Current Methods and Tools. *Brief. Bioinform.* **2012**, 13 (3), 350–364. https://doi.org/10.1093/bib/bbr060.

6. Janeway, C. A.; Travers, P.; Walport, M.; Shlomchik, M. J. Immunobiology: The Immune System in Health and Disease, 6th ed.; CRC Press: Boca Raton, FL, **2004**.

7. Unanue, E. R.; Turk, V.; Neefjes, J. Variations in MHC Class II Antigen Processing and Presentation in Health and Disease. *Annu. Rev. Immunol.* **2016**, 34, 265–297. https://doi.org/10.1146/annurev-immunol-041015-055420.

8. Hunt D.F., Henderson R.A., Shabanowitz J., Sakaguchi K., Michel H., Sevilir N., Cox A.L., Appella E., Engelhard V.H. Characterization of peptides bound to the class I MHC molecule HLA-A2.1 by mass spectrometry. *Science*. **1992**;255:1261–1263. doi: 10.1126/science.1546328.

9. Engelhard V.H., Appella E., Shabanowitz J., Sevilir N., Slingluff C.L., Hunt D.F., Benjamin D.C., Bodnar W.M., Cox A.L., Chen Y., et al. Mass Spectrometric Analysis of Peptides Associated with the Human Class I MHC Molecules HLA-A2.1 and HLA-B7 and Identification of Structural Features that Determine Binding. *Chem. Immunol. Allergy*. **1993**;57:39–62. doi: 10.1159/000319180.

10. Kowalewski D.J., Stevanović S. Biochemical Large-Scale Identification of MHC Class I Ligands. *Methods Mol. Biol.* **2012**;960:145–157. doi: 10.1007/978-1-62703-218-6_12.

11. Ghosh M., Gauger M., Marcu A., Nelde A., Denk M., Schuster H., Rammensee H.-G., Stevanović S. Guidance Document: Validation of a High-Performance Liquid Chromatography-Tandem Mass Spectrometry Immunopeptidomics Assay for the Identification of HLA Class I Ligands Suitable for Pharmaceutical Therapies. *Mol. Cell. Proteom*. **2020**;19:432–443. doi: 10.1074/mcp.C119.001652.

12. Bian, Y.; Zheng, R.; Bayer, F. P.; Wong, C.; Chang, Y.-C.; Meng, C.; Zolg, D. P.; Reinecke, M.; Zecha, J.; Wiechmann, S.; Heinzlmeir, S.; Scherr, J.; Hemmer, B.;

Baynham, M.; Gingras, A.-C.; Boychenko, O.; Kuster, B. Robust, Reproducible and Quantitative Analysis of Thousands of Proteomes by Micro-Flow LC-MS/MS. *Nat. Commun.* **2020**, 11 (1), 157. https://doi.org/10.1038/s41467-019-13973-x.

13. De Hoffmann, E.; Stroobant, V. *Mass Spectrometry: Principles and Applications*, 3rd ed.; Wiley-Blackwell: Hoboken, NJ, **2013**.

14. Olsen, J. V.; Macek, B.; Lange, O.; Makarov, A.; Horning, S.; Mann, M. Higher-Energy C-Trap Dissociation for Peptide Modification Analysis. *Nat. Methods* **2007**, 4 (9), 709–712. https://doi.org/10.1038/nmeth1060.

15. Ho, C. S.; Lam, C. W. K.; Chan, M. H. M.; Cheung, R. C. K.; Law, L. K.; Lit, L. C. W.; Ng, K. F.; Suen, M. W. M.; Tai, H. L. Electrospray Ionisation Mass Spectrometry: Principles and Clinical Applications. *Clin. Biochem. Rev*. **2003**, 24 (1), 3–12.

16. Fenn, J. B.; Mann, M.; Meng, C. K.; Wong, S. F.; Whitehouse, C. M. Electrospray Ionization for Mass Spectrometry of Large Biomolecules. *Science* **1989**, 246 (4926), 64–71. https://doi.org/10.1126/science.2675315.

17. Cech, N. B.; Enke, C. G. Practical Implications of Some Recent Studies in Electrospray Ionization Fundamentals. *Mass Spectrom. Rev.* **2001**, 20 (6), 362–387. https://doi.org/10.1002/mas.10008.

18. Konermann, L.; Ahadi, E.; Rodriguez, A. D.; Vahidi, S. Unraveling the Mechanism of Electrospray Ionization. *Anal. Chem*. **2013**, 85 (1), 2–9. https://doi.org/10.1021/ac302789c.

19. Macias, L. A.; Santos, I. C.; Brodbelt, J. S. Ion Activation Methods for Peptides and Proteins. *Anal. Chem.* **2020**, 92 (1), 227–251. https://doi.org/10.1021/acs.analchem.9b04859.

20. Earley, L.; Anderson, L. C.; Bai, D. L.; Mullen, C.; Syka, J. E. P.; English, A. M.; Dunyach, J.-J.; Stafford, G. C., Jr; Shabanowitz, J.; Hunt, D. F.; Compton, P. D. Front-End Electron Transfer Dissociation: A New Ionization Source. *Anal. Chem*. **2013**, 85 (17), 8385–8390. https://doi.org/10.1021/ac401783f.

21. Haag, A. M. Mass Analyzers and Mass Spectrometers. *Adv. Exp. Med. Biol*. **2016**, 919, 157–169. https://doi.org/10.1007/978-3-319-41448-5_7.

22. Schwartz, J. C.; Senko, M. W.; Syka, J. E. P. A Two-Dimensional Quadrupole Ion Trap Mass Spectrometer. *J. Am. Soc. Mass Spectrom.* **2002**, 13 (6), 659–669. https://doi.org/10.1016/S1044-0305(02)00384-7.

23. Savaryn, J. P.; Toby, T. K.; Kelleher, N. L. A Researcher's Guide to Mass Spectrometry-Based Proteomics. *Proteomics* **2016**, 16 (18), 2435–2443. https://doi.org/10.1002/pmic.201600113.

24. Hu, Q.; Noll, R. J.; Li, H.; Makarov, A.; Hardman, M.; Graham Cooks, R. The Orbitrap: A New Mass Spectrometer. *J. Mass Spectrom*. **2005**, 40 (4), 430–443. https://doi.org/10.1002/jms.856.

25. Bern M.; Kil YJ.; Becker C. Byonic: advanced peptide and protein identification software. *Curr Protoc Bioinformatics*. **2012**; Chapter 13:13.20.1-13.20.14. doi:10.1002/0471250953.bi1320s40

26. Andreatta M, Alvarez B, Nielsen M. GibbsCluster: unsupervised clustering and alignment of peptide sequences. *Nucleic Acids Res*. **2017**;45(W1):W458-W463. doi:10.1093/nar/gkx248

27. Reynisson B, Alvarez B, Paul S, Peters B, Nielsen M. NetMHCpan-4.1 and NetMHCIIpan-4.0: improved predictions of MHC antigen presentation by concurrent motif deconvolution and integration of MS MHC eluted ligand data. *Nucleic Acids Res*. **2020**;48(W1):W449-W454. doi:10.1093/nar/gkaa379

28. Reynisson, B.; Barra, C.; Kaabinejadian, S.; Hildebrand, W. H.; Peters, B.; Nielsen, M. Improved Prediction of MHC II Antigen Presentation through Integration and Motif Deconvolution of Mass Spectrometry MHC Eluted Ligand Data. *J. Proteome Res*. **2020**, 19 (6), 2304–2315. https://doi.org/10.1021/acs.jproteome.9b00874.

29. Alvarez B, Reynisson B, Barra C, et al. NNAlign_MA; MHC Peptidome Deconvolution for Accurate MHC Binding Motif Characterization and Improved T-cell Epitope Predictions. Mol Cell Proteomics. 2019;18(12):2459-2477. doi:10.1074/mcp.TIR119.001658

Chapter 2: Characterization of MHC Class I-Associated Peptides in a Mouse Cell Line

2.1 Introduction

2.1.1 Protein Post Translational Modification (PTM)

PTMs play a crucial role in regulating protein function by activating or deactivating enzymes, signaling pathways, and controlling cellular processes such as apoptosis, immune responses, and gene expression. Through PTMs, cells can rapidly adjust protein activity in response to environmental changes. PTMs are changes made to proteins after they have been synthesized by ribosomes. These modifications, such as phosphorylation, acetylation, glycosylation, acetylation, ubiquitination, and methylation, can result in peptides that might be recognized as foreign by the immune system, potentially leading to autoimmune diseases. [1,2]

PTMs can significantly influence how peptides are presented by Major Histocompatibility Complex (MHC) molecules to T-cells by increased or decreased binding affinity of peptides to the MHC molecules, thereby impacting immune recognition and response. [3]

Evaluating MHC-peptide binding can be a demanding process in terms of both time and resources. Competitive binding assays are commonly used to estimate binding affinity

by determining the concentration (IC50) at which a peptide of interest displaces 50% of a radiolabeled probe peptide from MHC class I molecules. [4,5,6]

2.1.2 RMA-S Stabilization Assay

In addition to competitive binding assays, which are frequently used to assess MHC-peptide binding affinity, thermal stability assays offer another valuable approach for evaluating MHC-peptide interactions. The RMA-S stabilization assay is an example of thermal stability assay used to evaluate the binding of peptides to MHC class I molecules on the surface of cells. RMA-S cells, mutant derivatives of the RMA T-lymphoma cell line in mice, lack the transporter associated with antigen processing (TAP). This defect prevents peptide transport into the ER for MHC class I loading, resulting in low surface expression of stable MHC class I molecules at 37°C. These unique characteristics make RMA-S cells valuable for studying peptide-MHC interactions.

However, when exogenous peptides are introduced into RMA-S cells, they can bind directly to the MHC class I molecules, stabilizing them and allowing their expression on the cell surface, especially when the cells are exposed to lower temperatures.

RMA-S cells are incubated with a test peptide at a lower temperature, typically 26°C. This lower temperature stabilizes the MHC class I molecules in a peptide-receptive conformation but does not result in stable surface expression. The test peptide is allowed to bind to MHC class I molecules during this incubation period.

After peptide incubation, the cells are shifted to 37°C, which induces the MHC molecules to become stable if they are bound by a peptide. Without peptide binding, the MHC molecules will not be stabilized and will be degraded.

The amount of MHC class I molecules on the surface of the RMA-S cells is measured using flow cytometry with antibodies specific to the MHC class I molecule (such as anti-H-2Kb or anti-H-2Db in mouse models).

The increase in surface expression of MHC molecules correlates with the ability of the test peptide to bind and stabilize the MHC class I molecules. If a peptide binds effectively to MHC class I, it will stabilize the MHC molecule at 37°C, resulting in higher surface expression. This is detected as an increase in fluorescence signal during flow cytometry. If the test peptide does not bind or bind weakly to MHC class I, the MHC molecules will not be stabilized, and the surface expression will remain low. [7]

2.1.3 Peptide SIINFEKL (ovaWT)

SIINFEKL (ovaWT) is commonly used to study post-translational modifications (PTMs) because it is a well-characterized peptide derived from ovalbumin, which is widely recognized for its ability to bind to MHC class I molecules, specifically H-2Kb in mice.[8] This peptide is used extensively in immunological research due to it having well-studied peptide epitopes[9,10], elicits robust T-cell responses[11,12], and serves as a baseline to study how various PTMs affect its binding affinity to MHC class I molecules[16].

These advantages make ovaWT an ideal peptide for studying how PTMs influence antigen presentation and immune recognition, as it provides a clear, measurable outcome through T-cell activation assays.

2.1.4 Mild Acid Elution (MAE)

In addition to MHC-peptide complex enrichment by immunoprecipitation method described in the introduction chapter, mild acid elution is an alternative method used to gently dissociate peptides bound to MHC molecules on the surface of antigen-presenting cells (APCs) without significantly damaging the cells or MHC molecules. This technique is commonly employed in immunology to study the peptide repertoire presented by MHC molecules, especially in the context of antigen processing and immune recognition. [13]

Cells expressing MHC molecules on their surface, such as dendritic cells, macrophages, or B cells, are prepared. These cells present peptides derived from degraded proteins, bound to MHC class I or II molecules.

The cells are exposed to a mild acidic solution, typically with pH 3-4. This pH is low enough to disrupt the weak interactions between the MHC molecule and the bound peptide, causing the peptides to dissociate from the MHC.

Importantly, the acid used is mild to prevent denaturation of the MHC molecules and preserve their integrity. The MHC molecules remain intact on the cell surface but are now peptide-free.

The acid-washed solution containing the eluted peptides is collected for further analysis. These peptides represent the natural repertoire of antigens that were bound to the MHC molecules before elution.

The cells themselves can survive mild acid treatment and can even be loaded with new peptides after the elution. [13,14]

2.1.5 Challenges and Project Goal

This project aims to analyze how post-translational modifications (PTMs) affect the presentation of tumor neoantigens—novel peptides from tumor-specific mutations—by introducing chemical modifications to synthetic peptides and studying their interactions with MHC molecules and TCRs.

The RMA-S stabilization assay can indicate that a peptide is associated with an MHC molecule, but it cannot determine which peptide is bound. Therefore, LC-MS analysis of the peptides eluted from the MHC molecules is necessary to confirm that the peptide on the MHC molecules is indeed the one loaded onto the MHC molecules to affect stabilization.

This project also aims to compare two MHC-associated peptide purification methods: immunoprecipitation and mild acid elution.

2.2 Materials

AnaSpec (Fremont, CA)

• Angiotensin II Phosphate >95% HPLC (Angio)

• Vasoactive Intestinal Peptide (VIP) 1-12 human, porcine, rat >95% HPLC (Vaso)

Synthesized in-house by Joey Kelly

• SIINFEKL(Ova)

• SIINFEK*L (trimethylated K, OvaK7m3)

Thermo Scientific (Waltham, MA)

• Pierce™ Water, LC-MS grade

Sigma Aldrich (St Louis, MO)

• Acetic acid, glacial, ≥99.99% trace metals basis

Fisher Chemical (Ward Hill, NJ)

• Methanol, Optima™ LC/MS grade

Honeywell (Charlotte, NC)

• Acetonitrile, B&J Brand™ LC-MS, for LC-MS and HPLC, >99.9%

PQ Corporation (Malvern, PA)

• Kasil 1624 potassium silicate solution

PolyMicro Technologies, Inc. (Phoenix, AZ)

• 100 µm i.d. x 360 µm o.d. Polyimide coated fused silica nano-capillary tubing

• 75 µm i.d. x 360 µm o.d. Polyimide coated fused silica nano-capillary tubing

Dr Maisch GMBH (Ammerbuch, Germany)

• Reprosil Pur 120 C18 AQ 3 µm

• Reprosil Pur 120 C18 AQ 10 µm

Eppendorf (Enfield, CT/Hamburg, Germany)

• Model 5424 Centrifuge

• Protein LoBind Microcentrifuge tubes (1.5 mL)

Zeus Industrial Products, Inc. (Orangeburg, SC)

• Teflon tubing, 0.012-inch i.d. x 0.060-inch o.d.

Sutter Instrument Co. (Novato, CA)

• P-2000 microcapillary laser puller with fused silica adapter

Agilent Technologies (Palo Alto, CA)

• 1100 Agilent high performance liquid chromatograph (HPLC)

Thermo-Fisher Scientific (San Jose, CA/Bremen, Germany)

• LTQ-Orbitrap Classic mass spectrometer

2.3 Methods

2.3.1 Class I MHC-associated Peptide Purification: Immunoprecipitation

Cell lysis and MHC class I peptide purification from RMA-S mouse cell line samples. 1e9 RMA-S cells are used as the starting material.

Immunoprecipitation of class I MHC-peptide complexes was accomplished with the mouse antibodies (**Table 3-1**) conjugated to NHS-Sepharose beads. For every 1E9 RMA-S cells, 3 mg of antimouse H-2Kb antibody was conjugated to 300 µL of beads.

For bead activation, all buffers were stored at 4°C or kept on ice during the procedure. The beads were washed twice with 1 mL of 1 mM HCl, followed by three washes with 1 mL of PBS. The antibody was added to the beads at a concentration of approximately 1 mg per gram of tissue. The mixture was rotated at 4°C for 3 hours to overnight. After the incubation, the supernatant was removed. 1 mL of 100 mM Tris was added to the beads and the mixture was rotated at 4°C for 1 hour, then the supernatant was removed, and the step was repeated once more.

Next, the beads were rinsed three times with 1 mL of acetate buffer, followed by three rinses with 1 mL of 20 mM Tris. These two rinsing steps were repeated two more times. Afterward, the beads were rinsed once with 20 mM Tris and 150 mM NaCl. Finally, a volume of 20 mM Tris and 150 mM NaCl equal to the volume of the beads was added. The activated beads were stored at 4°C for up to two months.

For the lysis step, the lysis buffer was prepared on the same day as the immunoprecipitation (IP) to prevent detergent degradation. All samples and solutions were kept ice-cold throughout the procedure. Inhibitors, 100 U/mL of Benzonase, and 5 mM $MgCl_2$ were added to the lysis buffer. Approximately 1 mL of buffer was added for every 100 mg of tissue. The sample was sonicate on ice for 5 cycles of 3-second pulses with 5-second waiting periods. The sample was rotated at 4°C for 15 minutes, then the sonication and rotation steps were repeated. The sample was centrifuged at 4°C at 15,000–20,000 rcf for 20 minutes. The supernatant was transferred to a new LoBind microcentrifuge tube. If the supernatant is cloudy, it was filtered through a 0.2 µm syringe filter. For larger samples, 5 mL LoBind tubes or 15 mL conical tubes were used, and they were filled no more than two-thirds full.

For purification, IgG antibody beads was added to the supernatant and rotated at 4°C for 2–6 hours, followed with centrifugation at 5,000 rcf for two minutes and then the supernatant was transferred to a new LoBind tube. Next, conjugated antibody beads was added to the supernatant and rotated for 2 hours or overnight at 4°C. The sample was centrifuged again at 5,000 rcf for two minutes and the supernatant was transferred to a new tube. This process was repeated as needed with additional antibodies. The final supernatant was stored at 4°C or freeze until needed.

For rinsing the beads, a rinse solution was added, the sample was vortexed to resuspend the beads, followed with centrifugation at 5,000 rcf for two minutes to pellet, and the supernatant was discarded. Three rinses were performed with 2 mL of PBS

containing 1% OG, followed with three rinses with 2 mL of high salt buffer, and three rinses with 2 mL of low salt buffer. On the final rinse, the beads were transferred to a new, labeled LoBind tube, followed with three rinses using 2 mL of no salt buffer and two rinses using 1 mL of H_2O.

For elution, 100 µL of 10% acetic acid (HOAc) was added to the beads and incubates on a shaker for five minutes. The sample was centrifuged and the supernatant was transferred to a new, labeled LoBind tube. The elution process was repeated twice, all the supernatant was collected into the same tube. Finally, the supernatant was dried using a speedvac and the dried sample was sent on dry ice for further analysis.

2.3.2 Class I MHC-associated Peptide Purification: Mild Acid Elution

RMA-S cells (1e9 cells) were incubated with 20 µM peptide at 26 °C for 1 hour, followed by a 6-hour incubation at 37 °C. The cells were collected by centrifugation at 500 x g for 5 minutes, then thoroughly washed with 1X PBS buffer (pH 7.4, Invitrogen). Afterward, the cells were incubated with 15 mL of mild acid elution (MAE) buffer (0.131 M citric acid, 0.066 M Na2HPO4, 150 mM NaCl, 0.3 µM aprotinin, and 5 mM iodoacetamide, pH 3.3) for 2 minutes. The cells were centrifuged at 4000 x g for 5 minutes, and the supernatant was collected, lyophilized, and stored at −80 °C for later use.

2.3.3 Sample Desalting and Contaminants Removal

PHEA clean-up protocol is based on the original protocol from Keira Mahoney.[15]

To begin, a concentrator was used to dry down the sample. A 250 µL pipette tip was prepared by plugging it with glass filter membranes (Whatman GF/F). A syringe was used to push the filter into the tip. Approximately 100 mg of PHEA (20 µm, 300Å) was added to a tube, followed by 250 µL of 100 mM ammonium formate. The packing slurry was mixed by pipetting up and down several times, then 50 µL of the slurry was added to the pipette tip. The tip was spun until the solution was almost gone and then continued by adding 50 µL increments of slurry until a packed bed of 5 mm to 1 cm was achieved.

The packed PHEA resin was rinsed by adding 100 µL of 100 mM ammonium formate at pH 3, then followed with 100 µL of H_2O, making sure that less than 1 mm of liquid is left at the resin-air surface. Proceeded with three rinses using 100 µL of 10 mM ammonium formate and 90% acetonitrile. The rinsing steps were repeated when working with real samples.

Next, the dried sample was reconstituted with 50 µL of 10 mM ammonium formate and 90% acetonitrile (for larger samples, reconstituted in 150 µL and loaded 50 µL at a time). 0.5 µL of phosphopeptide standards were added to each 50 µL sample. The sample was spun at approximately 100 rcf and was loaded it into the pipette tip, and the eluate was collected in a “Sample Load/Rinse tube.” After loading, the tip was rinsed once with 50 µL of 10 mM ammonium formate and 90% acetonitrile, the flow-through of the rinse was collected in the same “Sample Load/Rinse tube.”

For peptide samples, the tip was rinsed 2–3 times using 20 µL of 50% acetonitrile and 0.1% acetic acid, the eluate was collected in a “Peptide Elution tube.” For protein

samples, the tip was rinsed 2–3 times with 20 µL of 0.2% formic acid, the eluate was collected in a “Protein Elution tube.”

2.3.4 Liquid Chromatography-Mass Spectrometric (LC-MS) Analysis

The samples were analyzed using in-house prepared analytical columns and precolumns as described previously [1]. The emitter tip of the analytical column was laser-pulled to produce an opening of 2-5 µm, and a 2 mm Kasil frit was used in place of the irregular reverse phase (RP) resin. The precolumns (100 µm i.d. x 360 µm o.d. fused silica) were packed to 7 cm with 10 µm C18 beads, and the analytical columns (75 µm i.d. x 360 µm o.d. fused silica) were packed to 10 cm with 3 µm C18 beads.

Ova peptide samples and 100 fmol each of the internal standards (Angio and Vaso) were loaded onto the precolumn using a pressure vessel for a 15-minute desalting rinse at a flow rate of 100 nL/min with 0.1% acetic acid (AcOH) in water before connecting to the analytical column with a 2 cm Teflon tubing.

Reverse phase separation was conducted at a flow rate of 100 nL/min by HPLC using 0.1% AcOH for solvent A and 0.1% AcOH in 60% acetonitrile for solvent B, and a gradient as follows: 0% to 60% solvent B in 60 minutes, 60% to 100% solvent B in 2 minutes, 100% solvent B for 4 minutes, 0% in 2 minutes, followed by a 22-minute equilibration with 100% solvent A.

The peptides eluted from the analytical column were electrosprayed into an LTQ-Orbitrap Classic mass spectrometer. MS1 spectra were acquired with a resolution of

60,000, an AGC target of 5e5 and scan range of 300 to 2000 m/z in the Orbitrap analyzer, followed by low resolution data-dependent MS2 acquisition in the ion trap with a normal scan rate. Only precursor ions with charge +2 and +3 were selected for fragmentation. The top 2 most abundant precursor ions, as well as the targeted +1 and +2 precursor masses for the OvaWT and OvaK7m3 peptides, were selected for collision-induced dissociation (CID) in the ion trap analyzer with an AGC target of 1e4, a normalized collision energy of 35%, an activation time of 30 ms, and a 2.0 m/z isolation window. If the same precursor ion was selected three times or was detected twice within a 20-second repeat duration, the ion was dynamically excluded for 15 seconds.

The presence of OvaWT and OvaK7m3 peptides was verified by manually inspecting the targeted MS2 spectra for the expected fragment ion masses.

2.4 Results and Discussion

2.4.1 Validation of the RMA-S Stabilization Assay

Through the use of LC-MS, I successfully confirmed that the peptides associated with the MHC molecules were indeed the synthetic peptides that were introduced to the cells. This process involved several steps to ensure the correct identification and validation of the peptides.

Figure 2-1 shows the chromatographic profile of the LC-MS run for mild acid elution (MAE) of OvaWT, where the retention time and peak intensity provide an initial indication of the peptide's presence. To further validate this, extracted ion chromatograms (EICs)

corresponding to the known mass of the synthetic OvaWT peptide were generated. These EICs help isolate and visualize the specific ion corresponding to OvaWT, confirming its presence in the sample at the expected retention time.

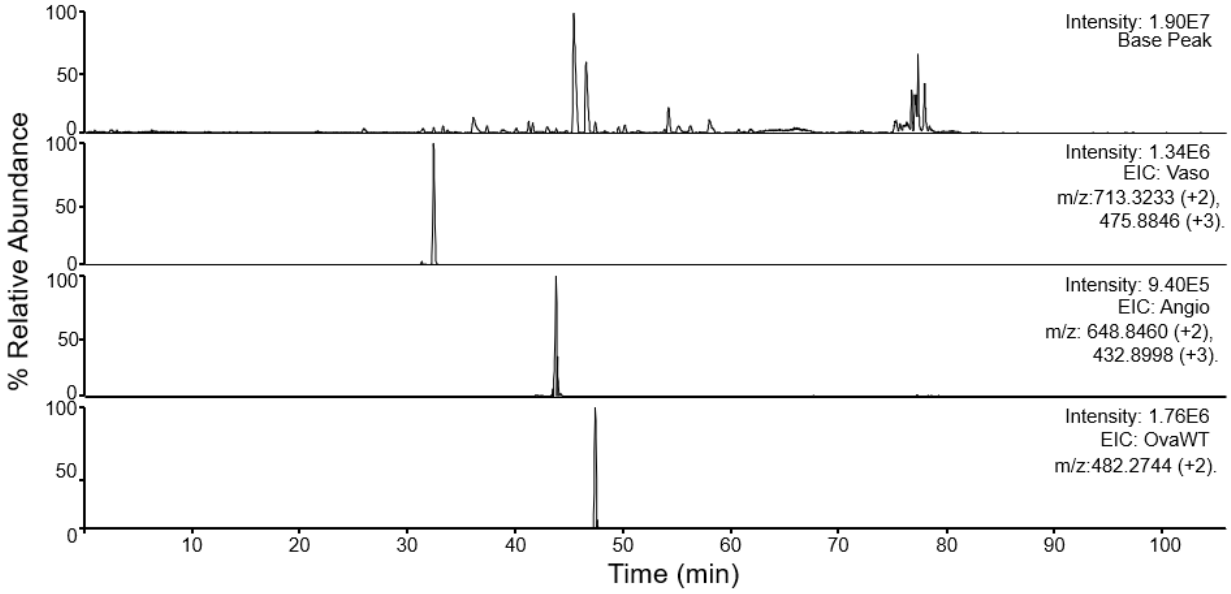


Figure 2- 1: *LC-MS analysis of OvaWT and spiked internal standard peptides (MAE).*

Extracted ion chromatograms (EIC) are shown for the base peak and selected ion m/z corresponding to the OvaWT, 100 fmol each of Vaso and Angio peptides.

Subsequently, the MS2 spectrum of the precursor ion corresponding to OvaWT, as shown in **Figure 2-2**, was used to confirm that the chromatographic peak corresponds to the correct peptide sequence. By analyzing the fragmentation pattern in the MS2 spectrum, key fragment ions characteristic of the OvaWT peptide were identified, further validating the identification of this peptide. The fragmentation pattern helps to confidently assign the peak to OvaWT, ensuring that the signal observed in the chromatography corresponds to the correct synthetic peptide.

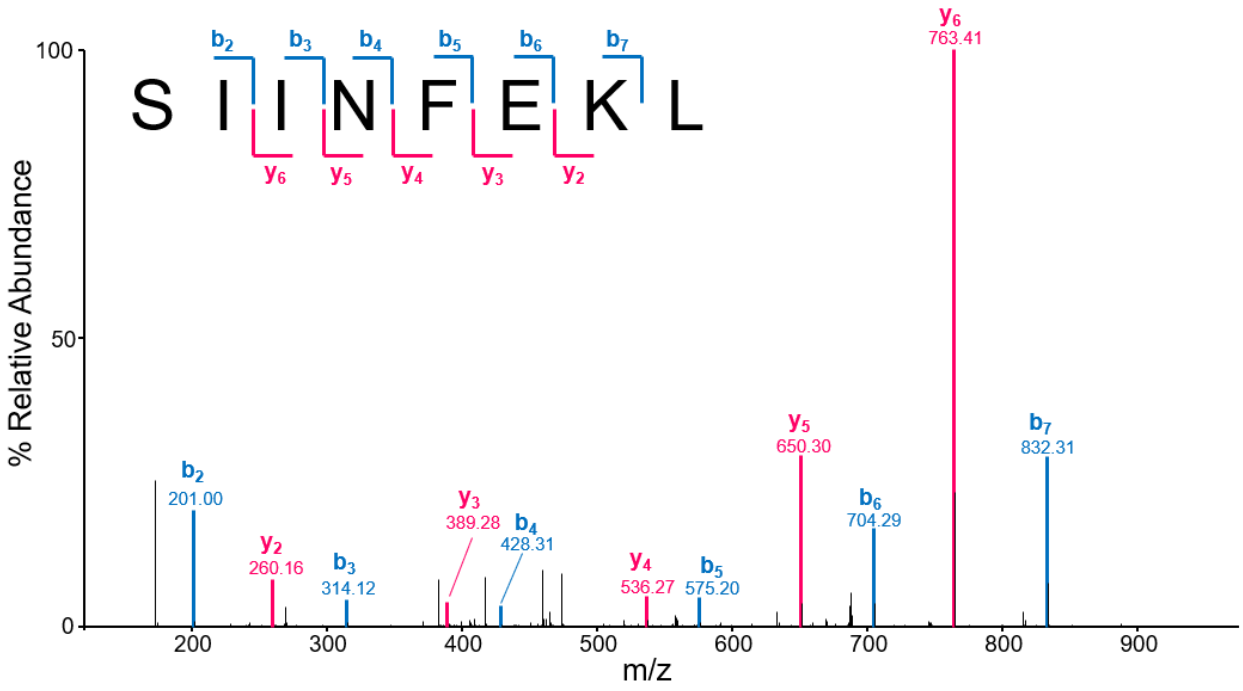


Figure 2- 2: *CID tandem mass spectrometry (MS/MS) spectrum acquired in the ion trap showing fragment masses of OvaWT [M+2H]2+ precursor ion with m/z 482.2744 (MAE).*

Sequence coverage for b+ ion fragments are labeled in blue, and y+ ion fragments are labeled in pink. The identified ions are sufficient for complete sequence coverage.

Next, the Ova peptide with a trimethylated lysine (OvaK7m3) modification was analyzed under similar conditions to confirm its binding to MHC molecules. The chromatography for this modified peptide is shown in **Figure 2-3**, and like with OvaWT, an MS2 spectrum of the precursor ion (**Figure 2-4**) was generated to validate the identification. The presence of diagnostic ions corresponding to the trimethylation modification in the MS2 spectrum

helped confirm that the modified peptide had successfully bound to the MHC molecules.

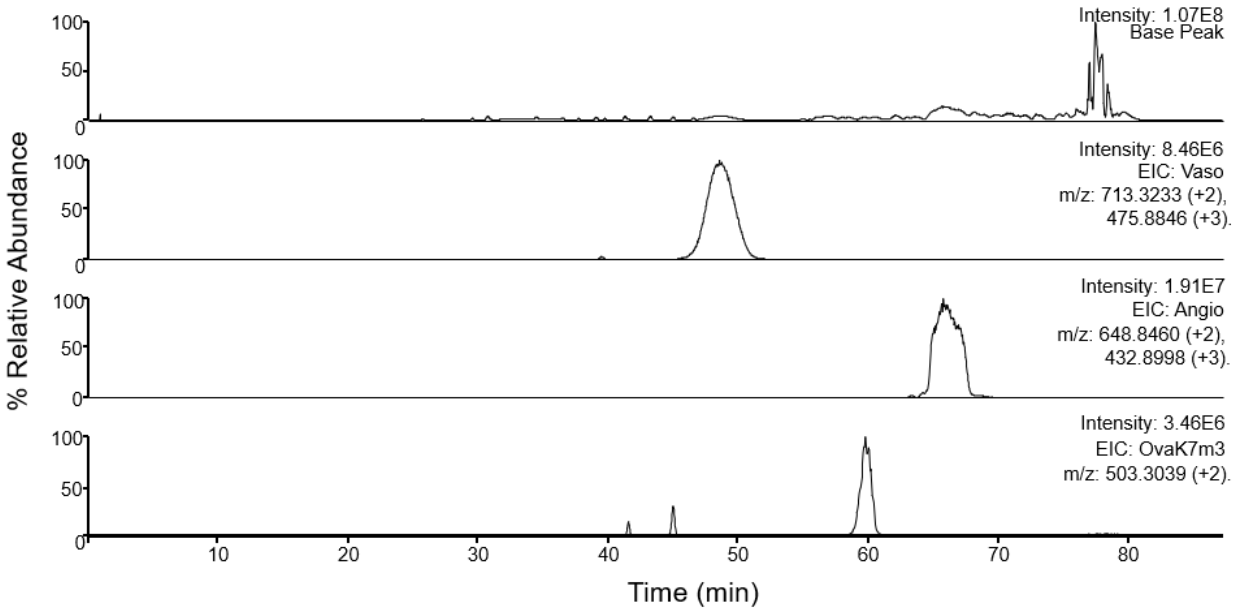


Figure 2- 3: *LC-MS analysis of OvaK7m3 and spiked internal standard peptides.*

Extracted ion chromatograms (EIC) are shown for the base peak and selected ion m/z corresponding to the OvaK7m3, 100 fmol each of Vaso and Angio peptides.

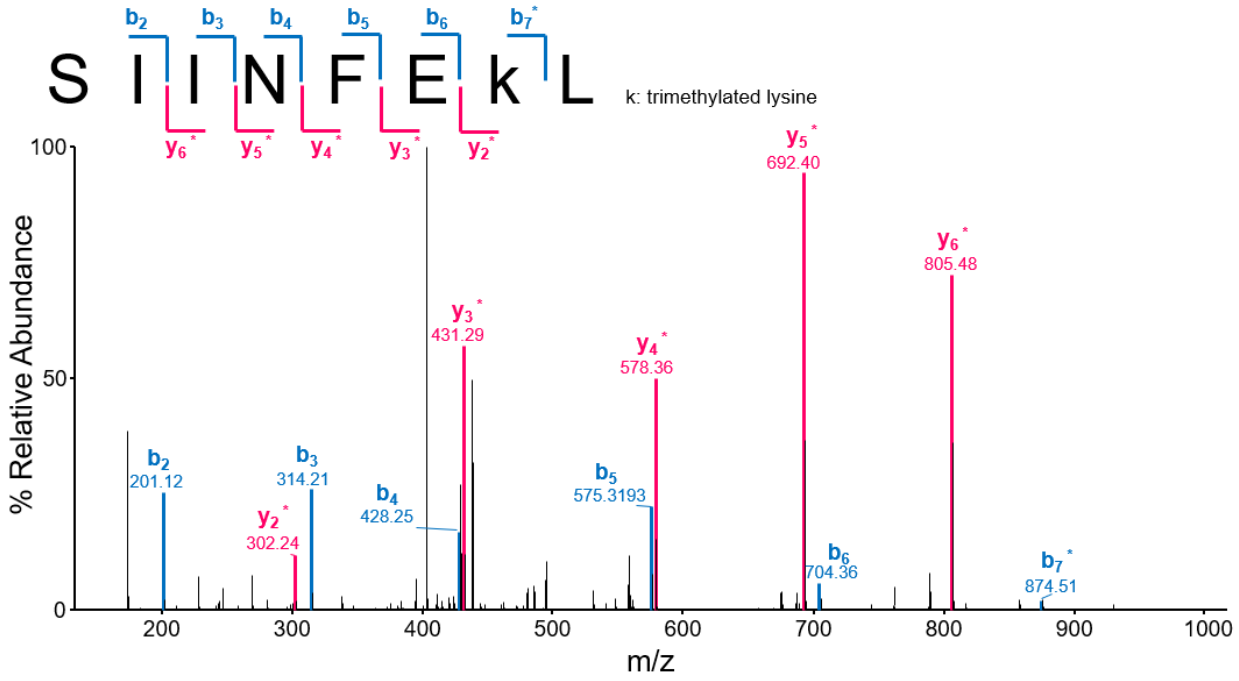


Figure 2- 4*: CID tandem mass spectrometry (MS/MS) spectrum acquired in the ion trap showing fragment masses of OvaK7m3 [M+2H]2+ precursor ion with m/z 503.3039 (MAE).*

Sequence coverage for b+ ion fragments are labeled in blue, and y+ ion fragments are labeled in pink. The fragments containing trimethylated lysine are marked by *. The identified ions are sufficient for complete sequence coverage.

This comprehensive analytical approach, combining both chromatographic and MS2 data, was critical for ensuring that the correct synthetic peptide—OvaWT or the trimethylated lysine variant—was presented on the MHC molecules. This analytical approach was crucial in verifying that the expected peptides were presented on the MHC, ensuring the accuracy of the experimental outcomes and validating the RMA-S stabilization assay. By confirming the peptide identity and MHC binding, this approach strengthened the conclusions drawn from the experiment, ensuring that the expected peptides were accurately detected and analyzed.

2.4.2 Comparison Between the Immunoprecipitation and Mild Acid Elution

In this research, the peptides I aimed to detect were those that had been deliberately introduced into the cells. Both immunoprecipitation (**Figure 2-5, 2-6**) and mild acid elution (**Figure 2-1, 2-2**) were effective in identifying these peptides. While mild acid elution has the drawback of introducing a high level of peptide and non-peptide contaminants[16], these contaminants might be peptides released from surrounding apoptotic cells. This limitation was not a significant issue in our case. Importantly, mild acid elution does not require the use of anti-MHC antibodies, which substantially reduces both the cost and time associated with the experiment. By addressing the contamination

challenge inherent to mild acid elution, I was able to maximize its advantages and make it a valuable method in this study.

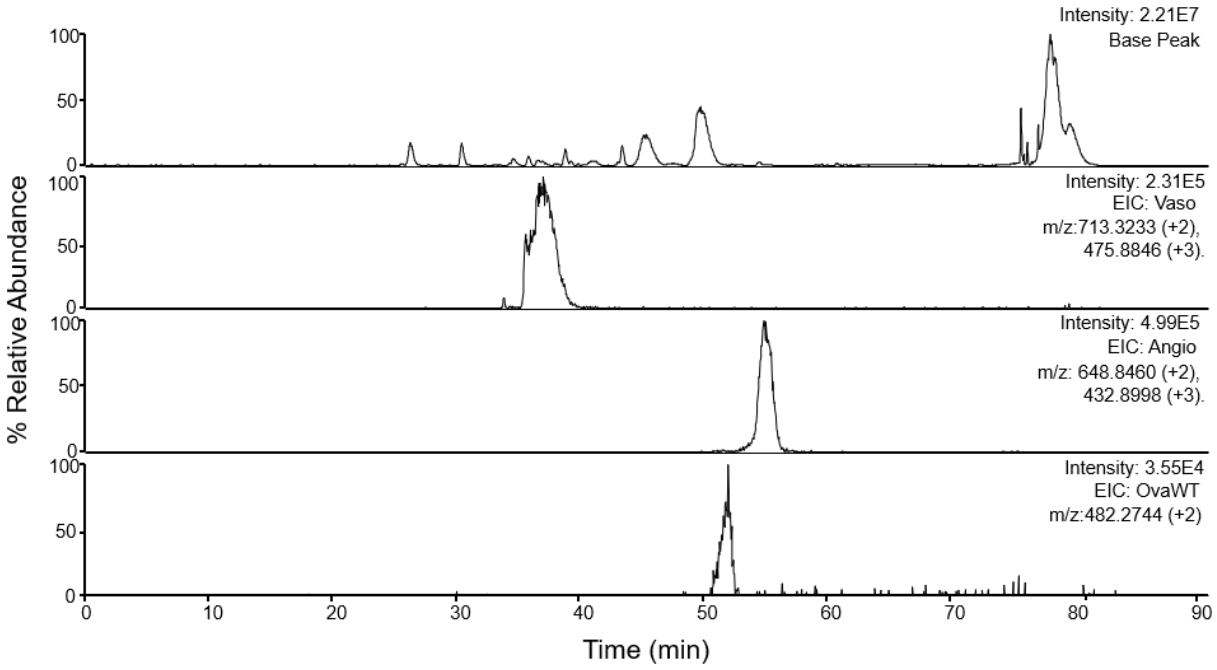


Figure 2- 5: *LC-MS analysis of OvaWT and spiked internal standard peptides (IP).*

Extracted ion chromatograms (EIC) are shown for the base peak and selected ion m/z corresponding to the OvaWT, 100 fmol each of Vaso and Angio peptides.

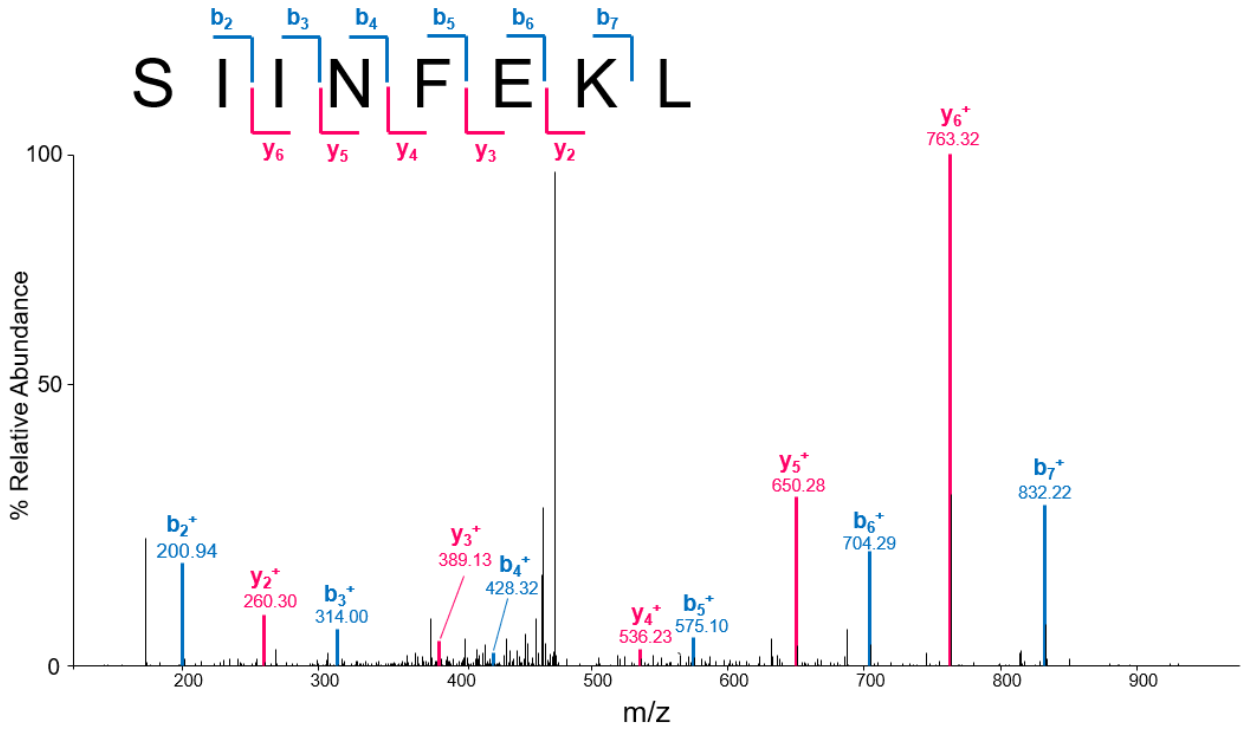


Figure 2- 6: *CID tandem mass spectrometry (MS/MS) spectrum acquired in the ion trap showing fragment masses of OvaWT [M+2H]2+ precursor ion with m/z 482.2744 (IP).*

Sequence coverage for b+ ion fragments are labeled in blue, and y+ ion fragments are labeled in pink. The identified ions are sufficient for complete sequence coverage.

2.5 Future Directions

Through the combination of LC-MS and the RMA-S stabilization assay, it was demonstrated that PTMs can alter the MHC binding affinity of peptides. [17]

One class of compounds that influences PTMs is histone deacetylase (HDAC) inhibitors, which are known to induce cancer cell death through multiple mechanisms. Previous experiments have shown that HDAC inhibitors increases lysine-acetylated peptide in the cell, therefore it is highly likely that HDAC inhibitors can increase the presentation of lysine-acetylated peptides by class I MHC molecules, which could

subsequently trigger a T-cell immune response against cancer cells. This project aims to investigate the effect of post-translational modifications (PTMs), such as lysine acetylation, on MHC binding affinity and T-cell recognition of the presented peptides. The results of this project are relevant to understanding how PTMs influence immune responses and can help guide further research into cancer immunotherapy, where modulating immune recognition is crucial, or into autoimmune diseases, where altered antigen presentation may contribute to pathogenesis.

The role of PTMs in modulating immune recognition is particularly significant in cancer, where changes in antigen presentation can impact immune surveillance and response. This connection between PTMs and immune function is further highlighted by the actions of histone deacetylases (HDACs) and their inhibitors.

HDACs play a critical role in regulating gene expression by removing acetyl groups from histones, leading to a more condensed chromatin structure and repression of gene transcription. In contrast, HDAC inhibitors prevent this deacetylation, thereby promoting a more relaxed chromatin state that allows the activation of tumor suppressor genes, such as p21 and p53. In addition to directly activating tumor suppressor pathways, HDAC inhibitors also inhibit angiogenesis—the formation of new blood vessels—which is essential for tumor growth and metastasis. Another important effect of HDAC inhibitors is their ability to enhance the immune response by increasing the expression of class I MHC molecules on the surface of cancer cells. This upregulation may lead to improved recognition and destruction of cancer cells by cytotoxic T-cells. [18,19,20]

Given these mechanisms, the hypothesis that HDAC inhibitors increase the presentation of lysine-acetylated peptides by class I MHC molecules is of significant interest. This could provide a novel link between epigenetic modifications and immune recognition, offering new insight into how HDAC inhibitors may be harnessed to enhance anti-tumor immunity. [20]

2.6 References

1. Zhong, Q.; Xiao, X.; Qiu, Y.; Xu, Z.; Chen, C.; Chong, B.; Zhao, X.; Hai, S.; Li, S.; An, Z.; Dai, L. Protein Posttranslational Modifications in Health and Diseases: Functions, Regulatory Mechanisms, and Therapeutic Implications. *MedComm* **2023**, 4 (3), e261. https://doi.org/10.1002/mco2.261.

2. Bagwan, N.; El Ali, H. H.; Lundby, A. Proteome-Wide Profiling and Mapping of Post Translational Modifications in Human Hearts. *Sci. Rep*. **2021**, 11 (1), 2184. https://doi.org/10.1038/s41598-021-81986-y.

3. Lee, J. M.; Hammarén, H. M.; Savitski, M. M.; Baek, S. H. Control of Protein Stability by Post-Translational Modifications. *Nat. Commun*. **2023**, 14 (1), 201. https://doi.org/10.1038/s41467-023-35795-8.

4. Peters, B.; Bui, H.-H.; Sidney, J.; Weng, Z.; Loffredo, J. T.; Watkins, D. I.; Mothé, B. R.; Sette, A. A Computational Resource for the Prediction of Peptide Binding to Indian Rhesus Macaque MHC Class I Molecules. *Vaccine* **2005**, 23 (45), 5212–5224. https://doi.org/10.1016/j.vaccine.2005.07.086.

5. Sette, A.; Sidney, J.; Bui, H.-H.; del Guercio, M.-F.; Alexander, J.; Loffredo, J.; Watkins, D. I.; Mothé, B. R. Characterization of the Peptide-Binding Specificity of Mamu-A*11 Results in the Identification of SIV-Derived Epitopes and Interspecies Cross-Reactivity. *Immunogenetics* **2005**, 57 (1–2), 53–68. https://doi.org/10.1007/s00251-004-0749-z.

6. Sette, A.; Vitiello, A.; Reherman, B.; Fowler, P.; Nayersina, R.; Kast, W. M.; Melief, C. J.; Oseroff, C.; Yuan, L.; Ruppert, J.; Sidney, J.; del Guercio, M. F.; Southwood, S.; Kubo, R. T.; Chesnut, R. W.; Grey, H. M.; Chisari, F. V. The Relationship between Class I Binding Affinity and Immunogenicity of Potential Cytotoxic T Cell Epitopes. *J. Immunol*. **1994**, 153 (12), 5586–5592. https://doi.org/10.4049/jimmunol.153.12.5586.

7. Ross, P.; Holmes, J. C.; Gojanovich, G. S.; Hess, P. R. A Cell-Based MHC Stabilization Assay for the Detection of Peptide Binding to the Canine Classical Class I Molecule, DLA-88. *Vet. Immunol. Immunopathol*. **2012**, 150 (3–4), 206–212. https://doi.org/10.1016/j.vetimm.2012.08.012.

8. Dersh, D.; Yewdell, J. W.; Wei, J. A SIINFEKL-Based System to Measure MHC Class I Antigen Presentation Efficiency and Kinetics. *Methods Mol. Biol.* **2019**, 1988, 109–122. https://doi.org/10.1007/978-1-4939-9450-2_9.

9. Beck, L.; Spiegelberg, H. L. The Polyclonal and Antigen-Specific IgE and IgG Subclass Response of Mice Injected with Ovalbumin in Alum or Complete Freund's Adjuvant. *Cell. Immunol*. **1989**, 123 (1), 1–8. https://doi.org/10.1016/0008-8749(89)90263-3.

10. Ke, Y.; Li, Y.; Kapp, J. A. Ovalbumin Injected with Complete Freund's Adjuvant Stimulates Cytolytic Responses. Eur. *J. Immunol*. **1995**, 25 (2), 549–553. https://doi.org/10.1002/eji.1830250237.

11. Rötzschke, O.; Falk, K.; Stevanović, S.; Jung, G.; Walden, P.; Rammensee, H. G. Exact Prediction of a Natural T Cell Epitope. Eur. J. Immunol. 1991, 21 (11), 2891–2894. https://doi.org/10.1002/eji.1830211136.

12. Dudani, R.; Chapdelaine, Y. Multiple Mechanisms Compensate To Enhance Tumor-Protective Cd8+ T Cell Response in the Long-Term Despite Poor Cd8+ T Cell Priming Initially: Comparison Between an Acute Versus a Chronic Intracellular Bacterium Expressing a Model Antigen. *J. Immunol*. **2002**, 168, 5737–5745.

13. Sturm, T.; Sautter, B.; Wörner, T. P.; Stevanović, S.; Rammensee, H.-G.; Planz, O.; Heck, A. J. R.; Aebersold, R. Mild Acid Elution and MHC Immunoaffinity Chromatography Reveal Similar Albeit Not Identical Profiles of the HLA Class I Immunopeptidome. *J. Proteome Res*. **2021**, 20 (1), 289–304. https://doi.org/10.1021/acs.jproteome.0c00386.

14. Sugawara, S.; Abo, T.; Kumagai, K. A Simple Method to Eliminate the Antigenicity of Surface Class I MHC Molecules from the Membrane of Viable Cells by Acid Treatment at pH 3. *J. Immunol. Methods* **1987**, 100 (1–2), 83–90. https://doi.org/10.1016/0022-1759(87)90175-x.

15. Mahoney, Keira. (**2022**). Improved methods for identification of MHC Class I-associated peptides [Doctoral dissertation, Chemistry - Graduate School of Arts and Sciences, University of Virginia].

16. Caron, E.; Kowalewski, D. J.; Chiek Koh, C.; Sturm, T.; Schuster, H.; Aebersold, R. Analysis of Major Histocompatibility Complex (MHC) Immunopeptidomes Using Mass Spectrometry. *Mol. Cell. Proteomics* **2015**, 14 (12), 3105–3117. https://doi.org/10.1074/mcp.m115.052431.

17. Kelly, J. J.; Bloodworth, N.; Shao, Q.; Shabanowitz, J.; Hunt, D.; Meiler, J.; Pires, M. M. A Chemical Approach to Assess the Impact of Post-Translational Modification on MHC Peptide Binding and Effector Cell Engagement. *ACS Chem. Biol.* **2024**, 19 (9), 1991–2001. https://doi.org/10.1021/acschembio.4c00312.

18. Ho, T. C. S.; Chan, A. H. Y.; Ganesan, A. Thirty Years of HDAC Inhibitors: 2020 Insight and Hindsight. *J. Med. Chem*. **2020**, 63 (21), 12460–12484. https://doi.org/10.1021/acs.jmedchem.0c00830.

19. Slaughter, M. J.; Shanle, E. K.; Khan, A.; Chua, K. F.; Hong, T.; Boxer, L. D.; Allis, C. D.; Josefowicz, S. Z.; Garcia, B. A.; Rothbart, S. B.; Strahl, B. D.; Davis, I. J. HDAC Inhibition Results in Widespread Alteration of the Histone Acetylation Landscape and BRD4 Targeting to Gene Bodies. *Cell Rep*. **2021**, 34 (3), 108638. https://doi.org/10.1016/j.celrep.2020.108638.

20. Eckschlager, T.; Plch, J.; Stiborova, M.; Hrabeta, J. Histone Deacetylase Inhibitors as Anticancer Drugs. *Int. J. Mol. Sci*. **2017**, 18 (7), 1414. https://doi.org/10.3390/ijms18071414.

Chapter 3: Characterization of Hamster MHC for the Study of COVID

3.1 Introduction

3.1.1 COVID-19

COVID-19, caused by the SARS-CoV-2 virus, is a respiratory illness that emerged in late 2019 and led to a global pandemic. The virus spreads primarily through respiratory droplets and close contact. Severity ranges from mild to severe, with common symptoms including fever, cough, and fatigue, while severe cases can result in difficulty breathing, organ failure, or death. COVID-19 has prompted widespread public health measures, including vaccinations, testing, mask-wearing, and social distancing, to control its spread. Various vaccines have been developed to help prevent severe illness and reduce transmission. [1,2]

3.1.2 Hamster as a COVID Model

Animal models are essential for studying COVID-19 due to their ability to mimic various aspects of the disease in humans, such as understanding disease progression, evaluating therapeutics and vaccines, studying immune responses, modeling severe cases and long COVID, and studying variants and viral evolution.

SARS-CoV-2 infects cells by binding to the human ACE2 (hACE2) receptor, but it is unable to infect mice because it cannot bind to the mouse ACE2 (mACE2) receptor. [3] Syrian hamsters are naturally susceptible to SARS-CoV-2 without genetic modification,[4] providing a more straightforward and physiologically relevant infection model.

Previous studies have used humanized hACE2 mice, meaning that the mice are genetically modifying to express a human version of a specific gene.[5] Humanized mice often exhibit acute infections, but the ability to study long-term COVID-19 effects (e.g., long COVID, which is known as post-acute sequelae of SARS-CoV-2 infection (PASC), refers to a range of symptoms that persist for weeks or months after the initial acute phase of a COVID-19 infection has resolved.) is limited. Hamsters, particularly Syrian hamsters, have a more human-like immune response to SARS-CoV-2.

Hamsters develop a human-like disease course, including respiratory symptoms, lung pathology, and immune responses. They also develop mild to severe disease with clear clinical signs, significant weight loss, viral shedding, lung pathology, and a measurable immune response. Unlike ferrets and non-human primates, which usually exhibit mild disease, the hamster model provides a wider range of measurable outcomes, offering strong discriminatory power to evaluate the efficacy of vaccines, drugs, and therapeutics against SARS-CoV-2, making them a better model for studying the immune system aspects of infection. [6,7,8,9]

3.1.3 Murine MHC Polymorphism

When studying MHC molecules, polymorphism plays an important role in determining the diversity of peptides that can be presented to T-cells, thereby influencing the strength and breadth of immune responses.

In contrast, hamsters and Golden Syrian hamsters in particular have been found to exhibit limited variability in their MHC genes and might have low degree of polymorphism for MHC class I. [10,11,12,13] The low polymorphism in hamsters may be due to historical genetic bottlenecks and inbreeding. For example, many laboratory strains of Syrian hamsters originate from a small population, which likely reduced genetic diversity, including in their MHC loci. [12] On the other hand, considerable polymorphism is observed in hamster class II MHC molecules. [14]

3.1.4 Significance and Challenges

3.1.4.1 Study of MHC Binding Motif in General

The study of MHC motifs and the ability to characterize MHC-associated peptides are crucial, not only for building animal models, as discussed in the previous section, but also for drug development. In the context of drug development, MHC Class I molecules provide insights into potential drug toxicity, while Class II molecules help predict whether anti-drug antibodies might be produced, which could ultimately reduce the drug's efficacy.

As discussed in previous chapters, antibodies specific to MHC molecules are required to perform immunoprecipitation for capturing these molecules. However, the

inherent polymorphism of MHC molecules complicates antibody production. One might assume that several rounds of affinity maturation would increase both the specificity and affinity of the antibody toward the MHC molecule. Yet, in practice, affinity maturation can result in antibodies losing affinity to some sub-alleles within the same MHC allele. For example, when generating an H2-IA mouse MHC-specific antibody, the initial antibody might cross-react with H2-IEk. Through affinity maturation, one might aim to enhance specificity for H2-IA, but this process could inadvertently make the antibody too specific to particular H2-IA alleles, such as H2-IA b, f, and p, while losing affinity for others, like H2-IA s, u, and v.

Once a potential antibody is generated, flow cytometry is often employed to assess its binding to MHC molecules. However, this technique does not always confirm whether the antibody is binding specifically to the MHC molecule or to another component, such as the Class II-associated invariant chain peptide (CLIP) peptide. The CLIP is a segment of the invariant chain that occupies the peptide-binding groove of MHC Class II molecules during their assembly. It temporarily blocks the groove to prevent premature binding of endogenous peptides, playing a crucial role in stabilizing the MHC Class II molecule. Despite CLIP's important role in stabilizing MHC molecules, using antibodies to study MHC interactions carries the risk that the antibody may bind to the CLIP peptide rather than the MHC molecule itself. As a result, an antibody that initially appears to have promising specificity could fail in real-world applications, where the CLIP peptide is eventually replaced by exogenous peptides meant for antigen presentation. This could render the antibody ineffective when it comes to recognizing the actual MHC-peptide complexes.

Therefore, to successfully elucidate the MHC motif in hamsters, we cannot rely solely on immunologists to produce the perfect antibodies, because they may not exist. It falls on mass spectrometrists to determine the MHC motif despite the limitations posed by antibodies.

In this study, I will explain how I overcame the challenge of immunological limitations and successfully characterized hamster MHC-associated peptides.

3.1.4.2 Study of Hamster MHC

The MHC peptide binding motif of the golden Syrian hamster remains largely understudied. This project aims to characterize hamster class I and II MHC-associated peptides to identify the binding motif, which is essential for the understanding of immune system responses to SARS-CoV-2 infection.

Previous studies have experienced challenges of characterizing hamster MHC molecules. Earlier studies in hamsters have indicated irregularities in the structure, expression, and/or function of the hamster MHC and its associated peptides. [15]

There are no hamster MHC-specific antibodies available, therefore mouse antibodies (**Table 3-1**) were used for these studies to purify hamster MHC-peptide complexes. Given the predicted low polymorphism of hamster class I MHC molecules,[10,11,12,13] I anticipate that some class I MHC-specific mouse antibodies may fail to capture hamster class I MHC molecules, as the corresponding mouse MHC alleles may be absent in hamsters.

Table 3- 1: *Antibodies and Their Specific MHC Alleles*.

Mouse Antibody	MHC Class	Mouse MHC Allele
34-5-8s	I	H2-Dd
B8-24-3	I	H2-B
2G5	I	H2-Kb, H2-Db
14-4-4s	II	H2-IEk
Y3JP	II	H2-IA b,f,p,q,r,s,u,v

Given the considerable polymorphism observed in hamster class II MHC molecules,[14] I anticipate that some class II MHC-specific mouse antibodies may fail to capture any MHC molecules, as the corresponding alleles in hamsters differ from those in mice.

3.2 Materials

Aldrich Chemical Co. (Milwaukee, WI)

• Ethylenediaminetetraacetic acid (EDTA), 99.99% purity

Bio-Rad Labs (Hercules, CA)

• Tris base >99.8% pure

Fisher Chemical (Ward Hill, NJ)

• Methanol, Optima™ LC/MS grade

Honeywell (Charlotte, NC)

• Acetonitrile, B&J Brand™ LC-MS, for LC-MS and HPLC, >99.9%

Oxford Glyco Systems (Oxfordshire, England)

• Tris HCl, Electrophoresis grade

PQ Corporation (Malvern, PA)

• Kasil 1624 potassium silicate solution

Roche (Basel, Switzerland)

• cOmplete™, Mini Protease inhibitor tablet, EDTA-free

Sigma Aldrich (St Louis, MO)

• Acetic acid, glacial, ≥99.99% trace metals basis

• L-Ascorbic Acid, ACS reagent, ≥99%

• Azulene

• CHAPS, ≥98%

• Iron (III) chloride, anhydrous, powder, ≥99.99% trace metals basis

• Formamide, ≥99.5% (GC), BioReagent for molecular biology

• Hydrochloric Acid, ACS reagent, 37%

• Sodium Chloride, ACS reagent, ≥99.0%

• Sodium hydroxide, reagent grade, 97%, powder

• PhosSTOP™ phosphatase inhibitor tablets

• Triton X-100

Supelco (Bellefonte, PA)

Hydrogen chloride solution, 3M in methanol, for GC derivatization, LiChropur™

Thermo Scientific (Waltham, MA)

• Pierce™ Formic Acid, LC-MS grade

• Halt™ Protease and phosphatase inhibitor cocktail (100x)

• Pierce™ Water, LC-MS grade

Anaspec (Fremont, CA)

• Angiotensin II Phosphate >95% HPLC (Angio)

• Vasoactive Intestinal Peptide (VIP) 1-12 human, porcine, rat >95% HPLC (Vaso)

Sigma Aldrich (St Louis, MO)

• Angiotensin I human acetate salt hydrate >90% HPLC

Applied Biosystems TM (Carlsbad, CA)

• POROS® MC 20 metal chelating packing material, 20 µm diameter

Cytiva Life Sciences (Marlborough, MA)/GE Healthcare (Boston, MA)

• NHS-Activated Sepharose 4 Fast Flow

Dr Maisch GMBH (Ammerbuch, Germany)

• Reprosil Pur 120 C18 AQ 3 µm

• Reprosil Pur 120 C18 AQ 10 µm

QIAGEN (Hilden, Germany)

• Ni-NTA spin columns

2.2.4 Equipment and Instrumentation

Agilent Technologies (Palo Alto, CA)

• 1100 Agilent high performance liquid chromatograph

Boekel Scientific (Feasterville-Trevose, PA)

Digital Incubator 133000

Branson (Danbury, CT)

• Branson 1200 Ultrasonic Bath

Cadence Science (Cranston, RI)

• Blunt end pipetting needle (16G)

Eppendorf (Enfield, CT/Hamburg, Germany)

• Model 5424 Centrifuge

• Protein LoBind Microcentrifuge tubes (1.5 mL)

Fisherbrand (Pittsburgh, PA)

• Model 50 Sonic Dismembrator with CL-18 Probe

• Mini Tube Rotator

• Digital Vortex

• Vortex Genie 2

GL Sciences (Tokyo, Japan)

• Centrifuge adaptor 5010-21514

Labconco Corp. (Kansas City, MO)

• Centrivap centrifugal vacuum concentrator 7810016

Millipore (Billerica, MA)

• Amicon ultra, 10 kDa regenerated cellulose spin filter

• Amicon ultra, 3kDa regenerated cellulose spin filter

New Objective (Littleton, MA)

• PicoClear 360 Union

PolyMicro Technologies, Inc. (Phoenix, AZ)

• 360 µm o.d. x 75 µm i.d. polyimide coated fused silica capillary

• 360 µm o.d. x 100 µm i.d. polyimide coated fused silica capillary

• 360 µm o.d. x 150 µm i.d. polyimide coated fused silica capillary

Sarstedt Inc (Newton, NC)

• 1.5 mL screw top microtubes

Supelco (Bellefonte, PA)

• Empore octadecyl C18 extraction disks

• Screw top clear glass vials with green thermoset caps and PTFE liners, 2 mL

Sutter Instrument Co. (Novato, CA)

• P-2000 microcapillary laser puller with fused silica adapter

Thermo-Fisher Scientific (San Jose, CA/Bremen, Germany)

• Orbitrap Fusion Tribrid mass spectrometer with commercial front-end ETD

• LTQ mass spectrometer with commercial back-end ETD

• LTQ-Orbitrap Classic mass spectrometer (custom modified with front-end ETD)

Zeus Industrial Products, Inc. (Orangeburg, SC)

• Teflon tubing, 0.012-inch i.d. x 0.060-inch o.d.

3.3 Methods

3.3.1 Class I and II MHC-associated Peptide Purification

Cell lysis and MHC class I peptide purification from hamster cell line samples were performed by Paul Rubiro at the La Jolla Institute for Immunology. 1µg to 10µg of protein extracted from the B07-17-1113 hamster cell line is used as the starting material.

Immunoprecipitation of class I and class II MHC-peptide complexes was accomplished with the mouse antibodies (**Table 3-1**) conjugated to NHS-Sepharose beads. For each gram of tissue (or 1E9 cells), 3 mg of antibody were conjugated to 300 µL of beads.

For bead activation, all buffers were stored at 4°C or kept on ice during the procedure. The beads were washed twice with 1 mL of 1 mM HCl, followed by three washes with 1 mL of PBS. The antibody was added to the beads at a concentration of approximately

1 mg per gram of tissue. The mixture was rotated at 4°C for 3 hours to overnight. After the incubation, the supernatant was removed. 1 mL of 100 mM Tris was added to the beads and the mixture was rotated at 4°C for 1 hour, then the supernatant was removed, and the step was repeated once more.

Next, the beads were rinsed three times with 1 mL of acetate buffer, followed by three rinses with 1 mL of 20 mM Tris. These two rinsing steps were repeated two more times. Afterward, the beads were rinsed once with 20 mM Tris and 150 mM NaCl. Finally, a volume of 20 mM Tris and 150 mM NaCl equal to the volume of the beads was added. The activated beads were stored at 4°C for up to two months.

For the lysis step, the lysis buffer was prepared on the same day as the immunoprecipitation (IP) to prevent detergent degradation. All samples and solutions were kept ice-cold throughout the procedure. Inhibitors, 100 U/mL of Benzonase, and 5 mM $MgCl_2$ were added to the lysis buffer. Approximately 1 mL of buffer was added for every 100 mg of tissue. The sample was sonicate on ice for 5 cycles of 3-second pulses with 5-second waiting periods. The sample was rotated at 4°C for 15 minutes, then the sonication and rotation steps were repeated. The sample was centrifuged at 4°C at 15,000–20,000 rcf for 20 minutes. The supernatant was transferred to a new LoBind microcentrifuge tube. If the supernatant is cloudy, it was filtered through a 0.2 µm syringe filter. For larger samples, 5 mL LoBind tubes or 15 mL conical tubes were used, and they were filled no more than two-thirds full.

For purification, IgG antibody beads was added to the supernatant and rotated at 4°C for 2–6 hours, followed with centrifugation at 5,000 rcf for two minutes and then the supernatant was transferred to a new LoBind tube. Next, conjugated antibody beads was added to the supernatant and rotated for 2 hours or overnight at 4°C. The sample was centrifuged again at 5,000 rcf for two minutes and the supernatant was transferred to a new tube. This process was repeated as needed with additional antibodies. The final supernatant was stored at 4°C or freeze until needed.

For rinsing the beads, a rinse solution was added, the sample was vortexed to resuspend the beads, followed with centrifugation at 5,000 rcf for two minutes to pellet, and the supernatant was discarded. Three rinses were performed with 2 mL of PBS containing 1% OG, followed with three rinses with 2 mL of high salt buffer, and three rinses with 2 mL of low salt buffer. On the final rinse, the beads were transferred to a new, labeled LoBind tube, followed with three rinses using 2 mL of no salt buffer and two rinses using 1 mL of H_2O.

For elution, 100 µL of 10% acetic acid (HOAc) was added to the beads and incubates on a shaker for five minutes. The sample was centrifuged and the supernatant was transferred to a new, labeled LoBind tube. The elution process was repeated twice, all the supernatant was collected into the same tube. Finally, the supernatant was dried using a speedvac and the dried sample was sent on dry ice for further analysis.

3.3.2 Sample Desalting and Contaminants Removal

To begin, a concentrator was used to dry down the sample. A 250 µL pipette tip was prepared by plugging it with glass filter membranes (Whatman GF/F). A syringe was used to blow the filter into the tip. Approximately 100 mg of PHEA (20 µm, 300Å) was added to a tube, followed by 250 µL of 100 mM ammonium formate. The packing slurry was mixed by pipetting up and down several times, then 50 µL of the slurry was added to the pipette tip. The tip was spinned until the solution is almost gone and continue adding 50 µL increments of slurry until a packed bed of 5 mm to 1 cm was achieved.

The packed PHEA resin was rinsed by adding 100 µL of 100 mM ammonium formate at pH 3, then followed with 100 µL of H_2O, making sure that less than 1 mm of liquid is left at the resin-air surface. Proceeded with three rinses using 100 µL of 10 mM ammonium formate and 90% acetonitrile. The rinsing steps were repeated when working with real samples.

Next, the dried sample was reconstituted in 50 µL of 10 mM ammonium formate and 90% acetonitrile (for larger samples, reconstituted in 150 µL and load 50 µL at a time). 0.5 µL of phosphopeptide standards were added to each 50 µL sample. Spinned the sample at around 100 rcf and were loaded it into the pipette tip, eluate was collected in a “Sample Load/Rinse tube.” After loading, the tip was rinsed once with 50 µL of 10 mM ammonium formate and 90% acetonitrile, the flow-through of the rinse was collected in the same “Sample Load/Rinse tube.”

For peptide samples, the tip was rinsed 2–3 times using 20 µL of 50% acetonitrile and 0.1% acetic acid, the eluate was collected in a “Peptide Elution tube.” For protein samples, the tip was rinsed 2–3 times with 20 µL of 0.2% formic acid, the eluate was collected in a “Protein Elution tube.”

3.3.3 Liquid Chromatography-Mass Spectrometric (LC-MS) Analysis

Hamster MHC samples and 100 fmol each of the peptide internal standards (Angio and Vaso) were loaded onto the precolumn using a pressure vessel followed by a 15-minute column wash step at a flow rate of 100 nL/min with 0.1% acetic acid (AcOH) in water before connecting to the analytical column with 2 cm of Teflon tubing.

Reverse phase separation was accomplished at a flow rate of 100 nL/min using 0.1% AcOH for solvent A and 0.1% AcOH in 60% acetonitrile for solvent B, and a gradient as follows: 0% to 60% solvent B in 60 minutes, 60% to 100% solvent B in 2 minutes, 100% solvent B for 4 minutes, 0% in 2 minutes, followed by a 22-minute equilibration with 100% solvent A.

The peptides eluted from the analytical column were electrosprayed into an Orbitrap Fusion Tribrid mass spectrometer.

3.3.3.1 Class I-associated Peptides

For class I MHC peptides, MS1 spectra were acquired with a resolution of 60,000, RF lens of 60%, maximum injection time of 50 ms, an AGC target of 4e5 and scan range of 300 to 1500 m/z in the Orbitrap analyzer, an intensity filter of 3e4 is applied, followed by

MS2 acquisition. Dynamic exclusion of 7 seconds is applied to the same precursor ion when it was selected twice within 7-second repeat duration.

Precursor ions with charge +2 with a precursor mass range of 400-1000 m/z were selected for fragmentation to generate MS2 spectra first by CID with Orbitrap detection with resolution of 7500, an AGC target of 1e5, a collision energy of 30%, an activation time of 10 ms, maximum injection time of 250 ms, activation Q of 0.25, and a 1.6 m/z isolation window. Then the same precursor ions were selected for ETD with Orbitrap resolution of 15,000, AGC target of 2e4, maximum injection time of 50 ms, ETD reagent target of 2e5, maximum ETD injection time of 50 ms, ETD reaction time of 45 ms, and a 1.6 m/z isolation window.

Precursor ions with charge +3 to +4 with precursor mass range of 300-800 m/z were selected for fragmentation to generate MS2 spectra first by CID with Orbitrap resolution of 7500, an AGC target of 1e5, a collision energy of 30%, an activation time of 10 ms, maximum injection time of 250 ms, activation Q of 0.25, and a 1.6 m/z isolation window. Then the same precursor ions were selected for ETD with rapid ion trap scan rate, AGC target of 2e4, maximum injection time of 50 ms, ETD reagent target of 2e5, calibrated charge-dependent ETD parameters, and a 1.6 m/z isolation window.

Precursor ions with charge +1 with precursor mass range of 800-1500 m/z were selected for fragmentation to generate MS2 spectra by HCD with Orbitrap resolution of 7500, AGC target of 1e5, HCD collision energy of 30%, maximum injection time of 250 ms, and a 1.6 m/z isolation window.

3.3.3.2 Class II MHC-associated Peptides

For Class II MHC peptides, MS1 spectra were acquired with a resolution of 60,000, RF lens of 60%, maximum injection time of 50 ms, an AGC target of 4e5 and scan range of 300 to 1500 m/z in the Orbitrap analyzer, an intensity filter of 5e4 is applied, followed by data dependent MS2 acquisition. Dynamic exclusion of 6 seconds is applied to the same precursor ion when it was selected twice within 6-second repeat duration.

When precursor ions with charge +2 with precursor mass range of 400-950 m/z were selected for fragmentation to generate MS2 spectra, they were first selected for CID with orbitrap resolution of 15,000, an AGC target of 2e5, a collision energy of 30%, an activation time of 10 ms, maximum injection time of 120 ms, activation Q of 0.25, and a 2.0 m/z isolation window. Then the same precursor ions were selected for ETD with orbitrap resolution of 15,000, AGC target of 1e5, maximum injection time of 22 ms, ETD reagent target of 2e5, maximum ETD injection time of 50 ms, ETD reaction time of 45 ms, and a 2.0 m/z isolation window.

When precursor ions with charge +3 to +5 with precursor mass range of 300-800 m/z were selected for fragmentation to generate MS2 spectra, they were first selected for ETD with orbitrap resolution of 15,000, AGC target of 2e5, maximum injection time of 120 ms, ETD reagent target of 2e5, calibrated charge-dependent ETD parameters, and a 2.0 m/z isolation window. Then the same precursor ions were selected to generate MS2 spectra for HCD with orbitrap resolution of 15,000, AGC target of 1e5, HCD collision energy of 26%,

stepped collision energy with ± HCD collision energy of 3%, maximum injection time of 22 ms, and a 2.0 m/z isolation window.

3.3.4 Data Analysis

The mass spectra were searched against a combined hamster and mouse protein database using the Byonic software (Protein Metrics) with a 10 ppm precursor ion mass (MS1) tolerance and a 0.4 Da fragment ion mass (MS2) tolerance. The following amino acid modifications were used: fixed C-terminal, aspartic acid, and glutamic acid methylation, and common variable modifications of 1 methionine oxidation, and 1 cysteine cysteinylation. One asparagine or glutamine deamidation and 2 aspartic acid or glutamic acid dimethylations were allowed as rare modifications. A maximum of three common modifications and two rare modifications were allowed.

The search algorithm identifications were manually validated.

Byonic results were output in Excel, followed by an R program I wrote to convert the sequences into the desired format for GibbsCluster and NetMHCIIpan. Sequence logos were generated using GibbsCluster 2.0 for MHC Class I peptide samples, and binding affinity was predicted using NetMHCIIpan 4.1.

3.4 Result and Discussion

3.4.1 LC-MS Analysis of MHC Class I and Class II Associated Peptides

When analyzing MHC class I peptides, which are typically shorter in length (8–13 amino acids), they usually exhibit a charge state of +2 under standard mass spectrometry conditions. This lower charge state, combined with the shorter sequence, makes them well-suited for fragmentation by CID (**Figure 3-1**). CID is effective because it breaks peptide bonds in a predictable manner, producing enough fragment ions to achieve full sequence coverage, even for relatively short peptides like MHC class I epitopes. This allows for straightforward identification of the peptide sequence based on the generated fragment ions.

However, when analyzing MHC class II peptides, which are generally longer (10–25 amino acids with a median of 16), the situation becomes more complex. These peptides can carry higher charge states, often larger than +2, due to the increased number of basic residues capable of accepting protons. For example, MHC class II peptides can exhibit charge states such as +4, as shown in **Figure 3-2**. In such cases, CID often performs poorly because the energy required to induce fragmentation is distributed across the many protonation sites, leading to inefficient fragmentation and generating fewer, less informative fragment ions. This results in low sequence coverage and makes peptide identification more challenging.

On the other hand, electron transfer dissociation (ETD) is particularly advantageous for peptides with higher charge states, like the +4 charge state in MHC class II peptides (**Figure 3-3**). ETD works by transferring electrons to multiply charged peptides, which induces fragmentation at different bond sites compared to CID, primarily cleaving along the peptide backbone. This method preserves labile modifications and produces more extensive fragmentation patterns, even for longer, highly charged peptides. As a result, ETD provides much better sequence coverage for MHC class II peptides with charge states like +4, making it the preferred method for analyzing longer peptides where CID struggles to produce comprehensive data.

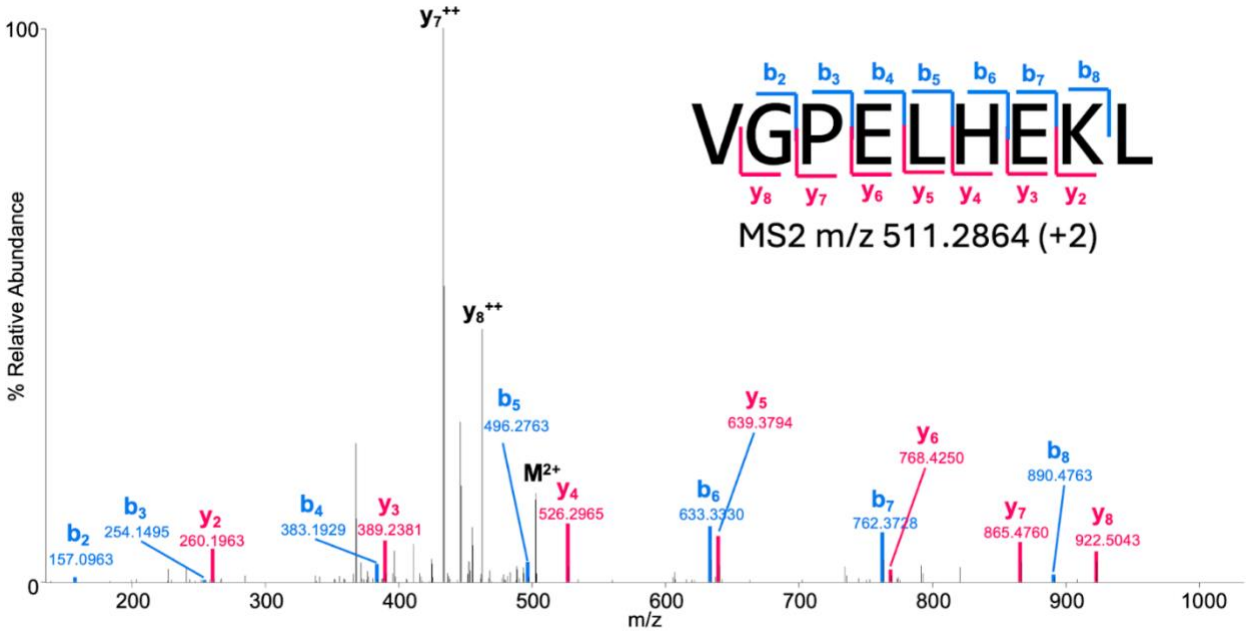


Figure 3- 1: *CID tandem mass spectrometry (MS/MS) spectrum showing fragment masses of [M+2H]2+ precursor ion with m/z 511.2864.*

Sequence coverage for b+ ion fragments are labeled in blue, and y+ ion fragments are labeled in pink. The identified ions are sufficient for complete sequence coverage.

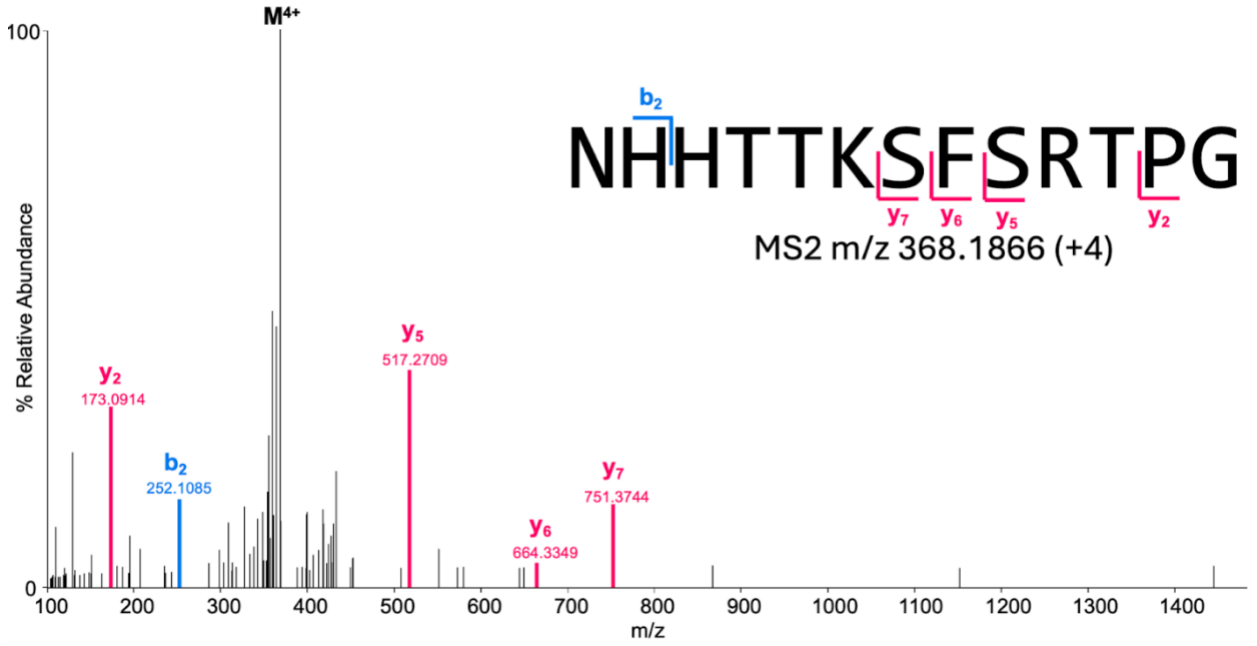


Figure 3- 2: *CID tandem mass spectrometry (MS/MS) spectrum acquired showing fragment masses of [M+4H]4+ precursor ion with m/z 368.1866.*

Sequence coverage for b+ ion fragments are labeled in blue, and y+ ion fragments are labeled in pink. The identified ions are sufficient for complete sequence coverage.

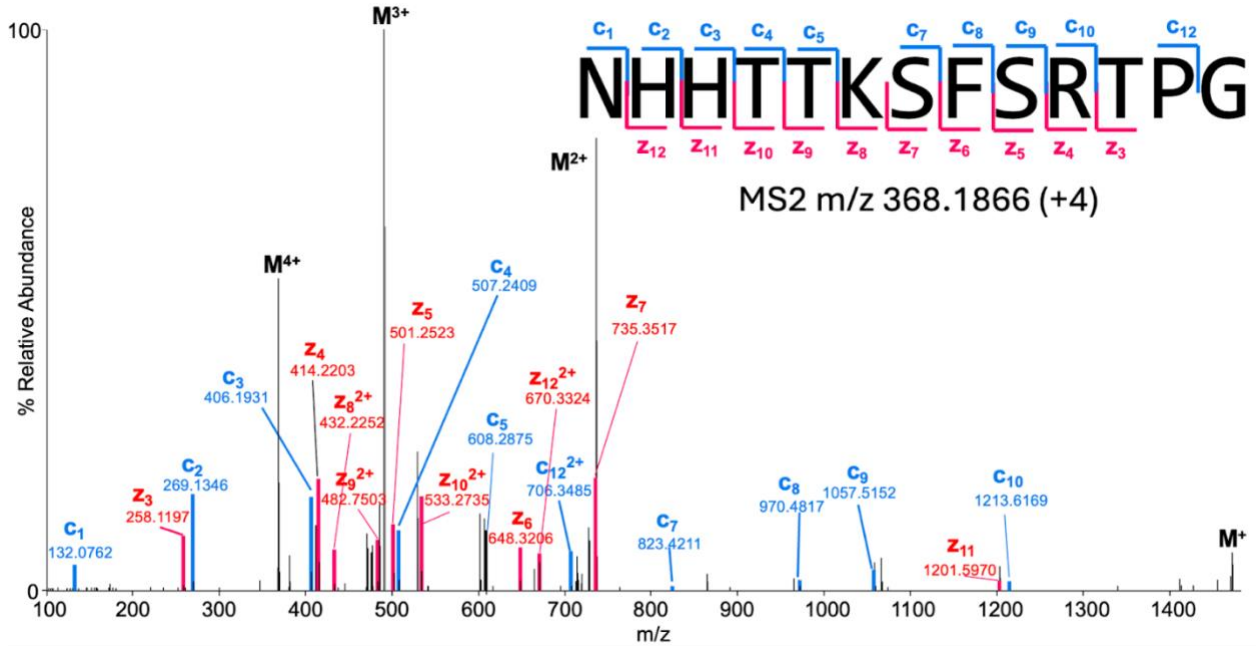


Figure 3- 3: *ETD tandem mass spectrometry (MS/MS) spectrum acquired showing fragment masses of [M+4H]4+ precursor ion with m/z 368.1866.*

Sequence coverage for b+ ion fragments are labeled in blue, and y+ ion fragments are labeled in pink. The identified ions are sufficient for complete sequence coverage.

3.4.2 Class I MHC-associated Peptides

After immunoprecipitation with the mouse antibody 34-5-8s, which specifically targets the mouse class I H2-Dd MHC allele, the motif identified from the hamster peptide sample (**Figure 3-4**) matches the published motif for the Class I H2-Dd mouse MHC allele (**Figure 3-5**). [16]

The other two mouse class I MHC antibodies failed to reveal any hamster MHC motifs, likely because the peptides are either not MHC-associated, or the sample contains a diverse mix of peptides with varying motifs, making it difficult to detect specific MHC-bound peptides. The mouse MHC alleles that are targets of these antibodies are H2-B, H2-Kb, and H2-Db, which indicates that these two mouse class I MHC alleles are not shared with the Golden Syrian hamster.

Knowing the peptide binding motif for Golden Syrian hamster (GSH) class I MHC molecules is significant for advancing immunological research in GSH models in infectious disease and cancer studies. Understanding the GSH class I MHC binding motif allows for better characterization of immune responses in this species, facilitating more precise experimentation, including vaccine development and pathogen-host interaction studies.

Knowing the MHC binding motif of GSH also helps in designing peptide-based vaccines or immunotherapies by predicting which viral, bacterial, or cancer peptides will be efficiently presented by GSH MHC molecules to T-cells. This is crucial for stimulating

effective cytotoxic T-cell responses in the hamster model, which can be translated into understanding immune responses in other species, including humans.

Having the GSH MHC class I binding motif allows for comparisons with the MHC binding motifs of other species, such as mice or humans. This knowledge can help researchers understand evolutionary differences in immune recognition and how different species respond to the same pathogens or immune challenges.

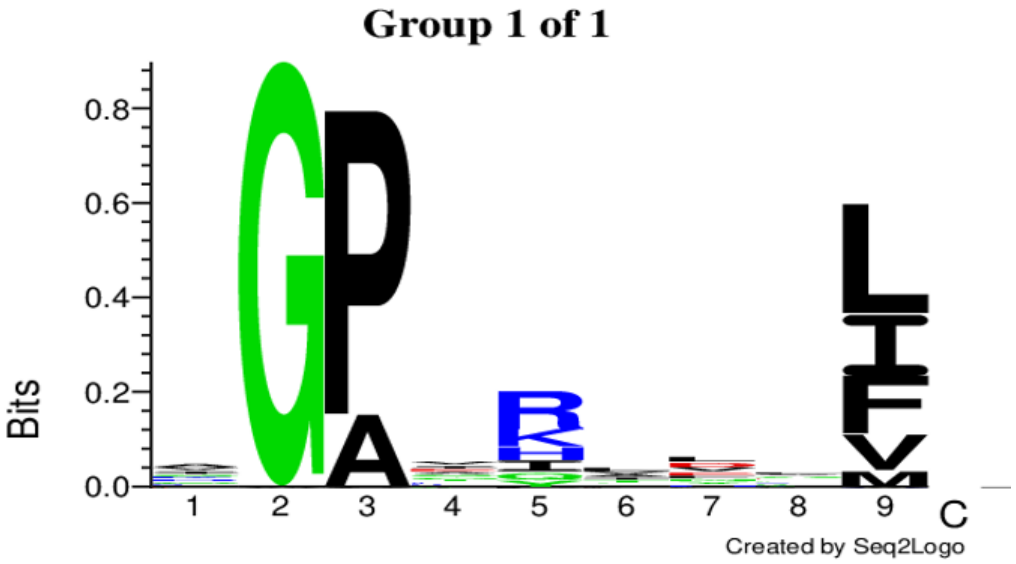


Figure 3- 4*: Motif of Hamster Class I MHC-bound Peptides.*

Mouse Antibody 34-5-8s was used.

Published motif

System	Allele	P2	P3	P5	P6	P9
H-2	Dd	G	P	RK		LFI

Figure 3- 5: *Published Class I H2-Dd Mouse MHC Allele Motif* [16]

3.4.3 Class II MHC-associated Peptides

The sequence alignment of the identified MHC class II peptides sorted by their source proteins is shown in **Figure 3-6**. This alignment revealed nested sets or common core sequences, which are consistent with the nature of MHC class II peptides (**Figure 1-1**, two open ends that allow for variability in peptide length). When proteins are degraded by proteases in the endosomal compartments, many series of peptides of varying lengths are produced. Portions of the peptide sequences overlap, revealing a "nested set" of shared peptide sub-sequences, from which the MHC peptide binding motif can be determined.

The “nested set” nature of the class II peptide sequences is due to the capacity of class II MHC molecules to bind peptides of variable lengths (10-25 residues with median of 16) in contrast to the relatively fixed and shorter lengths (8-13 residues) of class I MHC peptides (**Figure 1-1**). Proteins are cleaved at various positions by proteases, and MHC class II molecules typically bind a central 9-mer core sequence with terminal portions of the peptide extending over the open-ended MHC peptide binding cleft. Therefore, peptides surrounding this core may be of varying lengths but still capable of binding due to this conserved core sequence.

In addition, cathepsin S peptides were found in all GSH MHC class II samples. Cathepsin S is known to participate in the MHC class II peptide presentation pathway. [17] Therefore, it is likely a signature of the proteolytic cleavage processes involved in MHC class II peptide presentation, suggesting that the peptides isolated from these samples are MHC class II associated.

One limitation of using nested sets to predict binding core is that some peptides share more than 9 amino acids in common, while the core binding region of MHC class II molecules is limited to 9-10 amino acids (**Figure 1-1b**). This makes identifying the binding core challenging. To address this, I utilized the machine learning algorithm NetMHCpanIIpan 4.2, as described in Chapter 1. However, manually formatting and analyzing more than 2000 peptides (**Table 3-2**) is impractical. NetMHCpanIIpan 4.2 requires all peptides to be in FASTA format and listed one by one with their lengths, a process that would take weeks to complete manually. Additionally, processing such a large dataset often causes the algorithm to crash.

To overcome these challenges, I observed that NetMHCpanIIpan 4.2 operates by using a 9-mer sliding window to predict the binding affinity of peptides, moving sequentially through each peptide. However, since peptides in nested sets share identical sequences, the algorithm redundantly analyzes the same sequences multiple times, making it inefficient.

To optimize this process, I developed an R program that groups peptides from the same nested sets and compresses them into a single peptide that includes all amino acids. This allows NetMHCpanIIpan 4.2 to analyze each sequence only once, significantly reducing processing time and preventing crashes. My program automatically outputs the peptides in FASTA format, counts the amino acids, and organizes them in order. As a result, the weeks of manual analysis are reduced to less than ten seconds for large datasets.

NetMHCpanIIpan 4.2 predicted that some of the MHC class II-associated peptides characterized in this study have a strong binding affinity, with a very low % Rank Score (**Figure 3-6**).

Table 3- 2: *Total MHC-associated Peptides Characterized*.

Mouse Antibody	MHC class	# of peptides
34-5-8s	I	1592
B8-24-3	I	221
2G5	I	310
14-4-4s	II	1691
Y3JP	II	469

Complement C1q subcomponent subunit B (mouse & hamster)

H-2-IAq: LRPNQVIRF

%Rank Score: 4.43%

TINSPLRPNQVIRFE

TINSPLRPNQVIRFEK

Ig gamma-2A chain C region, A allele (mouse)

H-2-IAb: VRAPQVYVL, YVLPPPEEE

%Rank Score: 0.95%, 1.1%

H-2-IAk: VRAPQVYVL, QVYVLPPP

%Rank Score: 4.21%, 4.97%

H-2-IAq: VRAPQVYVL

%Rank Score: 1.82%

H-2-IAs: VRAPQVYVL, YVLPPPEEE, QVYVLPPPE

%Rank Score: 3.0%, 4.07%, 2.62%

H-2-IAu: VRAPQVYVL

%Rank Score: 2.04%

GSVRAPQVYVLPPPEEEMTKK

APQVYVLPPPEEEMT

APQVYVLPPPEEEMTK

APQVYVLPPPEEEMTKK

Fibrinogen beta (mouse)

H-2-IAb: YGGFTVQNE

%Rank Score: 3.9%

H-2-IAq: YGGFTVQNE

%Rank Score: 1.8%

H-2-IAk: NEASKYQVS

%Rank Score: 4.18%

H-2-IAs: YGGFTVQNE

%Rank Score: 0.73%

AHYGGFTVQNEASKYQVSVNK

AHYGGFTVQNEASKYQVSVNKY

Complement C1q subcomponent subunit A (mouse)

H-2-IAb: YQGTEADSI

%Rank Score: 0.67%

H-2-IAk: YQGTEADSI

%Rank Score: 0.9%

H-2-IAq: YQGTEADSI

%Rank Score: 1.78%

H-2-IAu: YQGTEADSI

%Rank Score: 0.95%

GRIYQGTEADSIFSG

GRIYQGTEADSIFSGF

IYQGTEADSIFSG

IYQGTEADSIFSGF

Ig heavy chain V region MOPC 21 (mouse)

H-2-IAk: SGGGLVQPG

%Rank Score: 1.92%

DVQLVESGGGLVQPGG

DVQLVESGGGLVQPGGSR

DVQLVESGGGLVQPGGSRK

Histone H2A type 1-G (mouse & hamster)

H-2-IAd: QGGVLPNIQ

%Rank Score: 2.93%

H-2-IAk: QGGVLPNIQ

%Rank Score: 1.91%

H-2-IAu: IAQGGVLPN

%Rank Score: 3.97%

VTIAQGGVLPNIQA

VTIAQGGVLPNIQ

IAQGGVLPNIQA

Figure 3- 6: *Sequence Alignment of Selected MHC class II Peptides with Predicted Strong Binding Affinity and the corresponding GSH proteins.*

The predicted binding core is in pink, an additional predicted binding core is underlined. Predicted Binding Affinity based on % Rank Score: Small % = high binding affinity. 1% Rank

Score means this peptide binds better than 99% of random peptide that binds to that specific MHC molecule.

3.5 Future Directions

Previous research suggests that Syrian hamsters express two monomorphic class I MHC molecules. [10] In this study, one of these molecules was found to share the same peptide binding motif as that of the mouse H2-Dd class I MHC molecule. The other GSH class I MHC could be an allele that differs from those in mice, and thus was not captured by the mouse MHC-specific antibodies tested here.

The next step will be to identify the class I and class II MHC alleles that are not common to or shared with those in mice. To achieve this, we first need to isolate and sequence the GSH MHC molecules. Because hamsters also express a homologue of beta-2-microglobulin (β2M),[10] an antibody specifically targeting hamster β2M should be sufficient to capture all GSH class I MHC molecules.

The hamster class II MHC molecules are considerably polymorphic.[14] This polymorphism results in a wide range of allele-specific MHC Class II variants, making it difficult to produce antibodies that can universally recognize all alleles or even target specific alleles.

One way to generate MHC Class II-specific antibodies is to immunize animals (e.g., mice or rabbits) with synthetic peptides that correspond to the polymorphic regions of the target MHC Class II molecule. However, there is a paradox associated with this approach:

antibodies are needed to isolate the target MHC Class II molecule to determine its polymorphic regions, creating a circular problem.

Therefore, plant-derived monoclonal antibodies might be a better alternative. The complementary DNAs for the light and heavy chains of a monoclonal antibody are synthesized by PCR amplification and introduced into tobacco plant cells. The plants then use their cellular machinery to produce the monoclonal antibodies. [18]

The forthcoming phase of this research will focus on evaluating the MHC binding affinity of the peptides characterized in this study. By performing detailed binding assays, we will quantify how strongly these peptides interact with the MHC molecules of interest, providing insights into their potential to elicit immune responses. This will involve the use of competition binding assays or thermal stability assays (detailed in Chapter 2) to further elucidate MHC-peptide interactions.

In addition to these experimental evaluations, another crucial future direction involves the development of an algorithm designed to predict potential epitopes from the spike protein of SARS-CoV-2 and other viruses. Leveraging the binding motifs identified in this study, the algorithm will predict which viral peptides are likely to be presented by MHC molecules, enabling the identification of immunogenic regions within viral proteins. This tool could greatly accelerate the design of vaccines and immunotherapies by pinpointing peptides with the highest likelihood of stimulating protective T-cell responses. Moreover, applying this algorithm to other viral pathogens could enhance preparedness for emerging infectious diseases by predicting potential epitopes for novel strains or future pandemics.

Taken together, these next steps will not only provide deeper mechanistic understanding of MHC-peptide binding but also create practical applications in the field of virology and immunotherapy, contributing to both basic science and translational research efforts.

3.6 References

1. Hekmatnia Y,; Rahmani F, Feili Z. Ebrahimzadeh F. A review of the effect of COVID-19 on immune responses of the body. *J Family Med Prim Care*. **2022**;11(5):1624-1632. doi:10.4103/jfmpc.jfmpc_677_21

2. Narayanan, S. A.; Jamison, D. A., Jr; Guarnieri, J. W.; Zaksas, V.; Topper, M.; Koutnik, A. P.; Park, J.; Clark, K. B.; Enguita, F. J.; Leitão, A. L.; Das, S.; Moraes-Vieira, P. M.; Galeano, D.; Mason, C. E.; Trovão, N. S.; Schwartz, R. E.; Schisler, J. C.; Coelho-Dos-Reis, J. G. A.; Wurtele, E. S.; Beheshti, A. A Comprehensive SARS-CoV-2 and COVID-19 Review, Part 2: Host Extracellular to Systemic Effects of SARS-CoV-2 Infection. Eur. *J. Hum. Genet*. **2024**, 32 (1), 10–20. https://doi.org/10.1038/s41431-023-01462-1.

3. Rawle DJ, Le TT, Dumenil T, et al. ACE2-lentiviral transduction enables mouse SARS-CoV-2 infection and mapping of receptor interactions. *PLoS Pathog*. **2021**;17(7):e1009723. Published 2021 Jul 2. doi:10.1371/journal.ppat.1009723

4. Mohl, BP.; Blaurock, C.; Riek, A. et al. Evidence for SARS-CoV-2 infected Golden Syrian hamsters (Mesocricetus auratus) reducing daily energy expenditure and body core temperature. *Sci Rep* **2024**, 14, 23263. https://doi.org/10.1038/s41598-024-73765-2

5. Bao, L.; Deng, W.; Huang, B.; Gao, H.; Liu, J.; Ren, L.; Wei, Q.; Yu, P.; Xu, Y.; Qi, F.; Qu, Y.; Li, F.; Lv, Q.; Wang, W.; Xue, J.; Gong, S.; Liu, M.; Wang, G.; Wang, S.; Song, Z.;

Zhao, L.; Liu, P.; Zhao, L.; Ye, F.; Wang, H.; Zhou, W.; Zhu, N.; Zhen, W.; Yu, H.; Zhang, X.; Guo, L.; Chen, L.; Wang, C.; Wang, Y.; Wang, X.; Xiao, Y.; Sun, Q.; Liu, H.; Zhu, F.; Ma, C.; Yan, L.; Yang, M.; Han, J.; Xu, W.; Tan, W.; Peng, X.; Jin, Q.; Wu, G.; Qin, C. The Pathogenicity of SARS-CoV-2 in hACE2 Transgenic Mice. *Nature* **2020**, 583 (7818), 830–833. https://doi.org/10.1038/s41586-020-2312-y.

6. Muñoz-Fontela, C.; Dowling, W. E.; Funnell, S. G. P.; Gsell, P.-S.; Riveros-Balta, A. X.; Albrecht, R. A.; Andersen, H.; Baric, R. S.; Carroll, M. W.; Cavaleri, M.; Qin, C.; Crozier, I.; Dallmeier, K.; de Waal, L.; de Wit, E.; Delang, L.; Dohm, E.; Duprex, W. P.; Falzarano, D.; Finch, C. L.; Frieman, M. B.; Graham, B. S.; Gralinski, L. E.; Guilfoyle, K.; Haagmans, B. L.; Hamilton, G. A.; Hartman, A. L.; Herfst, S.; Kaptein, S. J. F.; Klimstra, W. B.; Knezevic, I.; Krause, P. R.; Kuhn, J. H.; Le Grand, R.; Lewis, M. G.; Liu, W.-C.; Maisonnasse, P.; McElroy, A. K.; Munster, V.; Oreshkova, N.; Rasmussen, A. L.; Rocha-Pereira, J.; Rockx, B.; Rodríguez, E.; Rogers, T. F.; Salguero, F. J.; Schotsaert, M.; Stittelaar, K. J.; Thibaut, H. J.; Tseng, C.-T.; Vergara-Alert, J.; Beer, M.; Brasel, T.; Chan, J. F. W.; García-Sastre, A.; Neyts, J.; Perlman, S.; Reed, D. S.; Richt, J. A.; Roy, C. J.; Segalés, J.; Vasan, S. S.; Henao-Restrepo, A. M.; Barouch, D. H. Animal Models for COVID-19. *Nature* **2020**, 586 (7830), 509–515. https://doi.org/10.1038/s41586-020-2787-6.

7. Handley, A.; Ryan, K. A.; Davies, E. R.; Bewley, K. R.; Carnell, O. T.; Challis, A.; Coombes, N. S.; Fotheringham, S. A.; Gooch, K. E.; Charlton, M.; Harris, D. J.; Kennard, C.; Ngabo, D.; Weldon, T. M.; Salguero, F. J.; Funnell, S. G. P.; Hall, Y. SARS-

CoV-2 Disease Severity in the Golden Syrian Hamster Model of Infection Is Related to the Volume of Intranasal Inoculum. *Viruses* **2023**, 15 (3). https://doi.org/10.3390/v15030748.

8. Sia, S. F.; Yan, L.-M.; Chin, A. W. H.; Fung, K.; Choy, K.-T.; Wong, A. Y. L.; Kaewpreedee, P.; Perera, R. A. P. M.; Poon, L. L. M.; Nicholls, J. M.; Peiris, M.; Yen, H.-L. Pathogenesis and Transmission of SARS-CoV-2 in Golden Hamsters. *Nature* **2020**, 583 (7818), 834–838. https://doi.org/10.1038/s41586-020-2342-5.

9. Chan, J. F.-W.; Zhang, A. J.; Yuan, S.; Poon, V. K.-M.; Chan, C. C.-S.; Lee, A. C.-Y.; Chan, W.-M.; Fan, Z.; Tsoi, H.-W.; Wen, L.; Liang, R.; Cao, J.; Chen, Y.; Tang, K.; Luo, C.; Cai, J.-P.; Kok, K.-H.; Chu, H.; Chan, K.-H.; Sridhar, S.; Chen, Z.; Chen, H.; To, K. K.-W.; Yuen, K.-Y. Simulation of the Clinical and Pathological Manifestations of Coronavirus Disease 2019 (COVID-19) in a Golden Syrian Hamster Model: Implications for Disease Pathogenesis and Transmissibility. *Clin. Infect. Dis.* **2020**, 71 (9), 2428–2446. https://doi.org/10.1093/cid/ciaa325.

10. Darden, A. G.; Streilein, J. W. Syrian Hamsters Express Two Monomorphic Class I Major Histocompatibility Complex Molecules. *Immunogenetics* **1984**, 20 (6), 603–622. https://doi.org/10.1007/bf00430319.

11. Lobigs, M.; Rothenfluh, H. S.; Blanden, R. V.; Müllbacher, A. Polymorphic Peptide Transporters in MHC Class I Monomorphic Syrian Hamster. *Immunogenetics* **1995**, 42 (5), 398–407. https://doi.org/10.1007/bf00179402.

12. McGuire, K. L.; Duncan, W. R.; Tucker, P. W. Syrian Hamster DNA Shows Limited Polymorphism at Class I-like Loci. *Immunogenetics* **1985**, 22 (3), 257–268. https://doi.org/10.1007/bf00404485.

13. Yuhki, N.; O'Brien, S. J. DNA Variation of the Mammalian Major Histocompatibility Complex Reflects Genomic Diversity and Population History. *Proc. Natl. Acad. Sci. U. S. A.* **1990**, 87 (2), 836–840. https://doi.org/10.1073/pnas.87.2.836.

14. Duncan, W. R.; Streilein, J. W. Genetic Analyses of Alloreactions between Recently Wild and Classical Inbred Strains of Syrian Hamsters: Evidence in Favor of a Major Histocompatibility Complex. *Immunogenetics* **1981**, 13 (5), 393–403. https://doi.org/10.1007/bf00346020.

15. Phillips, J. T.; Streilein, J. W.; Proia, D. A.; Duncan, W. R. Immunochemical Characterization of Syrian Hamster Major Histocompatibility Complex Homologues. Adv. Exp. Med. Biol. 1981, 134, 69–85. https://doi.org/10.1007/978-1-4757-0495-2_7.

16. Peptide MHC binding motif. https://www.iedb.org/mhc/173 *Immunome Res*. **2008** Jan 25;4:2. doi: 10.1186/1745-7580-4-2

17. Biniossek, M. L.; Nägler, D. K.; Becker-Pauly, C.; Schilling, O. Proteomic Identification of Protease Cleavage Sites Characterizes Prime and Non-Prime Specificity of Cysteine Cathepsins B, L, and S. *J. Proteome Res*. **2011**, 10 (12), 5363–5373. https://doi.org/10.1021/pr200621z.

18. Hong, S.-Y.; Kim, T.-G.; Kwon, T.-H.; Jang, Y.-S.; Yang, M.-S. Production of an Anti-Mouse MHC Class II Monoclonal Antibody with Biological Activity in Transgenic Tobacco. *Protein Expr. Purif*. **2007**, 54 (1), 134–138. https://doi.org/10.1016/j.pep.2007.02.015.

19. Jia, Y.; Wang, Y.; Dunmall, L. S. C.; Lemoine, N. R.; Wang, P.; Wang, Y. Syrian Hamster as an Ideal Animal Model for Evaluation of Cancer Immunotherapy. *Front. Immunol.* **2023**, 14, 1126969. https://doi.org/10.3389/fimmu.2023.1126969.

Chapter 4: Characterization of MHC Class I-Associated Phosphopeptides

4.1 Introduction

4.1.1 PP2A Pathway

Peptide vaccines against cancer induce tumor-specific immune responses with long-lasting memory to prevent a recurrence. Previous peptide vaccines rely on neoantigens, which are the result of sequence mutations and are difficult to identify. The identification of neoantigens requires previous knowledge from RNA sequencing and are mostly specific to the tissue and patient.[1,2] Since the first class I major histocompatibility complex (MHC)-associated phosphopeptide was discovered in 1998, 3 phosphopeptides have offered promise for the development of vaccines against multiple cancer and viral infectious diseases. These immunogenic phosphopeptides were identified in hepatocellular cancer, breast cancer, colorectal cancer, leukemia, and other cancers.[1,4] The most recent preclinical trial evaluates class I MHC phosphopeptides as neoantigens as immunotherapeutic for treatment of melanoma.[5]

This commonality that different cancer and viral infectious diseases generate the same class I MHC-associated phosphopeptides could be due to a shared common pathology involving protein phosphatase 2A (PP2A), retinoblastoma protein (pRb), and protein 53 (p53).

Protein phosphatases are enzymes that play a critical role in regulating cellular processes by removing phosphate groups from serine, threonine, or tyrosine residues on proteins. They are essential in cell signaling because phosphorylation by kinases and dephosphorylation by phosphatases act as molecular switches. Dephosphorylation can turn off signaling pathways activated by kinases, modulate signal strength or duration, and reset proteins to their unphosphorylated states, preparing them for further rounds of signaling.

PP2A is the most abundant phosphatase in human cells. The second most abundant phosphatase is protein phosphatase 1 (PP1), which has a similar catalytic site compared to that of PP2A. PP2A can be inhibited by a mutation in the catalytic domain but is more commonly inhibited by SET (also known as inhibitor 2 of PP2A) and cancerous inhibitor of phosphatase 2A (CIP2A). Both SET and CIP2A inhibit PP2A by binding to the catalytic domain of PP2A, PP1 with a similar catalytic site structure is susceptible to SET and CIP2A as well. [1,6,7]

In a normal cell, phosphorylation is transient. However, inhibition of PP2A and PP1 in cancer cells causes aberrant signaling by disrupting the normal balance of kinase and phosphatase activity, which results in accumulation of phosphorylated proteins. This extended lifetime of protein phosphorylation provides sufficient time for the phosphorylated protein to undergo ubiquitination and degradation by proteasomes, generating phosphopeptides to be presented by the class I MHC molecules. This difference in the lifetime of phosphorylation between diseased and healthy cells is shared across

multi-types of cancer and virally infected cells, allowing for the generation of the same phosphopeptides.

The inhibition of PP2A is triggered by dysregulation of pRb and p53 in cancer.

The tumor suppressor protein pRb and two closely related proteins p107 and p130 are responsible for gene transcription repression and cell regulation at the G1 regulation point. Their dysregulation has been discovered in multiple cancers. PP1 is responsible for the dephosphorylation of pRb, while PP2A is responsible for the dephosphorylation of Retinoblastoma-like 1 (p107, or RBL1) and Retinoblastoma-like 2 (p130, or RBL2). In cancer, kinase inhibitors such as INK4A(p16) lose their function, and cell damage leads to stress that promotes upregulation of cyclin-dependent kinases (CDKs), which in turn hyperphosphorylate pRb, p107, and p130. [8,9] Once phosphorylated, pRb p107, and p130 dissociate from the E2 family transcription factors (E2Fs). The released E2Fs promote cell cycle progression, resulting in dysregulated transcription. The dissociation of E2Fs also leads to upregulation of CIP2A and SET that inhibit PP2A,[10] resulting in a vicious cycle that prevents reactivation of pRb p107, and p130 by PP1 and PP2A. [1,7,8,9,11]

The tumor suppressor protein p53 regulates cell cycle arrest at the G1/S checkpoint in response to abnormalities in the cell. Dysregulated cell proliferation caused by dysfunctional E2F leads to cellular stress that activates p53.[1] p53 responds to cellular stress by activating protein p21.[12] Protein p21 is responsible for triggering cell growth arrest at the G2/M checkpoint via inhibition of CDK2 and CDK1.[8,9,13] PP2A contributes to the activation of p53 by dephosphorylating Thr55 and regulates p53 through phosphorylation

sites on Mouse Double Minute 2 homolog (MDM2).[1] Inhibition of p53 in cancer cells leads to high levels of gene transcription and upregulation of CIP2A, resulting in inhibition of PP2A,[1] leading to another vicious cycle that prevents p53 from restoration to its normal function. (**Figure 4-1**)

PP2A, pRb, and p53 are closely related, therefore, dysfunction of any one of them will disrupt the normal function of the other two. Furthermore, the limited redundancy of these three proteins further exacerbates the effect of their aberrant signaling,[1] generating more phosphopeptides.

In order for these cancer-associated phosphopeptides to be recognized by the immune system, transporters associated with antigen processing (TAP) proteins translocate the phosphopeptides into the endoplasmic reticulum (ER), where they bind to and stabilize class I MHC molecules. The stabilized class I MHC-peptide complex migrates to the cell surface via the Golgi. The interaction between CD8+ T cell receptors (TCR) and class I MHC-antigenic peptide complexes triggers T cell activation in the lymph nodes and subsequently generates immune responses.[14]

Among the class I MHC phosphopeptides identified in our lab, approximately 80% of those tested can trigger central memory T-cell responses in healthy donors, indicating that these phosphopeptides are antigenic.[1]

This chapter will focus on the identification of phosphopeptides presented by class I MHC molecules on head and neck squamous cell carcinoma (HNSCC) cells.

The hypothesis of my research is that the class I MHC phosphopeptides that have been identified in other cancers and virally infected cells can be identified in HNSCC.

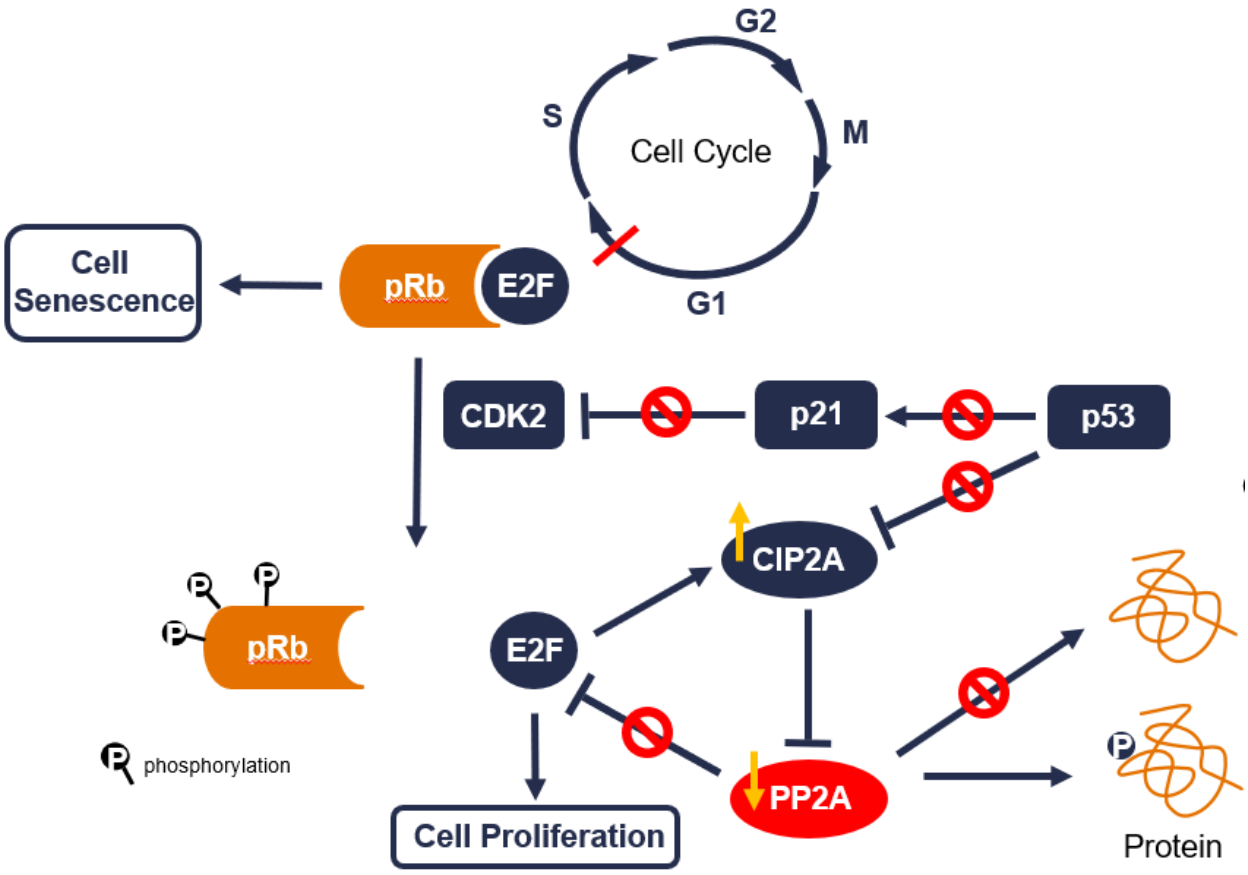


Figure 4- 1: *PP2A Pathway.*

4.1.2 Head and Neck Squamous Cell Carcinoma (HNSCC) Cancer

HNSCC are tumors found in the mucosal epithelium of the oral cavity, oropharynx, larynx, or hypopharynx.[15,16] They generally are associated with tobacco and alcohol consumption. The total incidence of HNSCC is approximately 947,000 cases as of March 15, 2024, which is anticipated to increase by 30% to 1.08 million new cases in 2030. [18-20]

Oropharyngeal infection by the human papillomavirus (HPV) leads to increased prevalence of HNSCC in the USA and Western Europe. [21] HNSCC cases are characterized by 20% HPV-positive and 80% HPV-negative tumors.[22] HPV-positive HNSCC involves an alternative mechanism in which the production of oncoprotein E6 and E7 protein inhibits p53 and pRb, respectively. E7 also directly binds and inhibits PP2A to prevent PP2A from dephosphorylating (hence inhibiting) proteins like protein kinase B (Akt), which inhibits apoptosis by transmitting cell signals to the cell nucleus.[23] HPV-negative HNSCC is the subject of this study for the identification of cancer-associated phosphopeptide antigens.

4.2 Materials

Aldrich Chemical Co. (Milwaukee, WI)

• Ethylenediaminetetraacetic acid (EDTA), 99.99% purity

Bio-Rad Labs (Hercules, CA)

• Tris base >99.8% pure

Fisher Chemical (Ward Hill, NJ)

• Methanol, Optima™ LC/MS grade

Honeywell (Charlotte, NC)

• Acetonitrile, B&J Brand™ LC-MS, for LC-MS and HPLC, >99.9%

Oxford Glyco Systems (Oxfordshire, England)

• Tris HCl, Electrophoresis grade

PQ Corporation (Malvern, PA)

• KaSil 1624 potassium silicate solution

Roche (Basel, Switzerland)

• cOmplete™, Mini Protease inhibitor tablet, EDTA-free

Sigma Aldrich (St Louis, MO)

• Acetic acid, glacial, ≥99.99% trace metals basis

• L-Ascorbic Acid, ACS reagent, ≥99%

• Azulene

• CHAPS, ≥98%

• Iron (III) chloride, anhydrous, powder, ≥99.99% trace metals basis

• Formamide, ≥99.5% (GC), BioReagent for molecular biology

• Hydrochloric Acid, ACS reagent, 37%

• Sodium Chloride, ACS reagent, ≥99.0%

• Sodium hydroxide, reagent grade, 97%, powder

• PhosSTOP™ phosphatase inhibitor tablets

• Triton X-100

Supelco (Bellefonte, PA)

Hydrogen chloride solution, 3M in methanol, for GC derivatization, LiChropur™

Thermo Scientific (Waltham, MA)

• Pierce™ Formic Acid, LC-MS grade

• Halt™ Protease and phosphatase inhibitor cocktail (100x)

• Pierce™ Water, LC-MS grade

Anaspec (Fremont, CA)

• Angiotensin II Phosphate >95% HPLC (Angio)

• Vasoactive Intestinal Peptide (VIP) 1-12 human, porcine, rat >95% HPLC (Vaso)

Sigma Aldrich (St Louis, MO)

• Angiotensin I human acetate salt hydrate >90% HPLC

Atlantic Peptides (Concord, NH)

• Custom phosphopeptide standards

• ST(pS)LVGGR

• FLG(pT)PIAKV

Thermo Scientific (Waltham, MA)

• Phosphopeptide Standards

• RVK(pS)PLFQF

• RTH(pS)LLLLG

Applied Biosystems TM (Carlsbad, CA)

• POROS® MC 20 metal chelating packing material, 20 µm diameter

Cytiva Life Sciences (Marlborough, MA)/GE Healthcare (Boston, MA)

• NHS-Activated Sepharose 4 Fast Flow

Dr Maisch GMBH (Ammerbuch, Germany)

• Reprosil Pur 120 C18 AQ 3 µm

• Reprosil Pur 120 C18 AQ 10 µm

QIAGEN (Hilden, Germany)

• Ni-NTA spin columns

2.2.4 Equipment and Instrumentation

Agilent Technologies (Palo Alto, CA)

• 1100 Agilent high performance liquid chromatograph

Boekel Scientific (Feasterville-Trevose, PA)

Digital Incubator 133000

Branson (Danbury, CT)

• Branson 1200 Ultrasonic Bath

Cadence Science (Cranston, RI)

• Blunt end pipetting needle (16G)

Eppendorf (Enfield, CT/Hamburg, Germany)

• Model 5424 Centrifuge

• Protein LoBind Microcentrifuge tubes (1.5 mL)

Fisherbrand (Pittsburgh, PA)

• Model 50 Sonic Dismembrator with CL-18 Probe

• Mini Tube Rotator

• Digital Vortex

• Vortex Genie 2

GL Sciences (Tokyo, Japan)

• Centrifuge adaptor 5010-21514

Labconco Corp. (Kansas City, MO)

• Centrivap centrifugal vacuum concentrator 7810016

Millipore (Billerica, MA)

• Amicon ultra, 10 kDa regenerated cellulose spin filter

• Amicon ultra, 3kDa regenerated cellulose spin filter

New Objective (Littleton, MA)

• PicoClear 360 Union

PolyMicro Technologies, Inc. (Phoenix, AZ)

• 360 µm o.d. x 75 µm i.d. polyimide coated fused silica capillary

• 360 µm o.d. x 100 µm i.d. polyimide coated fused silica capillary

• 360 µm o.d. x 150 µm i.d. polyimide coated fused silica capillary

Sarstedt Inc (Newton, NC)

• 1.5 mL screw top microtubes

Supelco (Bellefonte, PA)

• Empore octadecyl C18 extraction disks

• Screw top clear glass vials with green thermoset caps and PTFE liners, 2 mL

Sutter Instrument Co. (Novato, CA)

• P-2000 microcapillary laser puller with fused silica adapter

Thermo-Fisher Scientific (San Jose, CA/Bremen, Germany)

• Orbitrap Fusion Tribrid mass spectrometer with commercial front-end ETD

• LTQ mass spectrometer with commercial back-end ETD

• LTQ-Orbitrap Classic mass spectrometer (custom modified with front-end ETD)

Zeus Industrial Products, Inc. (Orangeburg, SC)

• Teflon tubing, 0.012-inch i.d. x 0.060-inch o.d.

4.3 Methods

4.3.1 Class I MHC-associated Peptide Purification

This collaboration involves Dr. Sara Pai from the Department of Medicine and Dr. James Heather from the Department of Surgery at Massachusetts General Hospital, Harvard Medical School. Cell lysis and MHC class I peptide purification from HNSCC samples were performed by Dr. James Heather.

Immunoprecipitation of class I MHC-peptide complexes was accomplished with the pan-human class I antibody W6/32 conjugated to NHS-Sepharose beads. For each gram of tissue (or 1E9 cells), 3 mg of antibody were conjugated to 300 µL of beads. The beads were

washed twice with PBS and incubated overnight with the antibody on a rotating platform to conjugate the antibody to the beads. The next day, the beads were pelleted, rotated for 1 hour with 100 mM Tris-HCl, and then washed twice with alternating solutions of 100 mM ammonium acetate (pH 5, containing 500 mM NaCl) and 100 mM Tris buffer (pH 8). Finally, the beads were resuspended in 20 mM Tris and 150 mM NaCl to a concentration of 1 mg/mL of antibody and stored at 4°C until use.

Tissue samples first were lysed in lysis buffer (20 mM Tris-HCl at pH 8, 1% CHAPS; 10 mL of lysis buffer per 1e9 cells or 1 gram of tissue) with protease inhibitors (5 µg/mL Aprotinin, 10 µg/mL Leupeptin, 10 µg/mL Pepstatin A, 0.04% Azide, 1mM Phenylmethylsulfonyl fluoride (PMSF)) and phosphatase inhibitors.

The cell lysate was ultracentrifuged at 36,000 rpm for 1 hour at 4°C. The supernatant was incubated with bead-conjugated antibody and rotated at 4°C for 16 hours (3 mg of the W6/32 antibody was used per gram of tissues). The tissues were kept at -80°C prior to analysis.

Following incubation, the supernatant was ultracentrifuged at 2000 rpm for 2 minutes, and the beads bound with MHC-peptide complexes were rinsed with following steps: 1 mL 20 mM Tris-HCl and 150 mM NaCl at pH8 twice, 1 mL 20 mM Tris-HCl and 1M NaCl at pH8 twice, and then 1 mL 20 mM Tris-HCl at pH 8 three times. The beads were transferred to a molecular weight cutoff (MWCO) filter (5kDa) and centrifuged for 30 minutes to remove the buffer. The beads in the MWCO filters were shipped on dry ice to the Hunt lab for analysis.

Cell samples were lysed in a buffer containing 20 mM Tris-HCl, 150 mM NaCl, 1% CHAPS, and protease and phosphatase inhibitors (pH 8). The buffer was added at a ratio of 5 mL per 1E9 cells or per gram of cells. Frozen cells were resuspended in the lysis buffer and rotated at 4°C for 1–2 hours. The lysate was ultracentrifuged at 36,800 rpm for 1 hour at 4°C. The supernatant was mixed with the antibody-conjugated beads and rotated overnight at 4°C to facilitate MHC-antibody binding.

4.3.2 Sample Desalting and Contaminants Removal

Samples were desalted and contaminants removed using a hydrophobic interaction liquid chromatography (HILIC)-based (cleanup) procedure. To perform polyhydroxyethyl-A (PHEA) cleanup (HILIC cleanup), a 250 µL pipette tip was plugged with 1 mm glass filter membranes from Whatman GF/F for securing PHEA resin in the tips. 20 µL 300 Å PHEA resin in 100 mM ammonium formate were packed 5 mm to 1 cm deep in the pipette tip using a centrifuge, followed by conditioning of the resin with 100 mM ammonium formate (pH 3), water, and 10 mM ammonium formate (pH 3) in 90% acetonitrile (ACN). A preparation gradient was performed by washing the column sequentially with 100 µL 100 mM ammonium formate (pH 3) once, 100 µL water once, and 100 µL 10 mM ammonium formate (pH 3) three times. This preparation gradient was repeated to further maximize contaminant and salt removal.

The sample was lyophilized using a concentrator and reconstituted in 10 mM ammonium formate (pH 3) with 90% ACN. After loading the sample into the pipette tip, the pipette was rinsed three times with 10 mM ammonium formate (pH 3) with 90% ACN.

The peptides were eluted from resin by rinsing the pipette with 50% ACN in 0.1% acetic acid (AcOH); the flow-through and the eluate from this step was collected and saved for esterification and phosphopeptide enrichment by IMAC. Finally, the pipette was rinsed with 0.2% formic acid to collect any remaining proteins. This step ensures complete elution from the column, ensuring that anything still bound to the column was collected.

4.3.3 Sample Screening

The samples were first screened on a Thermo Scientific Orbitrap Fusion Tribrid Mass Spectrometer, coupled with an Agilent 1100 HPLC system, to assess sample quality based on peptide and contaminant abundances prior to phosphopeptide enrichment.

The samples were loaded onto a precolumn (100 µm inside diameter (i.d.) x 360 µm outside diameter (o.d.) fused silica packed 7 cm long with 10 µm C-18 beads) that first was conditioned with 0.1% AcOH for 10 minutes. HNSCC Sample and 100 fmol each of the internal standards (Angio and Vaso) were loaded onto the precolumn using a pressure vessel for a 15-minute desalting rinse at a flow rate of 100 nL/min with 0.1% AcOH water before connecting to the analytical column (75 µm i.d. x 360 µm o.d. fused silica packed 10 cm long with 3 µm C-18 beads) with a 2 cm Teflon tubing.

Reverse phase separation was conducted at a flow rate of 100 nL/min by HPLC using 0.1% AcOH for solvent A and 0.1% AcOH in 60% ACN for solvent B, and a gradient as follows: 0% to 60% solvent B in 60 minutes, 60% to 100% solvent B in 2 minutes, 100%

solvent B for 4 minutes, 0% in 2 minutes, followed by a 22-minute equilibration with 100% solvent A.

The peptides eluted from the analytical column were electrosprayed into an LTQ-Orbitrap Classic mass spectrometer. MS1 spectra were acquired with 60,000 resolution, 5e5 AGC target and a 300 to 2000 m/z scan range in the Orbitrap analyzer, followed by low resolution data-dependent MS2 acquisition in the ion trap with normal scan rate. The top 3 most abundant precursor ions with a +2 or +3 charge were selected for collision-induced dissociation (CID) in the ion trap with a 1e4 AGC target, 35% normalized collision energy, 30 ms activation time, and a 2.0 m/z isolation window, and for electron transfer dissociation (ETD) with 50 ms activation time and 50 ms maximum injection time. If the same precursor ion was selected three times or was detected twice within a 20-second repeat duration, the ion was excluded for 15 seconds.

4.3.4 Fischer-Speier Esterification

The post-cleanup sample was transferred to a fresh 1.5 mL Lobind tube and lyophilized using a vacuum concentrator. Any remaining water was eliminated by subjecting the sample to two rounds of lyophilization following the addition of 50 µL methanol each time. Subsequently, 80 µL of 3M anhydrous hydrogen chloride in methanol was added to the sample. The sample was vortexed and sonicated in a water bath for approximately 10 seconds until vapors became visible at the side of the Lobind tube. It then was placed in a hood at room temperature for a reaction period of 1 hour before being lyophilized again. This process was repeated 2 times.

4.3.5 Phosphopeptide Enrichment

4.3.5.1 Preparation of NTA beads

5 mg Ni-NTA resin was removed from a Qiagen Ni-NTA spin column and placed in a Lobind tube. It was rinsed with 500 µL water, shaken with a vortex mixer, and then centrifuged. The supernatant was removed, followed by a 5-minute incubation with 500 µL EDTA with occasional shaking during the incubation. The incubation was repeated twice, and the supernatant was removed each time. The NTA resin was equilibrated with 500 µL water twice, and checked to ensure the pH has returned to that of water to confirm the removal of EDTA.

The Ni-NTA column was activated by 10-minute incubation with 500 µL of 130 mM $FeCl_3$ 3 times with occasional shaking during the incubation, and the supernatant was removed each time. The activated resin was rinsed with following steps: 500 µL of 0.01% AcOH once, 500 µL of 15% ACN in 0.01% AcOH twice, and 500 µL of 0.01% AcOH once again. The final NTA resin solution was stored in 1 mL of 0.01% AcOH in a new glass vial at 4 ℃ for up to 1 month.

4.3.5.2 Construction of Fe-NTA IMAC column

A Fe-NTA IMAC column was constructed using a 7-cm long piece of 150 µm i.d. x 360 µm o.d. fused silica with a 2 mm Kasil frit at one end and packed 2.5 cm of NTA beads. The column was conditioned with 0.01% acetic acid (AcOH) for 5 minutes at a flow rate of 10-20 µL/min; then, the column was activated by flowing 130 mM $FeCl_3$ for 5 minutes at a

flow rate of 10-20 µL/min, followed by a 5-minute incubation time. This process was repeated 3 times to ensure binding of iron (Fe^{3+}) to the NTA beads. Finally, the Fe-NTA IMAC column was rinsed with 25-30 µL of 0.01% AcOH at a flow rate of 0.5 µL/min. The pH of the flow through was checked to ensure a pH between 3-4, a pH > 4 indicates either the flowrate is too high during the rinse, and the iron has flushed off from the column, or that the $FeCl_3$ solution has degraded.

4.3.5.3 IMAC

For the IMAC enrichment of phosphopeptides, the dried sample was reconstituted in 50 µL of 60% of ACN and 40% of 0.01% AcOH with approximately pH 3-4. The sample was loaded on the Fe-NTA IMAC column at a flow rate of 0.5 µL/min. The column was rinsed with 25 µL of 60% of ACN and 40% of 0.01% AcOH followed by rinsing with 15 uL of 0.01% AcOH at a flow rate of 0.5 µL/min.

For phosphopeptide elution, the Fe-NTA IMAC column was attached to a homemade pre-column (100 µm i.d. x 360 µm o.d. fused silica with a 2 mm Kasil frit, packed 7 cm of 10µm C-18 resin). The columns were rinsed with 5 µL of 0.01% AcOH to check for leaks at the columns' junction. Phosphopeptides were eluted using 15-20 µL of 200 mM ascorbic acid at a flow rate of 0.5-1 µL/min. The ascorbic acid solution reduced the iron (Fe^{3+}) metal binding to the NTA beads and released phosphopeptides into the pre-column for separation. The columns were rinsed for 25 to 30 minutes with 5 µL of 0.01% AcOH.

The pre-column was detached from the Fe-NTA IMAC column and was used for high-performance liquid chromatography coupled mass spectrometry (HPLC-MS) analysis of the bound phosphopeptides. All IMAC flow-through eluates were collected and stored in the -35°C freezer.

4.3.6 Liquid Chromatography-Mass Spectrometric (LC-MS) Analysis

HNSCC MHC samples and 100 fmol each of the peptide internal standards (Angio and Vaso) were loaded onto the precolumn using a pressure vessel followed by a 15-minute column wash step at a flow rate of 100 nL/min with 0.1% AcOH in water before connecting to the analytical column with 2 cm of Teflon tubing.

Reverse phase separation was accomplished at a flow rate of 100 nL/min using 0.1% AcOH for solvent A and 0.1% AcOH in 60% acetonitrile for solvent B, and a gradient as follows: 0% to 60% solvent B in 60 minutes, 60% to 100% solvent B in 2 minutes, 100% solvent B for 4 minutes, 0% in 2 minutes, followed by a 22-minute equilibration with 100% solvent A.

The peptides eluted from the analytical column were electrosprayed into an Orbitrap Fusion Tribrid mass spectrometer.

4.3.7 Mass Spectrometric Analysis

MS1 spectra were acquired with a resolution of 60,000, using an RF lens setting of 60%, a maximum injection time of 50 ms, and an automatic gain control (AGC) target of 4e5. The scan range was set as 300 to 1500 m/z in the Orbitrap analyzer, with an intensity

filter threshold of 3e4. This was followed by data-dependent MS2 acquisition. A dynamic exclusion window of 7 seconds was applied, preventing the reselection of the same precursor ion if it was selected twice within that period.

For precursor ions with a charge state of +2 and a mass range of 400–1000 m/z, fragmentation was performed to generate MS2 spectra first by CID with detection in the Orbitrap at a resolution of 7500, an AGC target of 1e5, a collision energy of 30%, and an activation time of 10 ms. The maximum injection time was 250 ms, with an activation Q of 0.25, and a 1.6 m/z isolation window. Following CID, the same precursor ions were selected for ETD, with detection in the Orbitrap at a resolution of 15,000, an AGC target of 2e4, a maximum injection time of 50 ms, an ETD reagent target of 2e5, a 50 ms ETD injection time, a 45 ms reaction time, and a 1.6 m/z isolation window.

For precursor ions with charge states of +3 to +4 and a mass range of 300–800 m/z, fragmentation was initially conducted by CID to generate MS2 spectra with the same Orbitrap settings as above: 7500 resolution, an AGC target of 1e5, a 30% collision energy, a 10 ms activation time, a 250 ms maximum injection time, an activation Q of 0.25, and a 1.6 m/z isolation window. Subsequently, these precursor ions were analyzed by ETD using a rapid ion trap scan rate, an AGC target of 2e4, a maximum injection time of 50 ms, an ETD reagent target of 2e5, and charge-dependent calibrated ETD parameters, with a 1.6 m/z isolation window.

Precursor ions with a +1 charge state and a mass range of 800–1500 m/z were fragmented to generate MS2 spectra using HCD, with Orbitrap detection at a resolution of

7500, an AGC target of 1e5, a collision energy of 30%, a maximum injection time of 250 ms, and a 1.6 m/z isolation window.

4.3.8 Data Analysis

The mass spectra were searched against a human protein database (SwissProt with isoforms) using the Byonic software (Protein Metrics) with a 10 ppm precursor ion mass (MS1) tolerance and a 0.4 Da fragment ion mass (MS2) tolerance. The following amino acid modifications were used: fixed C-terminal, aspartic acid, and glutamic acid methylation, and common variable modifications of 1 methionine oxidation, 2 serine, threonine, and tyrosine phosphorylations, and 1 cysteine cysteinylation. One asparagine or glutamine deamidation and 2 aspartic acid or glutamic acid dimethylations were allowed as rare variable modifications. A maximum of three common variable modifications and two rare variable modifications were allowed.

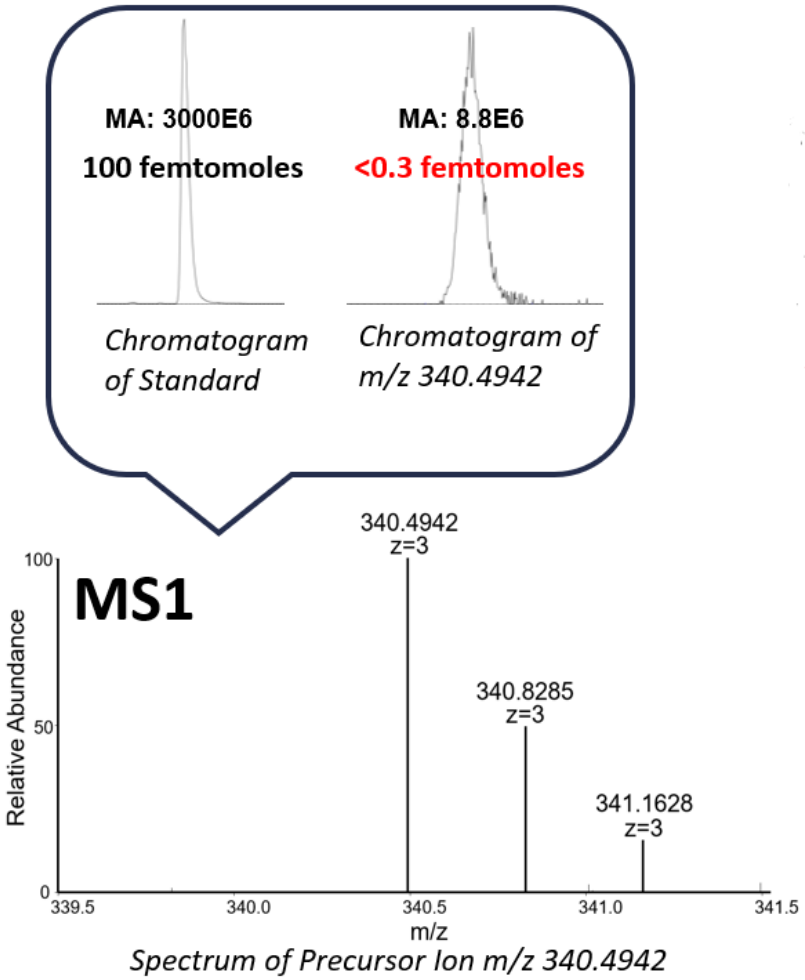
The search algorithm identifications were manually validated.

4.4 Results and Discussion

4.4.1 LC-MS Analysis of MHC Class I-Associated Peptides

By comparing the area under the curve (AUC) of the chromatographic peak for the ion at m/z 340.4942 to that of standards with known abundance, the relative abundance of this phosphopeptide was estimated to be less than 0.3 femtomoles (**Figure 4-2**). The AUC is a direct measure of ion intensity in the chromatogram, which reflects the concentration

of the peptide in the sample. By referencing this to standard peptides with known concentrations, it is possible to estimate the abundance of the target phosphopeptides. This approach not only confirms the presence of the peptide but also provides a quantitative measure, illustrating the sensitivity of the LC-MS method in detecting such low-abundance analytes. Achieving reliable detection at sub-femtomole levels underscores the robustness of the method in capturing trace amounts of phosphorylated species, which are often challenging to detect due to their low natural abundance and tendency to undergo modifications or degradation.



Spectrum of Precursor Ion m/z 340.4942

Figure 4- 2: *MS1 Spectrum and Chromatogram of Precursor Ion m/z 340.4942.*

Comparison of area under the peak of the standards to the ion with m/z 340.4942 is shown. MA: Mass Area, also refer to as area under the peak (AUC)

In **Figure 4-3**, a comparison between CID and ETD is presented, highlighting the differential fragmentation behavior of the phosphopeptides. CID, a common fragmentation method, often leads to the cleavage of labile bonds, including the phosphoester bond, resulting in the loss of the phosphate group. This loss complicates efforts to accurately localize the phosphorylation site within the peptide because the phospho-group is no longer retained in the fragment ions. As a result, CID spectra often lack the necessary information to localize the phosphorylation site.

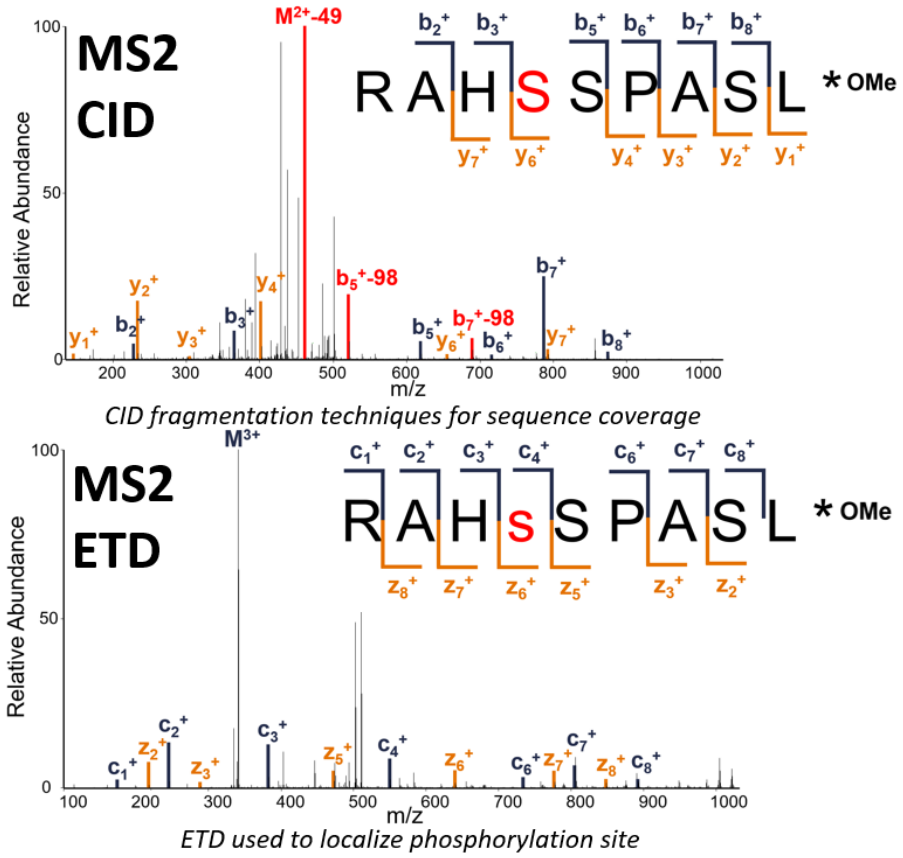


Figure 4- 3: *CID and ETD tandem mass spectrometry (MS/MS) spectra acquired showing fragment masses of precursor ion with m/z 340.4942.*

Sequence coverage for b+ and c+ ion fragments are labeled in blue, and y+ and z+ ion fragments are labeled in yellow. The identified ions are sufficient for complete sequence coverage. ETD successfully localized the phosphorylation site.

On the other hand, ETD provides a distinct advantage for phosphopeptide analysis. ETD preserves labile post-translational modifications like phosphorylation by cleaving the peptide backbone without disrupting the phospho-group. This allows for the generation of fragment ions that retain the modification, enabling more precise localization of the phosphorylation site. In this specific case, the use of ETD results in fragmentation that maintains the integrity of the phosphopeptide, as shown in **Figure 4-3**. The ability to confidently localize the phosphorylation site is critical for understanding biological signaling pathways and protein modifications, making ETD a more suitable method for phosphopeptide analysis in cases where site-specific information is essential.

4.4.2 Class I-associated Phosphopeptides

The challenges in identifying MHC-associated phosphopeptides include their low abundance, often at the attomole level, as well as the number of unique peptides and their similarity in size and molecular weight.

Although the levels of phosphopeptides are expected to be elevated in diseased cells, the abundances of phosphopeptides are much lower than that of unmodified peptides in a sample. Therefore, selective enrichment methods are needed prior to mass spectrometry analysis to prevent contaminants such as unmodified peptides, polymers, and salts from suppressing the signal of phosphopeptides. However, phosphopeptides are

difficult to enrich due to their negative charge, which makes them highly hydrophilic and susceptible to loss into the polar aqueous mobile phase (solvent A) of liquid chromatography while loading the sample into the column.

As a solution, hydrophilic interaction liquid chromatography (HILIC) with PHEA was used to remove ions and polymers from the sample, and iron (III)-nitrilotriacetic acid immobilized metal affinity chromatography (Fe-NTA IMAC) was used to enrich phosphopeptide via electrostatic interactions. The PHEA/HILIC method separates peptides based on peptide hydrophilicity, the separation is accomplished by the strength of hydrogen bonding interaction between peptides and the neutral stationary phase composed of PHEA.[24] Prior to IMAC, esterification modifies negatively charged C-terminus and amino acids such as aspartic and glutamic acid by converting their carboxylic side chain to corresponding methyl esters, which prevent them from competing with the phosphate group of phosphopeptides for binding to IMAC resin (Fe-NTA). IMAC uses positively charged metal ions (Fe+3) to bind negatively charged phosphate groups, enriching phosphopeptides from other non-modified peptides.[25] The IMAC column is connected to the pre-column for sample loading to reduce sample loss and contamination. Subsequently, the pre-column was connected to the analytical column for liquid-chromatography coupled mass spectrometry (LC-MS) analysis. This methodology enables the enrichment and identification of phosphopeptide at the attomole level. [25]

Application of the method to the analysis of P01127B HNSCC samples enabled the identification of 27 class I MHC-associated phosphopeptides that have been reported in

analysis of other cancer cells including breast cancer (**Figure 4-4**), colorectal cancer, esophageal cancer, lung cancer, cervical cancer, liver cancer, leukemia, and melanoma. In addition, 2 potentially new cancer-associated phosphopeptides were identified. Thus far, 6 of these phosphopeptides have been validated manually, and the others are in the process of verification.

#	Peptide	Related Tumor									Healthy Donor Test
		B	C	E	H	L	M	T	U	V	
1	**DSsEEKFL**	B	C		H	L	M				Not tested
2	FASPTsPPVL	B			H	L					Not tested
3	KPFKLSGLsF				H	L	M				Not tested
4	LPVsPGHRKT					L					Not tested
5	RPKP(sss)PVIF				H	L					Not tested
6	RPKtPPPAP					L					Not tested
7	RPIsPRIGAL	B	C		H	L	M				Not tested
8	RPIsVIGGVSLY			E	H						Not tested
9	RPIsVIGGVSL			E	H						Not tested
10	**RAHsSPASL**	B	C	E	H		M		U	V	No
11	RPRsPGSNSKV	B	C			L	M	T			No
12	RPRsPPGGP		C		H						Not tested
13	RPVsPAGPP					L					Not tested
14	RPVsPGKDITA					L					Not tested

#	Peptide	Related Tumor									Healthy Donor Test
		B	C	E	H	L	M	T	U	V	
15	RPVsPPQKA					L					Not tested
16	RRIsDPQVF	B	C	E	H	L	M		U	V	Not tested
17	RTIsQSSSL				H	L					Not tested
18	SAIsPKSSL				H	L					Not tested
19	SPKPPTRsP					L					Not tested
20	TPIsPLKTGV									V	Not tested
21	VPSsPLRKA					L					Not tested
22	ITQGtPLKY	B	C	E	H	L	M				Yes
23	SPKsPGLKA		C		H	L	M	T			Yes
24	RSHsSPASL		C	E			M		U	V	Yes
25*	RPIsPSPSA										Not tested
26*	RPARsVPSIAA										Not tested
27*	LPtSPRLTA										Not tested

*New phosphopeptides identified

B Breast
C Colorectal
E Esophageal
H Liver
L Leukemia/Lymphoma
M Melanoma
T Tonsil
U Lung
V Cervical

Figure 4- 4: *Phosphopeptides Identified in HNSCC Sample P01127B.*

Among the 29 phosphopeptides identified, 22 are phosphorylated at position 4 (**Figure 4-5**). This P4 orientation is structurally conserved for the phosphopeptides. When a phosphopeptide binds to the class I MHC molecule, position 4 is oriented away from the MHC molecule and cell surface, where it is solvent-exposed, hence available for direct contact with the T cell receptor. The negatively charged phosphate modification also allows for multiple stabilizing interactions with nearby MHC residues.[26]

The Class I MHC alleles of the patients are unknown but could be determined using the combination of the R program I wrote and GibbsCluster 2.0 (**Figure 4-6**).

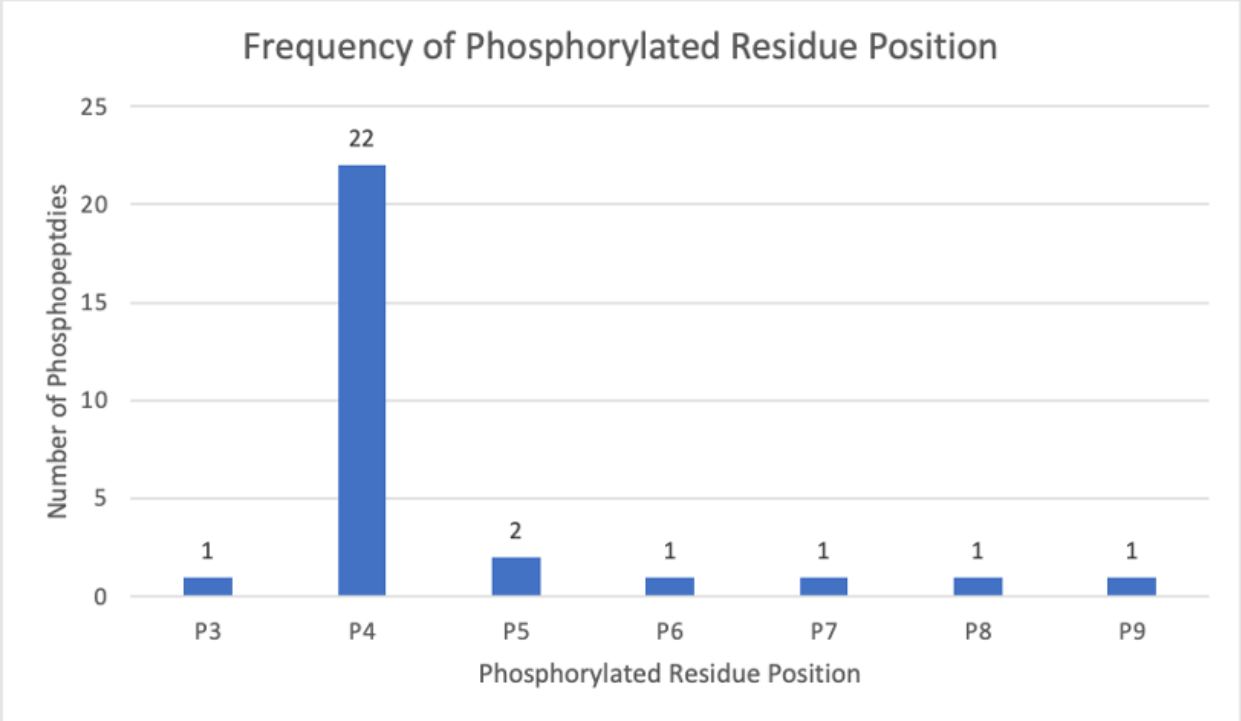


Figure 4- 5: *Frequency of Phosphorylated Residue Position.*

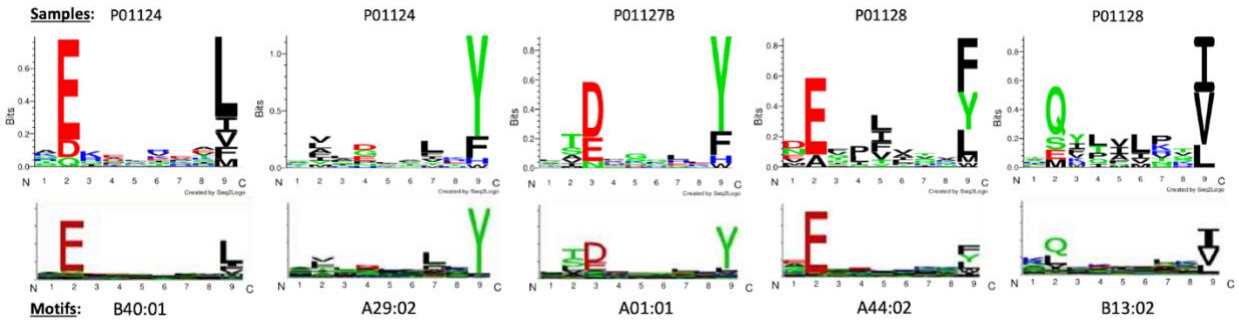


Figure 4- 6: *Class I HLA Alleles Determined from Selected HNSCC Samples.*

Sequence alignment of peptides found in HNSCC samples (top panel) in comparison with known HLA motifs (bottom panel).

4.5 Future Directions

The IMAC-LC-MS method enabled identification of 27 class I MHC-associated phosphopeptides in HNSCC tissue. The identification of phosphopeptides that have been identified in other cancers supports the hypothesis that cancer and virally infected cells can generate the common class I MHC-associated phosphopeptide, resulting from common PP2A downregulation.

One follow-on study is to analyze class I MHC-associated phosphopeptides presented on HPV-positive HNSCC cells. Because HPV infection involves an alternative mechanism, in which the production of oncoprotein E6 and E7 proteins inhibit p53 and pRb, respectively. The inhibition of p53 and pRb by E6 and E7 causes downstream changes in phosphorylation-based signaling, which is tightly regulated in normal cells but becomes dysregulated in HPV-positive cells. Phosphorylation is a critical regulatory mechanism in the cell, controlling various processes such as cell cycle progression, DNA repair, and stress responses. When p53 and pRb are inactivated, compensatory pathways are

activated, often involving increased kinase activity or reduced phosphatase activity, which results in aberrant or excessive phosphorylation events. [27] With this altered cellular signaling, additional phosphopeptides are expected to be found in HNSCC samples that are HPV-positive.

Modifications to the phosphopeptide enrichment method could be tested in attempts to identify more phosphopeptides in the cancer samples. A possible direction for improvement of phosphopeptide enrichment is the use of other chromatography methods orthogonal to the HILIC and IMAC. For example, strong cation exchange chromatography (SCX) has been reported as a highly effective method for phosphopeptide enrichment;[28] however, SCX does not retain multi-phosphorylated peptides. Yet another method, the electrostatic repulsion-hydrophilic interaction chromatography (ERLIC) has been reported as highly efficient in retaining multiply-phosphorylated peptides, and it can be further improved by coupling ERLIC with titanium dioxide (TiO2).[29] So far, the phosphopeptides identified in our lab are mostly singly phosphorylated; therefore, the addition of ERLIC to our current phosphopeptide enrichment method might help identify multi-phosphorylated peptides.

Among these phosphopeptide enrichment methods, SCX has a high potential for efficient removal of contaminants such as salts, polymers, and ions in a sample. This is because SCX separates peptides based on their charge states; therefore, most polymers and salts will not be retained in the SCX column due to their net-zero charge and can be removed as a result.

At the same time, the current PHEA/HILIC method can be improved to be more suitable for different sample types. Previous studies have shown optimized phosphopeptide enrichment with increasing salt concentration,[30,31] different tissues also have varying levels of ions (e.g., sodium, potassium, and chloride), which can contribute to the baseline salt content. For instance, muscle tissues contain high levels of potassium and sodium, naturally contributing to increased salt content. Kidneys, which regulate electrolyte balance, may also have elevated salt concentrations. [32]

However, the salt is not compatible with mass spectrometry and will result in clogging of the column tip and contaminating the ion source components. Instead, volatile salts such as ammonium acetate or ammonium formate could be tested.

After the identification of phosphopeptides, the next step is to evaluate whether these phosphopeptides are capable of triggering central memory T-cell responses in healthy donors using an enzyme-linked immunospot (ELISpot) assay. The ELISpot assay uses monoclonal antibodies to capture cytokines secreted by the cell when a MHC-bound phosphopeptide triggers an T-cell-based immune response. A phosphopeptide that can elicit an immune response makes it a promising candidate for immunotherapy.[33]

Notably, results from a recent preclinical trial demonstrated the efficacy of two of our class I MHC phosphopeptides in treating high-risk melanomas. One peptide elicited T cell responses in 5 out of 12 patients, while another peptide elicited T cell responses in 2 out of 12 patients.[34]

4.6 References

1. Mahoney, K. E.; Shabanowitz, J.; Hunt, D. F. MHC Phosphopeptides: Promising Targets for Immunotherapy of Cancer and Other Chronic Diseases. *Molecular & Cellular Proteomics* **2021**, *20*, 100112.

2. Slingluff, C. L. The Present and Future of Peptide Vaccines for Cancer. *The Cancer Journal* **2011**, *17* (5), 343–350.

3. Hogan K.T.; Eisinger D.P,; Cupp S.B 3rd.; Lekstrom K.J.; Deacon D.D.; Shabanowitz J., Hunt D.F.; Engelhard V.H.; Slingluff C.L. Jr.; Ross M.M. The peptide recognized by HLA-A68.2-restricted, squamous cell carcinoma of the lung-specific cytotoxic T lymphocytes is derived from a mutated elongation factor 2 gene. *Cancer Res*.**1998**, Nov 15;58(22):5144-50.

4. Cobbold, M.; De La Peña, H.; Norris, A.; Polefrone, J. M.; Qian, J.; English, A. M.; Cummings, K. L.; Penny, S.; Turner, J. E.; Cottine, J.; Abelin, J. G.; Malaker, S. A.; Zarling, A. L.; Huang, H.-W.; Goodyear, O.; Freeman, S. D.; Shabanowitz, J.; Pratt, G.; Craddock, C.; Williams, M. E.; Hunt, D. F.; Engelhard, V. H. MHC Class I–Associated Phosphopeptides Are the Targets of Memory-like Immunity in Leukemia. *Science Translational Medicine* **2013**, *5* (203).

5. Engelhard, V. H.; Obeng, R. C.; Cummings, K. L.; Petroni, G. R.; Ambakhutwala, A. L.; Chianese-Bullock, K. A.; Smith, K. T.; Lulu, A.; Varhegyi, N.; Smolkin, M. E.; Myers, P.; Mahoney, K. E.; Shabanowitz, J.; Buettner, N.; Hall, E. H.; Haden, K.; Cobbold, M.;

Hunt, D. F.; Weiss, G.; Gaughan, E.; Slingluff, Jr, C. L. MHC-Restricted Phosphopeptide Antigens: Preclinical Validation and First-in-Humans Clinical Trial in Participants with High-Risk Melanoma. *Journal for ImmunoTherapy of Cancer* **2020**, *8* (1).

6. Favre, B.; Turowski, P.; Hemmings, B. A. Differential Inhibition and Posttranslational Modification of Protein Phosphatase 1 and 2A in MCF7 Cells Treated with Calyculin-A, Okadaic Acid, and Tautomycin. *Journal of Biological Chemistry* **1997**, *272* (21), 13856–13863.

7. Sangodkar, J.; Farrington, C. C.; McClinch, K.; Galsky, M. D.; Kastrinsky, D. B.; Narla, G., All roads lead to PP2A: exploiting the therapeutic potential of this phosphatase. *Febs Journal* **2016**, 283 (6), 1004-1024.

8. Malumbres, M.; Barbacid, M., To cycle or not to cycle: A critical decision in cancer. *Nature Reviews Cancer* **2001**, 1 (3), 222-231.

9. Pezzella, F. T., Mahvash. Kerr, David J., *Oxford Textbook of Cancer Biology*. Oxford University Press: 2019.

10. Soofiyani, S. R.; Hejazi, M. S.; Baradaran, B. The Role of cip2a in Cancer: A Review and Update. *Biomedicine & Pharmacotherapy* **2017**, *96*, 626–633.

11. Kurimchak, A.; Graña, X. PP2A: More than a Reset Switch to Activate PRB Proteins during the Cell Cycle and in Response to Signaling Cues. *Cell Cycle* **2014**, *14* (1), 18–30.

12. Laine, A.; Sihto, H.; Come, C.; Rosenfeldt, M. T.; Zwolinska, A.; Niemelä, M.; Khanna, A.; Chan, E. K.; Kähäri, V.-M.; Kellokumpu-Lehtinen, P.-L.; Sansom, O. J.; Evan, G. I.; Junttila, M. R.; Ryan, K. M.; Marine, J.-C.; Joensuu, H.; Westermarck, J. Senescence Sensitivity of Breast Cancer Cells Is Defined by Positive Feedback Loop between CIP2A and E2F1. *Cancer Discovery* **2013**, *3* (2), 182–197.

13. Polager, S.; Ginsberg, D. P53 And E2F: Partners in Life and Death. *Nature Reviews Cancer* **2009**, *9* (10), 738–748.

14. Neefjes, J.; Jongsma, M. L. M.; Paul, P.; Bakke, O., Towards a systems understanding of MHC class I and MHC class II antigen presentation. *Nature Reviews Immunology* **2011**, 11 (12), 823-836.

15. Leemans, C. R.; Braakhuis, B. J. M.; Brakenhoff, R. H., The molecular biology of head and neck cancer. *Nature Reviews Cancer* **2011**, 11 (1), 9-22.

16. Johnson, D.E., Burtness, B., Leemans, C.R. et al. Head and neck squamous cell carcinoma. *Nat Rev Dis Primers* 6, 92 (**2020**). https://doi.org/10.1038/s41572-020-00224-3

17. Ferlay J, Ervik M, Lam F, Laversanne M, Colombet M, Mery L, Piñeros M, Znaor A, Soerjomataram I, Bray F (2024).

18. Global Cancer Observatory: Cancer Today. Lyon, France: International Agency for Research on Cancer. Available from: https://gco.iarc.who.int/today, accessed [15 Mar 2024].

19. Sung H, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A, Bray F. Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA *Cancer J Clin*. **2021** Feb 4. doi: 10.3322/caac.21660. Epub ahead of print. PMID: 33538338.

20. Ferlay J, Colombet M, Soerjomataram I, Parkin DM, Piñeros M, Znaor A, Bray F. Cancer statistics for the year **2020**: An overview. *Int J Cancer*. 2021 Apr 5. doi: 10.1002/ijc.33588. Epub ahead of print. PMID: 33818764.

21. Mehanna, H. et al. Prevalence of human papillomavirus in oropharyngeal and nonoropharyngeal head and neck cancer–systematic review and meta-analysis of trends by time and region. Head Neck 35, 747–755 (**2013**).

22. Leemans, C. R.; Braakhuis, B. J. M.; Brakenhoff, R. H., The molecular biology of head and neck cancer. *Nature Reviews Cancer* 2011, 11 (1), 9-22.

23. Pim, D.; Massimi, P.; Dilworth, S. M.; Banks, L., Activation of the protein kinase B pathway by the HPV-16 E7 oncoprotein occurs through a mechanism involving interaction with PP2A. *Oncogene* 2005, 24 (53), 7830-7838.

24. Qing G.; Lu Q.; Xiong Y.; Zhang L.; Wang H.; Li X.; Liang X.; Sun T. New Opportunities and Challenges of Smart Polymers in Post-Translational Modification Proteomics. *Adv. Mater.* **2017**; 29, 1604670.

25. Abelin J.G.; Trantham P.D.; Penny S.A.; Patterson A.M.; Ward S.T.; Hildebrand W.H.; Cobbold M.; Bai D.L.; Shabanowitz J.; Hunt D.F. Complementary IMAC enrichment

methods for HLA-associated phosphopeptide identification by mass spectrometry. *Nat Protoc*. **2015** Sep;10(9):1308-18.

26. Mohammed, F.; Cobbold, M.; Zarling, A. L.; Salim, M.; Barrett-Wilt, G. A.; Shabanowitz, J.; Hunt, D. F.; Engelhard, V. H.; Willcox, B. E. Phosphorylation-Dependent Interaction between Antigenic Peptides and MHC Class I: A Molecular Basis for the Presentation of Transformed Self. *Nature Immunology* **2008**, *9* (11), 1236–1243.

27. Jabbar, S.; Strati, K.; Shin, M. K.; Pitot, H. C.; Lambert, P. F. Human Papillomavirus Type 16 E6 and E7 Oncoproteins Act Synergistically to Cause Head and Neck Cancer in Mice. *Virology* **2010**, 407 (1), 60–67. https://doi.org/10.1016/j.virol.2010.08.003.

28. Macek B.; Mann M.; Olsen J.V. Global and site-specific quantitative phosphoproteomics: principles and applications. *Annu Rev Pharmacol Toxicol*. **2009**; 49:199-221.

29. Zarei M.; Sprenger A.; Metzger F.; Gretzmeier C.; Dengjel J. Comparison of ERLIC–TiO_2, HILIC–TiO_2, and SCX–TiO_2 for Global Phosphoproteomics Approaches. *Journal of Proteome Research*. **2011**;10(8), 3474-3483.

30. Alpert A.J. Effect of salts on retention in hydrophilic interaction chromatography. *Journal of Chromatography A*. **2018**;1538:45-53.

31. Cui Y.; Tabang D.N.; Zhang Z.; Ma M.; Alpert A.J.; Li L. Counterion Optimization Dramatically Improves Selectivity for Phosphopeptides and Glycopeptides in Electrostatic Repulsion-Hydrophilic Interaction Chromatography. *Anal Chem*. **2021** Jun 8;93(22):7908-7916.

32. Rossitto, G.; Mary, S.; Chen, J. Y.; Boder, P.; Chew, K. S.; Neves, K. B.; Alves, R. L.; Montezano, A. C.; Welsh, P.; Petrie, M. C.; Graham, D.; Touyz, R. M.; Delles, C. Tissue Sodium Excess Is Not Hypertonic and Reflects Extracellular Volume Expansion. *Nat. Commun*. **2020**, 11 (1), 4222. https://doi.org/10.1038/s41467-020-17820-2.

33. Slota, M.; Lim, J.-B.; Dang, Y.; Disis, M. L. Elispot for Measuring Human Immune Responses to Vaccines. *Expert Review of Vaccines* **2011**, *10* (3), 299–306.

34. Engelhard, V. H.; Obeng, R. C.; Cummings, K. L.; Petroni, G. R.; Ambakhutwala, A. L.; Chianese-Bullock, K. A.; Smith, K. T.; Lulu, A.; Varhegyi, N.; Smolkin, M. E.; Myers, P.; Mahoney, K. E.; Shabanowitz, J.; Buettner, N.; Hall, E. H.; Haden, K.; Cobbold, M.; Hunt, D. F.; Weiss, G.; Gaughan, E.; Slingluff, C. L., Jr. MHC-Restricted Phosphopeptide Antigens: Preclinical Validation and First-in-Humans Clinical Trial in Participants with High-Risk Melanoma. *J. Immunother. Cancer* **2020**, 8 (1), e000262. https://doi.org/10.1136/jitc-2019-000262.

Chapter 5: Future Directions

5.1 Single Cell Proteomics (SCP)

Characterizing MHC-associated peptides presents a unique challenge due to their very low abundance, requiring highly refined skills and methods. These skills are directly transferable to the field of single-cell proteomics, where precision and sensitivity are equally critical.

One promising approach is the use of advanced mass spectrometry techniques, such as Single Cell ProtEomics (SCoPE2).[1] Unlike scRNA-seq methods, which, while powerful, often fail to accurately reflect protein expression due to the complexities of post-translational regulation, SCoPE2 provides a more direct measure of protein abundance and function. [1,2,3]

Traditional antibody-based proteomics methods, although widely used, suffer from limitations such as signal suppression, particularly in highly multiplexed settings. Single-cell proteomics offers significant advantages, especially in cancer research. For example, it allows for the detailed characterization of tumor heterogeneity, enabling researchers to understand the diverse cellular landscape of tumors. Additionally, it can elucidate underlying tumor mechanisms and pathways, offering insights into disease progression and potential therapeutic targets. [3]

5.2 Structural Proteomics

Besides its applications in single-cell proteomics, mass spectrometry plays a pivotal role in structural proteomics, providing detailed insights into protein structure, dynamics, and interactions. One such technique is Hydrogen-Deuterium Exchange Mass Spectrometry (HDX-MS), which is widely utilized to investigate protein folding, conformational changes, and binding interactions at a high level of detail. [4]

HDX-MS measures the rate at which hydrogen atoms in a protein exchange with deuterium when exposed to deuterated water (D_2O).

Typically, a buffer containing 90%-95% D_2O by volume is used. Under these conditions, hydrogens on backbone amides and side chains begin to swap with deuterium. After a defined exchange period, ranging from milliseconds to hours, the reaction is stopped by rapidly lowering the pH to ~2.5 and the temperature to 0°C, shifting from physiological conditions (pH 7 and room temperature). The protein is then subjected to proteolytic digestion using acid-tolerant proteases like pepsin, allowing for the analysis of deuterium incorporation. However, challenges such as background exchange necessitate precise control of temperature and pH to retain the incorporated deuterium. [4]

HDX-MS can provide insights into protein folding and unfolding by tracking deuterium incorporation into newly exposed regions of a protein during conformational changes.[5] Additionally, it can identify binding sites by comparing the deuterium exchange profile of a protein alone versus in complex with ligands or other proteins.[6]

5.3 References

1. Petelski, A. A.; Emmott, E.; Leduc, A.; Huffman, R. G.; Specht, H.; Perlman, D. H.; Slavov, N. Multiplexed Single-Cell Proteomics Using SCoPE2. *bioRxiv*, **2021**. https://doi.org/10.1101/2021.03.12.435034.
2. Slavov, N. Single-Cell Protein Analysis by Mass-Spectrometry. *arXiv [q-bio.QM]*, **2020**. http://arxiv.org/abs/2004.02069.
3. Tan, Y. C.; Low, T. Y.; Lee, P. Y.; Lim, L. C. Single-Cell Proteomics by Mass Spectrometry: Advances and Implications in Cancer Research. *Proteomics* **2024**, *24* (12–13), e2300210. https://doi.org/10.1002/pmic.202300210.
4. Masson, G. R.; Burke, J. E.; Ahn, N. G.; Anand, G. S.; Borchers, C.; Brier, S.; Bou-Assaf, G. M.; Engen, J. R.; Englander, S. W.; Faber, J.; Garlish, R.; Griffin, P. R.; Gross, M. L.; Guttman, M.; Hamuro, Y.; Heck, A. J. R.; Houde, D.; Iacob, R. E.; Jørgensen, T. J. D.; Kaltashov, I. A.; Klinman, J. P.; Konermann, L.; Man, P.; Mayne, L.; Pascal, B. D.; Reichmann, D.; Skehel, M.; Snijder, J.; Strutzenberg, T. S.; Underbakke, E. S.; Wagner, C.; Wales, T. E.; Walters, B. T.; Weis, D. D.; Wilson, D. J.; Wintrode, P. L.; Zhang, Z.; Zheng, J.; Schriemer, D. C.; Rand, K. D. Recommendations for Performing, Interpreting and Reporting Hydrogen Deuterium Exchange Mass Spectrometry (HDX-MS) Experiments. *Nat. Methods* **2019**, *16* (7), 595–602. https://doi.org/10.1038/s41592-019-0459-y.

5. Englander, S. W.; Mayne, L. The Nature of Protein Folding Pathways. *Proc. Natl. Acad. Sci. U. S. A.* **2014**, *111* (45), 15873–15880. https://doi.org/10.1073/pnas.1411798111.
6. Chalmers, M. J.; Busby, S. A.; Pascal, B. D.; West, G. M.; Griffin, P. R. Differential Hydrogen/Deuterium Exchange Mass Spectrometry Analysis of Protein-Ligand Interactions. *Expert Rev. Proteomics* **2011**, 8 (1), 43–59. https://doi.org/10.1586/epr.10.109.