**Enrichment and Characterization of Post-Translationally Modified Peptides
for the Development of Novel Cancer Immunotherapeutics**

Stacy Alyse Malaker
Sterling Heights, Michigan

B.S., Biochemistry, University of Michigan, 2009
B.S., Anthropology-Zoology, University of Michigan, 2009

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
December, 2014

_____

_____

_____

_____

_____

This dissertation is dedicated to Ryan Patrick,
for all he taught me that cannot be learned in a classroom.

Acknowledgements

This dissertation would not be possible without the help and support of so many individuals. I cannot possibly express how thankful I am for everyone in my life that has helped me along the way, but I suppose I can try…

First, I need to thank the man, the myth, and the legend, Dr. Donald F. Hunt. I took a very convoluted path to find myself in this laboratory, but I thank some sort of deity every day for bringing me to study here. Don, I have to thank you for every crazy idea you have had. For every time you said, "This is the last experiment" (which you said way too often). For every time you looked at me, squinted skeptically, and questioned my thought process. For every time you laughed when I walked in your office. For every time I got a high-five or an A+ after successful experiments. For the flowers! Mostly, I have to thank you for making me a better scientist, and for giving me a sincere confidence in myself and my laboratory skills. I am eternally grateful for everything you have taught me during my graduate career.

Next, I need to thank Dr. Jeff Shabanowitz, the other man in charge. When I joined the lab, I was so scared of Jeff I wouldn't step foot in his office. After bonding over our brothers with Down syndrome, this fear broke down into enduring respect. Jeff, thank you all that you've done for me. I love all of your stories – from those making fun of Don to those about the "old days" in the lab. I appreciate all of your advice – from science and failed experiments to life and failed relationships. I admire your strength, passion, and remarkable intelligence. Thank you for picking me up when I thought I'd hit

rock bottom. That conversation was more important than I think you will ever know, and I cannot thank you enough for that. I hope that one day I can return the favor.

I must also thank my amazing collaborators, Dr. Mark Cobbold, Dr. Sarah Penny, and Lora Steadman, who have provided me with all of the samples and biological data necessary to make this dissertation possible. Mark and Sarah taught me nearly everything I know about immunology, seeing as I joined this lab with essentially zero knowledge on the subject. I had the wonderful opportunity to study with them in the UK, which was one of the most excellent experiences in my life. Mark, thank you for having me, for dinners and beers, and for allowing me to sway you in favor of O-GlcNAc's importance. Sarah, I can't even begin. Thank you for letting me stay with you, thank you for showing me around Britain, thank you for all the tea and conversations. Thank you for your insight, for your immunological knowledge, for your constant support. Most importantly, thank you for your friendship. [And thank Richard, too!]

Of course, I must thank my committee members: Drs. Ian Harrison, James Demas, Rebecca Pompano, Cameron Mura, and MD Kenneth Tung. Thank you so much for taking the time to read my dissertation and assist in this process. I also have to thank Drs. Mura, Harrison, and Demas specifically for being a part of my candidacy exam. Thank you for your thoughtful questions… and for passing me. That really made this whole thing possible. Additionally, I want to thank Dr. Harrison for the ARCS fellowship nomination. Receiving the fellowship allowed me to visit and study with my collaborators, which assisted infinitely in my education and research.

Then there are the numerous Hunt lab members that have served to make this experience memorable and educational. First, I want to thank the former Hunt-labbers,

especially Dr. Jeremy Balsbaugh, since a very insightful conversation with him at recruitment weekend landed me in this laboratory. I need to thank Dr. Josh Nicklay for his very astute advice, which was simply: "Jeff." I have to thank Dr. Aaron Bailey for all of his histone guidance, which essentially drove my candidacy work and discoveries. Many thanks go to Dr. Jessica Chapman-Lim for showing me how to make columns, which really made all of my work possible. I could not possibly forget to thank Dr. Sushmit Maitra and Dr. Andrew Dawdy for their work with O-GlcNAc enrichment and for their eccentric teaching methods. Dawdy, thank you for every quote I still have on my white board. Thanks to Dr. Joe Strukl for being a constant source of optimism, as there were so many times I pulled strength from his kind words. I must also thank Dr. Jenn Abelin, for her friendship and mentorship throughout my time in the Hunt lab. Jenn, thank you for taking me under your wing from day one, for training me on everything IMAC-related, and for going on countless walks to let off steam. I will always appreciate my partner in crime. Additionally, thank you to other members including: Dr. Patrick James, Dr. Michelle English, Dr. Jenny Cottine, Dr. Erin Jeffery, Dr. Lissa Anderson, Dr. Lichao Zhang, Dr. Weihan Wang, and Marshall Chaffee, for their various teachings and support throughout the years.

The current Hunt lab members also deserve thanks. In particular, I must thank the other half of the "two best friends anyone could have." Amanda, you have been a ray of sunshine on days it seems impossibly cloudy. I have to thank you for being my desk-mate, my classmate, and my friend. I also have to thank you for the infinite amount of laughs we've shared over the last 4+ years. I also have to thank Paisley, who is recently engaged. Does everyone know she's engaged? She's engaged. Paisley, I have to thank

you for all the laughs, for listening to me sound off, and for being a great friend. And for keeping your bronchitis away from me. Scott, I want to thank you for all the not-so-subtle under-the-table texts that really make group meeting tolerable. I can't forget about Dr. Dina Bai, either. Dina, thank you for being such a great listener and the best HR representative a girl could ask for. Thank you for editing my dissertation, and thank you for your support and encouragement over the years. Additionally, I want to thank Ben Barnhill, Stephanie Miller-Lehmann, Ellen Speers, and Garrett Tanner. Overall, the Hunt lab members have made my experience amazing; one that I will never forget.

Before I got to Charlottesville, I spent five long years in Ann Arbor at the great University of Michigan. There, I met some people that have greatly influenced me as a person and as a scientist. I owe a ton of thanks to my "biochem group" Nichole Hansen, Pradeep Poonnenn, and Admir Mesanovic. Without you guys, I'm not sure I could have made it through biochemistry, and I'm absolutely positive I would not have survived analytical chemistry. Thank you for your enduring friendships and your help along the way. I also must thank all of the people that I worked for/with at the Protein Structure Facility at U of M. First, to Dr. Henriette Remmer, who took a chance on a college freshman and kept me around for five years. Secondly, thanks to Jen Fisher (now Page), a great friend and teacher in my years at the PSF. I must also thank the various people who came through the doors and influenced my work, including: Amber Peariso, Chris Schneider, Al Webb, Katy Hout, and Will Colin. And finally, I must thank my amazing fraternity, Alpha Chi Sigma (AB), for being a source of infinite happiness and craziness in my time there. I am lucky to have so many amazing brothers.

I have to thank my friends in Charlottesville and in the chemistry department, since without them I doubt I would have made it through graduate school. To begin, I want to thank my kickball team. All my roommates, plus Ambria, Aaron, Josh, Shannon, Brittany, Putnam, Biz, Bubba, Lindz, Will, Rob, Will, Merrick, Monica, Sammy, Bri, and Katy: thank you for making Sundays something I look forward to. And a special thanks to the championship team, since that was one of the best days I've had in Charlottesville. I also want to thank Taylor, as even though he moved 10 hours away he is still a daily source of laughs and feels just as close as when he was here.

Then I need to thank my current/former roommates, my closest friends in town, and their significant others. Mike/Bro Kid/Yellow Mike/Regular Mike: thank you for being a big brother, for your relentless teasing, and for your (usually) sound advice. I have to thank the Nguyens (Janet and Dr. Nguyen, especially) for being a second family while in Virginia. Thanks for making me feel like a part of the family while I am so far away from my own. Thanks to my pseudo-roommate Michelle, a sweetheart by all accounts. Thank you for being a great friend, and for inspiring the rest of us to be more social. Thank you to Chad Bro Chill and Garrett G-rock Tanner for being the comedic relief that I often need, and thank you for imparting on me a vocabulary that I never could have dreamed of myself. Of course, thank you to Mike/White Mike/Whitey for infinite amounts of coffee, organic chemistry advice, and hours of sounding off about how much my project sucks. [An aside: thank you to Professor MacDonald who unknowingly provided me with a lot of DMF.] Thank you to Mike's other half, Vicky, for being a great voice of reason in almost any situation. I must thank the both of them for helping me shed my Lean Cuisine addition, teaching me how to cook, and having me over for home-

cooked family dinners on holidays. Finally, I have to thank Dan for putting up with me for over the last three years. Dan, I doubt I would have made it through graduate school without you. Thank you for all your support, your love, and your friendship. Thank you for talking sense into me when I need it most. Thank you for every hug on the bad days, and thanks for every cheers on the good ones. All of you have become family to me, and I feel lucky to have you in my life. I cannot thank you enough for everything.

Sometime during my first year of graduate school, an email chain started with three (and then four) of my best friends from college. To date, we have now sent over 5,600 emails to each other. I have to thank each member of the "chain gang," as they have supported me through my entire graduate experience. To all of you, I cannot thank you enough for everything you've done for me. Thank you for your love, your friendship, for yelling at me when it was necessary, and for being on my side when it wasn't. Kim, thank you for getting married, because without that Oklahoma wedding, the chain wouldn't have started! Serb and Janell, thank you (obviously) for the greatest 25th birthday celebration in the history of the world. Thank you for the daily laughs, conversations, and voices of reason. Lauren, thank you for being the best friend a girl could ask for. Thank you for being my rock. Thank you for always having my back, no matter what. Thank you for being someone I constantly look up to. I love you all tremendously and truly cherish your friendship.

Unquestionably, I must thank my amazing significant other. Derrick, in a short amount of time, you have drastically changed my life. Where do I even begin? Thank you for London. Thank you for Birmingham. Thank you for moving to Charlottesville and making this relationship possible. Thank you for sticking around, for hanging your things

and staying. Thank you for being the man of my dreams, just by being yourself. Thank you for being so proud of me, for being so supportive. Thank you for inspiring me to do things, especially those that make me (and us) better. Thank you for every laugh, for every gold star, for every song, for every wink, for every kiss, for every adventure. Thank you for always coming back to me, for loving me entirely. And finally, thank you for inviting me into your story. I hope to continue it for years and years to come. I look forward to the day we go dancing. I love you to the moon and back.

Last, but definitely not least, I have to thank my wonderful and supportive family. Thank you to all of my extended family. All of my cousins, aunts, and uncles have showered me with support and love as I went through this process. Then I have to thank the people closest to me. First, to my Ry-Ry. As the dedication says, you have taught me so much more than any teacher ever could. Thank you for your truly unconditional love, for all the dances in the kitchen, for every time you say, "I love you Dace, I miss you Dace. I love you more! I miss you more!" Thank you for being the best brother a little sister could ask for. Second, thank you to my amazing little sister, Megan/Meg/ Geeg/Geegjumah/MeggieMoo/Moodums. Thank you for being my confidante, someone I can always count on and complain to. Thank you for acting like my big sister sometimes, doing my make-up and picking out my clothes. Thank you for cuddles and your dumb cheek-kisses and your love. Thank you for being a great friend. Next, I have to thank my dad. Daddy, thank you for pushing me to do my best, always. Thank you for being an inspiration and an example. Thank you for the nearly 9 years of sobriety. Thank you for reaching out to me, and thank you for allowing us to have a fresh start. Thank you for being so proud of me. Even though I pretend to hate it, I am secretly appreciative every

time you brag about me. And finally, thank you to my momma. Momma bear, I don't know what I would do without you. You are a pillar of strength and love, one that I constantly lean on. Thank you for every phone call, for listening to every complaint, for every day you listened to me talk about science and understood none of it. Thank you for being not only my mother, but my best friend. Thank you for all the laughs, the hugs, the tears, and the love. I love all of you so much it makes my heart ache. Thank you for everything.

Table of Contents

Chapter 1

1 Introduction to the Dissertation

Chapter 2

2 Identification of HLA-associated Phosphopeptides from Esophageal Adenocarcinoma

    2.1 Introduction

Chapter 3

3 New Methodology for the Enrichment and Characterization of O-GlcNAcylated Peptides

3.1 Introduction

3.2 Materials

3.3 Methods

Chapter 4

4 Enrichment and Identification of HLA-associated O-GlcNAcylated Peptides from Various Cancers

Chapter 5

List of Figures and Tables

Chapter 1

Chapter 2

Chapter 3

Chapter 4

Chapter 5

<u>Abbreviations</u>

°C   degrees Celsius

•   radical species

A   solvent A, 0.1 M acetic acid in water

AC   analytical column

AD   Alzheimer's disease

AIDS   acquired immune deficiency syndrome

AIIP   phosphorylated angiotensin II phosphate

ACN   acetontrile

AcOH   acetic acid

ACT   adoptive cell transfer therapy

AGC   automatic gain control

AL   aldehyde

Ala, A   alanine

ALL   acute lymphoblastic lymphoma

Ambic   ammonium bicarbonate

AML   acute myeloid leukemia

amu   atomic mass unit

angio   angiotensin I (DRYIHPFHL)

APBA   amino phenylboronic acid

Arg, R   arginine

Asn, N   asparagine

Asp, D   aspartic acid

ATP   adenosine triphosphate

AIRE   auto immune regulator

B   solvent B, 0.1 M acetic acid in 70% acetonitrile

Bar   100,000 Pascals; 14.5038 pounds per square inch

BATDA   bis(acetoxymethyl) 2,2':6',2"-terpyridine-6,6"-dicarboxylate

BE   Barrett's Esophagus

BEMAD   β-elimination Michael addition chemistry

BLAST   Basic Local Alignment Search Tool

BSA   bovine serum albumin

BTTAA   2-[4-({bis[(1-tert-butyl-1H-1,2,3-triazol-4-yl)methyl]amino}methyl)-1H-1,2,3-triazol-1-yl]acetic acid

BTTP   3-[4-({bis[1-*tert*-butyl-1H-1,2,3triazol-4-yl)methyl]amino}methyl)-1H-1,2,3triazol-4-yl]propanol

C18   C18 bonded silica stationary phase

Ca^{+2} calcium (II) cation

CAD   collision activated dissociation

CAR   chimeric antigen receptor

CD8+   cluster of differentiation 8; T-cell presenting surface glycoprotein CD8

CDR   complementarity determining reunions

CEq   cell equivalents, number of MHC-associated peptides from equivalent number of cells

CHAPS   3-[(3-cholamidopropyl) dimethylammonio]-1-propane sulfonate

CLL   chronic lymphocytic leukemia

cm   centimeter

CO_2 carbon dioxide

CRM   charged residue model

C-term   C-terminus of peptide

CTL   cytotoxic T lymphocyte

$CuSO_4$ copper sulfate

Cys, C   cysteine

Da   Dalton, 1 amu

DCM   dichloromethane

D-MEM   Dulbecco's Modified Eagle Medium

DMF   dimethylformamide

DMSO   dimethylsulfoxide

DNA   deoxyribonucleic acid

DTT   dithiothreitol

EBV   Epstein-Barr virus

EDC   1-Ethyl-3-(3-dimethylaminopropyl)carbodiimide

EDTA   ethylenediaminetetraacetic acid

EGF   epidermal growth factor

ELISpot   enzyme-linked immunospot assay

EMT   epithelial-to-mesenchymal transition

ER   endoplasmic reticulum

ERBB4   Receptor tyrosine-protein kinase erbB-4

ESI   electrospray ionization

ETD   electron transfer dissociation

eV   electron volt, 1.602×10^{-19} J

FBS   fetal bovine serum

Fe^{2+} iron (II) cation

Fe^{3+} iron (III) cation

$FeCl_3$ iron (III) chloride

FETD   front-end electron transfer dissociation

FHIOSE   Simian virus 40 transformed nontumorigenic ovarian cell line

FMOC   fluorenylmethyloxycarbonyl chloride

fmol   femtomole, 1×10^{-15} moles

FT   Fourier transform

FTICR   Fourier transform ion cyclotron resonance

FT-MS   Fourier transform mass spectrometry

g   gram

gal   galactose

GalNAz   N-azidoacetylgalactosamine

GalT1   galactosyltransferase

Gln, Q   glutamine

Glu, E   glutamic acid

Gly, G   glycine

h, hr   hour

HBP   hexosamine biosynthetic pathway

HCl	hydrochloric acid
HCL	hairy cell leukemia
HD	healthy donor
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
His, H	histidine
HLA	human leukocyte antigen
HPLC	high performance liquid chromatography
IAA	iodoacetamide
ICS	intracellular cytokine staining
IDA	iminodiacetic acid
i.d.	inner diameter
IL2,10	interleukin 2 or 10
Ile, I	isoleucine
IMAC	immobilized metal affinity chromatography
IFN-γ	interferon gamma
ITMS	ion-trap mass spectrometer
JY	Epstein-Barr virus immortalized B-cell line
Kasil	potassium silicate solution
kDa	kiloDalton ($1x10^3$ Da)
K_{eq}	association constant
kV	kilovolt ($1x10^3$ V)
L	liter

LC	liquid chromatography
Leu, L	leucine
LOD	limit of detection
LTQ	linear ion trap quadupole
Lys, K	lysine
m	mass, meter
M	molar (moles/liter)
m/z	mass to charge ratio
M+nH	molecular ion with n charges
mAbs	monoclonal antibodies
MeOH	methanol
MES	2-(N-morpholino)ethanesulfonic acid
Met, M	methionine
MHC	major histocompatibility complex
min	minute
mg	milligram ($1x10^{-3}$ g)
µg	microgram ($1x10^{-6}$ g)
mL	milliliter ($1x10^{-3}$ L)
µL	microliter ($1x10^{-6}$ L)
mm	millimeter ($1x10^{-3}$ m)
µm	micrometer ($1x10^{-6}$ m)
mM	millimolar ($1x10^{-3}$ M)

mol	mole, 6.022×10^{23} molecules
MS	mass spectrometry
MS1	full mass spectrum
MS2	tandem mass spectrum
MW	molecular weight
MWCO	molecular weight cut off
NAAG	N-acetyl-aspartyl-glutamate
NaCl	sodium chloride
NaCNBH$_3$	sodium cyanoborohydride
NCOR2	nuclear receptor co-repressor 2
NHS	N-hydroxysuccinimide
NK	natural killer cells
nL	nanoliter (1×10^{-9} L)
nmol	nanomole (1×10^{-9} moles)
NSB	non-specific binding
N-term	amino-terminus of peptide
o.d.	outer diameter
OEAC	esophageal cancer; adenocarcinoma
OESCC	esophageal cancer; squamous cell carcinoma
OGA	O-GlcNAcase
O-GlcNAc	O-linked β-N-acetylglucosamine
OGT	O-GlcNAc transferase

OMSSA	Open Mass Spectrometry Search Algorithm
PAGE	polyacrylamide gel electrophoresis
PBA	phenylboronic acid
PBMC	peripheral blood mononucleated cells
PC	pre-column
PHA	phytohemagglutinin
Phe, F	phenylalanine
pmol	picomole ($1x10^{-12}$ moles)
PMSF	phenylmethylsulfonyl fluoride
PNGaseF	peptide *N*-glycosidase F
ppm	parts per million
Pro, P	proline
psi	pounds per square inch
PTM	post-translational modification
PVDF	Polyvinylidene fl023roride
RB	retinoblastoma
ROBO1	roundabout homolog 1
RP	reverse phase
RT	room temperature
s, gS	O-GlcNAcylated serine
s, pS	phosphorylated serine
s, sec	seconds
scFv	single chain variable fragment

SDS	sodium dodecyl sulfate
SELH	selenoprotein H
sHLA	soluble human leukocyte antigen
Ser, S	serine
SKOV3	human ovarian cancer cell line
STS	synthetic phosphorylated peptide (STsLVEGR)
t, gT	O-GlcNAcylated threonine
t, pT	phosphorylated threonine
T-ALL	T-cell acute lymphoblastic leukemia
TAP	transporter associated with antigen processing
TBTA	tris[(1-benzyl-1H-1,2,3-triazol-4-yl)methyl]amine
TCA	tricarboxylic acid
TCR	T-cell receptor
TDA	2,2':6',2''-terpyridine-6,6''-dicarboxylic acid
TFA	trifluoroacetic acid
TGFβ	transforming growth factor β
Thr, T	threonine
TIC	total ion current
TIL	tumor infiltrating lymphocytes
TNFα	tumor necrosis factor α
T$_{reg}$	regulatory T-cells
Tris	tris(hydroxymethyl)aminomethane HCl

Trp, W	tryptophan
Tyr, Y	tyrosine
Ub	ubiquitin
UDP	uridine diphosphate
UV	ultraviolet
V	volt
Val, V	valine
Vaso	vasoactive intestinal peptide (HSDAVFTDNYTR)
VEGF	vascular endothelial growth factor
WGA	wheat germ agglutinin
X	any amino acid
z	charge

Abstract

Phosphorylation and O-linked β-N-acetylglucosamine (O-GlcNAc) are post-translational modifications (PTMs) that modulate cellular functions through extensive cross-talk with signaling cascades. As such, these PTMs are directly involved in the regulation of many cellular processes, and dysregulation of these pathways can result in malignancy. Changes in protein expression and/or metabolism that accompany cellular transformation can be conveyed to the immune system via the MHC class I processing pathway. This system involves intracellular proteins that get degraded to 8-10 amino acid peptides. These peptides are then loaded onto MHC-class I molecules and presented onto the cell surface. If the peptide is antigenic, circulating CD8+ T lymphocytes can bind to the MHC-peptide complex and stimulate an immune response. Since dysregulated cell signaling in malignant cells leads to aberrant enzymatic modification of proteins, we believe transformed cells will present post-translationally modified peptides unique to cancer. We aim to identify tumor-specific peptides on the cell surface that can stimulate immune action in order to develop specific cancer immunotherapeutics.

In this dissertation, we first identify HLA-associated tumor-specific phosphopeptides from the surface of esophageal tissues. We then demonstrate the ability of these peptides to generate a response in healthy individuals, suggesting that they may be excellent candidates for further study. We then go on to present the first HLA class I associated O-GlcNAcylated peptide antigens detected on the surface of tumor cells. Additionally, because these peptide antigens are estimated to be present at 0.1-1% of the total peptide pool (10-100 copies/cell), we develop an enrichment technique to help

characterize the peptides. This technique allows us to sequence a total of 19 O-GlcNAcylated peptide antigens present at the attomole level. Finally, we show that healthy individuals have T-cell responses against the peptide antigens located in the central memory compartment, suggesting that the O-GlcNAcylated peptides may be targets of cancer immune surveillance and could be of importance in the design of cancer immunotherapies. Ideally, these peptide antigens will be utilized in the development of specific cancer immunotherapies such as adoptive T-cell therapy, T-cell receptor gene transfer, and/or immune-mobilizing monoclonal TCRs. Ultimately, we believe that post-translationally modified peptides are key to unlocking this potential and can be exploited in the development of cancer immunotherapies that will eradicate this devastating disease.

Chapter 1: Introduction to the Dissertation

1.1 Cancer and the Immune System

Despite advances in cancer detection and treatment, cancer mortality is projected to continue rising worldwide with an estimated 13.1 million deaths in 2030.[1] Traditional cancer therapies, such as radiation and chemotherapy, are nonspecific and carry several deleterious side effects, thus demanding the need for more specific and effective treatments. Advancement in this attempt is both undeniably and critically tied to the human immune system. To be sure, a healthy individual's immune system can eliminate cells that become malignant. However, when the immune system is overwhelmed by increased disease load and/or immunosuppression, cells can transform and escape the immune system. Once they escape, the transformed cells are allowed to proliferate rapidly, eventually leading to the disease state recognized as cancer.[2] Understanding how these malignant cells are eliminated by the immune system before they result in cancer is key in developing more effective cancer treatments. Thus, we aim to develop a treatment that will specifically target transformed, or malignant, cells. To do this, we plan to harness the ability of the immune system to clear transformed cells from the body.

Cells are equipped with an efficient system of displaying their health status to the immune system, called the major histocompatibility (MHC) class I processing pathway. Here, worn-out proteins are marked for degradation by ubiquitin. These proteins are then unfolded and fed into proteasomes, where they are chopped up into short fragments and then either recycled into the cytosol or transferred to transporters associated with antigen processing (TAPs). The TAPs function to funnel the peptides into the endoplasmic reticulum (ER), where the peptides are shortened to 8-10 amino acid peptides and loaded

onto MHC class I molecules. The MHC-peptide complex is then transported through the Golgi membrane to the cell surface, where the peptides are displayed to circulating immune cells.[3-5] This pathway is summarized in Figure 1.1.[5]

Figure 1.1 MHC class I processing pathway. Proteins marked for degradation by Ub are fragmented into 8-10 amino acid peptides in the proteasome, then transported into the ER by TAP. The peptides are loaded onto MHC class I molecules and the resulting complex is presented on the cell surface for immune recognition.
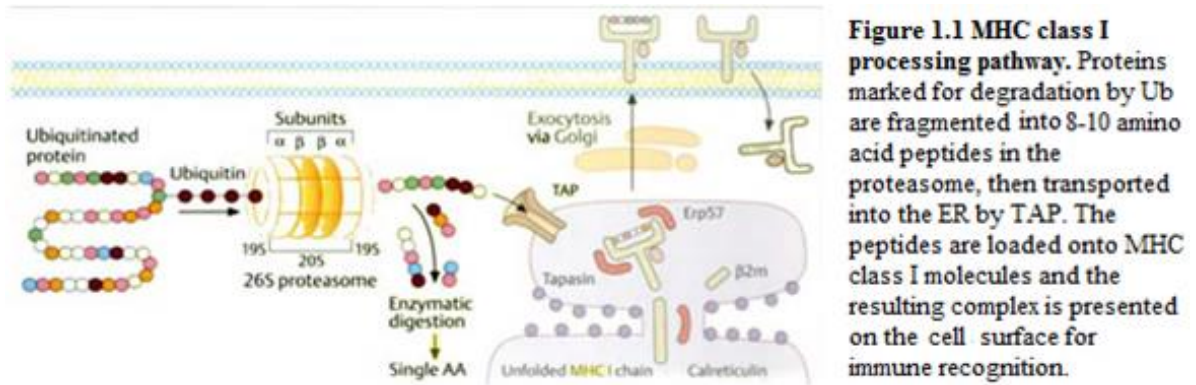
The human version of MHC class I is called the human leukocyte antigen (HLA) class I system, which is encoded by a highly polymorphic genomic region on chromosome 6. The classic HLA alleles, HLA-A, B, and C, are all encoded by this region. All of the class I molecules are heterodimers consisting of a transmembrane heavy α-chain associated with a smaller β2 microglobulin protein (Figure 1.2).[4,6]

Figure 1.2. Structure of HLA class I molecules. Left panel: tertiary structure of class I heterodimer, including 3 alpha chains and beta2-microglobulin. Right panel: Secondary structure of peptide binding groove. Interactions within this pocket define the amino acid binding motif.
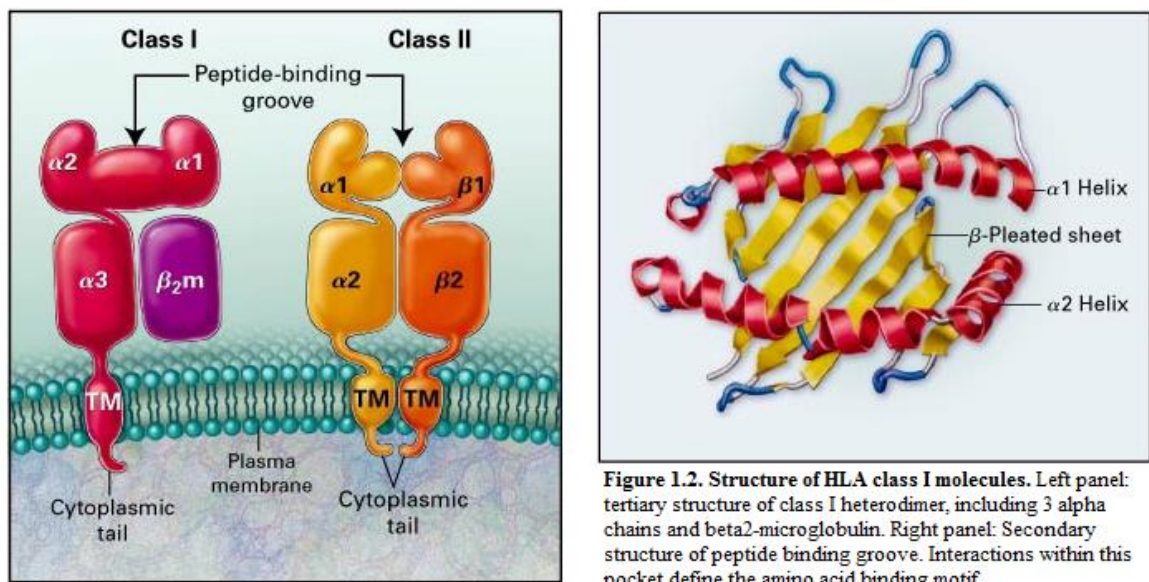
Variability between the different alleles occurs within the peptide binding groove of the

α-chain, which results in diverse peptide binding motifs as depicted in Figure 1.3.[4] The

HLA molecule relies on interactions with amino acids at positions 2 and 9 of the peptide,

which serve to anchor the peptide in the binding groove. Despite the limitations placed on

peptide binding by the allelic motifs, each class I molecule can bind approximately

10,000 peptides.[7] Additionally, each individual can express up to 6 different allotypes

(two of A, B, C), as one is inherited from each parent. Since each cell can present roughly

100,000-300,000 HLA molecules, each uninfected cell can display an extremely

heterogeneous population of peptides on its surface.[8]
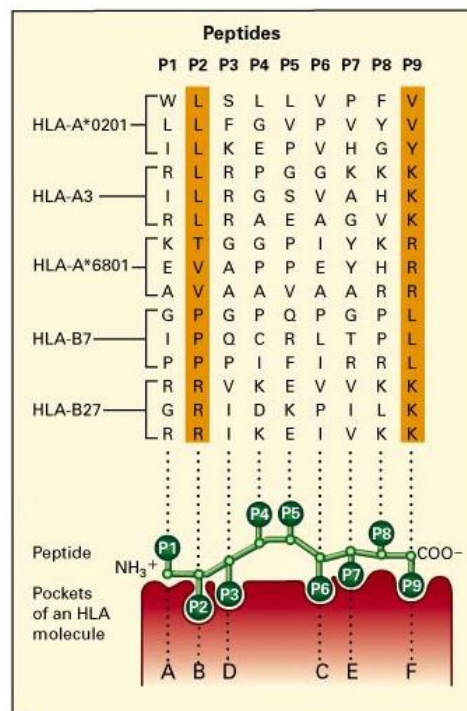
Figure 1.3. Common HLA binding motifs. The pocket created by the alpha chain of MHC class I molecule defines the peptide binding motif. Interactions occur between the pocket and peptide anchor residues, generally at positions 2 and 9. Importantly, residues 4 and 5 are often solvent exposed for interaction with circulating T-cells. Shown are common HLA alleles and their corresponding peptide binding motifs.

Residues 3 through 8 of the HLA-associated peptide are solvent exposed and

capable of forming interactions with circulating CD8+ cytotoxic T-cells (CTLs).[9-11]

These members of the adaptive immune response are extensively trained to differentiate

between self and non-self peptides via thymic education.[12-15] This process begins as progenitor T-cells, also called thymocytes, migrate from bone marrow to the thymus, where they start to divide and differentiate. Concurrently, genes encoding the T-cell receptor (TCR) rearrange to generate clones with different specificities capable of recognizing up to 10^8 distinctive antigens.[12] After TCR rearrangement, T-cell selection occurs, which first involves a round of positive selection. Here, thymocytes are given the opportunity to bind to HLA-peptide complexes. Those that cannot interact die by neglect, while those that bind weakly to the complexes receive a signal that blocks the pathway to apoptosis. Positive selection eliminates over 95% of all thymocytes.[13,14] Subsequently, negative selection occurs, which functions to prevent the formation of self-reactive T-cells. Two types of cells test the thymocytes for tolerance of self. The first, thymic dendritic cells, assess reactivity to HLA-peptide complexes presented on their cell surfaces. These HLA-peptide complexes represent shared proteins that all cell types present. The second, medullary thymic epithelial cells expressing the autoimmune regulator (*AIRE*) gene, express tissue-specific proteins and present them to create a "self-shadow" in the thymus. T-cells that react to self-antigens presented by these two types of cells are deleted by induction of apoptosis. Cells that survive both rounds of selection (approximately 3% of progenitors) are released into the secondary lymphoid organs as mature, naïve T-cells.[15]

The naïve T-cells are then primed by professional antigen presenting cells (APCs) such as dendritic cells and macrophages.[16] Once this occurs, they can become activated by cells displaying peptides that originate from microbes, viruses, and/or aberrant proteins found in cancer, thereby allowing them to distinguish between self and non- or

altered-self. Upon exposure to peptide antigens, properly educated CTLs release cytokines, which are small proteins that affect the behavior of other cells. Specifically, CTLs emit perforin, granzymes, and granulysin that stimulate apoptosis of the target cell while recruiting other immune cells to the infected area.[15,16]

The idea that CTLs can identify transformed cells underlies the immunoediting theory, which suggests that the immune system harnesses the ability to eliminate cancer before it becomes a recognizable tumor. According to Schreiber *et al* (2011) immunoediting has three phases: elimination, equilibrium, and escape (Figure 1.4).[2] The first, elimination, occurs when the innate and adaptive immune system work together to detect and destroy cancer. Experimental evidence for this was first described in the 1970s when mice were successfully immunized with irradiated tumor cells. When the same mice were injected with live cells from an independently induced tumor, the tumor cells proliferated, suggesting an immune response to a specific type of tumor.[17] This experiment has been corroborated by a number of investigations that show earlier onset of tumors or neoplasia in mice lacking certain immune subsets and/or immunodeficiency.[18-20] Understanding how the immune system can identify and eliminate transformed cells is of utmost importance in developing an effective immunotherapeutic.

The second phase of immunoediting, equilibrium, occurs when the immune system is unable to clear the transformed cells from the body completely. Instead, the adaptive immune system prevents tumor growth and maintains residual tumor cells in a state of dormancy.[4] An example of experimental evidence for this comes from mice treated with carcinogens that harbor cancer cells without having apparent tumors. When

these mice are treated with monoclonal antibodies (mAbs) against T-cells and/or certain cytokines, tumors appear rapidly at the injection site.[21] Additionally, equilibrium has been observed in humans, especially in immunodeficient individuals such as AIDS patients or organ recipients. For instance, a patient with melanoma that had been in remission for 16 years donated both kidneys post-mortem to two individuals. Both of the immunodepleted recipients developed melanoma, suggesting that the donor's immune system was maintaining the cancer in a dormant state, and the lack of immunity in the recipients allowed for the melanoma to grow into clinically apparent cancer.[22]

The last stage of immunoediting, escape, transpires when transformed cells acquire the ability to circumvent the immune system and emerge as rapidly growing tumors.[4] This is related to a number of hallmarks of cancer, including a loss of MHC class I molecules, genetic instability, and immunoselection.[23,24] Tumors establish an immunosuppressive state, recruiting regulatory T-cells (T_{reg}) that produce cytokines such as interleukin-10 (IL10) and transforming growth factor β (TGFβ) which downregulate anti-tumor CTL responses.[25] Tumors can also construct a physical barrier to prevent
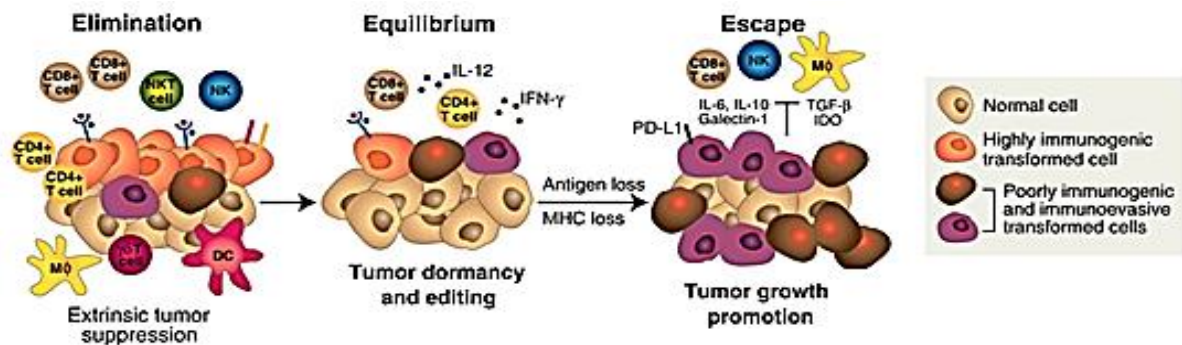
Figure 1.4 The immunoediting theory. Elimination: tumor cells are recognized by the innate immune system (natural killer cells) which secrete cytokines that activate dendritic cells, which then initiate an adaptive immune response. Equilibrium: maintenance of adaptive immune cells at the site of transformation without actively growing tumors. Escape: evasion of the immune system and rapidly dividing transformed cells.(From reference 4)

detection by immune cells.[26] Once the tumor acquires the ability to evade CTL detection

by any or all of these techniques, the tumor is allowed to expand rapidly.[24]

Overall, it is apparent that cancer and the immune system are intimately related,

even to the extent that patient immune responses can predict patient prognosis better than

current staging measures. The question remains, then, how CTLs can identify

transformed versus normal cells, especially since the presented peptides originate from

the host. The answer can be found within the hallmarks of cancer, defined by Hanahan

and Weinberg (2000, 2011), which describe the multistep development of tumors. These

include: sustaining proliferative signaling, evading growth suppressors, avoiding immune

destruction, enabling replicative immortality, inducing angiogenesis, genomic

instability/mutation, resisting cell death, and deregulating cellular energetics.[23,24] These

hallmarks allow cells to acquire traits that render them tumorigenic and eventually

malignant.

Underlying each of these hallmarks is the ability to change the internal cellular

environment, which will ultimately alter the repertoire of peptides presented on the cell

surface for recognition by CTLs.[27] For instance, genomic instability lends itself to genetic

mutation.[23] Failure of the tumor cell to repair DNA damage can result in mutated

proteins, and, consequently, presentation of mutated peptides on the surface of tumor

cells. Additionally, several of the hallmarks require an upregulation of proteins that may

not be expressed as often in a normal cell. For example, angiogenesis requires the

upregulation of vascular endothelial growth factor (VEGF) to increase vasculature within

a tumor, which allows for increased blood flow and growth in the tumor.[24] The increased

copy number of VEGF may be represented on the cell surface at highly abnormal levels.

Finally, and arguably most importantly, tumors acquire the ability to sustain chronic proliferation. Underlying this capacity is dysregulated cell signaling, which is very tightly controlled in normal cells and relies on post-translational modifications (PTMs) to mediate transcription and translation of cell cycle proteins.[24] In transformed cells, the PTMs are upregulated, occurring more often and/or on amino acids they may not normally modify. These modified proteins can be presented on the cell surface, resulting in an abnormal repertoire of peptides on the surface of the tumor cell.[8,27,28]

The overarching purpose of this work is to identify tumor-specific post-translationally modified peptides for use in cancer immunotherapeutics. We believe that these peptides are the best immunotherapeutic candidates for a number of reasons. First, a hallmark of cancer is dysregulated cell signaling, which leads to aberrant post-translational modifications that can be displayed by the HLA-class I system. These PTMs will occur more frequently on the surface of tumor cells and we can sequence these tumor-specific peptides by high-resolution mass spectrometry.[8] Further, proteins containing PTMs may undergo antigen processing differently, so that the resulting HLA-peptide may be different from that tolerized during thymic education.[29] This idea is corroborated by the fact that many post-translationally modified peptides have demonstrated an ability to stimulate a T-cell response that is specific for the modification.[8,30] Without central tolerance, the immune system should mount a substantial response against the modified peptides. Together, this leads us to believe that tumor-specific modified peptides will be excellent candidates for immunotherapy.

1.2 Introduction to Methods

1.2.1 HLA-associated peptide isolation (Performed by SAP, University of Birmingham)

To identify HLA-associated peptides, cancer cells from cell lines or tissues must be harvested and lysed. Then, HLA class I molecules are immunoaffinity-purified from cell lines or tissues and their associated peptides are extracted. 10^8-10^{10} cells or cell equivalents are lysed in CHAPS buffer and the lysate is subject to centrifugation. Supernatants are passed over protein A-sepharose preloaded with either HLA-A2 specific antibody (BB7.2), HLA-B7 specific antibody (ME1), or HLA-A/B/C nonspecific antibody (W6/32). Peptides remain within the HLA binding groove during this process, and can be eluted from the purified class I molecules with 10% acetic acid. This disrupts the non-covalent interactions between the HLA molecule and the bound peptide, allowing for separation by ultrafiltration with a molecular weight cut off (MWCO) filter.[8,31] The peptide mixture is then subjected to HPLC-ESI-MS/MS analysis to determine total peptide content per 10^7 cell equivalents, which allows us to determine the amount of sample to use in the next step of analysis.

1.2.2 Post-translational modification enrichment

Peptides with PTMs are estimated to be present at approximately 0.1-1% of the total HLA-associated peptide population.[28] To put this in perspective, if 10^5 peptides exist on the surface of a given cell, only 1-10 copies of a modified peptide may be present. Due to their extremely low abundance coupled with the large complexity and number of unmodified peptides present on the surface of cells, enrichment procedures are necessary to isolate the post-translationally modified peptides from the pool of unmodified peptides.

Given that this dissertation deals with multiple modifications, the individual enrichment

methods will be extensively discussed in subsequent chapters.

1.2.3 HPLC-ESI-MS/MS

The peptides isolated from enrichment are then separated by reverse-phase HPLC,

which operates on the basis of hydrophobicity. Hydrophobic peptides interact strongly

with the stationary phase and are retained longer than hydrophilic peptides over the

course of a reversed phase gradient. The elution of an enrichment is loaded onto an

irregular C18 (5–20 μm diameter) capillary precolumn (360-μm outer diameter, 75-μm

inner diameter) and washed with 0.1% acetic acid before connecting the precolumn to a

C18 (5 μm diameter) analytical capillary column (360-μm outer diameter, 50-μm inner

diameter) equipped with an electrospray emitter tip.[32] Peptides are gradient eluted by an

increasing concentration of organic solvent from 100% solvent A (0.1M acetic acid in

water), 0% solvent B (70% acetonitrile and 0.1M acetic acid) to 60% solvent B over a 60-

120 minute time period. In each chromatographic separation, commercially available

peptides (angiotensin and vasoactive peptide) are spiked in at known concentrations as

internal standards. The peak areas of these standard peptides are compared with peptide

peak areas to obtain a relative measure of the amount of peptide present.[32]

Coincident with elution from the column, ions are formed via electrospray

ionization. The solvent line which introduces the mobile phase to the HPLC column is

split to reduce flow rate from 0.2 mL/min to 60 nL/min. A 2 kV voltage is applied to the

waste line, which provides the electrospray emitter tip with a significant electric

potential.[33] This causes a build-up of positively charged ions at the tip that migrate to the

counter electrode, the inlet of the mass spectrometer (held at 0V). Movement of ions toward the inlet is counterbalanced by the surface tension of the liquid, giving rise to a Taylor cone. Since the voltage applied is sufficiently high, a cylinder extends from the cone until the Rayleigh instability limit is reached. Once this occurs, solvated ion droplets begin to break away.[34] Although the mechanism of complete ionization is disputed, it is generally believed that peptides follow the Charge Residue Model (CRM), which suggests that the droplets begin to shrink due to solvent evaporation until the charge density becomes too high. At this point, the Coulombic charges exceed the surface tension, and the droplet "explodes" into smaller and more highly charged droplets. This continues until the droplet contains only one solute molecule, which becomes gaseous while retaining the positive charge imparted on it.[35] At this point, ions are focused with RF-only quadrupoles to a high resolution mass analyzer.

Two stages of mass analysis occur in the mass spectrometers – a full-scan high-resolution MS1 that identifies the precursor masses present at a given time of the chromatographic elution, and a fragmentation MS2 that allows for sequencing a mass selected parent peptide. The first half of this is accomplished by focusing a packet of ions into either an FT-ICR or an Orbitrap mass analyzer. In an FT-ICR, ions held in a magnetic field absorb energy from an excitation pulse and precess in phase-coherent packets; the frequency of orbit is characteristic to the ions' mass-to-charge (m/z) ratios.[36,37] In an Orbitrap, ions are trapped in a purely electrostatic field around a central electrode and move in the axial direction with a frequency related to their m/z.[38,39] In both types of instruments, the frequencies are obtained from an image current induced

between the ions and outer electrodes, are converted to m/z, and give rise to a mass spectrum (MS1) using fast Fourier Transform.

The most abundant precursor masses present in the MS1 are then selectively fragmented in a linear ion trap (LTQ)[40] either by collision activated dissociation (CAD)[41] or electron transfer dissociation (ETD)[42,43]. In CAD fragmentation, the parent ion, which is protonated randomly along the peptide backbone, is subjected to low energy collisions with helium atoms. Multiple collisions occur, transferring translational energy to the peptide in a stepwise fashion. The translational energy is converted to vibrational energy, which is dispersed throughout all covalent bonds. When the vibrational energy exceeds the activation energy necessary to break the weakest bond, fragmentation occurs.[41] The proposed mechanism for CAD is shown in Figure 1.5. A peptide sequence can be determined when backbone fragmentation allows for assignment of complementary b- (N-terminal) and y- (C-terminal) type ions. The mass difference between members of the same series permits identification of a particular amino acid. CAD works best for short (<15 residues), low charge (z<+3) peptides. Because CAD is dependent on random backbone protonation, longer peptides with higher charge states tend to display suboptimal fragmentation since areas of high charge limit proton transfer along the backbone. Additionally, the most labile bond is preferentially cleaved during CAD fragmentation.[42] Since many PTMs are highly labile, the fragmentation spectrum of a post-translationally modified peptide will be dominated by the loss of the modification from the parent ion. This makes sequencing and site-localizing a modified peptide fairly difficult. However, the ions generated from PTM neutral losses can be used for diagnostic purposes, and will be discussed in more detail later.
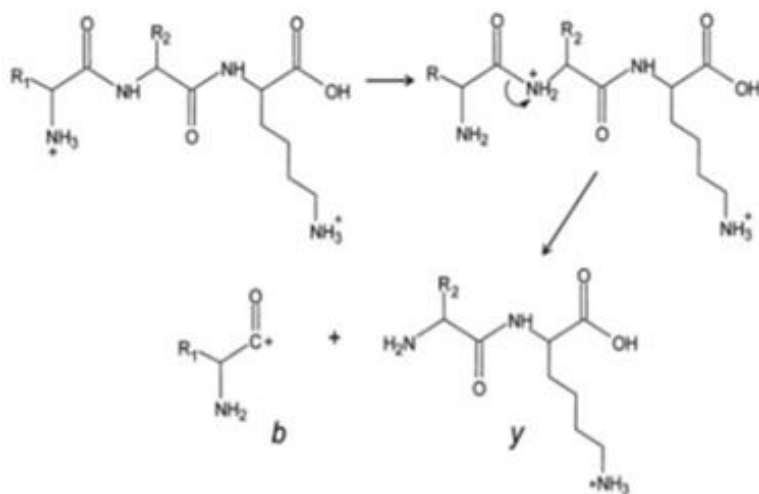
Figure 1.5. Proposed mechanism for CAD fragmentation. Multiple collisions with helium excite precursor ions to higher energy states through the conversion of translational energy to vibrational energy. This leads to fragmentation of the amide bond, resulting in the formation of b- and y-type ions that contain the N- and C-termini, respectively.

ETD, developed in the Hunt Laboratory, fragments peptides by transferring an electron from a radical anion to a protonated peptide. Capture of this thermal electron (<0.2 eV) is exothermic by ~6 eV and causes random fragmentation of the peptide backbone at the $N-C_\alpha$ bond. The proposed mechanism for ETD is shown below (Figure 1.6). This reaction creates fragments analogous to b/y type ions, but instead are called c-(N-terminal) and z•- (C-terminal) type ions, which are again used to assign peptide sequences. The ETD mechanism causes more backbone cleavages than CAD, resulting in higher sequence coverage of longer, highly charged peptides. Also, since ETD maintains labile bonds, PTMs are preserved and site-localization of the modified residue can be confidently assigned.[42]

Both CAD and ETD provide important information when attempting to sequence and site-localize the HLA-associated post-translationally modified peptides. Thus, both techniques are used to fragment parent ions to obtain as much sequence information as

possible. In most cases, the five most abundant parent ions present in the MS1 are

subjected to CAD and ETD fragmentation.

Figure 1.6 Proposed mechanism for ETD fragmentation. Capture of an electron by a backbone carbonyl leads to the formation of a radical anion, which results in the cleaveage of the N-C(alpha) bond. This creates complementary c- and z- type ions that contain the N- and C- termini, respectively.

The sensitivity allowed by enrichment technology followed by high-resolution

mass spectrometry is unparalleled by any other analytical method. These techniques

allow us to regularly detect attomole levels of peptide material, which correspond to

approximately 1-10 peptide copies per 10^9 cells.[43] This permits the detection, sequencing,

and site-localization of extremely low level analytes.

1.2.4 Data processing

After a data file is acquired, the MS2 spectra are searched against existing

databases such as RefSeq or Uniprot using the Open Mass Spectrometry Search

Algorithm (OMSSA).[44] This program takes experimental spectra, filters noise peaks,

extracts charge-to-mass (m/z) ratios, then compares the m/z values to those derived from

peptides produced by an *in silico* digestion of a protein sequence library. The resulting

search hits are then statistically scored; the peptide matches having the lowest score have the lowest probability of occurring randomly. OMSSA also allows for variable and static chemical modifications on the peptide. In this way, we can search through those peptides that have post-translational modifications. Candidate peptides that have a high probability of being correctly identified by OMSSA are manually verified using accurate mass (± 5 ppm), CAD, and ETD spectra.

Occasionally, OMSSA fails to match a spectrum to a known database protein. This is often the case with low-level and/or post-translationally modified peptides that produce weak MS2 spectra. In these cases, we utilize neutral loss programs that identify signature ions for individual PTMs. We then characterize the peptide using a combination of *de novo* sequencing and bioinformatic technologies. For instance, when a partial sequence can be assigned, Mascot (matrixscience.com) can be utilized to assign the rest of the sequence. This program allows users to input intact and fragment ion masses in order to match a peptide in a protein database. These scores are also assigned by probability, whereby a high score indicates a low probability that the assignment was made by chance. These hits are also manually validated.

1.2.5 Biological testing (Performed by SAP, University of Birmingham)

Once the HLA-associated peptides are identified, they are synthesized using Fmoc chemistry, purified via HPLC to >90% purity, and confirmed by mass spectrometry. The synthetic peptides are then subjected to biological tests in order to determine if they can stimulate immune responses in healthy individuals. We use healthy donor (HD) blood because we believe that transformation events are common

occurrences, as predicted by the immunoediting theory. If this is the case, HDs should have an adaptive/memory response to the tumor-specific peptides. Thus, testing these peptides in HDs allows us to identify the best candidates for cancer immunotherapies. While several other screening techniques are available, we have mainly focused on two: enzyme-linked immunospot (ELISpot) assays[45,46] and intracellular cytokine staining (ICS).[47]

50 mL of blood is collected from each HD using syringes into lithium heparin solution, which helps prevent blood clotting. Blood is centrifuged and the plasma is removed. Peripheral blood mononucleated cells (PBMCs) are then extracted from the blood. A PBMC is any blood cell that has a round nucleus, including: lymphocytes (CD4+ and CD8+ T-cells, B-cells, and natural killer (NK) cells). Blood is mixed with fetal bovine serum (FBS) and layered onto Ficoll Paque, a polysaccharide that separates layers of blood by density. After centrifugation, PBMCs are collected from the surface and resuspended in FBS for use in experimental studies.

ELISpot is used to assess the production of a cytokine, specifically interferon γ (IFNγ), by PBMCs (Figure 1.7). Polyvinylidene fluroride (PVDF)-backed microplate strips pre-coated with an antibody against human IFNγ are washed with sterile phosphate buffered saline (PBS). PBMCs isolated from healthy donors are resuspended in AIM-V medium with 10% human AB serum. Synthetic peptide antigens are added to the wells (2μg/mL) along with phytohemagglutinin (PHA) as a positive control. Most peptides bind to HLA with low affinity and often dissociate once they get to the cell surface. Presumably, these exogenously added synthetic peptides can associate with empty HLA molecules on the PBMCs. The resulting HLA-(synthetic) peptide complexes then become

targets for the corresponding CTLs in the HD PBMCs, if they contain TCRs specific to the synthetic peptides. [45,46]

After 7 days the wells are re-stimulated with peptides and incubated for 16-24 hours. Cells are then harvested, washed with AIM-V, and allowed to react with the detection reagent (mAb 7-B6-1). Subsequently, substrate solution (BCIP-NBT-plus) is added to each well until spots start to appear, then washed extensively. Dry plates are read by an ELISpot reader, where one spot corresponds to an IFNγ secreting cell. Results are displayed as a number of spot-forming units per 10^6 PBMCs. Peptide responses are compared to a positive control (PHA) and viral peptides that most individuals have been
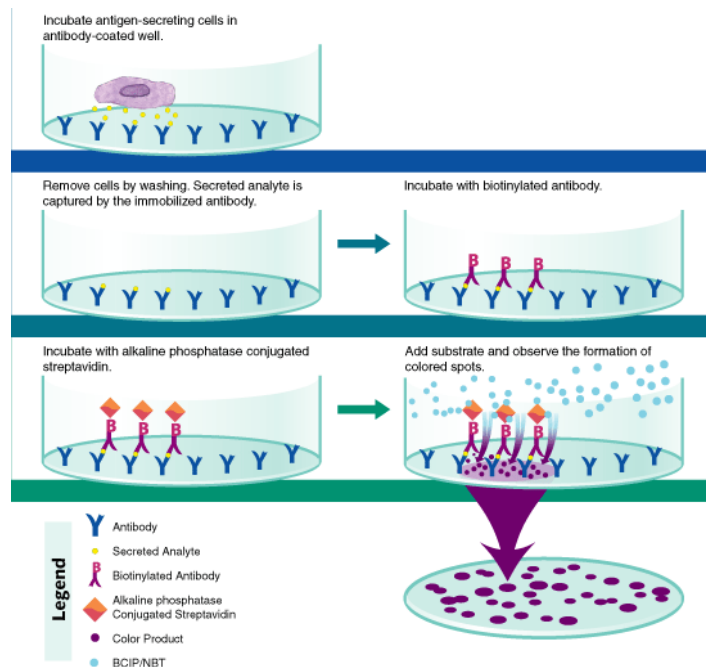
Figure 1.7 Enzyme-linked immunospot assay (ELISpot). Here, antibodies against IFNγ are coated onto a PVDF membrane and stimulated with synthetic peptides. The PBMCs release IFNγ (secreted analyte, yellow) and attach to the antibodies. Then a biotinylated polyclonal antibody specific for IFNγ is added to the wells, then alkaline-phosphatase conjugated to streptavidin, and finally a substrate solution (BCIP/NBT). A blue-black colored precipitate forms and appears as spots at the sites of cytokine localization, with each individual spot representing an individual analyte-secreting cell. The spots can be counted with an automated ELISpot reader system. (Image taken with permission from R&D system protocol)

exposed to and will have a response against. Those peptides showing a response with a magnitude similar to those viral and positive controls are considered strong immunotherapeutic candidates.

Cytokine production in response to specific stimuli can be measured using ICS. Golgi inhibitors are used to prevent cytokine release from the cells. Cells are then fixed and permeabilized to allow antibodies against individual cytokines to access the intracellular compartments. The antibodies are conjugated to fluorescent tags that can be monitored via flow cytometry, which permits us to measure individual (and total) cytokine responses.[47] We tested HD immunity against peptide antigens using three cytokines: (1) IFN γ, responsible for activation of innate immune system cells including macrophages and NK cells, (2) tumor necrosis factor α (TNFα), which also activates macrophages, and (3) interleukin-2 (IL-2), necessary for induction of T-cell and NK proliferation. Additionally, we monitored for CD107a, a cell surface protein indicative of degranulation and suggestive of T-cell killing.

Again, PBMCs are isolated from HDs and resuspended in AIM-V medium with 10% human AB serum. Synthetic peptide antigens are added to the wells (10μg/mL) and cells are expanded for 7 days. The positive control well is stimulated with 1 μg/mL PHA. On day 6, cells are washed and restimulated with peptide antigen overnight. The cells are harvested, washed, and stained with fixable viability dye (APC-Cy7) and surface antibodies: anti-CD3 (APC) and anti-CD8 (PerCP). Cells were fixed using formaldehyde, permeabilized, and stained with anti-IFNγ (PE), anti-IL2 (Pacific blue), and anti-TNFα (PE-Cy5.5). Cells are washed and analyzed on the flow cytometer.
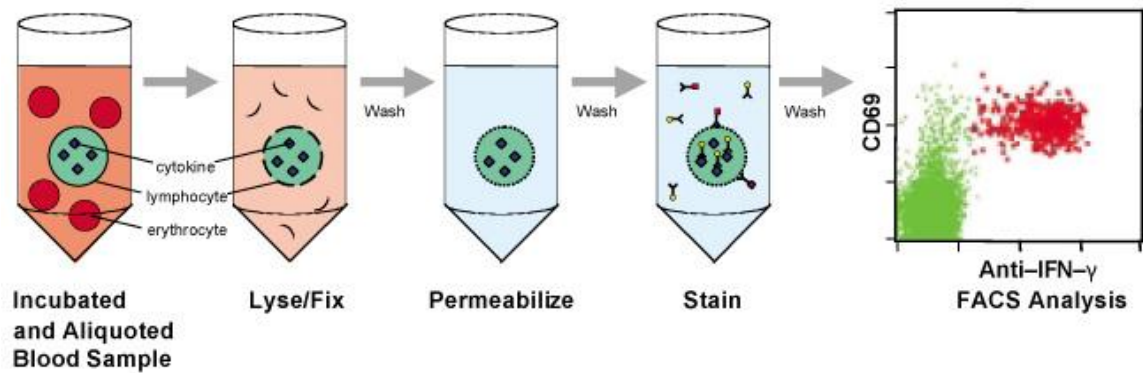
Figure 1.8 Intracellular cytokine staining. This workflow involves first isolating PBMCs from red blood cells (erythrocytes), then stimulating the cells with synthetic peptides. After stimulation, the cells are lysed, fixed, and permeabilized. Then, cytokine stains are added to infiltrate the cell membrane and attach to cytokines. The stains can be visualized on a flow cytometer to measure amount of cytokine produced in response to an antigen. (Image reprinted with permission from BD Biosciences protocol)

The ICS data can corroborate the ELISpot assay results by demonstrating which peptide antigens are capable of stimulating an immune response. However, the ICS data have the power to measure multiple cytokine responses while specifically determining whether the responses reside in the central memory compartment. The CD8+ T-cell specific response suggests a post-translationally modified peptide specific immunity.[47] Together, ELISpot and ICS are very powerful tools in screening large numbers of modified peptide antigens to determine the best candidates for cancer immunotherapeutics.

1.3 References

1. Ferlay J, Shin HR, Bray F, Forman D, Mathers C, and DM Parkin. Estimates of worldwide burden of cancer in 2008: GLOBOCAN. *Int J Cancer.* 2008; 127: 2893-2917.

2. Schrieber RD, Old LJ, and MJ Smyth. Cancer immunoediting: Integrating immunity's roles in cancer suppression and promotion. *Science.* 2011; 331(6024): 1565-1570.

3. Klein J and A Sato. The HLA system, first of two parts. *N Engl J Med*. 2000; 343(10): 702-709.

4. Engelhard VH, Brickner AG, and AL Zarling. Insights into antigen processing gained by direct analysis of the naturally processed class I MHC associated peptide repertoire. *Mol Immunol.* 2002; 39(3-4): 127-137.

5. Burmester G, Pezzutto A, Ulrichs T, and A Aicher. *Color Atlas of Immunology*. Thieme, Stuttgart; New York. 2002.

6. Bjorkman PJ, Saper MA, and B Samraoui. Structure of the human class I histocompatibility antigen, HLA-A2. *Nature.* 1987; 329(6139): 506-512.

7. Marsh SG. Nomenclature for factors of the HLA system, update January 2013. *Tissue Antigens*. 2013; 81: 471-473.

8. Zarling AL, Ficarro SB, White FM, and VH Engelhard. Phosphorylated peptides are naturally processed and presented by major histocompatibility complex class I molecules in vivo. *J Exp Med.* 2000; 192(12): 1755-1762.

9. Wucherpfennig KW. The first structures of T-cell receptors bound to peptide-MHC. *J Immunol.* 2010; 185(11): 6391-6393.

10. Mohammed F, Cobbold M, Zarling AL, and VH Engelhard. Phosphorylation-dependent interaction between antigenic peptides and MHC class I: a molecular basis for the presentation of transformed self. *Nat Immunol.* 2008; 9(11): 1236-1243.

11. Rammensee HG, Falk K, and O Rotzsche. Peptides naturally presented by MHC class I molecules. *Annual Rev. Immunol.* 1993; 11: 213-244.

12. Allison TJ, Winter CC< Fournie JJ, Bonneville M, and DN Garboczi. Structure of a human gamma-delta T-cell antigen receptor. *Nature.* 2001; 411: 820-824.

13. Baldwin KK, Reay PA, Wu L, Farr A, and MM Davis. A T-cell receptor specific blockade of positive selection. *J Exp Med*. 1999; 189: 13-24.

14. Starr TK, Jameson SC, and KA Hogquist. Positive and negative selection of T-cells. *Ann Rev Immunol*. 2003; 21: 139-176.

15. Stritesky GL, Jameson SC, and KA Hogquist. Selection of self-reactive T-cells in the thymus. *Ann Rev Immunol.* 2012; 30(1): 95-114.

16. Derbinski J, Schulte A, Kyewski B, and L Klein. Promiscuous gene expression in medullary thymic epithelial cells mirrors the peripheral self. *Nat Immunol*. 2001; 2: 1032-1039.

17. Usubuchi I, Nishimura S, Kudo H, Ito I, and Y Sobajima. Further studies on cross-immunity among syngeneic tumors in mice. *Tohoku J Exp Med.* 1975; 116: 373-377.

18. Dighe AS, Richards E, Old LK, and RD Schreiber. Enhanced in vivo growth and resistance to rejection of tumor cells expressing dominant negative IFN gamma receptors. *Immunity*. 1994; 1: 447-456.

19. Kaplan DH, Shankaran V, Dighe AS, Stockert E, Aguet M, Old LJ, and RD Schreiber. Demonstration of an interferon gamma-dependent tumor surveillance system in immunocompetent mice. *Proc Natl Acad Sci*. 1998; 95: 7556-7561.

20. Shankaran V, Ikeda H, Bruce AT, White JM, Swanson PE, Old LJ, and RD Schreiber. IFNgamma and lymphocytes prevent primary tumor development and shape tumor immunogenicity. *Nature.* 2001; 410: 1107-1111.

21. Bhatia A and Y Kumar. Cancer-immune equilibrium: questions unanswered. *Cancer Microenvironment.* 2011; 4(2): 2011.

22. Mackie RM, Reid R, and B Junor. Fatal melanoma transferred in a donated kidney 16 years after melanoma surgery. *N Engl J Med*. 2003; 348: 567-568.

23. Hanahan D and RA Weinberg. The hallmarks of cancer. *Cell.* 2000; 100(1): 57-70.

24. Hanahan D and RA Weinber. Hallmarks of cancer: the next generation. *Cell.* 2011; 144(5): 646-674.

25. Liu VC, Wong LY, Jang T, Shah AH, Park I, Yang X, Zhang Q, Lonning S, Teicher BA, and C Lee. Tumor evasion of the immune system by converting CD4+CD25- T-cells into CD4+CD25+ T regulatory cells: role of tumor-derived TGF beta. *J Immunol*. 2007; 178: 2883-2892.

26. Hagenaars M, Ensink NG, Basse PH, Hokland M, Nannmark U, Eggermont AM, Van de Velde CK, Fleuren GJ, and PJ Kuppen. The microscopic anatomy of experimental rat CC531 color tumor metastases: consequences for immunotherapy. *Clin Exp Metastasis.* 2000; 18: 189-196.

27. Finn OJ. Cancer immunology. *N Engl J Med.* 2008; 358(25): 2704-2715.

28. Engelhard VH, Altrich-Vanlith M, Ostankovitch M, and AL Zarling. Post-translational modifications of naturally processed MHC-binding epitopes. *Current Opinion Immunol.* 2006; 18: 92-97.

29. Manoury B, Mazzeo D, Fugger L, Viner N, Ponsford M. Streeter H, Mazza G, Wraith DC, and C Watts. Destructive processing by asparagine endopeptidase limits presentation of a dominant T-cell epitope in MBP. *Nat Immunol.* 2002; 3: 169-174.

30. Haurum JS, Tan L, Arsequell G, Frodsham P, Lellouch AC, Moss PAH, Dwek RA, McMichael AJ, and T Elliot. Peptide anchor residue glycosylation: effect on class I major histocompatibility complex binding and cytotoxic T lymphocyte recognition. *Eur J Immunol.* 1995; 25: 3270-3276.

31. Zarling AL, Polefrone JM, Evans AM, Hunt DF and VH Engelhard. Identification of class I MHC-associated phosphopeptides as targets for cancer immunotherapy. *Proc Natl Acad Sci.* 2006; 103(40): 14889-14894.

32. Udeshi ND, Compton PD, Shabanowitz J, Hunt DF, Rose KL. Methods for analyzing peptides and proteins on a chromatographic timescale by electron-transfer dissociation mass spectrometry. *Nat Protoc.* 2008; 3(11): 1709-1717.

33. Martin SE, Shabanowitz J, Hunt DF, Marto JA. Subfemtomole MS and MS/MS peptide sequence analysis using nano-HPLC micro-ESI fourier transform ion cyclotron resonance mass spectrometry. *Anal Chem.* 2000 Sep 15; 72(18): 4266-4274.

34. Dass C. Fundamentals of contemporary mass spectrometry. Hoboken, N.J.: Wiley-Interscience; 2007.

35. Kebarle P, Verkerk UH. Electrospray: From ions in solution to ions in the gas phase, what we know now. *Mass Spectrom Rev.* 2009 Nov-Dec; 28(6): 898-917.

36. Heeren RM, Kleinnijenhuis AJ, McDonnell LA, Mize TH. A mini-review of mass spectrometry using high-performance FTICR-MS methods. *Anal Bioanal Chem.* 2004 Feb; 378(4): 1048-1058.

37. Syka JE, Marto JA, Bai DL, Horning S, Senko MW, Schwartz JC, Ueberheide B, Garcia B, Busby S, Muratore T, Shabanowitz J, Hunt DF. Novel linear quadrupole ion trap/FT mass spectrometer: Performance characterization and use in the comparative analysis of histone H3 post-translational modifications. *J Proteome Res.* 2004 May-Jun; 3(3): 621-626.

38. Perry RH, Cooks RG, Noll RJ. Orbitrap mass spectrometry: Instrumentation, ion motion and applications. *Mass Spectrom Rev.* 2008 Nov-Dec; 27(6): 661-699.

39. McAlister GC, Phanstiel D, Good DM, Berggren WT, Coon JJ. Implementation of electron-transfer dissociation on a hybrid linear ion trap-orbitrap mass spectrometer. *Anal Chem.* 2007 May 15; 79(10): 3525-3534.

40. Schwartz JC, Senko MW, Syka JE. A two-dimensional quadrupole ion trap mass spectrometer. *J Am Soc Mass Spectrom.* 2002 Jun; 13(6): 659-669.

41. Wysocki VH, Tsaprailis G, Smith LL, Breci LA. Mobile and localized protons: A framework for understanding peptide dissociation. *J Mass Spectrom.* 2000 Dec; 35(12): 1399-1406.

42. Syka JE, Coon JJ, Schroeder MJ, Shabanowitz J, Hunt DF. Peptide and protein sequence analysis by electron transfer dissociation mass spectrometry. *Proc Natl Acad Sci U S A.* 2004 Jun 29; 101(26): 9528-9533.

43. Coon JJ, Shabanowitz J, Hunt DF, Syka JE. Electron transfer dissociation of peptide anions. *J Am Soc Mass Spectrom.* 2005 Jun; 16(6): 880-882.

44. Geer LY, Markey SP, Kowalak JA *et al*. Open Mass Spectrometry Search Algorithm. *J Proteome Res.* 2004; 3(5): 958-964.

45. Zhang W, Caspell R, and AY Karulin. ELISpot assays provide reproducibly results among different laboratories for T cell immune monitoring. *J Immunotoxicol.* 2009; 6(4): 227-234.

46. Anthony DD and PV Lehmann. T-cell epitope mapping using the ELISpot approach. *Methods.* 2003; 29(3): 260-269.

47. Cobbold M, De La Pena H, Norris A, Polefrone JM, Qian J, English AM, Cummings KL, Penny S, Turner JE, Cottine J, Abelin JG, Malaker SA, Zarling AL, Huang HW, Goodyear O, Freeman SD, Shabanowitz J, Pratt G, Craddock C, Williams ME, Hunt DF, and VH Engelhard. MHC class I-associated phosphopeptides are the targets of memory-like immunity in leukemia. *Sci Trans Med.* 2013; 5(203): 1-10.

Chapter 2: Identification of HLA-associated Phosphopeptides from Esophageal Adenocarcinoma

2.1 Introduction

2.1.1 Esophageal Cancer

The incidence of esophageal cancer has risen more rapidly than any other cancer over the last 40 years, particularly among white males.[1,2] The prognosis is particularly poor for this disease, as the five year survival rate is between 10-35% even if detected before metastasis.[3] After diagnosis, only 20-30% of patients are eligible for treatment with a curative intent, which usually involves chemotherapy followed by removal of the esophagus. Of these patients, 78% survive for one year, and only 46% survive for three years. Palliative chemotherapy is the only other option for the remaining 70-80% of patients to improve quality of life, but only 24% undergo the treatment. Of those, only 53% complete their therapy.[4] Thus, esophageal cancer carries an extremely terrible prognosis, even in those that undergo curative surgeries. An effective and specific therapy is sorely needed to help treat this cancer.

The two main types of esophageal cancer are squamous cell carcinoma (OESCC) and adenocarcinoma (OEAC), the latter being the focus of this study.[2] The two main risk factors for the disease include obesity and reflux disease.[5,6] The chronic inflammation brought by reflux disease can lead to replacement of the squamous epithelium by metaplasic columnar epithelium, called Barrett's Esophagus (BE). Metaplasia can then lead to dysplasia, which involves a change in size, shape, and organization of the esophageal tissue. Dysplasia is naturally brought about by cell cycle abnormalities,

specifically with known cancer pathways like EGF and p53, giving way adenocarcinoma and eventually metastasis.[7] The progression of the disease is depicted in Figure 2.1.[7]

Cell division of surface cells

Goblet cells

Glands

CJHU

Normal Lining   Barrett's Esophagus   with low-grade dysplasia   with high-grade dysplasia   Invasive carcinoma
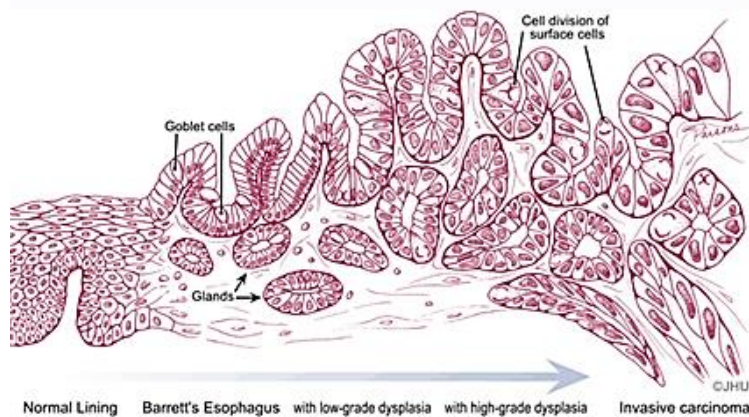
Figure 2.1 Progession of Esophageal Adenocarcinoma. Normal tissue is replaced by metaplasic columnar epithelium (Goblet cells), resulting in Barrett's Esophagus. Metaphasia leads to low-grade dysplasia and phenotypic changes in size, shape, and organization of tissue. Finally, high grade dysplasia is accompanied by cell cycle and signaling irregularities that eventually develop into invasive carcinoma.

2.1.2 Phosphorylation

Phosphorylation is the most common post-translational modification (PTM), as nearly one-third of proteins are phosphorylated in their lifetime. Its supreme utility arises from its reversible and transient nature, allowing for intricate control of cellular processes. Kinases serve to phosphorylate proteins on serine, threonine, and tyrosine residues, whereas phosphatases reverse the reaction. The interplay between these enzymes is of paramount importance to normal cell functions, including: transcription, differentiation, cell cycle regulation, migration, and metabolism, among others.[8] Aberrant phosphorylation results in dysregulation of these cell signaling pathways, which contributes to several hallmarks of cancer.[9,10]

We hypothesize that dysregulation of cell signaling in tumor cells will generate phosphorylated proteins that are unique to cancer. As a corollary, we believe that the

degradation of these cancer-specific phosphoproteins will result in their presentation via

the HLA class I processing pathway. By analyzing the peptides presented on both normal

and tumor tissue, we can identify tumor-specific phosphopeptides for use in novel cancer

immunotherapeutics. To be sure, we have identified phosphopeptides on the surface of

several cancers, including: colorectal, ovarian, breast, melanoma, and leukemia.[9-11]

Additionally, we have demonstrated that several of those phosphopeptides are capable of

eliciting an immune response through recognition of the transformed cells. Further, we

have generated CD8+ T-cells that specifically recognize a phosphorylated epitope but not

the unmodified version of the same peptide.[9-11] However, none of this work has been

done with esophageal cancer. Thus, the goal of this work is to identify tumor-specific

phosphopeptides isolated from esophageal tumors that are not present on normal tissue in

the same individual.

2.1.3 Phosphopeptide Enrichment

Despite being one of the most common post-translational modifications,

phosphorylation is often found at sub-stoichiometric levels when compared to

unmodified proteins. This issue is exacerbated when dealing with the high number of

HLA-associated peptides on the surface of cells. Thus, enrichment must be performed on

the samples in order to isolate the phosphopeptides from complex mixtures. To do this,

we utilize iron (III) immobilized metal affinity chromatography (IMAC).[12,13] In this

method, depicted in Figure 2.2, activated iron is immobilized onto iminodiacetic acid

(IDA) resin packed into a fused silica column. The positively charged iron forms a salt

bridge with the negatively charged phosphate moiety. All of the unmodified peptides can

be washed through the column, leaving only phosphopeptides bound to the resin. However, all acidic groups present on peptides (C-terminus, glutamic acid, aspartic acid) will bind to the resin. Thus, they are first neutralized via Fischer esterification.[12,13]
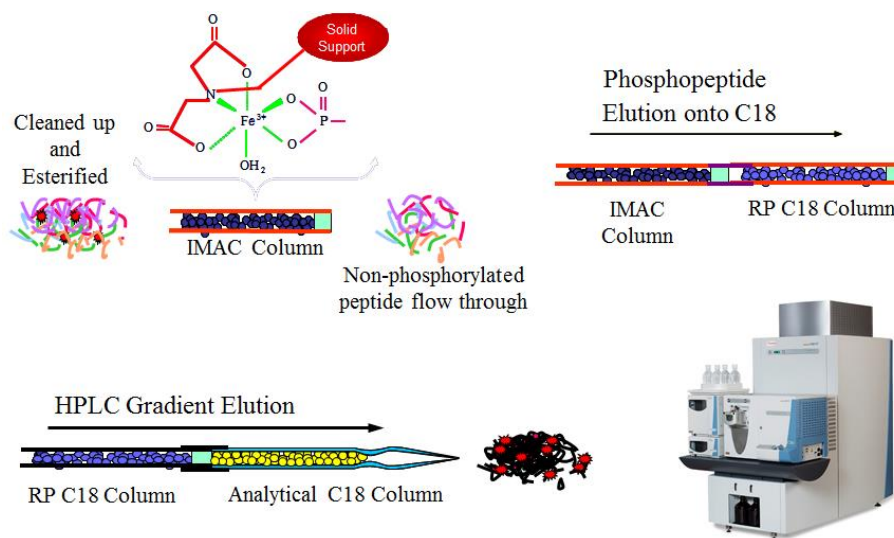
Figure 2.2 Immobilized metal affinity chromatography. Samples are cleaned up and esterified to reduce non-specific binding, then loaded onto the IMAC column where they bind to activated iron. Non-phosphorylated peptides are washed away while phosphopeptides are eluted with ascorbic acid onto a C18 precolumn. Samples are eluted via RP-HPLC and analyzed via high-resolution MS. [Adapted from JGA]

Once the samples are esterified, they are loaded onto the IMAC column. After unmodified peptides have been washed away, a C18 column is connected directly to the IMAC column. Then, ascorbic acid is added to reduce the iron from (III) to (II) thus neutralizing the positive charge and releasing the negatively charged phosphopeptide.[12,13] The phosphopeptides flow onto the C18 column where they can be eluted via reverse phase HPLC and analyzed with high-resolution mass spectrometry.

2.1.4 Phosphopeptide Fragmentation and Sequencing

As discussed in section 1.2.3, the CAD fragmentation spectrum of a post-translationally modified peptide will be dominated by the loss of the modification from

the parent ion. In the case of phosphorylation, this results in a CAD spectrum

overpowered by the loss of phosphoric acid (98 Da) from serine or threonine, since this is

the lowest energy fragmentation pathway.[14] Figure 2.3a shows an example CAD

spectrum from the peptide SVKPRRTpSL. In this case, few informative sequence ions

are present, thus it is possible to see how difficult sequencing and site-localizing a

modification can be. We can take advantage of this spectrum, though, by utilizing an in-

house developed program called "CAD Neutral Loss Finder" (credit to Dina Bai). This

program outputs all possible spectra that have a neutral loss of phosphoric acid from the

parent ion. Once a parent ion has been identified, we can use the corresponding ETD

spectrum (Figure 2.3b), the Mascot Sequence Query option, and *de novo* sequencing to

fully characterize and site-localize the phosphopeptide.
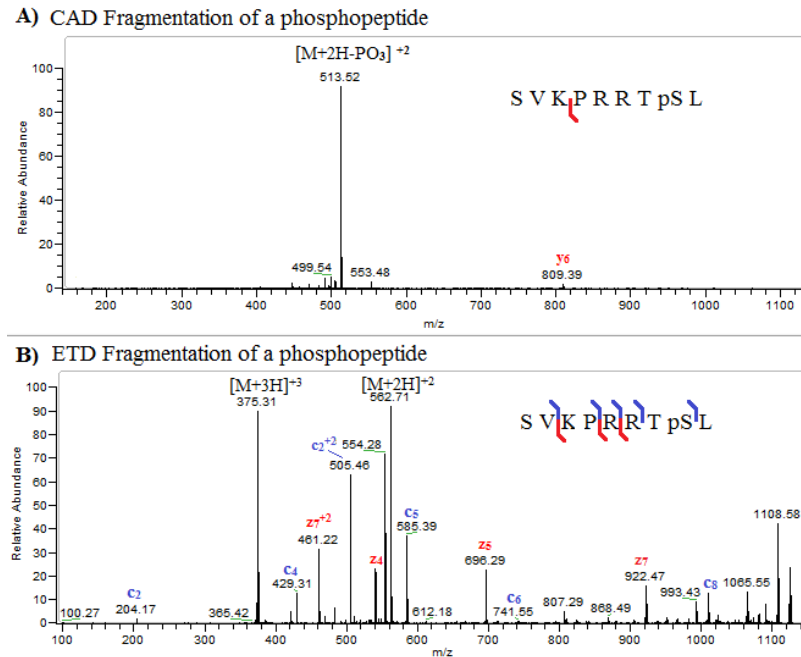
Figure 2.3. Fragmentation of phosphopeptides. A) CAD fragmentation results in the domination of a spectrum by the neutral loss of phosphoric acid, which can be used as a diagnostic in "CAD Neutral Loss Finder." **B)** ETD fragmentation preserves the post-translational modification, which allows us to sequence the phosphopeptide.

2.1.5 Project Aim

In this chapter, we seek to identify phosphopeptides isolated from the surface of human tumor tissue that are not present on the surface of matched normal tissue for use in cancer immunotherapeutics. To do this, we will isolate phosphopeptides via IMAC and high-resolution mass spectrometry, characterize them using a combination of *de novo* sequencing and bioinformatics, and screen the immunotherapeutic candidates using ELISpot assays.

2.2 Materials

2.2.1 Reagents

Amersham Pharmacia Biotech (Amersham, UK)

NHS-activated sepharose magnetic beads

Applied Biosystems^{TM} (Carlsbad, CA)

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

Grace Davison Discovery Science (Deerfield, IL)

Acetyl chloride, anhydrous

D_0 methanol, anhydrous

Honeywell (Morristown, NJ)

Acetonitrile, HPLC grade, ≥99.8% purity

J.T. Baker (Phillipsburg, NJ)

Glacial acetic acid, ≥99.9% purity

Mabtech (Sweden)

ELISpot PRO kit (human IFNγ)

Pierce (Rockford, IL)

LC-MS grade water

PQ Corporation (Valley Forge, PA)

Kasil ® 1624 potassium silicate solution

Sigma Aldrich (Saint Louis, MO)

Angiotensin I acetate salt hydrate, ≥90% purity

Angiotensin II phosphate

Aprotinin

Ascorbic acid

Azulene

CHAPS buffer

Dimethyl sulfoxide

Dulbecco's Modified Eagle Medium

Ethylenediaminetetraacetic acid (EDTA), analytical grade

Iron (III) chloride

HEPES buffer

Pepstatin A

Phenylmethylsulfonyl fluoride (PMSF)

Sodium chloride

Tris-HCl

Vasoactive intestinal peptide fragment 1-12, ≥97% purity

YMC Company, LTD (Kyoto, Japan)

ODS-AQ, C18 5 μm spherical silica particles, 120 A pore size

ODS-AQ, C18 5-20 μm spherical silica particles, 120 A pore size

2.2.2 Equipment and Instrumentation

Advanced Imaging Devices (Strasberg, Germany)

AID ELISpot Reading HR XL

Agilent Technologies (Palo Alto, CA)

1100 Agilent high performance liquid chromatograph

Branson (Danbury, CT)

Branson 1200 Ultrasonic Bath

Beckman Coulter (Pasadena, CA)

Optima LE-8K ultracentrifuge (Ti70 rotor)

Labconco Corp. (Kansas City, MO)

Centrivap centrifugal vacuum concentrator

Millipore (Billerica, MA)

Amicon ultra, 5 kDa regenerated cellulose spin filter

PolyMicro Technologies, Inc. (Phoenix, AZ)

360 μm o.d. x 50 μm i.d. polyimide coated fused silica capillary

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

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

Sutter Instrument Co. (Novato, CA)

P-2000 microcapillary laser puller with fused silica adapter

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

LTQ mass spectrometer (back end ETD)

LTQ FTICR hybrid mass spectrometer (custom modified with front end ETD)

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

LTQ-Orbitrap Velos 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.

2.3 Methods

2.3.1 Cell and tissue culture (Performed by SAP)

The OE19 cell line was grown up from a laboratory stock which was established in 1993 from an adenocarcinoma of gastric cardia/esophageal gastric junction of a 72 year old male patient. The cells were stored in fetal bovine serum (FBS) with 10% dimethylsulfoxide (DMSO) at -80ºC, were managed in sterile conditions in a laminar airflow hood, and incubated at 37ºC with 5% CO_2. The media used was Dulbecco's Modified Eagle Medium (D-MEM), 10% FBS, 1% penicillin/streptomycin and the cells were grown in vented tissue culture flasks with 20 mM HEPES buffer. Cell counts were determined by exclusion of Trypan blue dye under a light microscope.

Matched tissue samples (normal/tumor) from OE07, OE20, OE30, and OE32 were collected one-hour post-removal. Each tumor was at T3N1M0 staging, meaning that the patients had stage 3 sizing of tumor, regional lymph node metastases, but no distant metastases. Tumor and normal tissues ranged in size from 0.2 g to 2.2 g, and the entire sample was prepared for HLA-associated phosphopeptide analysis. We assume that 1 gram is equivalent to 1e9 cell equivalents (CEq).

2.3.2 HLA-associated peptide isolation (Performed by SAP)

HLA class I molecules were immunoaffinity purified from cell lines and tissues, and their peptides were extracted as previously described.[9-11] Cells were lysed in 10 mL CHAPS buffer, comprised of: 20mM Tris-HCl pH 8.0, 150 mM NaCl, 1% 3-[(3-cholamindopropyl) dimethylammonio]-1-propane sulfonate (CHAPS). Additionally,

protease inhibitors were added to prevent degradation of HLA molecules: 1 mM PMSF, 5 μg/mL aprotinin, and 10 μg/mL pepstatin A. Finally, phosphatase inhibitor cocktails II and III are added in 1:100 dilutions to prevent dephosphorylation of peptides during extraction. Lysates are subject to centrifugation at 100,000 xg for 1 hour at 4ºC in an ultracentrifuge. Supernatants were incubated with W6/32 antibody-bound NHS sepharose beads, which is specific for all classic HLA class I molecules. The beads were then subjected to a series of washes with lysis buffer, TBS (20 mM Tris-HCl, 150 mM NaCl pH 8), TBS 2 (20 mM Tris-HCl, 1 M NaCl pH 8), and 20 mM Tris-HCl pH 8. Peptides were eluted from the HLA class I molecules with 10% acetic acid and separated using a 5 kDa MWCO Ultrafree MC filter. Extracted peptides were stored at -80ºC before being shipped to the University of Virginia for analysis.

2.3.3 Peptide content determination by HPLC ESI-MS/MS

HLA-associated peptide mixtures were dried to completion in a speedvac and reconstituted in 0.1% acetic acid to a final concentration of 1e7 cell equivalents per μL. 1 μL of sample (1e7 CEq) was spiked into 4 μL of 0.1% acetic acid. 100 fmol of two standard peptides (angio and vaso) were spiked into the 5 μL solution and this was pressure loaded onto a C18 microcapillary precolumn (PC) (360 μm o.d. x 75 μm i.d.) with a 2 mm Kasil frit packed with 6-8 cm of C18 RP packing material (5-20 μm diameter particles). The PC was rinsed for 10-15 minutes with solvent A (0.1 M acetic acid), dried, and connected with Teflon to a bottleneck-fritted analytical column (AC)

(360 μm o.d. x 50 μm i.d) packed with C18 packing material (5 μm diameter particles) equipped with a laser-pulled electrospray emitter tip.

Peptides were eluted over a 40 minute gradient of 0-60% solvent B (70% acetonitrile, 0.1 M acetic acid) at 60 nL/min and electrospray ionized into the mass spectrometer. Intact peptide masses were acquired in the FTICR or the Orbitrap, and MS2 spectra (CAD and ETD) were acquired in the LTQ. Mass spectra were acquired using a method consisting of one high resolution scan (resolving power of 60,000 at 400 m/z) followed by 10 data dependent low resolution scans; the 5 most abundant parent ions were subject to both CAD and ETD. Data dependence parameters were set at a repeat count of 3, repeat duration of 10, and exclusion list length of 10. ETD parameters included: 35-45 ms reaction time (instrument dependent), FTMS automatic gain control (AGC) target of 2e5 charges, ITMS AGC target of 1e4, and ETD reagent target of 2e5.

Data analysis was performed using Xcalibur software. Raw data files were searched against the RefSeq database using OMSSA. MS2 searches used the following parameters: no enzyme specificity, e-value cutoff of 1, and variable modifications of Met (oxidation), Ser, Thr, and Tyr (phosphorylation). Mass tolerances for intact and product ion masses were set at ±0.1 Da and ±0.35 Da, respectively. Additionally, MS2 data was searched using either c- and z•- fragments (ETD) or b-and y- fragments (CAD). All peptide hits were subject to manual interpretation of MS1 (accurate mass) and MS2 (fragment ion) spectra.

Relative abundances of peptides were calculated by comparing the peptide peak area to the average peak areas of two internal standard peptides (angio and vaso), which we assume have a fixed concentration of 100 fmol. These peptide abundances were then

summed and compared to total peptide content present in a well characterized sample that had previously been analyzed by IMAC, which allows us to calculate the amount of sample necessary for performing IMAC on the current sample. In some cases, OE tissue (normal or tumor) had a negligible amount of peptide content in the screen. In these instances, amount of material sent determined the amount used for IMAC.

2.3.4 C18 cleanup of samples

OE-derived peptides in 0.1% acetic acid were combined with an equal volume of 0.1% and 100 fmol of an internal standard (angiotensin II phosphate (A2P), DRVpYIHPF). This solution was pressure loaded onto a fused silica clean-up column (360 μm o.d. x 150 μm i.d.) packed to 5 cm with C18 resin (5-20 μm diameter particles) at a flow rate of ~0.5 μL/min. The column loaded with sample was rinsed with 25 μL 0.1% acetic acid at the same flow rate. Then the column was moved to an HPLC where it was rinsed with solvent A for 10 minutes at 12-15 bar. Peptides were eluted into an Eppendorf tube using a 40 minute gradient (0-80% solvent B) that was held at 80% B for 30 additional minutes. The column was washed with 80% B and 20% A for an additional 20 minutes to clear any remaining contaminants from it. 200 fmol of another internal standard, STsLVEGR (STS), was then added to the sample tube and taken to dryness in the speedvac. The resulting tube was stored at -35ºC until further use.

2.3.5 Esterification of samples

Samples were esterified using a Fischer esterification. First, samples were dried in triplicate using 50 μL of completely anhydrous methanol. Concurrently, methanolic HCl was prepared by adding 160 μL acetyl chloride dropwise to 1 mL of methanol. 80 μL of this solution was added to the dried sample and the reaction was allowed to proceed for 1 hour. At this time, the sample was dried and rinsed once with 50 μL methanol. The entire process was repeated to assure completeness. The esterified and dried sample was stored at -35ºC until further use.

2.3.6 Phosphopeptide enrichment by IMAC

For each IMAC experiment, a 360 μm o.d. x 75 μm i.d. polyimide coated fused silica capillary fritted with a 2 mm Kasil frit was pressure loaded (200-500 psi) with 5 cm of POROS20 iminodiacetate packing material. The IMAC column was then prepared for iron (III) activation by the following rinse steps: 10 minute H_2O at 20 μL/min, 20 minute 50 mM EDTA at 20 μL/min, and 5 minute H_2O at 20 μL/min. At times that the frit did not allow for 20 μL/min flow rate, the rinse time was adjusted to allow the same amount of column volumes to pass through the IMAC column. For instance, a 10 μL/min flow rate would require the first H_2O rinse to be 20 minutes, totaling 200 μL through the column. The IMAC column was then activated using filtered iron (III) chloride ($FeCl_3$). Each activation step included a 10 minute $FeCl_3$ rinse at 20 μL/min followed by 3 minutes without pressure to allow the $FeCl_3$ to diffuse into the column particles. This was

repeated three times to assure complete activation. The activated column was equilibrated

with 0.01% acetic acid at 0.5 μL/min for 50 minutes.

Dried, esterified samples were brought up in 50 μL 1:1:1 0.01% acetic acid:

methanol: acetonitrile. This solution was loaded onto the column at 0.5 μL/min. A final

rinse of 0.01% acetic acid at 0.5 μL/min was performed for 30 minutes. At this point, a

PC was disconnected from an AC and rinsed with solvent A for 10 minutes at 30 bar. The

PC was then butt-connected with Teflon to the top of the IMAC column. The IMAC-PC

was allowed to rinse for 10 minutes with 0.01% acetic acid at a flow rate of 1 μL/min. If

a leak exists during this time, the pressure was let off and a new Teflon sleeve was

created. Phosphopeptides were eluted with 10 μL of 250 mM ascorbic acid solution for

10 minutes (1 μL/min flow rate). The IMAC-PC was rinsed for another 5 minutes with

0.01% acetic acid. The PC was then removed, rinsed with solvent A for 20 minutes at 30

bar, dried, and re-attached to an AC. The column was allowed to rehydrate for 20-30

minutes before the addition of 100 fmol of internal standards (angio and vaso). The

column was rinsed again for 10 minutes before elution.

2.3.7 HPLC-ESI-MS/MS analysis of phosphopeptides

Peptides were eluted at a flow rate of 60 nL/min using a 60 minute 0-60% B

gradient. Mass spectra were acquired as previously described within this chapter (section

2.3.3), with one high-resolution scan followed by data-dependent fragmentation of the

top 5 most abundant peptides using alternating CAD and ETD. Data dependence

parameters were set at a repeat count of 3, repeat duration of 10, and exclusion list length

of 10. ETD parameters included: 35-45 ms reaction time (instrument dependent), FTMS

AGC target of 2e5 charges, ITMS AGC target of 1e4, and ETD reagent target of 2e5.

Data analysis was carried out using Thermo Xcalibur software. Raw data files

were searched against the RefSeq database using OMSSA. MS2 searches used the

following parameters: no enzyme specificity, e-value cutoff of 1, fixed modifications of

the C-terminus, Asp, and Glu (methyl esters), and variable modifications of Met

(oxidation), Ser, Thr, and Tyr (phosphorylation). Mass tolerances for intact and product

ion masses were set at ±0.1 Da and ±0.35 Da, respectively. The data files were also

searched using the program described in section 2.1.4, "CAD Neutral Loss Finder,"

which identifies a neural loss of 98 from the parent ion. All peptide hits were subject to

manual interpretation of MS1 (accurate mass) and MS2 (fragment ion) spectra. In those

cases that Mascot was employed to assist in *de novo* sequencing, the following

parameters were used: ±6 ppm precursor mass tolerance, ±0.6 Da fragment ion mass

tolerance, no enzyme restriction, fixed methyl esters on the C-terminus, Asp, and Glu,

and variable phosphorylation on Ser, Thr, and Tyr.

Relative abundances were calculated by comparing peak areas of each

phosphopeptide to internal standards (angio and vaso) which are at a fixed concentration

of 100 fmol each. The abundances are then adjusted to account for sample loss through

the enrichment process by utilization of two phosphopeptide standards (A2P and STS)

spiked in at various points. For comparison purposes, all samples are normalized to

fmol/g, where we assume 1 gram = 1e9 cell equivalents.

2.3.8 Isolation of healthy donor PBMCs (Performed by SAP)

Peripheral blood mononucleated cells (PBMCs) are isolated from healthy donors (HDs) for use in biological testing. 50 mL of blood is removed from each HD into 50 mL syringes with lithium heparin solution (20 IU/mL) added. The blood was centrifuged for 10 min at 1200 x g and the plasma extracted from the top layer. PBMCs were then extracted from the blood. Blood was mixed with an equal volume of RPMI 10% FBS, layered onto ficoll paque, and centrifuged at 500 x g for 20 minutes. The PBMCs were collected from the interface, washed twice in RPMI 10% FBS, counted, and resuspended in RPMI 10% FBS for use in experimental studies.

2.3.9 Enzyme-linked immunospot (ELISpot) assays (Performed by SAP)

The enzyme linked immunospot (ELISpot) PRO human IFNγ kit was used for these experiments. Here, polyvinylidene difluoride (PVDF) membrane backed microplate strips pre-coated with monoclonal antibody (mAb) specific for human IFNγ were washed four times with sterile PBS; 200 μL/well. The wells were then blocked by incubation with RPMI 10% FBS for 30 minutes at RT. The medium was removed and a cell suspension of the PBMCs (5×10^5/well) in RPMI 10% FBS was added. The stimulating peptides were then added to the test wells along with phytohemagglutinin (PHA) to the positive control. The microplate was incubated for 7 days, then washed and restimulated with peptides for 16-24 hours at 37ºC. The cells were then removed and the plate was washed five times with PBS; 200 μL/well. The one-step detection agent (mAb 7-B6-1) was diluted 1:200 in PBS 0.5% FBS and 100 μL was added to each well. The plate was

allowed to incubate for 2 hours before washing five times with PBS. The substrate

solution (BCIP/NBT-plus) was filtered through a sterile 0.45 μm filter and 100 μL was

added to each well. When spots began to appear, the plate was extensively washed with

H_2O and dried. The spots representing individual IFNγ secreting cells were counted using

an ELISpot reader.

2.4 Results

2.4.1 Identification of HLA-associated phosphopeptides derived from esophageal samples

HLA class I associated phosphopeptides were isolated, enriched, and characterized from an esophageal cell line, primary tumors, and matched normal esophagus tissues. The HLA types of each of these samples are shown in Table 2.1.

	Sample code	Tumour Position	TNM	Sampled from	A locus	B locus	C locus
	OE07	Distal	T3N0M0	1° tumour	A*02	B*15 B*44	C*03 C*05
	OE20	Type I-II	T3N1M0	1° tumour	A*02 A*29	B*51 B*57	C*06 C*14
Primaries	**OE30**	Type II-III	-	1° tumour	A*01 A*24	B*07 B*08	C*07
	OE32	Type I	T3N1M0	1° tumour	A*01 A*29	B*35 B*57	C*04 C*06
Cell line	**OE19**	OGJ	T3N-M-	1° tumour	A*02 -	-	-

Table 2.1 HLA types of esophageal samples. Primary tumors were resected from four patients along with matched normal tissue. The tumor position, TNM staging, and HLA-alleles are listed from each tumor. The OE19 cell line originated from an adenocarcinoma of gastric cardia/esophageal gastric junction of unknown staging and HLA-type. (Figure credit to SAP)
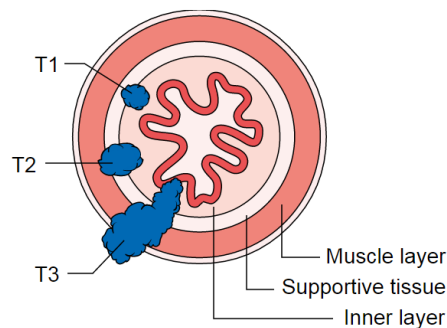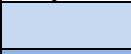
Figure 2.4 Types of esophageal cancer. T1 spans just the muscle layer and supportive tissue, T2 stretches into the muscle layer, and T3 portrudes from the muscle. Prognosis is made based on these criteria.

The following tables summarize those phosphopeptides isolated from the OE19 cell line and four primary tumors. The tables are separated by predicted HLA allele: -A (Table 2.2), -B (Table 2.3), and -C (Table 2.4). One peptide predicted to originate from HLA-E was detected and is depicted in Table 2.4 In total we identified 71

phosphopeptides, 68 of which were tumor-specific. In each table, we list the phosphopeptide sequence, protein source, HLA-binding prediction, relative abundance in each of the 5 samples, and relationships to known cancer pathways. Protein sources for each phosphopeptide were determined using OMSSA,[15] Mascot (matrixscience.com), and/or BLAST (blast.ncbi.nlm.nih.gov), and are listed using UniProt gene names (www.uniprot.org). HLA-binding was predicted using the "MHC-I binding prediction" program from the immune epitope database (IEDB; tools.immuneepitope.org), along with information about the HLA type of the patient. Abundances are only considered for those that match HLA-type of the peptide. If a peptide does not match a patient HLA-type, the corresponding square is gray. If the peptide matches the patient HLA-type, but does not

Key	
	0.1-1 fmol/g tissue
	1-10 fmol/g tissue
	10-100 fmol/g tissue
	>100 fmol/g tissue
	Not detected
	HLA mismatch

display a peptide, the corresponding square is white. Otherwise, the abundance is displayed in a heat map ranging from light to dark blue; a key is displayed at left. The relative abundances were calculated by comparing known levels of two internal standards (angio and vaso; 100 fmol). These values are adjusted for losses by comparing to phosphopeptide standards (A2P and STS). Finally, relationship to cancer pathways was determined using the NCI pathway interaction database site (pid.nci.nih.gov), and are listed next to the corresponding peptides.

Phosphopeptides	Protein of Origin	HLA type	OE07	OE20	OE19	OE30	OE30N(G)	OE32	Cancer Pathway
ITQGtPLKY	NCOR2	A*01	gray	gray	blue	blue			Notch/ER Signaling
NTDsPLRY	RPSA	A*01	gray	gray	blue			blue	
STDsETLRY	SYTL2	A*01	gray	gray				dark blue	
TMAsPGKDNY	KPNA6	A*01	gray	gray	light blue				
RmYsFDDVL	LMO7	A*01	gray	gray				light blue	
RSLsVEIVY	TP53	A*01	gray	gray	light blue				p53 binding protein
YPLsPAKVNQY	HIPK2	A*01	gray	gray				blue	
YPLsPTKISEY	HIPK1	A*01	gray	gray					
YPLsPTKISQY	HIPK1	A*01	gray	gray				blue	
KAFsPVRSV	ID2	A*02		blue		gray	gray	dark blue	c-Myc target/RB regulation/hypoxia response/ β-catenin signaling
RAHsEPLAL	DCAF15	A*02	blue			gray	gray		
RAHSsPASL	YAP1	A*02	blue			gray	gray		ERBB4/Notch/p73/TGFβ signaling
RLDsYVRSL	TRAPPC11	A*02			light blue	gray	gray		
RQIsQDVKL	AMPD2	A*02	light blue		light blue	gray	gray		
RSHsSPASL	WWTR1	A*02	blue			gray	gray		
RTFsPTYGL	SYNM	A*02	blue			gray	gray		
SIPtVSGQI	ZNF384	A*02			blue	gray	gray		
SMTRsPPRV	SRSF7	A*02			blue	gray	gray		
YAVPRRGsL	SVIL	A*02			gray	gray	gray	blue	regulation of androgen receptor activity
VYTyIQSRF	DYRK4	A*24	gray	gray	gray	blue		gray	
KSLsPSLLGY	NFATC2	A*29	gray		gray	gray	gray	blue	AP-1 TF regulation/wnt signaling
RPIsVIGGVSLY	SPATA13	A*29	gray		gray	gray	gray	light blue	CDC42/RAC regulation

Table 2.2 Phosphopeptides identified from OEAC samples predicted to bind HLA-A motifs. 22 peptides were detected with predicted binding alleles from HLA-A, 6 of which are from proteins that are known to interact with cancer pathways.

Phosphopeptides	Protein of Origin	HLA type	OE07	OE20	OE19	OE30	OE30N(G)	OE32	Cancer Pathway
APRRYsSSL	ARHGAP17	B*07	gray	gray	gray				CDC42/RAC regulation
FPLsPLRKY	SYTL2	B*07	gray	gray	gray	blue			
GPRsASLLSL	GPSM3	B*07	gray	gray	gray				
KPRsPVVEL	GRK2	B*07	gray	gray	gray	dark blue	blue		
LPAsPRARL	MAP7D1	B*07	gray	gray	gray	blue			
LPIFSRLsI	ZFP36L2	B*07	gray	light blue	gray			light blue	
QPRsPGPDYSL	GFI1	B*07	gray	gray	gray	blue			c-Myc target

Phosphopeptides	Protein of Origin	HLA type	Cancer Pathway
RPAsARAQPGL	SAYSD1	B*07	
RPFsPREAL	LUZP1	B*07	
RPIsPGLSY	CCDC6	B*07	
RPIsPPHTY	ROBO1	B*07	Adherens junction regulation/ RAC activation
RPmsESPHm	ZFP36L1	B*07	c-Myc target
RPRARsVDAL	LISCH7	B*07	
RPRGsQSLL	ERBB3	B*07	Integrin signaling
RPRPVsPSSL	SIK1	B*07	CREB regulation
RPRsPRQNSI	ATXN2	B*07	
RPSRsSPGL	SYNPO	B*07	
SPFKRQLsL	NUMBL	B*07	
TPRsPPLGL	MAPK11	B*07	RB regulation/p73/p38/FAS signaling
VPLIRKKsL	PGA5	B*07	
LKLsYLTWV	CAT4	B*15	
RQIsQDVKL	AMPD2	B*15	
KRYsGNMEY	LATS1	B*27	
LPKsPPYTAF	EIF4B	B*35	mTOR signaling
RPTsFADEL	FAM21C	B*35	
RPVsPFQEL		B*35	
RPVsTDFAQY	ABLIM1	B*35	
SFDSGsVRL	IRF6	B*44	
EERRsPPAP	FRA2	B*45	
KSGELLAtW	SIK2	B*57	mTOR signaling

Table 2.3 Phosphopeptides identified from OEAC samples predicted to bind HLA-B motifs. 30 peptides were detected with predicted binding alleles from HLA-B, 9 of which are from proteins that are known to interact with cancer pathways.

Phosphopeptides	Protein of Origin	HLA type	OE07	OE20	OE19	OE30	OE30N(G)	OE32	Cancer Pathway
RSDsYVEL	NCBP2	C5							
KRAsFAKSV	WNK4	C6/7							
KRYsEPVSL	FHDC1	C6/7							
KRYsGNMEY	LATS1	C6/7							
RKPsIVTKY	ATRX	C6/7							
RRFsGTAVY	ZNF518	C6/7							
RRFsPPRRM	H3BPG5	C6/7							
RRGsFEVTL	SELH	C6/7							

Peptide	Protein	Allele							Pathway
RRIsDPEVF	FILIP1L	C6/7							
RRIsDPQVF	FILIP1L	C6/7							
RRKsQVAEL	MK167IP	C6/7							
RRLsFLVSY	QARS	C6/7							tRNA aminoacylation
RRMsLLSVV	RIMKLB	C6/7							
RRNsAPVSV	RMD3	C6/7							Integrin signaling
RRPsLLSEF	NCOR1	C6/7							Notch/ER/ c-Myb signaling/ differentiation
RRSsSVAQV	UBD	C6/7							
SRSSSVLsL	PLEKHG3	C6/7							
TRKtPESFL	EPN1	C6/7							EGF signaling
RTHsLLLLL	RNASE4	E							

Table 2.4 Phosphopeptides identified from OEAC samples predicted to bind HLA-C and HLA-E motifs. 18 peptides were detected with predicted binding alleles from HLA-C, 4 of which are from proteins that are known to interact with cancer pathways. Additionally, one peptide from HLA-E was detected.

2.4.2. Peptide synthesis and biological testing

We selected a total of 40 phosphopeptides for synthesis and biological testing on the basis of: relationship to cancer pathways, tumor-associated expression, presence on multiple tumor types, and HLA allelic frequency in the human population. All of these are imperative in developing an effective immunotherapy. As seen in Tables 2.2-2.5, several of the peptides originate from proteins with relationships to cancer pathways. Those related to cancer were given priority because we believe they originate from dysregulated cell signaling pathways and should not be found on normal tissues. Thus, they should not mount an auto-immune response, which would render an immunotherapy dangerous to all individuals.

Additionally, an ideal cancer immunotherapy can be used across multiple cancer types. Thus, we created a database of every phosphopeptide we have detected in our laboratory, spanning several cancers including: melanoma, colorectal, breast, ovarian,

leukemias, lymphomas, etc. Using this database, we can select those peptides that have occurred frequently in our phosphopeptide analyses.

Only 3 peptides from one patient (OE30) were detected from the surface of normal tissue. The phosphopeptides from this sample were identified as: VPLIRKKsL from pepsinogen 5, KPRsPVVEL from β-andrenergic receptor kinase 1, and RPFsPREAL from leucine zipper protein 1. All of these were also detected on the surface of tumor tissue, but were only considered for biological testing if their presence on normal tissue was significantly lower than that on tumor.

Peptides that bind to the most commonly expressed HLA molecules were considered high priority for biological testing, since cancer immunotherapies using these peptides can be applied to the majority of the population. The expression frequencies of the ten most common HLA alleles found in the U.S. Caucasian population are shown in Table 2.5. This table does not account for HLA-C alleles, since they have a lower

Allele	Phenotype Frequency (%)
A*02	50.7
B*07	30.2
A*01	28.7
B*44	27.1
B*08	22.5
A*24	21.3
B*35	20.9
A*03	20.6
B*40	14.7
A*11	14

Table 2.5 Allele frequences of the US Caucasian population. Frequencies of the 10 most common alleles found in the US Caucasian population are shown as a percentage. A peptide predicted to bind those high frequency alleles are considered high priority for biological testing.

expression level than HLA-A and B. However, a significant number of HLA-C peptides were detected on the surface of tumors, suggesting their importance. We synthesized peptides from HLA-A*01, A*02, B*07, B*44, B*45, C*06, C*07, and one peptide from HLA-E.

After careful selection of phosphopeptides, they were synthesized and shipped to the University of Birmingham for biological testing. Here, Sarah Penny and colleagues performed 7-day cultured IFNγ ELISpot analyses to assess which phosphopeptides

stimulated healthy donor immune responses. We use these results to distinguish which

peptides are strong immunotherapeutic candidates for further biological testing. The

results are shown on the following pages. In total, 22 individuals were tested, but only

those HDs that matched peptide allotype were tested. The key is shown below.

Key	
Responses (day7)	
■	>100 SFU/500, 000 PBMC
■	25-100 SFU/500,000 PBMC
■	5-24 SFU/500,000 PBMC
	0-5 SFU/500,000 PBMC
■	Mismatched HLA
Peptides	
■	High response/moderate response and responses in more than 1 HD
■	Viral Controls

Total responses were heterogeneous between and within donors, with some

individuals having more and larger responses than others. Overall, 18 of the 37

phosphopeptides stimulated a high to moderate response and/or responses in more than

one HD, suggesting their ability to elicit an immune response. Some of those responses

were on the order of or higher than those to viral peptides. The implications of this will

be discussed in the next section.

	AR	SM	SL	JB	LQ	DL	CD	JZ	RA	CM	LM	LJI	AA	RM	DR	GR	AII	MR	MC	LQ	AM	SR	Total
HLA-A*01																							
ITQGtPLKY																							219
NTDsPLRY																							377
STDsETLRY																							177
TMAsPGKDNY																							5
YPLsPAKVNQY																							25
YPLsPTKISEY																							9
VTEHDTLLY																							1734
HLA-A*02																							
KAFsPVRSV																							33
RAHSsPASL																							26
RLDsYVRSL																							33
RQIsQDVKL																							56
RSHsSPASL																							103
SIPtVSGQI																							32
FLYALALLL																							637
GILGFVFTL																							2300
HLA-B*07																							
FPLsPLRKY																							10
KPRsPVVEL																							3
LPIFSRLsI																							1
QPRsPGPDYSL																							177
RPIsPGLSY																							41
RPIsPPHTY																							234
RPmsESPHm																							168
RPRARsVDAL																							64
RPRGsQSLL																							28
RPRsPRQNSI																							126
TPRsPPLGL																							147
RPPIFIRRL																							742
TPRVTGGGAM																							47
HLA-B*44																							
SFDSGsVRL																							76
HLA-B*45																							
EERRsPPAP																							244

Table 2.6 HLA-A and HLA-B peptide 7-Day ELISpot Responses from healthy donors. 12 HLA-A and 13 HLA-B peptides were tested in healthy donors to assess their ability to stimulate an immune response in HLA-matched individuals. Responses are depicted in heat-map style, ranging from white (no response) to dark red (high response). Those with gray boxes indicate an allelic mismatch. Total responses from all 22 individuals are shown by the blue bars at right. 5 viral peptides (highlighted in blue) were also tested in order to compare typical viral responses to the phosphopeptide responses. Phosphopeptides highlighted in orange indicate a moderate-high response and/or responses in more than one HD.

HLA-C6/7

Peptide	Total
KRAsFAKSV	73
KRYsGNMEY	201
RKPsIVTKY	36
RRGsFEVTL	59
RRIsDPEVF	17
RRLsFLVSY	16
RRMsLLSVV	514
RRNsAPVSV	34
RRPsLLSEF	185
RRSsSVAQV	10
TRKtPESFL	244
CRVLCCYVL	240

HLA-E

Peptide	Total
RTHsLLLLL	157

Table 2.7 HLA-C and HLA-E peptide 7-Day ELISpot Responses from healthy donors. 11 HLA-C and 1 HLA-E peptides were tested in healthy donors to assess their ability to stimulate an immune response in HLA-matched individuals. Responses are depicted in heat-map style, ranging from white (no response) to dark red (high response). Those with gray boxes indicate an allelic mismatch. Total responses from all 22 individuals are shown by the blue bars at right. 1 viral peptide (highlighted in blue) was also tested in order to compare typical viral responses to the phosphopeptide responses. Phosphopeptides highlighted in orange indicate a moderate-high response and/or responses in more than one HD.

2.5 Discussion

To our knowledge, this is the first study to identify phosphopeptides presented by HLA molecules on normal and tumor esophageal tissue. Through utilization of iron (III)-IMAC enrichment followed by HPLC-ESI-MS/MS, we have identified 71 phosphopeptides, 68 of which were tumor specific. We then selected 37 of these for biological testing after considering the parent proteins' relationship to cancer pathways, tumor-associated expression, presence on multiple tumor types, and HLA allelic frequency in the human population. These peptides were subjected to a 7-day ELISpot assay to assess their ability to stimulate an immune response in various healthy donors. The PBMCs from these HDs were capable of recognizing the tumor-specific (and tumor-associated) peptides as being indicative of malignant transformation, as evidenced by Tables 2.6 and 2.7. Comparing these responses to responses to common viral peptides demonstrates that the phosphopeptides are immunologically relevant, and suggests these peptides have potential as immunotherapeutic agents.

Interestingly, we only identified 3 phosphopeptides from one normal esophageal sample. These phosphopeptides were detected on the surface of OE30N(G), which was actually taken from gastric tissue. The normal esophageal tissue (OE30N(OE)) was only 0.3 g and did not provide enough peptide content to perform IMAC. Thus, we did not actually detect any phosphopeptides from the surface of normal esophageal tissue. This is partially explained by the small tissue size obtained in each matched set. OE30N(G) was only 0.5g, as was OE32N. OE07N was even smaller, weighing only 0.2g. The exception to this is OE20N, which weighed 2.2g. Even then, no phosphopeptides were detected in this sample. This is not to say that phosphopeptides are not present, however, if they exist

they are likely below our limit of detection. The lack of phosphopeptides detected on normal tissue corroborates our hypothesis that phosphorylation is an indicator of transformation events within a cell.

The three peptides we identified on OE30N(G) were VPLIRKKsL from pepsinogen 5, KPRsPVVEL from β-andrenergic receptor kinase 1, and RPFsPREAL from leucine zipper protein 1. The former, VPLIRKKsL, was detected 2x more on normal tissue than tumor tissue, thus it is likely a gastric-normal phosphopeptide. The latter phosphopeptides were detected at least 2x higher on tumor tissue, so we considered them tumor-associated peptides. KPRsPVVEL has been detected by our laboratory on several normal tissues, tumors, and cell lines, including: colorectal cancer and normal tissue, colorectal cancer cell lines, liver metastases, melanoma, EBV-transformed cell lines, among others. Interestingly, though, this peptide did not stimulate a response in any of the HDs we tested by ELISpot. This leads us to believe this peptide is also a normal phosphopeptide found on the surface of a plethora of cells. RPFsPREAL was not synthesized for further testing; however, it would be interesting to see if we obtained the same or similar results since we have also detected this phosphopeptide on several normal and transformed samples. This would help solidify the idea that phosphopeptides identified on the surface of normal tissue should not be considered high priority for further investigation.

Based on the biological data shown in Tables 2.6-2.7 along with the peptide expression data summarized in Tables 2.2-2.4, several phosphopeptides were identified that are of high interest for further investigation. For instance, the HLA-A*01 phosphopeptide ITQGtPLKY is derived from the protein nuclear receptor co-repressor 2

(NCOR2). This protein is a transcriptional co-repressor that can block the cellular differentiation of hematopoietic cells and is related to the initiation and progression of solid tumors.[16] Further, expression of NCOR2 has been shown to be associated with poor outcomes in patients with breast cancer.[17] We detected this peptide expressed on 2/3 HLA-matched esophageal tumors, and the phosphosite was previously unreported in Uniprot. However, our laboratory has detected this phosphopeptide on the surface of melanoma, colorectal cancer, and now esophageal cancer. We believe this suggests its importance in signaling a transformation event indicative of cancer. Further, the ELISpot data (Table 2.6) shows that the peptide was capable of stimulating a response in at least two individuals. Both of these HDs responded to the A*01 viral peptide at a similar magnitude as they did to the phosphopeptide. Taken together, we believe this peptide may be a good target for therapeutic development.

Another example of a strong candidate for immunotherapy is RPIsPPHTY. This peptide stimulated responses in two HLA-matched individuals (Table 2.6), which resulted in the highest cumulative response for that HLA type. This peptide is particularly attractive for therapies because it has an HLA-B*07 binding motif, which would affect approximately 30% of the U.S. Caucasian population (Table 2.5). The originating protein is called "roundabout homolog 1" (ROBO1) which acts as a molecular guide in cellular migration. Dysregulation of cell motility is a hallmark event in cancer cell invasion and metastasis.[18,19] ROBO1 is specifically responsible for adherens junction regulation and Rac1 activation.[20] Adherens junctions are comprised of cell-cell adhesion molecules and loss of adhesion integrity plays a key role in epithelial-to-mesenchymal transition (EMT) during tumor metastasis. Rac1 is a small G protein responsible for cell

cycle, cell-cell adhesion, motility, and epithelial differentiation. Interestingly, a few recent studies have attempted to inhibit Rac1 activity to suppress tumor growth in metastatic melanoma and liver cancer.[21] Due to the ability of this phosphopeptide to elicit strong immune responses, its broad applicability to the human population, and its parent protein's clear association with cancer, we believe this peptide may be an excellent candidate for future study.

Previous studies in our lab, specifically those that utilized W6/32 (non-specific) antibody to isolate HLA molecules, identified several tumor-specific phosphopeptides with supposedly unknown HLA types. However, upon further investigation, we believe these peptides likely originate from HLA-C alleles. HLA-C was often overlooked by these studies because it is expressed at 10% of those more common allotypes (HLA-A and HLA-B). This is likely due to the HLA-C molecule's low affinity for β_2-microglobulin, association with which is necessary for presentation on the surface. Thus, HLA-C likely gets degraded in the endoplasmic reticulum prior to surface presentation. Despite its lower abundance on the cell surface, HLA-C plays an integral role in disease.[22] HLA-C molecules bind to natural killer (NK) cells, and the NK cells use inhibiting receptors to recognize a repertoire of self-peptides in the HLA-C pocket, therefore preventing cell death. If the peptide is recognized as non-self or the HLA-C molecules are absent, the NK cell activates and kills the infected cell.[23] As described in the introduction, one of the hallmarks of cancer is an ability to evade the immune system. One way transformed cells can do this is by a complicated mechanism that downregulates HLA-A and -B but maintains normal levels of HLA-C and -E, which minimizes cytotoxic T-cell action while circumventing NK action on the infected cells.[19] This may be the case

for the tumor samples we investigated, as HLA-C-associated phosphopeptides represented >25% of total phosphopeptides detected.

Another reason why HLA-C peptides may be attractive candidates for cancer immunotherapies is because the HLA-C molecules are less polymorphic than the classic HLA molecules. According to Gonzalez-Galarza *et al*, a potential treatment targeting HLA-C may affect up to 78% of the population. This is evidenced by our 22 HDs, 10 of whom (45%) were HLA-C6/7 positive, second only to HLA-A*02 (50%). (Note: C6 and 7 have similar binding motifs and often cross-present peptides, hence they are often listed together.)[24] Its broad applicability can also be seen in Table 2.7, as five individuals responded, albeit weakly, to the phosphopeptide RRGsFEVTL. No other peptide stimulated a response in as many individuals. Interestingly, this peptide originates from selenoprotein H (SELH), which has a redox function and is overexpressed in several cancer cell lines.[25]

Along with RRGsFEVTL, several other HLA-C phosphopeptides stand out as strong therapeutic candidates. One peptide in particular, RRMsLLSVV, stimulated the highest response for any of the phosphopeptides that we investigated. The magnitude of this response was nearly double that seen against the viral peptide of the same allele, pointing to its extreme potency. The peptide derives from β-citryl-glutamate synthase B, which catalyzes the synthesis of β-citrylglutamate and N-acetyl-aspartyl-glutamate (NAAG). NAAG has been shown to stimulate growth in glioma stem cells.[26] Further, both of these molecules, when catabolized, result in the production of glutamate. One emerging hallmark of cancer is the reprogramming of energy metabolism within a tumorigenic environment. Most cancers depend on a high rate of aerobic glycolysis for

their continued growth and survival, which results in a decrease in the amount of

pyruvate entering in the tricarboxylic acid (TCA) cycle because upstream metabolites are

fed to other biosynthetic pathways.[19] Because the TCA cycle is responsible for the

synthesis of lipids, some amino acids, and nucleotides, these rapidly dividing cells

require a mechanism for replenishing the cycle. Thus, many cancer cells rely on

glutamate to support the TCA cycle. This peptide, along with other overlooked HLA-C-

associated phosphopeptides, shows incredible promise as we move forward in developing

targeted cancer therapies.

Future work on this project will involve further biological testing to determine

which candidates are the best subjects for cancer immunotherapeutics. Toward this end,

we plan on performing intracellular cytokine staining (ICS) with HD PMBCs to assure

that the immune response lies within the central memory component; that is, whether the

CD8+ T-cells are those stimulating an immune response.[11] In this same assay, we can

measure multifunctional cytokine responses (IFN γ, TNFα, and IL2) to a specific peptide.

This test can also be performed on tumor infiltrating lymphocytes (TILs) extracted from

the same tumors involved in this study. Here, TILs allow us to identify which peptides

are important to the functional immune response against esophageal cancer. Finally, we

plan to grow T-cells that are specific to our phosphopeptides and perform Europium

release assays, which are demonstrated in Figure 2.5. This will allow us to determine how

efficiently T-cells will kill cells in a phosphopeptide-specific manner.[27] Ultimately, we

believe that our phosphopeptide identification and candidate validation process will lead

to the development of targeted cancer immunotherapeutics.

Assay Principle
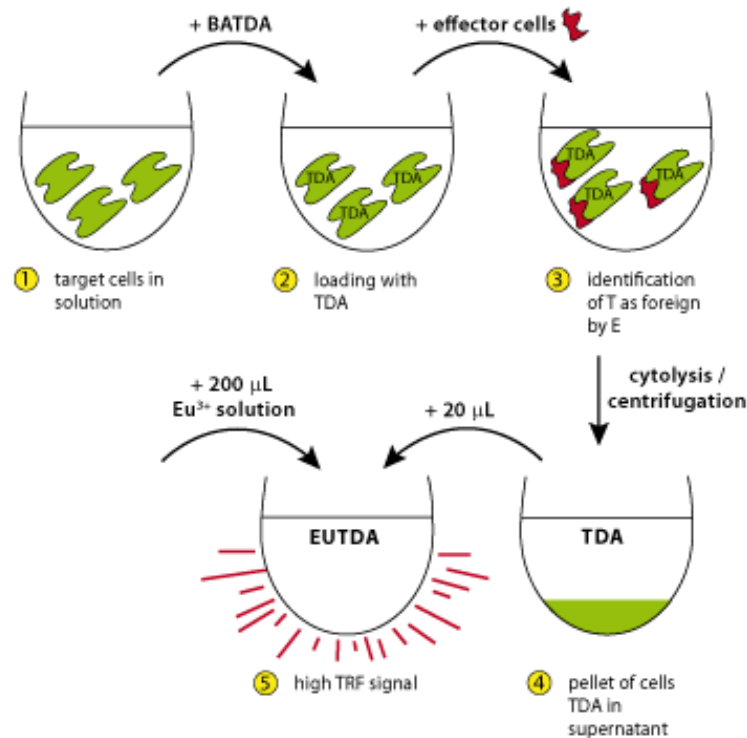
Figure 2.5 Europium release assays. A cancer cell line is loaded with BATDA (a fluorescence enhancing ligand), which penetrates the cell membrane. BATDA hydrolyzes to TDA within the cell. Then the T-cells grown up against a phosphopeptide are added. The T-cells kill the cells presenting the phosphopeptides, which releases TDA into solution. Europium is added, and the fluorescent chelate (EuTDA) levels are measured to determine the efficiency of T-cell killing. [Image taken with permission from BMG Labtech webite,]

2.6 References

1. Lepage C, Rachet B, Jooste V, Faivre J, and MP Coleman. Continuing rapid increase in esophageal adenocarcinoma in England and Wales. *Am J Gastroenterol.* 2008; 103: 2694-2699.

2. Cook MB, Chow WH and SS Devesa. Esophageal cancer incidence in the United States by race, sex, and histologic type, 1977-2005. *Br J cancer.* 2009; 101: 855-859.

3. CRUK 2013c. Oesophageal cancer key facts [online]. Available: http://www.cancerresearchuk.org/cancer-info/cancerstats/keyfacts/oesophageal-cancer/

4. Groene OCD, Hardwick R, Riley S, Crosby T, and K Greenaway. National Esophageogastic Cancer Audit, 2012. [online]. Available: https://catalogue.ic.nhs.uk/publications/clinical/oesophago-gastric/nati-clin-audi-supp-prog-oeso-gast-canc-2012/clin-audi-supp-prog-oeso-gast-2012-rep.pdf: Royal College of Surgeons.

5. Lagergren J, Bergstrom R, Lindgren A, and O Nyren. Symptomatic gastroesophadeal reflux as a risk factor for esophageal adenocarcinoma. *N Engl J Med.* 1999; 340: 825-831.

6. Lagergren J, Bergstrom R, and O Nyren. Association between body mass and adenocarcinoma of the esophagus and gastric cardia. *Ann Intern Med.* 1999; 130: 883-890.

7. Tselepis C, Perry I, and J Jankowski. Barrett's esophagus: dysregulation of cell cycling and intercellular adhesion in the metaplasia-dysplasia-carcinoma sequence. *Digestion.* 2000; 61: 1-5.

8. Ozlu N, Akten B, Timm W, Haseley N, Steen H, and J Steen. Phosphoproteomics. *Wiley Interdiscip Rev system Biol Med.* 2010; 2: 255-276.

9. Zarling AL, Ficarro SB, White FM, and VH Engelhard. Phosphorylated peptides are naturally processed and presented by major histocompatibility complex class I molecules in vivo. *J Exp Med.* 2000; 192(12): 1755-1762.

10. Zarling AL, Polefrone JM, Evans AM, Hunt DF and VH Engelhard. Identification of class I MHC-associated phosphopeptides as targets for cancer immunotherapy. *Proc Natl Acad Sci.* 2006; 103(40): 14889-14894.

11. Cobbold M, De La Pena H, Norris A, Polefrone JM, Qian J, English AM, Cummings KL, Penny S, Turner JE, Cottine J, Abelin JG, Malaker SA, Zarling

AL, Huang HW, Goodyear O, Freeman SD, Shabanowitz J, Pratt G, Craddock C, Williams ME, Hunt DF, and VH Engelhard. MHC class I-associated phosphopeptides are the targets of memory-like immunity in leukemia. *Sci Trans Med.* 2013; 5(203): 1-10.

12. Ficarro SB, McCleland ML, and Stukenberg PT. Phosphoproteome analysis by mass spectrometry and its application to Saccromyces cerevisiae. *Nat Biotech.* 2002; 20(3): 301-305.

13. Ficarro S, Chertihin O, and VA Westbrook. Phosphoproteome analysis of capacitated human sperm. *J Biol Chem.* 2003; 278(13): 11579-11589.

14. Udeshi ND, Compton PD, Shabanowitz J, Hunt DF, Rose KL. Methods for analyzing peptides and proteins on a chromatographic timescale by electron-transfer dissociation mass spectrometry. *Nat Protoc.* 2008; 3(11): 1709-1717.

15. Geer LY, Markey SP, Kowalak JA *et al*. Open Mass Spectrometry Search Algorithm. *J Proteome Res.* 2004; 3(5): 958-964.

16. Battaglia, S. Maguire O, and MK Campbell. Transcription factor co-repressors in cancer biology: roles and targeting. *Int J Cancer.* 2010; 126(11): 2511-2519.

17. Green AR, Burney C, Granger CJ, Paish EC, El-Sheikh S, Rakha EA, Powe DG, Macmillan RD, Ellis IO, and E Stylianou. The prognostic significance of steroid receptor co-regulators in breast cancer: co-respressor NCOR2/SMRT is an independent indicator of poor outcome. *Breast Cancer Res Treat.* 2008; 110(3): 427-437.

18. Hanahan D and RA Weinberg. The hallmarks of cancer. *Cell.* 2000; 100(1): 57-70.

19. Hanahan D and RA Weinber. Hallmarks of cancer: the next generation. *Cell.* 2011; 144(5): 646-674.

20. Dallol A, Forgacs E, Martinez A, Sekido Y, Walker R, Kishida T, Rabbitts P, Maher ER, Minna JD, and F Latif. Tumour specific promoter region methylation of the human homologue of the Drosophila Roundabout gene DUTT1 (ROBO1) in human cancers. *Oncogene.* 2002; 21(19): 3020-3028.

21. Parri M and Paola C. Rac and Rho GTPAses in cancer cell motility control. *Cell Comm & Signaling.* 2010; 8(23): 1-14.

22. Blais ME, Dong T, and S Rowland-Jones. HLA-C as a mediator of natural killer and T-cell activation: spectator or key player? *Immunology*. 2011; 13(1): 1-7.

23. Fan QR, Long EO, and DC Wiley. Crystal structure of the human natural killer cell inhibitory receptor KIR2DL-HLA-Cw4 complex. *Nat Immunol*. 2001; 2(5): 452-460.

24. Gonzalez-Galarza FF, Christmas S, Middleton D, and AR Jones. Allele frequency net: a database and online repository for immune gene frequencies in worldwide populations. *Nucleic Acids Res.* 2011; 39: D913-919.

25. Novoselov SV, Kryukov GV, Xu XM, Carlson BA, Hatfield DL, and VN Gladyshev. Selenoprotein H is a nucleolar thioredoxin-like protein with a unique expression pattern. *J Biol Chem.* 2007; 282: 11960-11968.

26. Long PM, Moffett JR, Namboodiri AM, Viapiano MS, Lawler SE, and DM Jaworski. N-acetylaspartate (NAA) And N-acetylaspartylglutamate (NAAG) promote growth and inhibit differentiation of glioma stem-like cells. *J Biol Chem.* 2013; 288(44): 31916-31917.

27. Jiang XR, Song A, Bergelson S, Arroll T, Parekh B, May K, Chung S, Strouse R, Mire-Sluis A, and M Schenerman. Advances in the assessment and control of the effector functions of therapeutic antibodies. *Nat Rev Drug Discov.* 2011; 10(2): 101-111.

Chapter 3: New Methodology for the Enrichment and Characterization of O-GlcNAcylated Peptides

3.1 Introduction

3.1.1 O-GlcNAcylation

O-linked β-N-acetylglucosamine (O-GlcNAc) is a monosaccharide that can be added to serine or threonine residues on a protein. O-GlcNAc is ubiquitous, as it has been found in all multicellular organisms on proteins from virtually all cellular compartments.[1] Unlike most protein glycosylation, though, this sugar is not elongated into higher-order structures. Additionally, O-GlcNAc is attached and removed several times in the life of a protein, cycling rapidly and occurring at various residues. In this sense, the O-GlcNAc modification is seemingly closer to phosphorylation than it is to typical protein glycosylation.[2] Similar to the kinase/phosphatase dynamic described earlier in the dissertation, this moiety is added to an amino acid by O-GlcNAc transferase (OGT) and removed by O-GlcNAcase (OGA), depicted in Figure 3.1.[2] The OGT gene is one of the most highly conserved proteins from worm to man. Mammals and insects appear to have only a single gene encoding the catalytic subunit of OGT, and knockout of the gene is embryonic lethal at the single-cell level in mammalian cells.[3]
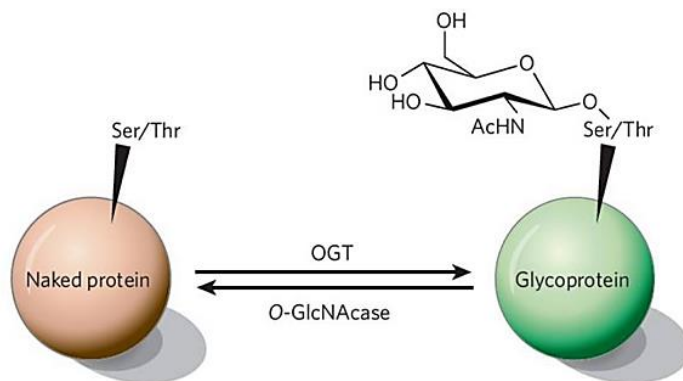

Figure 3.1 Dynamic cycling of O-GlcNAc. The addition of O-GlcNAc onto serine and threonine is catalyzed by OGT, whereas its removal is catalyzed by OGA. Control of this modification by these two enzymes is similar to that of phosphorylation.

O-GlcNAc is the endpoint of the hexosamine biosynthetic pathway, which culminates in the production of UDP-GlcNAc, the high-energy donor substrate for OGT. As demonstrated by Figure 3.2,[2] the biosynthesis of UDP-GlcNAc is affected and regulated by nearly every metabolic pathway in the cell, which renders OGT-catalyzed O-GlcNAcylation sensitive to insulin, nutrients, and cellular stress.[4] Thus, it is proposed that O-GlcNAcylation serves primarily to modulate cellular signaling and transcription in response to these factors.[5-7] This is supported by OGT's catalytic activity and peptide specificity, which are responsive to an extremely wide range of nucleotide concentrations, from nanomoles/L up to 100 millimoles/L. Additionally, when cells are exposed to almost any type of stress, including heat, high salt, heavy metals, UV light, and/or hypoxia, O-GlcNAcylation levels rapidly increase.[8]
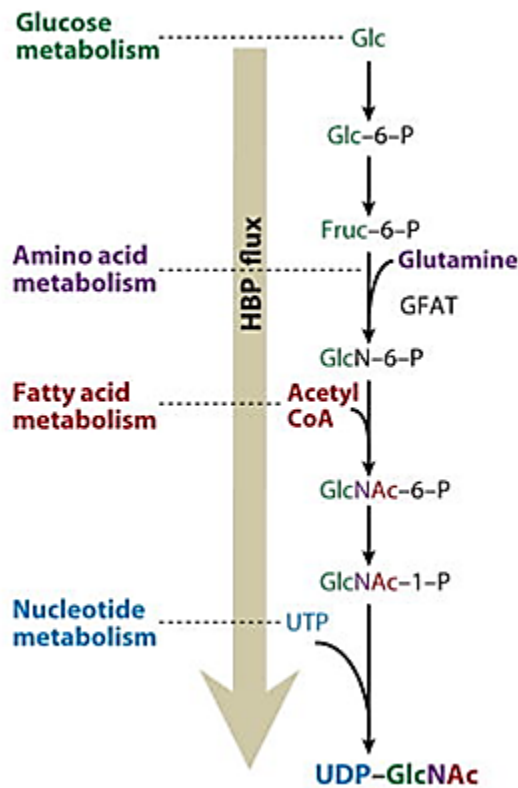
Figure 3.2 The hexosamine biosynthetic pathway. When glucose enters the cell, a small percentage (2-5%) is funneled to this pathway, where it is converted into UDP-GlcNAc. OGT then uses this substrate to modify proteins in response to cellular stress and nutrient availability. This figure highlights all the necessary metabolic components in the pathway, highlighting why O-GlcNAc is thought of as an extremely sensitive nutrient sensor.

O-GlcNAc modulates cellular functions through extensive cross-talk with signaling cascades that are also regulated by phosphorylation.[2,3] Site-mapping studies indicate a complex relationship exists between the two modifications, since they can compete for occupancy at the same amino acid, modify nearby sites, or influence protein activity while on distant residues. Notably, some phosphatases complex with OGT, suggesting the same enzyme complex can remove phosphate and attach O-GlcNAc.[3] Further, the two modifications have similar effects on cell cycle progression, signal transduction, and metabolism.[9] Normal cellular functioning relies on the tight control of these cellular processes and the post-translational modifications that regulate them. It follows that dysregulation of O-GlcNAcylation and its interplay with phosphorylation can play a significant role in chronic diseases such as diabetes, neurodegenerative diseases, cardiovascular disease, and cancer.[10-12] Its role in cancer will be discussed in Chapter 4.1.

Perhaps the best understood function of O-GlcNAc and disease is its role in diabetes.[12] Why hyperglycemia is deleterious to cells was previously an enigma since glucose itself is non-toxic. We now understand that many of the toxic effects of glucose require its conversion to glucosamine, which subsequently increases UDP-GlcNAc, and therefore O-GlcNAcylation.[13] Thus, current models suggest that prolonged, abnormal increases in O-GlcNAcylation due to hyperglycemia and/or hyperinsulinemia disturb the normal balance between O-GlcNAc and phosphorylation, ultimately leading to the toxicity associated with diabetes.[14,15]

O-GlcNAc is also implicated in neurodegenerative diseases such as Alzheimer's disease (AD). Tau, a microtubule-associated protein, is required for polymerization and

stability of microtubules in neurons. In normal brain, tau is extensively O-GlcNAcylated, as up to 12 sites of modification have been reported.[16] On the other hand, hyper-phosphorylated tau aggregates to form neurofibrillary tangles in AD neurons, promoting plaque formation.[17] Recent studies on human brain tau have established that O-GlcNAcylation negatively regulates phosphorylation in a site-specific manner, acting as a "cap" for serine and threonine residues. Additionally, quantitative studies between normal and AD brains showed a global decrease in O-GlcNAcylation and a concomitant increase in phosphorylation.[18]

Due to its ubiquitous nature and intimate association with human disease, it is understandable that O-GlcNAc is an attractive target of investigation. Despite this, the modification remains elusive and is often overlooked by those studying transcription and cell signaling. This can be explained by a number of technical challenges associated with studying protein O-GlcNAcylation that have severely hampered its detection and characterization. First, all cells contain high levels of hydrolases that rapidly remove the PTM from intracellular proteins when the cell is damaged or lysed. Thus, the modification is often lost during protein isolation. Secondly, O-GlcNAcylation generally does not affect the migration of polypeptides on SDS-PAGE or 2D electrophoresis, making its detection difficult by this process.[2,3,6]

More importantly, O-GlcNAc is particularly difficult to identify by traditional mass spectrometric methods since O-GlcNAcylated peptides are almost always sub-stoichiometric in comparison to the unmodified peptide or protein. To begin, they may actually be lower than the limit of detection (LOD), rendering them impossible to distinguish from background. Even if this is not the case, the instrument is often operated

in data-dependent mode, meaning that it will selectively fragment the top five most abundant species in a given MS1. Since unmodified peptides are usually much more abundant, they are the species that are selected for CAD and ETD. Without MS2 spectra, we are generally unable to diagnose that an O-GlcNAcylated species is present. This issue is exacerbated by ionization suppression, whereby the O-GlcNAcylated peptide can be partially or completely suppressed by the unmodified peptide. An example of this phenomenon is illustrated in Figure 3.3. Here, we hypothesize that during the ion desorption process (section 1.2.3), O-GlcNAc peptides favor the aqueous interior of the electrospray droplet, whereas unmodified peptides reside at the hydrophobic liquid-gas interphase, which is where desorption and ionization occur.[19]
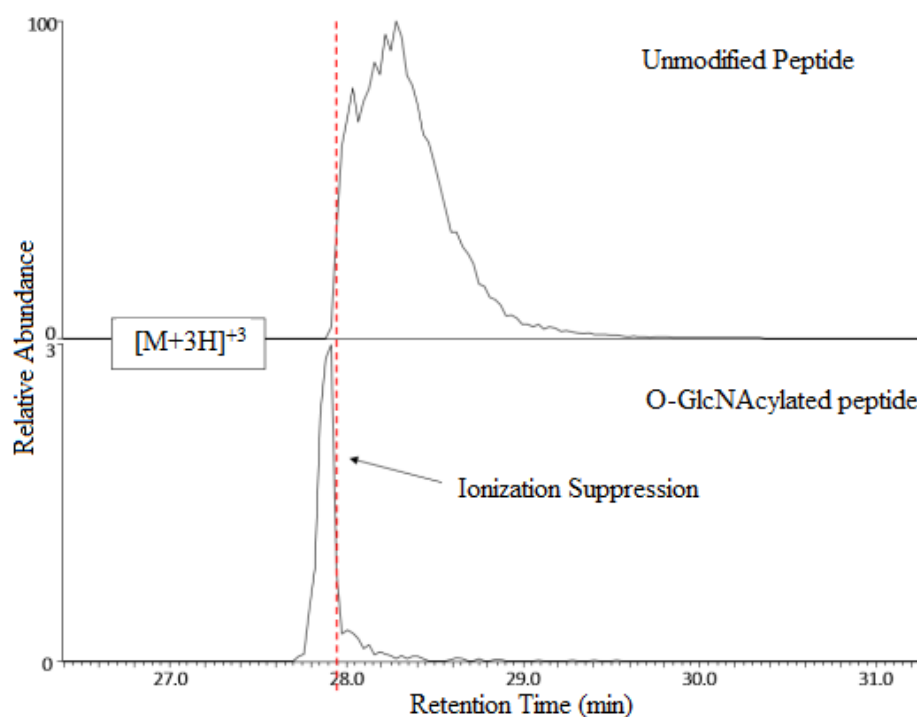
Figure 3.3 Ionization suppression of an O-GlcNAcylated peptide. In the bottom panel, singly O-GlcNAcylated peptide FGSSGLAPAHLGSNAFK begins to elute at 27.8 minutes, but is suppressed at 28 minutes by the unmodified version of the peptide. We believe this is due to the preference of O-GlcNAc for the hydrophilic interior of the electrospray droplet. (Adapted from AWD data)

Moreover, like phosphorylation, O-GlcNAc is highly labile during collision activated dissociation (CAD).[20] Figure 3.4A shows a typical CAD spectrum from an O-GlcNAcylated peptide. The most abundant signals in the spectrum correspond to charge-reduced peptide ions generated from the loss of GlcNAc that provide no sequence information. The product ion signal at m/z 204 corresponds to protonated O-GlcNAc which has been dissociated from the peptide. This information can be used for diagnostic purposes analogous to the "CAD Neutral Loss" program used for phosphorylation. Here, the program measures the abundance of the unmodified neutral loss peak ([M+H-GlcNAc]$^+$) and the oxonium ion. The user can then rank the abundances of these peaks,
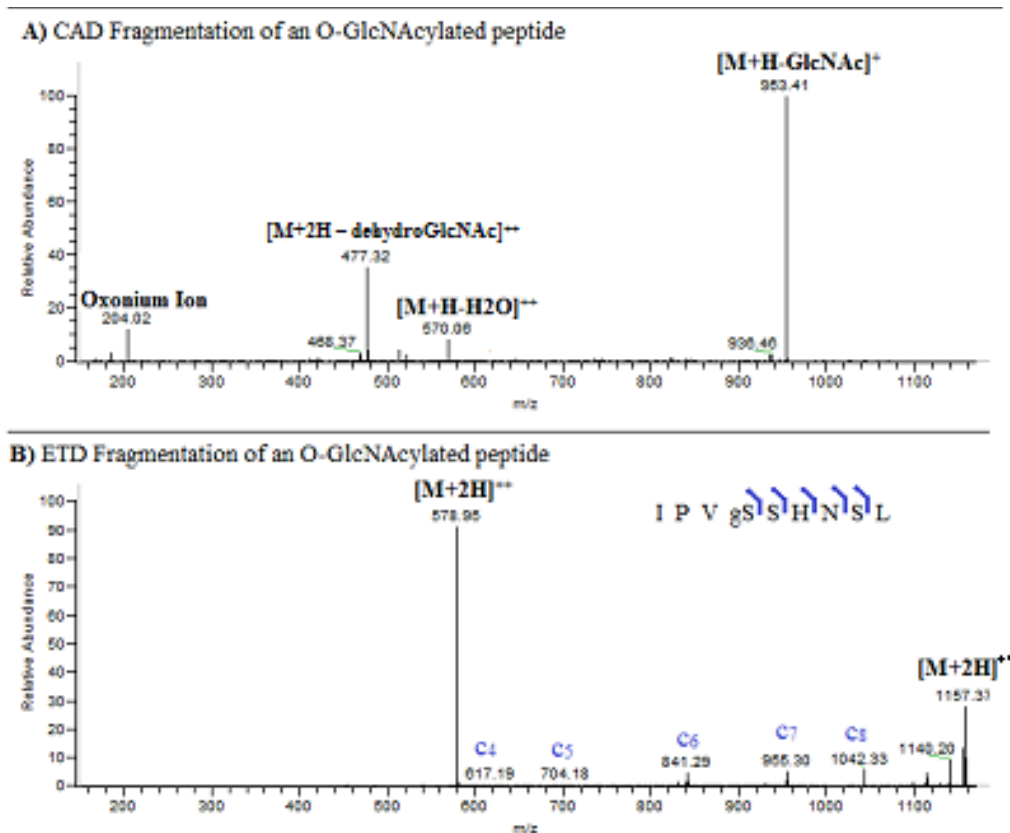
Figure 3.4 O-GlcNAcylated peptide fragmentation spectra. A) CAD fragmentation results in charged reduced species plus the oxonium ion at m/z 204. These can be used for diagnostic purposes, but do not allow for sequencing and/or site-localization of the modification. **B)** ETD fragmentation preserves the post-translational modification, which allows us to sequence the peptide. (Note: this is a +2 fragmentation spectrum; a +3 would provide even more information)

thus identifying likely O-GlcNAcylated peptide candidates. However, sequencing and site-localizing the modification with only a CAD spectrum is extremely difficult. Since ETD maintains the labile modification, we can utilize this fragmentation technique to glean more information about the parent peptide. Unfortunately, since O-GlcNAc usually resides in areas of high serine and threonine content, there is often low charge density in those peptides we detect. This results in incomplete ETD fragmentation and therefore incomplete sequence coverage of the peptide (shown in Figure 3.4B). Together, these issues make the identification, site-mapping, and study of O-GlcNAcylated proteins tremendously challenging.

3.1.2 O-GlcNAc Enrichment Methods

To ameliorate the above issues, enrichment methods are employed which serve to isolate O-GlcNAcylated peptides while reducing or eliminating non-O-GlcNAcylated peptide content. Several attempts have been made in this pursuit, but each enrichment technique is accompanied by a certain number of drawbacks. The common enrichment procedures include, but are not limited to: immunoprecipitation with a pan-O-GlcNAc antibody,[21-23] β-elimination Michael addition chemistry (BEMAD),[22] lectin weak affinity chromatography,[21,24,25] and chemoenzymatic tagging.[25-27] Common downfalls exist between all of these procedures, including: low reproducibility, low sensitivity, necessity of a high amount of starting material, significant non-specific binding, and either low recovery or a complete lack of internal standards. We seek to develop a robust, sensitive, and selective enrichment method that overcomes these difficulties.

Perhaps the most common enrichment technique utilizes lectin weak affinity chromatography. This involves assembling a ~12 meter column packed with wheat germ agglutinin (WGA), a carbohydrate-binding protein that is specific for sugar moieties. Often, the procedure involves loading milligrams of material onto the WGA column and fractionating the elution into 10-60 fractions, which are then individually subjected to RP-HPLC-MS/MS. [21,24,25] This is incredibly labor-intensive, requiring up to 60 hours of chromatography alone. Further, its specificity is greater for sialic acids (complex, branched saccharides) than O-GlcNAc residues, leading to significant non-specific binding.[24] Due to these downfalls coupled with the necessity for large amounts of starting material, WGA affinity chromatography was not considered for our studies.

Our laboratory has investigated another emergent technology that involves copper-catalyzed azide-alkyne cycloadditions ("click" chemistry), depicted in Figure 3.5.[26] Here, O-GlcNAc can be tagged by a modified galactosyltransferase, GalT1, which
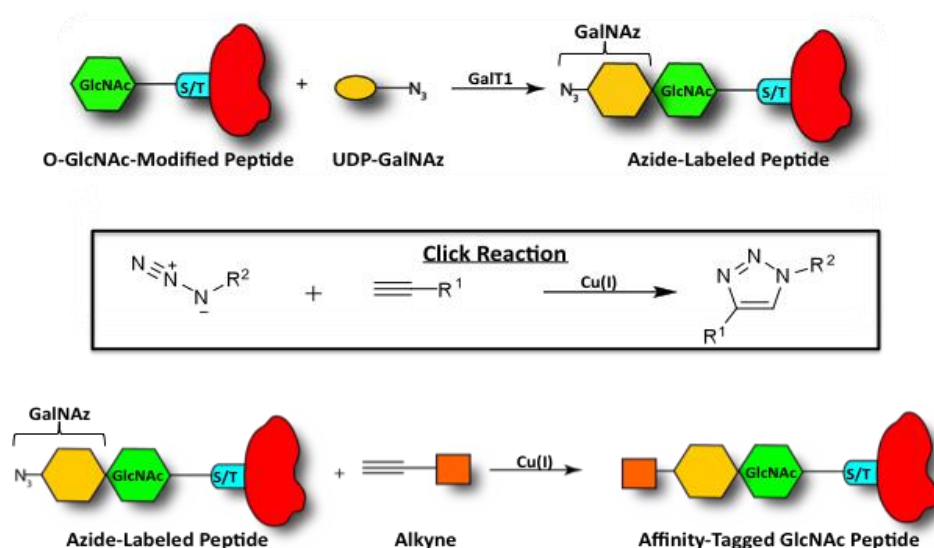
Figure 3.5 Click chemistry and affinity tagging. This figure depicts a schematic for the two-step labeling procedure for O-GlcNAc peptide enrichment. In the first reaction, O-GlcNAc is labeled with GalNAz using a modified galactosyltransferase enzyme. In the second step, click chemistry is used to label the GalNAz moiety with an alkyne attached to various affinity tags. (Taken from NDU dissertation)

enzymatically labels O-GlcNAc with N-azidoacetylgalactosamine (GalNAz). Once tagged, the azide group can react with an alkyne coupled to a number of affinity tags including: biotin, UV-cleavable biotin, and phosphate.[2,25-27] The biotin affinity tag was problematic due to low efficiency of elution from avidin, leading to extremely low yields. Additionally, the biotin tag readily fragments by ETD, making site-localization of O-GlcNAc difficult. The photocleavable biotin technique was carried out by the Hart lab and allowed for the detection of 141 O-GlcNAc sites in a complex sample.[2] Unfortunately, our laboratory was unable to recreate these results, presumably due to inefficient cleavage by UV light when biotin is attached to avidin beads. The final method used by Parker *et al* (2011) "clicks" O-GlcNAc-GalNAz with a phospho-alkyne. The phosphorylated peptides can then be enriched by standard phosphorylation-enrichment methods (i.e. IMAC).[26] We attempt to replicate and optimize the technique described in their publication.

Finally, we looked to boronate affinity chromatography, which has been utilized for decades in the enrichment of carbohydrates. In theory, under basic conditions (>pH 8) boronate is hydroxylated, yielding a tetrahedral boronate anion, which allows for the formation of esters with cis-diols. The bond can be reversed under acidic conditions (<pH 3) through hydrolysis of the covalent bond.[28-31] A schematic of this mechanism shown in Figure 3.6.
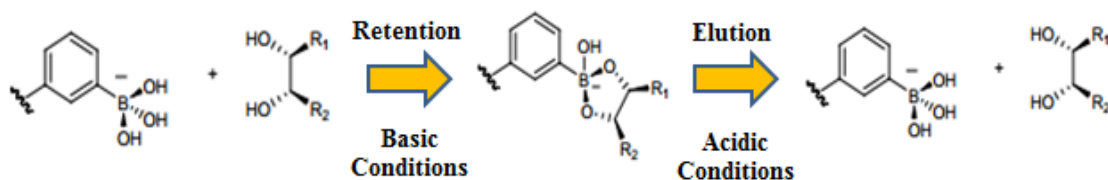
Figure 3.6 Proposed boronate affinity chromatography mechanism. Under basic conditions, a tetrahedral boronate ion can form a covalent bond with cis-diols. Non-glycosylated material can be washed away, and glycopeptides can be eluted by lowering the pH, which causes hydrolysis of the covalent bond.

Boronate affinity chromatography has been used for the separation of all four major groups of biomolecules, including glycoproteins. Several groups have reported successful separation of glycated hemoglobin from nonglycosylated hemoglobin.[32-34] Other groups have described isolation of glycosylated peptides from whole cell lysate.[29-31] Due to the success perpetuated in the literature coupled with the ease of use, we will attempt to use boronate affinity chromatography to enrich for O-GlcNAcylated peptides.

3.1.3 Project Aim

We seek to develop a robust, sensitive, and selective O-GlcNAc enrichment method that achieves quantitative yields of internal standards, thus overcoming difficulties associated with other techniques. To do this, we first plan to attempt optimization of the click chemistry reaction described by Palmer *et al*, where we will enzymatically label O-GlcNAc with GalNAz, then react this with a phospho-alkyne to yield a phospho-species that can be enriched by IMAC. Alternatively, we plan to isolate O-GlcNAcylated peptides via boronic acid chromatography.

3.2 Materials

3.2.1 Reagents

Amersham Pharmacia Biotech (Amersham, UK)

   NHS-activated sepharose magnetic beads

Applied Biosystems^{TM} (Carlsbad, CA)

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

   POROS® AL 20 perfusion chromatography bulk media, 20 μm diameter

Bruker (Billerica, MA)

   Boronic CLINPROT® Glyco-kit

Grace Davison Discovery Science (Deerfield, IL)

   Acetyl chloride, anhydrous

   D_0 methanol, anhydrous

Honeywell (Morristown, NJ)

   Acetonitrile, HPLC grade, ≥99.8% purity

Invitrogen (Carlsbad, CA)

   Click iT® Kit

   Galactosyltransferase (GalT1)

   Magnesium chloride

   Uridine diphosphate N-azidoacetylgalactosamine (UDP-GalNAz)

J.T. Baker (Phillipsburg, NJ)

   Glacial acetic acid, ≥99.9% purity

Millipore (Billerica, MA)

ProSep-PB Media

Pierce (Rockford, IL)

LC-MS grade water

PQ Corporation (Valley Forge, PA)

Kasil ® 1624 potassium silicate solution

Sigma Aldrich (Saint Louis, MO)

1-ethyl-3-(3-dimethylaminopropyl)carbodiimide

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

Ammonium hydroxide

Angiotensin I acetate salt hydrate, ≥90% purity

Ascorbic acid

Azulene

Ethanolamine

Iron (III) chloride

Phosphate buffered saline

Pyridine

Sodium cyanoborohydride

Vasoactive intestinal peptide fragment 1-12, ≥97% purity

Thermo-Fisher Scientific (Waltham, MA)

Aminophenyl boronic acid

Dimethylformamide

Hydrochloric acid

Immobilized boronic acid resin

YMC Company, LTD (Kyoto, Japan)

ODS-AQ, C18 5 μm spherical silica particles, 120 A pore size

ODS-AQ, C18 5-20 μm spherical silica particles, 120 A pore size

3.2.2 Equipment and Instrumentation

Agilent Technologies (Palo Alto, CA)

1100 Agilent high performance liquid chromatograph

Branson (Danbury, CT)

Branson 1200 Ultrasonic Bath

Labconco Corp. (Kansas City, MO)

Centrivap centrifugal vacuum concentrator

PolyMicro Technologies, Inc. (Phoenix, AZ)

360 μm o.d. x 50 μm i.d. polyimide coated fused silica capillary

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

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

Sutter Instrument Co. (Novato, CA)

P-2000 microcapillary laser puller with fused silica adapter

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

LTQ mass spectrometer (back end ETD)

LTQ FTICR hybrid mass spectrometer (custom modified with front end ETD)

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

LTQ-Orbitrap Velos 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 Peptide synthesis

O-GlcNAcylated peptides were synthesized using Fmoc chemistry, purified via HPLC to >90% purity, and confirmed by mass spectrometry. Once confirmed they were utilized as internal standards for all enrichment procedures. These included: IPVsSHNSL, RPPItQSSL, VLTsNVQTI, VLTsNVQTK, TVAtQASGLLSNK, APAVQHIVVtAAK, and APAtQAYGHQIPLR.

3.3.2 GalNAz addition reaction

The modified galactosyltransferase, GalT1, was provided by Peng Wu at Albert Einstein College of Medicine (NY, NY). 30 μL of the enzyme solution (concentration unknown) was dried to 5 μL in a vacuum concentrator at 40ºC. An excess of internal standard peptides (25 picomoles each) were dried down and brought up in 1 μL $MnCl_2$, 5 uL UDP-GalNAz, and 5 μL enzyme. The GalNAz addition reaction was then vortexed, centrifuged, and put on a shaker for 5 hours. The reaction was quenched by the addition of 0.2 μL glacial acetic acid and brought to a final volume of 15 μL with 0.1% acetic acid and. 1-5% of this reaction was loaded directly onto the precolumn C18 (5 μm diameter) connected to an analytical capillary column (360-μm outer diameter, 50-μm inner diameter) equipped with an electrospray emitter tip and analyzed via RP-HPLC ESI-MS/MS. Peptides were eluted over a 40 minute gradient of 0-60% solvent B (70% acetonitrile, 0.1 M acetic acid) at 60 nL/min and electrospray ionized into the mass spectrometer. Relative abundances were calculated by comparing peak areas of each

phosphopeptide to internal standards (angio and vaso) which are at a fixed concentration of 100 fmol each. Percent recoveries were determined by comparing the relative abundance to expected values.

3.3.3 C18 cleanup of samples

The above reaction was taken to dryness and brought up in 50 μL 0.1% acetic acid. This solution was pressure loaded onto a fused silica clean-up column (360 μm o.d. x 150 μm i.d.) packed to 5 cm with C18 resin (5-20 μm diameter particles) at a flow rate of ~0.5 μL/min. The column loaded with sample was rinsed with 25 μL 0.1% acetic acid at the same flow rate. Then the column was moved to an HPLC where it was rinsed with solvent A for 10 minutes at 12-15 bar. Peptides were eluted into an Eppendorf tube using a 40 minute gradient (0-80% solvent B) that was held at 80% B for 30 additional minutes. The column was washed with 80% B and 20% A for an additional 20 minutes to clear any remaining contaminants from it. The sample was taken to dryness in the speedvac and was stored at -35ºC until further use.

3.3.4 Click chemistry

The dried O-GlcNAc-GalNAz standard peptides were then brought up in 6 μL HEPES buffer (pH 8), 5 nmol alkyne (propargyl phosphate; 1 μL), 10 nmol sodium ascorbate (1 μL), 1 nmol coordinating compound (1 μL), and 10 nmol CuSO4 (1 μL) for a total reaction volume of 10 μL. The coordinating compounds tested included TBTA (tris[(1-benzyl-1H-1,2,3-triazol-4-yl)methyl]amine), BTTP (3-[4-({bis[1-*tert*-butyl-1H-

1,2,3triazol-4-yl)methyl]amino}methyl)-1H-1,2,3triazol-4-yl]propanol), and BTTAA (2-[4-({bis[(1-tert-butyl-1H-1,2,3-triazol-4-yl)methyl]amino}methyl)-1H-1,2,3-triazol-1-yl]acetic acid). BTTP and BTTAA were provided by Peng Wu at Albert Einstein College of Medicine (NY, NY). In the case of BTTAA, 100 μL agarose beads were washed 3x with 1 mL H_2O before addition to reaction. After 1.5 hours the reaction was quenched by the addition of 0.2 μL glacial acetic acid. 5% of this reaction was loaded directly onto an AC-PC and analyzed in a similar manner described above. Peptides were eluted over a 40 minute gradient of 0-60% solvent B (70% acetonitrile, 0.1 M acetic acid). Relative abundances were calculated by comparing peak areas of each phosphopeptide to internal standards (angio and vaso) which are at a fixed concentration of 100 fmol each. Percent recoveries were determined by comparing the relative abundance to expected values. After screening the reaction, samples were again subjected to a C18 cleanup (section 3.3.3).

3.3.5 Esterification and IMAC

Once the O-GlcNAc-GalNAz has been clicked with a phospho-alkyne and cleaned up, the resulting peptides were then esterified and enriched via IMAC. Samples were esterified using a Fischer esterification. First, samples were dried in triplicate using 50 μL of completely anhydrous methanol. Concurrently, methanolic HCl was prepared by adding 160 μL dropwise to 1 mL of methanol. 80 μL of this solution was added to the dried sample and the reaction was allowed to proceed for 1 hour. At this time, the sample was dried and rinsed once with 50 μL methanol. The entire process was repeated to

assure completeness. The esterified and dried sample was stored at -35ºC until further use.

For each IMAC experiment, a 360 μm o.d. x 75 μm i.d. polyimide coated fused silica capillary fritted with a 2 mm Kasil frit was pressure loaded (200-500 psi) with 5 cm of POROS20 iminodiacetate packing material. The IMAC column was then prepared for iron (III) activation by the following rinse steps: 10 minute H_2O at 20 μL/min, 20 minute 50 mM EDTA at 20 μL/min, and 5 minute H_2O at 20 μL/min. At times that the frit did not allow for 20 μL/min flow rate, the rinse time was adjusted to allow the same amount of column volumes to pass through the IMAC column. For instance, a 10 μL/min flow rate would require the first H_2O rinse to be 20 minutes, totaling 200 μL through the column. The IMAC column was then activated using filtered iron (III) chloride ($FeCl_3$). Each activation step included a 10 minute $FeCl_3$ rinse at 20 μL/min followed by 3 minutes without pressure to allow the $FeCl_3$ to diffuse into the column particles. This was repeated three times to assure complete activation. The activated column was equilibrated with 0.01% acetic acid at 0.5 μL/min for 50 minutes.

Dried, esterified samples were brought up in 50 μL 1:1:1 0.01% acetic acid: methanol: acetonitrile. This solution was loaded onto the column at 0.5 μL/min. A final rinse of 0.01% acetic acid at 0.5 μL/min was performed for 30 minutes. At this point, a PC is disconnected from an AC and rinsed with solvent A for 10 minutes at 30 bar. Then, a PC was disconnected from an AC and rinsed with solvent A for 10 minutes at 30 bar. The PC was then butt-connected with Teflon to the top of the IMAC column. The IMAC-PC was allowed to rinse for 10 minutes with 0.01% acetic acid at a flow rate of 1 μL/min.

If a leak exists during this time, the pressure was let off and a new Teflon sleeve was created. Phosphopeptides were eluted with 10 μL of 250 mM ascorbic acid solution for 10 minutes (1 μL/min flow rate). The IMAC-PC was rinsed for another 5 minutes with 0.01% acetic acid. The PC was then removed, rinsed with solvent A for 20 minutes at 30 bar, dried, and re-attached to an AC. The column was allowed to rehydrate for 20-30 minutes before the addition of 100 fmol of internal standards (angio and vaso). The column was rinsed again for 10 minutes before elution.

3.3.6 Phenylboronic acid solid phase extraction

For the following analyses, all supernatants were removed using gel-loading tips and extreme care was taken so as not to disturb the bead pellets. Further, all fractions (flow-through, wash, and elution, where applicable) were taken to dryness in a vacuum concentrator, brought up in 10 μL 0.1% acetic acid, and loaded onto a PC. The PC was rinsed for 10-15 minutes with solvent A (0.1 M acetic acid), dried, and connected with Teflon to a bottleneck-fritted analytical column (AC) (360 μm o.d. x 50 μm i.d) packed with C18 packing material (5 μm diameter particles) equipped with a laser-pulled electrospray emitter tip. Peptides were eluted over a 40 minute gradient of 0-60% solvent B (70% acetonitrile, 0.1 M acetic acid). Intact peptide masses were acquired in the FTICR or the Orbitrap, and MS2 spectra were acquired in the LTQ. Mass spectra were acquired using a method consisting of one high resolution scan (resolving power of 60,000 at 400 m/z) followed by 2 data dependent low resolution scans; the most abundant parent ion was subject to both CAD and ETD. Relative abundances were calculated by

comparing peak areas of each phosphopeptide to internal standards (angio and vaso) which were at a fixed concentration of 100 fmol each. Percent recoveries were determined by comparing the relative abundance of O-GlcNAcylated peptides to expected values. In later experiments, unmodified peptides were added and recoveries were measured in order to assess non-specific binding.

3.3.6.1 Procedure I: On-column, pH elution

A 360 μm o.d. x 150 μm i.d. polyimide coated fused silica capillary fritted with a 2 mm Kasil frit was pressure loaded (200-500 psi) with 1 cm of Thermo boronate coated polyacrylamide beads and washed with H_2O for 5 minutes. The column was then equilibrated with buffer A (250 mM ammonium acetate, pH 8.1-8.3). Dried sample (peptides with either O-GlcNAc (Ia) or O-GlcNAc-GalNAz (Ib)) was brought up in buffer A and loaded slowly onto the column (<0.5 μL/min). The flow-through was collected and dried for later analysis if necessary. The column was rinsed with buffer A for 10 minutes at a similar flow rate. The sample was eluted with 0.1 M acetic acid, and the elution was collected in an Eppendorf tube.

3.3.6.2 Procedure II: Batch, pH elution

10 μL of Millipore ProSep (IId) or Thermo boronate coated polyacrylamide beads (IIa,b,c) were diluted with 100 μL of buffer A (250 mM ammonium acetate, pH 8.1-8.3) with or without the addition of 50mM $MgCl_2$ (IIc) and spun down for 5 minutes at 7500 rpm. The supernatant was removed, and the wash was repeated twice more. Dried sample (IPV, RPP peptides with either O-GlcNAc or O-GlcNAc-GalNAz) were brought up in 20

µL buffer A and added to the pre-washed beads. The reaction was allowed to progress for 1 or 24 hours (IIb) with agitation, and then spun down for 5 minutes at 7500 rpm. The supernatant was removed and saved for future analysis. Beads were then resuspended in 100 µL buffer A and spun for 5 minutes at 7500 rpm. This was repeated, and wash supernatants were saved as well. Finally, the beads were resuspended in 20 µL elution buffer (0.1 M acetic acid) and agitated for 30 minutes. The reaction was spun down for 5 minutes at 7500 rpm and rinsed once with an additional 20 µL elution buffer.

3.3.6.3 Procedure III: Batch, water elution

10 µL of Millipore ProSep silica or Thermo boronate coated polyacrylamide beads were diluted with 100 µL of anhydrous DMF and spun down for 5 minutes at 7500 rpm. The supernatant was removed, and the wash was repeated twice more. Dried sample (IPV, RPP peptides with either O-GlcNAc or O-GlcNAc-GalNAz) were brought up in 20 µL DMF and added to the pre-washed beads. The reaction was allowed to progress for 1, 4, 8, or 24 hours with agitation, and then spun down for 5 minutes at 7500 rpm. The supernatant was removed and saved for future analysis (if necessary). Beads were then resuspended in 100 µL acetonitrile and spun for 5 minutes at 7500 rpm. This was repeated, and wash supernatants were saved for future analysis as well. Finally, the beads were resuspended in 20 µL elution buffer (0.1 M acetic acid) and agitated for 30 minutes. The reaction was spun down for 5 minutes at 7500 rpm and rinsed once with an additional 20 µL elution buffer.

In some cases, the sample was esterified before PBA enrichment. In these instances, the Fisher esterification protocol was used (section 3.3.5). Once complete,

samples were brought up in H_2O, pH was adjusted to >8 with ammonium hydroxide, and they were again taken to dryness.

3.3.6.4 Procedure IV: Batch, magnetic beads, water elution

NHS-activated magnetic beads were derivatized as described. First, 14.6 mg MES was dissolved in 3 mL H_2O and pH adjusted to 5 with ammonium hydroxide. 100 μL of the NHS activated beads were removed and beads were separated from the supernatant using an external magnet. The beads were then washed 2x with 25 mM MES. 1 mg of aminophenyl boronic acid (APBA) was brought up in 100 μL 25 mM MES and added to the washed beads. This solution was slow vortexed overnight at 4ºC. The following day, 0.05 M ethanolamine in PBS was prepared (90.7 μL ethanolamine in 2.9 mL PBS, pH adjusted to 8 with HCl). MES solution was removed from the beads using the external magnet, and 100 μL PBS solution was added to the beads. This mixture was incubated at room temperature for 60 minutes, then washed 2x with PBS solution. The beads were stored at 4ºC in DMF until enrichment was performed. Enrichments were performed as described in section 3.3.6.3; however, instead of centrifuging the Eppendorf to remove supernatant, the external magnet was used.

Later experiments involved capping underivatized CO_2 groups on the magnetic beads using EDC-mediated derivatization with ethanolamine. Here, pyridine HCl buffer was made using 836 μL H_2O, 81 μL pyridine, and was pH adjusted with HCl to 5.5. This was used to prepare 1 M ethanolamine solution in the pyridine HCl buffer (60.5 μL ethanolamine into 939.5 μL pyridine HCl), and was again pH adjusted with HCl to 5.5. Four vials of 0.1M EDC HCl were made by dissolving 4.79 mg EDC in 250 μL H_2O. To

100 μL magnetic beads washed 3x with DMF, 20 μL 1M ethanolamine and 5 μL 0.1M

EDC were added and allowed to react with agitation. Every 30 minutes, a fresh 5 μL

aliquot of 0.1M EDC was added to the reaction. After 2 hours, the reaction was quenched

by removing excess solution and rinsing 3x with DMF. Beads were stored at 4ºC in DMF

until enrichment was performed.

3.3.6.5 Procedure V: Batch/on-column, POROS 20 beads, water elution

A Schiff's base reaction was performed on POROS AL 20 beads to couple the

free amine on APBA with the aldehyde moiety on the POROS beads. First, PBS buffer

was prepared and pH adjusted to 6-7. Then, 1 mL 0.2M APBA (30.99 mg) was prepared

in PBS buffer and pH adjusted with NaOH to 6-7. Additionally, 100 μL of 80mg/mL

$NaCNBH_3$ was made by dissolving $NaCNBH_3$ in the PBS buffer. Finally, approximately

7 mg of POROS AL beads were dispersed into 200 μL 0.2M APBA solution, followed by

the addition of 1 μL of 80mg/mL $NaCNBH_3$. The reaction was allowed to progress for 2

hours at room temperature with agitation, and was quenched by washing the beads on a

spin column (pore size <20 μm) with water filtration 3x. After removing water and

capping the filtration column, beads were stored dry at 4ºC. Once needed, the beads were

brought up in 500 μL water, and 100μL of this solution was added to 400 μL water. 100

μL of the dilution solution was spun down, water was removed, and the beads were

washed 3x with DMF. Some experiments involved derivatizing the APBA-POROS 20

beads using the EDC chemistry described in 3.3.6.4. Enrichments were then performed as

described in section 3.3.6.3, except elution fractions were spun down at 9500 rpm.

The beads were also tested on-column by pressure loading (200-500 psi) a 360 μm o.d. x 75 μm i.d. polyimide coated fused silica capillary fritted with a 2 mm Kasil frit with 5 cm of APBA-derivatized POROS 20 beads. The column was rinsed for 5 minutes at a flow rate of 20 μL/minute at 400 psi, then equilibrated with DMF for 25 minutes at a flow rate of 16 μL/minute. Dried sample was brought up in 50 μL DMF and loaded slowly onto the equilibrated column (< 1μL/minute). An additional 25 μL DMF was added to the sample tube and loaded onto the column, again at the slow flow rate. The column was then rinsed with acetonitrile for 30 minutes at < 1μL/minute. Finally, samples were eluted from the column with 15μL 0.1M acetic acid, again at a flow rate of < 1μL/minute. The elution was collected and dried for analysis.

3.3.7 Phenylboronic acid enrichment of digested proteins

To test the PBA enrichment procedure, we used enzymatically digested proteins including Lamin A (an O-GlcNAcylated protein) and BSA (a non-GlcNAcylated protein). The digest of Lamin A was prepared by Amanda Wriston, wherein 5 pmol protein was digested by AspN for 8 hours at room temperature. Wriston also performed a 1 pmol screen of the sample; these values were used to calculate expected values in enrichment experiments. 50 pmol BSA was digested for 8 hours at room temperature using a 1:20 ratio of trypsin:BSA in 100 mM ammonium bicarbonate. Again, a 1% (500 fmol) screen was run on this sample to determine expected values in enrichment experiments.

After several failed attempts at enrichment of either protein, a C18 clean up step was added prior to enrichment to remove any salts present in the digests. Dried digested protein was brought up in 50 µL 0.1% acetic acid. This solution was pressure loaded onto a fused silica clean-up column (360 µm o.d. x 150 µm i.d.) packed to 5 cm with C18 resin (5-20 µm diameter particles) at a flow rate of ~0.5 µL/min. The column loaded with sample was rinsed with 25 µL 0.1% acetic acid at the same flow rate. Then the column was moved to an HPLC where it was rinsed with solvent A for 10 minutes at 12-15 bar. Peptides were eluted into an Eppendorf tube using a 40 minute gradient (0-80% solvent B) that was held at 80% B for 30 additional minutes. The column was washed with 80% B and 20% A for an additional 20 minutes to clear any remaining contaminants from it. The resulting tube was stored at -35ºC until enrichment was performed.

Enrichment on these samples was performed using procedure IV and V with and without ethanolamine capping. Additionally, BSA was tested using APBA-derivatized and aldehyde ("bare") POROS AL 20 beads, with and without esterification.

3.4 Results

3.4.1 Click chemistry enrichment

A schematic for click chemistry O-GlcNAc enrichment is shown in Figure 3.7.[26]

Here, a GlcNAcylated peptide is labeled with GalNAz by GalT1 to yield an O-GlcNAc-GalNAz peptide. We then perform a C18 clean up before utilizing copper (I) click chemistry to add a phosphate onto the GalNAz moiety. Once clicked, the phosphate-containing peptides are subjected to another C18 clean up, Fischer esterification, and iron (III) IMAC. Finally, peptides are analyzed by high-resolution mass spectrometry.

The first goal of this study was to optimize this procedure for small (8-15 amino acid) GlcNAcylated peptides. To be useful for future experiments, the overall recovery must be at least 30% from start to finish. Thus, each step

Figure 3.7 Click chemistry enrichment procedure. The GlcNAcylated peptides are first enzymatically labeled with GalNAz by GalT1, and then clicked with propargyl phosphate to yield a terminal phosphate. IMAC is performed to enrich phosphopeptides prior to analysis by mass spectrometry. The goal of this section is to optimize this procedure and see recovery of at least 30% from start to finish.

was tested for ideal conditions and reagents. Following each experiment, we screened 1% of the reaction to assess conversion and recovery. Conversion percentage is defined as the amount of desired product divided by desired product plus the starting material. Recovery is the amount of desired product divided by the amount of starting material. We started with the GalT1 enzymatic addition of GalNAz and found that using 50-100 μL of

enzyme resulted in a 100% conversion; however, the yield of desired product was only 60-80% in all conditions tested. We hypothesized this is due to a decrease in ionization efficiency caused by the addition of another sugar moiety. We decided to move forward, and cleaned up the O-GlcNAc-GalNAz peptides. Again, we screened 1%, but we only recovered 60-80% of all internal standards. Thus, the recovery of O-GlcNAcylated peptides up to this point is only 36-64%.

Despite this mediocre recovery, we moved on to optimization of the click reaction. Namrata Udeshi and Andrew Dawdy, former members of the Hunt laboratory, had previously attempted to identify optimal conditions for this method. They discovered ratios of azide, alkyne, coordinating compound, sodium ascorbate, and $CuSO_4$ (described in section 3.3.4) that consistently result in 100% conversion of starting material. These reaction conditions were repeated, but two downsides became apparent. First, the Click IT ® kit coordinating compound, TBTA (Figure 3.7, top right panel) cannot be purified via C18 cleanup and is present in the elution. This could potentially suppress any co-eluting O-GlcNAcylated peptides. Additionally, while the conversion is 100%, the recovery of starting material is only 30-40%. Thus, side reactions and/or loss of material are occurring throughout this process. Due to these downfalls, we sought to test other coordinating compounds in search of one that converts and recovers a higher percentage of O-GlcNAc-GalANAz to a phosphate containing peptide.

Depicted in Figure 3.7 are the various tris(triazolylmethyl)amine-based ligands that we attempted to use: TBTA (as discussed above), BTTP, and BTTAA. The latter molecules were synthesized by Peng Wu's lab at Albert Einstein College of Medicine (NY, NY). According to Wang *et al,* BTTP is a highly efficient and reliable catalyst for

this reaction.[35] Unfortunately, under all conditions tested (200x, 500x, 1000x, 2000x catalyst:azide ratio), the conversion was once again nearly 100%, but the recovery ranged from 20-40%. Similar to TBTA, BTTP binds to a C18 column, but BTTP elutes in the center of the gradient, making the likelihood of suppression even higher than that of TBTA. The Wu lab synthesized another ligand (BTTAA) attached to agarose beads in order to minimize C18 contamination. Unfortunately, even after washing the beads extensively, most peptides were lost in the reaction. Severe contamination was discovered in the screen, and only one peptide was found with a 57% conversion and a 13% recovery. Thus, in spite of its drawbacks, TBTA proved to be the best coordinating ligand of those investigated.

To measure the overall recovery from start to finish, we did an experiment using optimized conditions. We performed the GalT1 reaction, click reaction, two clean ups, esterification, and IMAC. The recovery from the IMAC alone was 20%. Thus, including all losses, the average recovery for the 3 O-GlcNAcylated internal standards was 2%. This is an unacceptable figure in any situation where the starting material is on the order of femtomoles. We abandoned this procedure in favor of one with fewer steps, higher recovery, and less chemical contamination.

3.4.2 Phenylboronic acid enrichment

In theory, under basic conditions boronate is hydroxylated, yielding a tetrahedral boronate anion, which allows for the formation of esters with cis-diols. The bond can be reversed under acidic conditions through hydrolysis of the covalent bond. Three commercially available solid phases derivatized with phenylboronic acid (PBA) were

tested on-column (procedure I) and in batch (procedure II): polyacrylamide (Thermo), silica (Millipore), and magnetic beads (Bruker). These protocols bound peptides in 250 mM ammonium acetate (pH 8.5), washed twice, and eluted using 0.1M acetic acid (pH 3.7). Under every condition tested using this technique, <3% of standard O-GlcNAc peptides were detected in the elution. Specific trial conditions are listed in section 3.3.

Procedure	Trial	Flow-through	Recoveries (%) Wash	Elution
I	a	97	nd	0
	b	100	nd	0
II	a	92	nd	3
	b	65	nd	0
	c	67	nd	0
	d	88	nd	1

Table 3.1 Recoveries for procedures I and II. Using the published method of basic pH binding/acidic pH elution, the elution recovery was always less than 3%. Naerly all of the remaining peptides were found in the flow-through. Washes were not investigated (nd = no data).

Due to reasons discussed in section 3.5, we revised our protocol to bind and wash under anhydrous conditions and elute with water. The recoveries dramatically improved, as illustrated by Table 3.2. The best recoveries were achieved with Thermo polyacrylamide beads, a 1 hour bind time, and without esterification. No difference in recovery was noted between O-GlcNAcylated peptides compared to those tagged with GalNAz. However, in each instance, five unmodified peptides were added to the reaction in various amounts. Even with the optimized conditions, all of the unmodified peptides were found in the elution, suggesting that the Thermo beads afford some amount of nonspecific binding. We then decided to explore other bead types to decrease the frequency and level of interfering peptides.

Procedure	Trial	Flow-through	Wash	Elution	Notes
III	a	63	nd	39	ProSep
	b	41	7	52	Thermo
	c	nd	nd	3	with esterification
	d	39	0	2-6*	with esterification and pH adjustment
	e	nd	nd	73	1 hour bind
	f	nd	nd	46	4 hour bind
	g	nd	nd	46	8 hour bind

(Column header group: Recoveries (%) spanning Flow-through, Wash, Elution)

Table 3.2 Procedure III O-GlcNAc peptide recoveries. Using the new protocol, we consistently achieve >40% yield of O-GlcNAc peptides. We attempted to optimize the procedure. Of those condiditons tested, Thermo beads with 1 hr bind time and no esterification provided the highest recovery.

We first looked to NHS-activated magnetic beads derivatized in-house with aminophenyl boronic acid. The beads are derivatized as shown in Figure 3.8 and described in section 3.3.6.4. Results from these experiments are shown in Table 3.3, including elution recovery of O-GlcNAcylated peptides and non-specific binding (NSB). Trial IVa sought to demonstrate that O-GlcNAcylated peptides are actually binding to the APBA and are not
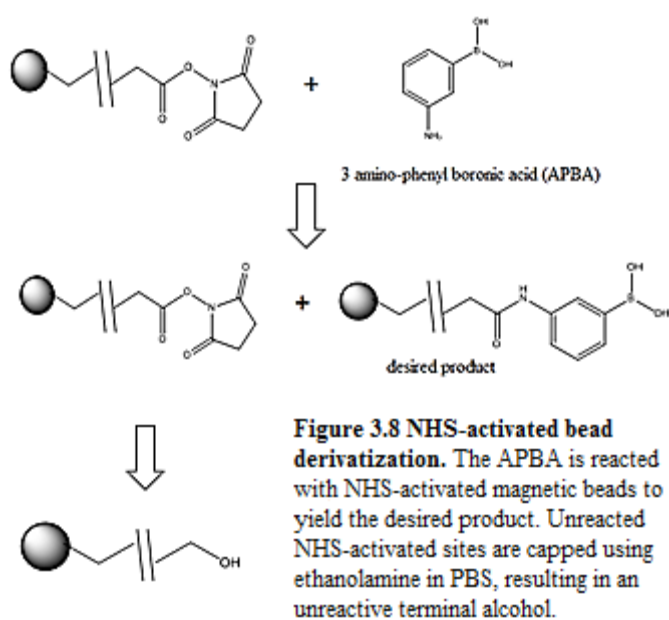
Figure 3.8 NHS-activated bead derivatization. The APBA is reacted with NHS-activated magnetic beads to yield the desired product. Unreacted NHS-activated sites are capped using ethanolamine in PBS, resulting in an unreactive terminal alcohol.

non- specifically binding to the solid phase. On the derivatized beads, there is a significant increase in O-GlcNAc binding (28%), whereas the unmodified peptide recovery did not change. We then wanted to optimize the magnetic bead procedure in Trials IVb, c, and d. Recoveries were highest for 10 μL beads, 1 hour bind time, and pH 3 elution. However, in each of these cases, the nonspecific binding was very high.

Procedure	Trial	Recoveries (%)				
		Flow-through	Wash	Elution	NSB	Notes
IV	a	nd	nd	13	27	underivatized beads
		59	0	41	28	derivatized beads
	b	43	nd	57	40	10 µL beads
		57	nd	43	30	20 µL beads
		48	nd	52	30	30 µL beads
	c	nd	nd	33	32	1 hour bind
		nd	nd	43	38	4 hour bind
		nd	nd	42	32	8 hour bind
	d	nd	nd	60	35	pH 3 elution
		nd	nd	34	35	pH 5 elution
		nd	nd	39	34	pH 8 elution
	e	nd	nd	37	19	EtNH2 capping

Table 3.3 Recoveries of O-GlcNAcylated peptides from procedure IV experiments. These experiments show (1) O-GlcNAcylated peptides bind specifically to APBA magnetic beads, (2) 10 uL beads, 1 hour bind time, and pH 3 elution are optimal, and (3) EDC-mediated EtNH2 capping can decrease non-specific binding by approximately 50%.

We hypothesized this may be due to free carboxylic acids on the beads, which form as the NHS ester hydrolyzes in water. Thus, we decided to cap these sites with EDC-mediated derivatization with ethanolamine. This reaction is shown in Figure 3.10 and described in section 3.3.6.4. After derivatization, non-specific binding decreased by approximately 50%.

Magnetic beads have several advantages over polyacrylamide or silica beads. To begin, they do not require centrifugation and can be separated from solvent by an external magnet. Additionally, the background is much lower, which is demonstrated in Figure 3.10. In the left panel (polyacrylamide beads), the base peak contains several contaminants that may suppress O-GlcNAcylated peptides in a real sample. However, in the right panel (magnetic beads) only the peptides being investigated are seen in the base peak. Thus, we decided to move forward with derivatized magnetic beads on a real sample.

Figure 3.9 Mechanism for EDC-mediated ethanolamine capping. Free carboxylic acids are capped via EDC-mediated derivatization with ethanolamine. This yields another unreactive alcohol and an urea, which is removed by washing the beads. This reaction allowed us to decrease non-specific binding by ~50%.
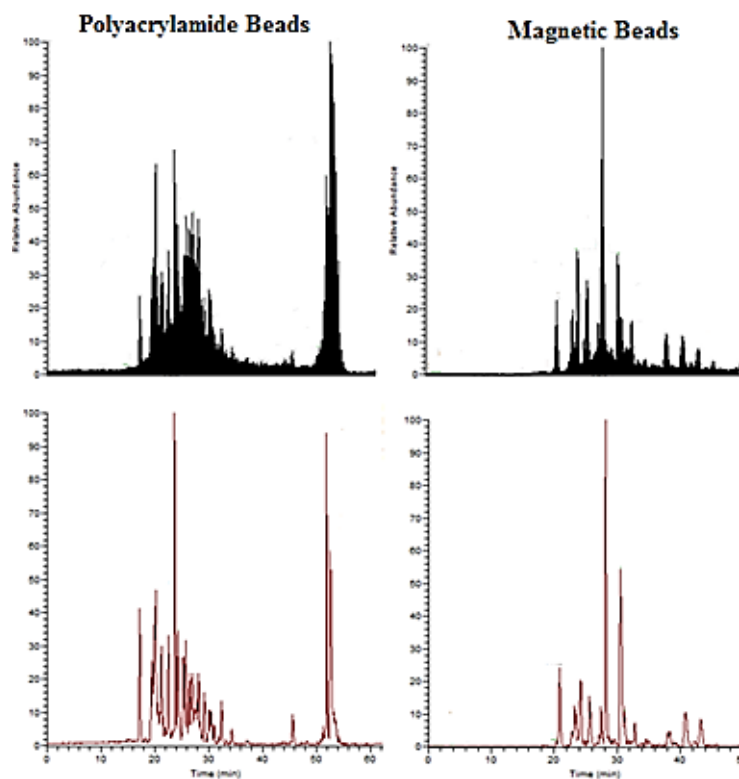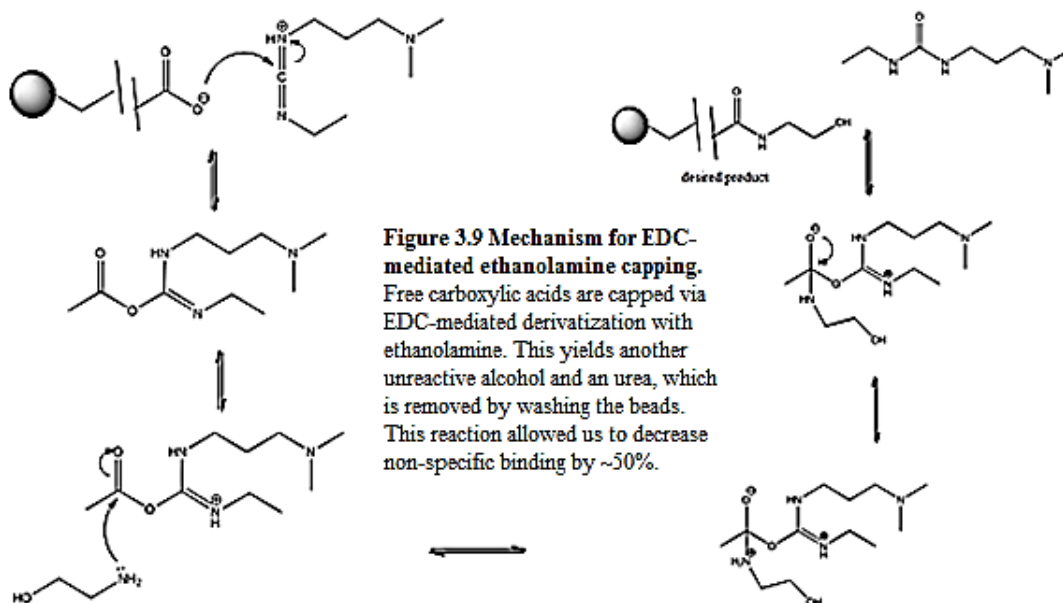
Figure 3.10 Total ion current and base peak chromatographs for polycrylamide vs magnetic beads. Top panel: TIC, bottom panel: base peak. Left panel: polyacrylamide beads, right panel: magnetic beads. The polyacrylamide beads have several contaminating agents that could decrease the likelihood of detecting O-GlcNAcylated peptides

To test our method, we digested an O-GlcNAcylated protein, Lamin A, with endoproteinase AspN. AspN generates large peptides with an N-terminal aspartic acid residue. Previously, Amanda Wriston identified several O-GlcNAcylated residues in this digested sample. We sought to identify known Lamin A O-GlcNAc sites, detect more O-GlcNAc sites, and recover O-GlcNAcylated internal standards. Unfortunately, only 5% of standards were recovered, and only one O-GlcNAcylated peptide was detected. Thus, a C18 clean up step was added prior to enrichment to remove any salts present in the digest. After cleaning up, we obtained a 60% average recovery of O-GlcNAcylated standards, but only found 4 O-GlcNAcylated Lamin A peptides at an average recovery of 9%. We concluded that these magnetic beads were not optimal for real samples.

Finally, we derivatized POROS 20 AL beads with APBA using reductive amidation. This reaction is depicted in Figure 3.11. On-column we obtained a 22%
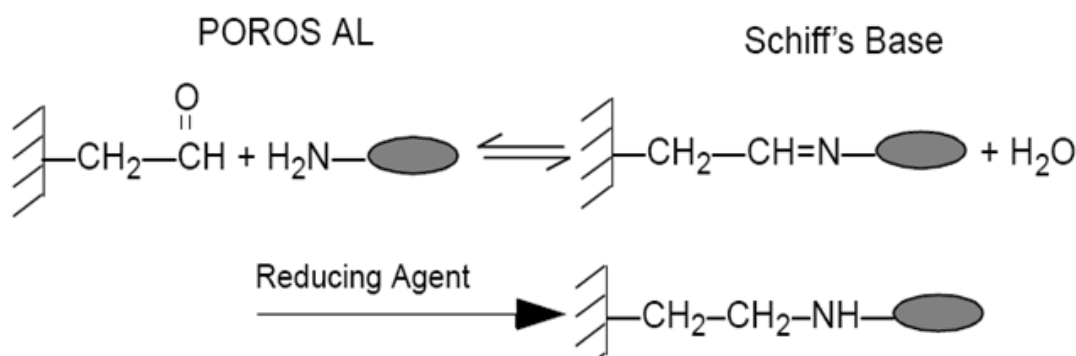
Figure 3.11 POROS20 AL derivatization. Free amine on aminophenylboronic acid (APBA) reacts with free aldehyde on POROS20 AL beads to form a Schiff's base. This is reduced to form the final product, with PBA present on the surface of the POROS20 beads.

recovery of O-GlcNAc standards. In batch, though, we attained a 44% average recovery of standards with only 12% nonspecific binding. Due to the extremely low background, we decided to move forward in testing these beads with AspN digested Lamin A. Again,

only 4% of standards were recovered, so a C18 clean up step was added prior to enrichment to remove any salts present in the digest. After cleaning up, we obtained a 69% average recovery of O-GlcNAcylated standards. More importantly, we detected all of the O-GlcNAcylated sites previously sited in Wriston's work and identified three new O-GlcNAc sites. Interestingly, the recovery tracked with the number of O-GlcNAc sites on a peptide. For instance, on peptide DKASASGSGAQVGGPISSGSSAS-SVTVTRSYRSVGGSGGGSFG the recoveries were as follows: unmodified (13%), mono- (15%), di- (23%), tri- (38%), quad- (60%), and penta- (79%). A hexa-O-GlcNAcylated was also detected that was not seen in previous experiments. Taken together, this suggests that the O-GlcNAc enrichment was successful.

The one downside of this experiment was nonspecific binding. To test and decrease non-specific binding, we performed a trypsin digestion on bovine serum albumin (BSA), a non-O-GlcNAcylated protein. We added O-GlcNAcylated peptides, performed a C18 clean up, and used the resulting elution for a number of tests. First, we tried esterifying peptides prior to enrichment. However, this procedure once again diminished all peptide binding to beads. We then tested the elution on underivatized ("bare" POROS 20) and derivatized (APBA POROS 20) beads. Here, we discovered that the O-GlcNAcylated peptide binding decreased by 50%, whereas nonspecific binding remained constant. This suggests a necessity for protocols that increase the amount of APBA present and/or that derivatize unreacted aldehydes on the surface of POROS20 beads without affecting O-GlcNAcylated peptide binding.

3.5 Discussion

The goal of this chapter was to develop a robust, sensitive, and selective O-GlcNAc enrichment method that achieves quantitative yields of internal standards. We have successfully established a procedure that achieves all of these requirements. The phenylboronic acid enrichment allows for up to 70% recovery of synthetic O-GlcNAcylated peptides when starting with only femtomoles of material. Additionally, it permits the enrichment of digested O-GlcNAcylated proteins at the sub-femtomole level. Previously, O-GlcNAcylated peptide and protein enrichment has been attempted in several ways, including: lectin-affinity chromatography, chemical derivatization, and chemoenzymatic labeling.[21-27] However, to our knowledge, none of these techniques produce quantitative yields of internal standards. Additionally, the number of steps involved and high level of contaminants severely decrease the likelihood of allowing for enrichment on the sub-femtomole level. The phenylboronic acid procedure overcomes these issues and has the potential to revolutionize our ability to identify and site-localize carbohydrate modifications.

In order to make the boronic acid procedure work for our purposes, we challenged a number of commonly accepted aspects of boronic acid solid phase extraction. As described in the previous section, the published mechanism for PBA enrichment requires boronate to be hydroxylated under basic conditions, which yields a tetrahedral boronate anion that allows the formation of esters with 1,2 cis-diols. The bond is reversed under acidic conditions through hydrolysis of the covalent bond.[28-34] However, under every condition we tested with this technique, we achieved a 0-3% recovery of our internal standards. This prompted us to reconsider the published mechanism, since it has been

perpetuated mainly by laboratories that utilize grams of starting material and are unconcerned with recovery of internal standards.

In the original 1959 paper by Lorand and Edwards, a pH depression experiment is performed in order to determine K_{eq} of ester formation with various diols. The experimental design was based on the fact that upon formation of the boronate ester, pK_a decreases compared with the boronic acid itself, resulting in a decrease of pH. The magnitude of pH depression under different conditions is directly proportional to the binding constants.[36] Despite being sited nearly 660 times, the study has several limitations that lead us to question the validity of their findings. First, the pH changes in these experiments are dependent upon the amount of diol added; in this case the pH varied by 2-3 units.[37] Springsteen and Wang later showed that changes in the pH by 2 units could change the binding constants by several-fold.[38] Thus, binding constants determined in the Lorand experiment can only be used as a general estimate of affinity.

Secondly, the pH depression experiment assumes the boronate anion is the only reactive species that can form the boronate ester. Further, it is assumed *none* of the trigonal boronic ester exists in solution.[37,38] Migamoto *et al* directly contradicted this finding by measuring rate constants with a stopped-flow spectrophotometer under pseudo-first order conditions. They found that the trigonal species is at least several thousand times more reactive than the boronate anion. They also suggest that boronic acid is always a reactive species, regardless of solution pH.[39] This higher reaction rate of boronic acid structure is supported by the proposed mechanism of diol-boronate complexation, which would require an intermediate pentavalent boron structure. However, only a few pentavalent boronic acid structures are known, rendering this

intermediate structure extremely unlikely.[40] Thus, the boronic acid is likely the reactive

species based on kinetic and steric observations.

Finally, and perhaps most importantly, boronic esters are extremely susceptible to

hydrolysis; in some cases even by ambient moisture. The overall process is an

equilibrium, and the forward reaction is favored when the boronate product is insoluble in

the reaction solvent. Otherwise, ester formation can be driven by distillation of water or

dehydrating conditions.[40] This observation is corroborated by the use of boronic acids to

protect diol units, which was first demonstrated in carbohydrate chemistry decades ago.

In every example we investigated, this protection reaction is done under anhydrous

conditions (i.e. benzene, dimethylchloride, dimethylformamide).[41-43]

Taken together, these factors caused us to propose a new mechanism and

procedure; this is shown in Figure 3.12. Here, dimethylformamide (DMF) acts as a

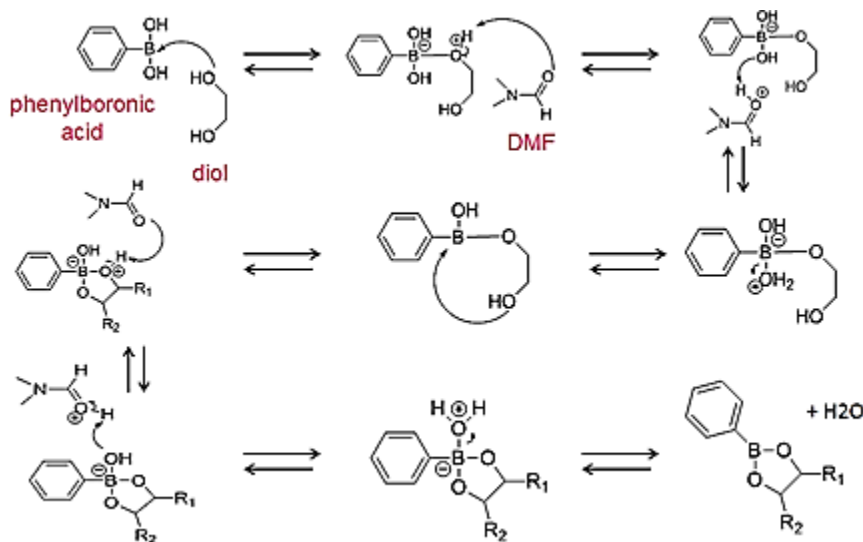dipolar aprotic solvent to mediate the reaction and push the equilibrium toward covalent

Figure 3.12 Proposed mechanism for boronic acid binding diols under anhydrous conditions. DMF acts as a dipolar aprotic solvent to mediate the reaction and push the equilibrium toward covalent bond formation. All reactions are reversible in water, highlighting the importance of completely anhydrous solvents.

bond formation. Once the bond is formed, the boronic esters may be washed with an equally anhydrous solvent (acetonitrile) to remove nonspecific binding. Finally, the bond may be hydrolyzed by the simple addition of water. Changing the protocol allowed us to dramatically increase recoveries, supporting our theory that sugars must be bound in anhydrous conditions.

Another aspect of this commonly held notion is that boronate only binds to vicinal 1,2 cis-diols. O-GlcNAc has an equatorial hydroxyl on the 4' carbon, whereas GalNAz has an axial hydroxyl at this positon.[28-34] We thus believed that only the O-GlcNAc tagged with GalNAz would bind to the PBA beads. However, after several trials using the new protocol, we found that O-GlcNAcylated peptides with and without GalNAz were recovered at nearly the same levels. Although, when taking into account an 80% recovery from both the GalT1 reaction and the C18 clean-up, the O-GlcNAcylated peptides have a higher overall recovery. This directly opposes the 1,2 cis-diol notion, so we investigated the literature to see how different the K_{eq} is between glucose and galactose. According to Springsteen and Wang, the association constant (K_{eq}, M^{-1}) of phenylboronic acid with galactose is 15, whereas glucose is 4.6.[38] Since they are separated by less than an order of magnitude, we hypothesize that by performing this reaction under anhydrous conditions we are simply pushing the equilibrium toward PBA-O-GlcNAc ester formation. As a corollary, it may be advantageous to add a drying agent to DMF to assure complete binding. We will attempt this in upcoming experiments.

Future directions will be devoted to optimizing this enrichment process even further. Specific efforts will be focused on attempting (a) on-column reactions, (b)

different magnetic beads, and (c) decreasing non-specific binding on POROS20 beads. We would like to move this reaction on-column so that we can directly connect the PBA column with a C18 column, minimizing sample loss. As discussed in the previous section, we have tried on-column reactions with three types of beads. The first type (Thermo polyacrylamide) was unsuccessful since we were still using the pH elution protocol, and was never revisited. The second type (Millipore ProSep) was unsuccessful for similar reasons, but also because the bead diameter is too large to pack in 360 o.d. x 150 i.d. fused silica columns. The third type (APBA-derivatized POROS20 AL beads) was packed in a 360 μm o.d. x 75 μm i.d. capillary to 5 cm. Using the anhydrous binding/water elution protocol, we attained a 22% recovery of internal standards (compared to 44% in-batch). We believe that this can be attributed to swelling of the beads during the acetonitrile wash. As the beads swelled, the flow rate decreased from ~20 μL/min to ~1 μL/min and was difficult to measure accurately. Because of this, the elution may not have been complete even after 30 minutes. Thus, we seek to find another anhydrous solvent that will not swell the beads but that is also compatible with connection to a C18 column. After identifying this solvent, we can continue to optimize the on-column procedure (amount of beads, flow rates, times) and elute directly from PBA to C18, ultimately minimizing sample loss and increasing likelihood of identifying low-level O-GlcNAcylated peptides.

Should the on-column optimization not work as planned, we would like to find a magnetic bead that we may use for real samples. As described in the previous section, even after C18 clean up the NHS-activated APBA derivatized magnetic beads only recovered 9% of digested Lamin A O-GlcNAcylated peptides. However, magnetic beads

are advantageous because they do not require centrifugation, which saves a significant

amount of time in the enrichment process. Further, the magnetic beads are not

accompanied by a high level of background (Figure 3.10), which increases the likelihood

of identifying low level O-GlcNAcylated peptides. Thus, we sought to collaborate with

Li *et al*, who published about magnetic beads coated with an extremely high density of

boronic acid sites. Their production method is shown in Figure 3.13. These beads are

advertised as having higher binding capacity and selectivity than other commercially

available PBA beads.[44] On first pass, we obtained a 99% recovery of standards

accompanied by a 50% recovery of non-specific binding. We then proceeded to test the
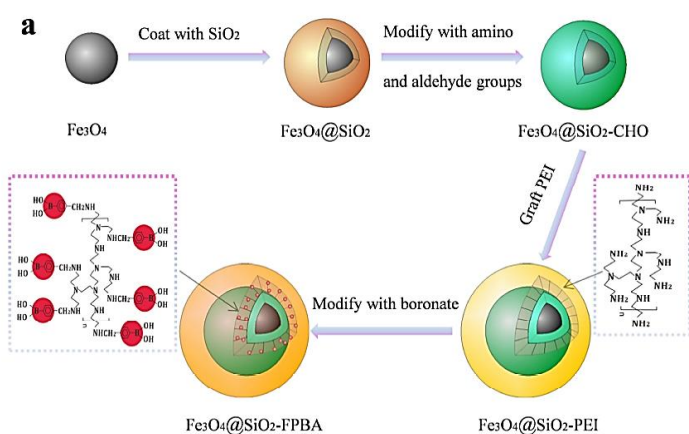
Figure 3.13 Li *et al* magnetic bead derivatization process. Iron beads are first coated with silica, then modified with amino and aldehyde groups. The aldehydes are then coated with PEI and modified with boronic acid. The boronic acids are free to interact with O-GlcNAc moieties on peptides.

beads with a BSA digest and found that nearly everything, O-GlcNAcylated and

unmodified peptides, bound with a 70-80% recovery. Due to the beads' initial success

with standard peptides, we plan to collaborate with this laboratory to ensure complete

boronic acid derivatization and to decrease non-specific binding significantly.

Alternatively, we may proceed with the APBA-derivatized POROS20 AL beads

while continuing to optimize the derivatization and batch enrichment process. These

beads were successful in enriching for Lamin A O-GlcNAcylated peptides with the sole

downside that a significant amount of nonspecific binding was apparent. Thus, we seek to decrease nonspecific binding by one of two ways. First, since we did not identify any particular bias for non-specific binding, we may attempt a large-scale study to find whether basic, acidic, hydrophilic, or hydrophobic peptides have a preferential affinity for the PBA beads. Depending on which (if any) emerge as a favorite, this can direct our experiments to exclude certain residues or peptides. Second, we wish to increase the amount of PBA present and/or that derivatize unreacted aldehydes on the surface of POROS20 beads without affecting O-GlcNAcylated peptide binding. To do this, we will modify the derivatization process by trying different pH 6-7 buffers, binding times, reducing agents, and temperatures. We may also perform the derivatization protocol twice to assure the highest level of conversion possible. Finally, we wish to cap any remaining free aldehydes by derivatizing them in another manner. We attempted EDC-mediated ethanolamine derivatization, but this only served to decrease O-GlcNAcylated peptide binding while keeping unmodified peptide binding constant. Thus, we must look to other derivatization methods for this purpose. For example, we can react the support with another low molecular weight primary amine such as 0.2M Tris buffer with the addition of reducing agent. Ultimately, we desire a solid support that will maximize O-GlcNAcylated peptide binding while minimizing unmodified peptide adherence.

3.6 References

1. Varki A, Cummings RD, Esko JD, *et al.,* editors. Essentials of Glycobiology. 2nd edition. Cold Spring Harbor (NY): Cold Spring harbor Laboratory Press; 2009.

2. Wang Z, Udeshi ND, Slawson C, Compton PB, Sakabe K, Cheung WD, Shabanowitz J, Hunt DF, and GW Hart. Extensive crosstalk between O-GlcNAcylation and phosphorylation regulates cytokinesis. *Biochem*. 2010; 3(104): 1-13.

3. Hart GW, Slawson C, Ramirez-Correa G and O Lagerlof. Cross talk between O-GlcNAcylation and phosphorylation: roles in signaling, transcription, and chronic disease. *Annu. Rev. Biochem*. 2011; 80: 825-858.

4. Ruan H, Singh JP, Li M, Wu J and X Yang. Cracking the O-GlcNAc code in metabolism. *Trends in Endocrinology & Metabolism.*2013; 24: 301-309.

5. Wells L, Vosseller K, and GW Hart. Glycosylation of nucleocytoplasmic proteins: signal transduction and O-GlcNAc. *Science.* 2001; 446: 2376-2378.

6. Hart GW, Housley MP, and C Slawson. Cycling of O-linked beta N-actetylglucosamine on nucleocytoplasmic proteins. *Nature.* 2007; 446: 1017-1022.

7. Love DC and JA Hanover. The hexosamine signaling pathway: deciphering the "O-GlcNAc code." *Sci STKE*. 2005; re13.

8. Hanover JA. Glycan-dependent signaling: O-linked N-acetylglucosamine. *FASEB j.* 2001; 15:1865-1876.

9. Wang Z, Gucek M, and GW Hart. Crosstalk between GlcNAcylation and phosphorylation: site-specific phosphorylation dynamics in response to globalled elevated O-GlcNAc. *Proc Nat Acad Sci.* 2008; 105: 13793-13798.

10. Copeland RJ, Han G, and GW Hart. Revealing Roles of O-GlcNAcylation in disease mechanisms and development of potential diagnostics. *Proteomics Clin Appl.* 2013; epub.

11. Slawson C and GW Hart. O-GlcNAc signaling: implications for cancer cell biology. *Nat Rev.* 2011; 11: 678-684.

12. Whelan SA, Lane MD, and GW Hart. Regulation of the O-linked b-N-acetylglucoasmine transferase by insulin signaling. *J Biol Chem*. 2008; 283: 21411-21417.

13. Slawson C, Copeland RJ, and GW Hart. O-GlcNAc signaling: a metabolic link between diabetes and cancer? *Trends Biochem Sci*. 2010; 35: 547-555.

14. McClain DA. Hexosamines as mediators of nutrient sensing and regulation in diabetes. *J Diabetes Complicat.* 2002; 16: 72-80.

15. Wells L, Vosseller K, and GW Hart. A role for N-acetylglucosamine as a nutrient sensor and mediator of insulin resistance. *Cell Mol Life Sci.* 2003; 60: 222-228.

16. Arnold CS, Johnson GV, Cole RN, Dong DL, Lee M, and GW Hart. The microtubule-associated protein tau is extensively modified with O-linked N-acetylglucosamine. *J Biol Chem.* 1996; 271: 28741-28744.

17. Liu F, Iqbal K, Grundke-Iqbal I, Hart GW, and CX Gong. O-GlcNAcylation regulates phosphorylation of tau: a mechanism involved in Alzheimer's disease. *Proc Natl Acad Sci.* 2004; 101: 10804-10809.

18. Deng Y, Li B, Liu Y, Grundke-Iqbal I, and CX Gong. Dysregulation of insulin signaling, glucose transporters, O-GlcNAcylation, and phosphorylation of tau and neutrofilaments in the brain: implication for Alzheimer's disease. *Am J Pathol.* 2009; 175: 2089-2098.

19. Iribarne JV and BA Thomson. On the evaporation of small ions from charged droplets. *J Chem Phys.* 1976; 64: 2287-2294.

20. Syka JE, Coon JJ, Schroeder MJ, Shabanowitz J, Hunt DF. Peptide and protein sequence analysis by electron transfer dissociation mass spectrometry. *Proc Natl Acad Sci U S A*. 2004 Jun 29; 101(26): 9528-9533.

21. Mortezai N, Wagener C, and F Buck. Combining lectin-affinity chromatography and immunodepletion – a novel method for the enrichment of disease-specific glycoproteins in human plasma. *Methods.* 2012; 56: 254-259.

22. Wang Z, Udeshi ND, O'Malley M, Shabanowitz J, Hunt DF, and GW Hart. Enrichment and site mapping of O-linked N-acetylglucosamine by a combination of chemical/enzymatic tagging, photochemical cleavage, and electron transfer dissociation mass spectrometry. *Mol Cell Proteome.* 2010; 9: 153-160.

23. Comer FI, Vosseller K, Wells L, Accavitti MA, and GW Hart. Characterization of a mouse monoclonal antibody specific for O-linked N-acetylglucosamine. *Anal Biochem.* 2001; 293: 923-934.

24. Darula Z and K Medzihradszky. Affinity enrichment and characterization of mucin core-1 type glycopeptides from bovine serum. *Mol Cell Proteome.* 2009; 8(11): 2515- 2526.

25. Chalkley RJ, Thalhammer A, Schoepfer R, and AL Burlingame. Identification of protein O-GlcNAcylation sites using electron transfer dissociation mass spectrometry on native peptides. *Proc Natl Acad Sci*. 2009; 106(22). 8894-8899.

26. Parker BL, Gupta P, Cordwell SJ, Larsen MR, Palmisano GP. Purification and Identification of O-GlcNAc-Modified Peptides Using Phosphate-Based Alkyne CLICK Chemistry in Combination with Titanium Dioide Chromatography and Mass Spectrometry. *J Proteome Res.* 2011; 10: 1449-1458.

27. Banerjee PS, Hart GW, and J Won Cho. Chemical approaches to study O-GlcNAcylation. *Chem Soc Rev.* 2013; 42: 4345-4357.

28. Liu XC and WH Scouten. Boronate affinity chromatography. *Methods in Molecular Biology, vol 147: Affinity chromatography: methods and protocols.* Humana Press Inc Totowa, NJ.

29. Zhang Q, Schepmoes AA, Brock JWC, Wu S, Moore RJ, Purvine SO, Baynes JW, Smith RD, and TO Metz. Improved methods for the enrichment and analysis of glycated peptides. *Anal Chem.* 2008; 80(24): 9822-9829.

30. Wimmer MA, Lochnit G, Bassil E, Muhling KH, and HE Goldbach. Membrane-associated boron-interacting proteins isolated by boronate affinity chromatography. *Plant Cell Physiol.* 2009; 50(7): 1292-1304.

31. Sparbier K, Wenzel T, and M Kostrzewa. Exploring the binding profiles of ConA, boronic acid and WGA by MALDI-TOF/TOF MS and magnetic particles. *J Chrom B.* 2006; 840: 29-36.

32. Li YC, Jeppsson JO, Jornten-Karlsson M, Larsson EL, Jungvid H, Galev IY, and B Mattiasson. Application of shielding boronate affinity chromatography in the study of the glycation pattern of haemoglobin. *J Chrom B.* 2002; 776: 149-160.

33. Zhang Q, Tang N, Brock JW, Mottaz HM, Ames JM, Baynes JW, Smith RD, and TO Metz. Enrichment and analysis of nonenzymatically glycated peptides: boronate affinity chromatography coupled with electron-transfer dissociation mass spectrometry. *J Proteome Res*. 2007; 6: 2323-2330.

34. Zhang Q, Tang N, Schepmoes AA, Phillips LS, Smith RD, and TO Metz. Proteomic profiling of nonenzymatically glycated proteins in human plasma and erythrocyte membranes. *J Proteome Res.* 2008; 7: 2025-2032.

35. Wang W, Hong S, Tran A, Jiang H, Triano R, Liu Y, Chen X, and Peng Wu. Sulfated ligands for the copper(I)-catalyzed azide-alkyne cycloaddition. *J Chem Asian.* 2011; 6: 2796-2802.

36. Lorand JP, and JO Edwards. Polyol complexes and structure of the benzeneboronate ion. *J Org Chem.* 1959: 24; 769-774.

37. Yan J, Springsteen G, Deeter S, and B Wang. The relationship among pKa, pH and binding constants in the interactions between boronic acids and diols – it is not as simple as it appears. *Tetrahedron.* 2004; 60: 11205-11209.

38. Springsteen G and B Wang. A detailed examination of boronic acid-diol complexation. *Tetrahedron.* 2002; 58: 5291-5300.

39. Miyamoto C, Suzuki K, Iwatsuki S, Inamo M, Takagi HD, and K Ishihara. Kinetic evidence for high reactivity of 3-nitrophenylboronic acid compared to its conjugate boronate ion in reactions with ethylene and propylene glycols. *Inorg Chem.* 2008; 47: 1417-1419.

40. Hall DG. Chapter 1: Structure, properties, and preparation of boronic acid derivatives. Overview of their reactions and applications. From *Boronic Acids*. Edited by DG Hall. 2005 Wiley-VCH. KGaA, Weinheim.

41. Brown HC, Phadke AS and MV Rangaishenvi. Ring enlargement of boracyclanes via sequential one carbon omologation. The first synthesis of bracyclanes in the strained medium ring range. *J Am Chem Soc.* 1988; 110(18): 6263-6264.

42. Matteson DS and GY Kim. Asymmetric alkyldifluroroboranes and their use in secondary amine synthesis. *Org Lett.* 2002; 4(13): 2153-2155.

43. Inglis SR, Woon ECY, Thompson AL, and CJ Schofield. Observations on the deprotection of pinanediol and pinacol boronate esters via fluorinated intermediates. *J Org Chem.* 2010; 75: 468-471.

44. Li H, Shan Y, Qiao L, Dou A, Shi X, and G Xu. Facile synthesis of boronate decorated polyethyleneimine grafted hybrid magnetic nanoparticles for the highly selective enrichment of modified nucleosides and ribosylated metabolites. *Anal Chem.* 2013; 85(23): 11585-11592.

Chapter 4: Enrichment and Identification of HLA-associated O-GlcNAcylated Peptides from Various Cancers

4.1 Introduction

4.1.1 O-GlcNAcylation and Cancer

As described in Chapter 1, the multistep development of human tumors involves several hallmarks of cancer including: reprogramming of energy metabolism, angiogenesis, sustained proliferative cell signaling, and evasion of apoptosis. All of these contribute to the ability of transformed cells to continue rapid growth and proliferation.[1,2] Changes in protein expression and/or metabolism that accompany cellular transformation can be conveyed to the immune system via the MHC class I processing pathway[3] This system involves intracellular proteins that are degraded to 8-10 amino acid peptides. These peptides are then loaded onto MHC-class I molecules and presented onto the cell surface. If the peptide is antigenic, circulating CD8+ T lymphocytes can bind to the MHC-peptide complex and stimulate an immune response.[4-6] Since dysregulated cell signaling in malignant cells leads to aberrant enzymatic modification of proteins, we believe transformed cells will present post-translationally modified peptides unique to cancer. We aim to identify tumor-specific peptides on the cell surface that can stimulate immune action in order to develop specific cancer immunotherapeutics.

In Chapter 2, we demonstrated that phosphorylation is preserved on peptides during antigen presentation by the MHC class I processing pathway. Additionally, we have established the ability of tumor-specific phosphopeptides to stimulate a memory-like immune response in healthy individuals. Another modification, O-linked β-*N*-acetylglucosamine (O-GlcNAc), is involved in cross-talk with phosphorylation in several

signaling pathways that are known to be dysregulated in cancer.[7] For example, c-Myc is O-GlcNAcylated at Thr58 in non-dividing cells. However, in several human lymphomas, kinase GSK3β is mutated causing Thr58 to be phosphorylated extensively, stimulating cellular growth.[8] Another instance of this O-GlcNAc/phosphate interplay occurs within the IκB-NF-κB pathway, which plays an important role in cancer progression. O-GlcNAcylation of inhibitor nuclear factor kB kinase at Ser733 (normally phosphorylated) enhances its activity, especially in response to high glucose levels.[9] Thus, O-GlcNAcylation is important for regulating signaling and transcriptional processes relevant to cancer.

Further, tumor cells have altered glucose metabolism, as cancer cells preferentially utilize glycolysis to produce ATP in order to satisfy their increased energetic and biosynthetic requirements.[1,2] Although aerobic glycolysis is an inefficient process compared to oxidative phosphorylation, in a nutrient rich environment, efficient energy production is less important than biosynthesis. And since cancerous cells are programmed to rapidly proliferate, they must consume carbon and nitrogen-rich nutrients to biosynthesize metabolites needed for cell division. This metabolic shift, termed the "Warburg effect," has been confirmed in human cancers *in vivo,* as these tumors require up to 10x higher glycolytic flux compared to adjacent normal tissue.[10-12] Interestingly, the rate of glucose metabolism directly correlates with tumor aggressiveness.[13]

These changes in glucose uptake alter downstream nutrient signaling pathways. The hexosamine biosynthetic pathway (HBP), whose endpoint is UDP-GlcNAc, is particularly affected since it shares its first two steps with glycolysis. O-GlcNAc transferase (OGT) activity responds to physiological changes in UDP-GlcNAc, which

leads to elevated O-GlcNAc modifications on proteins.[7] It follows, then, that OGT and O-GlcNAc levels should be upregulated in cancers. This is the case in numerous cancers, including: breast,[14,15] chronic lymphocytic leukemia,[16] liver,[17] esophageal,[18] prostate,[19] lung, and colorectal,[20] among others. Figure 4.1 summarizes the modification and its enzymatic dynamics in different tumors.[21]
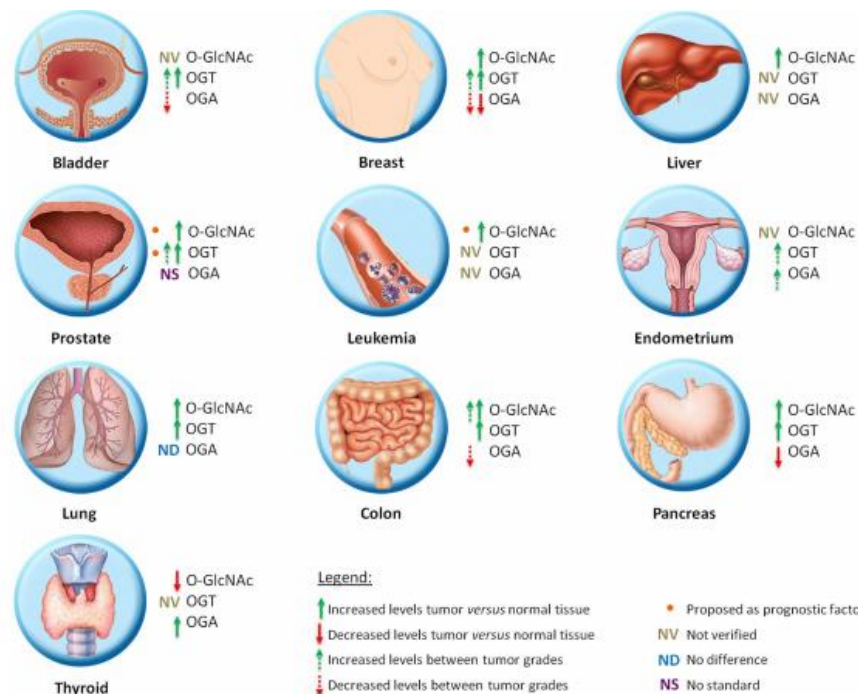
Figure 4.1 Dynamics of O-GlcNAc, OGT, and OGA in various cancers. An increase in O-GlcNAc and/or OGT has been reported in bladder, breast, prostate, chronic lymphocytic leukemia, liver, esophageal, prostate, lung, and colorectal cancers. Interestingly, only thyroid cancer saw the converse effect, but the implications of this were not studied.[21]

O-GlcNAcylation is implicated in other aspects of cancer development and progression, including: evasion of apoptosis, metastasis, and promoting aneuploidy. The reduction of O-GlcNAcylation in prostate tumor cells decreased the expression of an anti-apoptotic protein (Bcl-XL) and induced the pro-apoptotic cleavage of caspases 3 and 9.[19] Since reducing O-GlcNAcylation activates intrinsic apoptotic pathways, O-GlcNAc may

have a role in protecting tumor cells from apoptosis. Further, O-GlcNAcylation affects the trafficking of cell adhesion molecules important to metastasis. For example, O-GlcNAcylation of β-catenin not only regulates its nuclear localization and transcriptional activity, but also blocks its association with E-cadherin, a molecule critical to epithelial cell adhesion. Loss of E-cadherin transport to the cell surface of epithelial cells is important to mechanisms underlying metastasis. This was demonstrated in breast cancer cell lines, where OGA inhibition markedly increased cell migration and invasion.[14,15] Finally, overexpression of OGT by only two-fold promotes aneuploidy through disruptions in mitotic progression. This is likely because OGT performs multiple important functions at the mitotic spindle, including O-GlcNAcylation of spindles and midbodies, along with promotion of protein-protein interactions. The disruption of these function shows that O-GlcNAc is a key regulator of multiple pathways involved with proper function of M phase, and by association, aneuploidy.[22]

Taken together, O-GlcNAcylation plays an integral role in tumor formation, growth, and survival. Similar to phosphorylation, O-GlcNAcylation is implicated in the dysregulated cell signaling underlying many hallmarks of cancer. O-GlcNAcylation is also related to cancer due to reprogramming of energy metabolism and its roles in evasion of apoptosis and metastasis. Thus, we hypothesize that tumor-specific MHC-associated O-GlcNAcylated peptides exist and can be identified utilizing high-resolution mass spectrometry coupled with novel enrichment techniques. We further hypothesize that these peptides will stimulate immune action and can be used to develop specific cancer immunotherapeutics.

4.1.2 MHC class I-associated O-GlcNAcylated peptides

Approximately twenty years ago, MHC class I-associated O-GlcNAcylated peptides were investigated extensively by the Haurum lab and its collaborators. They began by showing that mice MHC class I molecules (allele HLA-2D[b]) could bind synthetic O-GlcNAcylated peptides and that the glycosylation does not affect peptide-MHC binding. In the same study, they showed that this peptide complex could elicit a strong cytotoxic T-cell (CTL) response.[23] They went on to find that T-cell recognition is glycopeptide-specific, sensitive to the fine structure of the sugar, and that the sugar may be recognized directly.[24] Then, they solved the crystal structure of the MHC-class I-glycopeptide complex, demonstrated in Figure 4.2. In this structure, the glycan is solvent exposed and available for direct recognition by the T-cell receptor (TCR). In the same publication, they also modeled the interaction between the TCR and the MHC-peptide complex, showing that at least one saccharide residue can be accommodated in the standard TCR-MHC geometry (also seen in Figure 4.2).[25]

The recognition of O-GlcNAc modified peptides by T-cells was recognized by these laboratories as being highly significant, since glycosylation can create a T-cell neo-epitope without requiring a change at the genetic level. They also acknowledged that alterations in intracellular metabolism lead to changes in PTMs, including O-GlcNAcylation, and these changes are related to autoimmunity and malignancy. But at the time of their investigations, it was not known whether post-translationally modified peptides could be presented by MHC class I molecules *in vivo*. Thus, they went on to acquire biological data that suggested this was the case for O-GlcNAcylated peptides. They first demonstrated that O-GlcNAc substituted peptides can be efficiently
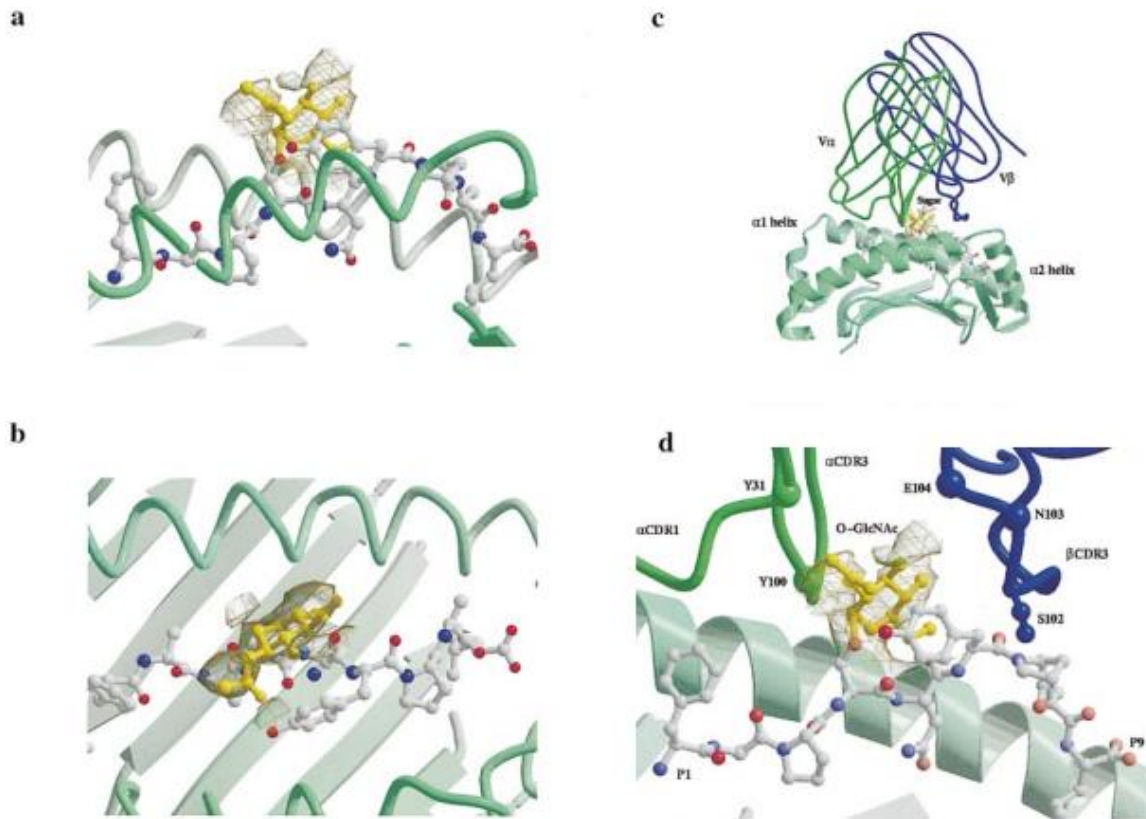
Figure 4.2 Crystal structure of H-2D^b complexed with an O-GlcNAcylated peptide K3G and modeled with TCR. K3G sequence is FAPgSNYPAL, used in A-D; HLA-2D^b is a mouse MHC class I allele. (A) The main chain of H-2D^b is highlighted in green, the peptide is in ball-and-stick form, and the O-GlcNAc is highlighted in yellow. Electron density of the glycan is shown as chicken wire. (B) Same as A, but viewed down onto the binding groove. (C) A model for glycopeptide-specific CTL clone is shown positioned over the peptide. TCR a chain is green, TCR b is blue, and H-2Db is green. Peptide is in ball-and-stick form with glycan highlighted in yellow. (D) a zoomed in view of the interaction in C.

translocated across the endoplasmic reticulum (ER) membrane by the transporter associated with antigen processing (TAP).[26] Further, by labeling O-GlcNAc via enzymatic transfer of [^{3}H]Gal onto terminal O-GlcNAc residues and measuring radioactivity, they estimated that 0.1-1% of all human MHC class I-associated peptides were O-GlcNAcylated. They also treated these samples with peptide *N*-glycosidase F (PNGaseF) which cleaves N-linked carbohydrate structures, and demonstrated that the labeled peptides were not sensitive to this treatment. This indicated the MHC-derived peptides contain O-β-linked GlcNAc residues.[27] Finally, they isolated peptides presented

by human HLA-A2, purified the peptides via lectin affinity chromatography, and subjected the eluate to Edman degradation and pool sequencing. They once again labeled these peptides and measured the approximate abundance of O-GlcNAcylated peptides.[28]

However, due to instrument sensitivity issues at the time of publication, this group was unable to sequence any O-GlcNAcylated peptides naturally presented by the MHC class I processing pathway. Thus, no class I MHC reactivity has ever been identified toward a defined, naturally processed glycopeptide. This can be explained by the sub-stoichiometric levels of the modification coupled with the fact that the unmodified peptide often co-elutes with the modified peptide, which suppresses ionization of the latter.[7] Additionally, traditional fragmentation techniques such as collision activation dissociation (CAD) causes fragmentation of the labile O-GlcNAc bond, making site localization near impossible.[30] In the Hunt laboratory, we employ cutting-edge fragmentation (electron transfer dissociation, ETD) and enrichment techniques (phenylboronic acid enrichment, PBA) that allow for the identification and sequencing of low-level post-translationally modified proteins and peptides.

4.1.3 Project Aim

In this chapter, we seek to identify the first MHC class I-associated O-GlcNAcylated peptides from the surface of tumor cells. To do this, we will isolate O-GlcNAcylated peptides using the PBA enrichment procedure described in the previous chapter. We will utilize high-resolution mass spectrometry in order to characterize and sequence the peptides using a combination of *de novo* sequencing and bioinformatics. Finally, we will test the tumor-specific O-GlcNAcylated peptides in order to assess their immunogenicity.

4.2 Materials

4.2.1 Reagents

Amersham Pharmacia Biotech (Amersham, UK)

    NHS-activated sepharose magnetic beads

Applied Biosystems^{TM} (Carlsbad, CA)

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

    POROS® AL 20 perfusion chromatography bulk media, 20 μm diameter

Becton Dickinson (Oxford, UK)

    APC-Cy7 fixable visibility dye

    APC anti-CD3 dye

    Pacific Blue anti-IL2 dye

    PE anti-IFNγ dye

    PE Cy-5.5 anti-TNFα dye

    PerCP anti-CD8 dye

Grace Davison Discovery Science (Deerfield, IL)

    Acetyl chloride, anhydrous

    D_0 methanol, anhydrous

Honeywell (Morristown, NJ)

    Acetonitrile, HPLC grade, ≥99.8% purity

Invitrogen (Carlsbad, CA)

    AIM-V media

J.T. Baker (Phillipsburg, NJ)

Glacial acetic acid, ≥99.9% purity

Mabtech (Sweden)

ELISpot PRO kit (human IFNγ)

Pierce (Rockford, IL)

LC-MS grade water

PQ Corporation (Valley Forge, PA)

Kasil ® 1624 potassium silicate solution

Sigma Aldrich (Saint Louis, MO)

1-ethyl-3-(3-dimethylaminopropyl)carbodiimide

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

Ammonium hydroxide

Angiotensin I acetate salt hydrate, ≥90% purity

Angiotensin II phosphate

Aprotinin

Ascorbic acid

Azulene

CHAPS buffer

Dimethyl sulfoxide

Dulbecco's Modified Eagle Medium

Ethanolamine

Ethylenediaminetetraacetic acid (EDTA), analytical grade

Iron (III) chloride

HEPES buffer

Phosphate buffered saline

Pyridine

Pepstatin A

Phenylmethylsulfonyl fluoride (PMSF)

Sodium chloride

Sodium cyanoborohydride

Tris-HCl

Vasoactive intestinal peptide fragment 1-12, ≥97% purity

Thermo-Fisher Scientific (Waltham, MA)

Aminophenyl boronic acid

Dimethylformamide

Hydrochloric acid

Immobilized boronic acid resin

YMC Company, LTD (Kyoto, Japan)

ODS-AQ, C18 5 μm spherical silica particles, 120 A pore size

ODS-AQ, C18 5-20 μm spherical silica particles, 120 A pore size

4.2.2 Equipment and Instrumentation

Advanced Imaging Devices (Strasberg, Germany)

AID ELISpot Reading HR XL

Agilent Technologies (Palo Alto, CA)

1100 Agilent high performance liquid chromatograph

Branson (Danbury, CT)

Branson 1200 Ultrasonic Bath

Becton Dickinson (Oxford, UK)

FACSCantoII flow cytometer

Beckman Coulter (Pasadena, CA)

Optima LE-8K ultracentrifuge (Ti70 rotor)

Labconco Corp. (Kansas City, MO)

Centrivap centrifugal vacuum concentrator

Millipore (Billerica, MA)

Amicon ultra, 5 kDa regenerated cellulose spin filter

PolyMicro Technologies, Inc. (Phoenix, AZ)

360 μm o.d. x 50 μm i.d. polyimide coated fused silica capillary

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

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

Sutter Instrument Co. (Novato, CA)

P-2000 microcapillary laser puller with fused silica adapter

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

LTQ mass spectrometer (back end ETD)

LTQ FTICR hybrid mass spectrometer (custom modified with front end ETD)

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

LTQ-Orbitrap Velos 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 Cell and tissue culture (Performed by SAP, MC, KC, AP)

Cell lines were grown up from a laboratory stock and were stored in FBS, DMEM/F12K, Medium 199, and/or MCDB105 media with 10% dimethylsulfoxide (DMSO) at -80ºC, were managed in sterile conditions in a laminar airflow hood, and incubated at 37ºC with 5% CO_2. The media used was Dulbecco's Modified Eagle Medium (D-MEM), 10% FBS, 1% penicillin/streptomycin and grown in vented tissue culture flasks with 20 mM HEPES buffer. Cell counts were determined by exclusion of Trypan blue dye under a light microscope. Tissue samples (normal/tumor) were collected one-hour post-removal. The entire tumor sample was prepared for HLA-associated phosphopeptide analysis. We assume that 1 gram is equivalent to 1e9 cell equivalents (CEq).

4.3.2 sHLA-A*0201 transfection and peptide isolation (Performed by AP)

SKOV3 and FHIOSE were transfected with the sHLA-A*0201 construct by electroporation using a BIORAD GenePulser Xcell. The cell lines were grown in their respective media to 80-90% confluency, trypsinized and placed in electroporoation cuvettes (750 μL at 5e6/mL) along with 30 μg of plasmid DNA. Cells were pulsed with a voltage of 250 V and a capacitance at 940 μF. Electroporated cells were plated into 96 well plates with their respective complete growth media, and after 3 days successfully transfected cells were selected with G418.

sHLA-A*0201 molecules with immunoaffinity purified from SKOV3 and FHIOSE and their associated peptides were extracted. Approximately 16L of cell media (~30 mg sHLA) were passed over protein A sepharose pre-loaded with the anti-VLDLr antibody and washed with 20 mM sodium phosphate buffer, pH 7.2. Peptides were eluted from immunopurified sHLA molecules with 0.2N acetic acid and brought up in 10% acetic acid. The elution was boiled at 76°C for 10 minutes, cooled to RT, and passed through a 3 kDa MCWO filter. Filtered peptides were flash frozen, lyophilized, and reconstituted in 10% acetic acid. The second elution from this procedure containing approximately 1.5μg of total peptide was used for analysis.

4.3.3 HLA-associated peptide isolation (Performed by SAP, MC, AZ)

HLA class I molecules were immunoaffinity purified from cell lines and tissues, and their peptides were extracted as previously described.[3] Cells were lysed in 10 mL CHAPS buffer, comprised of: 20mM Tris-HCl pH 8.0, 150 mM NaCl, 1% 3-[(3-cholamindopropyl) dimethylammonio]-1-propane sulfonate (CHAPS). Additionally, protease inhibitors are added to prevent degradation of HLA molecules: 1 mM PMSF, 5 μg/mL aprotinin, and 10 μg/mL pepstatin A. Finally, phosphatase inhibitor cocktails II and III are added in 1:100 dilutions to prevent dephosphorylation of peptides during extraction. Lysates are subject to centrifugation at 100,000 xg for 1 hour at 4ºC in an ultracentrifuge. Supernatants were incubated with W6/32 antibody-bound NHS sepharose beads, which is specific for all classic MHC class I molecules. The beads are then subject to a series of washes with lysis buffer, TBS (20 mM Tris-HCl, 150 mM NaCl pH 8), TBS

2 (20 mM Tris-HCl, 1 M NaCl pH 8), and 20 mM Tris-HCl pH 8. Peptides were eluted from the HLA class I molecules with 10% acetic acid and separated using a 5 kDa MWCO Ultrafree MC filter. Extracted peptides were stored at -80ºC before being shipped to the University of Virginia for analysis.

4.3.4 Peptide content determination by HPLC ESI-MS/MS

HLA-associated peptide mixtures were dried to completion in a speedvac and reconstituted in 0.1% acetic acid to a final concentration of 1e7 cell equivalents per µL. 1 µL of sample (1e7 CEq) was spiked into 4 µL of 0.1% acetic acid. 100 fmol of two standard peptides (angio and vaso) were spiked into the 5 µL solution and this was pressure loaded onto a C18 microcapillary precolumn (PC) (360 µm o.d. x 75 µm i.d.) with a 2 mm Kasil frit packed with 6-8 cm of C18 RP packing material (5-20 µm diameter particles). The PC was rinsed for 10-15 minutes with solvent A (0.1 M acetic acid), dried, and connected with Teflon to a bottleneck-fritted analytical column (AC) (360 µm o.d. x 50 µm i.d) packed with C18 packing material (5 µm diameter particles) equipped with a laser-pulled electrospray emitter tip.

Peptides were eluted over a 40 minute gradient of 0-60% solvent B (70% acetonitrile, 0.1 M acetic acid) at 60 nL/min and electrospray ionized into the mass spectrometer. Intact peptide masses were acquired in the FTICR or the Orbitrap, and MS2 spectra (CAD and ETD) were acquired in the LTQ. Mass spectra were acquired using a method consisting of one high resolution scan (resolving power of 60,000 at 400 m/z) followed by 10 data dependent low resolution scans; the 5 most abundant parent

ions were subject to both CAD and ETD. Data dependence parameters were set at a

repeat count of 3, repeat duration of 10, and exclusion list length of 10. ETD parameters

included: 35-45 ms reaction time (instrument dependent), FTMS automatic gain control

(AGC) target of 2e5 charges, ITMS AGC target of 1e4, and ETD reagent target of 2e5.

Data analysis was performed using Xcalibur software. Raw data files were

searched against the RefSeq database using OMSSA. MS2 searches used the following

parameters: no enzyme specificity, e-value cutoff of 1, and variable modification of Met

(oxidation). Mass tolerances for intact and product ion masses were set at ± 0.1 Da and

± 0.35 Da, respectively. Additionally, MS2 data was searched using either c- and z-

fragments (ETD) or b-and y- fragments (CAD). All peptide hits were subject to manual

interpretation of MS1 (accurate mass) and MS2 (fragment ion) spectra.

Relative abundances of peptides were calculated by comparing the peptide peak

area to the average peak areas of two internal standard peptides (angio and vaso), which

we assume have a fixed concentration of 100 fmol.

4.3.5 C18 cleanup of samples

HLA-associated peptides in 0.1% acetic acid were combined with an equal

volume of 0.1% and various amounts of internal standards. Depending on the HLA allele,

2-5 of the following peptides were used: IPVsSHNSL, RPPItQSSL, VLTsNVQTI,

VLTsNVQTK, TVAtQASGLLSNK, APAVQHIVVtAAK, and APAtQAYGHQIPLR.

This solution was pressure loaded onto a fused silica clean-up column (360 μm o.d. x 150

μm i.d.) packed to 5 cm with C18 resin (5-20 μm diameter particles) at a flow rate of

~0.5 μL/min. The column loaded with sample was rinsed with 25 μL 0.1% acetic acid at

the same flow rate. Then the column was moved to an HPLC where it was rinsed with

solvent A for 10 minutes at 12-15 bar. Peptides were eluted into an Eppendorf tube using

a 40 minute gradient (0-80% solvent B) that was held at 80% B for 30 additional minutes.

In later experiments to exclude CHAPS buffer, peptides were collected in 3 fractions: 0-

23 minutes, 23-40 minutes, 40-70 minutes. In these instances, 1% of the elution was

screened to assure peptides were present but CHAPS was not. In either situation, the

elutions were taken to dryness and stored at -35ºC until further use.

4.3.6 Phenylboronic acid solid phase extraction

For the following analyses, all supernatants were removed using gel-loading tips

and extreme care was taken so as to not disturb the bead pellets. Further, all fractions

(flow-through, wash, and elution, where applicable) were taken to dryness in a vacuum

concentrator, brought up in 10 μL 0.1% acetic acid, and loaded onto a PC. The PC was

rinsed for 10-15 minutes with solvent A (0.1 M acetic acid), dried, and connected with

Teflon to a bottleneck-fritted analytical column (AC) (360 μm o.d. x 50 μm i.d) packed

with C18 packing material (5 μm diameter particles) equipped with a laser-pulled

electrospray emitter tip. Peptides were eluted over a 60 or 120 minute gradient of 0-60%

solvent B (70% acetonitrile, 0.1 M acetic acid). Intact peptide masses were acquired in

the FTICR or the Orbitrap, and MS2 spectra were acquired in the LTQ. For most

analyses, mass spectra were acquired using a method consisting of one high resolution

scan (resolving power of 60,000 at 400 m/z) followed by 5 data dependent low resolution

scans; the most abundant parent ions were subjected to both CAD and ETD. In later

experiments, a neutral loss method was developed whereby the top 10 ions are subjected

to CAD fragmentation. Then, if a neutral loss of 101 ([GlcNAc]$^{++}$), 110 ([GlcNAc+

H$_2$O]$^{++}$), or 67 ([GlcNAc]$^{+3}$) was detected from the parent ion (±1 Da), ETD is triggered

on that parent ion. Parameters for both CAD and ETD were the same, but the analysis

was done on the LTQ-Velos Orbitrap due to its upgraded software. Since the instrument

parameters are optimized for whole-protein work, the pressure within the HCD cell was

lowered so that Δp is approximately 0.8 Torr. In either case, relative abundances were

calculated by comparing peak areas of each O-GlcNAcylated peptide to internal

standards (angio and vaso) which were at a fixed concentration of 100 fmol each. Percent

recoveries were determined by comparing the relative abundance of O-GlcNAcylated

peptides to expected values.

   Data analysis was carried out using Thermo Xcalibur software. Raw data files

were searched against the RefSeq database using OMSSA. MS2 searches used the

following parameters: no enzyme specificity, e-value cutoff of 1, and variable

modifications of Met (oxidation), Ser and Thr (O-GlcNAcylation). Mass tolerances for

intact and product ion masses were set at ±0.01 Da and ±0.35 Da, respectively. The data

files were also searched using the program described in section 3.1.1, "Find GlcNAc,"

which identifies relative abundances of oxonium ion and charge reduced species. All

peptide hits were subject to manual interpretation of MS1 (accurate mass) and MS2

(fragment ion) spectra. In those cases that Mascot was employed to assist in *de novo*

sequencing, the following parameters were used: ±6 ppm precursor mass tolerance, ±0.6

Da fragment ion mass tolerance, no enzyme restriction, and variable O-GlcNAcylation on Ser and Thr.

4.3.6.1 Procedure I: Thermo polyacrylamide beads

10 μL of Thermo boronate coated polyacrylamide beads were diluted with 100 μL of anhydrous DMF and spun down for 5 minutes at 7500 rpm. The supernatant was removed, and the wash was repeated twice more. Dried sample was brought up in 20 μL DMF and added to the pre-washed beads. The reaction was allowed to progress for 5 or 60 minutes with agitation, and then spun down for 5 minutes at 7500 rpm. The supernatant was removed and saved for future analysis (if necessary). Beads were then resuspended in 100 μL acetonitrile and spun for 5 minutes at 7500 rpm. This was repeated, and wash supernatants were saved for future analysis as well. Finally, the beads were resuspended in 20 μL elution buffer (0.1 M acetic acid) and agitated for 30 minutes. The reaction was spun down for 5 minutes at 7500 rpm and rinsed once with an additional 20 μL elution buffer.

4.3.6.2 Procedure II: Derivatized magnetic beads

NHS-activated magnetic beads were derivatized as described. First, 14.6 mg MES was dissolved in 3 mL H_2O and pH adjusted to 5 with ammonium hydroxide. 100 μL of the NHS activated beads were removed and beads were separated from the supernatant using an external magnet. The beads were then washed 2x with 25 mM MES. 1 mg of aminophenyl boronic acid (APBA) was brought up in 100 μL 25 mM MES and added to the washed beads. This solution was slow vortexed overnight at 4ºC. The following day, 0.05 M ethanolamine in PBS was prepared (90.7 μL ethanolamine in 2.9 mL PBS, pH

adjusted to 8 with HCl). MES solution was removed from the beads using the external magnet, and 100 μL PBS solution was added to the beads. This mixture was incubated at room temperature for 60 minutes, then washed 2x with PBS solution. The beads were stored at 4°C in DMF until enrichment was performed. Enrichments were performed as described in section 3.3.6.3; however, instead of centrifuging the Eppendorf to remove supernatant, the external magnet was used.

Later experiments involved capping underivatized CO_2 groups on the magnetic beads using EDC-mediated derivatization with ethanolamine. Here, pyridine HCl buffer was made using 836 μL H_2O, 81 μL pyridine, and was pH adjusted with HCl to 5.5. This was used to prepare 1 M ethanolamine solution in the pyridine HCl buffer (60.5 μL ethanolamine into 939.5 μL pyridine HCl), and was again pH adjusted with HCl to 5.5. Four vials of 0.1M EDC HCl were made by dissolving 4.79 mg EDC in 250 μL H_2O. To 100 μL magnetic beads washed 3x with DMF, 20 μL 1M ethanolamine and 5 μL 0.1M EDC were added and allowed to react with agitation. Every 30 minutes, a fresh 5 μL aliquot of 0.1M EDC was added to the reaction. After 2 hours, the reaction was quenched by removing excess solution and rinsing 3x with DMF. Beads were stored at 4°C in DMF until enrichment was performed.

4.3.6.3 Procedure III: Derivatized POROS AL 20 beads

A Schiff's base reaction was performed on POROS AL 20 beads to couple the free amine on APBA with the aldehyde moiety on the POROS beads. First, PBS buffer was prepared and pH adjusted to 6-7. Then, 1 mL 0.2M APBA (30.99 mg) was prepared in PBS buffer and pH adjusted with NaOH to 6-7. Additionally, 100 μL of 80mg/mL

NaCNBH$_3$ was made by dissolving NaCNBH$_3$ in the PBS buffer. Finally, approximately 7 mg of POROS AL beads were dispersed into 200 μL 0.2M APBA solution, followed by the addition of 1 μL of 80mg/mL NaCNBH$_3$. The reaction was allowed to progress for 2 hours at room temperature with agitation, and was quenched by washing the beads on a spin column (pore size <20 μm) with water filtration 3x. After removing water and capping the filtration column, beads were stored dry at 4ºC. Once needed, the beads were brought up in 500 μL water, and 100μL of this solution was added to 400 μL water. 100 μL of the dilution solution was spun down, water was removed, and the beads were washed 3x with DMF. Some experiments involved derivatizing the APBA-POROS 20 beads using the EDC chemistry described in 3.3.6.4. Enrichments were then performed as described in section 3.3.6.3, except elution fractions were spun down at 9500 rpm.

4.3.7 Esterification and IMAC

In later experiments we tested whether an IMAC flow-through could be used for GlcNAc enrichment, so as to save precious tumor sample. In these instances, both phosphorylated and O-GlcNAcylated peptides were spiked into samples, then cleaned up using the procedure described in section 4.3.4. Samples were then esterified using a Fischer esterification. First, samples were dried in triplicate using 50 μL of completely anhydrous methanol. Concurrently, methanolic HCl was prepared by adding 160 μL dropwise to 1 mL of methanol. 80 μL of this solution was added to the dried sample and the reaction was allowed to proceed for 1 hour. At this time, the sample was dried and

rinsed once with 50 μL methanol. The entire process was repeated to assure completeness. The esterified and dried sample was stored at -35ºC until further use.

For the IMAC experiment, a 360 μm o.d. x 75 μm i.d. polyimide coated fused silica capillary fritted with a 2 mm Kasil frit was pressure loaded (200-500 psi) with 5 cm of POROS20 iminodiacetate packing material. The IMAC column was then prepared for iron (III) activation by the following rinse steps: 10 minute H_2O at 20 μL/min, 20 minute 50 mM EDTA at 20 μL/min, and 5 minute H_2O at 20 μL/min. At times that the frit did not allow for 20 μL/min flow rate, the rinse time was adjusted to allow the same amount of column volumes to pass through the IMAC column. For instance, a 10 μL/min flow rate would require the first H_2O rinse to be 20 minutes, totaling 200 μL through the column. The IMAC column was then activated using filtered iron (III) chloride ($FeCl_3$). Each activation step included a 10 minute $FeCl_3$ rinse at 20 μL/min followed by 3 minutes without pressure to allow the $FeCl_3$ to diffuse into the column particles. This was repeated three times to assure complete activation. The activated column was equilibrated with 0.01% acetic acid at 0.5 μL/min for 50 minutes.

Dried, esterified samples were brought up in 50 μL 1:1:1 0.01% acetic acid: methanol: acetonitrile. This solution was loaded onto the column at 0.5 μL/min. A final rinse of 0.01% acetic acid at 0.5 μL/min was performed for 30 minutes. All of the flow-through was collected and dried for future GlcNAc enrichment. Then, a PC was disconnected from an AC and rinsed with solvent A for 10 minutes at 30 bar. The PC was then butt-connected with Teflon to the top of the IMAC column. The IMAC-PC was allowed to rinse for 10 minutes with 0.01% acetic acid at a flow rate of 1 μL/min. If a

leak exists during this time, the pressure was let off and a new Teflon sleeve was created.

Phosphopeptides were eluted with 10 µL of 250 mM ascorbic acid solution for 10

minutes (1 µL/min flow rate). The IMAC-PC was rinsed for another 5 minutes with

0.01% acetic acid. The PC was then removed, rinsed with solvent A for 20 minutes at 30

bar, dried, and re-attached to an AC. The column was allowed to rehydrate for 20-30

minutes before the addition of 100 fmol of internal standards (angio and vaso). The

column was rinsed again for 10 minutes before elution.

Data analysis was carried out using Thermo Xcalibur software. Raw data files

were searched against the RefSeq database using OMSSA. MS2 searches used the

following parameters: no enzyme specificity, e-value cutoff of 1, fixed modifications of

the C-terminus, Asp, and Glu (methyl esters), and variable modifications of Met

(oxidation), Ser, Thr, and Tyr (phosphorylation). Mass tolerances for intact and product

ion masses were set at ±0.1 Da and ±0.35 Da, respectively. The data files were also

searched using the program described in section 2.1.4, "CAD Neutral Loss Finder,"

which identifies a neutral loss of 98 from the parent ion. All peptide hits were subject to

manual interpretation of MS1 (accurate mass) and MS2 (fragment ion) spectra. In those

cases that Mascot was employed to assist in *de novo* sequencing, the following

parameters were used: ±6 ppm precursor mass tolerance, ±0.6 Da fragment ion mass

tolerance, no enzyme restriction, fixed methyl esters on the C-terminus, Asp, and Glu,

and variable phosphorylation on Ser, Thr, and Tyr.

4.3.8 Preparing PBMCs (Performed by SAP/SAM)

PBMCs are separated from RBCs and granulocytes using density gradient separation on Ficoll Paque. 50 mL of blood is collected from healthy donors and centrifugred at 1100 x g for 10 minutes. The plasma is removed, EDTA is added to 2 mM, and the plasma is frozen. The remaining blood is mixed with an equal volume of plain RPMI medium and slowly layered over Ficoll Paque. The mixture is then spun at 800 x g for 20 minutes and washed twice with RPMI 10% FCS. The cells are counted and resuspended at 5×10^6/mL in AIM-V medium for future use.

4.3.9 Enzyme-linked immunospot (ELISpot) assays (Performed by SAP/SAM)

The enzyme linked immunospot (ELISpot) PRO human IFNγ kit was used for these experiments. Here, polyvinylidene difluoride (PVDF) membrane backed microplate strips pre-coated with monoclonal antibody (mAb) specific for human IFNγ were washed four times with sterile PBS; 200 μL/well. The wells were then blocked by incubation with RPMI 10% FBS for 30 minutes at RT. The medium was removed and a cell suspension of the PBMCs (5×10^5/well) in RPMI 10% FBS was added. The stimulating peptides were then added to the test wells along with phytohemagglutinin (PHA) to the positive control. The microplate was incubated for 7 days, then washed and restimulated with peptides for 16-24 hours at 37ºC. The cells were then removed and the plate was washed five times with PBS; 200 μL/well. The one-step detection agent (mAb 7-B6-1) was diluted 1:200 in PBS 0.5% FBS and 100 μL was added to each well. The plate was allowed to incubate for 2 hours before washing five times with PBS. The substrate

solution (BCIP/NBT-plus) was filtered through a sterile 0.45 μm filter and 100 μL was added to each well. When spots began to appear, the plate was extensively washed with H_2O and dried. The spots representing individual IFNγ secreting cells were counted using an ELISpot reader.

4.3.10 Intracellular cytokine staining (ICS) (Performed by SAP/SAM)

Cells were plated to $1x10^6$ per well in a 48 well plate. Synthetic peptide antigens were added to the wells (10μg/mL) and cells are expanded for 7 days. Positive control was stimulated with 1 μg/mL PHA, negative control stimulated with DMSO only. On day 6 cells were washed and restimulated with peptide antigen overnight. The cells were harvested, washed, and stained with fixable viability dye (APC-Cy7) and surface antibodies: anti-CD3 (APC) and anti-CD8 (PerCP; BD). Cells were fixed using 2% paraformaldehyde, permeabilized using 0.5% saponin, and stained with anti-IFNγ (PE), anti-IL2 (Pacific blue), and anti-TNFα (PE-Cy5.5) for 30 minutes at RT. Cells were washed, lightly fixed, and analyzed on the FACSCantoII flow cytometer.

4.4 Results

4.4.1 Sequencing the first HLA-associated O-GlcNAcylated peptides

Propelled in part by Haraum's observation of O-GlcNAcylated peptides in MHC class I coupled with our success in phosphopeptide analysis, we decided to search for O-GlcNAcylated peptides in tumor samples. Thus, we affinity-isolated human leukocyte antigen (HLA)-B*0702 peptide complexes from a primary acute lymphoblastic lymphoma (ALL), three primary chronic lymphocytic leukemia tumors (CLL), a primary hairy cell leukemia (HCL), Epstein-Barr virus transformed B lymphoblastoid cell lines (JY), and a primary acute myeloid leukemia (AML). In these samples, we identified the first MHC-associated O-GlcNAcylated peptide. Figure 4.3 shows the CAD (a) and ETD (b) spectra that allowed us to sequence this peptide and site-localize the modification. IPVsSHNSL originates from myocyte-specific enhancer factor 2C, which is a transcriptional activator implicated in vascular development. The protein is known to interact with CARM1, an enzyme with OGT activity. Despite this, the O-GlcNAc site is novel.

Additionally, in the ALL sample, three more O-GlcNAcylated peptides were found. All of these peptides have the same sequence (RPPItQSSL), but were detected with three different methylation states on Arg: unmodified, monomethylation, and an asymmetric dimethylation. Table 4.1 summarizes abundances of the four peptides and the samples in which they were detected. Figure 4.4 demonstrates the small molecule fragmentation pathways in ETD that allow us to make the dimethylation symmetry assignment. To our knowledge, these are the first methylated and O-GlcNAcylated peptides isolated from an HLA class I complex. One study reported an MHC class I

asymmetrically di-methylated peptide from an HLA-B*39 restricted sample; however, no

further tests were done to investigate its relevance.[30]
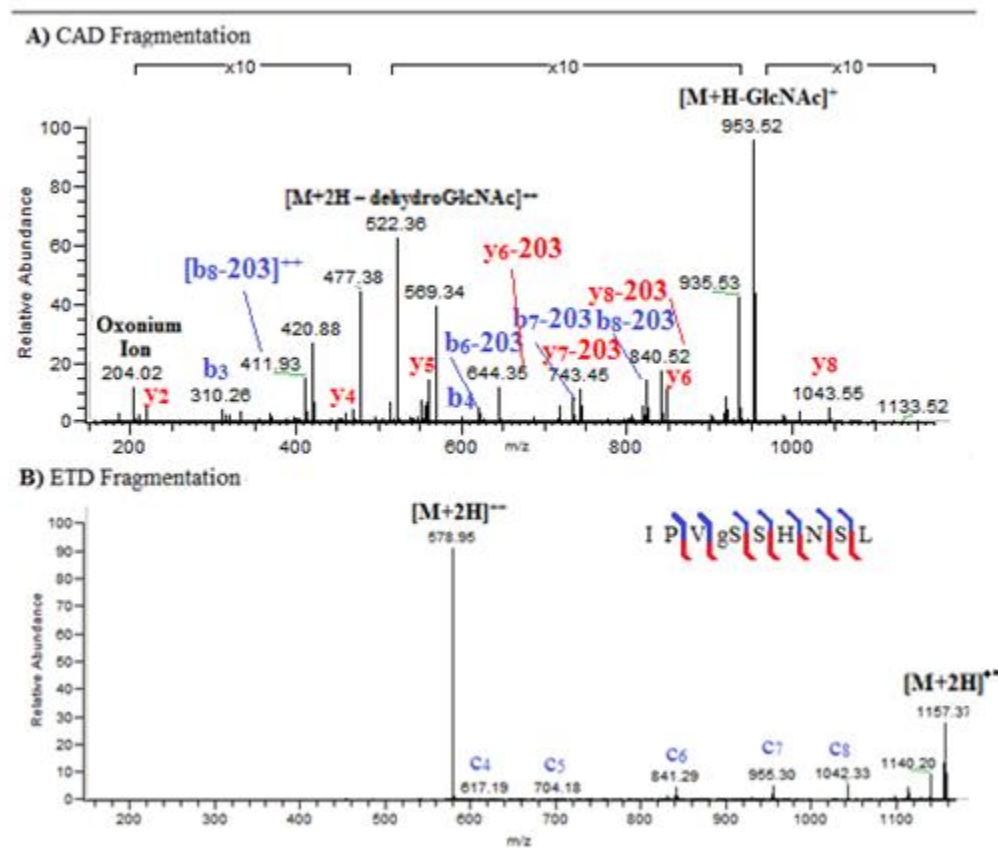
Figure 4.3 MS2 spectra of IPVsSHNSL, the first fully characterized HLA-associated O-GlcNAcylated peptide. (A) CAD fragmentation spectrum of the peptide shows signature ions as well as neutral losses allowing for assignment of most b and y ions. (B) ETD fragmentation spectrum of the peptide gives additional sequence information. Together, these spectra allow for nearly full sequence coverage.

Peptide	JY (2002)	JY (2011)	ALL	AML	CLL (donor 1)	CLL (Donor 2)	HCL	Relative Abundance
IPV(gS)SHNSL	■	■	■	■	■	■	■	*
RPPI(gT)QSSL			■					**
(CH₃) RPPI(gT)QSSL			■					***
(CH₃)₂RPPI(gT)QSSL			■					****

Table 4.1 First identified HLA-associated O-GlcNAcylated peptides in various samples. Abundances are shown in heat map format, with the key at right. IPVsSHNSL was detected in 7 samples with varying abundances. RPPItQSSL was detected in three different methylation states in the ALL sample only.
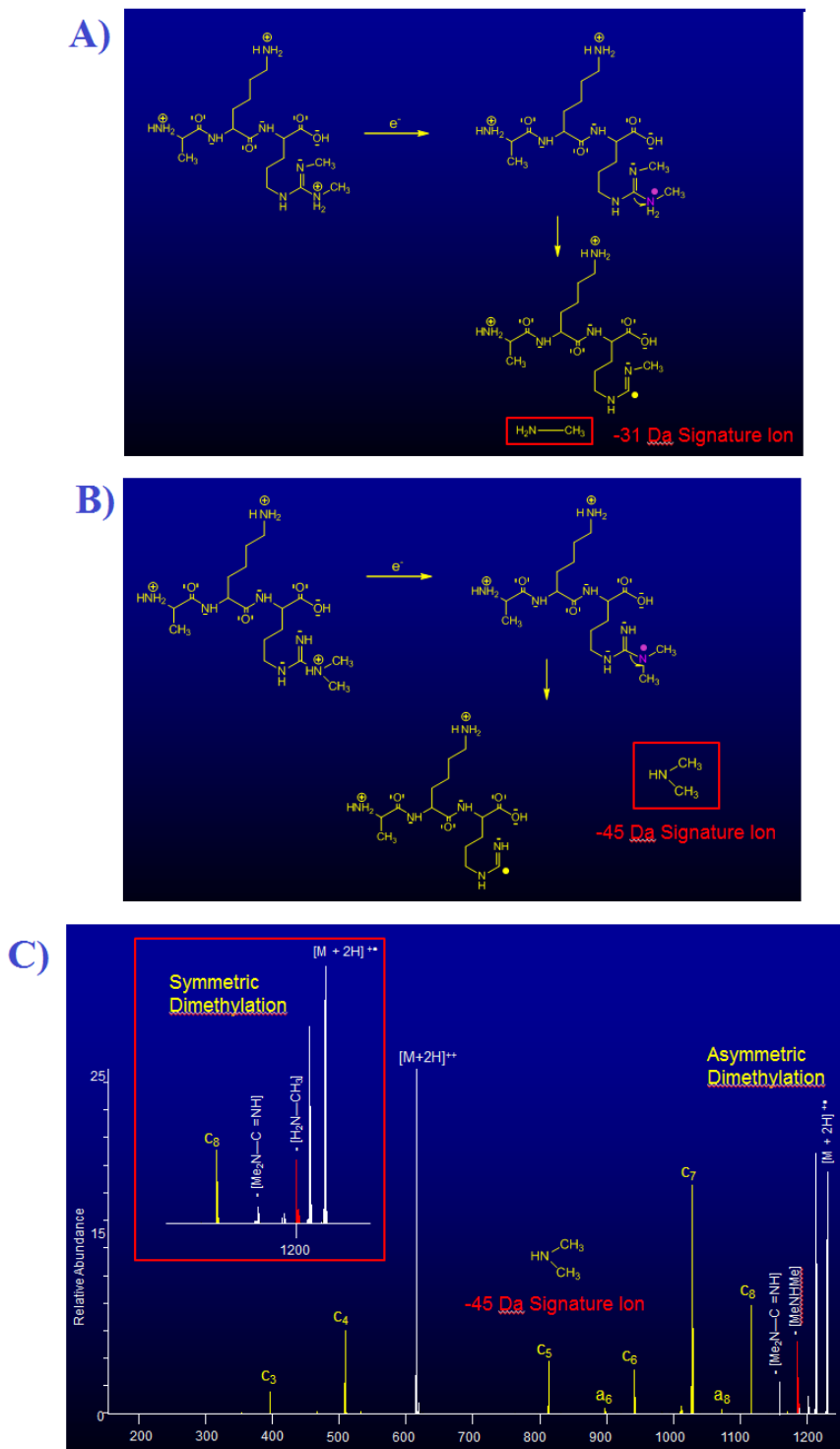
Figure 4.4 Small molecule fragmentation via ETD for the assignment of dimethylation states. (A) Symmetric di-methylation fragmentation pathway causes a -31 Da apparent loss from the charge reduced species. (B) Asymmetric di-methylation fragmentation pathway causes a -45 loss from the charge reduced species. (C) (Me₂)RPPItQSSL ETD spectrum allowing for the assignment of asymmetric dimethylation. Inset: synthetic peptide with symmetric dimethylation, used for confirmation purposes.

4.4.2 Enrichment of HLA-associated O-GlcNAcylated peptides

Several attempts were made to isolate additional O-GlcNAcylated peptide antigens from tumors without success. Thus, we concluded that an enrichment technique was necessary to detect them within the extremely large pool of unmodified peptides. To do this, we looked to our phenylboronic acid (PBA) enrichment procedure developed in the previous chapter.

4.4.2.1 Procedure I, Thermo polyacrylamide beads

After our initial success with synthetic peptide enrichment on Thermo polyacrylamide beads, we attempted enrichment of tumor samples with these beads. We started with a T-cell acute lymphoblastic leukemia (T-ALL; prepared by SAP) sample. This is a rare type of leukemia affecting T-cells instead of B-cells, usually occurring in older children and teenagers. After enrichment, we achieved a 34% recovery of internal standards (IPV-GalNaz, RPP-GalNAz). More importantly, we discovered a new O-GlcNAcylated peptide with sequence VLTsNVQTI, originating from ETS-related transcription factor Elf-1. We then decided to enrich an Epstein-Barr virus transformed B lymphoblastoid cell line (JY-B7-2002). We detected IPVsSHNSL in the 1e7 cell equivalent screen of this sample and wanted to determine if any other O-GlcNAcylated peptides were present. Even though we attained a 25% recovery of internal standards (IPV-GalNaz, RPP-GalNAz) and a 22% recovery of IPVsSHNSL, we did not detect any new O-GlcNAcylated peptides. Additionally, we had significant amounts of non-specific binding, which interferes with O-GlcNAc detection by lowering the probability of being

selected for MS2 fragmentation. Thus, we derivatized magnetic beads in order to lower this nonspecific binding.

4.4.2.2 Procedure II: Derivatized magnetic beads

After synthetic optimization of the APBA-derivatized magnetic beads, we enriched SKOV-3, an ovarian cancer cell line. Unfortunately, we only attained a 4% recovery of standards (IPV-Glc, RPP-Glc). We then attempted to enrich another Epstein-Barr virus transformed B lymphoblastoid cell line (JY-A2). Here we attained a 0% recovery of standards in the elution. The magnetic beads were abandoned after this venture, as we believe an interferent in the samples must affect sample binding.

4.4.2.3 Procedure I + C18 clean-up

Prior to phosphopeptide enrichment, a C18 clean-up is performed to assure salts and contaminants are removed from the sample. By analogy, we postulated that we could increase O-GlcNAcylated peptide recovery by doing a clean-up before enrichment. Thus, we cleaned up the SKOV-3 sample and enriched with polyacrylamide beads. Here, we attained a 31% recovery of internal standards (IPV-Glc, RPP-Glc). More importantly, we sequenced another O-GlcNAcylated peptide with two O-GlcNAcs present: ALTtsAHSV, originating from homeodomain-interacting protein kinase 3 (HIPK3). We have evidence for the singly-modified peptide by MS1 accurate mass, but due to ionization suppression by an interferent, we were unable to attain MS2 confirmation. We went on to enrich a melanoma sample (DMM122, B7B18) using this altered protocol. Here, we attained a 75% average recovery of internal standards (VLT-Glc, RPP-Glc) and were able to sequence another new O-GlcNAcylated peptide antigen PTASQARsTI, from SH3 domain-containing RING finger protein 3.

Based on these successes, we went on to re-enrich the JY-A2 sample. Prior to C18

clean-up, I came to the realization that CHAPS buffer is always present in the elution of

the enrichment and is usually the most abundant ion present (Figure 4.5). We postulated

that CHAPS may be occupying free boronic acid sites and interfering with O-

GlcNAcylated peptide recovery. Thus, in order to eliminate CHAPS, the clean-up elution

was collected for the first 23 minutes only. By doing this, we completely eliminated
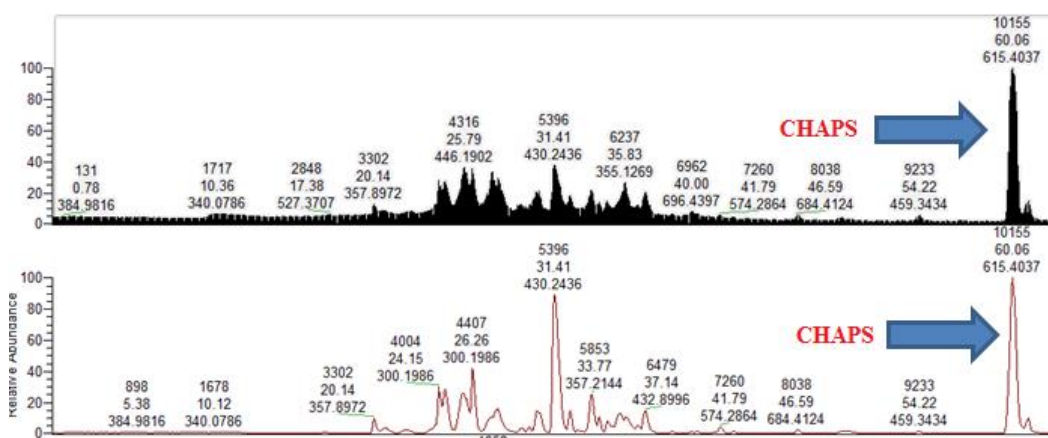
Figure 4.5 CHAPS buffer interference. Top panel: total ion current, bottom panel: base peak from T-ALL enrichment elution. The ion at 615 m/z corresponds to CHAPS and is the most abundant ion present in the elution. Thus, the clean-up procedure was altered to exclude CHAPS but collect all peptide content.

CHAPS binding, and attained a 49% average recovery of internal standards (IPV-Glc,

RPP-Glc). However, we detected VLTsNVQTI by MS1 in the screen of this cell line, and

did not attain an MS2 of the peptide in the elution. Similarly, an AML cell line THP-1

showed evidence for VLTsNVQTI along with another O-GlcNAcylated peptide in the

screen; the latter was too weak to sequence. After enrichment, we attained a 47% average

recovery of standards, but did not attain an MS2 of either O-GlcNAcylated peptide.

To combat this issue, we performed two experiments: extending the gradient from 60 to 120 minutes, and performing a top-10 CAD (instead of a top 5 CAD/ETD toggle). Here, JY-A2 was again cleaned up and enriched. In both experiments, we recovered approximately 60% of internal standards (Me-RPP-Glc, RPP-Glc), but only in the extended gradient experiment did we obtain an MS2 of VLTsNVQTI. Thus, we continued to use this extended gradient for all subsequent experiments.

4.4.2.4 Utilization of IMAC flow-through

Two problems were evident in the majority of these experiments: (a) tumor samples are limited and (b) nonspecific binding interferes with detection of O-GlcNAcylated peptides. Utilizing the flow-through from an IMAC experiment could combat both enduring issues. This experiment would allow for two tandem enrichment procedures, saving precious tumor sample while depleting a large amount of interfering peptides. We started by cleaning up and esterifying JY-A2, then performed IMAC. The flow-through was collected during IMAC and dried down for subsequent O-GlcNAc enrichment. The IMAC elution was not analyzed in this experiment. In the O-GlcNAc enrichment elution, we recovered 32% of the internal standard (RPP-Glc) spiked in pre-clean up, and 73% of the internal standard (IPV-Glc) spiked in post-IMAC. This experiment confirmed proof-of-principal and stimulated a repeat experiment using ALL01-B7. In this experiment, the IMAC elution was analyzed. Here, 0% of any O-GlcNAcylated peptides were detected, a 37% recovery of phosphopeptide internal standards was achieved, and 31 phosphopeptides were sequenced (Table 4.2). Four of these phosphopeptides had never been detected before, 4 were new to the sample, and the rest were seen previously in an IMAC experiment performed by Andrew Norris (2006).

In the O-GlcNAc enrichment elution, 28% average recovery of known O-GlcNAcylated peptides were detected (IPV-Glc, RPP-Glc, along with various methylation states). Thus,

Phosphopeptides	Protein of Origin	HLA type	ALL01
KPAVSRsRSSSL	KIAA0556	B*07	
RGDsRPRLV	PIEZO1	B*07	
RPLsATRKTL	FAM65B	B*07	
RPRDtRRISL	CIC	B*07	
KPFKLSGLsF	RSRC2	B*07	
KPRRFsRSL	MARKSL1	B*07	
KPVsPKSGTL	ARHGEF7	B*07	
KPYsPLASL	NFATC2	B*07	
KLSGLsF	MARKSL1	B*07	
KPPHsPLVL	MYC	B*07	
KPPYRSHsL	CEP95	B*07	
RAPsPSSRM	SRRM2	B*07	
RGIsPIVF	YHDF2	B*07	
RPAKsmDSL	ARHGAP30	B*07	
RPAsAGAmL	MEF2D	B*07	
RPDsRLGKTEL	SETD2	B*07	
RPFsDSDAL	*	B*07	
RPFsPREAL	LUZP1	B*07	
RPHsPEKAF	C17orf85	B*07	
RPPSsSQQL	KATbB	B*07	
RPQRATsNVF	MYL6	B*07	
RPsSLPDL	ARID1B	B*07	
RPTsRLNRL	NCOR1	B*07	
RPVsPFQEL	*	B*07	
RPVsPGKDI	HIVEP2	B*07	
RPVsPGKDV	*	B*07	
RPVtPVSDL	KLF10	B*07	
RPYsPPFFSL	FAM53C	B*07	
RSMsGGHGL	TSC2	B*07	
RSVsPTFL	AKAP11	B*07	
TPIsPGRASGm	RUNX1	B*07	

Key

	0.1-1 fmol/g tissue
	1-10 fmol/g tissue
	10-100 fmol/g tissue
	>100 fmol/g tissue
	New phosphopeptide
	New to sample

Table 4.2 Phosphopeptides identified from ALL01-B7 sample. 31 phosphopeptides were detected with predicted binding alleles from HLA-B*07. 4 of these phosphopeptides (pink) had never been sequenced before, 4 of them (green) are new to the sample, and the rest were previously detected by AN. Abundances are shown in heat-map format; key at right. Starred proteins were not found in the Uniprot database.

we can enrich O-GlcNAcylated peptides from IMAC flow-throughs, effectively enriching

for two modifications in one experiment. However, nonspecific binding remains as the

largest issue with the O-GlcNAc enrichment.

4.4.2.5 Procedure III: Derivatized POROS 20 AL beads

Due to the success achieved with APBA-derivatized POROS beads using

synthetic peptide mixtures, we decided to enrich ALL01-B7 sample on these beads.

Using these beads, we attained an average internal standard recovery of 59% (VLT,

TVA, APAV, APAT) and an average recovery of 90% for those four known O-

GlcNAcylated peptides present in the sample (IPV-Glc, RPP-Glc, along with various

methylation states). The POROS 20 beads carry the benefit of less nonspecific binding,

which allowed us to sequence 3 new O-GlcNacylated peptide antigens: APVsSKSSL

(homeodomain interacting protein kinase 1), DPRAtAQQL (coiled-coil domain

containing protein 142), and RPVtASITTM (MKL/myocardin-like protein 2).

Nearly all of the aforementioned experiments were performed with a top 5

CAD/ETD toggle instrument method, but even with enrichment, the O-GlcNAcylated

peptide antigens are rarely in the top 5 most abundant species for any given MS1. Thus, a

neutral loss method was developed whereby the top 10 ions are subjected to CAD

fragmentation. Then, if a neutral loss of 101 ($[GlcNAc]^{++}$), 110 ($[GlcNAc+ H_2O]^{++}$), or

67 ($[GlcNAc]^{+3}$) was detected from the parent ion (± 1 Da), ETD is triggered on that

parent ion. This instrument method increases the sensitivity achieved by a single

experiment. We again decided to enrich ALL01-B7 sample on derivatized POROS 20

beads using this instrument method. In this experiment, we attained a 59% average

recovery of the four internal standards and a 47% recovery of the known O-

GlcNAcylated peptides in the sample. Each of these 8 peptides had a CAD and an ETD spectrum, fulfilling the goal of the experiment. More importantly, the experiment led to the identification of 5 new O-GlcNAcylated peptides, including: TPIsQAQKL (E1A binding protein p400), RPPVtKASSF (R3H containing protein #2), IPIsLHTSL (protein PRRC2B), APTsAAAL (HIPK1), and a di-GlcNAcylated peptide HPSstAAVL (Pro/Ser rich protein 1). Also present were approximately 10 peptides with MS2s displaying signature ions but were mixed or were too low level to sequence. To assure that this method would work with other samples, O-GlcNAc enrichment was performed on a T-ALL02 plasma sample. We attained an 83% recovery of the four internal standards, and once again detected new O-GlcNAcylated peptide antigens: TAStRPSV (protein strawberry notch homolog 1) and VASsRSASM (transcription factor 12). Thus, this instrument method can aid in the detection and sequencing of O-GlcNAcylated peptides.

To this point, method optimization of O-GlcNAcylated peptide enrichment on HLA-associated peptides included: adding a C18 clean up to deplete salt and buffer contaminants, derivatizing POROS-20 AL beads to lower nonspecific binding, manipulating instrument methods to increase apparent sensitivity, and utilization of tandem IMAC/GlcNAc enrichment to conserve precious tumor tissue. Using these various methods, we identified and sequenced a total of 19 peptide antigens in several modification states from 10 types of cancer. These novel O-GlcNAcylated peptide antigens and their corresponding samples, proteins, alleles, and abundances are depicted in Table 4.3. The key is shown below:

Key

Shade	Meaning
(lightest blue)	0.1-1 fmol/g tissue
(light blue)	1-10 fmol/g tissue
(medium blue)	10-100 fmol/g tissue
(dark navy)	>100 fmol/g tissue
(gray)	HLA mismatch

Legend for table cells below: L = 0.1-1 fmol/g, M = 1-10 fmol/g, D = 10-100 fmol/g, N = >100 fmol/g, mm = HLA mismatch, (blank) = not detected.
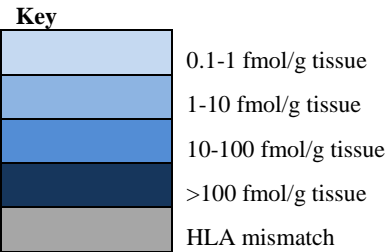
O-GlcNAcylated peptides	Protein of Origin	HLA type	ALL-01	AML01	CLL GF*	CLL05*	CLL07*	CLL09*	DMM122	FHIOSE*	HCL01*	JY-A2	JY-B7	SKOV-3	T-ALL01	T-ALL02 plasma	THP-1
ALTTs*AHSV	Q9H422	A*02	mm	mm		mm	mm	mm		mm	mm		mm	M		mm	
ALTtsAHSV	Q9H422	A*02	mm	mm		mm	mm	mm		mm	mm		mm	M		mm	
VLTsNVQTI	P32519	A*02	mm	mm		N	mm	mm	N	mm	M	mm		N	mm	mm	L
TAStRPSV	A3KN83	A*29	mm	mm	mm	mm	mm	mm	mm	mm	mm	mm	mm	mm	mm	L	mm
VASsRSASM	Q99081	A*29	mm	mm	mm	mm	mm	mm	mm	mm	mm	mm	mm	mm	mm	L	mm
(diMe)RPPItQSSL	Q9P2N5	B*07	N			mm				mm		mm		mm	mm	mm	mm
(Me)RPPItQSSL	Q9P2N5	B*07	M			mm				mm		mm		mm	mm	mm	mm
APTsAAAL	Q86Z02	B*07	M			mm				mm		mm		mm	mm	mm	mm
APVsSKSSL	Q86Z02	B*07	M			mm				mm		mm		mm	mm	mm	mm
DPRAtAQQL	Q17RM4	B*07	M			mm				mm		mm		mm	mm	mm	mm
HPMsTASQV	Q8N6B4	B*07	L			mm				mm		mm		mm	mm	mm	mm
HPS(st)AAVL	Q86XN7	B*07	L			mm				mm		mm		mm	mm	mm	mm
IPIsLHTSL	Q5JSZ5	B*07	L			mm				mm		mm		mm	mm	mm	mm
IPVsSHNSL	Q06413	B*07	N	M	L	mm	N	M		mm	N	mm	N	mm	mm	mm	mm
RPPItQSSL	Q9P2N5	B*07	M			mm				mm		mm		mm	mm	mm	mm
RPPVtKASSF	Q9Y2K5	B*07	M			mm				mm		mm		mm	mm	mm	mm
RPVtASITTM	Q9ULH7	B*07	M			mm				mm		mm		mm	mm	mm	mm
TPIsQAQKL	Q96L91	B*07	L			mm				mm		mm		mm	mm	mm	mm
PTASQARsTI	Q8TEJ3	B*18	mm	mm	mm	mm	mm	mm	L	mm	mm	mm	mm	mm	mm	mm	mm

Table 4.3 HLA-associated O-GlcNAcylated peptide antigens identified from various tumors. 19 peptide antigens were detected with various modification states in a wide array of tumors, including: ALL, CLL, HCL, AML, T-ALL ovarian cancer, melanoma, and transformed cell lines. Relative abundances are corrected for amount of tissue used and recovered, and are shown in heat-map form.

4.4.3 Peptide synthesis and biological testing

4.4.3.1 Enzyme-linked immunospot assay (ELISpot; performed by SAP)

We then evaluated immunity to the O-GlcNAcylated peptide antigens in 10 healthy donors (HD) using a 7-day cultured interferon-γ (IFN γ) enzyme-linked immunospot (ELISpot) assay, which we previously used to assess phosphopeptide immunogenicity. We synthesized 6 tumor-associated O-GlcNAcylated peptides: 5 HLA-B*07 restricted peptides including IPVsSHNSL and 4 variations of RPPItQSSL. We also
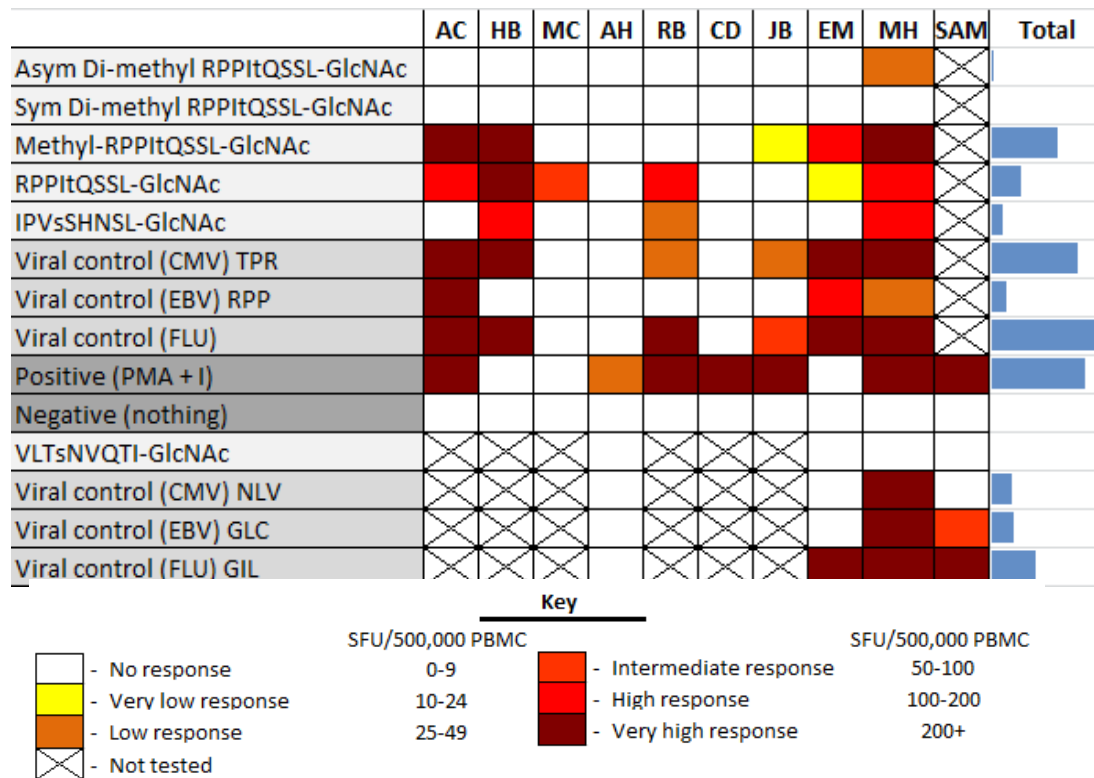
	AC	HB	MC	AH	RB	CD	JB	EM	MH	SAM	Total
Asym Di-methyl RPPItQSSL-GlcNAc									Low	X	
Sym Di-methyl RPPItQSSL-GlcNAc										X	
Methyl-RPPItQSSL-GlcNAc	Very high	Very high					Very low	High	Very high	X	▮
RPPItQSSL-GlcNAc	High	Very high	Intermediate		High		Very low	High		X	▮
IPVsSHNSL-GlcNAc		High		Intermediate						X	▮
Viral control (CMV) TPR	Very high	Very high		Intermediate		Low		Very high	Very high	X	▮
Viral control (EBV) RPP	Very high							High	Low	X	▮
Viral control (FLU)	Very high				High	Intermediate		Very high	Very high	X	▮
Positive (PMA + I)	Very high			Low				Very high	Very high		▮
Negative (nothing)											
VLTsNVQTI-GlcNAc	X	X	X		X	X	X				
Viral control (CMV) NLV	X	X	X		X	X	X		Very high		▮
Viral control (EBV) GLC	X	X	X		X	X	X		Very high	Intermediate	▮
Viral control (FLU) GIL	X	X	X		X	X	X	Very high	Very high		▮

Key

	SFU/500,000 PBMC			SFU/500,000 PBMC
□ - No response	0-9	▮ - Intermediate response		50-100
▮ - Very low response	10-24	▮ - High response		100-200
▮ - Low response	25-49	▮ - Very high response		200+
⊠ - Not tested				

Table 4.4 HD responses to O-GlcNAcylated peptide antigens. O-GlcNAcylated peptide antigens detected in tumor tissues were synthesized and donor responses were tested against them using 7-day IFN γ ELISpot assays. HDs were only tested if their HLA-type matched the peptide antigen. Responses across all donors are summed and shown by blue bars.

tested 1 HLA-A2 peptide, VLTsNVQTI. To compare relative immunogenicity of individual peptides, we utilized HLA-specific peptide viruses from cytomegalovirus,

Epstein-Barr virus, and influenza along with positive (PHA) and negative (DMSO) controls. As seen in Table 4.4, responses were heterogeneous, varying in type and magnitude between different donors. These data suggest that O-GlcNAcylated peptide antigens may be of great immunotherapeutic use.

The viral controls for both HLA allotypes provide a benchmark for strong immunotherapeutic candidates, since most adults have been exposed to the aforementioned viruses and should thus have a memory response to them. As evidenced by the large blue bars visible on the right of the chart, this is the case for all HLA-A*02 and HLA-B*07 viral peptide antigens. Using these values for comparison, it is evident that at least 3 of the 5 O-GlcNAcylated peptide antigens are immunogenic. Both IPVsSHNSL and RPPItQSSL stimulated a response in several B*07 positive HDs and the stimulation was quite potent in magnitude. Stronger yet was (Me)-RPPItQSSL, as this peptide stimulated a response in numerous B*07 positive HDs. The combined response against this peptide was even larger than the response against Epstein-Barr virus in these individuals. Interestingly, this peptide was the least abundant of the 5 peptides tested in HDs (see Table 4.3). Since T-cells can recognize antigens present as low as 5 copies/cell, it is not surprising that this peptide present at 10 copies/cell can still generate an intense immune response.[5] Overall, these data suggest that O-GlcNAcylated peptide antigens may be of great immunotherapeutic use.

4.4.3.2 Intracellular cytokine staining (ICS; performed by SAP)

Cytokine production in response to specific stimuli can be measured using intracellular cytokine staining (ICS).[31] Golgi inhibitors are used to prevent cytokine

release from the cells. Cells are then fixed and permeabilized to allow antibodies against individual cytokines to access the intracellular compartments. Here, we tested one HD's responses against the O-GlcNAcylated peptide antigens using three cytokines: (1) IFN γ, responsible for induction of anti-viral protein synthesis, (2) tumor necrosis factor α (TNFα), responsible for inhibition of viral gene expression, and (3) interleukin-2 (IL-2), necessary for induction of T-cell proliferation. Additionally, we monitored for CD107a, a cell surface protein indicative of degranulation and suggestive of T-cell killing. Results are displayed in Figure 4.6.

Here the ability of individual peptides to stimulate cytokine production was assessed and is reported as a percentage of CD8+ cells expressing each cytokine. (di-Me)RPPItQSSL (symmetric and asymmetric) and VLTsNVQTI, in agreement with the ELISpot assay, did not produce significant cytokine responses. However, RPPItQSSL and IPVsSHNSL generated a moderate level of cytokines above background. Again, (Me)-RPPItQSSL emerged as a key target of the memory cells. Once again, the responses were on the order of the viral controls, and were even higher than that of EBV. These CD8+ cells produced multifunctional cytokine responses and upregulated CD107a in response to the stimulus, demonstrating their ability to eliminate a cell presenting this O-GlcNAcylated peptide.

These data corroborate the ELISpot assay results in that the O-GlcNAcylated peptide antigens are capable of stimulating an immune response. The ICS data specifically demonstrates that the response resides in the central memory compartment, suggesting an O-GlcNAcylated peptide-specific immunity in this individual. The implications of these findings will be discussed in the following section.
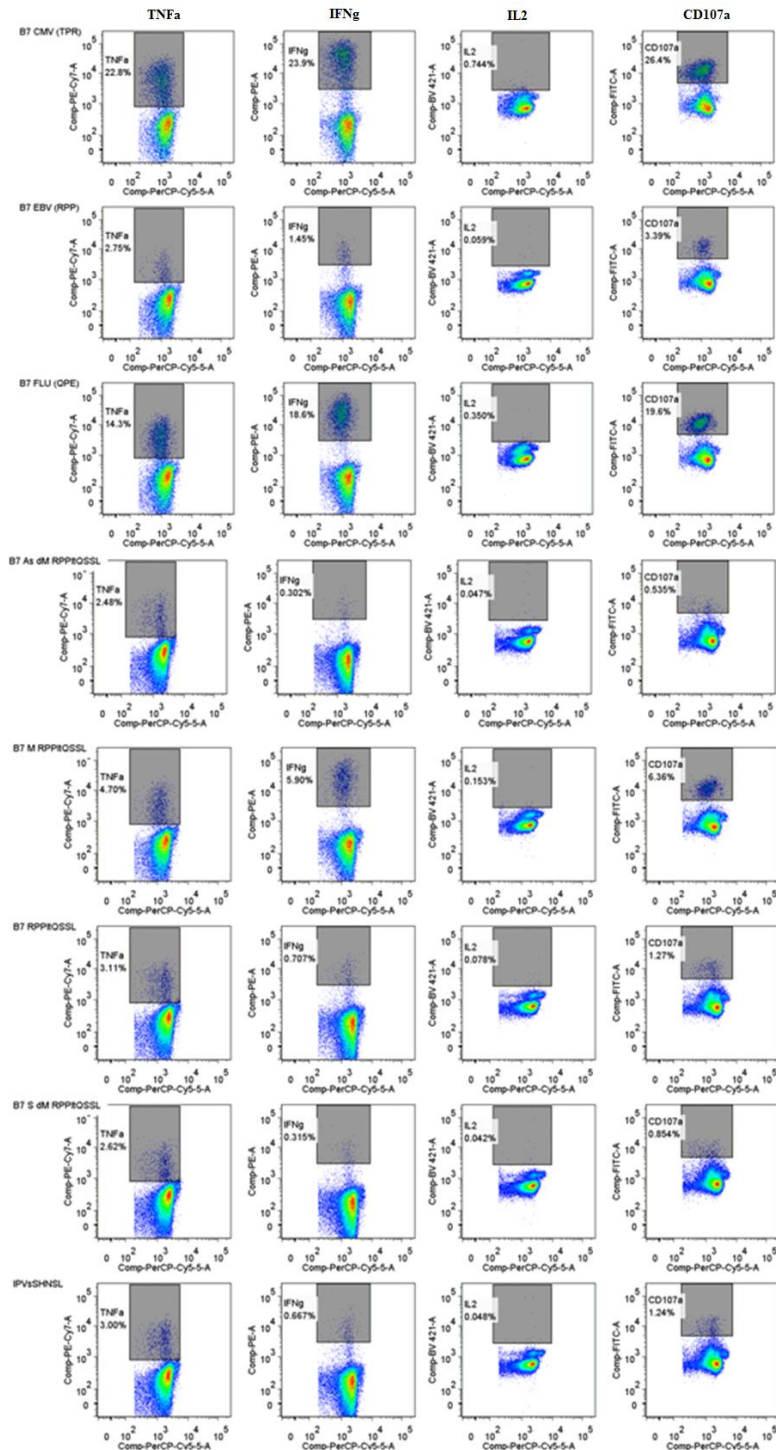
Figure 4.6 Healthy donor CD8+ T lymphocyte secretion of cytokines in response to HLA-B7 O-GlcNAcylated peptide antigens. ICS results from HD MH, measuring CD8+ production of 3 cytokines against 4 O-GlcNAcylated peptide antigens. Responses are shown as a percentage of total CD8+ cells. Percentage of CD8+ T-cells was determined by comparing the signals of cells displaying the CD3 to the signals of cells displaying both CD3 and CD8.
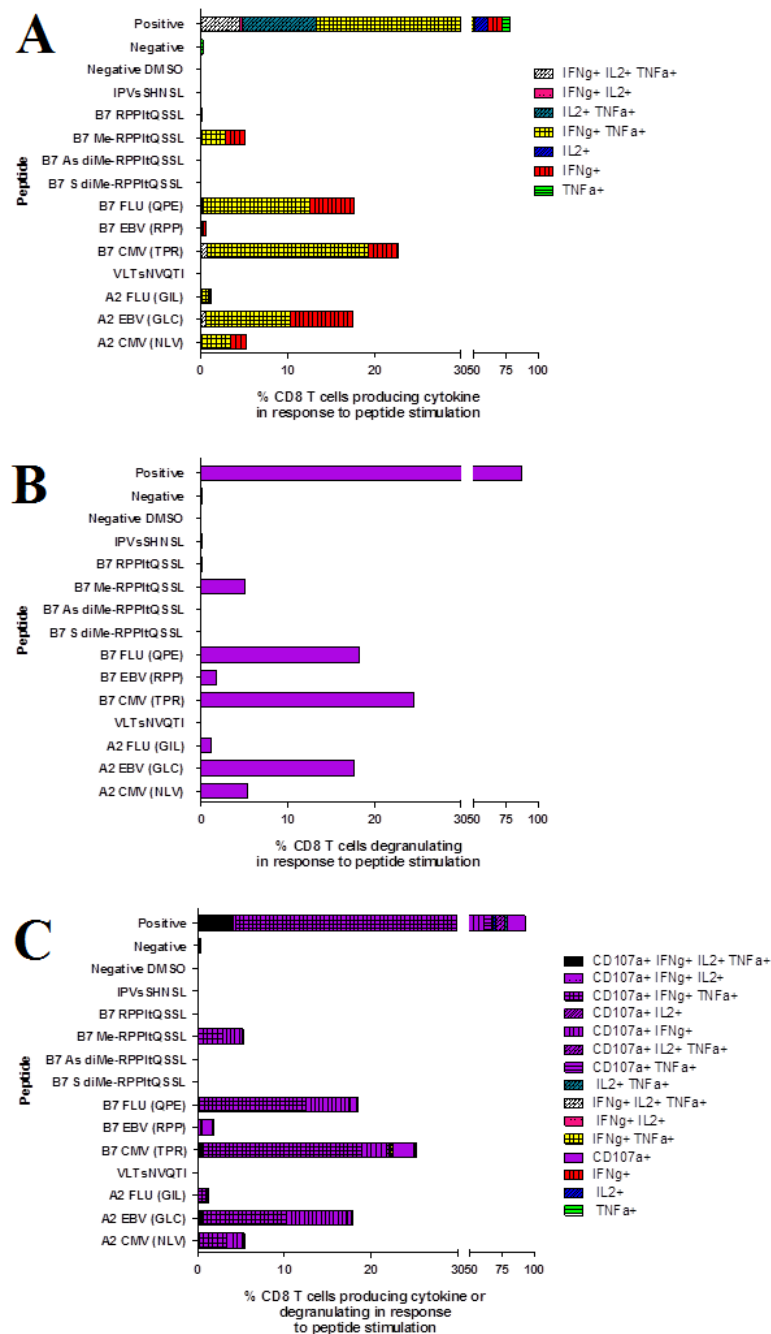
Figure 4.7 Compiled ICS Data. Cytokine responses of individual T-cells were measured using ICS. A) Each color represents a different combination of cytokines or a single cytokine released by activated T-cells. Multi-cytokine responses to the peptide antigens are seen and are shown as a percentage of total CD8+ cells. RPP(Me) shows a response on the same order of magnitude as those responses to viral controls. B)   CD107, a marker of T-cell killing and degranulation, is measured in response to various peptides and viral controls. Again, RPP(Me) shows a strong response. C) Figures 4.7A and 4.7B are combined to show total percentage of CD8+ T-cells producing cytokine or degranulating.

4.5 Discussion

In this study, we sequenced the first HLA-associated O-GlcNAcylated peptides. By optimizing the enrichment procedure for tumor samples, we were able to enrich for O-GlcNAcylated peptides using our phenylboronic acid enrichment methodology. Several cancers were investigated, including: ALL, T-cell ALL, melanoma, ovarian, and Epstein-Barr virus transformed B lymphoblastoid cell lines (JY). In total, we identified and sequenced 19 peptide antigens, shown in Table 4.3. To our knowledge, these are the only known HLA-associated O-GlcNAcylated peptides ever sequenced. Intriguingly, most of these antigens originate from pathways known to be dysregulated in malignant cells. Further, according to Uniprot and diBOGAP, none of these O-GlcNAcylation sites have been reported previously. Together, this suggests that aberrant O-GlcNAcylation caused by cancer progression and dysregulated cell signaling may be the cause of these post-translational modifications. Thus, we went on to synthesize these peptides in order to test their immunogenicity. Several peptide antigens stimulated a strong immune response in healthy individuals (Table 4.4), and these responses were largely within the central memory compartment (Figure 4.7). This suggests that healthy individuals may have been exposed to a stimulus that establishes immunological memory to these tumor-associated peptides.[31]

Notably, many of these peptides have multiple modifications present. Two of the 19 peptides have two adjacent O-GlcNAc sites. Until the crystal structure of an O-GlcNAcylated peptide in an MHC binding pocket was solved, many believed that the O-GlcNAc modification would not be conducive to T-cell receptor binding.[25] Modeling studies may elucidate the ability of the second O-GlcNAc modification to be

accommodated in the standard TCR-MHC geometry. Additionally, two peptides have methylated arginines present; one with a mono-methylation, the other with an asymmetric di-methylation. To our knowledge, these are the first methylated and O-GlcNAcylated peptides isolated from an MHC class I complex. One study reported an MHC class I asymmetrically di-methylated peptide from an HLA-B*39 restricted sample; however, no further tests were done to investigate its relevance.[30] We synthesized all variations of this peptide, including a symmetric dimethylated version, and tested its ability to stimulate an immune response. Interestingly, the symmetric dimethylated peptide did not generate a response in any of the individuals tested. However, all of the other forms (asymmetrically dimethylated, methylated, and unmethylated) stimulated responses in ELIspot and ICS that were inversely proportional to their detected abundances. Based on these responses, we are interested in investigating these multiply modified HLA-associated peptides further.

Also important, several of these peptides originate from transcription factors and/or kinases related to essential pathways for malignant transformation. For instance, TPIsQAQKL derives from E1A-binding protein p400, a SWI/SNF component of the NuA4 histone acetyltransferase complex. This complex acetylates histones H2A and H4 in order to upregulate expression of E2F1 and Myc target genes, thereby affecting cell cycle progression and proliferation.[32-24] Additionally, 2 of the peptide antigens (APTsAAAL and APVsSKSSL) originate from homeodomain interacting kinase 1 (HIPK1), a serine/threonine protein kinase involved in transcription regulation and TNF-mediated cellular apoptosis.[35-37] Interestingly, mice without HIPK1 developed fewer and smaller tumors than those with HIPK1.[38] Preliminary evidence generated by our

laboratory suggests that O-GlcNAcylation on HIPK1 prevents apoptosis, suggesting a

role in tumorigenesis and cancer progession (unpublished; JLB dissertation). Finally, a

doubly-O-GlcNAcylated   peptide   from   ovarian   cancer   was   sequenced   from

homeodomain-interacting protein kinase 3 (HIPK3), a relative of HIPK1. This kinase is

involved with negatively regulating apoptosis through the Fas-mediated pathway.[39,40]

Here, HIPK3 phosphorylates FADD, increasing evasion of apoptosis and causing drug

resistance in cancer treatments that target FasL receptors.[41] As demonstrated by Figure

4.7, the parent proteins of the O-GlcNAcylated peptides interact with or modify several

key cell signaling proteins. Many of these proteins reside within the nuclear envelope,

where O-GlcNAcylation levels are the highest. Taken together, it is apparent that these

peptides likely originate from dysregulated pathways present in transformed cells.
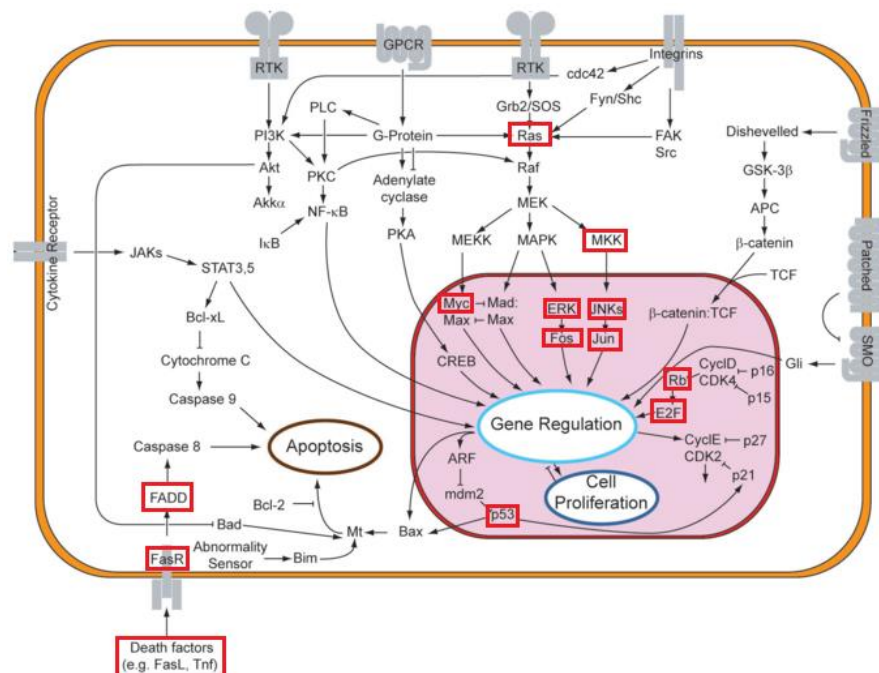

Figure 4.8 Key cell signaling pathways. The parent proteins identified in this study interact with or directly modify those proteins highlighted by red boxes. Most of these are within the nuclear envelope, where O-GlcNAcylation plays an important role in gene regulation and cell proliferation. Also, O-GlcNAc seems to downregulate apoptosis, leading to the evasion of death in tumor cells.

Compared to the phosphopeptide enrichments described in Chapter 2, HLA-associated O-GlcNAcylated peptide antigens appear far less abundant. One reason for this observation may be that the O-GlcNAc modification prevents protein degradation.[42] If this is the case, fewer O-GlcNAcylated peptides will be generated for presentation on the cell surface. Another reason could simply be that the modification is less abundant than phosphorylation. However, current estimates suggest that the two modifications are of similar abundance within the cell.[7] That being said, the list of O-GlcNAcylated peptide antigens generated in this study is almost certainly not exhaustive. This is evidenced by approximately 10 spectra in ALL01 enrichment elution that contained signature ions for O-GlcNAcylated peptides but were either (a) mixed with other species or (b) too low-level to sequence. Additionally, most of the O-GlcNAcylated peptide antigens we identified were isolated from HLA-B7 individuals. This may be due to the higher number of basic residues (Arg, Lys), and thus higher charges, on these peptides. The higher charge contributes to better ETD spectra, which are helpful in sequencing and site-localizing the modification. Since the O-GlcNAc modification often occurs in regions of low charge density (high Ser/Thr content), especially in HLA-A2 samples, many of the peptides generated have a single charge on the amino terminus and will not fragment well by CAD or ETD. Efforts are underway to increase charge density by derivatizing O-GlcNAcylated residues. Hopefully, this will enhance our ability to detect these peptides and help elucidate their true abundance.

Forthcoming studies will focus on continual optimization in order to identify more HLA-associated O-GlcNAcylated peptide antigens. Most of the optimization plans were outlined in section 3.5, including: moving the reaction on-column, attempting

different magnetic beads, and decreasing nonspecific binding. Specific optimization procedures for HLA-associated O-GlcNAcylated peptides will involve (a) changing the clean-up procedure and (b) utilization of a new instrument with enhanced detection abilities. To begin, the current clean-up gradient ramps from 0-80%B in 40 minutes, meaning that every minute is accompanied by a 2% change in solvent B. Peptides and CHAPS elute within a very narrow window of this gradient, making it exceptionally difficult to separate peptides from CHAPS routinely and cleanly. Thus, we plan to optimize this gradient to make this procedure more robust by extending the gradient linearly (i.e. 0-80% in 80 minutes) or slowing the gradient where peptides elute (i.e. 0-20% in 10 minutes, 20-60% in 60 minutes, 60-80% in 10 minutes). This should make the overall procedure more routine for both O-GlcNAc enrichment and tandem IMAC/O-GlcNAc enrichment.

Additionally, our laboratory has acquired a new high-resolution mass spectrometer with novel detection capabilities. The Thermo Orbitrap Fusion Tribrid, unlike the Orbitrap Classic or Orbitrap Velos, can perform up to 50 fragmentation events per MS1 due to parallelization of spectra acquisition. In other words, while the instrument is still acquiring an MS1, it can also fragment parent ions, making the entire process faster. Seeing as one of the most difficult aspects of O-GlcNAcylated peptide antigen identification is actually obtaining MS2 spectra for the peptides, this instrument opens up possibilities in detection limits that were previously unavailable. Future experiments using this instrument will utilize a similar method as described in section 4.3.5, whereby ETD is triggered only when a neutral loss ion is detected in CAD. However, instead of a top 10 method, we can likely obtain CAD spectra on the top 30-40

precursor ions present in each MS1. Alternatively, the Fusion allows for triggered ETD when product ions, such as the oxonium ion at 204 m/z, are present. This may be an even more reliable method, since occasionally the neutral loss ion is absent in O-GlcNAcylated peptide spectra. Overall, this instrument increases the limit of detection and sensitivity of our current methods, likely augmenting our ability to identify novel O-GlcNAcylated peptides.

Finally, additional peptides have been synthesized and shipped to our collaborators for continued biological testing. The O-GlcNAcylated peptide antigens VASsRSASM, ALTstAHSV, APVsSKSSL, RPVtASITTM, TPIsQAQKL, IPIsLHTSL, and PTASQARsTI were selected due to their known associations with dysregulated cell signaling pathways. These peptides are currently undergoing ELISpot analyses with HLA-matched individuals, and those demonstrating significant immunogenicity will be subjected to ICS. Moreover, non-O-GlcNAcylated forms of RPPITQSSL and (Me-R)PPITQSSL were synthesized to assure the strong responses for the corresponding O-GlcNAcylated peptides (Table 4.4, Figure 4.6) is due to the sugar moiety. Similarly, O-GalNAcylated versions of IPVsSHNSL, RPPItQSSL, and (Me-R)PPItQSSL were synthesized to verify the responses are O-GlcNAc specific. Together, we hope these studies will show that the CD8+ T-cell responses to the O-GlcNAcylated peptides are directed toward this specific peptide and type of glycosylation. Ultimately, we believe that our O-GlcNAcylated peptide enrichment, identification, and candidate validation process will lead to the development of novel and effective cancer immunotherapeutics.

4.6 References

1. Hanahan D and RA Weinberg. The hallmarks of cancer. *Cell.* 2000; 100(1): 57-70.

2. Hanahan D and RA Weinber. Hallmarks of cancer: the next generation. *Cell.* 2011; 144(5): 646-674.

3. Zarling AL, Polefrone JM, Evans AM, Hunt DF and VH Engelhard. Identification of class I MHC-associated phosphopeptides as targets for cancer immunotherapy. *Proc Natl Acad Sci.* 2006; 103(40): 14889-14894.

4. Klein J and A Sato. The HLA system, first of two parts. *N Engl J Med*. 2000; 343(10): 702-709.

5. Engelhard VH, Brickner AG, and AL Zarling. Insights into antigen processing gained by direct analysis of the naturally processed class I MHC associated peptide repertoire. *Mol Immunol.* 2002; 39(3-4): 127-137.

6. Burmester G, Pezzutto A, Ulrichs T, and A Aicher. *Color Atlas of Immunology*. Thieme, Stuttgart; New York. 2002.

7. Hart GW, Slawson C, Ramirez-Correa G and O Lagerlof. Cross talk between O-GlcNAcylation and phosphorylation: roles in signaling, transcription, and chronic disease. *Annu. Rev. Biochem*. 2011; 80: 825-858.

8. Wang X, Pandey A, and GW Hart. Dynamic interplay between O-linked N-acetylglucosaminylation and glycogen synthase kinase-3-3dependent phosphorylation. *Mol Cell Proteomics.* 2007; 6: 1365-1379.

9. Kawauchi K, Araki K, Tobiume K, and N Tanaka. Loss of p53 enhances catalytic activity of IKKb through O-linked b-N-acetyl glucosamine modification. *Proc Natl Acad Sci.* 2008; 106(9): 3431-3436.

10. Warburg O. On respiratory impairment in cancer cells. *Science.* 1956; 123: 309-314.

11. Lynch TP and MJ Reginato. O-GlcNAc transferase: a sweet new cancer target. *Cell Cycle.* 10(11): 1712-1713.

12. Dang CV and GL Semenza. Oncogenic alterations of metabolism. *Trends Biochem Sci.* 1999; 24(2): 68-72.

13. Morrish F, Isern N, Sadilek M, Jeffrey M, and DM Hockenbery. c-Myc activates multiple metabolic networks to generate substrates for cell-cycle entry. *Oncogene.* 2009; 28(27): 2485-2491.

14. Caldwell SA, Jackson SR, Shahriari KS, Lynch TP, Sethi G, Walker S, Vosseller K, and MJ Reginato. Nutrient sensor O-GlcNAc transferase regulates breast cancer tumorigenesis through targeting of the oncogenic transcription factor FoxM1. *Oncogene.* 2010; 29: 2831-2842.

15. Gu Y, Mi Q, Ge Y, Liu H, Fan Q, Han C, Yang J, Han F, Lu X, and W Yu. GlcNAcylation plays an essential role in breast cancer metastasis. *Cancer Res.* 2010; 70(15): 6344-6351

16. Shi Y, Tomic J, Wen F, Shah S, Bahlo A, Harrison R, Dennis JW, Williams R, Gross BJ, Walker S, Zuccolo J, Deans JP, Hart GW, and DE Spaner. Aberrant O-GlcNAcylation characterizes chronic lymphocytic leukemia. *Leukemia.* 2010; 24: 1588-1598.

17. Zhu Q, Zhou L, Yang Z, Lai M, Xie H, Wu L, Xing C, Zhang F, and S Zheng. O-GlcNAcylation plays a role in tumor recurrence of hepatocellular carcinoma following liver transplantation. *Med Oncol.* 2013; 29: 985-993.

18. Qiao Z, Dang C, Zhou B, Li S, Zhang W, Jiang J, Zhang J, Kong R, and Y Ma. O-linked N-acetylglucosamine transferase (OGT) is overexpressed and promotes O-linked protein glycosylation in esophageal squamous cell carcinoma. *J Biomed Res.* 2012; 26(4): 268-273.

19. Lynch TP, Ferrer CM, Jackson SR, Shahriari KS, Vosseller K, and MJ Reginato. Critical role of O-linked b-N-acetylglucosamine transferase in prostate cancer invasion, angiogenesis, and metastasis. *J Biol Chem.* 2012; 287: 11070-11081.

20. Mi W, Gu Y, Han C, Liu H, Fan Q, Zhang X, Cong Q, and W Yu. O-GlcNAcylation is a novel regulator of lung and color cancer malignancy. *Biochim Biophys Acta.* 2011; 1812: 514-519.

21. De Queiroz RM, Carvalho E, and WB Dias. O-GlcNAcylation: the sweet side of cancer. *Frontiers in Oncology.* 2014; 4(132): 1-10.

22. Wang Z, Udeshi ND, Slawson C, Compton PB, Sakabe K, Cheung WD, Shabanowitz J, Hunt DF, and GW Hart. Extensive crosstalk between O-GlcNAcylation and phosphorylation regulates cytokinesis. *Biochem.* 2010; 3(104): 1-13.

23. Haurum JS, Arsequell G, Lellouch AC, Wong SYC, Dwek RA, McMichael AJ, Elliot T. Recognition of Carbohydrate by Major Histocompatibility Complex Class I-restricted, Glycopeptide-specific Cytotoxic T Lymphocytes. *J. Exp. Med.* 1994; 180: 739-744.

24. Haurum JS, Tan L, Arsequell G, Frodsham P, Lellouch AC, Moss PAH, Dwek RA, McMichael AJ, Elliot T. Peptide anchor residue glycosylation: effect on class I major histocompatibility complex binding and cytotoxic T lymphocyte recognition. *Eur J Immunol.* 1995; 25: 3270-3276.

25. Glithero A, Tomro J, Haurum JS, Arsequell G, Valencia, G, Edwards J, Springer S, Townsend A, Pao YL, Wormald M, Dwek RA, Jones EY, Elliot T. Crystal Structures of Two H-2Db/Glycopeptide Complexes Suggest a Molecular Basis for CTL Cross-Reactivity. *Immunity*. 1999; 10: 63-74.

26. Haurum JS. Recognition of peptides carrying post-translational modifications by class I MHC restricted cytotoxic T cells. 1996. D. Phil. Thesis, University of Oxford. Oxford, United Kingdom.

27. Haurum JS, Hoier IB, Arsequell G, Neisig A, Valencia G, Zeuthen J, Neefjes J, Elliot T. Presentation of Cytosolic Glycosylated Peptides by Human Class I Major Histocompatibility Complex Molecules In Vivo. *J Exp Med.* 1999; 190(1) 145-150.

28. Kastrup IB, Stevanovic S, Arsequell G, Valencia G, Zeuthen J, Rammensee HG, Elliot T, and JS Haurum. Lectin purified human class I MHC-derived peptides: evidence for presentation of glycopeptides *in vivo. Tissue Antigens.* 2000; 56: 129-135.

29. Syka JE, Coon JJ, Schroeder MJ, Shabanowitz J, Hunt DF. Peptide and protein sequence analysis by electron transfer dissociation mass spectrometry. *Proc Natl Acad Sci U S A*. 2004 Jun 29; 101(26): 9528-9533.

30. Yague J, Vazquez J, and JL de Castro. A post-translational modification of nuclear proteins, N^G, N^G-dimethyl-Arg, found in a natural HLA class I peptide ligand. *Protein Science.* 2000; 9: 2210-2217.

31. Cobbold M, De La Pena H, Norris A, Polefrone JM, Qian J, English AM, Cummings KL, Penny S, Turner JE, Cottine J, Abelin JG, Malaker SA, Zarling AL, Huang HW, Goodyear O, Freeman SD, Shabanowitz J, Pratt G, Craddock C, Williams ME, Hunt DF, and VH Engelhard. MHC class I-associated phosphopeptides are the targets of memory-like immunity in leukemia. *Sci Trans Med.* 2013; 5(203): 1-10.

32. Fuchs M, Gerber J, Drapkin R, Sif S, Ikura T, Ogryzko V, Lane WS, Nakatani T, and DM Livingston. The p400 complex is an essential E1A transformation target. *Cell.* 2001; 106: 297-307.

33. Doyon Y, Selleck W, Lane WS, Tan S, and J Cote. Structural and functional conservation of the NuA4 histone acetyltransferase complex from yeast to humans. *Mol Cell Biol.* 2004; 24(5): 1884-1896.

34. Doyon Y and J Cote. The highly conserved and multifunctional NuA4 HAT complex. *Current Opinion in Genetics and Development.* 2004; 12: 147-154.

35. Matre V, Nordgard O, Alm-Kristiansen AH, Ledsaak M, and OS Gabrieldsen. HIPK1 interacts with c-Myb and moedulates its activity through phosphorylation. *Biochem and Biophys Res Comm.* 2009; 288:150-154.

36. Li X, Zhang R, Luo D, Park SJ, Wang Q, Kim Y, and W Min. Tumor necrosis factor alpha-induced desumoylation and cytoplasmic translocation of homeodomain-interacting protein kinase 1 are critical for apoptosis signal-regulating kinase 1-JNK/p38 activation. *J Biol Chem.* 2005; 280(15): 15061-15070.

37. Ecsedy JA, Michaelson JS, and P Leder. Homeodomain-interacting protein kinase 1 modulates Daxx localization, phosphorylation, and transcriptional activity. *Mol Cell Biol.* 2003; 23(3): 950-960.

38. Kondo S, Lu Y, Debbas M, Lin AW, Sarosi I, Itie A, Wakeham A, Tuan J, Saris C, Elliot G, Ma W, Benchimol S, Lowe SW, Mak TW, and SK Thukrai. Characterization of cells and gene-targeted mice deficient for the p53-binding kinase homeodomain-interacting protein kinase 1 (HIPK1). *Proc Natl Acad Sci.* 2003; 100(9): 5431-5436.

39. Rochat-Steiner V, Becker K, Micheau O, Schneider P, Burns K, and J Tschopp. FIST/HIPK3: a Fas/FADD-interacting serine/threonine kinase that induces FADD phosphorylation and inhibits Fas-mediated Jun NH2-terminal kinase activation. *J Exp Med.* 2000; 192(8): 1165-1174.

40. Lan HC, Li HJ, Lin G, Lai PY, and BC Chung. Cyclic AMP stimulates SF-1 dependent CYP11A1 expression through homeodomain-interacting protein kinase 3-mediated jun N-terminal kinase and c-Jun phosphorylation. *Mol Cell Biol.*2007; 27(6): 2027-2036.

41. Curtin JF and TG Cotter. JNK regulates HIPK3 expression and promotes resistance to Fas-mediated apoptosis in DU 145 prostate carcinoma cells. *J Biol Chem.* 2004; 279(17): 17090-17100.

42. Wells L, Vosseller K, and GW Hart. Glycosylation of nucleocytoplasmic proteins: signal transduction and O-GlcNAc. *Science.* 2001; 446: 2376-2378.

Chapter 5: Conclusions and Future Applications

The research presented in this dissertation focuses on enrichment and identification of HLA-associated, tumor-specific, post-translationally modified peptides for the development of novel cancer immunotherapeutics. These peptides are paramount immunotherapeutic candidates for a number of reasons. First, one of the most important hallmarks of cancer is dysregulated cell signaling, which leads to aberrant post-translational modifications that can be displayed by the HLA-class I system. These PTMs will occur more frequently on the surface of tumor cells and we can sequence these tumor-specific peptides by high-resolution mass spectrometry. Further, proteins containing PTMs may undergo antigen processing differently, so that the resulting HLA-peptide may be different from that tolerized during thymic education. This idea is corroborated by the fact that many post-translationally modified peptides have demonstrated an ability to stimulate a T-cell response that is specific for the modification. Without central tolerance, the immune system should mount a substantial response against the modified peptides. Together, this leads us to believe that tumor-specific modified peptides will be excellent candidates for immunotherapy.

Efforts to identify these post-translationally modified peptides began by using iron(III) immobilized metal affinity chromatography to enrich for phosphopeptides presented on the surface of esophageal tumor tissue. Here, we identified 71 phosphopeptides, 68 of which were tumor specific. We carefully selected 37 phosphopeptides and subjected them to ELISpot analysis, which demonstrated their immunological relevance and ability to stimulate a response in healthy donors. Then, after identifying the first fully characterized HLA-associated O-GlcNAcylated peptide,

we sought to develop a novel enrichment technology to characterize more of these peptides. We then succeeded in developing a robust, sensitive, and selective O-GlcNAc enrichment method that achieves quantitative yields of internal standards. By performing phenylboronic acid enrichment under anhydrous conditions and with in-house derivatized beads, we regularly achieve up to 70% recovery of synthetic O-GlcNAcylated peptides when starting with only femtomoles of material. Once optimized with synthetics, we moved to enrichment of tumor-specific HLA-associated O-GlcNAcylated peptides. After adding a C18 clean-up to deplete salt and buffer contaminants and manipulating instrument methods to increase apparent sensitivity, we identified and sequenced a total of 19 peptide antigens in several modification states from 10 types of cancer. As with the esophageal tumor phosphopeptides, we then demonstrated the ability of these O-GlcNAcylated peptide antigens to stimulate an immune response using ELISpot. Additionally, we performed ICS analysis to show that the response resides in the central memory compartment. Thus, we have made significant strides in identifying peptide candidates for the development of novel cancer immunotherapeutics. Efforts in our laboratory will continue to expand on the findings in this dissertation.

The question remains as to how we plan to use these peptide antigens for cancer immunotherapeutics. The ultimate goal is to develop a specific therapy, unlike chemotherapy and radiation, which can target transformed cells before they become medically detectable tumors. A number of methods are being explored for this purpose including: peptide vaccines,[1-3] adoptive T-cell therapy,[4] engineered T-cells,[5] TCR-like biomolecules,[6] and immune-mobilizing monoclonal TCRs against cancer (ImmTACs).[7] Perhaps the most obvious use for the peptides is in peptide vaccines, which serve to

activate CTLs to target and kill cells displaying those peptides. While seemingly straightforward, many peptide vaccines have been successful in generating an immune response, but have failed to convert to clinical responses.[8] This has largely been attributed to interpatient variability in the display of known antigens. To overcome this, multi-peptide vaccines have been developed that encompass several HLA-restrictions.[9] The other challenge in peptide vaccines is eliciting an immune response in the harsh tumor microenvironment. Regulatory T-cells (T_{reg}), which normally serve to modulate immune responses, are found at high concentrations within solid tumors and disrupt CTL responses.[10] This may be surmounted by coupling the peptide vaccine with immunomodulators, which are thought to selectively reduce T_{reg} activity while enhancing NK cell and CTL activities.[1] Thus, a multi-peptide vaccine containing several HLA-restrictions along with immunomodulators may be a good option for specific therapies. However, small peptides are quickly degraded by proteases in the plasma and have very short half-lives, so other options must be considered.[9]

Adoptive T-cell transfer (ACT) has been successful in treating many cancer patients, and may make good use of tumor-specific post-translationally modified peptides. This procedure involves the selection of lymphocytes with antitumor activity from patients followed by *ex vivo* expansion of these cells. The individuals are then immunodepleted before the expanded lymphocytes are infused back into their system.[11,12] This procedure has been performed with durable regression in melanoma and renal cell carcinoma patients, holding great promise within the oncological community.[4] We postulate that our post-translationally modified peptides may have an even better clinical outcome because they are biologically active antigens and many have been found on

multiple types of tumors. Thus, if patients have T-cells targeting our post-translationally modified peptides, it may be possible to select those from the blood using HLA tetramers and selectively expand them with cytokines like IL-2. However, this requires robust multimers which have proved much more difficult to refold than viral peptide tetramers.[13]

An extension of ACT utilizes engineered TCRs and TCR-like biomolecules to target tumor specific antigens. In these cases, the isolated T-cells can be genetically modified to express a transgene encoding a tumor-targeting TCR or chimeric antigen receptor (CAR). The latter receptors bind to tumor antigens via antibody-derived complementarity determining reunions (CDRs), yet are coupled to the intracellular machinery for TCR signaling. This can mediate the functional activation of T-cells even when antigens are not displayed in the context of MHC molecules. In these instances, the choice of tumor antigen for either TCRs or CARs is critical since it dictates the efficacy and the safety of ACT.[5,6] The post-translationally modified peptide antigens we identified could be used as bait for these receptors, because most were truly tumor-specific and thus should not instigate an autoimmune response. This specificity would combat one of the most prominent side effects of ACT, making the treatment an even safer clinical procedure. Overall, this procedure provides a promising approach for the treatment of solid and hematological malignancies, but is far from becoming a routine procedure due to the labor and costs involved.

Finally, the post-translationally modified peptides may be used to develop ImmTACs, engineered by Immunocore (Abingdon, UK). These comprise a tumor-specific TCR covalently linked to a CD3-specific single chain antibody fragment, as demonstrated by Figure 5.1. The ImmTACs engage specifically with tumor cells and

induce responses from all CD3+ T-cells in the local environment.[7] The downside is that

the epitopes need to be entirely tumor-specific. Therefore, we need to assess HLA
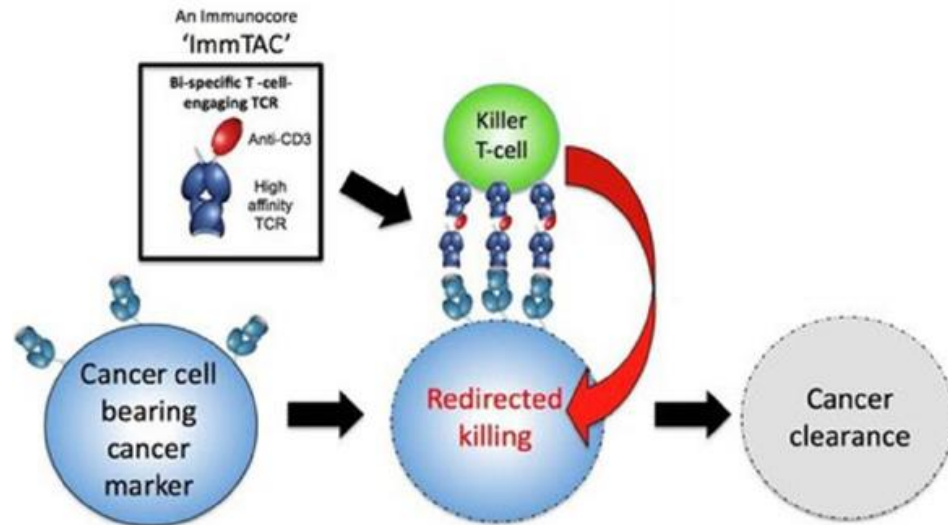

Figure 5.1 Peptide-specific ImmTACs as potential immunotherapeutics. Cancer cells express HLA-associated post-translationally modified peptides on the cell surface. ImmTACs specifically target these complexes with high-affinity TCRs and stimulate all local T-cells with the anti-CD3 scFv, redirecting T-cell killing toward the cancer cells. This stimulates the clearance of cancer.

class I-associated expression of the post-translationally modified peptides in a range of

healthy tissues to assure the peptides are entirely tumor-specific. We have already made

attempts toward this, but only in a few esophageal samples. If we could assure a subset of

peptides fit this criterion, we could use them to treat a range of cancers. To this end, we

are attempting to clone post-translationally modified peptide specific T-cells in order to

sequence their TCRs. Collaboration between Immunocore and our laboratories may

prove to be extremely beneficial so that we may develop ImmTACs directed toward

peptide antigens we have identified.

The application of HLA-associated, tumor-specific, post-translationally modified

peptides for the development in cancer immunotherapies has immense promise due to

those advances outlined above. Our investigation of phosphorylated and O-GlcNAcylated

peptide antigens identified nearly 100 candidates for further analysis. Additionally, we

developed a technique to identify O-GlcNAcylated peptide antigens, which were

previously unknown. We went on to test these peptide antigens with ELISpot and ICS,

which demonstrated the immunogenicity of several peptides and allowed for the selection

of several strong vaccine candidates. While strides have been made toward understanding

how best to harness the immune response to fight cancer, much research remains to be

done in the field. Ultimately, we believe that post-translationally modified peptides are

key to unlocking this potential and can be utilized to develop cancer immunotherapies

that will eradicate this devastating disease.

5.2 References

1. Walter S, Weinschenk T, Stenzl A, Zdrojowy R, Pluzanska A, Szczylik C, Staehler M, Brugger W, Dietrich PY, Mendrzyk R, Hilf N, Schoor O, Fritsche J, Mahr A, Maurer D, Vass V, Trautwein C, Lewandrowski P, Flohr C, Pohla H, Stanczak JJ, Bronte V, Mandruzzato S, Biedermann T, Pawelec G, Derhovanessian E, Yamagishi H, Miki T, Hongo F, Takaha N, Hirakawa K, Tanaka H, Stevanovic S, Frisch J, Mayer-Mokler A, Kirner A, Rammensee HG, Reinhardt C, and H Singh-Jasuja. Multipeptide immune response to cancer vaccine IMA901 after single-dose cyclophosphamide associates with longer patient survival. *Nat Med.* 2012; 18(8): 1254-1261.

2. Filipazzi P, Pilla L, Mariani L, Patuzzo R, Castelli c, carnisaschi C, Maurichi A, Cova A, Rigamonti G, Giardino F, Di Florio A, Asioli M, Frati P, Sovena G, Squarcina P, Maio M, danielli R, Chiarion-Sileni V, Villa A, Lombardo C, Tragni G, Santinami M, Parmiami G, and L Rivoltini. Limited induction of tumor-cross reactive T cells without a measurable clinical benefit in early melanoma patients vaccinate with HLA class I moedified peptides. *Clin Cancer Res.* 2012;

3. Dosset M, Vauchy C, Beziaud L, Adotevi O, and Y Godet. Universal cancer peptide-base therapeutic vaccine breaks tolerance against telomerase and eradicates established tumor. *Clin Cancer Res.* 2012;

4. June CH. Adoptive T cell therapy for cancer in the clinic. *J Clin Invest.* 2007; 117(6): 1466-1476.

5. Porter DL, Levine BL, Kalos M, Bagg A, and CH June. Chimeric antigen receptor-modified T cells in chronic lymphoid leukemia. *N Engl J Med.* 2011; 365(8): 725-733.

6. Weidanz JA, Hildebrand WH. Expanding the targets available to therapeutic antibodies via novel disease-specific markers. *Int Rev Immunol.* 2011; 30(5-6): 312-327.

7. Liddy N, Bossi G, Adams KJ, Lissina A, Mahon TM, Hassan NK, Gavarret J, Bianchi FC, Pumphrey NJ, Ladell K, Gostick E, Deqell AK, Lissin NM, Harwood NE, Molloy PE, Li Y, Cameron BJ, Sami M, Baston EE, Todorov PT, Paston SJ, Dennis Re, Harper JV, Dunn SM, Ashfield R, Johnson A, McGrath Y, Plesa G, June CH, Kalos M, Price DA, Vuidepot A, Williams DD, Sutton DH, and BK Jakobsen. Monoclonal TCR-redirected tumor cell killing. *Nat Med.* 2012; 18(6): 980-987.

8. Rosenberg SA and ME Dudley. Cancer regression in patients with metastatic melanoma after the transfer of autologous antitumor lymphocytes. *Proc Natl Acad Sci.* 2004; 101(2): 14639-14645.

9. Slingluff CL. The present and future of peptide vaccines for cancer: single or multiple, long or short, alone or in combination? *Cancer J.* 2011; 17: 343-350.

10. Berd D and M Mastrangelo. Effect of low dose cyclophosphamide on the immune system of cancer patients: depletion of CD4+, 2H4+ suppressor-inducer T-cells. *Cancer Res.* 1988; 48: 1671-1675.

11. Dudley ME and SA Rosenberg. Adoptive cell transfer therapy. *Semin Oncol.* 2007; 34(6): 524-531.

12. Rosenberg SA, Restifo NP, Yang JC, Morgan RA, and ME Dudley. Adoptive cell transfer: a clinical path to effective cancer immunotherapy. *Nat Rev Cancer.* 2008; 8(4): 299.308.

13. Toebes M, Coccoris M, Bins A, Rodenko B, Gomez R, Nieuwkoop NJ, van de Kasteele W, Rimmelzwaan GF, Haanen JB, Ovaa H, and TN Schumacher. Design and use of conditional MHC class I ligands. *Nat Med.* 2006; 12: 246-251.