DISSERTATION

submitted to the

Combined Faculties for the Natural Sciences and for Mathematics

of the Ruperto-Carola University of Heidelberg, Germany

for the degree of

Doctor of Natural Sciences

presented by

Renata Blatnik, Dipl. Biochemist

born in: Novo mesto, Slovenia

Oral examination: 20.11.2015

Development of mass spectrometry methodology for direct identification of viral epitopes from MHC I molecules

Referees: Prof. Dr. Matthias Mayer

PD. Dr. Dr. Angelika Riemer

The work presented in this thesis was performed from October 2010 to April 2015 under supervision of PD. Dr. Dr. Angelika Riemer in the research group Immunotherapy and -prevention at the German Cancer Research Center (DKFZ), Heidelberg, Germany and during a lab visit from October 2014 to February 2015 at the University of Southern Denmark, Odense, Denmark, in the Department for Biochemistry and Molecular Biology, in Prof. Dr. Martin R. Larsen's subgroup within the Protein Research group.

Publications based on this work:

Dekhtiarenko I, Fischer S, **Blatnik R**, Holzki JK, Bokner L, Marandu TF, Hoppe S, Lisnić B, May T, Lemmermann NAW, Holtappels R, Reddehase MJ, Riemer AB, Cicin-Sain L. C-terminal epitope localization facilitates antigen processing, direct presentation and T-cell memory inflation. *Submitted manuscript*

Conference presentations based on this work:

Blatnik R, Hoppe S, Wühl M, Rösli C, Riemer AB (2015) A mass spectrometry method to identify naturally processed and presented HPV epitopes. Proteomic Forum 2015 (printed abstract and poster)

Blatnik R, Hoppe S, Küpper M, Grabowska AK, Scharf A, Warnken U, Schnölzer M, Ruppert T, Rösli C, Riemer AB (2014) A nano-UPLC-ESI-MS^3 method to directly identify low abundant HPV T cell epitopes on the cell surface. 62nd ASMS conference on mass spectrometry and allied topics (printed abstract and poster)

Blatnik R, Hoppe S, Winter J, Küpper M, Grabowska AK, Warnken U, Ruppert T, Schnölzer M, Rösli C, AB Riemer (2013) A mass spectrometry approach for low abundant epitope identification – applied for detection of HPV epitopes for therapeutic vaccine design. Annual meeting of German society for immunology (DGfI) (printed abstract and poster)

Blatnik R, Winter J, Warnken U, Hoppe S, Grabowska AK, Link S, Wühl M, Ruppert T, Schnölzer M, Riemer AB (2012) HPV16 E6 and E7 T cell epitope identification by mass spectrometry. 10^{th} annual Cancer immunotherapy meeting (printed abstract and poster)

Blatnik R, Hoppe S, Wühl M, Grabowska AK, Riemer AB (2012) A direct mass spectrometry approach for HPV T cell epitope identification. American association for cancer research (AACR) annual meeting (printed abstract and poster)

Acknowledgements

I would like to express my deepest gratitude to everyone who contributed to this work, be it by helping me in the lab, through scientific discussions or personal support.
First of all, I would like to thank **PD. Dr. Dr. Angelika Riemer,** for giving me the chance to come to Heidelberg and start working with her. Her support, thrust and belief, throughout my whole PhD period, were of immense value.
Special thanks goes to TAC members: **Prof. Dr. Ralf Bartenschlager**, **Dr. Martina Schnölzer, Dr. Uwe Warnken** and **Dr. Christoph Rösli** for their valuable advice during our advisory meetings. I would like to thank **Prof. Dr. Matthias Mayer,** for extremely productive discussion and being my PhD examiner. I would also like to thank **PD Dr. Suat Ötzbeck** and **Dr. Steeve Boulant** for being my PhD examiners.

I would like to extend my appreciation to all past and present F130 members; **Dr. Stephanie Hoppe**, **Dr. Agnieszka Grabowska, Dr. Hadeel Khallouf, Alina Steinbach, Martin Wühl, Alexandra Klevenz** and **Sebastian Kruse**. Thank you for contributing to a positive work environment and providing support during my PhD period. Thank you **Steffi,** for helping me out every time I needed a second hand; thank you for the abundant productive scientific discussions, for all the support and fun moments we shared. It was a pure joy to have you as a PhD twin-sister. Thank you for correcting my thesis, contributing valuable advice and also for translating the Abstract into German. **Martin,** thank you for the great maintenance of the cell culture and uncountable IP samples, without your help this thesis would take way longer. **Agi,** thank you for your help, your scientific support, the positive attitude and a warm welcome in Stuttgart, when we visited Wasen. I would like to thank **Hadeel,** for being a pleasant and cheerful desk neighbor. Thank you for all of your support and shared fun moments, be it at work or privately. And thank you for the best humus. Thank you **Alina** and **Alex** for jumping in to help me with the cell culture and performing FACS stainings for me. Thank you for the best home-made cheesecake. **Alina,** thank you also for all productive scientific discussions.
A big thank you goes to all students, who contributed directly or indirectly to my project; **Marc Pfeil, Aakanksha Jain, Victoria Hentschke, David Kranz, Fanny Georgi, Kinga Grabowska, Sina Knapp, Jasmin Mangei, Miriam Bertrant, Sebastian Utz, Julia Schessner, Lisa Dreßler, Johanna Dickmann**. Special thanks to **Jan Winter, Johannes Stortz** and **Marius Küpper,** for their numerous binding assays.

I would like to thank the DKFZ and ZMBH Proteomics core facility members who always warmly welcomed me and helped me with the mass spec analysis and all of the troubleshooting. Special thanks to **Dr. Uwe Warnken**, **Dr. Thomas Ruppert** and **Dr. Martina Schnölzer,** for introducing me to the proteomics mass spec and for the fruitful scientific discussions. Thank you, **Sebastian Link** for the first IP analysis and for introducing me to the QTrap5500 instrument. I would like to thank **Dr. Annette Scharf** for great advice and scientific contributions to my project, and **Lilyana Behrens,** for helping me with the maintenance of the QTrap5500. Thank you **Dr. Rober Hardt** and **Dr. Bernd Heißling,** for letting me use your computers for the MaxQuant analysis and for the advice on how to perform them better. Thank you **Patrick Winterhalter,** for great scientific debates and making my stay at the ZMBH fun.

My sincere appreciation is extended to **Dr. Christoph Rösli,** for welcoming me to his group. Christoph, thank you for your scientific support and for helping with the QTrap6500 troubleshooting. Thanks to all other members of Rösli's group, especially: **Dr. Sabrina Hanke, Dr. Laura Kuhlmann** and **Wiebke Nadler,** for their valuable scientific advice and fun moments in the lab and at conferences.

I would like to express my gratitude to **Prof. Dr. Martin R. Larsen,** for welcoming me in his group at the University of Southern Denmark, Odense, Denmark, from October 2014 to February 2015 and for generously sharing his knowledge. Martin, without your help this thesis would be very different. Special thanks to **Dr. Lasse G. Falkenby,** for technical and experimental support during my lab visit in Denmark. I would especially like to thank you for all troubleshooting, when my samples caused the column clogging. Thank you **Ass. Prof. Dr. Giuseppe Palmisano,** for answering my chemical questions. I would like to thank all the other members of the Protein Research Group at the University of Southern Denmark for making my stay in Odense a memorable life experience.

I am also extremely grateful to the collaborators; **Prof. Dr. Čičin-Šain's group** at the Helmholtz Centre for Infection Research (Braunschweig, Germany), especially **Dr. Iryna Dekhtiarenko** for providing me with mCMV IP samples, and **Dr. LeGall's group** at the Ragon Institute of MGH, MIT and Harvard, (Cambridge, MA, USA), especially to **Dr. Marijana Ručević,** for providing HIV IP samples, extremely fruitful discussions and most valuable advice.

I would like to thank our **neighboring groups: Hoppe Seylers** and **de Villiers** for their support and help. A special thanks goes to **Krammers,** who warmly welcomed us to their journal club.

Thank you **Angelika Schmidt-Zitouni** and **Monika Bock** for dealing with the German bureaucracy for me.

I would like to express my gratitude to everyone who helped proofreading this work; **Dr. Stephanie Hoppe, Dr. Stefan J. Kempf, Dr. Marijana Ručević, Dr. Iryna Dekhtiarenko, Dr. Laura Kuhlmann and Marijana E. Stanko.**

Thank you **Roya and Kevin Moore** from Odense, for being so open, welcoming and friendly. Roya, you were the best landlady I could wish for and staying with you was truly enjoyable.

Thanks go to all my Heidelberg friends, especially: **Jana, Alex, Regina, Ioana, Anagha, Jason, Vinay, Paola, Zhiqi, Naveen, Frank, Lav, David and Taylor**. You were truly a great company and support. Thank you for keep reminding me that there is a life beyond long hours in the lab. I will keep wonderful memories of things we did together.

Anja K., thank you for introducing me to HPV, when I was still in Ljubljana, and your support during my PhD time in Heidelberg.

Marijana E. S., thank you for being my best friend. Thank you for being there for me, despite the distance of 1000 km. Thank you for listening, encouraging and supporting me. It would have been very different without you.

I would like to say a big thank you to my **family,** for their unconditional love, support and understanding. The next two lines for my family are in Slovenian: Hvala mami in oci, za brezpogojno ljubezen in podporo. Hvala, da sta mi vedno stala ob strani. Hvala tudi vsem ostalim sorodnikom za prijeten sprejem, ko sem bila doma.

List of abbreviations

2D	two dimentsional
3D	three dimensional
A	ampere
Å	angstrom
aa	amino acid
Ab	antibody
ACN	acetonitrile
AF2	excitation energy
AIDS	acquired immune deficiency syndrome
APC	antigen presenting cell
APM	antigen processing machinery
a.u.	arbitrary units
β_2M	beta 2-microglobulin
°C	degrees Celsius
C-terminus	carboxyl terminus
CAD	collision gas
CE	collision energy
CESI	capillary electrophoresis-electrospray ionization
CID	collision induced dissociation
CMV	cytomegalovirus
cps	counts per second
CTL	cytotoxic T lymphocytes
Cys	cystein, cystin
Da	Dalton
DC	dendritic cell (in immune system chapter) or direct current (in MS chapter)
DMSO	dimethyl sulfoxide
DNA	deoxyribonucleic acid
DP	declustering potential
DTT	dithiothreitol
e.g.	for example
ELISA	enzyme-linked immunosorbent assay
ESI	electron spray ionization
ER	endoplasmatic reticulum
ERAP	endoplasmatic reticulum aminopeptidase
F	phenylalanine
FA	formic acid
FACS	fluorescence-activated cell sorting
FDA	Food and Drug Administration (USA)
fmol	femtomol
G3P	glyceraldehyde 3-phosphate
h	hour(s)
hCMV	human cytomegalovirus
HILIC	hydrophilic interaction liquid chromatography
HIV	human immunodeficiency virus
HLA I	human leukocyte antigene
HPV	human papillomavirus
HRAM	high – resolution accurate mass
HSV-1	herpes simplex virus type 1
kDa	kilo Dalton
IAA	iodoacetamide
IEF	isoelectric focusing
Ig	immunoglobulin
IPG	immobilized pH gradient
IP	immunoprecipitation
L	liter
LC	liquid chromatography
LIT	linear ion trap

LOB	limit of blank
LOD	limit of detection
LSEC	liver sinusoidal endothelial cells
Lys	lysine
M	molar
mCMV	murine cytomegalovirus
mM	milimolar
MALDI	matrix-assisted laser desorption/ionization
MeOH	methanol
MHC	major histocompatibility complex
min	minute
MRM	multiple reaction monitoring
MS	mass spectrometry
m/z	mass-to-charge ratio
mg	milligram
mL	milliliter
µg	microgram
µL	microliter
N-terminus	amino terminus
NK	natural killer
nL	nanoliter
NP	normal phase
PBS	phosphate buffered saline
pH	pondus hydrogenii
RNA	deoxyribonucleic acid
PRM	parallel reaction monitoring
psi	pounds per square inch
PTM	posttranslational modification
PVDF	polyvinylidene fluoride
RF	radio-frequency
RP	reverse phase
RT	room temperature
Q1, Q2, Q3	first, second, third quadrupole
SD	standard deviation
SDS-PAGE	sodium dodecyl sulfate polyacrylamide gel electrophoresis
speLC	solid-phase extraction capillary liquid chromatography
SRM	selected reaction monitoring
TAP	transporter associated with antigen processing
TCR	T-cell receptor
TEAB	triethylammonium bicarbonate buffer
TFA	trifluoroacetic acid
TiO_2	titanium dioxide, titanium beads
TOF	time-of-flight
UPLC	ultra-high performance liquid chromatography
V	volt
v	volume
WB	western blot
WHO	World Health Organization
wt	wild type
ZIC HILIC	zwitterionic hydrophilic liquid chromatography

Table of Contents

List of Figures

Abstract

Persistent infections with viruses such as cytomegalovirus (CMV), human immunodeficiency virus (HIV) or human papillomavirus (HPV) can lead to serious illnesses or cancer development. There are no effective therapies available to permanently eliminate these infections and cure caused diseases. With advances in understanding viral biology and biology of immune responses, one could design therapies by which the immune system is manipulated in order to eradicate the virus or the virus-induced illness. To do so, one potential approach is direct identification of viral antigen-derived epitopes, which are presented on the surface of infected or diseased cells for immune recognition. These epitopes are, in most cases, of low abundance and therefore difficult to identify.

This work presents the development of a targeted highly specific liquid chromatography-mass spectrometry (LC-MS) methodology for detection of low abundant viral epitopes from the surface of infected cells. It also offers a solution for detection of epitopes from other complicated experimental set-ups, such as identification of low abundant tumor mutation-derived epitopes.

The methodology was developed first for the detection of human leukocyte antigen (HLA)-A2-restricted HPV16 E6 and E7 epitopes, and then applied to identify HIV-derived epitopes, and mouse (m)CMV-derived epitopes presented by the mouse major histocompatibility (MHC) I complex H-2D^b. The work describes the optimization of isolation, purification and enrichment of T cell epitopes for MS detection. First, HLA I-epitope complexes were immunopurified and treated with acid for dissociation of complexes. Next, epitope-containing eluates were subjected to various enrichment, purification and fractionation strategies, including ultrafiltration, normal and reverse phase chromatography, and a newly established chemical tagging strategy for epitope isolation by TiO_2 pull down. Finally, epitopes were analyzed with a targeted highly specific and sensitive nano-LC-MS^3 approach, where every measured peptide was manually optimized to generate the best possible spectrum.

The HPV16 E7$_{11-19}$ YMLDLQPET peptide was reported to be presented on HPV16-postive cell lines and tumor samples before. We were not able to identify it on the surface of HPV16-transformed cells. However, the H-2D^b-restricted mCMV epitope was successfully detected in high abundance on the surface of only 1x10^7 cells, which is the lowest cell number ever reported for an experiment like this. The cell number could even be further reduced. Moreover, three low abundant HLA-A2-restricted HIV-derived epitopes were successfully detected on the surface of HIV-transfected cells. One of them is the first directly identified Nef-derived epitope ever reported.

In conclusion, this work demonstrates that the developed strategy for direct identification of virus-derived epitopes on the cell surface is broadly applicable to various MHC I types and virus-infected target cells. The methodology can be extended to direct identification of low abundant tumor mutation-derived epitopes. In general, directly identified epitopes form a solid base of future immunotherapy design.

Zusammenfassung

Persistierende virale Infektionen, wie zum Beispiel durch das Cytomegalievirus (CMV), das humane Immundefizienz-Virus (HIV) oder das humane Papillomvirus (HPV), können zu schwerwiegenden Krankheiten oder Krebs führen. Bis heute gibt es keine effektiven Therapien, die diese Infektionen permanent beseitigen oder diese Krankheiten heilen können. Durch Fortschritte im Verständnis der Virus- und Immunbiologie können nun Strategien entwickelt werden, die das Immunsystem gezielt manipulieren, um das Virus oder die durch das Virus hervorgerufene Krankheit zu eliminieren. Ein möglicher Ansatz ist die direkte Identifizierung von Epitopen, die von viralen Antigenen stammen und dem Immunsystem auf der Zelloberfläche von infizierten Zellen präsentiert werden. Diese Epitope kommen in vielen Fällen nur in geringer Menge vor, und sind aus diesem Grund schwer zu detektieren.

In dieser Arbeit wurde eine hochspezifische massenspektrometrische Methodik (engl. *targeted liquid chromatography-mass spectrometry*, LC-MS) zur Detektion von viralen Epitopen entwickelt, die nur in geringer Menge auf der Zelloberfläche von infizierten Zellen vorhanden sind. Diese Methodik ermöglicht zudem die Detektion von anderen Epitopen, die in ähnlich komplexen experimentalen Ansätzen in geringer Menge präsentiert werden, wie zum Beispiel Epitope, die in Tumorzellen durch Tumor-spezifische Mutationen entstehen.

Die Methodik wurde zunächst für den Nachweis von Epitopen, die von den HPV-Proteinen E6 und E7 stammen und durch das humane Leukozytenantigen (HLA)-A2 restringiert sind, entwickelt. In der Folge wurde die Technik auf die Identifizierung von Maus-CMV- und HIV-abgeleiteten Epitopen umgelegt. Diese Arbeit beschreibt die Optimierung der Isolation, Aufreinigung und Anreicherung von T-Zell-Epitopen zur MS Detektion. Zuerst wurden HLA Klasse I (HLA I)-Epitop-Komplexe immunpräzipitiert und zur Dissoziierung des Komplexes mit Säure behandelt. Danach wurden die Eluate, die die Epitope enthalten, verschiedenen Strategien zur Anreicherung, Aufreinigung und Fraktionierung unterzogen. Diese umfassten Ultrafiltration, Normalphasen- und Umkehrphasen-Chromatographie, isoelektrische Fokussierung, sowie eine neu etablierte Strategie, welche auf chemischer Markierung und anschließender TiO_2-Fällung von zu isolierenden Epitopen basiert. Zuletzt wurden die Epitope mit einem hochspezifischen und empfindlichen *targeted nano-LC-MS*^{3}-Ansatz analysiert, wobei jedes gemessene Epitop-Peptid manuell optimiert wurde um das bestmögliche Spektrum zu generieren.

Das HPV16 $E7_{11-19}$-Peptid YMLDLQPET wurde in vorangegangen Studien auf der Zelloberfläche von HPV16-positiven Zellen und in Tumorproben entdeckt. In dieser Studie gelang es nicht, dieses Peptid auf HPV16-transformierten Zellen zu identifizieren. Das Maus-CMV Epitop, welches vom murinen Haupthistokompatibilitätskomplex (engl. *major histocompatibility complex*, MHC) H-2D^b präsentiert wird, konnte jedoch erfolgreich in großer Menge auf der Zelloberfläsche von nur $1x10^7$ Zellen detektiert werden. Dieses ist die geringste Zellzahl, die nach unserem Wissen jemals für ein vergleichbares Experiment benötigt wurde. Die Zellzahl könnte sogar noch weiter reduziert werden. Darüberhinaus wurden drei HIV-Epitope, die in geringer Menge von HLA-A2 auf der Zelloberfläche

von HIV-transfizierten Zellen präsentiert werden, erfolgreich detektiert. Eines davon ist das erste direkt identifizierte Epitop, das vom Nef-Protein abgeleitet ist.

Zusammenfassend zeigt diese Arbeit, dass die entwickelte Strategie zur direkten Identifizierung von viralen Epitopen auf Zelloberflächen breit anwendbar ist. Sie kann für die Epitop-Identifizierung auf verschiedenen MHC I Typen und in verschiedenen Virus-infizierten Zellen genutzt werden. Die Methodik kann zudem erweitert werden, um durch Tumor-spezifische Mutationen entstandene Neo-Epitope direkt zu identifizieren. Generell sind direkt identifizierte Epitope eine solide Basis für zukünftige Ansätze in der Immuntherapie.

1. Introduction

1.1. Viruses

Viruses are infectious agents which can only survive in a living host cell or organism. They were first described as a new type of infectious agent by the Dutch microbiologist Martinus Beijerinck in 1898 (1). Since then, numerous viruses have been characterized and it has been estimated that 10^{31} viruses exist on Earth, most of them being phages which infect bacteria (2).

The main component of a virus structure is genetic material packed in a protein capsid. Some viruses are additionally covered with a lipid bilayer (envelope). A common trait of all viruses is that they use the host cell machinery for their survival and replication. Thereby, they can kill the host or they stay in a so called latent phase where their host organism or cell shows only few or no signs of infection. Such viruses can persist in such states for months or even years (1).

Animal viruses can be divided in seven groups based on their genomic material; single and double stranded DNA viruses, double stranded RNA viruses, positive or negative single stranded RNA viruses, single stranded RNA retroviruses and double stranded DNA retroviruses (1).

Some examples of common virus-induced illnesses in humans are the flu or the seasonal cold, which are easily eliminated in a healthy individual. However, viruses causing persisting infection are not easily cleared. Furthermore, there is no effective cure for infection with many of these viruses. In some cases, persistent infection leads to cancer development. Some examples of cancer-inducing viruses are the human papillomavirus (HPV) causing cervical cancer (3), human hepatitis B and C viruses (HBV and HCV) causing hepatocellular carcinoma (4), human herpes virus 8 (HHV8) causing Kaposi's sarcoma (5) or Epstein-Barr virus (EBV) that causes different kinds of lymphomas and nasopharyngeal carcinoma (6).

There is also no cure but only limited therapeutic solutions to manage the infection with viruses causing some other serious illnesses, such as the human immunodeficiency virus (HIV) causing acquired immune deficiency syndrome (AIDS), the Ebola virus causing Ebola disease, or the SARS coronavirus causing severe acute respiratory syndrome (SARS) (1).

To prevent the infection with a virus, prophylactic vaccines have been a great leap forward in fighting viral illnesses. They have been successfully protecting against measles, mumps, rubella, polio, hepatitis A and B, rabies, smallpox, influenza and others. These vaccines induce an immune response mainly leading to antibody production. These so-called neutralizing antibodies recognize and eliminate viruses when they enter an individual, before target cells can be infected. However, most prophylactic vaccines have no therapeutic effects (7).

The most common therapeutic solutions for virus-induced illnesses are antiviral drugs. They have different mechanisms of action and are designed to help combating infections caused by a particular virus. Antiviral drugs target different steps in the virus life cycle, such as transcription of viral genetic material, viral protein translation and processing, or virus assembly. Despite specific targeting of viral biology, these drugs in most cases do not eliminate the virus after infection, but only limit its spreading and associated symptoms (1). Therefore, there is a great need to find better treatment options for viral diseases.

In the past decades, the field of immunology advanced considerably in understanding biological mechanisms in greater detail. It was shown that T cells are important to clear persistent viral infections (8). Therefore, today, modifying or precise targeting of immune responses can be a solution to combat persistent viral infections (9).

1.2. Short introduction to viruses used in this thesis

1.2.1. Human papillomavirus (HPV)

HPV is a commonly sexually transmitted virus and up to 95% of the sexually active population is or will be exposed to it during their lifetime (10, 11). Persistent infection with some viral HPV types can lead to cancer development (12), therefore they are grouped in low-risk or high-risk types. Low-risk types can cause warts, whereas persistent infection with high-risk types can lead to cervical cancer, which is the second most prevalent cancer in women. Furthermore, high-risk HPV types also cause other ano-genital cancers and cancer in the head and neck region (13-16).

HPV is a double stranded circular DNA virus (~ 8000 bp) and codes for eight proteins. The early proteins E1, E2, E4, E5, E6 and E7 are expressed throughout the course of infection, whereas the late proteins L1 and L2 are coding for the viral capsid proteins and are expressed only in the latest phase of the viral life cycle (Figure 1) (17, 18).

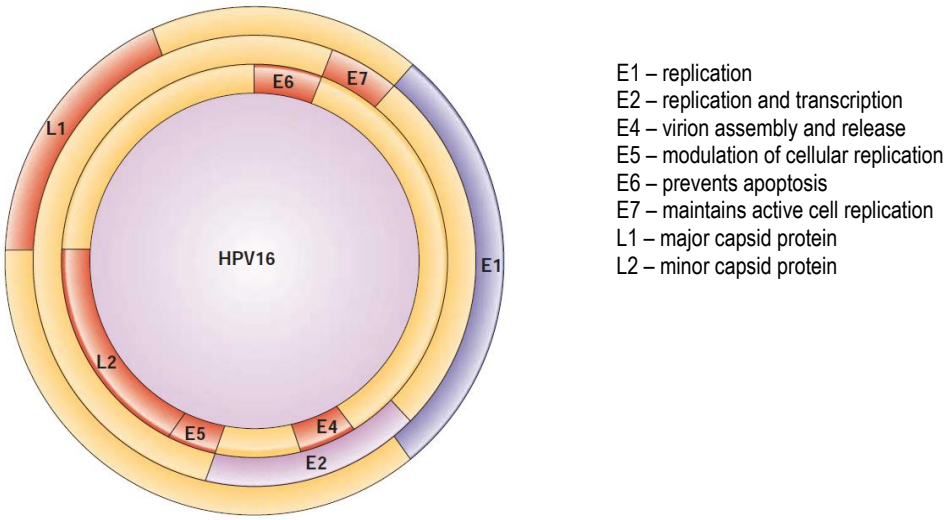


Figure 1. Schematic presentation of the HPV genome.
Genes are grouped in early (E) and late (L) genes. E1 and E2 are involved in DNA replication, E4 is involved in virion assembly and E5 in modulation of replication. The E6 and E7 oncoproteins are responsible for host cell immortalization. The late genes L1 and L2 code for the capsid proteins. Illustration adapted from (19).

The viral life cycle starts with infection of the host organism. The viral L1 capsid protein binds to the basement membrane under a previously (mechanically) damaged epithelial layer and the virus can be internalized into the basal keratinocytes via a so far uncharacterized endocytic pathway. The viral life cycle is tightly connected with the differentiation process of epithelial cells (17, 18). Its DNA is transferred to the nucleus, where RNA transcription is initiated (20, 21). First, E1 and E2 proteins that

work as replication factors are expressed. In the first stages of cell proliferation, the E5 protein, which acts as a modulator of cellular replication, is expressed as well (17, 18). Next, the E6 and E7 proteins are translated. They promote malignant transformation in cells infected with high-risk HPV. They are present during the whole viral life cycle. The E6 protein degrades the tumor suppressor protein p53, which regulates expression of proteins involved in control of the cell cycle. The E7 protein degrades the retinoblastoma protein (pRB) family and associates with cyclins to prevent cell cycle regulation. This leads to the accumulation of mutations in the host cell genome, genome instability, malignant transformation and eventually cancer development (3, 17).
In the last stages of the viral life cycle, E4, and the major (L1) and minor (L2) capsid proteins are expressed. Capsid proteins are packed together with the viral DNA with the help of the E4 protein to form new virions, which are released afterwards by normal cell shedding (17, 18, 22).

The chance to get infected with a high-risk HPV type and consequently develop cancer was reduced after prophylactic vaccines were available for immunization of adolescents (19, 23-25). They are virus-like particle based vaccines, which induce the production of neutralizing antibodies and ensure protection for at least 8.5 years with the trials still on-going (26). However, these vaccines have no therapeutic effects in already infected individuals (27). Due to socio-economic reasons they are not accessible to everyone, especially in developing countries. Therefore, therapeutic solutions are needed.
As it was shown that T cell responses are important to clear HPV infections (28), therapeutic vaccines targeting the oncoproteins E6 and E7, which are expressed throughout the whole viral life cycle, seems a reasonable approach. Vaccines based on viral proteins or DNA did not give promising results (29). However, vaccination with long synthetic peptides derived from E6 and E7 proteins gave encouraging results, however only in premalignant stages (30-32). For rational therapeutic vaccine design, it is important to identify which viral peptides are truly presented on the surface of the transformed cells for immune recognition. This question is addressed, amongst others, in this thesis.

1.2.2. Cytomegalovirus (CMV)

Human CMV is a highly prevalent and globally distributed virus, with up to 90% of the population being latently infected. The virus can infect nearly any cell type and it spreads through saliva, urine, semen, cervical secretions, blood transfusions or organ transplantation. Furthermore, it penetrates the placenta to infect embryos and fetuses. The infection in healthy immunocompetent individual is usually asymptomatic or causes a mild mononucleosis-like syndrome. After primary infection and its successful resolving, the virus enters a latency state. However, infection of newborns and immunocompromised individuals, such as allograft recipients or HIV infected individuals, can cause serious illnesses and health complications. The most common effects in CMV-infected newborns are hepatitis, anemia, growth retardation, brain damage and hearing loss. In immunocompromised adults, the most common illnesses and complications are hepatitis, colitis, organ dysfunction in allograft recipients or death (33).

The virus is species-specific, which leads to limitations in *in vivo* studies of human (h) CMV. Therefore, murine (m) CMV in mice is often used as an animal model. mCMV has similar structural and genetic characteristics as well as similar adaptation strategies to the host immune system as hCMV (34, 35).

CMV is a herpes virus. It has a linear double stranded DNA composed of app. 230 kbp (36) and it contains 165 to 252 open reading frames (37, 38). Based on the timing of the expression, proteins can be divided in intermediate early (IE), early (E) and late (L) proteins (39, 40). IE protein expression starts directly after viral DNA enters the host cell nucleus. IE proteins are responsible for expression of all later viral genes and for modulation of host cell functions (41, 42). E proteins are expressed 4 h after infection. They are responsible for host immune system evasion, regulation of viral DNA replication and repair. Late genes start to transcribe 24 h post infection and encode for structural components and virus assembly (36, 39).

Antiviral therapy with ganciclovir or valganciclovir reduces viral load and prevents development of CMV disease in immunocompromised adults. However, there is no effective cure to eliminate the virus (43). One of the suggested therapeutic solutions to eliminate the virus is a recombinant CMV vaccine expressing dominant viral peptides, which are primarily presented to the immune system on the surface of naturally CMV infected cells (44-46). To this end, the dominant peptides need to be identified and viruses engineered such that these peptides will be indeed presented by the immune system upon the vaccination (45, 46). This can be achieved by direct identification of CMV-derived epitopes from cells infected with different CMV mutants by mass spectrometry (MS) analysis, as it was demonstrated in this thesis.

1.2.3. Human immunodeficiency virus (HIV)

HIV is the leading killing infectious agent in the world. According to the World Health Organization (WHO), there were approximately 35 million people worldwide living with HIV/AIDS in 2013. The number of new infections in 2013 was estimated to be 2.1 million individuals worldwide (47). The virus primarily infects CD4+ T cells and macrophages (1, 48).

HIV is a small retrovirus, enclosing reverse transcriptase, protease, integrase, ribonuclease (RNase H), and two copies of positive single-stranded RNA. Its viral proteins can be divided in three groups based on their function; 1) structural proteins and viral enzymes, 2) regulatory proteins and 3) accessory and auxiliary proteins. The viral envelope is composed of glycoproteins (gp)120 and gp 41, which are responsible for viral entry. Other structural proteins are matrix protein p17, capsid protein p24 and nucleocapsid proteins p7 and p6, which protect RNA from degradation. The reverse transcriptase transcribes viral RNA to DNA, whereas the ribonuclease (RNase H) degrades the viral RNA template after reverse transcription for complete synthesis of viral double stranded DNA in the host cell. The integrase helps integrating viral DNA into the host genome after it has been transcribed from single stranded RNA to double stranded DNA. The trans-activator of transcription (Tat) and regulator of expression of virion (Rev) proteins are regulatory proteins responsible for viral gene

expression. Tat modulates HIV gene transcription through binding to cellular transcription factors, whereas Rev mediates the export of unspliced and partially spliced mRNAs from the nucleus. After transcription and translation of the viral genome, proteases are responsible to process viral polyproteins to functional proteins. Furthermore, the HIV genome codes for several accessory and auxiliary proteins, which are involved in various processes during the viral life cycle. The viral infectivity factor (Vif) directs cell proteins responsible for antiviral mechanisms into degradation, thereby promoting viral infectivity. The viral proteins R and X (Vpr, Vpx) are involved in trafficking of the viral genome to the nucleus for integration into the host genome. Moreover, the Vpr and Vpx proteins are involved in cell cycle arrest and induction of apoptosis in proliferating cells. The viral protein unique (Vpu) is involved in two different processes; first, it degrades CD4 molecules in the endoplasmatic reticulum and second, it enhances virion release. The negative factor (Nef) has several functions, namely downregulation of CD4 expression, modulation of the activation of T cell proliferation and suppression of apoptosis (1, 48-50).

The acute phase of HIV infection starts 2-3 weeks after infection with a high virus burst. Individuals experience flu-like symptoms which resolve spontaneously. Subsequently, the viral load in the blood is reduced to a steady state, which is maintained by the equilibrium between virus production and clearance. The steady state can persist for months or years. HIV is characterized by high replication rates and high error rates of reverse transcriptase, resulting in numerous new viral variants generated in individuals and across infected populations per day. This viral diversity hinders effective immune recognition by T cells and can result in CD4+ T cell loss. Moreover, high viral diversity renders antiviral therapy and vaccine development a challenging task (1, 49).

Numerous antiviral drugs targeting different steps in the HIV life cycle were developed. Standard therapies use a combination of several drugs, as mono- or dual-therapies induced drug resistance due to high viral mutation rates. However, these drugs only reduce the viral load in the blood, but they do not eliminate the virus completely (1, 49). Therefore, other therapeutic solutions are needed.
CD4+ and CD8+ T cell populations are activated for viral clearance in infected individuals (49, 51). Furthermore, vaccines designed based on peptides presented on the surface of infected cells to the immune system, induced responses against HIV. This shows that identifying conserved HIV-derived peptides presented to immune cells could contribute to therapeutic vaccine development (52).

1.3. The immune system

The immune system is a complex system, which recognizes danger signals originating from pathogens, such as bacteria or viruses, from damaged tissue, or from cancer cells. For the purpose of finding these danger signals, highly specialized immune system cells circulate through blood and tissues for immune surveillance and to trigger immune responses. Cells involved in these processes are B and T lymphocytes, dendritic cells (DCs), natural killer (NK) cells, macrophages, monocytes, neutrophils and granulocytes, which are all produced and maturated in the lymphoid organs, such as bone marrow, spleen, thymus, lymph nodes and tonsils (53).

1.3.1. The innate and adaptive immune system

Immune responses can be grouped into innate or adaptive immune responses, which are both closely connected through signaling and cell interactions. However, they differ in response mechanisms and kinetics.

The innate response is the fastest response to pathogens and is thus critical in the first hours or days of exposure to a new pathogen. The innate system is not specific for a particular pathogen and it does not provide long-lasting immunity. The first defense line are physical body barriers in the form of various epithelia. When pathogens enter a tissue, macrophages, monocytes, DCs or neutrophils remove them by pino- and phagocytosis. The trigger for pino- and phagocytosis is recognition of molecular patterns conserved across broad groups of pathogens. These so called pathogen-associated molecular patterns (PAMPs) are, for example, bacterial lipids (LPS-lipopolysaccharides), peptidoglycanes and mannose-rich oligosaccharides, flagellin and viral RNA. They are recognized by pattern recognition receptors, such as Toll-like receptors, which initiate signaling cascades for inflammatory cytokine production.
Professional antigen presenting cells (APCs), such as DCs can present fragments of pathogen antigens on the cell surface for interaction and priming of naïve T cells. Hence, they are a link between innate and adaptive immunity (53).

The innate response against pathogens triggers adaptive immunity. The key cells of the adaptive immune system are B and T lymphocytes. Each cell has a different specificity for antigens due to unique receptors on the cell surface developed during the lymphocyte maturation process. After being exposed to an antigen presented by a professional APC, naïve lymphocytes undergo clonal proliferation and differentiation.
B cells differentiate into effector cells called plasma cells. These are producing antibodies with the same binding properties as the B cell receptor that recognized the particular pathogen. After a pathogen is removed, most specific B cells are eliminated. However, a small pool of B cells differentiates into memory cells. Hence, the most important B cell functions are antibody production and memory B cell development.
Naïve T cells differentiate into one of the subsets of effector T cells. These are either cytotoxic T lymphocytes (CTLs), with the characteristic co-receptor CD8, recognizing peptides presented by major

histocompatibility complex I (MHC I) on nucleated, infected cells. The other subset of effector T cells are helper T cells, with the characteristic co-receptor CD4, recognizing MHC II-peptide complexes on the surface of antigen presenting cells, such as DCs. As with B cells after pathogen clearance, most effector T cells die. A small set of cells, however, remains in the body after pathogen elimination and forms a pool of so called T memory cells. These cells provide a faster immune response upon a second encounter with the same pathogen. The most important T cell functions are thus screening and killing of infected cells, interaction with innate immune cells and memory T cell development (53).
It was shown that CTLs are important for (spontaneous) virus and tumor eradication (8, 54, 55). As this work describes method development for direct detection of MHC I restricted peptides, only immunobiology connected to this thesis is described in more detail in the next sections.

1.3.2. Cytotoxic T lymphocytes (CTL)

CTLs originate from common lymphoid hematopoietic stem cells, which leave the bone marrow for the thymus for maturation. T cell precursor cells undergo T-cell receptor (TCR) gene rearrangement resulting in an α:β receptor, which consists of an α and β chain as the name suggests. The gene rearrangement is highly variable and is, thus, producing up to 10^{18} different TCRs. At this stage, cells express both CD8 and CD4 co-receptors simultaneously, and are called double positive thymocytes. Afterwards, cells are positively selected by recognizing self-peptide:MHC complexes. Concurrently, either CD8 or CD4 expression is terminated, resulting in single positive thymocytes. These cells undergo negative selection, which removes cells that react to self-antigen too strongly. Mature single positive T cells, which are MHC restricted and self-tolerant, migrate to the periphery for antigen recognition as naïve T cells (53).

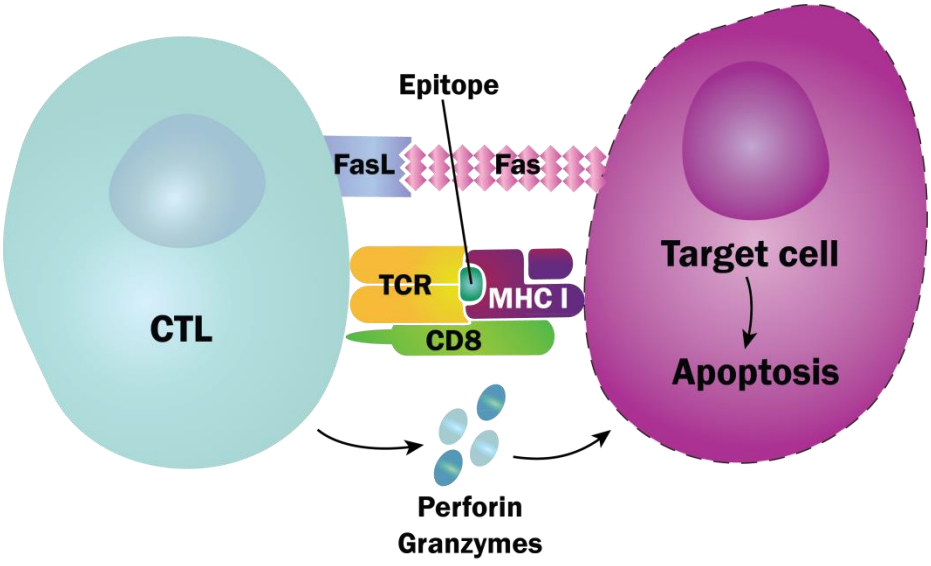


Figure 2. CTL mediated killing.
The TCR on the surface of an effector CD8+ T cell recognizes a MHC I-epitope complex on the infected cell. The binding of CD8 to the MHC I molecule stabilizes the interaction. Target cell killing is mediated by exocytosis of perforin and granzymes, and interaction of Fas ligand with its receptor Fas on the target cell. Adapted from (53).

Naïve T cells need to be primed by professional APCs presenting their cognate antigen, in combination with CD4+ T cell help, to result in fully activated cytotoxic T lymphocytes (CTLs). CTLs

can eliminate nucleated infected, diseased or damaged cells which present their specific epitope/antigen. Activated effector CTLs interact with their TCR to the MHC I-epitope complex on the target cell. The co-receptor CD8 is important for successful binding (53).

After successful activation, effector CTLs can eliminate their target cells either through perforin and granzyme release or/and through death receptor activation (Fas-Fas ligand interaction). Perforin molecules direct granzyme molecules to enter the cell (Figure 2). Granzymes target the electron transport chain in mitochondria for activation of the mitochondrial (intrinsic) pathway of apoptosis and/or they directly target procaspases. The Fas ligand on the surface of activated CTLs binds the death receptor (Fas) on the target cell, which results in activation of the extrinsic pathway of apoptosis (53).

1.3.3. The major histocompatibility complex

The major histocompatibility complex (MHC) is called human leukocyte antigen (HLA) complex in humans and was first described as one of the most important compatibility determinants for organ transplantation. Its true function, however, is to present peptides derived from antigens originating from the cell interior to T cells. There are two classes of MHC molecules, MHC I and MHC II. They differ by their structure, how they acquire peptides/epitopes, on which cells they are expressed and to which type of T cells they present epitopes.

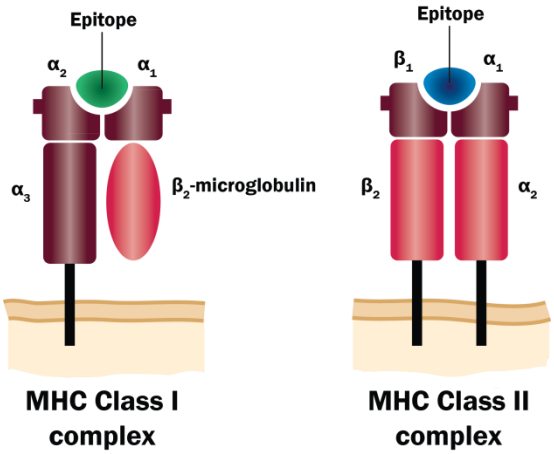


Figure 3. Schematic representation of MHC I and II molecules.
The MHC I molecule is composed of a heavy α chain with three Ig domains and a light non-covalently bound β chain. The α_1 and α_2 domains are the most polymorphic domains to ensure versatile binding properties of the peptide binding cleft. The MHC II molecule is composed of one α and one β chain. The binding cleft is formed by the α_1 and β_1 Ig domains. To ensure diverse binding properties of peptide binding cleft, these two domains have the highest polymorphism. Adapted from (53).

MHC I is a heterodimeric glycoprotein composed of a heavy α and light β chain. The heavy chain has 350 amino acids (aa) and a molecular mass of app. 44 kDa. It can be divided into three functional regions: external, transmembrane and intra-cytoplasmic. The external part is folded in three immunoglobulin (Ig) domains, α_1-α_3, each comprised of app. 90 aa. The interface of the α_1 and α_2 domains forms a closed epitope binding cleft, which non-covalently binds 8-11 aa long peptides. To ensure diverse MHC binding properties, the α_1 and α_2 domains are polymorphic, whereas the α_3

domain is conserved. A soluble and conserved β chain, called β_2-microglobulin (β_2M), with a molecular mass of app. 12 kDa, is non-covalently associated with the heavy chain (Figure 3). MHC I is expressed by all nucleated cells and is recognized by CD8+ T cells. Peptide presentation by MHC I is important for intracellular pathogen detection, including viruses.
MHC II is a heterodimeric glycoprotein composed of one α and β chain with a molecular weight of app. 30 kDa each. Each of the chains is comprised of two Ig domains. They form an open peptide binding cleft at the interface of the α_1 and β_1 domains. To assure a high binding diversity of the peptide binding cleft, these two domains are the most polymorphic. The binding cleft is open, therefore, the MHC II-restricted peptides have variable lengths spanning from 9 to 25 aa. The α_2 and β_2 domains are conserved (Figure 3). MHC II is only found on APCs, such as DCs, which present peptides derived from phagocytosed antigens. MHC II is recognized by CD4+ T cells (53). Since this thesis focuses on the detection of MHC I epitopes, only the MHC I antigen presentation process and MHC I polymorphism is discussed in more detail.

MHC I-peptide complex generation is a multi-step process and the involved components are termed antigen processing machinery (APM). It starts in the cytoplasm with the proteasomal degradation of poly-ubiquitinated misfolded or damaged self-proteins or foreign (e.g. viral) proteins (56). The proteasome has defined proteolytic activities maintained by the beta subunits 5-7, and proteosomal cleavage result in peptides with a length of 2-25 aa. Proteasomal degradation specificities and activity alter upon interferon-gamma (IFNγ) stimulation due to replacement of the catalytic proteasomal beta subunits 5-7 (57) by new beta subunits 8-10. This change produces peptides with different characteristics (58, 59). As many proteasomal peptides are too long to be successfully transported from the cytoplasm to the endoplasmatic reticulum (ER), where the loading of peptides to the MHC I complex takes place, they are trimmed by aminopeptidases in the cytosol (60, 61). Mostly 8-12 aa long peptides, but occasionally also longer ones, are selected and transported into the lumen of the ER through the transporter associated with antigen processing (TAP) (62, 63).

Loading of peptides on the MHC complex occurs via the peptide loading complex, consisting of the transporter TAP, the chaperons tapasin and calreticulin, the thiol oxidoreductase ERp57, and the MHC I molecule. A peptide is loaded into the peptide binding cleft and, if too long, trimmed by the endoplasmatic reticulum aminopeptidases 1 (ERAP1) and ERAP2 (64). Finally, the assembled MHC I-peptide complex is transported through the Golgi apparatus to the cell surface for CD8+ T cell recognition (65) (Figure 4).

MHC I molecules are encoded in a large chromosomal region with over 200 genes in humans. The MHC I α chain and both MHC II chains are located on chromosome 6 in humans and chromosome 17 in mice. MHC genes are called HLA in humans and H-2 in mice. The β_2M gene is located on chromosome 15 or 5 in humans or mice, respectively, and is not polymorphic. There are three class I genes; A, B and C in humans, and K, D, and L in mice, which are highly polymorphic. Until now, 3,192 HLA-A, 3,977 HLA-B and 2,740 HLA-C alleles have been identified in humans, with the numbers of

newly identified alleles still increasing. (66). Furthermore, every individual is diploid for MHC I genes, which results in three to six different MHC I molecules expressed on each cell. Thus, polygeny and polymorphism contribute to the high MHC I molecule variability among individuals of the same species (53). Most of the differences between MHC I alleles are located in the peptide binding cleft, resulting in different binding properties for peptides and TCR recognition (53).

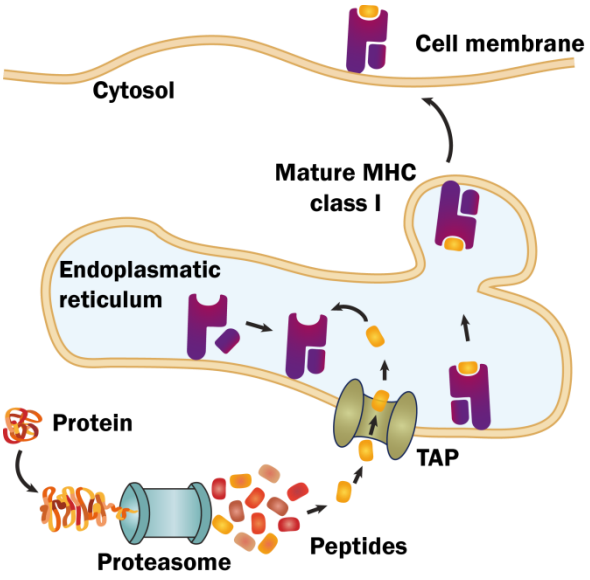


Figure 4. Schematic representation of the antigen processing machinery.
The proteasome degrades poly-ubiquitinated proteins into peptides in the cytoplasm. Peptides are transported through the transporter associated with antigen processing (TAP) into the lumen of the endoplasmatic reticulum and loaded on empty MHC I molecules with the assistance of the peptide loading complex. Subsequently, the MHC I-peptide complex is transported via the Golgi apparatus to the cell surface for CD8+ T cell recognition. Adapted from (53, 65).

The pioneering work of Rammensee and colleagues allowed for identification of the first binding motives and MHC I binding peptides in 1990 and 1991 (67-69). They used immunopurification of MHC I-peptide complexes and Edman degradation to determine which aa are the most prevalent in a particular place of the peptide sequence. In 1992, Hunt and colleagues performed similar experiments, only with Edman degradation being replaced with MS analysis yielding more identified peptides from one sample (70). Both groups reported comparable HLA-A2 binding patterns for the second aa in the peptide (position 2), leucine, and the aa at position 9, valine or leucine. The reported dominant peptide length was 9 aa (68, 70).

A peptide binds to the binding cleft non-covalently with hydrogen bonds and ion interactions. The interactions are usually strongest on both ends of the binding cleft. The aa of the peptide, which interact with the binding cleft the strongest, are called anchor residues and are different for every MHC allele. However, binding is influenced by interactions of the cleft with other aa in the peptide to a lesser extent as well (71).

Over the years, MS acquired data contributed to the establishing of peptide binding motives to MHC I molecules. They revealed that peptide binding motives are shared among some MHC I alleles, allowing their grouping in so called supertype families. For humans, nine supertype families were suggested, namely HLA-A1, -A2, -A3, -A24, -B7, -B27, -B44, -B58 and -B62 (72, 73). The grouping is

relevant especially for development of epitope-based vaccines, where the immune system can be stimulated with a small number of immunogenic peptides binding to all supertype family representatives. This enables the use of the same peptide vaccine in individuals bearing different HLA I alleles and, thus, a broad applicability across different populations.

The direct detection of MHC I peptides with MS enabled better understanding of the APM. It allowed the detection of differences in posttranslational modification (PTM), such as phosphorylation and glycosylation, in healthy and diseased cells (74) and finally the detection of viral (75) and tumor specific peptides (74). Furthermore, the direct MS detection of epitopes resulted in the development of the first therapeutic vaccine, IMA901, for renal cell carcinoma treatment by Rammensee and colleagues, which is currently in late-phase clinical studies (76-78).

Moreover, the information about peptide binding to MHC I molecules were used for development of MHC I peptide prediction algorithms, which are freely available on the internet, such as SYFPEITHI (79), immune epitope database (IEDB) (80-82) and NetMHC (83, 84) algorithms. However, prediction algorithms only work well for the most frequently studied alleles, where the data pool for algorithm training is big, whereas they are less reliable for rare alleles with less available information. Moreover, algorithms miss atypical peptides and they do not predict post-translation modifications. Thus, prediction algorithms are often combined with *in vitro* binding studies to confirm the binding properties of studied peptides (75). In the case of identifying viral MHC I peptides, the sequence information of a particular virus is used to predict possible binders (85, 86). In the case of cancer, the genomic sequencing data of cancer cells is compared with those of the healthy cells in order to identify mutated proteins. Protein regions containing mutations are used for MHC I prediction analysis and predicted binders are tested for their real binding potential *in vitro* afterwards (87-89).

Lately, a similar strategy of genetic sequencing and neoantigen (tumor specific epitope) predictions for tumor cells was employed in combination with the direct identification of predicted neoantigens with MS. Both studies reported the successful identification of novel neoantigens on the surface of tumor cells. Furthermore, identified neoantigens evoked T cell responses (90, 91). These studies confirmed the importance of direct identification of epitopes by MS analysis for successful design of anti-cancer treatment. Moreover, direct MS identification of MHC I epitopes can contribute to therapeutic vaccine design to cure persistent viral infections as well.

1.4. Introduction to proteomics mass spectrometry analysis

The goal of proteomics is a complete characterization of all proteins, their expression, turnover, localization, interaction, structure, modifications and activity in an organism, tissue, cell, organelle or specific signaling pathway. Thus, this makes proteomics more challenging than genomics, which does not vary from cell to cell or upon stimulus-induced changes. Several techniques can be employed for protein studies, such as protein microarrays, sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE), western blot or enzyme-linked immunosorbent assay (ELISA). However, MS has been the most widely used technique for large scale protein analysis, as it generates broader and deeper information than other technologies. This is mostly due to recent instrumentational advances, such as higher resolution power, better mass accuracy, sensitivity and scanning rates. A

prerequisite for successful MS analysis is suitable sample preparation and separation of proteins or peptides (92, 93). Basic liquid chromatography and mass spectrometry principles as well usual proteomics sample preparation workflows are introduced in this chapter.

1.4.1. Sample preparation for proteomics mass spectrometry analysis

The quality of MS data is heavily dependent on sample purity. Thus, protein isolation needs to be cell compartment-, cell-, tissue- and organism-specific. Furthermore, protein abundance is important for successful MS analysis as well. When a target protein is low abundant in a complex matrix, it can be enriched by immunoaffinity isolation or depletion of more abundant proteins. Sample complexity can be further reduced by protein fractionation strategies, such as one or two dimensional polyacrylamide gel electrophoresis (1D- or 2D-PAGE) or isoelectric focusing (94).

For effective protein solubilization, SDS is one of the best detergents. However, its usage in MS sample preparation is limited due to its influence on downstream sample processing, especially in LC-MS analysis. To overcome these problems, acid sensitive surfactants, such as Rapigest, have been developed. They contain an acid labile moiety, which is cleaved during or after enzymatic digestion and removed by short pelleting in a centrifuge (95). After successful solubilization, proteins are subjected to enzymatic degradation (96). Various proteolytic enzymes with diverse specificities for cleaving the amide bond are commonly used (92).

Sample complexity reduction on the peptide level can be conducted with one or combinations of more chromatographic fractionation strategies, such as strong or weak ion-exchange chromatography or isoelectric focusing. Any fractionation strategy used, should exploit different separation chemistry than the ones used in the final on-line LC-MS platform to successfully reduce sample complexity. Some of the separation strategies have a higher affinity for a particular PTM and are, thus, suitable for enrichment of peptides carrying that particular PTM (92).

1.4.2. Liquide chromatography

The most commonly used separation of peptides and proteins is reverse phase chromatography. Especially on the peptide level, reverse phase chromatography is the method of choice as its on-line connection to the ESI-MS allows for the automatization of analysis. Most of the LC-MS platforms are developed for this configuration. In contrast, hydrophilic interaction liquid chromatography (HILIC) is mostly used for pre-fractionation of peptides (97, 98). As these two types of chromatographic separation were used in the scope of this thesis, this chapter describes their separation principles.

1.4.2.1. Reverse phase liquid chromatography (RP-LC)

The stationary phase in RP-LC has a non-polar or hydrophobic surface, whereas the mobile phase is hydrophilic, allowing hydrophobic molecules to adsorb and to be retained longer. Usually the stationary phase for peptide separation is C18 material, which is silica covered with 18-carbon length alkyl groups. Other types of stationary phase include C8 (8-carbon alkyl chains) and C4 (4-carbon alkyl chains), which are more suitable for separation of proteins (99). Most reverse phase columns in

proteomics are particle-based. However, monolithic columns have been applied successfully as well (99). Monolithic columns are filled with a single piece of porous material prepared from silica or organic polymers. The advantage of monolithic columns is high permeability and low resistance compared to the particle filled columns, allowing the use of longer columns, which provide better separation. Monolithic material also allows filling capillaries with smaller inner diameters (e.g. 20 µm), compared to particle filled columns (where the smallest routinely used inner diameter is 75 µm) (100). However, they are not in common use as their performance is not yet comparable to particle filled columns (99).

Elution from the reverse phase material is usually conducted with a gradient of the mobile phase, beginning with high aqueous solvent content and ending with a high content of organic solvent. Besides the composition, also the pH of the mobile phase is important, as it can change the retaining properties of analytes (101). Chromatography is conducted with ultra/high performance liquid chromatography (U/HPLC) systems allowing the usage of long columns with small diameters for better separation of analytes. Reverse phase chromatography is usually used for on-line separation of peptides before MS analysis. If not specified differently, LC-MS in this thesis stands for reverse phase liquid chromatography-mass spectrometry analysis.

1.4.2.2. Hydrophilic interaction liquid chromatography (HILIC)

The stationary phase of normal phase liquid chromatography (NP-LC) is more hydrophilic than the mobile phase, causing more hydrophilic molecules to adsorb. HILIC is a variant of NP-LC with a polar stationary phase containing various types of ligands to achieve multiple-interaction solid phases, and thus combining characteristics of three major LC methods; normal phase, reverse phase and ion chromatography, resulting in different retention and separation selectivities (102). Here, the two most broadly used HILIC materials in proteomics are presented; TSKgel and zwitterionic (ZIC) HILIC.

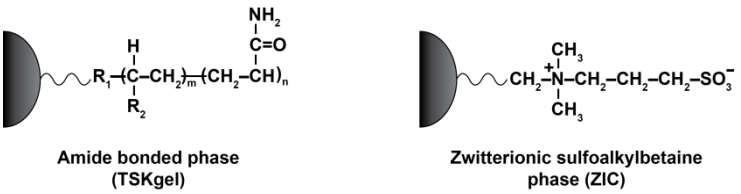


Figure 5. Schematic representation of TSKgel and ZIC HILIC stationary phases used in this project.
TSKgel is composed of an amide-silica phase, which binds analytes through hydrogen bonding and ion-exchange interactions. ZIC HILIC is composed of zwitterionic sulfoalkylbetaine bound to silica gel. The interactions of analyte and stationary phase are through hydrogen bonding and ion-exchange. Adapted from (97, 102).

TSKgel consists of carbamoyl groups bound to silica gel, known also as amide-silica phase. The carbamoyl groups bind carbonyl or hydroxyl groups of the sample through hydrogen bond interactions, whereas the silica gel binds through ion-exchange interactions and hydrogen bonding. TSKgel HILIC is suitable for separation of peptides, (oligo)saccharides and other polar compounds (Figure 5) (102).

ZIC HILIC has a zwitterionic sulfoalkylbetaine stationary phase with acidic sulfonic acid groups connected to quaternary ammonium groups through a short alkyl linker. Separation is based on hydrogen bonding and ion-exchange interactions. It is used for separation of peptides, proteins and other small organic molecules (Figure 5) (97, 102).

1.4.3. Mass spectrometry

Mass spectrometry is a method of choice for identification and quantification of biomolecules due to its sensitivity, accuracy, high speed and possibility to perform large-scale analysis. This method is based on measuring the mass-to-charge ratio (m/z) of analytes. For successful analysis, a generation of gas phase ions of the analyte is necessary. This is achieved in the ion source of the MS instrument either by loss or gain of charge (e.g. protonation or deprotonation) depending on the ionization technique. Next, ions enter the mass analyzer where they are separated based on the m/z and finally, they hit the detector, where they produce a record of relative abundance for mass spectrum generation (103).

Biomolecules, such as proteins and peptides, are sensitive molecules which can undergo fragmentation when exposed to an excess of energy. To preserve the bioanalyte structure before it enters a MS instrument, so called soft ionization methods are in use. The most common ion sources in proteomics are matrix-assisted laser desorption/ionization (MALDI), described by Tanaka (104) and Karas (105), and electrospray ionization (ESI) described by Fenn (106), all in the 1980s .
For MALDI, an analyte is first co-crystalized with a high excess of a MALDI matrix on a plate, which is afterwards subjected to laser desorption. A laser pulse results in sublimation of matrix and analyte into the gas phase, where molecules of matrix and analyte can interact. It is hypothesized that the matrix protonates the analyte (107, 108). MALDI ionization mostly produces singly-charged ions. However, larger proteins analyzed by MALDI acquire higher charge states. (109). Very high molecular mass analytes (up to 300,000 Da) can be ionization with MALDI. The main disadvantage of MALDI is the low reproducibility of signal due to non-homogenous distribution of matrix and analyte (110).
On the contrary, ESI ionization takes place in solution and is therefore suitable for a direct, on-line connection with liquid chromatography separation devices. The liquid containing the analyte flows through a capillary tip, which is under a strong electrical field. This results in liquid dispersion into droplets, and their charging. Charged droplets evaporate, resulting in analyte transfer to the gas phase and accumulation of charge on the analyte (111-113). ESI produces multiply charged ions, which is well suited for measurement of bigger molecules. As molecular masses are measured as mass-to-charge ratio, higher charge states reduce m/z and allow for measuring of molecules which would otherwise be outside the mass range of mass analyzers. All described traits made ESI-MS in combination with on-line LC separation one of the most widespread MS applications in proteomics (114).

As mentioned above, mass analyzers separate ions based on their m/z. They can operate in a scanning and a selecting mode. In the scanning mode, they monitor all molecules within a specified m/z range, which results in a full scan, whereas in the selecting mode, they isolate one ion and filter out all others.
Quadrupole, ion trap, orbitrap, Fourier transform ion cyclotron resonance (FT-ICR) and time-of-flight (TOF) are five basic types of mass analyzers currently used for proteomic research. They have different designs and performances, so each of them determines m/z ratios differently. For illustration, Fourier transform-based mass analyzers have high mass accuracy and resolution (up to 1,000,000)

but are slow, whereas quadrupole and ion traps are fast and have low mass accuracy and resolution (~1,000). However, different mass analyzers can be combined together in tandem or hybrid instruments to use the strengths of each of them (115).

A quadrupole consists of four parallel metal rods connected pairwise and with an angle of 180°, thereby forming a “tube”. Each pair of diagonally opposing rods has the same direct current (DC) or radio-frequency (RF) potential, generating an electric field. After the electric field is applied, an ion will oscillate depending on its m/z and the time-varying RF field in the quadrupole. The m/z of an ion is calculated from the electric field values applied to the quadrupole which allow an ion to pass through the analyzer. If the filtering mode of the quadrupole is applied, only ions with a particular m/z will be able to pass the analyzer (93).

The most widely used ion trap in MS is the linear ion trap (LIT), which has the same design as a quadrupole. The electric field is created by two-dimensional RF potential, and additionally axial stopping and axillary potentials are applied to end of the electrodes. With the stabilizing RF and axial potentials, ions are kept in the trap. LIT permits ion acceleration, thereby excitation and consequently fragmentation of trapped ions, resulting in MS^2/MS^3 spectra (116, 117)

The orbitrap is based on principles described by Makarov (118). In brief, it is an electrostatic ion trap, where ions are caught between a spindle-shaped inner electrode supplied with high voltage and an outer electrode at ground potential. The ions rotate around and oscillate along the inner electrode forming complex trajectories. The oscillating signature of an ion is used for calculation of the m/z ratio using the Fourier transformation (118, 119). Its advantages are high resolution with high mass accuracy and – in modern instruments – high sensitivity and speed.

FT-ICR was first described in 1974 by Comisarow and Marshall (120). The analyzer allows high resolution and high mass accuracy. The most common form of FT-ICR is an open cylinder. Electrodes are separated axially in two or four segments (semi cylinders). After ions enter the cell, a RF potential is applied to excite them, such that they start to orbit. The potential is turned off and ions move only in the magnetic field. Based on the cyclotron frequency in the magnetic field, the m/z of an ion is determined using the Fourier transformation (119, 121).

A TOF analyzer measures the time that an ion with a certain m/z takes to traverse a certain distance in a vacuum without an electric field after being accelerated through an electric field. With the introduction of a reflector (electric field ion mirror), the flight distance is elongated, which improves the resolving power of the instrument. The advantage of TOF analyzers is that their m/z range is almost unlimited. Usually TOF analysis is combined with MALDI ionization (122).

As outlined above, every MS instrument determines m/z for measured ions. MS^1 is the determination of the precursor ion m/z. MS^2 analysis monitors m/z of fragments of this precursor ion. MS^2 spectrum generation starts with m/z determination of the precursor ion by the first MS analyzer. Afterwards, the precursor is fragmented and the m/z for fragment ions is determined in the second analyzer for MS^2 spectrum generation. Fragments are generated through a process called collision induced dissociation (CID), where precursor ions are fragmented in the collision cell, which is a quadrupole filled with inert gas, such as helium or argon. The inert gas molecules collide with analytes, resulting in molecule breaks (115).

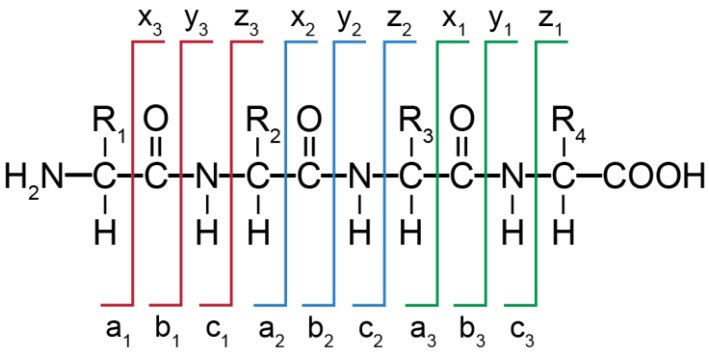


Figure 6. Nomenclature of peptide fragmentation by mass spectrometry (MS).
The charge is on the N-terminus for a-, b- and c-ions, whereas it is on the C-terminus for x-, y-, and z-ions.

Peptides and proteins will preferentially break at the most labile amide bond, resulting in a characteristic pattern of so called b- and y-ions. If the charge is on the N-terminus of the peptide, they are named b-ions. If the charge is located on the C-terminus of the peptide, ions are referred to as y-ions. If different types of fragmentation are employed (e.g. electron-transfer dissociation), the fragmentation results in a-x and/or c-z ion products (115). The peptide fragment nomenclature established by Roepstorff, Fohlmann (123) and Biemann (124) is presented in Figure 6. Mass differences between MS^2 ions allow for aa sequence determination of a peptide. This is achieved by comparison of an acquired peptide spectrum with theoretical spectra calculated from a proteomic or genomic database with dedicated software (125).

Proteomics MS analysis can be divided into top-down and bottom-up approaches. The top-down approach performs analysis of intact proteins, whereas the bottom-up approach requires enzymatic cleavage of proteins, which results in shorter peptides. When a bottom-up approach is performed on a mixture of proteins, it is called shotgun proteomics (126). Peptides are separated via on-line liquid chromatography and directly measured with MS^1 and/or MS^2 analysis (127). A database search is performed on the resulting MS^2 spectra for peptide identification. The advantages of the bottom-up approach are automatization of LC-MS analysis, well established peptide identification programs and quantification strategies. However, only a portion of the protein sequence is covered with this approach, meaning that one can miss identification of protein isoforms and localization of posttranslational modifications (PTM). Nevertheless, the bottom-up approach is most widely used among all proteomics MS analyses (92, 127, 128).
The top-down approach is a relatively young proteomics MS discipline; therefore it lacks many tools that are already available for the bottom-up approach. The top-down approach is mostly performed with off-line separation and direct infusion of single proteins or simple protein mixtures for MS analysis, whereas on-line LC separation solutions are limited so far. For top-down MS analysis, mostly FT-ICR or orbitrap analyzers with high resolution and mass accuracy have been in use, allowing charge state determination. Intact proteins can acquire variable charge states; therefore the MS spectrum interpretation is complicated with instruments with lower resolving power. The advantages of top-down approach are better sequence coverage, isoform detection and identification of more PTMs (127, 128).

MS^2 analysis on the peptide level is performed mostly with an untargeted approach (also referred to as shotgun and data-dependent acquisition). In the first analyzer, the m/z of all precursor ions is determined, which is referred to as the full scan mode. The most intense precursors are then selected for fragmentation and MS^2 generation. The dynamic exclusion function ensures that an already fragmented precursor is not fragmented again for a particular time period, which results in a good analytical coverage of peptides and proteins (Figure 7 A). However, an ion of a different peptide but with the excluded m/z will not be fragmented either. Thereby, its identity is missed, which makes this measuring approach less reproducible and low abundant peptides can be easily missed, especially in complex samples (92). Both, lower reproducibility and inability to detect low abundant peptides can be partially overcome by employment of long LC separation gradients and pre-fractionation of complex samples, which both have drawbacks. Long LC separation gradients reduce the throughput (129) and pre-fractionation introduces sample losses (130).

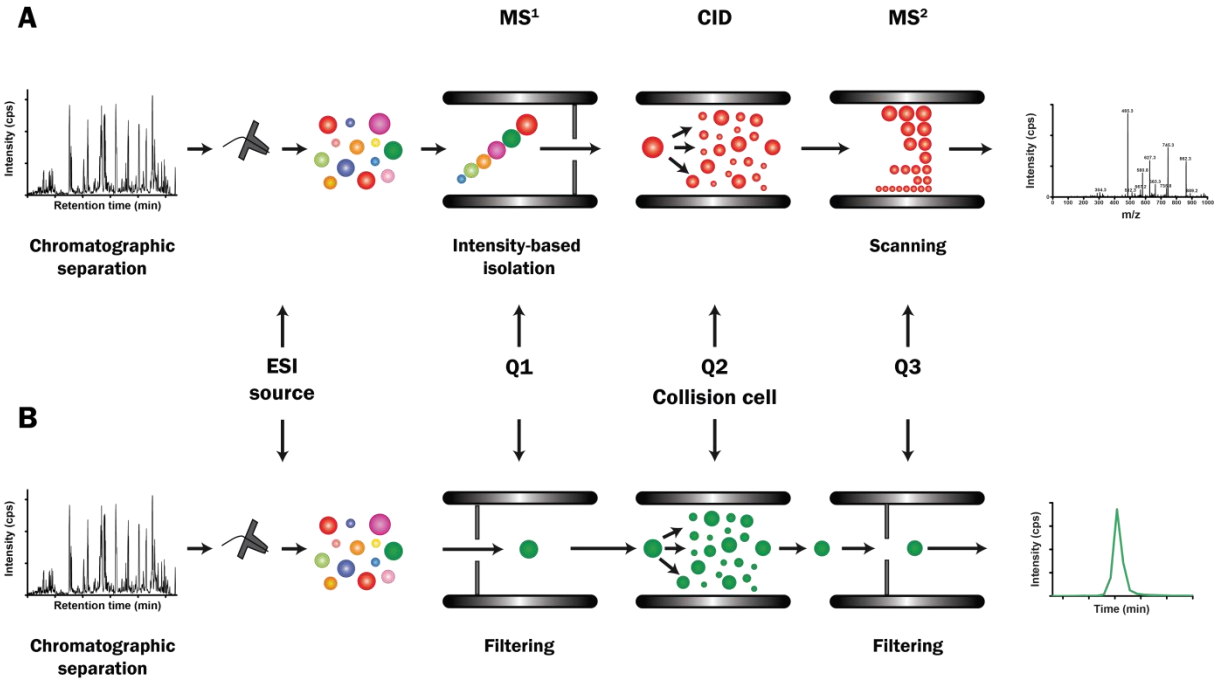


Figure 7. Comparison of untargeted and targeted MS^2 approaches of peptide analysis with an ESI-quadrupole system.
A) In untargeted approaches, all eluting peptides are scanned for their m/z in the first quadrupole to result in the so called full MS spectrum. The most intense precursors are selected for fragmentation in the collision cell and scanned in the third quadrupole for MS^2 spectrum generation. B) In targeted approaches, a predetermined ion precursor is filtered in the first quadrupole and fragmented in the collision cell. Selected fragments are filtered in the third quadrupole and only their intensities are measured, while all others are ignored. Quantification is possible by comparing peak intensities between samples or with internal standards Q: quadrupole, CID: collision induced dissociation. Adapted from (131).

After data acquisition, results are subjected to data base searches for peptide and protein identification, where an identity is assigned based on the comparison with theoretical spectra from a proteomic or genomic database. With introduction of labeling strategies and sophisticated bioinformatics tools, amounts of particular proteins or global proteome changes can be investigated. The advantage of data-dependent acquisition lies in its applicability to any sample without prior knowledge of its content (92, 127).

Targeted MS^2 analysis was established to improve sensitivity, selectivity, and reproducibility of sample analysis and quantification of sample analysis. Several strategies, such as selected reaction monitoring (SRM, known also as multiple reaction monitoring MRM), parallel reaction monitoring (PRM) and accurate inclusion mass screening have been developed. The common trait of all of them is that they measure pre-defined analytes and ignore others. Furthermore, they are able to collect target ions for a longer time to increase signal intensity. SRM is usually performed with hybrid triple quadrupole-LIT instruments, where the first quadrupole acts as a filter and selects a target ion. The selected ion is fragmented in the second quadrupole with CID. The selected, usually most intense, fragments are filtered by the third quadrupole before they hit the detector (Figure 7 B). The recorded signals allow precise quantification of a peptide and a corresponding protein (131-133).
Quantification is based on comparing peak intensities of transitions between samples and/or an added internal standard. The pair of precursor and fragment ion is called a transition. For reliable detection and quantification of a peptide at least three transitions are monitored, and for proteins at least three unique peptides are required. Experimental design and data processing of the SRM experiment is based on target peptide information obtained in data-dependent acquisition experiments and corresponding full MS^2 spectra. Therefore, detailed knowledge of the target peptide is a prerequisite for optimal SRM method design. The method disadvantage is that only a limited number of transitions can be measured in one analysis and that optimization of transition is not trivial, as it requires technical knowledge and is time consuming. When created, a SRM method can be applied to numerous samples with high reproducibility (131-133). Taken together, targeted and untargeted MS^2 are complementary approaches for protein analysis (93).

Equivalent to MS^2 analysis providing more information than the molecular mass in MS^1 analysis, additional fragmentation stages in MS^3 analysis provide further information about a MS^2 fragment. MS^3 analysis is possible with hybrid triple quadrupole-linear ion trap instruments and is widely used for small molecule structural studies (134). MS^3 has higher specificity compared to MS^2, which makes MS^3 scanning an attractive method also in proteomics as it reduces interferences. MS^3 in proteomics usually operates in a targeted mode and is sometimes also called MRM^3 (pronounced MRM cubed) (134).
When a selected ion enters the first quadrupole, the m/z is determined and it is allowed to pass into the collision cell. The ion is fragmented in the collision cell and the fragments with the most intense signal are allowed to enter the linear ion trap (LIT). In the LIT, selected fragments are collected and subjected to a second fragmentation. The resulting fragments hit the detector and generate a MS^3 spectrum (134) as illustrated in Figure 8.
During the second fragmentation, the most labile amide bond breaks. MS^3 y-ions from MS^2 y-ion fragmentation have the same m/z, whereas b-ions have a lower mass, which is reduced by the sum of the missing N-terminal aa in the y-precursor and the missing N-terminus. They are designated as bˆ (pronounced b hat). Similarly as for y-ions, MS^3 b-ions from MS^2 b-ion fragmentation have the same m/z, whereas y-ions have a lower mass, which is reduced for the sum of the missing C-terminal aa in the b-precursor and the missing C-terminus. They are designated as yˆ (pronounced y hat) (135).

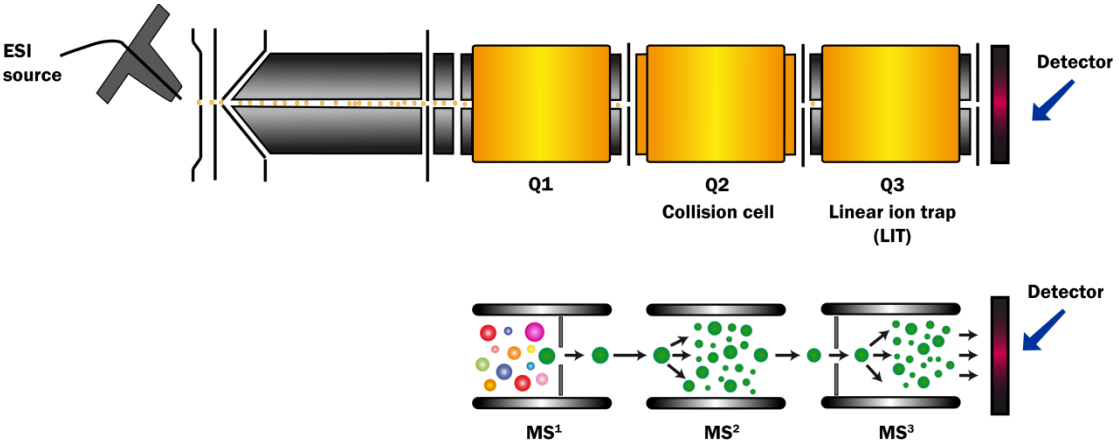


Figure 8. Schematic representation of a triple quadrupole-linear ion trap instrument and the MS³ spectrum generation principle. The upper part illustrates the instrument and the lower part the MS³ spectrum generation principle. After ions enter the instrument, selected ions are filtered in the first quadrupole and further fragmented in the collision cell. A selected fragment is filtered in the linear ion trap (LIT) and further fragmented for MS³ spectrum generation. Q: quadrupole. Adapted from (131, 134).

For every optimal MS^3 spectrum, all MS^2 fragments of interest need to be manually optimized for their fragmentation energies, meaning that knowledge about the measured peptide is a crucial prerequisite for this methodology. The additional fragmentation and the required energies for optimal spectrum generation make MS^3 scanning highly specific, as the interferences from other ions are minimized, and sensitive, as the signal-to-noise ratio is increased. However, not every MS^2 fragment ion can be used for MS^3 fragmentation. Their sequence needs to be long enough to generate a specific MS^3 spectrum (min. 5 aa), and the ion signal intense enough, such that the resulting MS^3 spectrum has intense signals as well (134-136).

Three types of MS instruments were employed during this thesis; ESI-triple quadrupole-linear ion trap hybrid instruments (QTrap5500 and 6500, AbSciex), an ESI-quadrupole-orbitrap hybrid instrument (Q-Exactive, Thermo Scientific) and a MALDI-TOF MS instrument (UltraFlextreme). All three instruments analyze in MS^1 and MS^2 mode, whereas the QTrap instruments allow measurements in MS^3 mode as well.

1.4.4. Isolation and detection of MHC I peptides with LC-MS

The MHC I-peptide complex isolation protocol has not changed significantly since it was published in the early 1990s (67, 68, 70). MHC I-peptide complexes are isolated by immunoprecipitation (IP) after cell lysis with detergents. The choice of available detergents is limited, as they need to preserve the 3D conformation of the MHC I complex and also be LC-MS friendly. Numerous studies reported the usage of CHAPS (67, 68, 91, 137, 138). After the IP, a sample is subjected to acidic elution of bound peptides, ultrafiltration and reverse phase purification for peptide extraction. A recent publication reported an isolation strategy with reverse phase isolation only (139), which minimizes peptide losses occurring during ultrafiltration steps (140, 141).

Isolated peptides are then measured with LC-MS^2 analysis in an untargeted approach, resulting in high numbers of identified epitopes and global peptidome analysis as in (139, 142), or in a targeted approach for a target peptide verification and quantification across multiple samples (91, 143).

2. Aim of this thesis and study design

The overall aim of the group Immunotherapy and -prevention at the DKFZ, Heidelberg, Germany, where the work for this thesis was performed, is to develop a therapeutic HPV vaccine.

As it was shown in the literature before (144), spontaneous regression of HPV induced malignancies is induced by T cell mediated immune responses. Thus, the identification of HLA-restricted HPV-derived T cell epitopes was chosen as the basis of the project. Up to now, epitope candidates for vaccine development in most cases have been indirectly identified with functional cellular assays testing the potential immunogenicity of candidate epitopes. However, a good immunogenicity profile does not mean that the epitope is actually presented on the surface of HPV-positive cells.

We addressed this problem by applying the reverse immunology principle (85-88) for direct identification of HPV epitopes on cell surfaces by mass spectrometry. We chose HLA-A2 for its high frequencies across all populations worldwide (145) and because the HLA-A2 restricted HPV16 $E7_{11-19}$ epitope was previously described to be presented on the cell surface of HPV16-transformed cells and tumors (146, 147).

The reverse immunology principle steps that we followed were, first *in silico* predictions of HPV16 E6 and E7 epitope binding properties to HLA-A2 molecules. Second, good predicted binders were tested for their actual binding *in vitro* in competition-based binding assays described by Kessler et al. (148, 149). This part of the study was performed by my colleagues Dr. Stephanie Hoppe and Marius Küpper. Finally, strong binding peptides were monitored in MS analysis after their isolation from HPV16-transformed cells in context of my thesis (Figure 9).

The amounts of epitopes presented on the cell surface were expected to be low, due to known HPV immune evasion mechanisms (17, 150, 151). Therefore, a targeted, highly sensitive MS^3 analysis was employed (134-136, 152). As the number of analytes that can be monitored during one MS^3 measurement is limited, reduction of potential targets for mass spectrometry analysis by preselection of potential epitopes – as described above – is essential.

The aim of this thesis was the development of a mass spectrometry methodology for direct identification of viral T cell epitopes from the surface of transfected or infected cells. This included isolation, extraction and purification, as well as the LC-MS detection of viral epitopes.

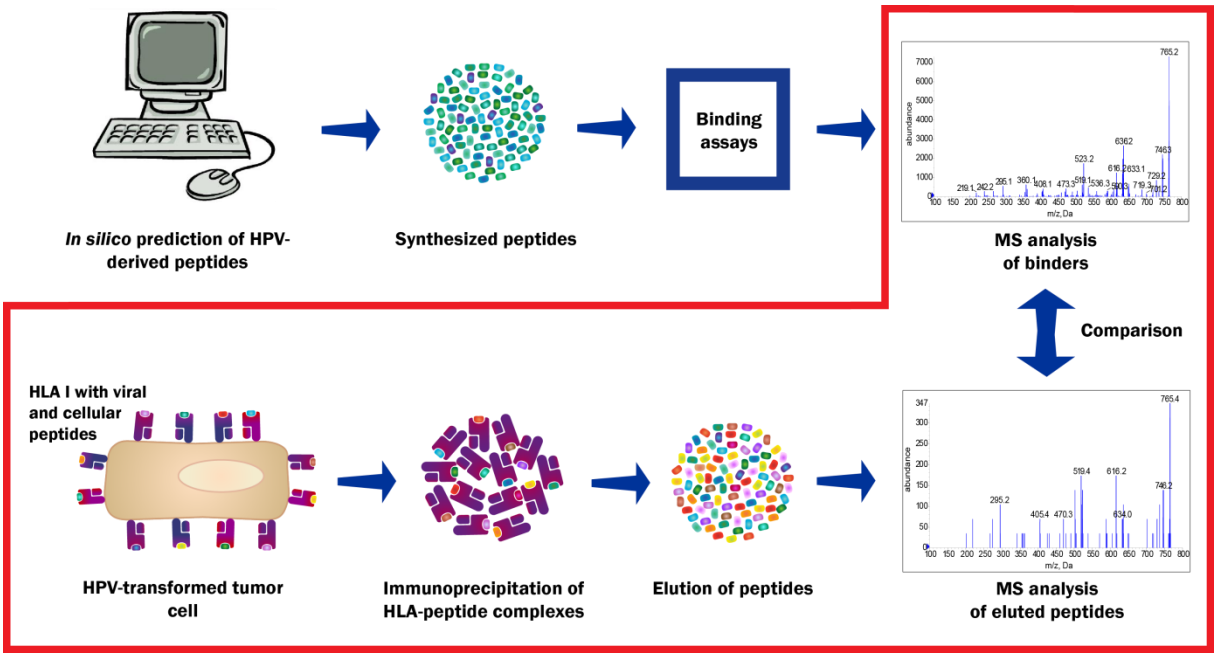


Figure 9. Strategy for therapeutic HPV vaccine development of the Immunotherapy and -prevention Research Group.
Epitopes were predicted in silico for their binding properties. Good predicted binders were tested for their actual binding properties in vitro and only good binders were monitored in mass spectrometry analysis after their isolation from the HPV-transformed cells. The red framed part of the illustration shows the optimization steps described in this project. The unframed part represents the parts of the strategy that were performed in a closely related PhD project within the group.

3. Materials and Methods

3.1. Materials

3.1.1. Chemicals

Chemical	Company
2-mercaptoethanol	Roth, Karlsruhe, Germany
2,2,2 trifluoroethanol (TFE)	Fluka Analytical, Schnelldorf, Germany
α-cyano-4-hydroxy-cinnamic acid	Sigma-Aldrich, Steinheim, Germany
Acetone	Sigma-Aldrich, Steinheim, Germany
Acetonitrile (ACN), HPLC/MS grade	Biosolve BV, CE Valkenswaard, The Netherlands Fluka Analytical, Schnelldorf, Germany Sigma-Aldrich, Steinheim, Germany
Acetic acid, 100 %, glacial	Merck Millipore, Billerica, MA, USA
Ammonium bicarbonate (ABC)	Sigma-Aldrich, Steinheim, Germany
Ammonium hydroxide solution, ≥25% NH_3 in H_2O	Sigma-Aldrich, Steinheim, Germany
CHAPS	AppliChem, Darmstadt, Germany
D-Glyceraldehyde 3-phosphate (G3P) 8 – 13 mg/mL in H_2O	Sigma-Aldrich, Steinheim, Germany
Dimethyl sulfoxide (DMSO)	Sigma-Aldrich, Steinheim, Germany
Dimethyl sulfoxide (DMSO), LC/MS grade	Sigma-Aldrich, Steinheim, Germany
Dithiothreitol (DTT)	Sigma-Aldrich, Steinheim, Germany
Ethyl acetate	Sigma-Aldrich, Steinheim, Germany
Ethanol (EtOH), 96% (v/v)	Sigma-Aldrich, Steinheim, Germany
ECL Western Blotting Substrate, Pierce	Thermo Fisher Scientific, Waltham, MA, USA
Formaldehyde solution, 37% in H_2O	Sigma-Aldrich, Steinheim, Germany
Formic acid (FA), LC/MS grade	ProteoChem, Inc., Denver, USA Biosolve BV, CE Valkenswaard, The Netherlands
Glycerol, PlusOne (87%)	GE Healthcare, Germany
Glycine	Sigma-Aldrich, Steinheim, Germany
Glycolic acid	Sigma-Aldrich, Steinheim, Germany
Hydrochloric acid (HCl)	VWR International, Fontenay-sous-Bois, France Sigma-Aldrich, Steinheim, Germany
IEF minearal oil, PlusOne DryStrip cover fluid	GE Healthcare, Germany
Iodoacetamide (IAA)	Sigma-Aldrich, Steinheim, Germany
IPG buffer, pH 3-10	GE Healthcare, Uppsala, Sweden
Isopropanol	Sigma-Aldrich, Steinheim, Germany
Laemmli sample buffer (2x)	Biorad, Hercules, USA
Methanol	Sigma-Aldrich, Steinheim, Germany
Methanol, LC-MS grade	Biosolve BV, CE Valkenswaard, The Netherlands Sigma-Aldrich, Steinheim, Germany
Phenylmethanesulfonylfluoride (PMSF)	Roth, Karlsruhe, Germany
Polyoxyethylene-sorbitan monolaureate (Tween20)	MP Biomedicals, Illkirch, France
Potassium chloride (KCl)	Roth, Karlsruhe, Germany
Potassium dihydrogen phosphate (KH_2PO_4)	Roth, Karlsruhe, Germany
Powdered milk, blotting grade	Roth, Karlsruhe, Germany
Protease inhibitor cocktail (PIC), cOmplete Mini	Roche Diagnostics, Mannheim, Germany
Sodium acetate ($NaCH_3COO$)	Sigma-Aldrich, Steinheim, Germany
Sodium azide (NaN_3)	AppliChem, Darmstadt, Germany
Sodium cyanoborohydride ($NaBH_3CN$)	Sigma-Aldrich, Steinheim, Germany
Sodium dodecyl sulfate (SDS), ultrapure	Roth, Karlsruhe, Germany
Sodium bicarbonate ($NaHCO_3$)	Sigma-Aldrich, Steinheim, Germany
Sodium chloride (NaCl)	Sigma-Aldrich, Steinheim, Germany
Sodium hydroxide (NaOH)	Sigma-Aldrich, Steinheim, Germany

Chemical	Company
Sodium phosphate dibasic dihydrate ($Na_2HPO_4 \cdot 2H_2O$)	Sigma-Aldrich, Steinheim, Germany
Triethylammonium bicarbonate (TEAB) buffer (1M)	Fluka Analytical, Schnelldorf, Germany
Trifluoroacetic acid (TFA), >99% pure, protein sequencing grade	Sigma-Aldrich, Steinheim, Germany
Trizmabase (Tris)	Sigma-Aldrich, Steinheim, Germany
Water (H_2O), LC/MS grade	Biosolve BV, CE Valkenswaard, The Netherlands

3.1.2. Buffers and solutions

Name	Components
10x PBS	1.37 M NaCl 27 mM KCl 100 mM Na_2HPO_4 (anhydrous) 20 mM KH_2PO_4 pH=7.2 adjusted with NaOH or HCl
10x SDS-PAGE running buffer	30 g Tris 144 g Glycine 10 g SDS ad. 1 L ddH_2O
2x CHAPS buffer	1.2% (w/v) CHAPS dissolved in 1x PBS
2x SDS-PAGE loading buffer	5% 2-mercaptoethanol (v/v) in Laemmli sample buffer (2x)
CHAPS lysis buffer	5.25 mL 2x CHAPS buffer 1 mM PMSF 1 cOmplete Mini PIC tablet
Direct elution buffer	10% acetic acid 1x cOmplete Mini PIC tablet in H_2O
IP elution buffer	0.3% TFA in H_2O (all LC-MS grade)
IP washing buffer	0.6% (w/v) CHAPS dissolved in 1x PBS
G3P reaction solution per a micro-column	40 µL 10 mM TEAB in H_2O 30 µL G3P (stock) 0.7 µL 0.6 M $NaBH_3CN$ pH=2 – 2.2 adjusted with TEAB or G3P
IEF rehydration solution	1x OFFGEL stock solution
IEF washing solution	50% methanol 49% H_2O 1% FA (all LC-MS grade)
OFFGEL stock solution (1.25x) for a sample	4 mL H2O (LC-MS grade) 40 µL glycerol (87%) 46 µL IPG buffer, pH 3-10 (stock)
PMSF stock	100 mM PMSF dissolved in 96% ethanol
TiO_2 elution buffer	30% ACN 60 µL ammonium hydroxide solution (≥25% NH_3 in H_2O) in H_2O pH>11.6
TiO_2 loading buffer	80% ACN 5% TFA 1 M glycolic acid in H_2O
TiO_2 washing buffer 1	80% ACN 1% TFA in H_2O
TiO_2 washing buffer 2	10% ACN 0.1% TFA in H_2O

Name	Components
Ultrafiltration tubes washing buffer	50% Methanol 0.1% TFA in H_2O
WB blocking buffer	5% (w/v) milk powder in 1x PBS
WB transfer buffer	9 g Tris 43.2 g Glycine 20% (v/v) methanol ad. 3 L ddH_2O
WB washing buffer	0.05% (v/v) Tween20 in 1x PBS

3.1.3. LC mobile phases and other MS solutions

All reagents were prepared with LC-MS grade.

Mobile phase	Composition	Company	Used with
A	0.1% FA	ProteoChem, Inc., Denver, USA	nanoAcquity UPLC system in DKFZ
	0,01% TFA	ProteoChem, Inc., Denver, USA	
	in H_2O	Biosolve BV, CE Valkenswaard, The Netherlands	
B	0.1% FA	ProteoChem, Inc., Denver, USA	nanoAcquity UPLC system in DKFZ
	0,01% TFA	ProteoChem, Inc., Denver, USA	
	in ACN	Biosolve BV, CE Valkenswaard, The Netherlands	
A	0.1% FA	Biosolve BV, CE Valkenswaard, The Netherlands	nanoAcquity UPLC system in ZMBH
	in H_2O	Biosolve BV, CE Valkenswaard, The Netherlands	
B	0.1% FA	Biosolve BV, CE Valkenswaard, The Netherlands	nanoAcquity UPLC system in ZMBH
	in ACN	Biosolve BV, CE Valkenswaard, The Netherlands	
A	0.1% FA in H_2O	Fluka Analytical, Schnelldorf, Germany	speLC
B	0.1% FA	Sigma-Aldrich, Steinheim, Germany	speLC
	in ACN	Fluka Analytical, Schnelldorf, Germany	
Solvent for manual injection of peptides	50% ACN	Biosolve BV, CE Valkenswaard, The Netherlands	QTrap5500 and QTrap6500
	0.1% FA	ProteoChem, Inc., Denver, USA	
	in H_2O	Biosolve BV, CE Valkenswaard, The Netherlands	
MALDI matrix solution	saturated α-cyano-4-hydroxycinnamic acid in	Sigma-Aldrich, Steinheim, Germany	UltrafleXtreme
	70% ACN	Sigma-Aldrich, Steinheim, Germany	
	0.1% TFA in H_2O	Fluka Analytical, Schnelldorf, Germany	

3.1.4. Liquid chromatography and mass spectrometry instrumentation

Name	Company	Location
QTrap5500	Ab Sciex, Foster City, CA, USA	ZMBH Proteomics mass spectrometry core facility, Heidelberg
QTrap6500	Ab Sciex, Foster City, CA, USA	DKFZ, Heidelberg
Q Exactive	Thermo Fisher Scientific, Bremen, Germany	University of Southern Denmark, Odense, Denmark

Name	Company	Location
UltrafleXtreme	Bruker Daltonik, Bremen, Germany	University of Southern Denmark, Odense, Denmark
nanoAcquity UPLC system	Waters, Milford, MA, USA	ZMBH Proteomics mass spectrometry core facility and DKFZ, Heidelberg
speLC system	Thermo Fisher Scientific, Odense, Denmark	University of Southern Denmark, Odense, Denmark

3.1.5. Other laboratory equipment and instrumentation

Equipment	Name	Company
Analytical balance	XS DualRange	Mettler Toledo, Glostrup, Denmark
		Ohaus, Nänikon, Switzerland
Cell counter	Countess™ Automated Cell Counter	Invitrogen, Carlsbad, USA
Cell freezing device	Mr. Frosty	Bel-Art Products, Wayne, NJ, USA
Cell culture incubator	Heracell 150i	Thermo Fisher Scientific, Waltham, MA, USA
Cell culture microscope	Wilovert Standard 30	Hund Wetzlar, Wetzlar, Germany
Centrifuges	Centrifuge 5417R	Eppendorf, Hamburg, Germany
	Centrifuge 5418	Eppendorf, Hamburg, Germany
	Megafuge 16R	Thermo Fisher Scientific, Waltham, USA
	Centrifuge 5424	Eppendorf, Hamburg, Germany
	Centrifuge 5824 R	Eppendorf, Hamburg, Germany
	Centrifuge 5810 R	Eppendorf, Hamburg, Germany
	Centrifuge 5418	Eppendorf, Hamburg, Germany
Glass pipets		Hirschmann Labortechnik, Eberstadt, Germany
Glassware	Duran	Schott, Mainz, Germany
	Fisherbrand	Thermo Fisher Scientific, Waltham, USA
Isoelectric focusing device	Agilent 3100 OFFGEL Fractionator	Agilent Technologies, Santa Clara, CA, USA
Laminar flow hood	SterilGard Class II laminar flow hood	The Baker company, Stanford, USA
Magnetic stirrer	CombimagRCO	IKA-Werke GmbH, Staufen, Germany
	MR-Hei-Standard	Heidolph, Instruments, Schwabach, Germany
Neubauer cell counting chamber	Profondeur	Brand, Wertheim, Germany
pH meter	766	Knick, Berlin, Germany
	MP220, InLab Microelectrode	Mettler Toledo, Glostrup, Denmark
Pipette-Boy	pipetus®-akku	Hirschmann Labortechnik, Eberstadt, Germany
Pipettes	Pipetman® Gilson, P2, P20, P200, P1000, P5000	Gilson, Bad Camberg, Germany
	FinnPipette F2	Thermo Scientific, Rockford, IL, USA
Power supply	EPS500-400, EPS3500	Pharmacia, Uppsala, Sweden
	MP250V	MS Major Science, Saratoga, CA, USA

Equipment	Name	Company
Rolling shaker	CATMR5	Zipperer, Staufen, Germany
Rotator		NeoLab, Heidelberg, Germany
Scale	Kern EG 4200-2NM	Kern&Sohn, Balingen, Germany
SDS-PAGE chamber	Mini-Protean® Tetra cell	Biorad, Hercules, USA
Thermomixer	Thermomixer compact	Eppendorf, Hamburg, Germany
Ultrasonic water bath	VWR Ultrasonic cleaner USC300TH	VWR International, Leuven, Belgium
Vacuum centrifuge	SpeedVac RVC 2-25 CD plus	Martin Christ Gefriertrocknungsanlagen GmbH, Osterode am Harz, Germany
	Speedvac concentrator	Bachofer Laboratoriumsgeräte, Reutlingen, Germany
Vortexer	Vortex-Genie 2	Scientific Industries, Bohemia, NY, USA
WB chamber	Mini Trans-Blot® cell	Biorad, Hercules, USA
WB imaging device	Fusion-SL	Vilber Lourmat, Eberhardzell, Germany

3.1.6. Peptide extraction materials

Material	Name	Company
Cellulose tissue plug for micro-columns	Kimtech Science precision wipes	Kimberly-Clark Professional, Reigate, Surrey, UK
C8 micro-column plug	Empore Octyl C8 extraction disk	3M company, St. Paul, MN, USA
Extraction pipette tips C18	Bond Elut OMIX pipette tips	Agilent Technologies, Santa Clara, CA, USA
Extraction cartridges C18	Seppak Vac 1cc (100 mg)	Waters, Milford, MA, USA
HILIC chromatographic packing	TSKgel Amide-80, 3µm	Tosoh corporation, Tokyo, Japan
HILIC chromatographic packing	SeQuant ZIC-HILIC, 3µm	Merck SeQuant AB, Umea, Sweden
RP chromatographic packing	Poros® 20 R2 reversed-phased resin, 20µm	Applied Biosystems, Bedford, USA
RP chromatographic packing	Oligo R3 reversed-phased resin, 50µm	Applied Biosystems, Bedford, USA
RP chromatographic packing	Zorbax SB-C18, 5 µm	Agilent Technologies, Santa Clara, CA, USA
Ultrafiltration tubes, 3 kDa	Amicon,	Merck Millipore, Cork, Ireland
Ultrafiltration tubes, 2 kDa	Vivacon 500®	Sartorius Stedim, Göttingen Germany
Ultrafiltration tubes, 10 kDa	Vivacon 500®	Sartorius Stedim, Göttingen Germany
Titanium dioxide TiO_2	Titansphere TiO	GL Sciences, Tokyo, Japan

3.1.7. Liquid chromatography-mass spectrometry consumables and equipment

Consumable	Name	Company	Used with
Ceramic capillary cutter		Thermo Fisher Scientific, Odense, Denmark	speLC
Diamond capillary cutter	Shortix	Scientific glass technology, Singapore	nanoAcquity UPLC system in DKFZ and ZMBH
ESI emitters	PicoTips	New objective, Woburn, USA	QTrap5500 and QTrap6500

Consumable	Name	Company	Used with
In-house packed analytical column	Tubing: 200 µm ID/350 µm OD, length 7 cm	Polymicro technologies; Molex, CM Scientific, Silsden, UK	speLC
	Chromatographic packing: RaprosilPurAQ 3 µm 120 C18 material	Dr. Maisch, Ammerbuch, Germany	
nanoAcquity analytical column	nanoAcquity UPLC column 75 µm ID, length 25 cm, 1,7 µm BEH130 C18 material	Waters, Milford, MA, USA	nanoAcquity UPLC system in DKFZ and ZMBH
nanoACQUITY trapping column	Symmetry C18 nanoACQUITY trap, 5 µm, 180 µm ID, length 2 cm, C18 material	Waters, Milford, MA, USA	nanoAcquity UPLC system in ZMBH
Stage tips	Stage tips C18 material, 200 µL	Thermo Fisher Scientific, Odense, Denmark	speLC
Syringes for direct injection to the MS instrument	Hamilton SYR 500 µL and 1mL	Hamilton, Reno, USA	QTrap5500 and QTrap6500
UPLC vials	Plastic LC vials	MZ-Analysentechnik, Mainz, Germany	nanoAcquity UPLC system in DKFZ and ZMBH

3.1.8. Other consumables

Product		Company
Aluminum foil		CeDo GmbH, Mönchengladbach, Germany
Cell culture dishes (10 cm Ø)		TPP, Trasadingen, Switzerland
Cell culture flasks (25 cm^2)		TPP, Trasadingen, Switzerland
Cell culture flasks (75 cm^2)		TPP, Trasadingen, Switzerland
Cell culture flasks (150 cm^2)		TPP, Trasadingen, Switzerland
Cell scraper		Sarstedt, Newton, NC, USA
Dynabeads®		Invitrogen GmbH, Karlsruhe, Germany
Falcon Tube (15mL and 50mL)		Nerbe Plus, Winsen, Germany
GammaBind™ Plus Sepharose™ beads		GE Healthcare, Uppsala, Sweden
Gloves		Microflex, Reno, NV, USA
IPG 24 cm strips, pH 3-10		GE Healthcare, Uppsala, Sweden
Mini-Protean® TGX™ Gels Any kD™, various combs		Biorad, Hercules, USA
Parafilm		Bemis, Neeah, WI, USA
Pipette tips 10, 200, 1000 µL		Starlab, Hamburg, Germany
Pipette tips 20, 200, 1000 µL, Diamond		Gilson, Bad Camberg, Germany
Plastic syringe (5,10 and 20 mL)		DB, Drogheda, Ireland
Polyvinylidene fluoride (PVDF) transfer membrane		Thermo Fisher Scientific, Waltham, USA
Reaction tubes:	1.5mL and 2 mL	Starlab, Hamburg, Germany
	Protein LoBind Tube (1.5 mL)	Eppendorf, Hamburg, Germany
	Safeseal Microcentrifuge Tubes, low binding (1.7mL)	Sorenson, Bioscience, Salt Lake City, UT, USA
Scalpel		Feather, Osaka, Japan
Western Blotting Filter Paper		Thermo Scientific, Rockford, IL, USA

3.1.9. Enzymes and markers

Name	Company
Alkaline phosphatase (AP), 20 U/µL	Roche, Mannheim, Germany
Precision Plus Protein™ Kaleidoscope marker	Biorad, Hercules, USA
Trypsin, suitable for protein sequencing	Sigma-Aldrich, Steinheim, Germany

3.1.10. Antibodies

Primary antibodies

Name	Company
mouse-anti-human HLA-A2 (monoclonal, clone BB7.2)	BD Biosciences, San Diego, CA, USA and a kind gift from Prof. Dr. Rammensee (University of Tübingen, Germany)
mouse-anti-mouse H-2D^{b} (monoclonal, clone 28-14-8)	BD Biosciences, San Diego, CA, USA
rabbit-anti-human MHC class I (monoclonal, clone EPR1394Y)	Epitomics, Burlingame, CA, USA

Secondary antibodies

Name	Company
goat-anti-mouse-IgG (polyclonal, horseradish peroxidase-conjugated)	Jackson ImmunoResearch Laboratories, PA, West Grove, USA
goat-anti-rabbit IgG (polyclonal, horseradish peroxidase-conjugated)	Rockland, Gilbertsville, PA, USA

3.1.11. Cell lines

Neme	Description	HPV-status	Source	Culture medium	Reference
BSM	human, B-LCL, suspension	negative	IHWG Cell Bank, Seattle, WA, USA	complete RPMI-P	
EA	human, B-LCL, suspension	negative	IHWG Cell Bank, Seattle, WA, USA	complete RPMI-P	
CaSki	human, cervical, adherent	HPV16 +	Kindly provided by Felix Hoppe-Seyler, F065, DKFZ	complete DMEM	(153)
SNU1000	human, cervical, adherent	HPV16 +	Korean Cell Line Bank, Seoul, South Korea	complete RPMI-H	(154)
SNU17	human, cervical, adherent	HPV16 +	Korean Cell Line Bank, Seoul, South Korea	complete RPMI-H	(154)

3.1.12. Cell culture basal media and supplements

Product	Company
Dulbecco's Modified Eagle Medium (DMEM)	Sigma-Aldrich, Taufkirchen, Germany
RPMI-1640 (1x) with L-glutamine	Sigma-Aldrich, Taufkirchen, Germany
Fetal Calf Serum (FCS)	Biowest, Nuaillé, France Sigma-Aldrich, Taufkirchen, Germany
HEPES	Life Technologies, Grand Island, NY, USA
Penicillin/Streptomycin-Solution (P/S)	Sigma-Aldrich, Taufkirchen, Germany
Sodium pyruvate	PAA Laboratories, Cölbe, Germany
Trypan blue solution	Sigma-Aldrich, Taufkirchen, Germany
Trypsin-EDTA solution	Sigma-Aldrich, Taufkirchen, Germany

3.1.13. Cell culture media

Medium	Components
Complete DMEM	DMEM, 1% (v/v) P/S, 10% (v/v) FCS
Complete RPMI-H	RPMI-1640, 1% (v/v) P/S, 10% (v/v) FCS, 25 mM HEPES, 25 mM $NaHCO_3$
Complete RPMI-P	RPMI-1640, 15% (v/v) FCS, 1% 100 mM sodium pyruvate

3.1.14. Synthetic peptides

All HPV, endogenous and mCMV lyophilized synthetic peptides with a purity of >95% were purchased from PSL GmbH (Heidelberg) and the DKFZ Peptide Core Facility (Dr. Rüdiger Pipkorn, Dr. Stefan Eichmüller). Peptides were dissolved in DMSO at 10 mg/mL and stored in small aliquots in -80°C. All HIV lyophilized synthetic peptides with >98% purity were purchased from Bio-Synthesis Inc. (Lewisville, TX, USA) and dissolved in 10% DMSO in RPMI-1640 medium at 2 mg/mL. They were stored in small aliquots in -20°C

3.1.14.1. HLA-A2 binding HPV16 E6 and E7 peptides

Name	Origin	Region	AA sequence
$E6_{9-17}$	HPV16 E6	9-17	FQDPQERPI
$E6_{9-19}$	HPV16 E6	9-19	FQDPQERPIKL
$E6_{25-33}$	HPV16 E6	25-33	ELQTTIHDI
$E6_{25-33}$ seq1	HPV16 E6	25-33	ELQTTIHEI
$E7_{7-15}$	HPV16 E7	7-15	TLHEYMLDL
$E7_{11-18}$	HPV16 E7	11-18	YMLDLQPE
$E7_{11-19}$	HPV16 E7	11-19	YMLDLQPET
$E7_{11-20}$	HPV16 E7	11-20	YMLDLQPETT
$E7_{11-21}$	HPV16 E7	11-21	YMLDLQPETTD
$E7_{12-19}$	HPV16 E7	12-19	MLDLQPET
$E7_{12-20}$	HPV16 E7	12-20	MLDLQPETT
$E7_{77-86}$	HPV16 E7	77-86	RTLEDLLMGT
$E7_{77-87}$	HPV16 E7	77-87	RTLEDLLMGTL
$E7_{78-86}$	HPV16 E7	78-86	TLEDLLMGT
$E7_{81-90}$	HPV16 E7	81-90	DLLMGTLGIV
$E7_{82-90}$	HPV16 E7	82-90	LLMGTLGIV

3.1.14.2. HLA-A2 binding HIV peptides

Name	Origin	Region	AA sequence
FK10	p15; pol polyprotein	56-65	FLGKIWPSYK
IV9	reverse transcriptase (RT); pol polyprotein	192-200	ILKEPVHGV
LI9	Pro; HIV protease, pol polyprotein	61-69	LVGPTPVNI
VL9	reverse transcriptase (RT); pol polyprotein	30-38	VIYQYMDDL
VL10	Nef protein	179-188	VLEWRFDSRL
VV9	p24; gag polyprotein, capsid protein	16-24	VLAEAMSQV
YI9	reverse transcriptase (RT); pol polyprotein	70-78	YTAFTIPSI
YL9	p24; gag polyprotein, capsid protein	51-59	YVDRFYKTL

3.1.14.3. HIV peptides with concurrent binding motives for HLA-A2 and -B7

Name	Origin	Region	AA sequence
AP10	reverse transcriptase (RT); pol polyprotein	182-191	AENREILKEP
AT9	p17; gag polyprotein, matrix protein	2-10	AVNPGLLET
EY11	p24; gag polyprotein, capsid protein	46-56	EPFRDYVDRFY
FF16	p15; pol polyprotein	1-16	FLGKIWPSHKGRPGNF
HN10	p24; gag polyprotein, capsid protein	72-81	HQAISPRTLN
IL10	p17; gag polyprotein, matrix protein	29-38	IEIKDTKEAL
KF13	p15; pol polyprotein	4-16	KIWPSHKGRPGNF
LQ12	p15; pol polyprotein	130-141	LRSLFGSDPSSQ
MV9	p24; gag polyprotein, capsid protein	5-13	MTNNPPIPV
QR14	p17; gag polyprotein, matrix protein	28-41	QLQPSLQTGSEELR
QS8	p17; gag polyprotein, matrix protein	42-49	QPSLQTGS
SE10	p15; pol polyprotein	3-12	SRPEPTAPPE
TE10	reverse transcriptase (RT); pol polyprotein	3-12	TEEKIKALVE

3.1.14.4. HLA-A2 binding endogenous peptides

Name	Origin	Region	AA sequence
HLA-A2 positive control (PC)1	p68 RNA helicase	96-104	YLLPAIVHI
HLA-A2 PC2	coatomer subunit gamma-1	147-155	AIVDKVPSV

3.1.14.5. H-2D^b and K^b peptides

Name	Origin	Region	AA sequence
HSV-1 peptide (H-2K^b binding)	HSV-1 glycoprotein B	499-506	SSIEFARL
mCMV peptide (H-2D^b binding)	mCMV M45 protein	985-993	HGIRNASFI
H-2D^b PC1	cyclin-dependent kinase inhibitor 1B	33-41	FGPVNHEEL
H-2D^b PC3	spectrin alpha chain	381-389	KALINADEL
H-2D^b PC4	mitotic spindle-associated MMXD complex subunit MIP18	144-152	AALENTHLL

3.1.15 Software

Name	Company
Adobe Illustrator CS5	Adobe Systems, San Jose, CA, USA
Adobe Photoshop CS5	Adobe Systems, San Jose, CA, USA
Analyst 1.5.2 and 1.6.2	Ab Sciex, Foster City, CA, USA
EndNote X7.2	Thomas Reuters, Philadelphia, PA, USA
FlexAnalysis 3.4	Bruker Daltonik, Bremen, Germany
FlexControl 3.4	Bruker Daltonik, Bremen, Germany

Name	Company
MaxQuant 1.5.2.8 (155, 156)	MPI of Biochemistry, Martinsried, Germany
nanoAcquity Console 1.42	Waters, Milford, MA, USA
MS Office 2010	Microsoft, Redmond, WA, USA
PRISM 5	GraphPad, LaJolla, CA, USA
Skyline 2.6.0 (157, 158)	University of Washington, Seattle, USA
Xcalibur 3.0.63	Thermo Fisher Scientific, Bremen, Germany

3.2. Methods

3.2.1. Cell culture methods

All cell lines were maintained in a sterile environment and only sterile equipment, materials, solutions and cell culture media were employed. Cells were kept under standard conditions in a humidified incubator at 37°C and under 5% CO_2. All cell lines were authenticated and regularly checked for contaminations by multiplex PCR (Multiplexion GmbH, Heidelberg).

3.2.1.1. Culturing of cell lines

Adherent cells were cultured in 75 cm^2 and 150 cm^2 flasks and suspension cell lines in 75 cm^2 flasks. For passaging of confluent adherent cell lines, medium was aspirated, cells washed with 1x PBS and incubated with trypsin-EDTA solution at 37°C until they detached. Trypsin was inactivated by adding approx. five-fold access of fresh medium. Cells were resuspended, transferred to 50 mL tubes, filled up with fresh medium and subjected to centrifugation at 1400 rpm for 5 min at room temperature (RT). Supernatant was discarded and cells were resuspended in fresh medium. The splitting ratio for CaSki and SNU17 cells was 1:5 to 1:10 and for SNU1000 1:3 to 1:6. Cells were split after they reached 85 – 100% confluence and cell culture media was exchanged every 2 to 4 days.

The suspension cell lines BSM and EA were resuspended by shaking to dissociate cell aggregates every day. For splitting and experiments, an aliquot with desired cell number was pelleted in a 15 mL or 50 mL tube at 1400 rpm, 5 min, RT. Cells were resuspended in the appropriate volume of medium. They were spit every 4 – 5 days.

3.2.1.2. Counting of cells

The cell suspension or its 1:10 dilution was mixed 1:1 with Trypan blue solution. Cell numbers were determined automatically using the Countess® automated cell counter or manually using a Neubauer counting chamber (0.1 mm depth). The number of cells determined with manual counting was calculated with the formula:

$$\text{Total cell number} = \frac{\text{counted cells}}{\text{number of large squares}} \times \text{dilution factor} \times \text{chamber factor } (10000)$$

3.2.1.3. Thawing and freezing of cells

For thawing, cryovials were taken from the liquid nitrogen tank, immediately incubated in a 37°C water bath and cells were then resuspended in a prepared 50 mL tube filled with warm (37°C) fresh medium. Cells were pelleted at 1400 rpm, 5 min, RT and supernatant was discarded. The washing step was repeated; cells were resuspended in 7 mL medium and seeded in a 25 cm^2 flask. The medium was exchanged 12 – 24 h afterwards.

In order to freeze cells, a cell suspension was pelleted at 1400 rpm, 5 min, RT; supernatant was discarded and cells were resuspended in freezing medium to result in a final concentration of up to 10^7 cells/mL. Freezing medium was pre-cooled cell line specific medium containing 10% DMSO and 10% FCS additionally to the required ingredients. Aliquots of 1 mL of cell suspension were distributed in cryovials, which were immediately placed into the pre-cooled cell freezing device (Mr. Frosty). The cell freezing device was left at -80°C for 3 days, when the cryovials were transferred to the liquid nitrogen tank for permanent storage.

3.2.1.4. External loading of BSM cells with the $E7_{11-19}$ peptide

In order to assess the immunoprecipitation (IP) protocol (see 3.2.3.1) and epitope extraction with ultrafiltration (see 3.2.4.1), an experiment was performed where HPV16- HLA-A2+ BSM cells were exogenously loaded with the peptide $E7_{11-19}$, which was described to be present on CaSki cells and cervical cancers from patients (146, 147). Cells were washed with 1x PBS and resuspended in 20 mL of fresh medium as described in section 3.2.1.1. Aliquots of $5x10^7$ cells were distributed in new 75 cm^2 flasks and different amounts of peptide (2 µg, 20 µg and 200 µg) were added to each flask. As the peptides were dissolved in DMSO, the first control was cells incubated with 20 µL DMSO, which corresponded to the highest amount of DMSO added to cells with 200 µg of peptide. The second control was untreated cells. After cells were incubated under standard cell culture conditions for 5 h, they were collected in 50 mL tubes and washed five times with 50 mL 1x PBS at 1400 rpm, 5 min, RT. Cell pellets were subjected to cell lysis for the IP experiment (see 3.2.3.1) after the last washing step.

3.2.2. Biochemical methods

3.2.2.1 Sodium Dodecyl sulfate Polyacrylamide Gel Electrophoresis (SDS-PAGE)

Equal volumes of sample and 2x SDS-PAGE loading buffer were mixed. They were heated at 95°C for 10 min in order to denature and reduce proteins in the sample. Samples and protein marker (Precision Plus Protein™ Kaleidoscope marker) were loaded on a SDS-PAGE gel in equal amounts. SDS-PAGE was performed in a Mini-Protean® Tetra cell using 1x SDS-PAGE running buffer at RT. The first 15 min of separation were conducted at 90 V, then the voltage was increased to 140 V for approximately 50 min or until proteins were separated.

3.2.2.2. Western Blot (WB) analysis

SDS-PAGE separated proteins were transferred from the gel to a polyvinylidene fluoride (PVDF) membrane, which was activated with 100% methanol for 1 min prior to transfer. The transfer was carried out in a Mini Trans-Blot® cell at 30 V for 1.5 h at 4°C in WB transfer buffer.

The membrane was blocked with 5% milk powder in 1x PBS at RT for 1 h, followed by incubation with the primary and secondary antibodies. The primary antibody was diluted 1:1500 in 1x PBS containing 1.5% milk powder (PBS/1.5%MP). Incubation was performed at RT for 1 h. Unbound primary antibody was washed away by incubating the membrane in the WB washing buffer three times for 10 min on a rocking platform. The secondary antibody was diluted 1:5000 in PBS/1.5%MP. Incubation was carried out at RT for 1 h. The membrane was washed as described above and subsequently incubated in 1 mL ECL WB substrate at RT in the dark for 4 min. Chemiluminescence was measured with a Fusion-SL imaging instrument. All incubation steps after protein transfer to the PVDF membrane were performed on a rocking platform, except the final incubation with the ECL WB substrate for WB development.

3.2.2.3. In-solution reduction, alkylation and trypsin digest of proteins

To investigate how proteins bind to the reverse phase material, dry protein fractions were resuspended in 40 µL 50% 2,2,2-trifluoroethanol (TFE) in water with rigorous mixing and incubated 1 h at 65ºC with 300 rpm mixing for protein denaturation. Subsequently, 3.5 µL of 100 mM dithiothreitol (DTT) in 25 mM ammonium bicarbonate were added. The sample was incubated for 30 min at 65 ºC with 300 rpm mixing, when 5 µL of 200 mM iodoacetamide (IAA) in 25 mM ammonium bicarbonate were added, followed by incubation for 30 min at 65 ºC with 300 rpm mixing in the dark. 180 µL of 25 mM ammonium bicarbonate and 0.2 µg of trypsin were added to the sample and incubated overnight at 37 ºC with 300 rpm mixing. The reaction was quenched by addition of 10% TFA in water for the final concentration of 1% TFA. Samples were analyzed by LC-MS^2.

3.2.3. Epitope isolation strategies

3.2.3.1. Immunoprecipitation (IP)

To directly detect naturally processed and presented HPV16 epitopes on the cell surface by MS, IP enrichment of MHC class I molecules was optimized. Various parameters were compared and optimized, such as the ratio of antibody and bead amounts, incubation times and input cell numbers. The optimal protocol is described below.

Cells were seeded in cell culture dishes and grown until they reached 90 – 100% confluence. They were grown under standard cell culture conditions as described in section 3.2.1.

2 mL of GammaBind™ Plus Sepharose™ bead suspension was pelleted at 5000 rpm, 3 min, RT in a swinging rotor, supernatant was discarded and 1 mL of 10% acetic acid was added. Beads were vigorously mixed and subjected to centrifugation at 5000 rpm, 3 min RT in a swinging rotor. Supernatant was removed and beads were washed five times with 1 mL 1x PBS by pelleting as

described above. Supernatant was removed and beads were resuspended in 1 mL 0.02% sodium azide in 1x PBS and stored at 4°C.

50 µL of bead suspension (corresponding to 25 µL dry breads) was washed three times with 500 µL IP washing buffer. Beads were coupled with 20 µg of the mouse-anti MHC I antibody in 500 µL IP washing buffer at RT for 2 to 3 h on a rotator. Beads were washed three times with IP washing buffer before incubation with the lysate.

Cells were washed two times with ice-cold 1x PBS. Suspension cells were harvested with centrifugation at 1400 rpm, 5 min, RT, supernatant was removed and cells were set on ice before 1 mL ice-cold IP lysis buffer was added. Cells were lysed by pipetting. Adherent cells from one 10 cm Ø cell culture dish were collected by scraping and lysed in 1 mL of the ice-cold IP lysis buffer. Cells from several dishes were pooled in one sample by transferring the lysate from one cell culture dish to the next cell culture dish without addition of new IP lysis buffer. The cell lysate was incubated on ice for 10 min and vortexing was performed every 3 min. Cell lysates were centrifuged for 30 min at 14000 rpm, at 4°C. Supernatants were pooled to achieve homogeneity and to be distributed equally to the coupled beads. Aliquots of supernatant were mixed with coupled Ab-beads and incubated 3 to 4 h at 4°C on the rotator. Beads were pelleted at 5000 rpm, 3 min, at 4°C, supernatant was removed and beads were washed twice with 500 µL ice-cold 1:1 diluted IP lysis buffer in 1x PBS, twice with 500 µL ice-cold IP washing buffer and three times with 500 µL ice-cold 10 mM Tris/HCl buffer, pH 8. Supernatant was removed and IP samples were stored at -80°C until further processing and analysis.

3.2.3.2. Epitope elution from an IP sample

Epitopes dissociate from the MHC I complex at a pH below 2.9, which is required for epitope dissociation as described elsewhere (148). Two elution buffers were tested, namely 10% acetic acid or 0.3% TFA in water, but the later was chosen for the standard protocol. The elution buffer was added to the IP sample immediately after the sample was taken from -80°C or immediately after the IP experiment. The elution buffer was added in a volume which corresponded to three volumes of the dry beads in an IP sample. Usually, 70 – 75 µL of elution buffer was used for one IP sample.

To assess IP protocol efficiency, two murine MHC I restricted peptides were added into the elution buffer, as internal standard controls, before the buffer was added to the IP sample. The peptides were SSIEFARL, originating from HSV-1, and FGPVNHEEL from mouse cyclin-dependent kinase inhibitor 1B, which were determined not to bind to HLA-A2 *in silico*.

3.2.3.3. Direct elution of MHC I bound peptides from the cell surface by acidic treatment

SNU17 cells were seeded in 10 cm Ø cell culture dishes and grown until they reached 90-95% confluence. They were grown under standard cell culture conditions as described in section 3.2.1. Every cell culture dish was washed extensively three times with 1x PBS. Cells were collected either by trypsin-EDTA detachment as described in 3.2.1.1 with three 1x PBS washing steps performed afterwards or they were collected by scraping on ice. All the remaining liquid was removed before 1

mL per 10^7 cells of ice-cold elution buffer composed of 10% acetic acid and 1x protease inhibitors in water was added. Cell suspension in the elution buffer was transferred in low binding Eppendorf tubes and incubated on ice for 10 min with rigorous mixing every 3 min. Subsequently, samples were centrifuged at 12000 g for 10 min at 4°C to pellet cells. Eventually, the supernatant was subjected to ultrafiltration with 2 kDa cut-off following the protocol described in 3.2.4.1 and then desalted using a Seppak cartridge, or it was directly subjected to desalting with the Seppak cartridge as described in section 3.2.4.2.2. The sample was vacuum dried prior to LC-MS^3 analysis or subjected to fractionation by isoelectric focusing.

3.2.4. Epitope enrichment and purification strategies

Several epitope extraction strategies were employed during this thesis project. They were combined in various experimental pipelines (Figure 10). This section describes the experimental procedures of every method, whereas pipelines are shortly outlined together with results in the Results section.

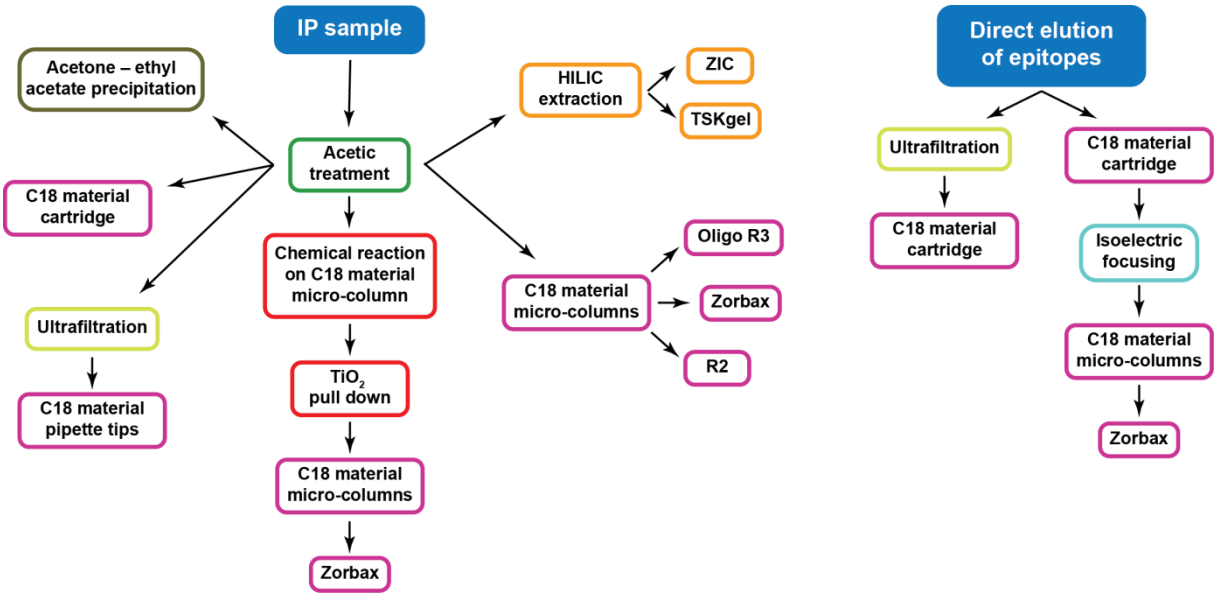


Figure 10. Schematic representation of epitope extraction pipelines.
Colors represent different chemical principles of extractions used in this thesis; green – acetic elution of epitopes from IP sample, army green – acetone – ethyl acetate precipitation, orange – hydrophilic liquid chromatography (HILIC), yellow – ultrafiltration, purple – all reverse phase (C18) methods, light blue – isoelectric focusing.

3.2.4.1. Ultrafiltration

The ultrafiltration step was performed in order to reduce IP sample complexity by removing proteins from the sample. Ultrafiltration was performed in a table-top centrifuge operating at 15000 g at RT. The duration of each centrifugation step depended on the pore size of the ultrafilter membrane. It ranged from 10 min for ultrafilters with a cut-off at 10 kDa to 30 min for those with a cut-off at 2 kDa. The membrane of the ultrafiltration device was first washed twice with 150 µL 50% MeOH/0.1% TFA in water. The device was centrifuged until all the liquid had passed the membrane in the last washing step. IP sample eluates were added on the ultrafilter membrane and centrifuged until all the liquid had

passed through the membrane. The remains of the IP sample on the membrane were washed twice with 150 µL 50% MeOH/0.1% TFA in water until all the liquid had passed through the membrane. The filtrate was transferred into a low binding Eppendorf tube and subjected to drying in a vacuum centrifuge. The sample was desalted with C18 extraction tips prior to LC-MS^3 analysis.

3.2.4.2. Reverse phase C18 material enrichment

To ensure reproducible results in LC-MS analysis, every sample needed to contain the minimum possible amount of contaminants, which were introduced during the preparation. The highest proportion of contaminants are usually various salts, which can be easily removed by reverse phase extraction. Peptides bind to the C18 material whereas salts are washed off. Several applications and chromatographic materials were used during this project.

3.2.4.2.1. C18 extraction pipette tip

IP eluates were directly subjected to C18 extraction with the pipette tip or dried ultrafiltrated IP samples were resuspended in 150 µL 3% ACN/0.1% TFA by vigorous mixing and incubation in an ultrasound bath for 5 – 10 min at RT before C18 material pipette tip extraction. Bond Elut OMIX pipette tips were wetted by aspirating and discarding 100 µL of 50% ACN/0.1% TFA three times. Tips were equilibrated three times by aspirating and discarding 100 µL 0.1% TFA. Peptides were bound to the C18 material in a tip during 35 cycles of aspirating and dispensing of the sample. Tips were rinsed three times by aspirating and discarding 100 µL 0.1% TFA. Peptides were eluted in a low binding Eppendorf tube containing 150 µL 50% ACN/0.1% TFA during 20 cycles of aspirating and dispensing of the liquid. The eluate from the C18 extraction pipette tip was subjected to drying in a vacuum centrifuge. The C18 material resin in the tip was not allowed to dry or to aspirate air in any of the steps.

3.2.4.2.2. Seppak C18 extraction cartridge

The Seppak C18 cartridge has a capacity to bind up to 20 mg of peptides. Due to its bigger size, larger volumes than 0.7 mL are required for the experimental procedure. The peptide solutions are applied onto the top of the cartridge and pressed through with an overpressure created by a 10 or 20 mL plastic syringe, which is mounted on the top of the cartridge.

Samples were prepared the same way as described in 3.2.4.2.1 using 800 µL 3%ACN/0.1% TFA instead. The cartridge was wetted three times with 1 mL 100% ACN, once with 1 mL 70% ACN/0.1% TFA and three times with 40% ACN/0.1% TFA. Wetting was followed by equilibrating three times with 1 mL 0.1% TFA. The sample was then loaded on the cartridge, the flow through collected in a low binding Eppendorf tube and dried in a vacuum centrifuge. The cartridge was washed three times with 1 mL 0.1% TFA and the sample was eluted with 1 mL 50% ACN/0.1% TFA into a low binding Eppendorf tube. The eluate was vacuum centrifuged to complete dryness. The flow rate was high during wetting, equilibration and washing steps, whereas it was slower than 1 drop/second for binding

and elution of the sample. All liquids used for wetting, equilibrating and washing of the Seppak C18 cartridge were discarded. The C18 material in the cartridge was not allowed to dry in any of the steps.

3.2.4.2.3. Self-packed reverse phase micro-columns

Micro-columns were prepared similarly as described in (159-161). Briefly, a small plug of the C8 material from the Empore Octyl C8 extraction disk was punched and packed into a 200 µL pipette tip. The Empore Octyl C8 material plug was later exchanged with a 2 mm x 2 mm piece of Kimtech cellulose tissue. A cellulose plug was cut out with a scalpel, wetted with a drop of 0.1% TFA in water and packed into a 200 µL pipette tip. The cellulose plug in the pipette tip was then washed with 80 µL of 0.1% TFA in water by applying the liquid on the top of the tip and pressing it through with an overpressure created by a 5 or 10 mL plastic syringe, which was mounted on the top of the pipette tip (Figure 11 A). Subsequently, a suspension of C18 chromatographic packing material (R2, Oligo R3 or Zorbax) in 50% ACN/0.1% TFA in water was added on the top of the plug and pressure was applied with a syringe to result in column packing. Usually, micro-columns with a length of 8 mm were produced. A batch of two to three samples was processed using a plastic syringe, whereas more samples were processed with a centrifuge for higher throughput. To this end, a micro-column was placed through a hole in the lid of a low binding Eppendorf tube with the tip facing down (shown in Figure 11 B) and placed in the centrifuge. All solutions (washing and equilibrating solvents and samples) were added on the top of the micro-column and span-down with slow speed. Solutions ran through the micro-column and were collected in the bottom of the Eppendorf tube. Spinning rates depended on the chromatographic material, but were never higher than 1200 g for washing and equilibration and 500 g for binding and elution of a sample. A micro-column was never allowed to dry out.

After the micro-column was assembled, it was washed once with 150 µL 50% ACN/0.1% TFA in water, once with 150 µL 0.1% TFA in ACN and once with 150 µL 50% ACN/0.1% TFA in water. It was equilibrated two times with 150 µL 0.1% TFA in water. Subsequently, a sample was added and the flow through was collected. The micro-column was washed two times with 150 µL 0.1% TFA in water and placed in a fresh low binding Eppendorf tube for elution of the sample from the micro-column. The elution was performed with 150 µL 30-35% ACN/0.1% TFA in water if not specified differently. The eluate and flow through were subjected to complete vacuum centrifuge drying before LC-MS^2/MS^3 analysis.

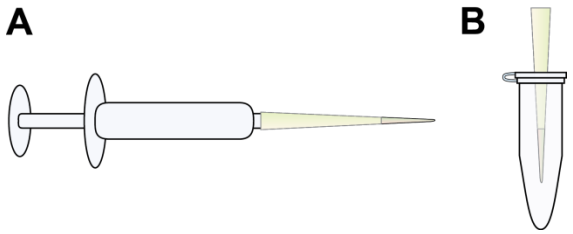


Figure 11. Illustration of a micro-column.
It can be mounted on the plastic syringe (A) or in an Eppendorf tube for loading in a centrifuge (B).

3.2.4.3 Hydrophilic interaction liquid chromatography (HILIC) enrichment

An IP eluate was vacuum centrifuged to complete dryness and resuspended in 100 µL 80% ACN/1% TFA in water with rigorous vortexing. A micro-column was prepared similarly as described in the previous subchapter 3.2.4.2.3 with some modifications required due to the chromatographic material chemistry. Either the C8 material from the Empore Octyl C8 extraction disk or the Kimtech cellulose tissue piece was used as a micro-column plug. A suspension of the TSKgel or the ZIC-HILIC chromatographic packing in 80% ACN/0.1% TFA in water was used instead of C18 material packing. A suspension of HILIC chromatographic packing was added on the top of the plug to produce an app. 6 mm long micro-column. The micro-column was washed once with 150 µL 0.1% TFA in water, once with 150 µL 50% ACN/0.1% TFA in water and equilibrated twice with 150 µL 80% ACN/0.1% TFA in water. A sample was added and the flow through fraction was collected. The micro-column was washed once with 150 µL 80% ACN/0.1% TFA in water and the wash fraction was collected. In usual HILIC experiments, the sample is eluted with 150 µL 30% ACN/0.1% TFA in water or with a solvent with a lower ACN content. For the purpose of this project, the sample was eluted gradually, starting with higher ACN contents in a solvent to lower ACN contents: 70% ACN/0.1% TFA → 60% ACN/0.1% TFA → 50% ACN/0.1% TFA → 40% ACN/0.1% TFA → 30% ACN/0.1% TFA → 20% ACN/0.1% TFA → 10% ACN/0.1% TFA → 0.1% TFA, all in water. Fractions were collected separately and a small aliquot was used for MALDI-MS analysis.

3.2.4.4. Acetone – ethyl acetate precipitation

In order to separate peptides from other proteins and the detergent in an IP sample, an acetone – ethyl acetate precipitation was performed. Acetone precipitation is used for protein purification (162), whereas ethyl acetate precipitation is described to remove detergents (163). The eluate of three BSM IP samples (200 µL) was mixed with 10 pmol of each HPV16 synthetic peptide from the list in 3.1.14.1, followed by the addition of 1400 µL of ice-cold acetone. An aliquot of 100 µL was taken for drying in a vacuum centrifuge. The rest of the sample was incubated at -20°C overnight. The sample was centrifuged at 20000 g at 4°C for 30 min to precipitate proteins. The supernatant was transferred to a new low binding Eppendorf tube. The pellet was completely dried in the vacuum centrifuge, whereas the supernatant was only partially dried until app. 50 µL of the solvent was left. The partially dried supernatant was subjected to the addition of 5.5 µL 10% DMSO, 0.85 µL 100% ACN and 110 µL ethyl acetate for precipitation of detergent.

The sample was immediately vigorously mixed for 90 s and subsequently centrifuged at 15000 g at RT for 5 min. 90% of the upper organic layer was transferred in a new Eppendorf tube. 100 µL of fresh ethyl-acetate was added into the sample tube, vigorously mixed by vortexing for 90 s and centrifuged at 15000 g at RT for 5 min. The upper layer was again added to the tube already containing the organic phase from the first extraction step. The bottom layer was transferred to a fresh low binding Eppendorf tube. Subsequently, samples were vacuum dried, resuspended in 20 µL 50 mM TEAB and analyzed with a MALDI-MS instrument. The workflow is illustrated in Figure 12.

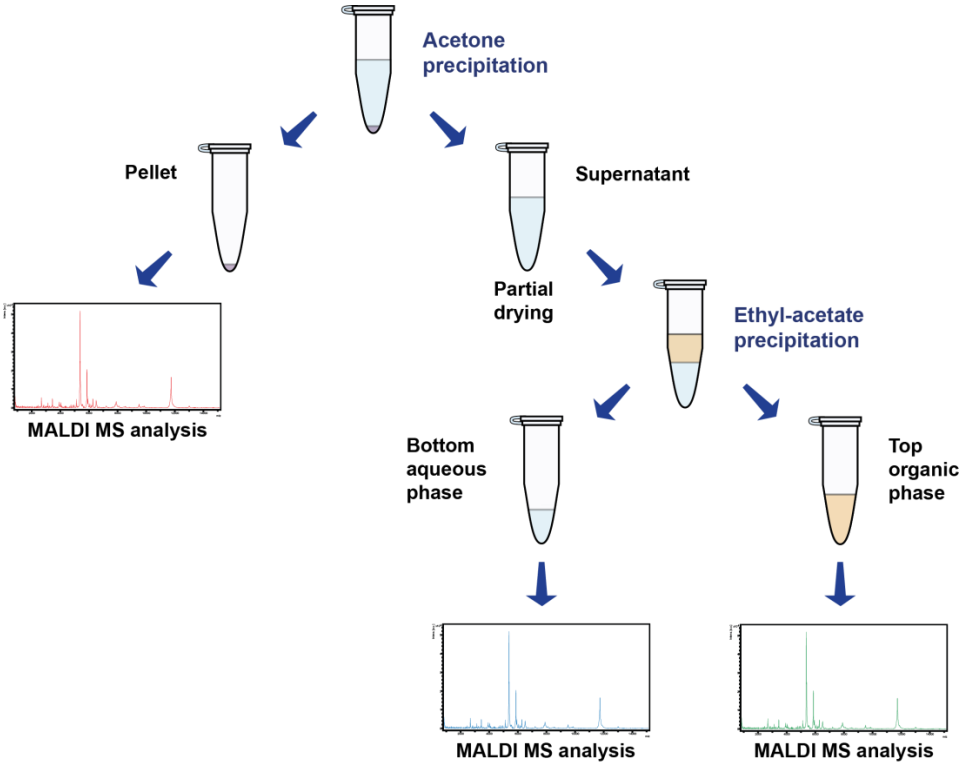


Figure 12. A schematic representation of the acetone – ethyl acetate precipitation workflow.
Dried IP eluate was mixed with acetone and left at -20°C overnight. After centrifugation, the pellet was dried in a vacuum centrifuge; whereas the supernatant was transferred into a new Eppendorf tube and partly dried. The partly dried supernatant was subjected to ethyl acetate precipitation, resulting in bottom aqueous and top organic phases. They were separated by transferring to new Eppendorf tubes and were subsequently vacuum centrifuged to complete dryness before analysis with MALDI-MS.

3.2.4.5. Chemical modification enrichment strategy

3.2.4.5.1. Chemical modification of primary amines by glyceraldehyde-3-phosphate (G3P) on a C18 micro-column

Synthetic peptides or IP eluates were bound and washed on a R2, Oligo R3 or Zorbax micro-column as described in section 3.2.4.2.3. The G3P reaction solution, containing 40 µL 0.01M TEAB, 30 µL stock G3P and 0.7 µL 0.6M $NaBH_3CN$, was prepared freshly for every experiment. The required pH 2 – 2.2 of the reaction mix was adjusted with a small volume of TEAB or G3P. The G3P reaction solution was added on top of the micro-column and a small aliquot (app. 10 µL) was pressed through with overpressure created by a syringe mounted on top of the micro-column. The micro-column was incubated at RT for 10 – 15 min before the reaction solution in the micro-column was refreshed by pressing a new aliquot through it. The cycles of incubation and reaction solution refreshment were repeated until the complete reaction solution was pushed through the micro-column, which took 1.5 to 2 h. The micro-column was washed with 0.1% TFA in water to remove the remaining reagents. Samples were eluted from the micro-column with 35% ACN/0.1% TFA in water and subjected to titanium dioxide (TiO_2) enrichment.

3.2.4.5.2. Titanium dioxide (TiO_2) enrichment and dephosphorylation of phospho-group containing peptides

TiO_2 enrichment was performed based on methods described in (164, 165). Briefly, corresponding amounts of 100% TFA, 100% ACN, water and glycolic acid were added to the micro-column eluate after G3P modification to result in final concentration of 80% ACN, 5% TFA and 1 M glycolic acid in water. 0.4 mg of dry TiO_2 beads were added to the sample and mixed vigorously for 15 min at RT. Subsequently, the TiO_2 beads were pelleted by a short spin in a centrifuge and the supernatant was transferred to new TiO_2 beads. The incubation was repeated two times with 0.2 mg TiO_2 beads each. They were pelleted and the supernatant was saved after the last incubation until after LC-MS^2/MS^3 analysis. If the reaction was successful, the supernatant was discarded. TiO_2 beads from all incubation steps were combined in a new low binding Eppendorf tube and washed once with 150 µL TiO_2 loading buffer, once with 150 µL TiO_2 washing buffer 1 and once with 150 µL TiO_2 washing buffer 2 by pelleting in the centrifuge. The supernatant was collected after every washing step. The TiO_2 beads were vacuum centrifuged to complete dryness and subjected to elution of peptides with 100 µL of TiO_2 elution buffer by vigorous mixing for 15 min at RT. TiO_2 beads were pelleted with a short spin in a table top centrifuge. Supernatant was transferred to a fresh Eppendorf tube and the elution from TiO_2 beads was repeated. Eluates were combined and 100% TFA was added for pH adjustment in the range between 8.5 and 9.2. Subsequently, 20 µL 10x alkaline phosphatase buffer and 0.2 µL alkaline phosphatase were added to the sample. The sample was incubated overnight at 37°C, at 400 rpm. Samples were dried in the vacuum centrifuge before they were analyzed with speLC-MS^2 or they were desalted on C18 material prior to LC-MS^3 analysis.

3.2.4.6. Isoelectric focusing (IEF) fractionation

In order to reduce the complexity of samples acquired during the direct elution of epitopes from the cell surface with acetic treatment, samples were fractionated by IEF. This was done on an Agilent 3100 OFFGEL Fractionator following the instructions of the device manufacturer. The fractionation was performed on a 24 cm immobilized pH gradient (IPG) gel strip with the pH ranging from 3 to 10. The strip was used together with a 24-well frame, resulting in 24 fractions at the end of the procedure. The IPG strip was rehydrated by adding 30 µL of rehydration buffer per well for 20 min. Each sample was desalted and vacuum dried before it was resuspended in a solution containing 0.72 mL water and 2.88 mL OFFGEL stock solution (1.25x) with vigorous mixing. Subsequently, 150 µL of the sample was added into each well. The fractionation was conducted with a current of 50 µA and a starting potential higher than 200 V. The fractionation was finished when the total energy exceeded 64 kWh, which took 36 to 48 h. Fractions were collected in fresh low binding Eppendorf tubes. Every well was incubated with 200 µL of the IEF washing solution for 35 min and the liquid was combined with the corresponding fraction. Fractions were subjected to complete dryness in a vacuum centrifuge and desalting using the Zorbax C18 material self-packed micro-columns before they were analyzed with LC-MS^3.

3.2.5. Liquid chromatography – mass spectrometry (LC-MS) instrumentation

Experiments described in this thesis were performed with three LC-MS platforms and one MALDI-TOF instrument. All LC-MS^3 experiments were performed with a nanoAquityUPLC system coupled to a QTrap5500 (ZMBH) or a QTrap6500 (DKFZ) mass spectrometer. LC-MS^2 experiments were done with a speLC system coupled to the Q-Exactive mass spectrometer, and MALDI TOF MS experiments were performed with the UltrafleXtrem instrument. All LC-MS^2 and MALDI TOF MS experiments were performed during a lab visit from October 2014 to February 2015 at the University of Southern Denmark, Odense, Denmark, in the Department for Biochemistry and Molecular Biology, in Prof. Dr. Martin R. Larsen's subgroup within the Protein Research group. Experimental and technical support was kindly given by Prof. Dr. Martin R. Larsen and Dr. Lasse G. Falkenby. Instrument specifications and measuring parameters for all platforms are described below.

3.2.5.1. Liquid chromatography separation

3.2.5.1.1. nanoAcquity UPLC system (DKFZ)

Dried, desalted samples were resuspended in 3% ACN/0.1% FA in water with 15 s vigorous mixing and subjected to incubation in an ultrasonic water bath for 10 min. Subsequently, samples were transferred into LC auto sampler vials. For every analysis, 10 µL were injected to the analytical column.

Chromatographic separation was performed on a nanoAquity BEH130 column with a particle size of 1.7 µm, an inner diameter (ID) of 75 µm and a length of 25 cm, filled with C18 material. Solvent A of the mobile phase was composed of 0.1% FA and 0.01% TFA in water. Solvent B consisted of 0.1% FA and 0.01% TFA in ACN. The flow rate during the separation was 300 nL/min. A steep linear gradient, starting with 3% mobile phase B and increasing to 10% by 1 min, was followed by a slower linear gradient, which started with 10% mobile phase B and increased to 40% at minute 50. The column oven temperature was set to 45°C.

3.2.5.1.2. nanoAcquity UPLC system (ZMBH)

Sample preparation was the same as described above for the nanoAcquity UPLC system at the DKFZ. For every analysis, 20 µL were injected to the pre-column.

After peptides were captured on the pre-column (nanoAcquity trap, particle size 5 µm, ID 180 µm, length 2 cm) at 99.5% mobile phase A with a flow rate of 10 µL/min for 7 min, they were chromatographically separated on the analytical column as described above. Solvent A was composed of 0.1% FA in water and solvent B of the mobile phase consisted of 0.1% FA in ACN. The flow rate during the separation was 225 nL/min. A steep linear gradient, starting with 3.5% mobile phase B and increasing to 15% by 2.8 min, was followed by a slower linear gradient, which started with 10% mobile phase B and increased to 40% at minute 70. The column oven temperature was set to 35°C.

3.2.5.1.3. Solid-Phase Extraction Capillary Liquid Chromatography (speLC)

The speLC system is a high throughput automated solid-phase extraction gradient LC system optimized for analysis of sub-complex samples with 100-1000 peptides. It employs gradients shorter than 10 min with 1 – 2 µL/min flow rates, which enables analysis of >100 samples per day. The system uses disposable micropipette solid phase extraction tips (StageTips) for sample loading. This application enables direct concentration and desalting on the StageTip, which makes sample preparation faster, as desalting of the sample is not required prior to loading (166).

Dried samples were resuspended in 5 µL 40% ACN/0.1% TFA in water, vigorously mixed for 15 s and diluted to a final volume of 100 µL with 0.1% TFA in water. The StageTip preparation steps and sample loading were performed by centrifugation in a spin centrifuge with a special adapter to hold 12 StageTips.
Each tip was first washed with 20 µL of 0.1% FA in ACN (solvent B), centrifuged for 20 s to push the liquid through the resin, followed by an equilibration with 20 µL of 0.1% FA in water (solvent A) and a centrifugation for 20 s. An aliquot of sample (usually 5 µL) was loaded on the tip together with additional 20 µL of solvent A and centrifuged for 30 s. Subsequently, 40 µL of solvent A were added on the tip and centrifuged for 20 s, such that a volume of liquid was left above the resin to prevent its drying (166).

SpeLC separation was performed with an in-house prepared column (ID 200µm, length 7 cm) packed with 3µm ReprosilPur-AQ 120 C18 material.
Separation started with an initial 1 min equilibration of the column with 4% solvent B of the mobile phase, followed by a linear gradient to 35% solvent B in 5 min. The flow rate was 1.5 µL/min.

3.2.5.2. Mass spectrometry analysis

3.2.5.2.1. QTrap5500 and QTrap6500 mass spectrometers

QTrap instruments are hybrid triple quadrupole-linear ion trap low resolution mass spectrometers mostly applied for targeted MS^2 and MS^3 analysis. The 6500 series is a newer model with improved sensitivity and faster LIT scanning speeds compared to the 5500 series. Thus, the QTrap6500 instrument allows targeted analysis of more analytes. Furthermore, the QTrap6500 permits analysis with a low mass profile, measuring until 1000 m/z, and a high mass profile, measuring until 2000 m/z, whereas the Qtrap5500 measures only until 1000 m/z.

3.2.5.2.1.1. Manual optimization of MS^3 spectra of synthetic peptides

In order to achieve the highest signals in the MS^3 spectra, peptide specific parameters, such as collision energy (CE), declustering potential (DP) and excitation energy (AF2), were adjusted for each peptide precursor and the most intense fragments comprised of ≥ 5 amino acids. The declustering potential is an electric potential between the orifice plate (MS instrument entrance) and the ground in the skimmer, which is located behind the orifice plate in the MS instrument. The DP minimizes the solvent cluster ions, which may be attached to the analyte. The CE accelerates the analyte into the

collision cell (2nd quadrupole) for the effective first fragmentation. The AF2 is applied to the MS^2 fragment in the linear ion trap (3rd quadrupole) to generate the MS^3 spectrum. To create a scheduled LC-MS^3 method, the retention time of each peptide was determined after manual MS^3 optimization.

For these optimizations, synthetic peptides were diluted to a final concentration of 0.5-2.0 µg/mL in 50% ACN/0.1% FA and injected into the QTrap5500 or QTrap6500 instrument with the built-in syringe pump maintaining a constant flow of 5 to 10 µL/min. The positive electron spray voltage on the TurboV ion source was set to 5500 V. The curtain gas was set at 20 or 30 psi, ion source gas at 15 psi, collision gas (CAD) high and the interface heater temperature at 150°C. The resolution of the first quadrupole (Q1) was set at low or unit resolution for the QTrap 5500 or the QTrap 6500, respectively. The linear ion trap (LIT) was set to LIT resolution. The LIT scan rate was 1000 Da/s. The QTrap6500 was operated in the high mass hardware profile measuring until 2000 m/z.
CE and DP were optimized using the manual compound optimization script in the Analyst 1.5.2 or the Analyst 1.6.2 program for the QTrap5500 or the QTrap6500, respectively. The script chooses defined numbers of the most intense fragments in the MS^2 spectra and determines the optimal energies required for the highest signals. The average parameter values for DP and CE of the three optimization replicates were used to optimize the AF2 value. The optimal AF2 value was determined by ramping of the potential from 0.0 to 0.2 V. The used AF2 was determined when the MS^2 precursor intensity was reduced to 5% of its starting intensity.
The DP, CE and AF2 parameters of a minimum of three fragments per peptide (except H-$2D^b$ PC1 and HSV-peptide for the QTrap5500) were used for building the LC-MS^3 method and determination of the retention time for scheduling of peptides. The optimized parameters for QTrap5500 and 6500 for different projects are listed in Tables 1 – 5.

3.2.5.2.1.2. QTrap5500 and QTrap6500 parameters for LC-MS^3 analysis

The mass spectrometers were equipped with a NanoESI III electron spray ionization source with PicoTip ESI emitters. Analysis was performed in positive mode with both instruments. Source voltage was set at 2700 V or 3000 V and curtain gas at 20 psi or 30 psi for QTrap5500 and QTrap6500, respectively. Ion source gas was at 15 psi, CAD high and interface heater temperature at 150°C. The resolution of the Q1 was set to low or unit resolution for QTrap5500 and QTrap6500, respectively, and the linear ion trap was set to LIT resolution. LIT scan rate was 10000 Da/s, fill time was dynamic and MS^3 excitation time was 25 ms. The QTrap6500 operated in the low mass hardware profile. Manually optimized MS^3 parameters listed in Tables 1 – 5 were used to create LC-MS^3 methods.

Peptide name	Peptide sequence	Time (min)	MS^1 precursor	MS^2 fragment	MS^2 fragment mass	DP	CE	AF2
E7 7-15	TLHEYMLDL	44.9	567.78	b_8^{2+}	502.23	86	19	0.06
				b_6	775.34	86	23	0.12
				b_7	888.43	86	19	0.11
E7 11-19	YMLDLQPET	39.4	555.26	b_5	636.30	105	20	0.09
				b_6	764.36	105	19	0.09
				b_8	990.46	105	19	0.11
E7 11-20	YMLDLQPETT	38.9	605.79	b_5	636.30	84	19	0.09
				b_6	764.36	84	19	0.10
				b_8	990.46	84	17	0.11
E7 12-20	MLDLQPETT	34.0	1047.50	y_5	575.27	208	48	0.08
				b_5	601.30	208	45	0.08
				b_7	827.40	208	42	0.10
				b_8	928.44	208	43	0.11
E7 82-90	LLMGTLGIV	54.3	916.55	b_7-H_2O	668.38	70	43	0.08
				b_8-H_2O	781.46	70	39	0.10
				b_8	799.47	70	35	0.09
H2-Db PC1	FGPVNHEEL	31.5	521.25	y_7^{2+}	419.20	100	27	0.10
				y_8	894.43	100	26	0.12
HSV-1 peptide	SSIEFARL	39.1	461.75	y_5	635.35	90	30	0.12
				y_6	748.44	90	22	0.12
HLA-A2 PC1	YLLPAIVHI	53.7	519.82	y_8	875.57	116	24	0.12
				y_6	649.40	116	19	0.10
				y_7	762.49	116	19	0.10
HLA-A2 PC2	AIVDKVPSV	32.4	464.28	y_6-H_2O	626.35	66	18	0.15
				y_6	644.36	66	19	0.09
				y_7	743.43	66	18	0.10

Table 1. List of parameters that were used to create a LC-MS^3 method for targeted analysis with Qtrap5500.

Peptide name	Peptide sequence	Time (min)	MS^1 precursor	MS^2 fragment	MS^2 fragment mass	DP	CE	AF2
E6 25-33	ELQTTIHDI	28.0/ 20.0	535.28	y_7-H_2O^{2+}	405.21	78	25	0.05
				MH-H_2O^{2+}	526.27	78	15	0.05
				y_6	699.37	78	23	0.09
				y_7	827.43	78	21	0.08
E7 7-15	TLHEYMLDL	39.4/ 31.1	567.78	b_8^{2+}	502.23	86	19	0.05
				b_6	775.34	86	19	0.10
				b_7	888.43	86	19	0.10
E7 11-18	YMLDLQPE	35.0/ 26.8	1008.47	b_5	636.30	10	47	0.08
				b_6	764.36	10	39	0.09
				b_7	861.42	10	41	0.09
				MH-H_2O	990.46	10	37	0.09
E7 11-19	YMLDLQPET	35.0/ 26.8	555.26	b_5	636.30	111	15	0.08
				b_6	764.36	111	13	0.10
				b_8	990.46	111	14	0.11
E7 11-19	YM(Ox)LDLQPET	30.9/ 22.7	563.26	b_5	652.30	111	15	0.08
				b_6	780.36	111	13	0.10
E7 11-20	YMLDLQPETT	34.7/ 26.4	605.79	y_5	575.27	86	19	0.08
				b_5	636.30	86	15	0.09
				b_6	764.36	86	15	0.09
				b_8	990.46	86	17	0.09
E7 11-20	YM(Ox)LDLQPETT	30.5/ 22.3	613.78	y_5	575.27	86	19	0.08
				b_5	652.30	86	15	0.09
				b_6	780.36	86	15	0.09
E7 11-21	YMLDLQPETTD	34.3/ 26.1	663.30	y_5-H_2O	544.22	100	33	0.09
				y_5	562.24	100	21	0.08
				b_5	636.30	100	23	0.08
				b_6	764.36	100	21	0.09
E7 12-20	MLDLQPETT	29.8/ 21.7	1047.50	y_5	575.27	216	47	0.07
				b_5	601.30	216	45	0.08
				b_7	827.40	216	43	0.09
				b_8	928.44	216	43	0.10
E7 77-87	RTLEDLLMGTL	49.1/ 41.1	631.34	b_{10}^{2+}	565.80	81	19	0.07
				MH-H_2O^{2+}	622.34	81	21	0.09
				b_7	841.48	81	29	0.11
				b_8	972.52	81	29	0.12
E7 82-90	LLMGTLGIV	47.5/ 39.5	916.55	b_6-H_2O	611.36	70	43	0.09
				b_7-H_2O	668.38	70	45	0.09
				b_8-H_2O	781.46	70	39	0.09
				b_8	799.47	70	35	0.08
HLA-A2 PC1	YLLPAIVHI	47.3/ 39.0	519.82	y_5	552.35	81	27	0.08
				y_6	649.40	81	21	0.09
				y_7	762.49	81	19	0.09
HLA-A2 PC2	AIVDKVPSV	27.9/ 19.8	464.28	y_6-H_2O	626.35	66	17	0.09
				y_6	644.36	66	19	0.09
				y_7	743.43	66	17	0.09

Table 2. List of parameters that were used to create a LC-MS^3 method for analysis of the HLA-A2 binding HPV16 derived peptides with Qtrap6500.

Peptide name	Peptide sequence	Time (min)	MS^1 precursor	MS^2 fragment	MS^2 fragment mass	DP	CE	AF2
FK10	FLGKIWPSYK	34.9	619.85	y_8	978.54	120	31	0.12
			413.57	y_8^{2+}	489.77	70	19	0.08
				b_9^{2+}	546.80	70	19	0.08
				y_5	680.34	70	23	0.10
IV9	ILKEPVHGV	23.3	496.30	b_7^{2+}	409.25	75	25	0.06
				y_5	508.29	75	31	0.09
			331.20	y_8^{2+}	439.76	60	15	0.10
				y_6-H_2O	619.32	60	17	0.12
LI9	LVGPTPVNI	34.4	909.54	y_6	640.37	150	45	0.11
				b_7	664.40	150	45	0.10
				y_7	697.39	150	45	0.11
				b_8	778.45	150	42	0.10
VL9	VIYQYMDDL	38.7	1159.53	b_5	667.35	150	48	0.11
				y_6	784.32	150	49	0.09
				b_6	798.38	150	47	0.10
				y_7	947.38	150	48	0.11
VL9	VIYQYM(Ox)DDL	30.8	1175.53	b_5	667.35	150	50	0.10
				b_6	814.38	150	47	0.10
				y_6	800.32	150	49	0.09
				y_7	963.38	150	48	0.11
VL10	VLEWRFDSRL	38.5	440.91	y_7^{2+}	490.26	80	24	0.07
				y_8-H_2O^{2+}	545.77	80	24	0.07
				y_8^{2+}	554.78	80	23	0.08
				y_9^{2+}	611.32	80	23	0.11
VV9	VLAEAMSQV	30.1	947.49	b_6	615.32	180	43	0.10
				b_8-H_2O	812.40	180	41	0.10
				y_8-H_2O	830.41	180	39	0.10
				MH-H_2O	929.48	180	39	0.11
VV9	VLAEAM(Ox)SQV	22.1	963.49	b_6	631.32	180	43	0.10
				b_8-H_2O	828.40	180	41	0.10
				y_8-H_2O	846.41	180	39	0.10
				MH-H_2O	945.48	180	39	0.11
YI9	YTAFTIPSI	46	1012.53	b_5-H_2O	566.26	160	52	0.09
				b_6-H_2O	679.34	160	41	0.11
				b_6	697.35	160	41	0.09
YL9	YVDRFYKTL	29.4	602.82	b_8^{2+}	537.27	120	31	0.09
				b_7	972.49	120	29	0.13
			402.22	y_7^{2+}	471.76	80	17	0.08
				y_8^{2+}	521.29	80	19	0.10
HLA-A2 PC1	YLLPAIVHI	47.3	519.82	y_5	552.35	81	27	0.08
				y_6	649.40	81	21	0.09
				y_7	762.49	81	19	0.09
HLA-A2 PC2	AIVDKVPSV	27.8	464.28	y_6-H_2O	626.35	66	17	0.09
				y_6	644.36	66	19	0.09
				y_7	743.43	66	17	0.09

Table 3. List of parameters that were used to create a LC-MS^3 method for analysis of the HLA-A2 binding HIV derived peptides with Qtrap6500.

Peptide name	Peptide sequence	Time (min)	MS^1 precursor	MS^2 fragment	MS^2 fragment mass	DP	CE	AF2
AP10	AENREILKEP	22.2	599.82	b_9-H_2O^{2+}	533.29	125	31	0.09
				b_9^{2+}	542.29	125	29	0.10
				b_7	826.44	125	31	0.12
			400.22	y_8^{2+}	499.79	55	17	0.08
AT9	AVNPGLLET	30.9	457.25	y_6-H_2O	611.34	51	19	0.10
				y_6	629.35	51	13	0.11
				b_7	665.40	51	15	0.10
				b_8	794.44	51	14	0.10
EY11	EPFRDYVDRFY	36.5	753.85	y_{10}^{2+}	689.33	195	39	0.10
				MH-H_2O^{2+}	744.85	195	39	0.11
			502.90	MH-H_2O^{3+}	496.90	70	13	0.07
				b_5-H_2O	627.29	70	21	0.13
FF16	FLGKIWPSHKGRPGNF	31.9	614.34	y_{11}^{2+}	641.83	120	31	0.09
				y_{12}^{2+}	698.37	120	29	0.10
				y_7	775.42	120	33	0.13
				y_{14}^{2+}	790.92	120	29	0.12
HN10	HQAISPRTLN	19.7	568.81	MN-NH_3^{2+}	560.30	130	27	0.09
				y_6	687.38	130	29	0.11
				y_8-NH_3	854.47	130	33	0.12
				y_8	871.50	130	28	0.12
IL10	IEIKDTKEAL	25.1	580.33	y_7	804.45	101	31	0.12
				y_8	917.53	101	27	0.13
			387.22	y_8^{2+}	459.27	40	17	0.09
				b_9^{2+}/y_9-H_2O^{2+}	514.78	40	19	0.09
KF13	KIWPSHKGRPGNF	22.6	508.61	MH-NH_3^{3+}	502.94	90	23	0.09
				y_{10}^{2+}	548.79	90	27	0.09
				y_{11}^{2+}	641.83	90	23	0.10
				y_6	647.33	90	29	0.12
				y_{12}^{2+}	698.37	90	29	0.10
LQ12	LRSLFGSDPSSQ	30.4	647.33	b_5	617.38	101	29	0.11
				b_7	761.43	101	28	0.11
				b_8	876.46	101	32	0.10
MV9	MTNNPPIPV	32.5	491.76	y_7-NH_3	733.39	40	21	0.11
				b_7-NH_3	751.34	40	17	0.10
				b_7	768.37	40	13	0.11
QR14	QLQPSLQTGSEELR	28	793.41	y_{11}^{2+}	608.81	140	35	0.10
				y_{12}^{2+}	672.84	140	31	0.09
				MH-H_2O^{2+}	784.41	140	33	0.11
				y_8	919.45	140	37	0.12
QS8	QPSLQTGS	17.4	409.21	b_5	554.29	40	13	0.09
				b_6-H_2O	637.33	40	19	0.10
				b_6	655.34	40	13	0.10
				b_7	712.36	40	13	0.10
SE10	SRPEPTAPPE	17.6	540.77	a_7-NH_3	694.35	75	29	0.13
				a_7	711.38	75	33	0.12
				b_7	739.37	75	23	0.13
				y_7	740.35	75	21	0.14
TE10	TEEKIKALVE	25.1	580.33	a_9^{2+}	492.81	101	31	0.08
				b_9^{2+}	506.81	101	27	0.07
				b7-H2O	782.44	101	29	0.12
			387.22	y_8^{2+}	465.29	30	23	0.07
HLA-A2 PC1	YLLPAIVHI	47.3	519.82	y_5	552.35	81	27	0.08
				y_6	649.40	81	21	0.09
				y_7	762.49	81	19	0.09
HLA-A2 PC2	AIVDKVPSV	27.8	464.28	y_6-H_2O	626.35	66	17	0.09
				y_7	743.43	66	17	0.09
				y_6	644.36	66	19	0.09

Table 4. List of parameters that were used to create a LC-MS^3 method for analysis of HIV peptides with concurrent binding motives for HLA-A2 and -B7 with the Qtrap6500.

Peptide name	Peptide sequence	Time (min)	MS^1 precursor	MS^2 fragment	MS^2 fragment mass	DP	CE	AF2
mCMV peptide	HGIRNASFI	18.7	507.78	a_8^{2+}	428.23	106	29	0.07
				b_8^{2+}	442.23	106	25	0.06
				b_7-H_2O	718.37	106	31	0.09
				y_8	877.49	106	27	0.12
HSV-1 peptide	SSIEFARL	24.6	461.75	MH-H_2O^{2+}	452.75	86	17	0.07
				y_5	635.35	86	21	0.09
				y_6	748.44	86	19	0.09
H2-Db PC1	FGPVNHEEL	19.4	521.25	y_7^{2+}	419.20	81	23	0.06
				b_8^{2+}	455.70	81	21	0.06
				y_7	837.41	81	23	0.10
H2-Db PC3	KALINADEL	22.9	493.78	MH-H_2O^{2+}	484.77	61	15	0.06
				b_7	726.41	61	21	0.09
				b_8	855.46	61	17	0.10
H2-Db PC4	AALENTHLL	22.9	491.27	y_7^{2+}	420.23	61	19	0.06
				y_5	597.33	61	21	0.07
				y_6	726.38	61	21	0.08
				y_7	839.46	61	19	0.10

Table 5. List of parameters that were used to create a LC-MS^3 method for analysis of H-$2D^b$ binding peptides with the Qtrap6500.

3.2.5.2.2. Q Exactive mass spectrometer

The Q Exactive is a hybrid quadrupole-orbitrap mass spectrometer which uses a quadrupole for precursor ion selection, and a high – resolution accurate mass (HRAM) orbitrap for detection. It can be used for targeted and untargeted MS analysis.

Measurements were performed using a Top2 data dependent acquisition method as described in (166). The instrument was measuring all ions in MS^1 mode (full-scan) and performed MS^2 measurement for two of the most intense doubly or triply charged ions from full-scan. The instrument was operating with mass resolutions of 35000 and 17500 for the full-scan and MS^2 measurement, respectively. Mass ranges were set to 400 to 1600 m/z, fill times were set to 50 ms for the full scan and to 100 ms for MS^2. Target values were set to $3x10^6$ for full scan and to $5x10^5$ for MS^2.

3.2.5.2.3. UltrafleXtreme mass spectrometer

The UltrafleXtreme instrument is a high resolution accurate mass matrix-assisted laser desorption/ionization – time-of-flight mass spectrometer (MALDI-TOF MS) equipped with a collision cell for MS^2 analysis. MS^2 analysis can also be performed in the LIFT mode where elevated laser intensity is used for ionization causing analyte fragmentation.

0.5 µL of MALDI matrix solution and 0.5 µl of sample were mixed on the MALDI target and dried before analysis. Protein measurements were performed in positive linear MS^1 mode in a mass range of 1000-25000 m/z and peptides were measured in positive reflector MS^1 mode in a mass range of 300-1500 m/z with mass resolution >10000. 500 laser shots with 60% laser power were accumulated per spectrum in both modes.

3.2.6. Data analysis and data processing

3.2.6.1. MS^2 data analysis with the MaxQuant software

The MaxQuant 1.5.2.8 software (155, 156) was used for spectrum identification and quantification of the experiments assessing the binding of proteins from the IP sample to the Seppak cartridge or the Zorbax micro-columns. Samples of all fractions and replicates were searched against human and mouse sequences from the Uniprot database from NCBI (retrieved December 10th 2014) and compared between samples, meaning that a peptide identified in one LC-MS analysis was searched for in other LC-MS analysis on MS^1 and MS^2 level. The mouse sequence was used to identify anti-HLA-A2 antibody derived peptides, whereas the human sequence was applied for the identification of other cell components. Database search was performed with trypsin as the digestion enzyme with zero miscleavages, carbamidomethyl as fixed modification on cysteine, oxidation on methionine as variable modification and with 1% false discovery rate. Intensities of unique peptides present in ≥ half of the samples and having a MaxQuant score ≥60 were used for quantification. Protein quantification was performed by summing peptide intensities of ≥ three unique peptides per protein.

3.2.6.2. MS^1 quantification with the Skyline software

Data processing with the Skyline program (157, 158) was performed for experiments analyzed with the HRAM Q Exactive instrument, allowing quantification of peptides on the MS^1 level. The program searched against the HLA-A2 HPV16 E6 and E7 peptides listed in 3.1.14.1. The specified chemical modifications were set as variable. The program extracted, assigned and integrated chromatographic peaks of ion currents from peptides and their modified counterparts, with all possible modification combinations. All peak assignments were manually corrected. The main criteria for manual peak selection were retention time and mass accuracy, which was ≤4 ppm. Most peptides were present in two different charge states; therefore their co-elution was a further criterion. The instrument resolved the isotopic distribution of peptide precursor ion, which was also considered during manual peak selection.

3.2.6.3. MS^3 data analysis

MS^3 data analysis was performed manually by comparing MS^3 spectra of IP samples with MS^3 spectra of synthetic peptides with the Analyst 1.5.2 and 1.6.2 programs for the QTrap5500 and the QTrap6500, respectively. Peak areas were determined using the quantification function of the Analyst software. Peak areas of transitions from the same peptide were summed and used for further calculations.

To assess IP protocol efficiency and cell input number for optimal IP yield, the signal areas of the naturally presented HLA-A2 binding peptides YLLPAIVHI and AIVDKVPSV were normalized to the signal areas of synthetic peptides SSIEFARL (originating from HSV-1 glycoprotein B) and FGPVNHEEL (from mouse cyclin-dependent kinase inhibitor 1B), which were added to the IP elution buffer. Both synthetic peptides are murine MHC class I restricted and were determined to be non binders to HLA-A2 *in silico*.

3.2.6.3. MS^1 MALDI TOF data analysis

The program FlexControl 3.4. was used for UltrafleXtreme instrument operation and data acquisition, FlexAnalysis 3.4 was used for data processing and visualization. Spectra were manually inspected for the presence of proteins and peptides and their masses manually assigned. Mass spectra with subtracted baseline were used for overlaid representation.

4. Results

The first part of the results section describes the optimization of epitope isolation with immunoprecipation (IP), followed by the results of different epitope extraction, enrichment and purification strategies from the IP. The IP epitope isolation is compared with direct elution of peptides from cell surfaces by acetic elution. All tested workflows are illustrated in Figure 10 and were first optimized to directly identify HLA-A2 HPV16 E6 and E7 T cell epitopes. They were later applied to two other viral infection models as a validation. Those results are described in the final part of the chapter.

4.1. Optimization of the immunoprecipitation (IP) protocol

To be able to directly detect only HLA-A2 binding peptides from HPV16 transformed cells by LC-MS, the immunoprecipitation protocol was optimized first. Immunoprecipitation is a method that isolates a protein or protein complexes of interest from a biological sample. An antibody (Ab), which is immobilized on protein A or G coated sepharose or magnetic beads, binds a particular antigen. Isolated bead-Ab-protein complexes are pelleted in the centrifuge or on a magnet, and the remaining biological sample is removed. The protocol can be performed in two ways. (1) The Ab is first incubated with the sample to bind target proteins and then beads are added to bind Ab-antigen complexes or (2) the Ab is first bound to the beads and Ab-bead complexes are incubated with the sample to affinity purify target proteins. In the first approach, the target protein remains longer within the biological matrix and thereby the chance for protein degradation by endogenous proteases is increased. Therefore, the second approach was selected. Optimization of the IP protocol was driven by the idea that every step has to provide the highest yields, which can only be obtained by determination of the right Ab:antigen:beads ratios to capture all target molecules from the sample, optimal incubation times and optimal buffers.

4.1.1. Optimization of coupling an antibody to sepharose beads

The first step in our IP protocol is the binding of the mouse anti-HLA-A2 antibody BB7.2 to sepharose beads coated with protein G. G protein is a streptococcal protein, which binds the Fc portion of IgG antibodies with high affinity. To assess the optimal amount of Ab for saturation of the beads, but on the other hand not losing Ab due to oversaturation, a titration experiment was performed. Furthermore, the Ab-bead binding kinetics were determined for estimation of the optimal time needed for efficient coupling. To do so, 7 aliquots of 25 µL of pelleted beads were incubated with increasing amounts (5, 10, 25, 50, 70 and 100 µg) of the mouse anti-HLA-A2 antibody. A small aliquot of the supernatant from every sample was collected at the beginning of the incubation, 30 min, 1 h, 2 h and 3 h later and analyzed by Western blot.

The bands corresponding to the Ab heavy chain at 50 kDa were not detected after 3 h incubation in supernatant when 5 or 10 µg Ab were used, whereas a slight band was detected in the supernatant when 25 µg Ab was used (Figure 13 A). Bands from 50, 70 and 100 µg Ab were present in the samples after 3 h incubation, but their intensities were decreased compared to the beginning of the

incubation, meaning that beads bound some of the Ab but not all because of oversaturation (Figure 13 A). It was concluded that 25 µL of dry breads bind approx. 20 µg of Ab.

The time course of the binding reveals that incubation for 2 h at RT is long enough to bind 5 and 10 µg of Ab and that saturation of beads was reached in 2 h also for 25 µg of Ab as the band intensities after 2 h and 3 h incubation were comparable (Figure 13 B).

From these results the optimal coupling conditions were determined to be 20 µg anti-HLA-A2 Ab with 25 µl dry beads incubated for 2 h to 3 h at RT in the IP washing buffer.

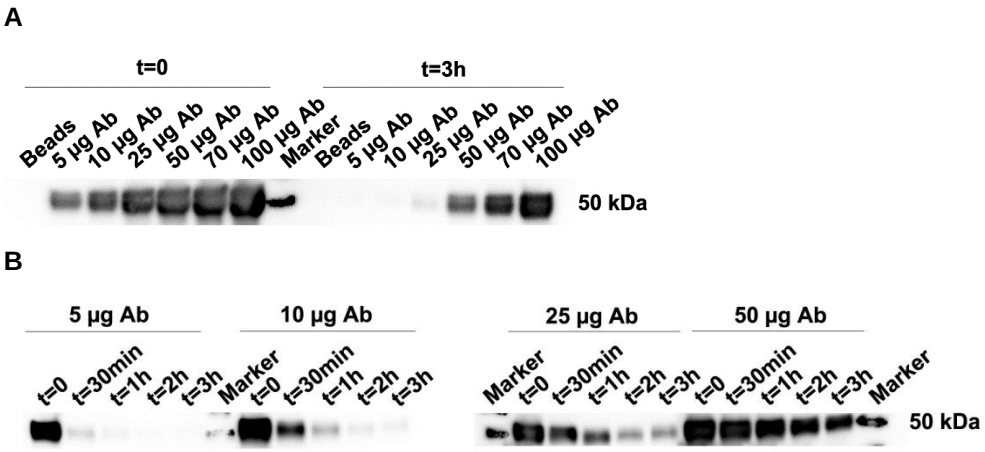


Figure 13. Optimization of Ab:beads ratio and incubation time.
A fixed amount of sepharose beads was incubated with 5 µg, 10 µg, 25 µg, 50 µg, 70 µg or 100 µg of the mouse anti-HLA-A2 antibody in the IP washing buffer or with the buffer alone as a control for 3 h. A small aliquot of the supernatant from every sample was collected at the beginning of the incubation, 30 min, 1h, 2h and 3h later. A) Titration of beads with anti-HLA-A2 Ab; B) Kinetics of anti-HLA-A2 Ab binding to beads. Equal amounts of supernatants were loaded for SDS-PAGE separation, followed by Western Blot analysis. Detection was performed with HRP-conjugated anti-mouse-IgG Ab.

4.1.2. Optimization of other IP protocol parameters

One important component in cell lysis is the detergent. The choice of possible detergents for our purpose was limited, as it needs to preserve the 3D structure of the MHC-peptide complex for a successful IP and epitope extraction. Additionally, it needs to be LC-MS friendly and not costly. We decided to use CHAPS, also used in the group of Prof. Rammensee and Prof. Stevanović at the University of Tübingen, Germany (67-69, 138, 167, 168), and took their protocol as a basis for our protocol with slight modifications. We did not perform the IP on a column but in suspension, as we wanted to keep all steps in a small scale for future small tumor specimens. Following the protocol described in section 3.2.3.1, experiments were performed to compare different parameters such as the time of incubation of the Ab-beads in the lysate, an additional sonication step of the lysate before the incubation with Ab-beads, and the use of magnetic beads instead of sepharose beads. Samples were analyzed with Western blot or by LC-MS^3 analysis.

We saw no improvement in IP yield if the incubation of Ab-beads with the lysate took longer than 3 h or if sonication was used to homogenate the lysate. IP with the magnetic beads, which were easier to handle, gave results comparable to those from sepharose beads (data not shown).

Due to the costs, sepharose beads were chosen over magnetic beads. Figure 14 shows a representative enrichment of HLA-A2 complexes with the optimized IP protocol.

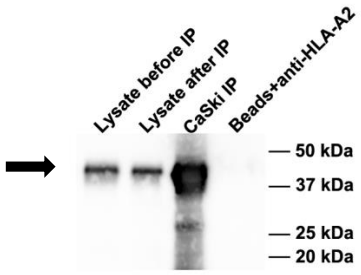


Figure 14. Successful enrichment of HLA-A2 molecules from CaSki cells with the optimized IP protocol.
The IP sample, cell lysate before and after incubation with Ab-beads and Ab-beads incubated with the IP lysis buffer alone were separated by SDS-PAGE and transferred to a PVDF membrane. Detection was performed with rabbit anti-MHC I Ab and HRP-conjugated anti-rabbit-IgG Ab. The strong bands marked with the arrow correspond to the mass of the HLA I heavy chain (44 kDa).

Western blot detection was performed with a rabbit anti-MHC I Ab, which binds all HLA-I types. A band at 44 kDa corresponding to the size of the HLA heavy chain was detected in the lysate before and after IP as well as in the IP sample. However, the intensity of the band in the IP sample is significantly stronger than that in the lysate before and after incubation with the coupled Ab-beads, indicating successful HLA-A2 enrichment. Of note, a band remains visible in the "Lysate after IP", which is due to only HLA-A2 molecules being taken out of the lysate, while all other HLA-I molecules (HLA-A3, -B7, -B37 and -C7) remain.

4.1.3. Optimization of the cell number per IP sample

The optimal IP capture of HLA-A2 molecules depends on the ratio of Ab-beads to the number of HLA-A2 molecules in the cell lysate, which is related to the number of cells and their HLA-A2 expression levels. In order to optimize the cell number for an IP sample, aliquots of 20 µg Ab coupled to 25 µL pelleted beads were incubated with increasing numbers of CaSki, SNU17 or SNU1000 cells. The IP experiment was performed using 2-5 or 2-6, 2-5 and 1, 2, 3 and 5 confluent CaSki, SNU17 and SNU1000 dishes, respectively. A 90-95% confluent cell culture dish contains 10^7 cells on average; therefore, dish number is an estimation of the input cell number. IP was performed following the protocol described in section 3.2.3.1. Epitopes were dissociated from HLA-A2-peptide complexes by acetic treatment with 0.3% TFA in water, subjected to ultrafiltration, desalting with OMIX tips and analysis with LC-MS^3, using the nanoAcquity UPLC-QTrap5500 system. In order to compare IP samples, total intensities of the HLA-A2 binding endogenous peptides AIVDKVPSV and YLLPAIVHI, which are presented by almost every HLA-A2+ cell (personal communication with Prof. Rammensee), were normalized to the total signals from the internal standard peptides FGPVNHEEL and SSIEFARL added to the IP elution buffer before epitope elution. The results from two biological replicates from CaSki cells and one replicate of SNU17 and SNU1000 cells are depicted in Figure 15.
For SNU17 and SNU1000, maximum signal intensity was reached with 3 dishes (equivalent to app. $3x10^7$ cells) per IP sample. Intensity changed <10% compared to IP samples of more cells, owning to experimental variance. Signal intensity gains for CaSki IP were highest when ≥4 (replicate I) or ≥5 dishes (replicate II) were used for an IP sample. The smaller signal increase (<10%) between the last

two samples in each replicate indicated that the optimal number of cells to saturate Ab-beads was approached.

That means that for an optimal HLA-A2 CaSki IP with 20 µg Ab coupled to 25 µL pelleted beads, 6 dishes (app. $6x10^7$cells) are necessary.
Importantly, these results demonstrate that the signal intensity is cell line dependent. Thus, the input cell number for optimal IP results needed to be adjusted for each cell line. MS based adjustment would be too time-consuming for all cells and all MHC I molecules in our cell line collection, therefore, FACS staining was performed to compare CaSki HLA-A2 expression levels with the investigated MHC I expression levels of other cell lines by colleagues in the group (S. Hoppe, A. Klevenz, M. Küpper and A. Steinbach). These results served as a reference for IPs within the group and for collaboration projects.

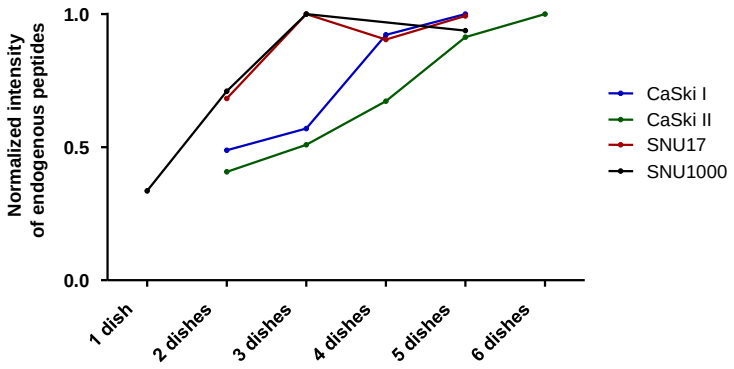


Figure 15. Optimization of the input cell number per IP sample.
Aliquots with fixed amount of Ab-beads were incubated with the lysate from increasing numbers of CaSki, SNU17 or SNU1000 cells (app. 10^7 cells/dish). Epitopes were eluted with 0.3% TFA in water containing murine MHC I restricted peptides FGPVNHEEL and SSIEFARL. Samples were ultrafiltrated, desalted with OMIX tips and analyzed with the nanoAcqutiy-QTrap5500 platform. Total intensities of the HLA-A2 endogenous peptides AIVDKVPSV and YLLPAIVHI were normalized to the total intensities of the added synthetic murine MHC I peptides. The optimal dish number was reached when the normalized signal changed for ≤10%. Results from two biological replicates of CaSki (I –first replicate, II –second replicate) and one of SNU17 and SNU1000 are shown.

4.1.4. LC-MS^3 detection of externally pulsed $E7_{11-19}$ on the cell surface

During all LC-MS^3 analyses described above, the presence of the HLA-A2 HPV16 $E7_{11-19}$ YMLDLQPET peptide, which was detected on CaSki cells and tumors before (146, 147), was examined. The peptide was not detected in any IP sample. To validate the experimental procedure, the target HLA-A2 HPV16 $E7_{11-19}$ YMLDLQPET peptide was externally loaded on the cell surface of HPV16 negative cells, BSM, which are homozygous for HLA-A2, -B15 and -C4. Prior to the experiment, it was checked *in silico* with different prediction servers that this peptide does not bind to the other two HLA molecules, HLA-B15 and -C4. A suspension containing $5x10^7$ cells was incubated either with 2 µg, 20 µg or 200 µg $E7_{11-19}$ peptide, 20 µL DMSO or 20 µL medium. Cells were harvested after 5 h incubation and the IP protocol was performed as described in section 3.2.3.1. IP samples were then treated with 0.3% TFA in water containing internal standards FGPVNHEEL and SSIEFARL, ultrafiltrated, desalted with OMIX tips and analyzed with two technical replicates using LC-MS^3 (nanoAcquity UPLC-QTrap6500 system).

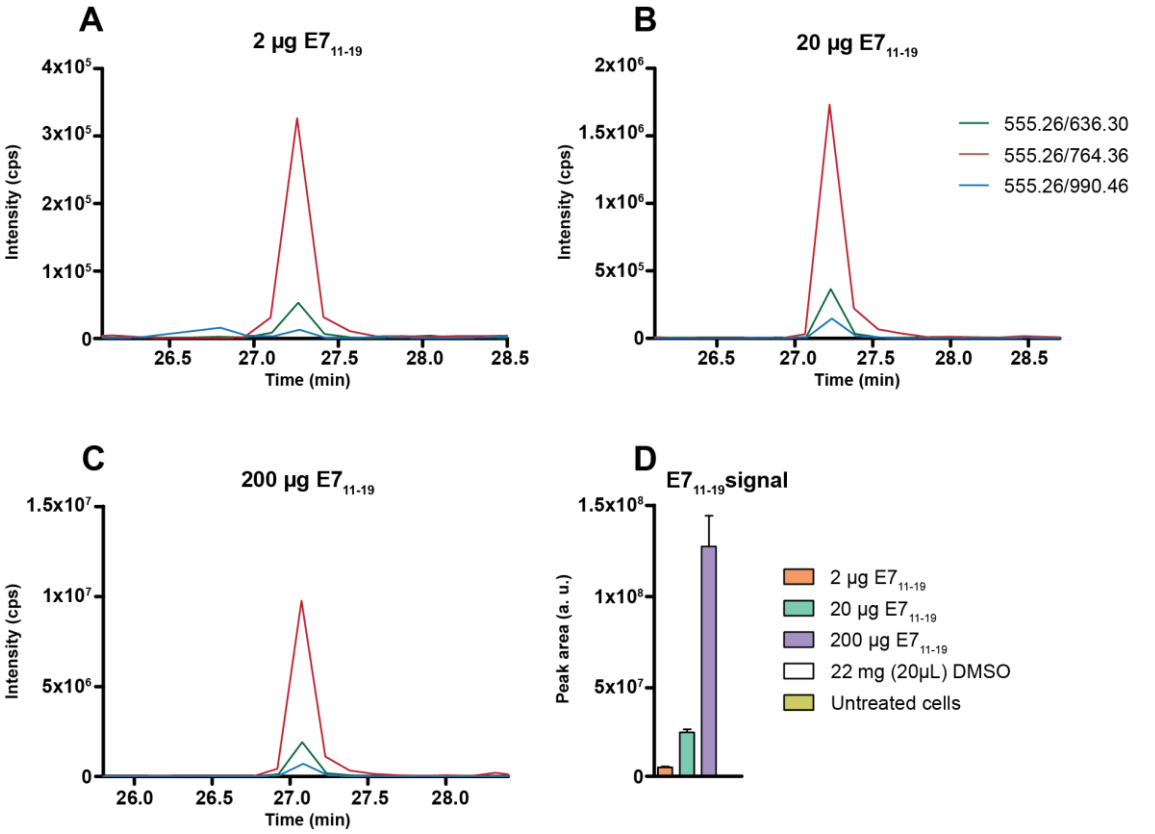


Figure 16. IP of HLA-A2+HPV- BSM cells externally loaded with the $E7_{11-19}$ YMLDLQPET peptide.
Cells were incubated with 2, 20, 200 µg $E7_{11-19}$, 20 µL DMSO or left untreated. IP was performed 5 h after incubation; peptides were eluted with pH reduction, subjected to ultrafiltration, desalting with OMIX tips and analysis with LC-MS^3 using the nanoAcquity UPLC-QTrap6500 system. All results are representative of one biological replicate. A) – C) Extracted ion chromatogram for transitions b_5 (555.26/636.30), b_6 (555.26/764.36) and b_8 (555.26/990.46) of synthetic peptide $E7_{11-19}$, representative of one technical replicate; D) Signal intensities for $E7_{11-19}$ for two technical replicates for IPs of cells treated with different amounts of $E7_{11-19}$. cps: counts per second, a.u.: arbitrary units

The peptide $E7_{11-19}$ YMLDLQPET was detected in all samples where cells were incubated with the peptide prior to the IP. The chromatographic profile of three measured transitions showed their co-elution and their relative intensities at the expected retention time (27.1 – 27.5 min) (Figure 16 A – C). The peptide $E7_{11-19}$ was not present in samples where cells were incubated with 20 µL DMSO or were left untreated. The detected signals ratios of the $E7_{11-19}$ YMLDLQPET peptide between IP samples with different amounts added to cells were around five-fold (Figure 16 D), whereas the ratios of actually added amounts of peptide to cells were ten-fold. This means that the signal intensities in the LC-MS^3 analysis of three different conditions partially correlated to the amount which was added to cells. The endogenous peptides AIVDKVPSV and YLLPAIVHI were detected in all samples.

These results show that the experimental setting is capable of capturing endogenous peptides as well as our target peptide $E7_{11-19}$ YMLDLQPET, when present in amounts that are high enough to be detected with LC-MS^3.

4.2. Epitope extraction, enrichment and purification strategies

The analysis to find the target peptide $E7_{11-19}$ YMLDLQPET on CaSki or SNU cell lines was repeated several times without success, although the same or higher sample input was used as reported elsewhere (146). We assumed that the peptide was lost during sample preparation due to absorption on the plasticware or on the ultrafiltration device. Therefore, the ultrafiltration step was omitted and materials with low binding affinities for proteins and peptides were used throughout the whole experimental workflow.

To exclude the possibility of the IP elution buffer not effectively dissociating epitopes from HLA molecules, two elution buffers described in the literature were compared, namely 10% acetic acid in water (70) and 0.2% TFA in water (67-69, 138). Two CaSki IP samples prepared in parallel from $6x10^7$ cells each, were treated either with 70 µL of 10% acetic acid in water or 70 µL of 0.3% TFA in water with gentle mixing at RT for 15 minutes. Samples were subjected to desalting with OMIX tips and analyzed using LC-MS^3 methodology (nanoAcquity UPLC-QTrap5500 system). The results showed that there was no difference between the two IP elution buffers, as comparable amounts of endogenous peptides AIVDKVPSV and YLLPAIVHI were detected in both IP samples (Figure 39 in the Appendix). The 0.3% TFA in water elution buffer was chosen for further experiments.

In this set of experiments, also the presence of the $E7_{11-19}$ YMLDLQPET peptide in the IP samples was assessed. The identification of HPV16 $E7_{11-19}$ peptide was not successful. It was assumed that the amount of the input material was not sufficient for detection of the target peptide $E7_{11-19}$. Therefore ten CaSki IP samples from $6x10^7$ cells each were prepared using five OMIX tips for desalting before LC-MS analysis. Increased sample amounts caused the LC column to clog, both on the nanoAquity-QTrap6500 and the speLC-Q Exactive instuments. Therefore, other enrichment and purification strategies were required.

4.2.1. Hydrophilic interaction liquid chromatography (HILIC)

To examine HILIC as a purification and enrichment strategy, two BSM IP samples were eluted with 70 µL 0.3% TFA in water each. The elution buffer contained all peptides from 3.1.14.1. IP eluates were pulled together to ensure uniform sample composition. An aliquot of 40 µL, taken as a starting IP eluate reference, and the rest of the eluate were subjected to vacuum drying. The sample was resuspended in 80% ACN/1% TFA and each half was loaded either on a TSKgel- or a ZIC-HILIC micro-column (see 3.2.4.3). Fractions were gradually eluted from the micro-columns with a starting elution buffer containing high to low % ACN. All fractions, including unbound flow through and washing fractions, were dried, resuspended and spotted on the MALDI target plate together with MALDI matrix solution.

Gradual elution from the ZIC-HILIC micro-column showed that peptides and proteins co-eluted, as peptides and proteins were present in the same fractions, whereas the TSKgel-HILIC material could separate proteins from peptides (Figure 17 and Figure 18).

Most peptides were contained in the flow through and the wash fractions of the micro-column (red and green lines in Figure 17), while most proteins were present in the fractions eluted with 60% and 70% ACN/0.1%TFA in water (red and turquoise lines in Figure 18).

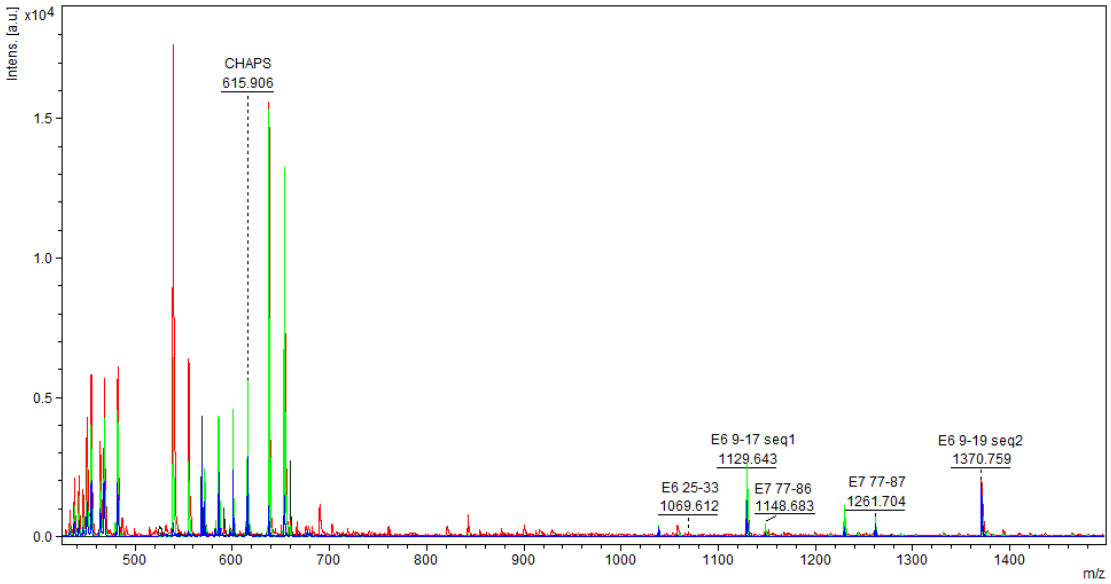


Figure 17. MALDI TOF MS peptide profile of BSM IP samples with added peptides, after fractionation with the TSKgel HILIC micro-column.
The IP eluate was subjected to vacuum drying prior to resuspension in 80% ACN/1% TFA. Gradual elution started with an elution buffer containing 70% ACN/0.1% TFA and ended with an elution buffer containing 0.1% TFA in water. The analysis was performed with the UltrafleXtreme MALDI-TOF MS instrument. Unfractionated IP sample is represented by the blue line. The flow through fraction is shown in green and the wash fraction by the red line. The fraction eluted with 70% ACN/0.1% TFA is shown by the grey line which has peaks only in the low mass range and no peaks corresponding to peptide masses. All other fractions were not plotted as they were of as low or lower intensity as the fraction eluted with 70% ACN/0.1% TFA. Peptide peaks are marked with their masses and names. Peptides were detected in the unfractionated IP sample, flow through and wash fraction.

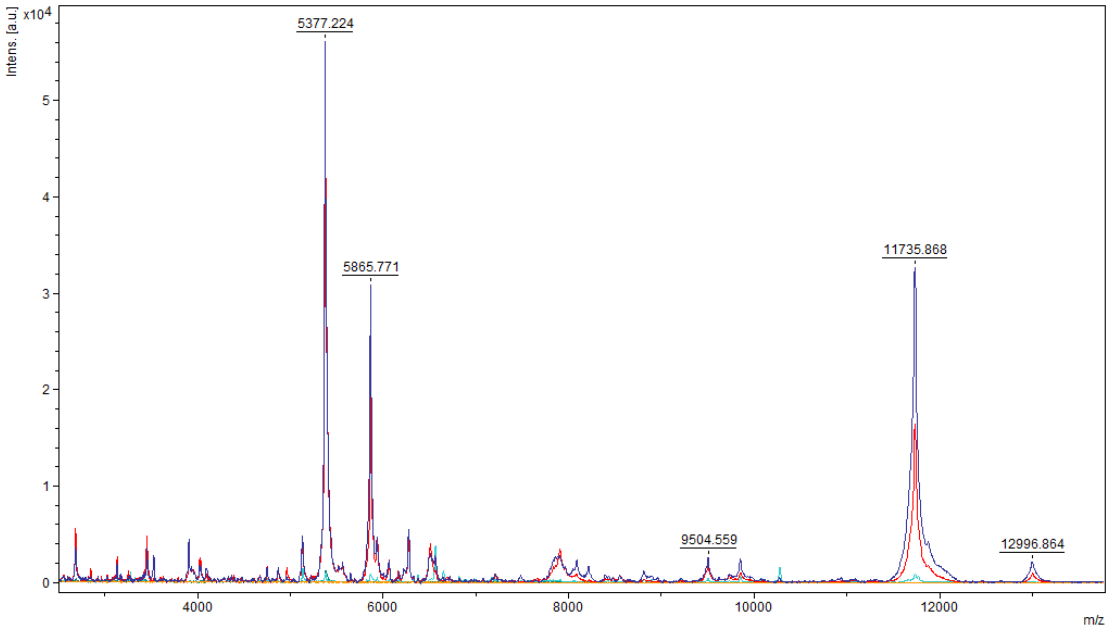


Figure 18. MALDI TOF MS protein profile of BSM IP samples with added peptides, after fractionation with the TSKgel HILIC micro-column.
The IP eluate was subjected to vacuum drying prior to resuspension in 80% ACN/1% TFA. Gradual elution started with an elution buffer containing 70% ACN/0.1% TFA and ended with an elution buffer containing 0.1% TFA in water. The analysis was performed with the UltrafleXtreme MALDI-TOF MS instrument. The fraction eluted with 70% ACN/0.1% TFA is shown by the red line and the fraction eluted with 60% ACN/0.1% TFA in turquoise. The flow through (orange) and the wash fraction (gray) are overlaid on the baseline of the profile. All other fractions were not plotted as they had intensity as low as the flow through and the wash fractions. Proteins were detected in the unfractionated IP sample and fractions eluting with 60% and 70% ACN/0.1% TFA. Peaks at masses 5865.771 and 11735.868 m/z most probably correspond to the β_2-microglobulin (β_2M) protein.

Due to poor ionization of peptides containing no basic amino acids, they were not detected with the MALDI-TOF MS. Only peptides $E6_{9-17}$ FQDPQERPI, $E6_{9-19}$ FQDPQERPIKL, $E6_{25-33}$ ELQTTIHDI, $E7_{77-86}$ RTLEDLLMGT and $E7_{77-87}$ RTLEDLLMGTL were observed.

Many low molecular mass contaminants (masses below 700 m/z), including CHAPS, were present in the fractions rich with peptides, but they were also present in lower intensities in fractions eluting with ≤70% ACN/0.1% TFA (Figure 17).

The MALDI TOF MS protein profile revealed the presence of several protein species, many of them being contaminants in the IP sample. The intense peaks at 5865.771 and 11735.868 m/z most likely corresponded to the β_2-microglobulin (β_2M) chain (Figure 18).

Taken together, the TSKgel HILIC is a good alternative for epitope purification and enrichment. However, when the experiment was up-scaled, the solubility of the material in the HILIC loading buffer, which contains high organic solvent, was limited or ineffective. Therefore, peptide/protein separation with HILIC chromatography was not considered to be a suitable sample purification and enrichment strategy.

4.2.2. Acetone – ethyl acetate precipitation

Acetone precipitation is a widely used method in proteomics to isolate proteins (162), whereas ethyl acetate precipitation was described as a method to remove a detergent from a sample (163).

In order to test combined acetone – ethyl acetate precipitation, a workflow was designed such that proteins are first removed from the sample by acetone precipitation, leaving unprecipitated peptides in the supernatant. Next, the supernatant is subjected to ethyl acetate precipitation to remove detergent and thereby achieving peptide enrichment and purification.

IP eluates of three BSM IP samples, peptide mixture of all peptides from 3.1.14.1 and ice-cold acetone were mixed. A portion of the sample was used for direct vacuum drying as an unprocessed reference. The rest of the sample was subjected to acetone precipitation. The supernatant was vacuum-dried until only 50 µL of liquid remained and subjected to ethyl acetate precipitation. The aqueous bottom layer and ethyl acetate top layer were separated by transferring into fresh Eppendorf tubes and dried in the vacuum centrifuge. All fractions generated in this experiment, namely acetone pellet, aqueous bottom and ethyl acetate top layers, as well as unprocessed sample were resuspended and spotted on the MALDI target plate. Samples were measured using the UltrafleXtreme MALDI-TOF MS instrument.

Peptides and CHAPS were present in; acetone pellet, the aqueous and the ethyl acetate top layer (Figure 41 in the Appendix). Surprisingly, proteins were detected in the aqueous ethyl acetate layer but not in the acetone precipitation pellet (Figure 42 in the Appendix).

Furthermore, the signals of unprocessed IP sample, which was approximately 7% of the whole sample subjected to the precipitation, gave higher results for proteins and peptides than any of the fractions. These results showed that the acetone – ethyl acetate precipitation is not a workflow suitable for the purification and enrichment of HLA epitopes.

4.2.3 Fractionation of IP samples on reverse phase chromatographic material

None of the strategies described above successfully purified and enriched epitopes from IP samples. Therefore, several reverse phase material approaches were examined. We aimed to determine the optimal concentrations of organic solvent needed for removing CHAPS and proteins from IP eluates for successful epitope enrichment.

All experiments on micro-columns packed with reverse phase materials were performed as described in section 3.2.4.2.3 and on the Seppak cartridge as described in section 3.2.4.2.2.

4.2.3.1 R2 and Oligo R3 micro-column

In order to determine optimal elution conditions to purify epitopes from the micro-column, three CaSki IP samples were eluted with 70 µL 0.3% TFA in water and shortly centrifuged. All eluates were combined to ensure sample homogeneity. The supernatant was loaded on the R2 or Oligo R3 micro-column with the Empore C8 material plug. Both types of micro-column clogged during loading of one IP sample eluate.

Next, the question whether the clogging happened on the C18 material or on the micro-column plug was examined. Furthermore, the possibility of using a different material for the plug was assessed. To this end, the CaSki IP eluate of four samples was combined and each half was loaded either on a tip containing only an Empore C8 material plug or a cellulose Kimtech tissue plug, but no C18 material. The IP eluate was easily passing through both materials, causing no clogging. It was concluded that the clogging of the micro-columns occurred within the C18 material and not on the plug.

Additionally, Empore C8 material peptide binding properties were compared with cellulose Kimtech tissues. The same amount of HLA-A2 HPV16 peptides (3.1.14.1) was bound to the Oligo R3 micro-column once with an Empore C8 material and once with an cellulose Kimtech tissue plug. The results revealed no difference in micro-column performance, meaning that both materials can be used as a micro-column plug. Thus, Kimtech tissue was chosen for further use due to its lower price and easier accessibility.

4.2.3.2 Seppak C18 cartridge

As outlined in the section above, R2 and Oligo R3 micro-columns could not be used for epitope purification from the IP sample. Therefore, a purification strategy with a Seppak cartridge was tested. The eluates of two CaSki IP samples were combined and loaded on the Seppak cartridge. Loading did not reduce the ease with which the liquid was passing through the resin and it did not cause clogging of the cartridge.

Next, an experiment was performed to address the question whether the Seppak cartridge did not clog due to bigger volume and larger binding capacities than micro-columns or due to the sample contaminants not binding to the Seppak resin. Furthermore, we tested when the components binding to the Seppak resin eluted during gradual elution and compared that with an Oligo R3 micro-column.

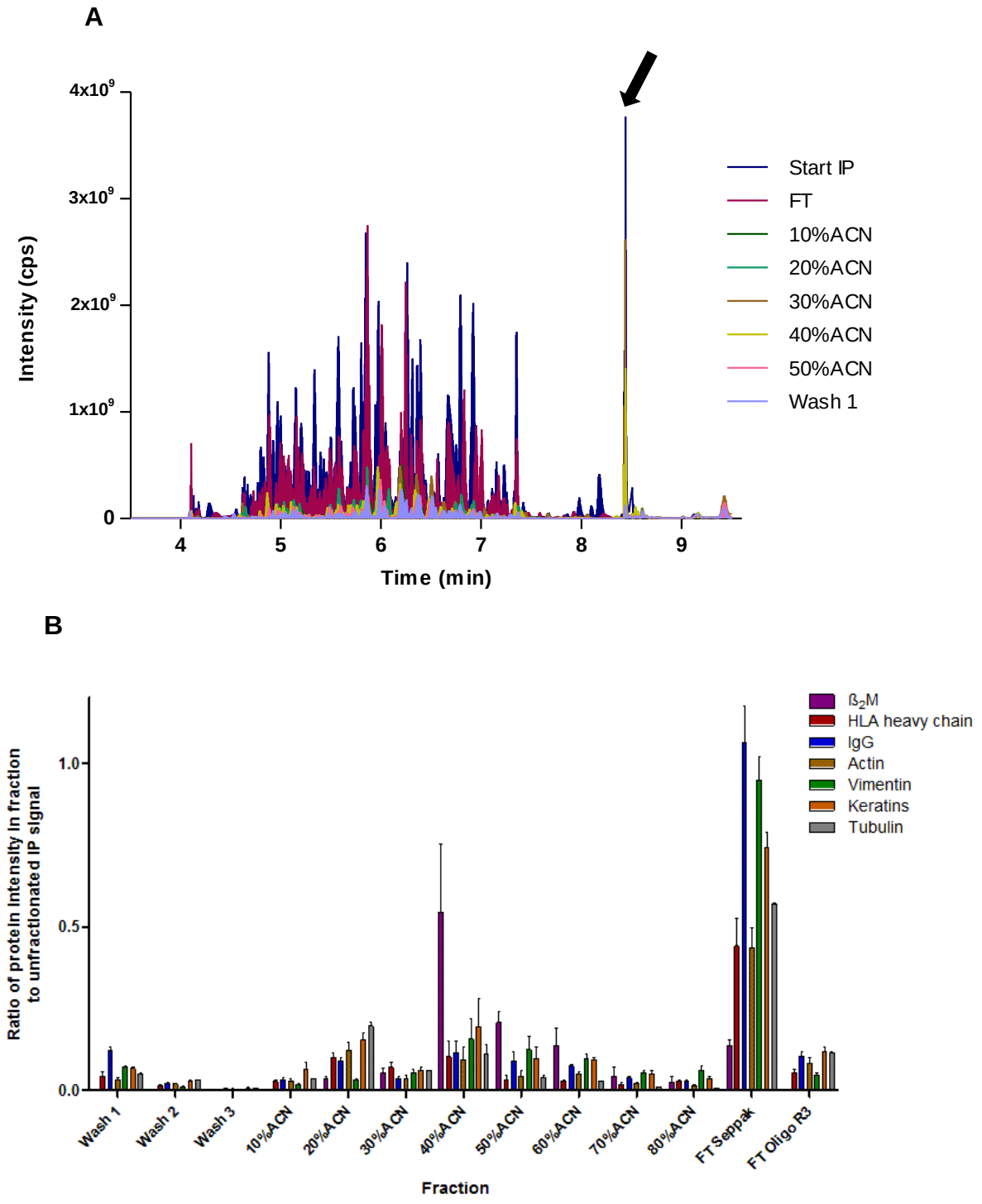


Figure 19. Chromatographic profile (A) and relative protein quantification of an IP sample (B) after fractionation on the Seppak cartridge or Oligo R3 micro-column.
The IP sample was bound to the Seppak resin and gradually eluted with solvents containing low to high % ACN. Fractions were collected and subjected to in-solution reduction, alkylation and digestion. Samples were analyzed with the speLC-Q-Exactive platform. A) Start IP-unfractionated reference IP eluate, FT-Seppak flow through fraction, 10% – 50% ACN-fractions eluted with 10% ACN/0.1% TFA – 50% ACN/0.1% TFA in water, wash1-first Seppak wash after IP binding. The black arrow shows the strong CHAPS peak. B) Data processing was performed with the MaxQuant program. Protein intensities in fractions were compared relative to the intensities in the unfractionated IP sample. FT Seppak-Seppak flow through fraction, FT Oligo R3-Oligo R3 micro-column flow through fraction, 10% – 80% ACN-Seppak fractions eluted with 10% ACN/0.1% TFA to 80% ACN/0.1% TFA in water, Wash1-3-first to third Seppak wash after IP binding. Quantified proteins are: IgG-sum of intensities for heavy and light chains, Actin and Tubulin-sum of intensities for alpha and beta protein, Keratins-sum of intensities for I 14, I 17, II 6 and II 7 keratins. Error bars are SD from three technical replicates, except the fractions wash1 and 40% ACN with two technical replicates.

To do so, three CaSki IP sample eluates were combined, the first aliquot was vacuum dried for a reference, the second was loaded on the Seppak cartridge and the third on the Oligo R3 micro-column with a Kimtech tissue plug. As expected, it was not possible to wash and elute the sample from the Oligo R3 micro-columns due to clogging. Therefore, only the flow through fraction was collected and dried in the vacuum centrifuge.
An aliquot of IP eluate was loaded on the Seppak cartridge without any difficulties. The fractionation from the Seppak cartridge was conducted with gradual elution with solvents, with increasing % ACN as follows: 10% ACN/0.1% TFA → 20% ACN/0.1% TFA → 30% ACN/0.1% TFA → 40% ACN/0.1% TFA → 50% ACN/0.1% TFA → 60% ACN/0.1% TFA → 70% ACN/0.1% TFA → 80% ACN/0.1% TFA, all in water. All fractions, including flow through and three wash fractions were collected and subjected to vacuum drying.
Fractions generated during the experiment (unprocessed IP eluate, Oligo R3 micro-column flow through and all Seppak fractions) were subjected to reduction with DTT, alkylation with IAA and in-solution trypsin digestion overnight (detailed protocol in section 3.2.2.3).

Equal aliquots of each fraction were measured with the speLC-Q-Exactive MS instrument. The chromatographic profile of the LC-MS^2 analysis for some fractions (the starting unprocessed IP eluate, the Seppak flow through, the first Seppak wash and the Seppak fractions eluted with 10% – 50% ACN/0.1% TFA) are depicted in Figure 19 A. The comparison of the chromatographic profiles indicated that the Seppak flow through contained most of the protein material relative to other fractions and that its profile matched well with those of the unprocessed IP eluate. The second and third Seppak wash fractions and fractions eluted with higher organic content (≥50% ACN/0.1% TFA) contained least detected proteins and their chromatographic profile was comparable to the one from the first Seppak wash fraction (light blue in Figure 19 A).

High singly charged peaks with the mass 615.4 m/z detected at a retention time of 8.4 min correspond to CHAPS in the sample (black arrow in Figure 19 A). As expected, the highest CHAPS signal was observed in the unfractionated sample. CHAPS was detected in the 20% ACN/0.1% TFA fraction in low intensity, whereas most of it was present in the fractions eluted with 30% and 40% ACN/0.1% TFA.
Identification and quantification of results was performed using the MaxQuant software version 1.5.2.8, as described in 3.2.6.1 (155, 156). The results of a data base search with the analysis software revealed that besides the expected antibody and HLA-A2 proteins, also cytoskeleton proteins (tubulin, actin, vimentin, plektin), keratin, ribosome subunits, heat shock proteins, histones and others, had bound unspecifically to the coupled Ab-beads and were present in the IP samples (8.7. in the Appendix). The results for selected proteins are presented as the ratio of intensities in a fraction to the intensity in the unprocessed IP eluate in Figure 19 B.
The results of relative comparison confirmed that most of the proteins were, indeed, contained in the Seppak flow through, indicating that the proteins mostly did not bind to the Seppak C18 material. In contrast, the Oligo R3 flow through contained little proteins, meaning that they bound to the resin and

thereby clogged it. β_2M bound the most compared to other proteins and it was present with the highest signal in the fraction eluted with 40% ACN/0.1% TFA.

When examining the producers' descriptions of tested materials, the key difference between the R2 and Oligo R3 materials and the Seppak cartridge was pore size. The pore size of the Seppak C18 material was app. 130 Å, whereas the pore size of the R2 and Oligo R3 materials was in the range of 300 Å to 3000 Å. The pore size of the R2 and Oligo R3 was big enough to enable diffusion and retention of proteins into pores, whereas this was prevented or hindered in the Seppak material.
The only disadvantage of the Seppak cartridge was its bigger volume and amount of the resin, which could cause sample losses on the plastic walls or losses due to irreversible binding on the cartridge resin. When one works with higher sample amounts such as whole cell lysates, these factors are not critical, whereas in our case they could contribute to the loss of already low abundant epitopes and preclude their detection with LC-MS.

4.2.3.3 Zorbax micro-column

In order to minimize possible losses during sample purification, we examined a reverse phase column packing the material Zorbax, which has a pore size of 80 Å. It could be easily packed into the micro-columns. If the hypothesis about the effect of the pore size was true, then the Zorbax material with smaller pore size should bind even less proteins from an IP sample than what was observed with the Seppak cartridge.
A similar experimental setting was performed as described above. The Zorbax material was packed in a 200 µL tip as described in section 3.2.4.2.3. IP eluate from four SNU17 IP samples was combined and the equivalent of three IP samples was loaded on one micro-column, causing no clogging. It is important to note that at the end of the loading, the force needed to push the liquid through the micro-column increased, indicating that with loading of higher amounts of IP eluate, also the Zorbax micro-column would have clogged.
The micro-column was washed and fractions were collected during gradual elution with solvents with increasing % ACN as follows: 15% ACN/0.1% TFA → 20% ACN/0.1% TFA → 25% ACN/0.1% TFA → 30% ACN/0.1% TFA → 35% ACN/0.1% TFA → 40% ACN/0.1% TFA → 50% ACN/0.1% TFA → 60% ACN/0.1% TFA → 70% ACN/0.1% TFA, all in water. All fractions including the flow through fraction, wash fractions after IP eluate loading, and the IP eluate equivalent to one IP sample were dried in a vacuum centrifuge. Subsequently, they were subjected to reduction with DTT, alkylation with IAA and in-solution trypsin digestion over night as described in section 3.2.2.3. Samples were measured with the speLC-Q-Exactive instrument. The chromatographic profile of the LC-MS^2 analysis for some fractions (one third of the unprocessed IP eluate relative to the amount used for processing with the Zorbax micro-column, the Zorbax flow through, the first Zorbax wash and the Zorbax fractions eluted with 15% – 50% ACN/0.1% TFA) are depicted in Figure 20 A.

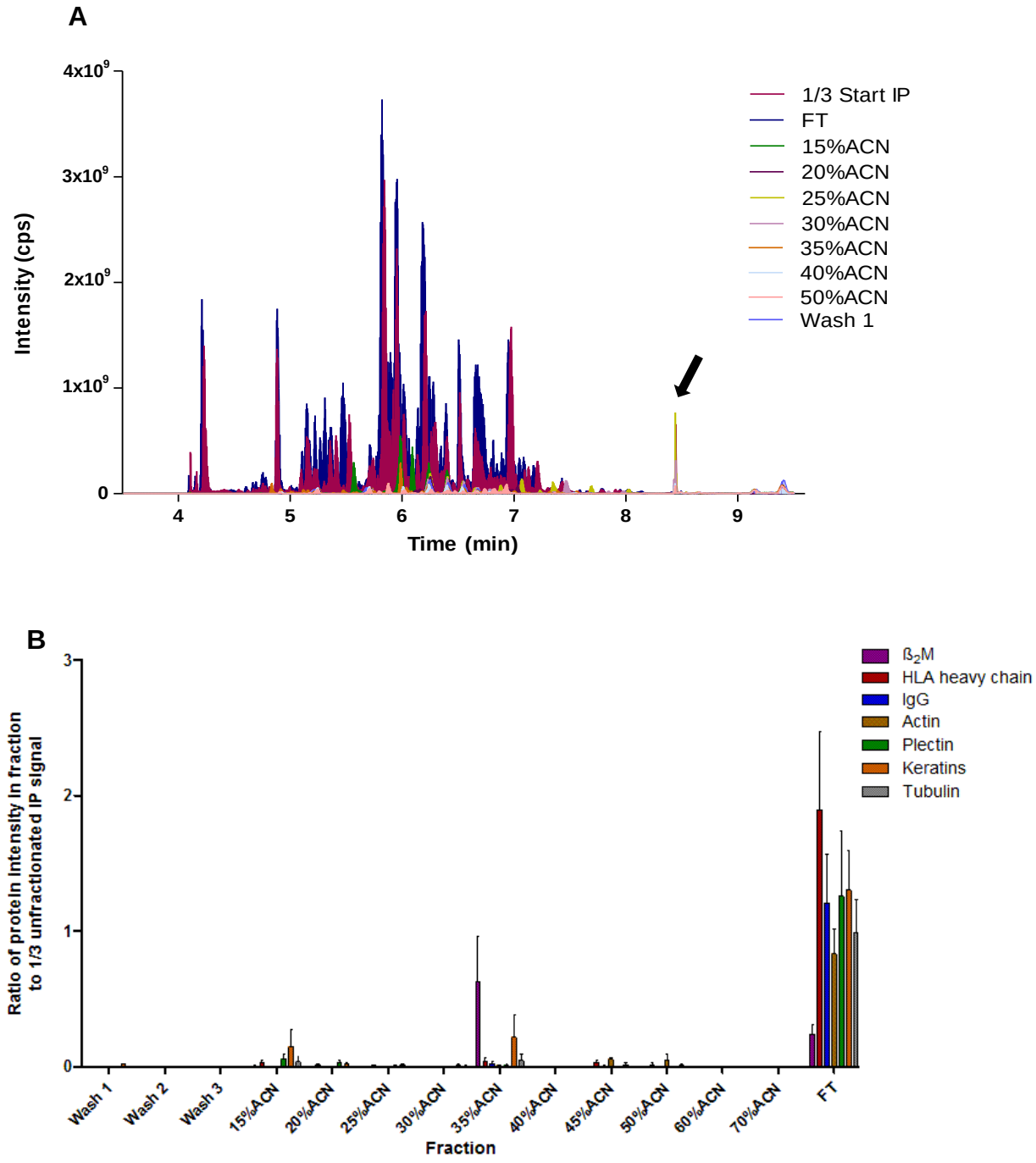


Figure 20. Chromatographic profile (A) and relative protein quantification of an IP sample (B) after fractionation on the Zorbax micro-column.
The IP sample was bound to the Zorbax resin and gradually eluted with solvents containing low to high % ACN. Fractions were collected and subjected to in-solution reduction, alkylation and digestion. Samples were analyzed with the speLC-Q-Exactive platform. A) 1/3 Start IP: 1/3 of unfractionated reference IP eluate, FT: flow through fraction, 15% – 50% ACN: fractions eluted with 15% ACN/0.1% TFA – 50% ACN/0.1% TFA in water, Wash1: first Zorbax wash after IP binding. The black arrow shows the CHAPS peak. B) Data processing was performed with the MaxQuant program. Protein intensities in fractions were compared to the intensities in the 1/3 of unfractionated IP sample. FT: flow through fraction, 15% – 70% ACN: Zorbax fractions eluted with 15% ACN/0.1% TFA to 70% ACN/0.1% TFA in water, Wash1-3: first to third Zorbax wash after IP binding. Quantified proteins are: IgG-sum of intensities for heavy and light chains, Actin and Tubulin-sum of intensities for alpha and beta protein, Keratins-sum of intensities for I 14, I 17, II 6 and II 7 keratins. Error bars represent the SD from a minimum of two technical replicates.

As hypothesized, the comparison of the chromatographic profiles shows that the Zorbax flow through contained most of the proteins material relative to other fractions. Expectedly, the profile of the flow

through contained more peaks than the unprocessed IP sample, as the latter corresponded to 1/3 of what was used for binding to the micro-column. The second and the third Zorbax wash fractions and fractions eluted with 60% or 70% ACN/0.1% TFA in water contained least detected peptides and their chromatographic profile was comparable to that from the fractions eluted with 50% ACN/0.1% TFA (light pink in Figure 20 A).

The singly charged CHAPS peak with 615.4 m/z was detected at 8.4 min (black arrow in Figure 20 A). Expectedly, CHAPS was detected in in the unfractionated sample as well as in the fractions eluted with 25% and 30% ACN/0.1% TFA.

The identification results (Figure 20 B) confirmed the presence of numerous protein contaminants such as cytoskeleton proteins (tubulin, actin, plektin), keratin, ribosome subunits, histones and others in the IP eluate. As seen before, the biggest proportion of proteins represented keratins, which is not surprising as the SNU17 cells used for this experiment derive from keratinocytes.

The relative quantification results for selected proteins are presented as the ratio of intensities in a fraction to the intensity of the same protein in one third of unprocessed IP eluate in Figure 20 B. Most of the proteins were present in the Zorbax flow through, indicating that the proteins mostly did not bind to the micro-column resin, or they bound irreversibly as their relative intensities did not result near 100% (ratio at 3) but more at 50% (ratio at 1.5). β_2M binds the strongest to the Zorbax resin compared to other proteins and it was present with the highest signal in the fraction eluted with 35% ACN/0.1% TFA. The amounts of proteins present in other fractions were so low that they could be considered as a background or blank sample signals, which was reflected in the results in Figure 20 A, B.

Taken together, the results of gradual elution of an IP eluate from the Zorbax C18 material confirmed the hypothesis that the pore size had an impact on binding of protein contaminants on the reverse phase material and clogging of the micro-column.

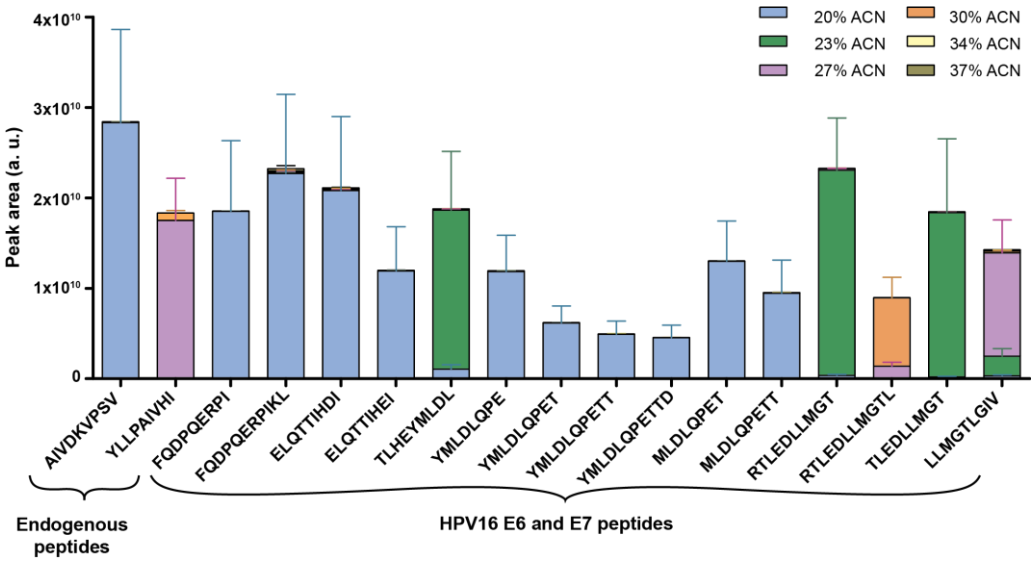


Figure 21. Gradual elution of HLA-A2 HPV16 synthetic peptides from the Zorbax micro-column.
100 pmol/peptide in the BSM IP eluate were bound to the micro-column and eluted with increasing % ACN. Fractions were collected, vacuum dried and analyzed with the speLC-Q-Exactive platform. Data was processed with the Skyline program. Signal intensities are plotted for every peptide in each fraction. 20% – 37% ACN-Zorbax fractions eluted with 20% ACN/0.1% TFA to 37% ACN/0.1% TFA in water. Error bars represent the SD from two technical replicates.

After finding an experimental set-up that separates proteins from peptides, we next wanted to examine the possibility of separating peptides from CHAPS by sequential elution. A similar gradual elution experiment was performed with 100 pmol per peptide of target HLA-A2 HPV16 E6 and E7 (3.1.14.1) and endogenous control synthetic peptides (3.1.14.4) in the BSM IP eluate from one IP sample. As seen in the previous experiment, CHAPS eluted from the Zorbax material with the solvent containing 25% to 30% ACN/0.1% TFA. Therefore, the gradual elution was performed with solvents composed of % ACN closer to that range, such that they followed the series: 20% ACN/0.1% TFA → 23% ACN/0.1% TFA → 27% ACN/0.1% TFA → 30 ACN/0.1% TFA → 34% ACN/0.1% TFA → 37% ACN/0.1% TFA → 40% ACN/0.1% TFA → 43% ACN/0.1% TFA, all in water. Fractions were vacuum dried, resuspended and analyzed on the speLC-Q-Exactive instrument in two technical replicates. The MS^1 intensities of each peptide in all fractions are depicted in Figure 21, except those eluted with 40% and 43% ACN/0.1% TFA, which contained no peptides.

Most of the HLA-A2 HPV16 peptides already eluted with the solvent containing 20% ACN/0.1% TFA in water. However, some of them were more hydrophobic and eluted with up to 30% ACN/0.1% TFA in water. Thus, effective elution of all HLA-A2 HPV16 peptides can be conducted with ≥30% ACN/0.1% TFA in water.

Narrowing down the ACN concentration range in the elution buffer showed that the highest CHAPS signals were in fractions eluted with 23% and 27% ACN/0.1% TFA in water (Figure 22), whereas other fractions contained minimal CHAPS signals. CHAPS eluted in the same fractions as some of the more hydrophobic epitopes, therefore, the sequential elution of epitopes from the Zorbax micro-column to reduce the amount of CHAPS in the IP eluate without losing some of more hydrophobic peptides was not possible. Thus, we did not employ this strategy.

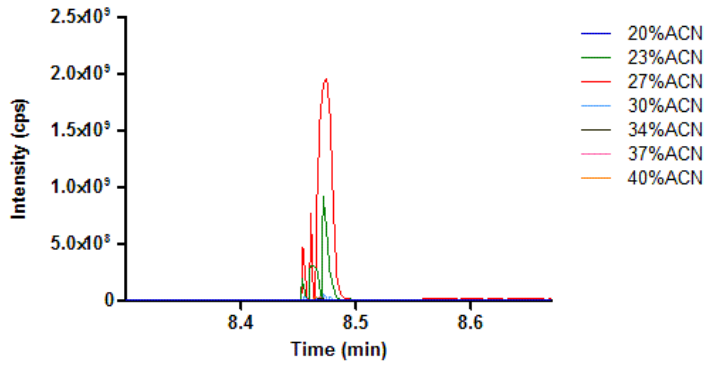


Figure 22. LC-MS chromatographic profile of CHAPS in different fractions after gradual elution from the Zorbax micro-column. 100 pmol/peptide in the BSM IP eluate were bound to the micro-column and eluted with increasing % ACN. Fractions were collected, vacuum dried and analyzed with the speLC-Q-Exactive platform. 20% – 40% ACN-Zorbax fractions eluted with 20% ACN/0.1% TFA to 40% ACN/0.1% TFA in water.

4.2.4. Comparison of RP extraction methods and ultrafiltration

To evaluate the peptide recoveries from the C18 material extraction methodologies established during this thesis, they were compared with ultrafiltration which is widely employed for separation of proteins and peptides in proteomics. This is also the case for the preparation of MHC I IP samples (70, 138). To this end, ten BSM IP sample eluates were combined, distributed in two aliquots and then mixed with the HLA-A2 HPV16 synthetic peptides (3.1.14.1) in low and high amounts. The first aliquot contained 600 fmol/peptide, whereas the second aliquot contained 600 pmol/peptide. Each of the aliquots was distributed equally among five extraction materials such that every sample at the beginning of the processing contained 120 fmol/peptide or 120 pmol/peptide. Five extraction options were used for processing: 2 kDa Vivacon ultrafilters, 10 kDa Vivacon ultrafilters, 10 kDa Amicon ultrafilters, Seppak C18 1 mL cartridges, and Zorbax C18 material micro-columns (as described in sections 3.2.4.1, 3.2.4.2.2 and 3.2.4.2.3, respectively). Elution of samples from the Seppak cartridge and the Zorbax micro-column was performed with 35% ACN/0.1% TFA. All the samples were dried in a vacuum centrifuge.

The unprocessed starting IP sample with added peptides would have caused clogging of the LC column, as experienced before, therefore, the pure peptide mixtures with the same concentration as those added to the IP eluates were used as a reference. All samples and pure peptides were analyzed using the speLC-Q-Exactive instrument.

All methionine containing peptides underwent oxidation on the methionine, resulting in two distinct chromatographic peaks with the unmodified peptide eluting later. These peptides were quantified as the sum of unoxidized and oxidized counterparts. Selected methionine containing peptides (on the right of the red line in Figure 23 A, B) are only shown in the unoxidized form for demonstration of methionine oxidation levels.

When bigger amounts of peptides were added into the sample (120 pmol/peptide), there were no differences between tested extraction strategies for most of the peptides (Figure 23 A). The degree of methionine oxidation was lowest for extraction with the Zorbax micro-column among all tested strategies, as the signal for unoxidized methionine was highest (right side of the graph in Figure 23 A). The unoxidized methionine signals for other extraction options were lower, indicating a higher oxidation degree.

The recovery of smaller amounts of peptides (120 fmol/peptide) after IP eluate processing was best with the Zorbax micro-columns for almost all peptides (Figure 23 B).

When small amounts of peptides were added to the IP eluate, the methionine oxidized in all peptides to high degrees after processing with ultrafilters and to a lesser degree with Seppak. Least oxidized peptides were detected after processing with the Zorbax micro-column (right side of the graph in Figure 23 B). The degree of methionine oxidation was in general higher in IP samples containing less peptides than in IP eluates containing higher peptide amounts, despite the same sample processing.

It is important to note that none of the tested strategies removed CHAPS, because its signal was detected in every measured sample.

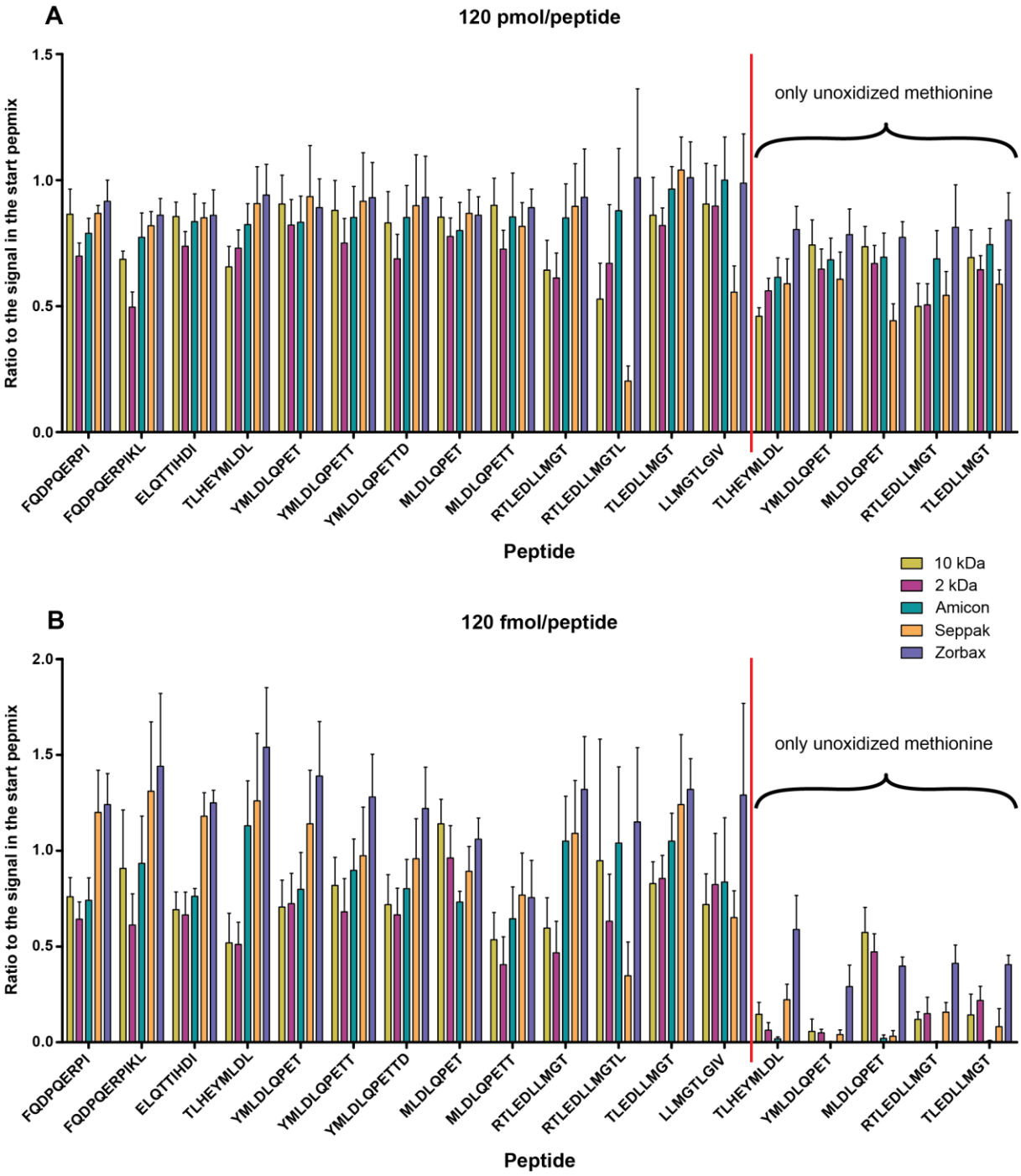


Figure 23. Comparison of different epitope extraction methods with high or low amounts of peptides in the IP eluate.
A) 120 pmol/peptide or B) 120 fmol/peptide were added into BSM IP eluates for each extraction method test. Samples were processed, vacuum dried and analyzed with the speLC-Q-Exactive instrument. Data processing and MS^1 relative quantification were performed with the Skyline program (157, 158). Peptide intensities in every sample were compared to the intensities in the unprocessed starting peptide mixture. Signal intensities of methionine containing peptides were calculated as intensity sum of oxidized and unoxidized counterparts. Results on the right of the red line represent the ratio of selected only unoxidized peptides to the total peptide signal in the starting mixture (oxidized and unoxidized). 10kDa: ultrafiltration with 10 kDa Vivacon ultrafilters, 2kDa: ultrafiltration with 2 kDa Vivacon ultrafilters, Amicon: ultrafiltration with 10 kDa Amicon ultrafilters, Seppak: processing with the Seppak cartridge, Zorbax: processing with the Zorbax micro-column. Error bars represent the SD from three technical replicates.

Taken together, in the case of high amounts of target peptide, all of the tested extraction strategies gave good results. When smaller amounts of target peptides were present in the IP sample, the Zorbax micro-column yielded a better performance than the other extraction strategies. Furthermore,

processing with the Zorbax micro-column caused least methionine oxidation, both with low and high peptide amounts.
The reasons for this are the smaller volume of the Zorbax resin and the smaller plastic surfaces where low abundant peptides could adsorb. This test showed the importance of experimental set up down-scaling for minimizing losses, when the amount of input material is low or investigated peptides are of low abundance.

4.2.5. Chemical tagging of primary amines for purification of epitopes

4.2.5.1. Method optimization on synthetic peptides

As shown in the previous chapter (4.2.3.4), the purification and enrichment strategy using Zorbax C18 micro-columns effectively separated peptides from most of the protein contaminants but not from CHAPS. To overcome this problem, we examined the possibility of chemical labeling of peptides with a tag containing a phospho-group, which can subsequently be enriched either by immobilized metal affinity chromatography (IMAC) (169-171) or titanium dioxide (TiO_2) pull down (164, 165). It was shown before that TiO_2 does not bind detergents. Moreover, detergents did not significantly influence the binding affinity of TiO_2 for phospho-peptides, whereas detergents affected the phospho-peptide isolation with the IMAC enrichment (172).

Chemical modifications on peptides most commonly target reactive groups, such as primary amine (-NH_2), carboxyl (-COOH) or thiol (-SH) group. As our target peptides do not contain any cysteines with thiol groups, a modification can be performed either on the primary amine or the carboxyl group. Most chemical tagging reactions for isotope labeled quantification strategies, such as iTRAQ, TMT or dimethyl labeling, are directed against the primary amine group. The yields of chemical labeling with these reagents are >97% (173-176). To be able to exploit similar reaction principles and to keep the cost of labeling reagents low, a reaction based on the dimethyl labeling mechanism was selected.

Relative quantification with dimethyl labeling in proteomics is based on the reaction of several isotopomers of formaldehyde and cyanoborohydride, with primary amines, which results in the introduction of two methyl groups. Biological samples to be compared are labeled on primary amines with different isotopomeric methyl groups and mixed in one sample afterwards. Thus, the relative comparison between different biological samples in one LC-MS analysis is possible.
Dimethyl labeling is based on the reaction of primary amines with aldehydes to form imine or Schiff bases, which can be reduced with sodium cyanoborohydride ($NaBH_3CN$) or sodium triacetoxyborohydride ($NaBH(OCOCH_3)_3$) to amine (Figure 24 A) to result in a new C-N molecular bond (177-179). The reaction takes place on primary amines of the N-terminus and lysine (Lys) side chains with formaldehyde following the scheme in Figure 24 B.

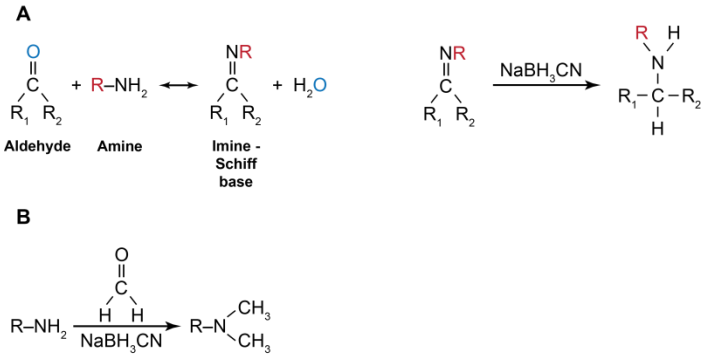


Figure 24. Formation of C-N bond through imine formation followed by reduction and chemical reaction of dimethyl labeling.
A) General chemical reaction of a new C-N bond formation through imine – Schiff base. An aldehyde reacts with an amine to form an imine or Schiff base, which can be reduced to a new amine with $NaBH_3CN$. Adapted from (173). B) A specific case of imine formation in dimethyl labeling reaction, which takes place on primary amines (peptide N-terminus and Lys side chains). Adapted from (176). R: remainder of the molecule.

The primary amines on the N-terminus and on the Lys side chain have different pKa values, meaning that they react in different pH environments. The N-terminal amine group readily reacts at lower pH, whereas the primary amine group on the Lys side chain at higher pH. The optimal pH for nearly 100% conversion of dimethyl labeling on both the N-terminus and on the Lys side chain was reported to be 5 to 8.5 (176).

The reaction can be performed in-solution, online on the LC column or on resin (176). The latter was chosen for this project for several reasons. Firstly, peptides from an IP eluate are bound on the reverse phase micro-column for peptide isolation anyway. Secondly, remaining reagents can be easily removed by washing the resin after the reaction and modified peptides subsequently eluted from the resin for further processing. Thus, at least one drying step of peptides from IP eluates is avoided and losses due to repetitive drying and dissolving steps minimized.

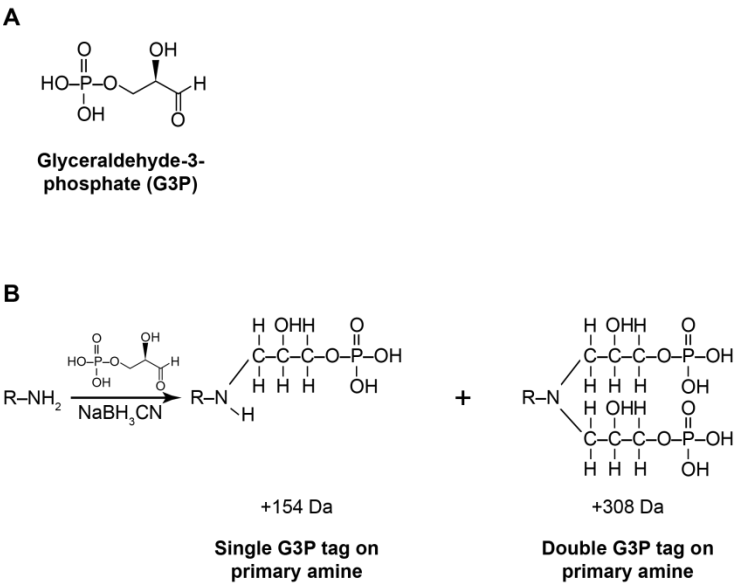


Figure 25. Chemical formula of glyceraldehyde-3-phosphate (G3P) and hypothesized products
A) Chemical formula of glyceraldehyde-3-phosphate (G3P) selected as the tagging reagent, B) Hypothesized products after G3P labeling of primary amines.

In order to exploit the dimethyl labeling chemical principles for this project, another molecule was needed instead of formaldehyde. The required functional groups in the tagging molecule were an aldehyde group, which was necessary for the reaction to occur, and a phospho-group necessary for TiO_2 separation of peptides from CHAPS. Moreover, the reagent should ideally be commercially available. We found glyceraldehyde-3-phosphate (G3P), whose formula is depicted in Figure 25 A. We hypothesized that the reaction will result in products carrying a single or a double G3P tag on the primary amine (Figure 25 B).
Before any modification was tested on the IP eluate, it needed to give optimal results with synthetic peptides. Therefore, all optimization experiments described below were conducted with the HLA-A2 HPV16 synthetic peptide mixture (3.1.14.1) and synthetic endogenous peptides (3.1.14.4).

Before the G3P labeling experiment was performed on the resin, the dimethyl labeling reaction was tested on the Oligo R3 micro-column. Furthermore, reaction buffers with different pH were compared to test effects on the reaction yields. 100 pmol/peptide of the HLA-A2 HPV16 synthetic peptides were bound on three micro-columns. The reaction was performed as described in section 3.2.4.5.1, with the only change that 60 µL 0.05 M $NaCH_3COO$ with pH 2.8, 5.5 or 8.2 were mixed with 0.6 µL 4% (vol/vol) formaldehyde and 0.6 µL 0.6 M $NaBH_3CN$ and added on the top of the micro-columns. An aliquot (app. 10 µL) of the reaction mixture was pressed through by overpressure created with a syringe mounted on the top of a micro-column. The micro-column was incubated at RT for 10 – 15 min before the reaction solution in the micro-column was refreshed with a new aliquot (app. 10 µL). The remaining reagents were washed away with 0.1% TFA in water, peptides eluted with 35% ACN/0.1% TFA, vacuum dried and analyzed with the speLC-Q Exactive instrument. Results were processed with the Skyline software, which searched for the HLA-A2 HPV16 and endogenous synthetic peptides with dimethyl modifications on the N-terminus and Lys. All peptides were modified only on the N-terminus in the reaction with pH 2.8, whereas at pH 5.5 the conversion partially happened also on Lys side chains, resulting in mixed products. Peptides fully converted on the N-terminus and Lys side chains at pH 8.2. The reaction gave almost 100% conversion rate.

Based on the results of the dimethyl labeling reaction, the reaction with the chosen labeling reagent G3P was performed, following the protocol described in the section 3.2.4.5.1 and similarly as described above. The pH values were adjusted to 3.5, 5.5 and 7.2. Peptides were eluted from the micro-columns after the reaction and analyzed on the speLC-Q Exactive instrument. Results were inspected manually and analyzed for product intensities using the Skyline program, searching against target peptides with possible modifications. The results showed the presence of the singly G3P labeled product (Figure 26) but also a product which had a 14 Da higher molecular mass than expected (G3P+14). The MS^2 spectra of a representative peptide $E6_{9\text{-}19}$ FQDPQERPIKL, showing a successful modification at pH 5.5, are depicted in Figure 26 A, B and C. It can be observed that the peptide was modified on the N-terminus, as the characteristic peak for N-terminal phenylalanine (F) at 120.08 m/z shifted for 154.01 to 274.09 or for 168.02 to 288.10 m/z (marked in red in Figure 26 A, B and C), whereas most other characteristic peaks remained unchanged. Surprisingly, the modification

was identified only on the N-terminus, but not on the Lys side chain, also in the sample with increased pH. This can be deducted from the dominant y_8 peak at 980.59 m/z, which did not shift for 154 or 168 m/z.

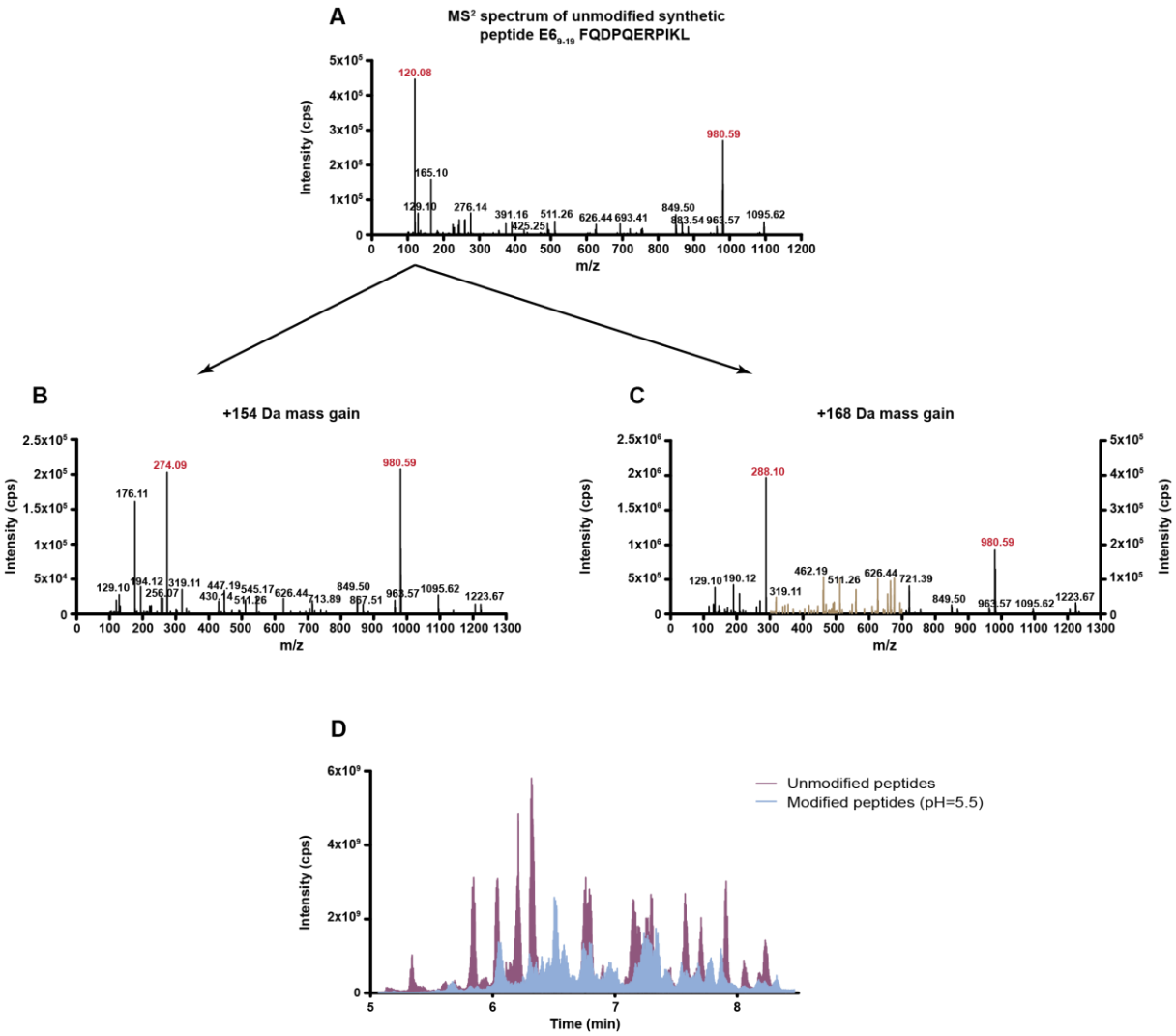


Figure 26. G3P tagging of the peptide E6$_{9-19}$, FQDPQERPIKL, resulted in successful modification on the N-terminus.
HLA-A2 HPV16 peptides were bound to the Oligo R3 micro-column and subjected to chemical tagging with G3P. Peptides were washed and eluted prior to speLC-Q-Exacitve analysis. A) MS^2 spectrum of unmodified FQDPQERPIKL, B) MS^2 spectrum of singly G3P modified peptide, C) MS^2 spectrum of G3P+14 modified FQDPQERPIKL. The light brown peaks are plotted to the axes on the right for better representation of characteristic low intensity peaks. N-terminal phenylalanine and its modified counterparts together with the y_8 ion, are marked in red in A-C. D) Overlaid LC-MS^2 chromatographic profile of unmodified peptides and peptides modified at pH 5.5, showing the decrease of signal intensity between samples.

The signals for doubly G3P labeled products were almost as low as those from the background, which could be due to the reaction predominantly resulting in singly labeled products, or impaired ionization of the doubly labeled molecule (180-182) due to the introduction of two phospho-groups. Furthermore, small amounts of peptides with a mass increase of 28 Da were detected in samples of the reaction at pH 7.2. The amounts of various species were pH dependent such that singly G3P (+154 Da) labeled product was most abundant at pH 3.5, and G3P+14 (+168 Da) modified products most abundant at pH 7.2.

All G3P modified peptides had reduced signal intensities relative to the unmodified counterpart, owing to the reduced ionization because of the introduced phospho-group (180-182). Therefore, the overall intensity of the LC-MS chromatographic profile after modification was lower. Furthermore, its complexity was increased, due to the presence of more products than in the sample of unmodified peptides (Figure 26 D).

To investigate how introduced phospho-groups hindered ionization of G3P modified peptides an additional step of dephosphorylation with alkaline phosphatase was introduced into the workflow (180-182). Ideally, enzymatic dephosphorylation would be performed directly after TiO_2 pull down in the TiO_2 elution buffer with appropriate pH adjustment to prevent additional losses due to drying of the sample prior to the enzymatic reaction. To investigate this possibility, a test was conducted where G3P-modified peptides after reaction at pH 5.5 were resuspended in TiO_2 elution buffer and the pH was adjusted to 9 with formic acid. A first aliquot was subjected to direct drying in the vacuum centrifuge as a reference, whereas a second aliquot was incubated with alkaline phosphatase for 1 h at 37 °C and with mixing at 400 rpm. Samples were subjected to analysis with the speLC-Q-Exactive. Data were inspected manually and analyzed for product intensities using the Skyline program, searching against target peptides with possible modifications.
The masses for modifications after dephosphorylation were expected to be increased by 74 Da (dephosphorylated single G3P), 88 Da (dephosphorylated G3P+14) or 148 Da (dephosphorylated double G3P) compared to the unmodified counterpart. Results showed that most of the peptides were dephosphorylated, resulting in two to ten-fold gains in signal intensities for the dephosphorylated single G3P (+74 Da) peptides (Figure 27) and similarly for dephosphorylated G3P+14 (+88 Da) peptides. Signal gains were highest for peptides containing no basic amino acids. Furthermore, peaks for doubly G3P modified dephosphorylated peptides (+148 Da) could now be observed. It is important to note that a small portion of the G3P modified peptides underwent spontaneous dephosphorylation prior to the enzymatic reaction (purple bars in Figure 27).
Manual inspection of spectra revealed the presence of additional side products which were not expected from the chemistry of the reaction and were not observed prior to enzymatic dephosphorylation. Peaks with the following masses were detected on primary amines: 44 Da, 58.1 Da and 118 Da. The masses correspond to the normal G3P expected modifications after dephosphorylation 74 Da, 88 Da and 148 Da, respectively, but with a mass loss of 30 Da each. MS^2 spectra of expected and side products are depicted in Figure 28 for the representative peptide $E6_{9\text{-}19}$ FQDPQERPIKL.

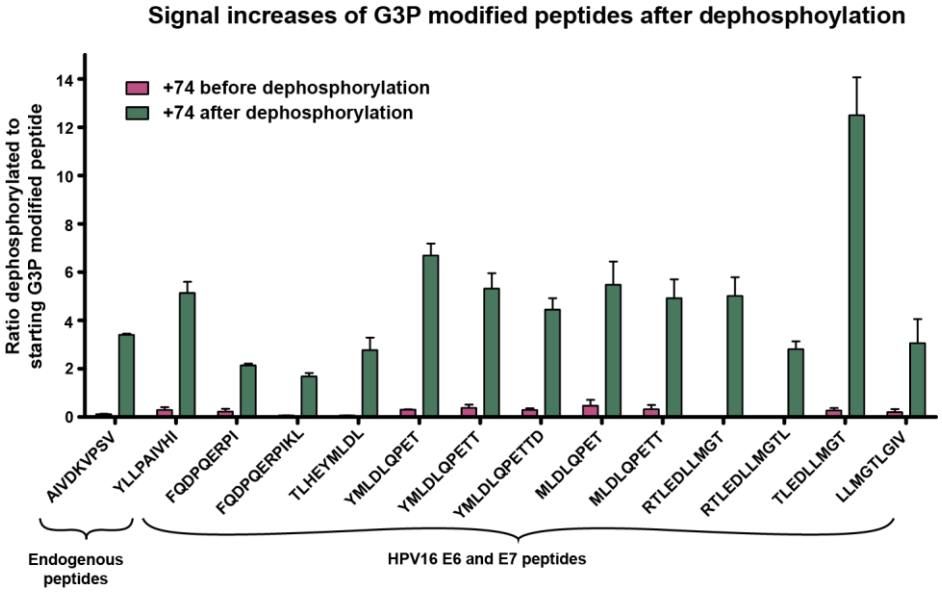


Figure 27. Enzymatic dephosphorylation of G3P modified peptides increased signal intensities.
G3P modified peptides were resuspended in TiO_2 elution buffer and pH was adjusted to 9. One aliquot was subjected to direct vacuum drying as a reference, whereas the second was enzymatically dephosphorylated with alkaline phosphatase for 1h, 37 °C, 400 rpm. Samples were analyzed with the speLC-Q-Exacitve instrument and data quantified with the Skyline program for peptides with the single dephosphorylated G3P tag (+74 Da). Before dephosphorylation: spontaneous dephosphorylation before enzymatic reaction; after dephosphorylation: product generated after enzymatic dephosphorylation. Error bars represent the SD from two technical replicates.

The single G3P dephosphorylated (+74 Da) peptides and side products with mass increases of 44 Da, 58 Da or 88 Da acquired the modification on the N-terminus (Figure 28 A – D), as the N-terminal phenylalanine mass changed for the respective mass, whereas the characteristic y_8 peak (980.59) did not undergo any mass shifts. Importantly, the mass increase of 88 Da can originate from dephosphorylated G3P+14 and also from a peptide with doubly G3P dephosphorylated modification and the mass loss of 30 Da on each of the G3P dephosphorylated tags (+2x44 Da). They could be located on the N-terminus and render the same MS^2 spectrum with the characteristic peak at 208.13. If the +88 Da modification occurred with a distribution of the G3P tags between the N-terminus and a Lys side chain, then the products would be distinguishable in the MS^2 spectra (Figure 28 F and I). The spectrum in Figure 28 F shows a +88 Da product on either the N-terminus alone or as a distribution of two 44 Da modifications on the N-terminus and Lys with mass shifts from 120.08 to 164.11 or 208.13 for +44 Da or +88 Da tags, respectively, and 980.59 to 1024.62 for a +44 Da tag. The spectrum in Figure 28 I shows a +88 Da product again on the N-terminus alone or as a distribution of +74 and +14 Da modifications on the N-terminus and Lys with mass shifts from 120.08 to 194.12 or 208.13 for +74 Da or +88 Da tags, respectively and 980.59 to 994.61 for a +14 Da tag. These results showed that the +14 Da group bound to the primary amine either in combination with the G3P tag or alone. We also observed a product with +28 Da on primary amines at higher pH.

The doubly G3P dephosphorylated peptide (+148 Da) appeared with the modification either only on the N-terminus as in Figure 28 G with a peak at 268.16 and no change at y_8 (980.59), or as a distribution of the single G3P dephosphorylated tag on the N-terminus and Lys (Figure 43 in the Appendix).

The side product with the mass increase of 118 Da was a mixture of three different products as seen in Figure 28 H. The +118 Da was a combination of a single G3P dephosphorylated (+74 Da) tag and its counterpart with the 30 Da mass lose (+44 Da). As seen for +88 and +148 Da tags, they were located either on the N-terminus or distributed among the N-terminus and Lys, which was also observed for the +118 Da modification. The observed mass shifts were from 120.08 to 164.11, 194.12 or 238.14 for +44 Da, +74 Da and +118 Da tag on the N-terminus, respectively, whereas the mass shifts from 980.59 to 1024.62 or 1054.63 for +44 Da or +74 Da tag on the y_8 fragment, respectively.

As reactions gave various side products, a new experiment was conducted in order to determine a most optimal pH where one of the species should have a dominant signal. The same experimental set up for the G3P chemical modification, as described above, was performed for the following pH conditions: 1.5, 2, 2.5, 3, 3.5, 4, 5.5, 7 and 8. 100 fmol/peptide were bound on each of nine Zorbax micro-columns and one aliquot was vacuum dried directly for relative comparison of signal intensities of unmodified and modified peptides. After G3P modification on the Zorbax micro-column, half of the eluate was subjected to direct vacuum drying and another half to peptide isolation with TiO_2 pull down and enzymatic dephosphorylation. All samples were measured with the speLC-Q Exactive instrument.
The most intense products after G3P modification on the Zorbax micro-column before and after TiO_2 pull down for two representative peptides $E6_{9-19}$ FQDPQERPIKL and $E7_{11-19}$ YMLDLQPET are depicted in Figure 29 A and B. The highest signal intensities were detected for singly G3P (+154 Da) modified peptides after elution from the Zorbax micro-column and dephosphorylated singly G3P (+74 Da) modified peptides after the TiO_2 isolation and dephosphorylation. In both cases, the intensities were higher when the G3P modification was performed at lower pH. The intensities decreased with increasing reaction pH. The intensities of dephosphorylated doubly G3P (+148 Da) modified peptides were low at low pH. The signal intensities for the singly G3P+14 Da (+168 Da) modified peptides after modification and elution from the Zorbax micro-column and dephosphorylated singly G3P+14 Da (+88 Da) modified peptides after TiO_2 isolation and dephosphorylation increased with the modification reaction performed with increasing pH until pH=5.5 and then the amounts decreased again. At pH >5.5 the product with 28 Da mass gain was observed for the $E7_{11-19}$ YMLDLQPET peptide. A proportion of $E7_{11-19}$ YMLDLQPET peptide was left unmodified after the reaction on the Zorbax micro-column at lower pH conditions, which was observed also for some other peptides.
As peptide $E6_{9-19}$ FQDPQERPIKL contains two primary amines, the possible products were more diverse than for the $E7_{11-19}$ peptide. Beside products described above, more of low intensity were observed at higher pHs, namely +162 Da (74+88 Da), +192 Da (44+148 Da), +132 Da (44+88=58+74 Da), +222 Da (3x74 Da), showing that one or two G3P molecules reacted with the same primary amine.

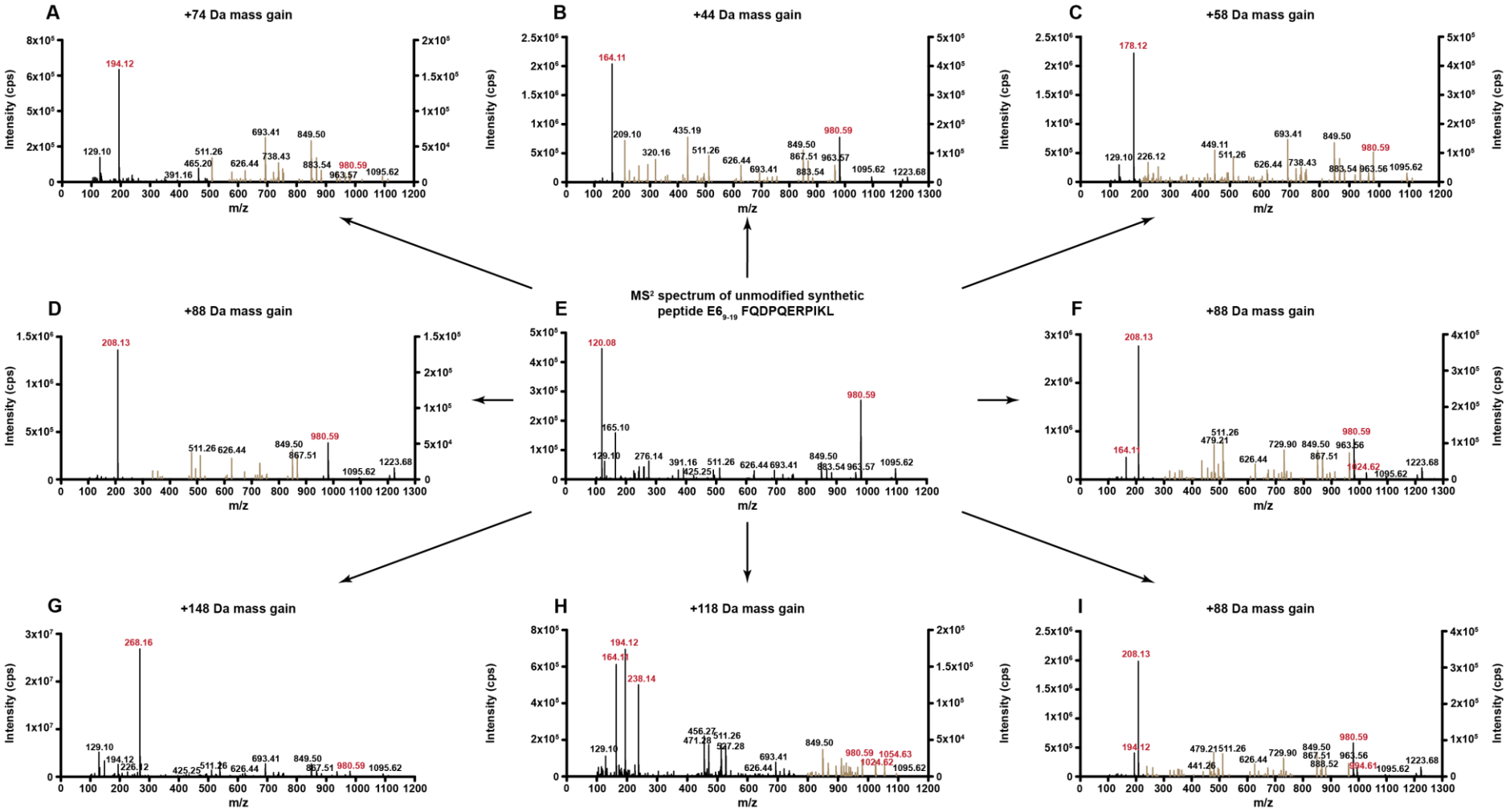
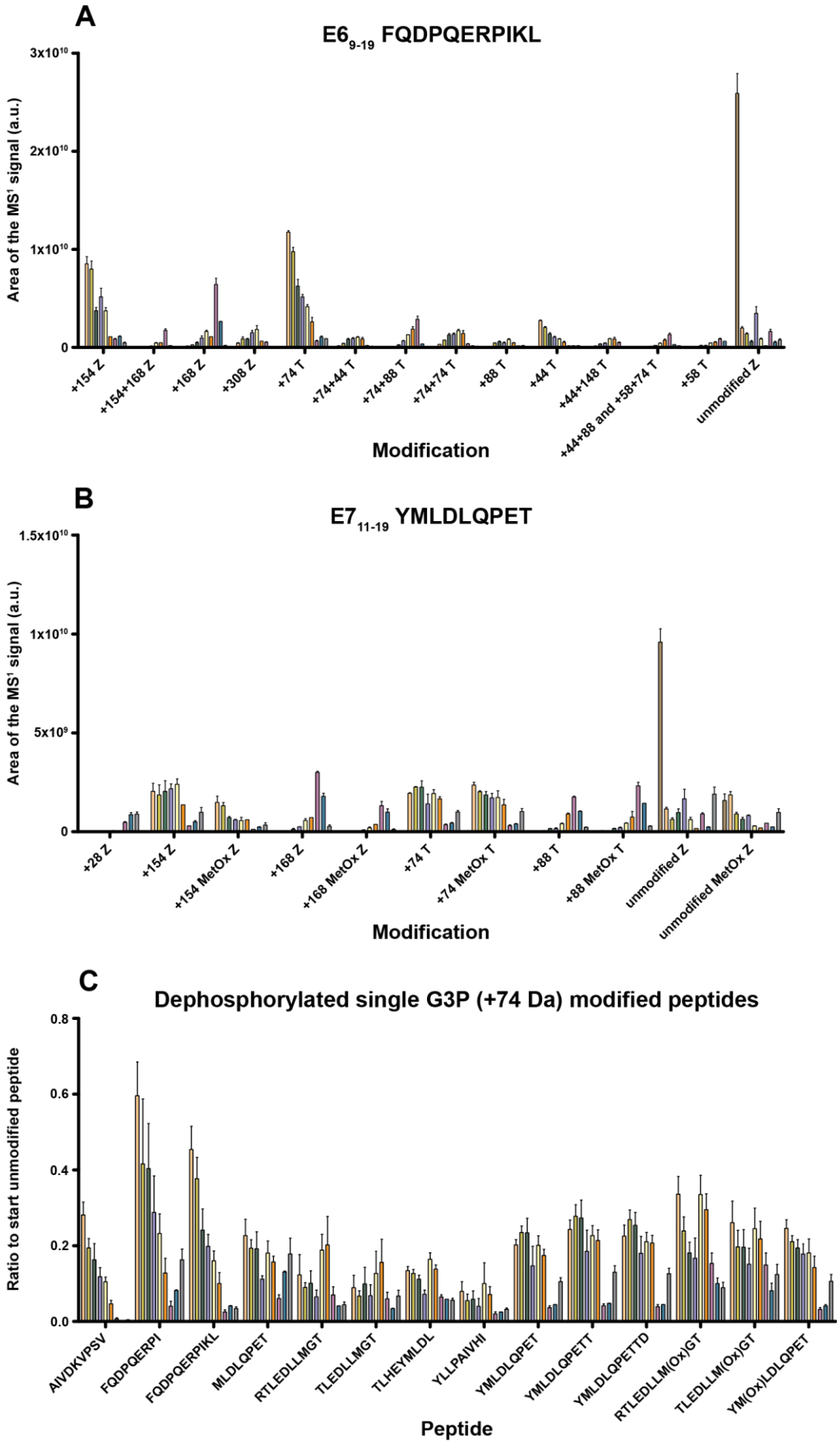


Figure 28. MS² spectra of observed products after dephosphorylation of G3P modified peptide E6$_{9-19}$ FQDPQERPIKL.
G3P modified peptides were treated with alkaline phosphatase before they were analyzed with the speLC-Q-Exactive instrument. The light brown peaks are plotted to the axes on the right for better representation of characteristic low intensity peaks. Characteristic peaks for N-terminal phenylalanine (120.08 m/z) and y_8 (980.59 m/z) together with their modified counterparts are marked in red to demonstrate the position of modification in the peptide.


A
E6 9-19 FQDPQERPIKL
Area of the MS¹ signal (a.u.)
3x10¹⁰
2x10¹⁰
1x10¹⁰
0
+154 Z
+154+168 Z
+168 Z
+308 Z
+74 T
+74+44 T
+74+88 T
+74+74 T
+88 T
+44 T
+44+148 T
+44+88 and +58+74 T
+58 T
unmodified Z
Modification
B
E7 11-19 YMLDLQPET
Area of the MS¹ signal (a.u.)
1.5x10¹⁰
1x10¹⁰
5x10⁹
0
+28 Z
+154 Z
+154 MetOx Z
+168 Z
+168 MetOx Z
+74 T
+74 MetOx T
+88 T
+88 MetOx T
unmodified Z
unmodified MetOx Z
Modification
C
Dephosphorylated single G3P (+74 Da) modified peptides
Ratio to start unmodified peptide
0.8
0.6
0.4
0.2
0.0
AIVDKVPSV
FQDPQERPI
FQDPQERPIKL
MLDLQPET
RTLEDLLMGT
TLEDLLMGT
TLHEYMLDL
YLLPAIVHI
YMLDLQPET
YMLDLQPETT
YMLDLQPETTD
RTLEDLLM(Ox)GT
TLEDLLM(Ox)GT
YM(Ox)LDLQPET
Peptide

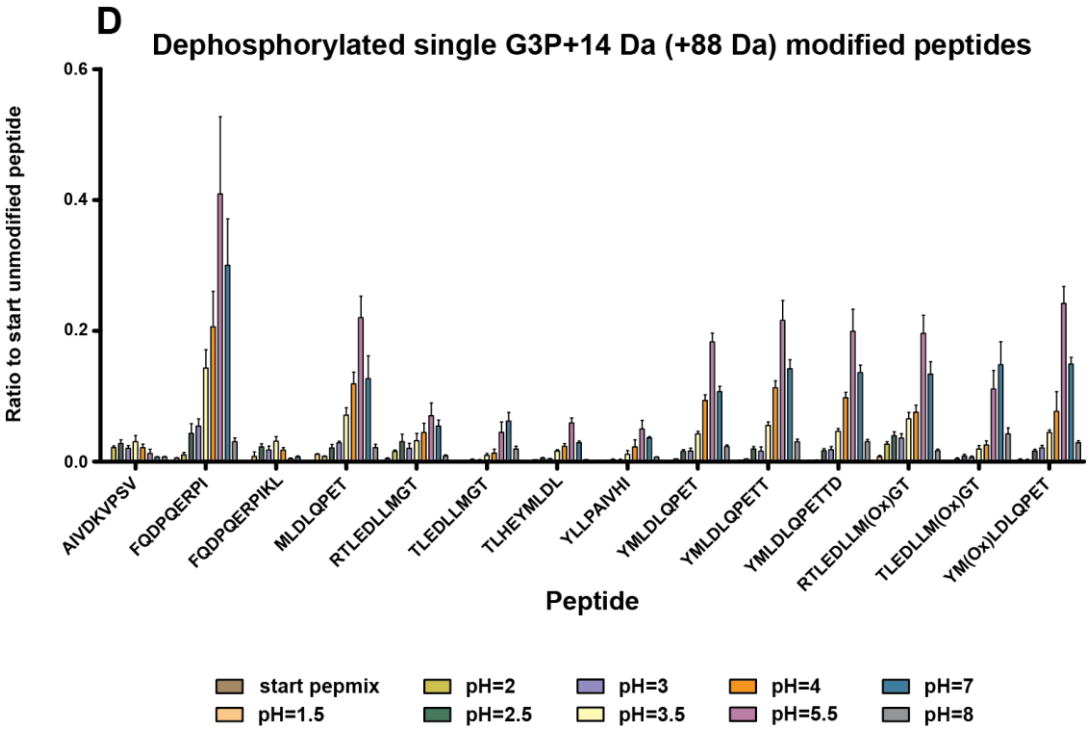


Figure 29. Signal intensities of various modifications on the representative peptides $E6_{9-19}$ and $E7_{11-19}$ and comparison of most intense signals for a selection of HLA-A2 peptides.
HPV16 HLA-A2 peptides were either bound on the Zorbax micro-column for G3P modification at different pH or vacuum dried for a reference. Peptides were eluted after the G3P reaction and one half was subjected to vacuum drying and the other half to TiO_2 pull down followed by dephosphorylation. Analysis was performed with the speLC-Q-Exactive instrument and data processed with the Skyline program. The MS^1 intensities of modified peptides tagged at different pH conditions were compared to the intensities of peptides before modification. A) $E6_{9-19}$ FQDPQERPIKL most intense peptide products; B) $E7_{11-19}$ YMLDLQPET most intense peptide products; C) A selection of other dephosphorylated singly G3P (+74 Da) modified peptides after TiO_2 isolation and dephosphorylation; D) A selection of other dephosphorylated singly G3P+14 Da (+88 Da) modified peptides after TiO_2 isolation and dephosphorylation; Z: peptide in a sample analyzed after the G3P reaction and elution from the Zorbax micro-column, T: peptide in a sample analyzed after TiO_2 isolation and dephosphorylation; M(Ox) and MetOx: oxidized methionine in a peptide. Error bars represent the SD from a minimum of two technical replicates.

However, the most dominant signals were those for singly G3P (+154 Da) modified and dephosphorylated singly G3P (+74 Da) modified peptides when the reaction was performed at lower pH. From these results, it was concluded that the most optimal pH for the reaction will be ~2 as at an even lower pH some of the peptides did not react effectively, leaving a higher proportion of unmodified peptides, whereas at higher pH the mixture of products got more complex.

The MS^1 signal intensities of other peptides with singly dephosphorylated G3P (+74 Da) or singly dephosphorylated G3P+14 (+88 Da) modifications were compared relative to the signal of unmodified counterparts (Figure 29 C, D). The trend of the pH effect on the amount of the dephosphorylated singly G3P (+74 Da) modified or of the dephosphorylated singly G3P+14 (+88 Da) modified peptides was similar as described above. The optimal pH for the reaction of other peptides was also in the low pH range between 1.5 and 2.5. Furthermore, sample processing introduced oxidation of methionine, which led to different retention times on the LC system, resulting in two separately eluting species. Thus, total signal intensities of methionine containing peptides were approximately equally distributed between the unoxidized and oxidized counterparts.

The final signal intensities for the pH conditions giving the highest signals were peptide dependent and were between 20% and 30% of those before the reaction for most of the peptides. Two peptides gave higher signal intensities and three had intensities below 20% (Figure 29 A, B).

The final experimental workflow for epitope extraction that was optimized through steps described in this subchapter is schematically presented in Figure 30. In brief, after isolation of HLA-peptide complexes from the sepharose beads, peptides are dissociated from the HLA complexes by acetic treatment. IP samples contain protein contaminants and detergent. The protein content can be reduced with sample processing on micro-colums packed with Zorbax resin, which does not bind bigger proteins, but still binds detergent and peptides. Peptides bound on the micro-column resin are then subjected to chemical modification with glyceraldehyde 3-phophate (G3P) at pH 2. Remaining reagents are removed by micro-column washing, and peptides are eluted with 30%-35% ACN/0.1% TFA in water. Peptides are separated from the detergent by binding to titanium dioxide (TiO_2) beads via a phospho-group introduced during chemical tagging. Subsequently, after elution from TiO_2 beads, peptides are subjected to enzymatic dephosphorylation, vacuum drying, desalting and LC-MS analysis.

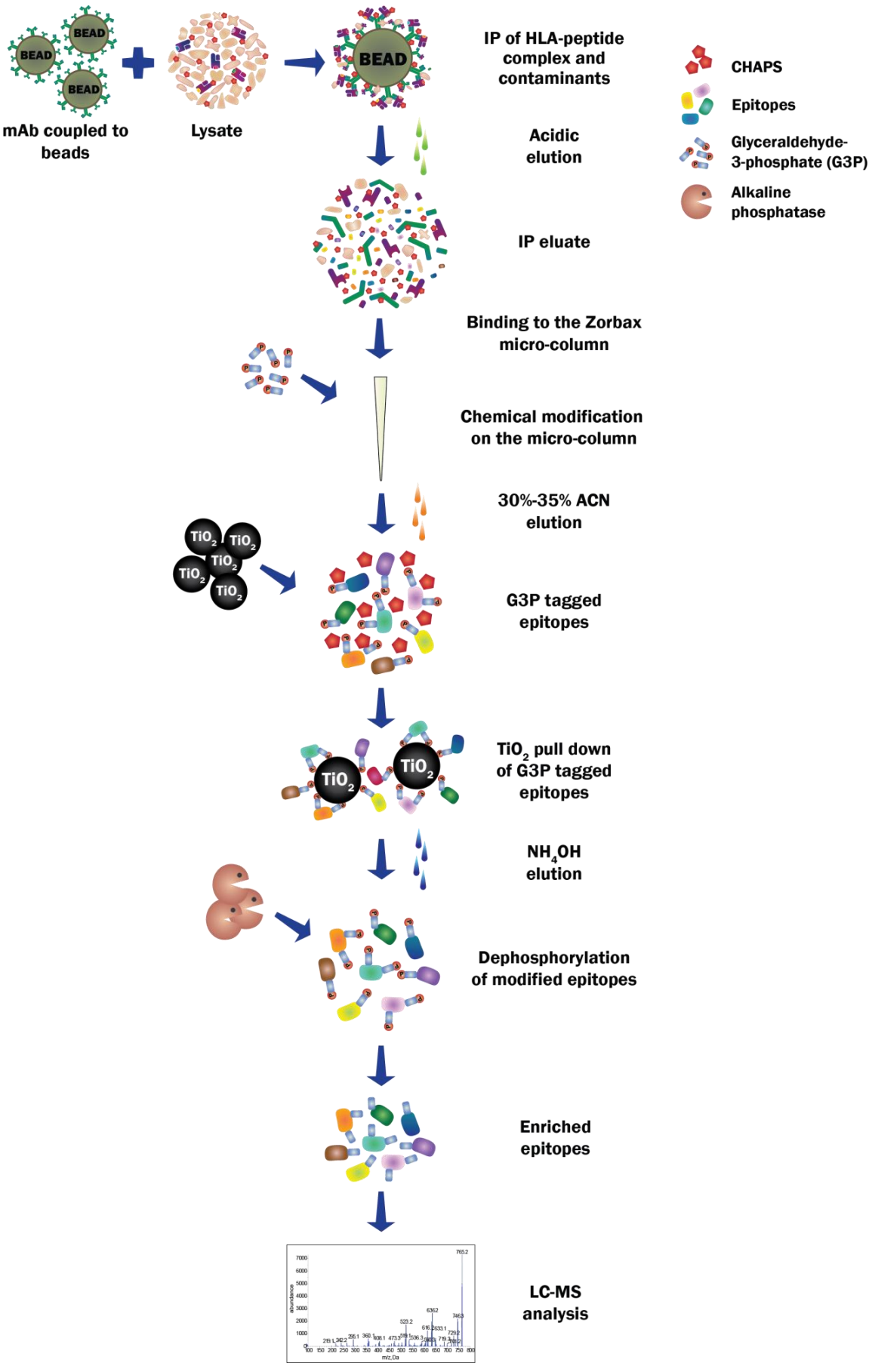


Figure 30. Experimental workflow for epitope extraction from an IP sample with the G3P tagging and TiO_2 pull down.

4.2.5.2. G3P chemical modification of IP samples

The optimized workflow for isolation of epitopes based on G3P chemical tagging described above was applied to IP samples of CaSki, SNU17 and SNU1000 cells. The experiment with CaSki cells was performed three times with 4, 8 or 18 IP samples and for SNU17 and SNU1000 cells once with 8 IP samples each. LC-MS^2 analysis was conducted with the speLC-Q-Exactive instrument.

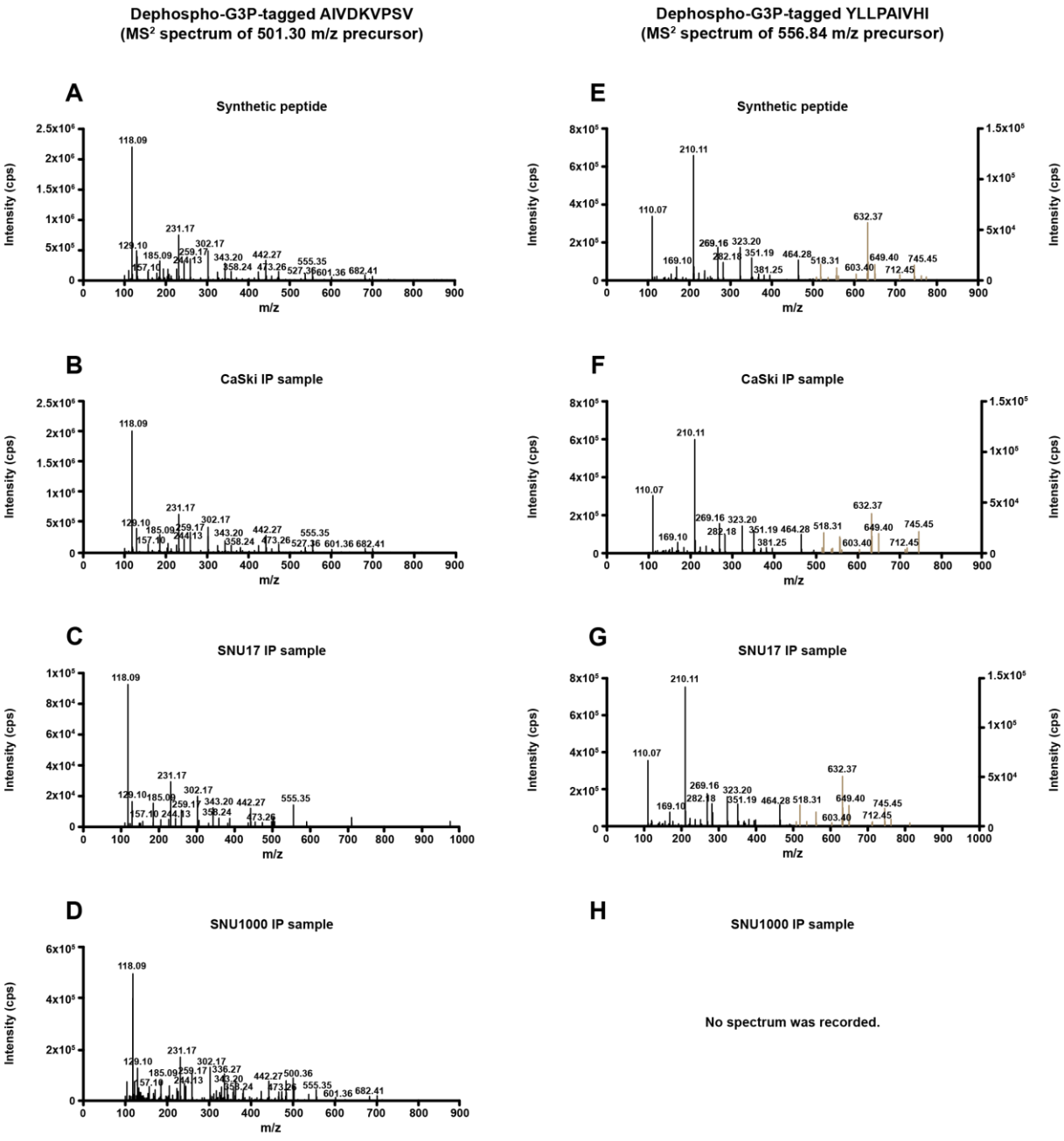


Figure 31. MS^2 spectra for endogenous peptides AIVDKVPSV and YLLPAIVHI successfully identified in HLA-A2 IP samples from CaSki, SNU17 and SNU1000 cells.

Samples were prepared following the established protocol for isolation of epitopes with G3P tagging, TiO_2 pull down and dephosphorylation. Subsequently, samples were analyzed with the speLC-Q-Exactive instrument. MS^2 spectra of IP samples (B – D, F – H) are compared to the MS^2 spectrum of the singly G3P-tagged dephosphorylated (+74 Da) synthetic peptide acquired in previous experiment (A and E). The light brown peaks are plotted to the axes on the right for better representation of characteristic low intensity peaks.

The signal for CHAPS was not detected in any of the measured IP samples. The endogenous AIVDKVPSV and YLLPAIVHI peptides were successfully identified in all CaSki and SNU17 IP samples as MS^2 spectrum for modified synthetic peptide match with the spectrum identified in the IP samples (Figure 31 A – C and E – G). Only peptide AIVDKVPSV was identified in SNU1000 IP samples, whereas peptides YLLPAIVHI was not detected (Figure 31 D, H).

None of the target HLA-A2 HPV16 peptides was detected in any of the IP samples analyzed. With this experiment we showed that the established experimental workflow successfully isolates epitopes from all IP contaminants in the IP sample.

4.3. Comparison of immunoprecipitation and direct elution of epitopes from the cell surface by acetic treatment

Direct elution of epitopes from the cell surface is another method of epitope isolation, which is performed easily and fast. On the other hand, it can be only employed for suspension cells and adherent monolayer cell culture cells, which can easily be collected by centrifugation or scraping. However, the method is not HLA-specific and elutes all epitopes and other cell surface molecules which dissociate in a low pH environment.

To assess the efficiency of the direct elution strategy, a comparison of HLA-A2 IP and direct elution of epitopes was conducted with 10^8 SNU17 cells each. Direct elution from cells and IP were performed as described in sections 3.2.3.3 and 3.2.3.1, respectively. Cells for direct elution of epitopes were once collected with trypsin treatment and once with gentle scraping on ice with 10% acetic acid.

Trypsin detaches cells but also cuts proteins and peptides from the cell surface, therefore less complex samples and lower amounts of eluted peptides were expected. Directly eluted samples were subjected to 2 kDa ultrafiltration and desalting on the Seppak cartridge as the volume of the sample in total was 9 mL. The IP samples were treated with 0.3% TFA in water and extraction was done on the Zorbax micro-column. IP or directly eluted samples was analyzed using $LC\text{-}MS^3$.

Directly eluted samples collected either by scraping or trypsin treatment resulted in high signals throughout the whole LC separation indicating the sample complexity (Figure 32), which is greatly reduced in the IP sample, as only minor background peaks and the two dominant peaks of endogenous peptides AIVDKVPSV and YLLPAIVHI were detected at 28.5 min and 47.2 min, respectively (Figure 32 B). It is important to note that the pressure on the LC device was elevated or reached its maximum during the separation of the directly eluted sample of cells collected by scraping, whereas this did not happen in the sample from trypsin treated cells or the IP sample.

The MS^3 methodology measures only pre-optimized transitions, resulting in high specificity, high sensitivity and low or no background or interferences. This was found to be true for the less complex IP sample, but not for the directly eluted samples, despite reducing their complexity with 2 kDa ultrafiltration. Literally, every programmed transition generated a complex MS^3 spectrum in the directly eluted samples (Figure 44).

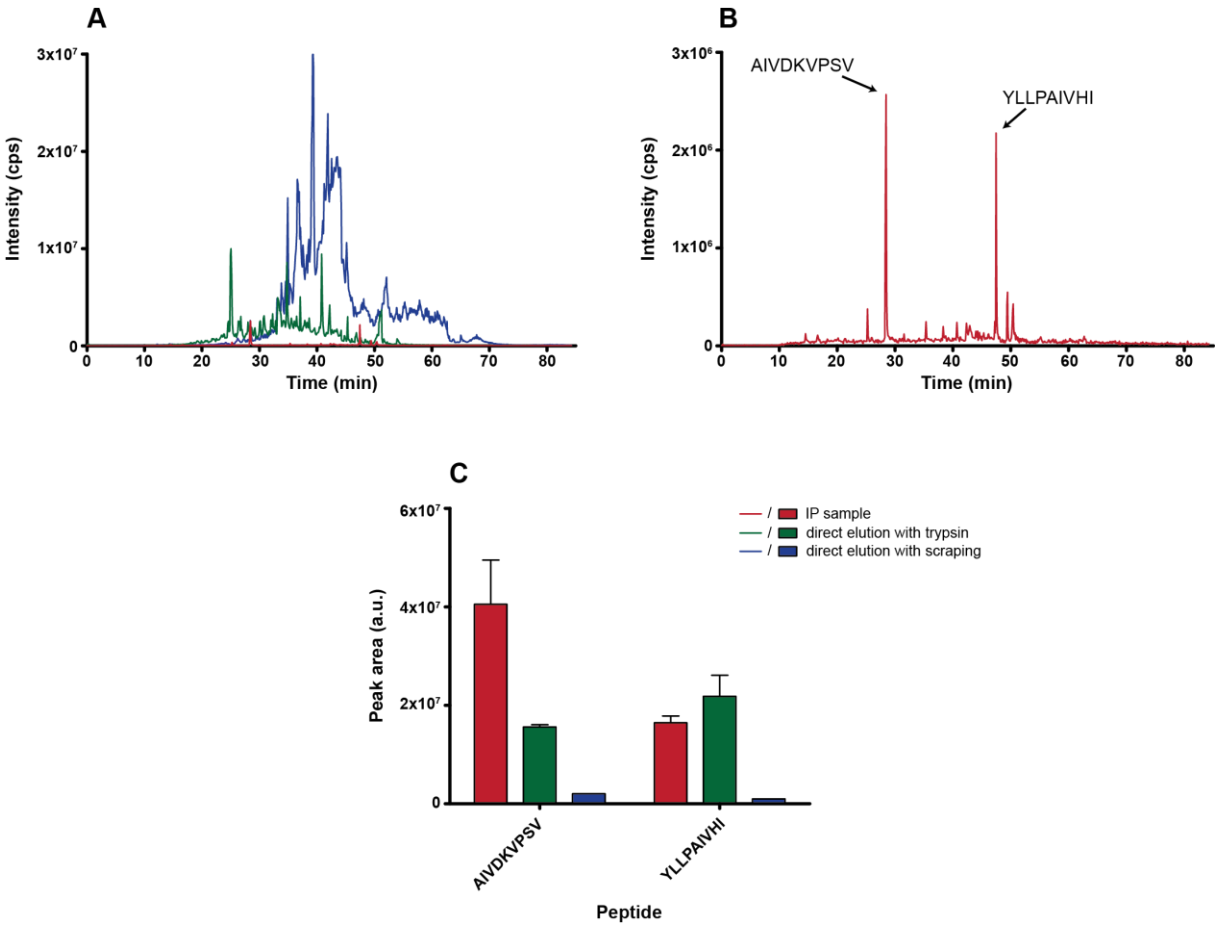


Figure 32. Comparison of direct elution of epitopes from the SNU17 cell surface and HLA-A2 IP.
Directly eluted samples were collected with scraping or trypsin treatment, subjected to 10% acetic acid treatment, 2kDa ultrafiltration and Seppak enrichment. The IP sample was eluted with 0.3% TFA and purified on the Zorbax micro-column. Samples were analyzed with the nanoAcquity UPLC-QTrap6500 platform. A) LC-MS^3 chromatographic profile of all three samples. B) LC-MS^3 chromatographic profile of the IP sample only. Peaks for the endogenous peptides AIVDKVPSV and YLLPAIVHI are marked. C) Signal intensities of the endogenous peptides AIVDKVPSV and YLLPAIVHI in all three samples. Error bars represent SD of minimal two technical replicates for the IP and directly eluted samples of trypsin collected cells and one technical replicate of directly eluted samples of cells collected with scraping.

The signal intensities of the endogenous peptides AIVDKVPSV and YLLPAIVHI were compared in all three samples (Figure 32 C). Intensities were highest in the IP sample and comparable or lower in the directly eluted sample of trypsin collected cells, whereas they were greatly reduced in the directly eluted sample of cells collected by scraping.

Based on the results described above, pre-fractionation strategies using different separation chemistry than the reverse phase materials are required for improved identification of target peptides from directly eluted samples. Isoelectric focusing (IEF), which separates peptides and proteins based on their isoelectric point (183, 184), was employed in this project. To this end, SNU17 cells were subjected to mild acetic treatment for direct elution of epitopes, desalting with a Seppak cartridge, IEF fractionation and LC-MS^3 analysis. The endogenous peptide YLLPAIVHI was successfully focused to fraction 15 with a pH between 7.9 and 8.2 (Figure 45), whereas the endogenous peptide AIVDKVPSV was detected in multiple fractions with low intensities. Target HPV16 E6 and E7 peptides were not detected in any of the fractions. As the endogenous peptide AIVDKVPSV was not successfully focused to one fraction, IEF was not considered optimal for pre-fractionation of directly eluted epitopes from the cell surface.

4.4. Identification of viral epitopes

4.4.1. HPV16 E6 and E7 epitopes in HPV16 transformed cells

As indicated in previous subchapters, surface presentation of peptide $E7_{11-19}$ in the CaSki, SNU17 and SNU1000 cell lines could not be confirmed with our experimental workflow. However, endogenous peptides AIVDKVPSV and YLLPAIVHI were detected in every HLA-A2 IP sample measured, indicating that the workflow for sample preparation and analysis worked. Moreover, the experiment with external pulsing of the HPV negative cell line BSM with the peptide $E7_{11-19}$ resulted in identification of the target peptide as well as endogenous peptides AIVDKVPSV and YLLPAIVHI, further confirming the efficiency of the workflow.

4.4.2. mCMV derived H-2D^b epitopes in the virus transfected cells

The aim of the cooperation project with the Prof. Čičin-Šain's group at the Helmholtz Centre for Infection Research (Braunschweig, Germany) was to confirm the presence of the known murine cytomegalovirus (mCMV) M45 protein derived epitope HGIRNASFI in complex with MHC I H-2D^b molecules in cells infected with different mCMV mutants: mCMV wild type (mCMV^{WT}), mCMV$^{M45I\rightarrow A}$ and mCMV$^{M45\ C\text{-}term}$. The mCMV mutants induced different HGIRNASFI-specific CD8+ T cell responses upon *in vivo* infection and differential activation of a M45-specific cytotoxic T-cell line (CTL) upon *in vitro* co-culture with virus-infected cells. As mCMV biology is not the main focus of this study, only a short description of the recombinant viruses used in the study is provided here. The mCMV^{WT} recombinant contained the HGIRNASFI epitope at its endogenous position within the M45 protein and it induced M45-specific CD8 T-cell responses *in vivo* but did not activate the CTL *in vitro*. The mCMV$^{M45I\rightarrow A}$ recombinant did not present the endogenous M45 epitope due to the introduced mutation. However, the C-term mutant expressed the HGIRNASFI epitope "ectopically" at the C-terminal end of the M45 protein. This mutant induced much stronger M45-specific CD8 T-cell responses *in vivo* and did activate the CTL *in vitro*.

The biological experiments and IP isolation, following our optimazed protocol, were performed in Braunschweig, whereas epitope elution, extraction, and analysis by LC-MS^3 were performed by me. The aforementioned mCMV viruses were used for infection of $1x10^7$ liver sinusoidal endothelial cells (LSEC) each. $1x10^7$ uninfected LSEC cells were used as a control. Cells were subjected to H-2D^b IP, acetic treatment, ultrafiltration, desalting with OMIX tips and LC-MS^3 data analysis.

Three endogenous H-2D^b restricted epitopes, AALENTHLL, FGPVNHEEL and KALINADEL, were monitored to ensure the quality of IP sample preparation and subsequent analysis. All three peptides were detected in every IP sample. FGPVNHEEL intensity was low, whereas AALENTHLL and KALINADEL, which co-eluted in the same chromatographic peak, had high intensity (Figure 33 A).

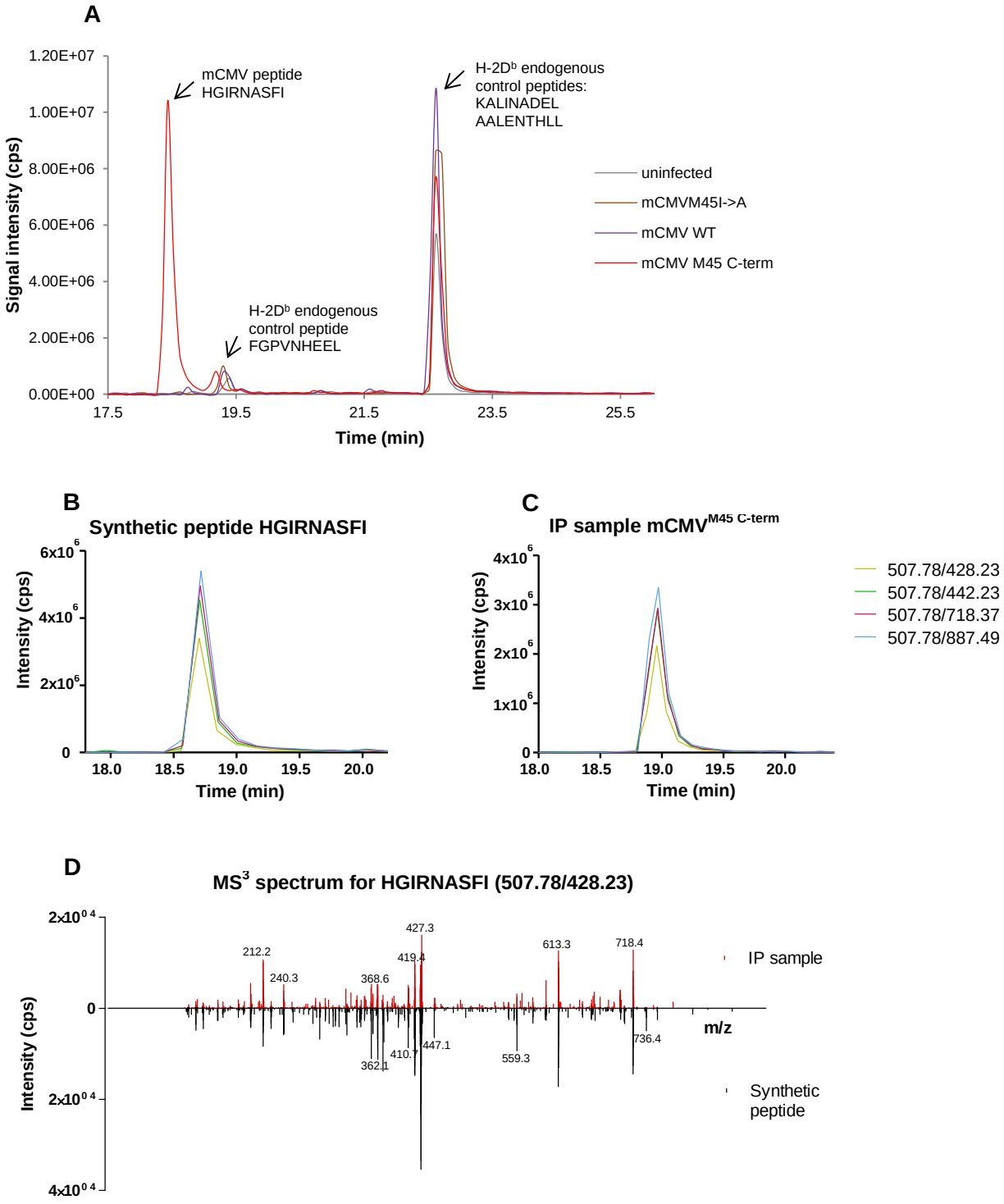


Figure 33. Detection of the mCMV H-2D^b restricted HGIRNASFI epitope.
LSEC cells were infected with one of three mCMV mutants or left untreated prior to H-2D^b IP. Samples were subjected to ultrafiltration and desalting with OMIX C18 tips prior to LC-MS^3 analysis with the nanoAcquity UPLC-QTrap6500 platform. A) LC-MS^3 chromatographic profile for all investigated samples. Arrows point to the mCMV (HGIRNASFI) and the H-2D^b endogenous peptides (low abundant FGPVNHEEL and co-eluting abundant AALENTHLL and KALINADEL); B) Extracted ion chromatogram for transitions: a_8^{2+} (507.78/428.23), b_8^{2+} (507.78/442.23), b_7-H_2O (507.78/718.37) and y_8 (507.78/877.49) for the synthetic peptide HGIRNASFI; C) Extracted ion chromatogram for transitions 507.78/428.23, 507.78/442.23, 507.78/718.37 and 507.78/877.49 for the peptide detected in the IP sample from cells infected with mCMV$^{M45\ C\text{-}term}$; D) MS^3 spectrum for transition a_8^{2+} (507.78/428.23) detected for the synthetic peptide HGIRNASFI (black) and for the same transition in the IP sample from cells infected with mCMV$^{M45\ C\text{-}term}$.

The mCMV H-2D^b restricted peptide HGIRNASFI was successfully identified in the IP sample from the LSEC cells infected with mCMV$^{M45\ C\text{-}term}$ in high abundance, whereas it was not detected in any other IP sample (Figure 33 A). The identity of the HGIRNASFI peptide in the LSEC mCMV$^{M45\ C\text{-}term}$ IP sample was confirmed with several criteria. First, the retention time for the peptide detected in the IP sample was within 15 s of that from the synthetic HGIRNASFI peptide. Second, the extracted ion chromatograms for all transitions in IP samples displayed the matching profile with those of the standard peptides (Figure 33 B,C). Finally, MS^3 spectra were monitored for four transitions and all of them matched between the synthetic HGIRNASFI peptide and the peptide identified in the IP sample. For clarity of presentation, only one of them is shown in Figure 33 D. MS^3 spectra for the IP sample and the synthetic peptide are presented on the same axis, with one turned upward and the other downward for better representation of the matching MS^3 fingerprint. Results were confirmed in three independent biological replicates.

4.4.3. Detection of HIV immunodominant epitopes

In a collaborative study with Dr. LeGall's group at the Ragon Institute of MGH, MIT and Harvard, (Cambridge, MA, USA) we aimed at targeted identification of well-known HLA-A2-restricted HIV-derived immunodominant epitopes, as reported in the Los Alamos Database HIV-1 (185). These epitopes were identified by ELIspot screening using long HIV peptides exogenously pulsed onto cells. The LeGall group has established an unbiased, untargeted MS-based approach to identify MHC-bound epitopes directly eluted from the surface of live HIV-positive cells (manuscript submitted[1]). They identified a number of potentially novel HIV epitopes presented by HIV transfected 293T cells expressing HLA-A2 and HLA-B7 molecules, and several of the previously reported HIV epitopes derived from the HIV-Gag and Pol proteins. Since an untargeted approach is less sensitive and non-specifically dissociated epitopes from all HLA class I molecules expressed on the surface are also contained in the sample, we applied our targeted LC-MS^3 analysis to investigate the surface presentation of the best known immunodominant HIV epitopes, specifically presented on HLA-A2 molecules. Furthermore, we investigated whether particular peptides bind to HLA-A2 or other HLA molecules with similar binding motives, particularly to HLA-B7. For our targeted approach, 1x10^8 293T cells were transfected with the HIV-R5 strain pseudotyped with VSVg or left untransfected as a control and subjected to HLA-A2 IP isolation. Epitopes were subsequently isolated by acetic treatment and binding to the Zorbax micro-column (section 3.2.4.2.3). The LC-MS^3 analysis was performed once monitoring peptides with HLA-A2 binding motives and once monitoring peptides with concurrent binding motives for HLA-A2 and HLA-B7. The isolated MHC peptides were analyzed in a single experiment. The endogenous HLA-A2 restricted epitopes AIVDKVPSV and YLLPAIVHI were monitored by LC-MS^3 to ensure the quality of IP sample preparation and subsequent analysis. The intensity of AIVDKVPSV was high in control 293T cells, whereas its abundance decreased to almost the detection limit upon HIV transfection, indicating displacement of self-epitopes by HIV-derived ones. YLLPAIVHI remained highly abundant in the IP samples obtained from control and HIV transfected 293T cells. Among the targeted immunodominant HIV HLA-A2 epitopes, EPFRDYVDRFY, FLGKIWPSYK and VLEWRFDSRL were successfully identified with low abundance in the IP sample

[2]Rucevic M, Kourjian G, Boucau J, Garcia Bertran W, Berberich MJ, Walker BD and LeGall S. MHC-bound HIV peptides identified from various cell types reveal common nested peptides and novel T cell responses.

of HIV transfected 293T cells, but not in the control 293T cells (Figure 34, Figure 35, Figure 36). Identities of these three HLA-A2 HIV epitopes were confirmed with the same criteria as for the mCMV HGIRNASFI peptide described in the previous subchapter. However, the retention times for peptides detected in the IP sample was shifted for up to 45 s relative to the synthetic counterparts. The extracted ion chromatograms for all transitions in IP sample displayed the matching profile with those of the standard peptides (Figure 34, Figure 35, Figure 36 A and D).

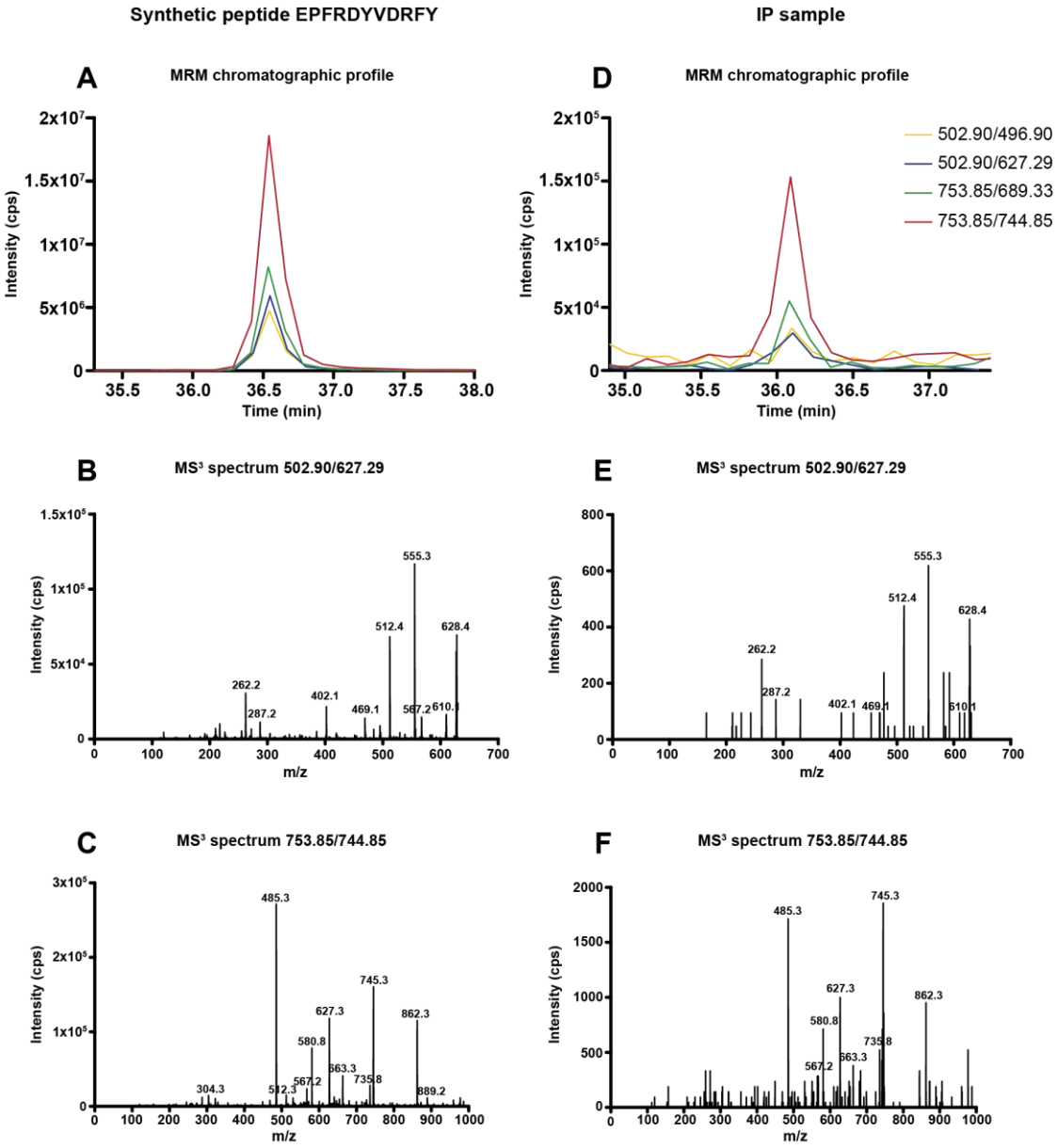


Figure 34. Detection of the HIV HLA-A2-restricted EPFRDYVDRFY epitope.
293T cells were transfected with the HIV-R5 strain pseudotyped with VSVg or left untransfected as a control prior to HLA-A2 IP. Samples were subjected to Zorbax micro-column purification and subsequently to LC-MS^3 analysis with the nanoAcquity UPLC-QTrap6500 platform. A) Extracted ion chromatogram for transitions: MH-H_2O^{3+} (502.90/496.90), b_5-H_2O (502.90/627.29), y_{10}^{2+} (753.85/689.33) and MH-H_2O^{2+} (753.85/744.85) for the synthetic peptide EPFRDYVDRFY; B) MS^3 spectrum for transition b_5-H_2O (502.90/627.29) for the synthetic peptide EPFRDYVDRFY; C) MS^3 spectrum for transition MH-H_2O^{2+} (753.85/744.85) detected for the synthetic peptide EPFRDYVDRFY; D) Extracted ion chromatogram for transitions: 502.90/496.90, 502.90/627.29, 753.85/689.33 and 753.85/744.85 for the peptide detected in the HLA-A2 IP sample from HIV transfected 293T cells; E) MS^3 spectrum for transition 502.90/627.29 detected in the IP sample; F) MS^3 spectrum for transition 753.85/744.85 detected in the IP sample of HIV transfected cells.

Overall intensity of all three identified HIV peptides was low. Therefore, not every MS^3 fragmentation resulted in a MS^3 spectra fingerprint with matching ratios of peak intensities equivalent to those obtained for the synthetic peptide. However, the most characteristic dominant peaks were detected in all spectra for all transitions. MS^3 spectra were monitored for four transitions per peptide but only two of them are shown in Figure 34, Figure 35, Figure 36 panels B, C, E and F for clarity reasons.

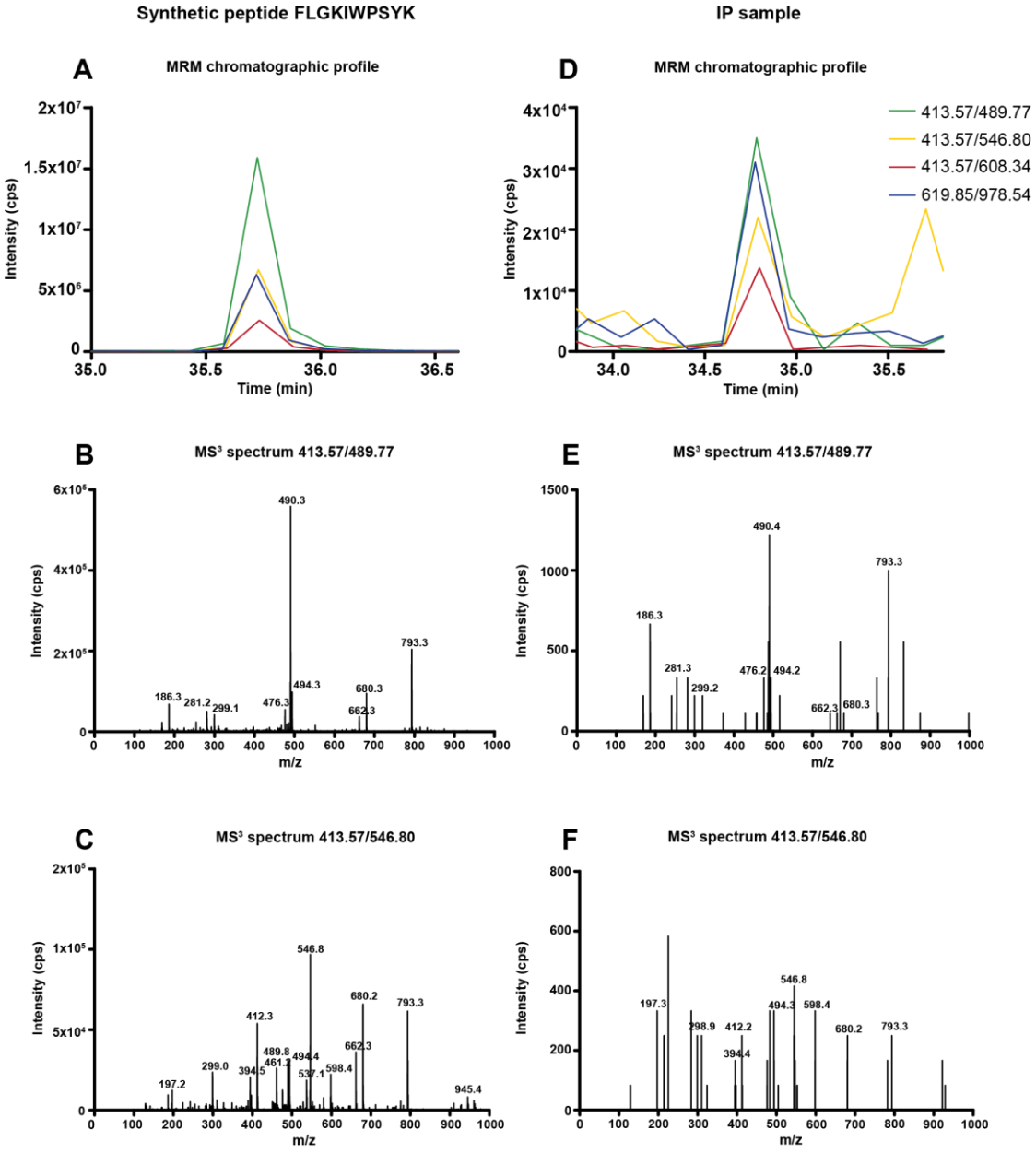


Figure 35. Detection of the HIV HLA-A2-restricted FLGKIWPSYK epitope.
293T cells were transfected with the HIV-R5 strain pseudotyped with VSVg or left untransfected as a control prior to IP. Samples were subjected to Zorbax micro-column purification and to LC-MS^3 analysis with the nanoAcquity UPLC-QTrap6500 platform. A) Extracted ion chromatogram for transitions: y_8^{2+} (413.57/489.77), b_9^{2+} (413.57/546.80), y_5 (413.57/680.34) and y_8 (619.85/978.54) for the synthetic peptide FLGKIWPSYK; B) MS^3 spectrum for transition y_8^{2+} (413.57/489.77) detected for the synthetic peptide FLGKIWPSYK; C) MS^3 spectrum for transition b_9^{2+} (413.57/546.80) detected for the synthetic peptide FLGKIWPSYK; D) Extracted ion chromatogram for transitions: 413.57/489.77, 413.57/546.80, 413.57/680.34 and 619.85/978.54 for the peptide detected in the HLA-A2 IP sample from HIV transfected cells; E) MS^3 spectrum for transition 413.57/489.77 detected in the IP sample; F) MS^3 spectrum for transition 413.57/546.80 detected in the IP sample of HIV transfected cells.

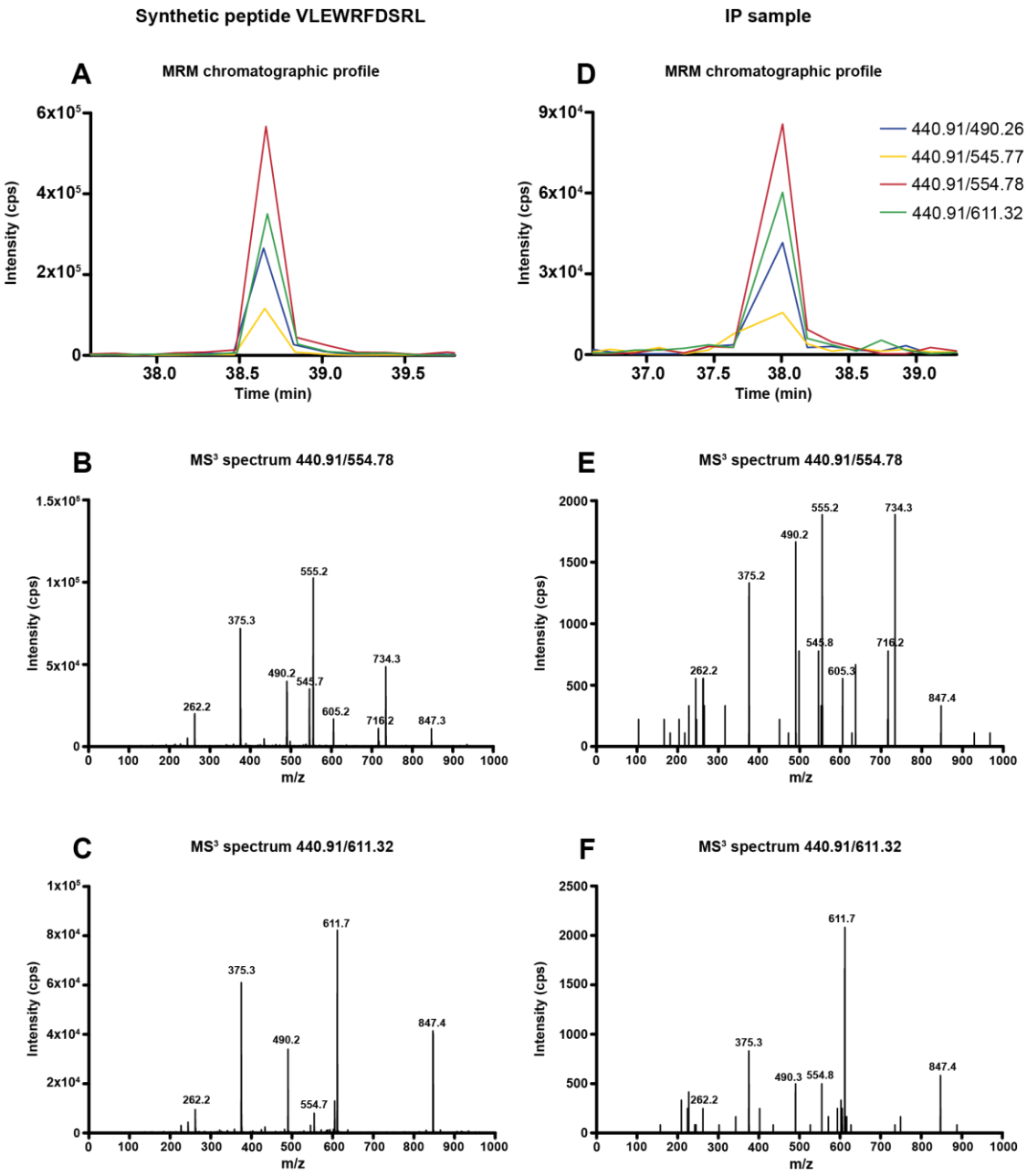


Figure 36. Detection of the HIV HLA-A2-restricted VLEWRFDSRL epitope.
293T cells were transfected with the HIV-R5 strain pseudotyped with VSVg or left untransfected as a control prior to IP. Samples were subjected to Zorbax micro-column purification and to LC-MS^3 analysis with the nanoAcquity UPLC-QTrap6500 platform. A) Extracted ion chromatogram for transitions: y_7^{2+} (440.91/490.26), y_8-H_2O^{2+} (440.91/545.77), y_8^{2+} (413.57/554.78) and y_9^{2+} (440.91/611.32) for the synthetic peptide VLEWRFDSRL; B) MS^3 spectrum for MS^3 transition y_8^{2+} (413.57/554.78) detected for the synthetic peptide VLEWRFDSRL; C) MS^3 spectrum for MS^3 transition y_9^{2+} (440.91/611.32) detected for the synthetic peptide VLEWRFDSRL; D) Extracted ion chromatogram for transitions: 440.91/490.26), 440.91/545.77, 413.57/554.78 and 440.91/611.32 for the peptide detected in the HLA-A2 IP sample from HIV transfected cells; E) MS^3 spectrum for MS^3 transition 413.57/554.78 detected in the IP sample; F) MS^3 spectrum for MS^3 transition 440.91/611.32 detected in the IP sample of HIV transfected cells.

Peptide FLGKIWPSYK was detected with the lowest abundance among the three identified peptides. Therefore, its spectra in Figure 35 F only contain the dominant peaks but their relative intensities do not match the relative intensities of the synthetic peptide, as the peptide was at the border of detection. Due to the better MS^3 spectrum quality in Figure 35 E and the presence of dominant peaks in the other two MS^3 transitions, this peptide was still confirmed to be present, but has to be validated in future experiments. The peptide EPFRDYVDRFY, which contained concurrent binding motives for HLA-A2

and HLA-B7 molecules, was identified in the HLA-A2 IP (Figure 34). This confirmed its surface presentation on HLA-A2 molecules on HIV transfected 293T cells. Whether it can also bind to the HLA-B7 molecule remains to be confirmed in subsequent independent experiments with HLA-B7-specific IP of HIV transfected 293T cells. The HIV Nef derived peptide VLEWRFDSRL, a well described immunodominant epitope, was also identified. It is the first directly identified Nef-derived epitope reported so far. Together, all three peptides were reliably detected.

5. Discussion

Major histocompatibility class I complexes (MHC I), in humans called human leukocyte antigens (HLA) I, present epitopes on surface of all nucleated cells to CD8+ T cells. Hence, they are a link between the interior of a cell and the immune system. They allow T cell recognition and discrimination between self and non-self in the case of infection and also between normal and tumor cells. Thus, activation of the immune system with peptides originally presented by tumor cells or persistently infected cells, as in the case of some viruses, is an attractive concept which became one of the pillars in modern immunotherapy approaches (89).

The first clinical trial based on the first epitope identified to be tumor-associated was initiated soon after their discovery and was administered as a peptide vaccination (186). MHC I epitopes were first identified indirectly in cell based assays, which did not provide proof that the immunologically active epitope is really presented on the cell surface. Therefore there was a great need for direct epitope identification. The first studies from 1990 were performed with the Edman degradation, where the binding motives of an allele were studied. Low numbers of peptide sequences were identified (67-69). The breakthrough happened two years later when the Edman degradation was replaced by mass spectrometry (MS) analysis, which allowed for identification of more epitopes (70, 187). Since then, leading laboratories in the field identified numerous tumor-associated epitopes and many of them entered clinical trials, with the IMA901 vaccination being the first therapeutic vaccine for renal cell carcinoma (76-78). Moreover, direct identification with MS analysis facilitated antigen processing machinery studies and contributed to the understanding of MHC binding motives (74).

Despite technological advances in the MS field, resulting in fast and sensitive high – resolution accurate mass (HRAM) instruments, samples today are measured in data-dependent MS acquisition with these instruments (known as shotgun), which lack reproducibility and can easily miss low abundant epitopes (92, 127).

To overcome the challenge of limited sample amounts and low epitope abundance, we aimed to develop a methodology for targeted identification of human papillomavirus (HPV)16 E6 and E7 epitopes presented on the cell surface. Furthermore, we aimed to extend the methodology to detect epitopes from other viruses to demonstrate its broad applicability. To this end, targeted MS analysis was employed, which is more sensitive than data-dependent analysis used in other studies. Our approach measures only pre-defined and pre-programed analytes, resulting in longer measuring times per analyte for more intense signals (114, 131-133, 188, 189). To improve the analysis further, we used an MS^3 scanning approach, which is more specific than MS^2 analysis due to minimized interferences from other co-eluting species and more sensitive due to increased signal-to-noise ratio (134-136).

High-risk types of HPV cause cervical cancer and other malignancies of anogenital and oropharyngeal epithelia. HPV16 is the causative agent in approximately 50% of all cervical cancers (190, 191), which is the third most common cancer in women worldwide (192-194). Furthermore, HPV16 has been associated with oropharyngeal head and neck squamous cell carcinoma, which incidence increased

over the past decade (195, 196). Although prophylactic vaccines were introduced, they do not protect against already existing infection. Moreover, they are not accessible to everyone, especially in the developing countries due to their high cost. Therefore, there is a need for effective treatment options.
HPV16-induced malignant transformation is mediated by two viral oncoproteins, E6 and E7, which are expressed in all stages of HPV-mediated malignancy and are foreign to the body. Thus, these proteins are attractive targets for therapeutic vaccine design. (25, 29, 197). Identification of HPV16 E6 and E7-derived epitopes presented on the surface of infected cells represents a step in the direction to reach this goal.

We approached the problem of targeting low abundant HPV16 E6 and E7 epitopes with a reverse immunology strategy (87), where the number of potential epitope candidates is reduced by *in silico* predictions and *in vitro* binding assay selection, such that only strong binding peptides are then monitored in a targeted fashion during MS analysis of isolated HLA I-epitope complexes from the surface of HPV16-transformed cells. In the scope of this project, a methodology for efficient epitope isolation, enrichment and purification was developed, together with a targeted highly specific and sensitive LC-MS^3 analysis.
HPV16 E6 and E7-derived T cell epitopes were expected to be of low abundance due to viral immune evading mechanisms, resulting in low expression levels of the oncoproteins E6 and E7 and down-regulation of HLA-peptides complexes on HPV16-transformed cells. Thus, special care was taken to maintain high experimental yields throughout the whole methodology development, be it for HLA-peptide complex isolation or epitope purification and enrichment. Furthermore, we aimed to establish an experimental pipeline with small material input amounts as it is planned to be extended to small biopsy specimens, allowing direct LC-MS identification of HPV epitopes in cancer patients.
As mentioned above, the developed methodology was later successfully applied to murine cytomegalovirus (mCMV) and human immunodeficiency virus (HIV)-derived epitope identification in two collaborative projects.

5.1. Optimization of immunoprecipitation

Immunoprecipitation (IP) of MHC peptides can be performed on columns or in suspensions in Eppendorf tubes. The on-column application is suitable for larger amounts of material, such as whole tumors (138, 198). As outlined above, our starting material amounts were expected to be low, therefore we decided to establish our workflow in the low-scale suspension-based platform.
For effective isolation of membrane proteins, the choice of the detergent in the lysis buffer is important. Cleavable detergents are employed in usual LC-MS proteomics studies. These detergents precipitate after sample acidification and can be easily pelleted, whereas the supernatant is used for downstream sample processing causing no interferences with the workflow. Unfortunately, all detergents from this group are denaturating, which would cause MHC I-peptide complex disassembly and peptide loss. Therefore, these detergents were not applicable to our project. The suitable detergent that we found, which is LC-MS friendly and maintains the protein 3D structure, was CHAPS. Thus, the employed lysis buffer was as in (67, 68, 91, 137, 138, 167, 168).

Other parameters for the optimal IP protocol, such as Ab:beads:MHC complex ratios and incubation times, were determined during this work. As shown in section 4.1, IP was successfully optimized for HLA-A2 molecules. The optimal ratio of Ab and beads was determined to be 20 µg Ab for 25 µL of pelleted beads. The ratio of Ab and beads was significantly lower than that reported by the producer, which was approximately 10 mg of murine Ab for 1 mL of pelleted beads, which equals 250 µg of Ab for 25 µL of pelleted beads. Presumably, this is due to higher affinity of the Ab used in the producer's study for sepharose beads. Moreover, efficient Ab-bead binding occurred after 2 h of incubation at RT (Figure 13). Sufficient incubation of the cell lysate with coupled Ab-bead complexes was determined to be 3 h, when the signal for the HLA heavy chain at 44 kDa was detected in the IP sample (Figure 14). It was also detected in the lysate before and after incubation with the Ab-bead complexes. This was due to WB staining performed with an anti-HLA I Ab that recognizes all types of HLA I molecules. On CaSki cells, HLA-A2, -A3, -B7, -B37 and -C7 are expressed. Thus, the band in the lysate before the IP consisted of all listed HLA molecules, and after IP of remaining HLA I molecules. However, the intensity of the 44 kDa band in the IP sample compared to the lysate was higher, showing successful enrichment of HLA-A2 (Figure 14). From this experiment, it is difficult to estimate how many HLA-A2 complexes were left uncaptured in the lysate after the IP procedure, as all HLA I molecules were stained in WB not only HLA-A2. Unfortunately, no suitable antibody to distinguish HLA-A2 from other alleles in WB experiments was available.

In order to capture most of the HLA-A2 molecules from the sample, a titration experiment was performed, where lysates of increasing numbers of cells were incubated with a fixed amount of coupled Ab-beads. The required number of cells to result in the highest LC-MS^3 signal intensities for endogenous peptides, thereby indicating saturation of most binding places on the Ab-bead complexes, was $6x10^7$ for CaSki cells and $3x10^7$ for SNU17 and SNU1000 cells (Figure 15). However, it may be possible that there are still some unbound HLA-A2 molecules left in the lysate after the IP.

HLA-epitope complexes dissociate after acetic treatment. Two acetic buffers described in the literature were compared, namely 0.3% TFA (67-69, 138) and 10% acetic acid, both in water (70, 187). The LC-MS^3 results were comparable (Figure 39 in the Appendix), meaning that the composition is not as important as the correct pH of the elution buffer, which has to be below pH 2.9 as described in (148, 149). It can be assumed that other buffers with the required pH would perform similarly. 0.3% TFA in water was chosen for our protocol.

All samples were subjected to ultrafiltration and desalting prior to LC-MS^3 analysis as described in (138, 198). The highly abundant HLA-A2-restricted endogenous peptides AIVDKVPSV and YLLPAIVHI, which were measured to control quality of sample preparation and LC-MS^3 analysis, were detected in all measured samples. In contrast, the target peptide HLA-A2 HPV16 $E7_{11-19}$ YMLDLQPET, which was identified by MS^3 on CaSki cells before (146), was not detected.

The IP protocol and epitope extraction with ultrafiltration and desalting was further validated by an experiment, where the target HLA-A2 HPV16 $E7_{11-19}$ YMLDLQPET peptide was externally loaded on the cell surface of HPV16- HLA-A2+ BSM cells. As demonstrated in Figure 16, the HLA-A2 HPV16 $E7_{11-19}$ YMLDLQPET target peptide was successfully detected in all samples where peptide was added to cells prior to the IP, indicating that the $E7_{11-19}$ peptide can be detected with the established

methodology. However, the ratios of identified target peptide between samples were app. fivefold (Figure 16 D), whereas the ratios of actually added amounts of peptide to cells were tenfold. This difference could be due to sample preparation or unexamined biological reasons.

We assumed that the $E7_{11-19}$ target peptide in the IP samples from HPV16+ cells adsorbed to the ultrafiltration membrane, which caused its loss and prevented LC-MS^3 identification due to low peptide amounts. Therefore the ultrafiltration step was omitted and the extraction was performed with OMIX tips. Unfortunately, the target peptide was still not detected. In order to reliably identify the target peptide $E7_{11-19}$, the sample input per one LC-MS^3 analysis was up-scaled, which lead to clogging of the LC column or severely elevated LC system pressure on all LC-MS platforms used in this study, indicating the necessity to examine other epitope extraction and purification methodologies. Moreover, plasticware with low affinity for protein was used to further minimize potential peptide losses in all sample preparation steps.

5.2. Epitope extraction, enrichment and purification strategies

In large scale epitome identification, epitopes are eluted from an immunoaffinity column, subjected to ultrafiltration and to desalting with C18 material in a pipette tip (138, 198). As described before, we aimed to replace the ultrafiltration step with other strategies which remove contaminating proteins and detergents, as we expected they were the major reason for clogging of the LC column. The widely applied ultrafiltration-desalting workflow removes proteins, but detergents remain in the sample despite extensive washing. It was reported that detergents influence reverse phase LC separation and MS detection of peptides (163, 199, 200), therefore we aimed to eliminate detergents and proteins with different approaches.

Hydrophilic interaction liquid chromatography (HILIC) is a variant of normal phase liquid chromatography which combines characteristics of three major LC methods; normal phase, reverse phase and ion chromatography. Hence, HILIC exhibits different retention and separation characteristics than reverse phase chromatography, which is the most widely used separation methodology for proteomics MS analysis (102). Two resins commonly used in proteomics, TSKgel and ZIC HILIC, with different chemistries of the stationary phase, were assessed (97, 102, 201). We assumed that the separation of epitopes from other sample components is possible, because they would differentially retain on the HILIC material.

The gradual elution from ZIC HILIC did not separate peptides from contaminants, as they were contained together in the same elution fractions, whereas the TSKgel HILIC separated peptides from proteins. Peptides were contained in the flow through and wash fractions, whereas proteins eluted in 60% or 70% ACN/0.1% TFA fractions (Figure 17, Figure 18). However, low mass contaminants and detergent were still present in the same fractions as the ones containing the peptides. Still, a reduction of sample complexity was achieved, making TSKgel HILIC a promising strategy for sample preparation to start with. However, up-scaling of the experimental workflow resulted in ineffective dissolving of dry IP eluate in the solvent with a high proportion of organic content, therefore, HILIC was not suitable for this project.

Acetone precipitation is a widely used method for protein isolation and buffer exchange in proteomics, which is required to minimize interferences of the downstream sample preparation steps, e.g. trypsin digest, where high urea contents would be hindering the reaction. Acetone precipitates proteins, which are pelleted in the centrifuge, and the supernatant containing undesirable compounds can be removed (162). We aimed to use this strategy for removing proteins from the peptides contained in the supernatant. Ethyl acetate precipitation was described as a method for detergent removal. Here, detergent diffuses from the aqueous solution into the organic solvent ethyl acetate. The two phases form distinct layers after short centrifugation. Layers can be easily transferred to new tubes, achieving the separation of detergent from components contained in the aqueous phase (163, 199, 200). We here aimed to remove peptides from detergent.
However, the result of the experiment was a distribution of peptides, detergent and other low mass contaminants over all fractions, whereas proteins were detected in the aqueous ethyl acetate bottom layer (Figure 41,Figure 42). Moreover, the intensities of all peptides and proteins in the fractions were significantly reduced compared to approximately 7% of unprocessed starting IP sample, indicating that this experimental procedure caused significant losses. The peptides probably remained in the small volume of the interface liquid which was left in the Eppendorf tube after separation of organic and aqueous layer after ethyl acetate precipitation. Taken together, the acetone – ethyl acetate precipitation was not suitable for extraction of peptides from an IP eluate due to high losses. The ineffectiveness of the experimental set up could be explained with the high hydrophobicity of our target peptides, which caused a distribution between organic and aqueous solvent.

Reverse phase chromatography is the method of choice for protein and peptide separation in proteomics. Reverse phase chromatography binds more hydrophobic peptides and proteins, whereas hydrophilic compounds, such as salts are not retained on the stationary phase. This allows for an easy removal of hydrophilic contaminants from the sample, which could cause interferences during LC-MS analysis. Furthermore, peptides and proteins are retained differently on the stationary phase, which facilitates separation (92). Therefore, we assessed several reverse phase applications to separate contaminants from epitopes by sequential elution.
As described before, OMIX C18 material pipette tips caused clogging of the LC analytical columns, when they were used without prior ultrafiltration of the IP eluate. Similarly, micro-columns packed with R2 and Oligo R3 reverse phase materials clogged after loading of small IP eluate amounts. Next, Seppak cartridges were assessed as they were used for solid phase extraction of peptides in (202). Loading of the IP eluate was easy and did not cause any backpressure increase or clogging. Seppak cartridges have bigger volumes and consequently higher binding capacities than the above described materials. To examine whether unproblematic loading was due to unsaturated Seppak resin or to the sample contaminants not being retained on the resin, IP eluate fractions were examined after gradual elution for their protein and detergent content. They revealed that cell components other than MHC I complexes unspecifically bound to the coupled Ab-beads (8.7. in the Appendix). In the ideal setting, this should not happen as a specific antibody was used to capture MHC I complexes only. Moreover, protein contaminants were mostly contained in the Seppak flow through fraction, whereas only small

amounts of proteins were identified in the flow through of the Oligo R3 material. This indicated that protein contaminants were not retained on the Seppak resin but they were on the Oligo R3 material, which clogged (Figure 19). Besides protein contaminants, also the detergent CHAPS was detected with an intense peak in fractions eluted with 30% and 40% ACN/0.1%TFA in water from the Seppak cartridge. These results showed that Seppak could be used as an alternative for reduction of IP eluate complexity as a big proportion of proteins was removed.

The main difference between both tested reverse phase materials was the pore size, which was 130 Å or 300 – 3000 Å for Seppak or R2 and Oligo R3, respectively. To be able to keep the experimental workflow in small scale to prevent sample losses on surfaces of the used equipment, a micro-column packed with Zorbax reverse phase material with a pore size of 80 Å was examined. Loading of the IP eluate was effortless for the amount equivalent to three IP eluates and could have been increased further. However, it is important to note that the backpressure increased slightly, indicating that eventually this material would have clogged as well. The amount of loaded material causing no clogging was significantly higher than that loaded on the R2 and Oligo R3 material. The loaded IP sample was gradually eluted and examined in a similar experiment as with the Seppak cartridge. It was hypothesized that Zorbax material should bind fewer proteins than Seppak due to its smaller pore size. This was indeed confirmed as all Zorbax fractions contained fewer proteins than the Seppak fractions (compare Figure 19 B and Figure 20 B)

The protein identifications results (8.7. in the Attachment) were comparable for Seppak and Zorbax, showing that numerous protein contaminants such as cytoskeleton proteins (tubulin, actin, plektin), keratin, ribosome subunits, histones and others were present in the IP sample. This further confirms the known fact that coupled Ab-beads have affinity not only for the target antigen but also for other proteins (203). The biggest proportions of identified proteins were represented by keratins. Keratins are known contaminants in MS proteomics introduced by insufficiently careful sample handling (204). In our case, the starting material was keratinocytes, which are rich in keratins anyway. Thus, we assume that most of the identified keratins in our samples stemmed from the sample source cells themselves.

β_2 microglobulin (β_2M) was the protein which was retained on both resins to the highest proportions. Its signal was highest in the fractions eluted with 35% and 40% ACN/0.1%TFA from Zorbax and Seppak. β_2M is the smallest detected protein in the IP eluate with a mass of 12 kDa. This indicates that the 80 Å pore size was big enough to bind β_2M. Presumably, materials with smaller pore size than the Zorbax material potentially bind less β_2M. This would allow processing of more IP samples through the same micro-column. One would need to determine the optimal pore size which will still allow epitopes to be retained but not β_2M. This approach could be a solution for separating peptides from proteins (instead of ultrafiltration) also in other applications, e.g. antimicrobial peptide detection from complex protein matrixes (205).

After removing proteins from the IP sample on the Zorbax micro-column, the detergent CHAPS still remained in the sample and eluted from the resin together with peptides at 23% – 27% ACN/0.1% TFA in water (Figure 22). Therefore, sequential elution of peptides and CHAPS was not possible as the

most hydrophobic peptides eluted with 30% ACN/0.1% TFA in water (Figure 21). The sequential elution strategy might be exploited for more hydrophilic epitopes, but only after empirical confirmation with synthetic peptides, proving that they all elute with the solvent containing less than 23% ACN/0.1% TFA in water. This criterion highly depends on HLA I binding properties, as some HLA I molecules are prone to bind more hydrophobic peptides (e.g. HLA-A2) than others (e.g. HLA-A3, -A11) (79).
We observed that CHAPS eluted later than all peptides from the speLC column (at 8.4 min) with an intense narrow peak, whereas it eluted earlier than the most hydrophobic peptides from the Zorbax material. Moreover, our collaboration partners from the Ragon Institute of MGH, MIT and Harvard (Cambridge, MA, USA) observed that CHAPS eluted in a broad peak over several minutes on their chromatographic system. This indicates that C18 materials from different producers have different retaining specificities and that examining C18 materials from several manufacturers could result in finding more suitable ones for sequential separation of all peptides from CHAPS.

In order to evaluate C18 extraction of peptides from IP samples, the different C18 materials were compared with ultrafiltration strategies which are commonly used for epitope analysis (138, 198). All samples gave intense CHAPS signals, confirming that these extraction strategies did not remove the detergent. Furthermore, the results showed that all separation strategies gave comparable outcomes with high input material amounts, whereas low input amounts had better recoveries with reverse phase isolation strategies. The best among those was separation with Zorbax micro-columns (Figure 23). The worst recovery was observed for small sample input on ultrafiltration devices, despite extensive washing with organic solvent composed of 50% methanol/0.1% TFA in water. This was probably due to high proportions of peptides absorbing to the ultrafilter surfaces as also seen in (140, 141). The effect was not as prominent when higher input material amounts were used.
Detergents are known to prevent peptide binding to (plastic) surfaces (172, 206), which could explain the observed higher intensity of peptide signals for the experimental set up with lower peptide amounts. This is the only positive trait of detergents, but their negative features outweigh this positive effect by far.
All extraction strategies increased methionine oxidation in methionine containing peptides. The proportion of oxidized peptides was higher for the series of experiments with low material amounts, but among those it was the lowest for samples processed with the Zorbax micro-column (Figure 23 B).
With this set of experiments we showed that the Zorbax micro-column had the lowest losses of peptides compared with other sample extraction solutions, and that less peptides underwent methionine oxidation. Furthermore, we showed the importance of downscaling the volumes and contact surfaces for sample processing when the expected starting amounts of target peptides are low.

5.2.1. Chemical tagging for removal of detergent

As outlined above, none of the strategies employed so far removed CHAPS from the IP eluate. To overcome this challenge, we established a chemical tagging strategy for peptides, which exploits the chemical principles of dimethyl labeling (173, 176) for introduction of a glyceraldehyde 3-phosphate

(G3P) tag. The reaction takes place on primary amines (N-terminal and on lysine (Lys) side chains) and results in single and double tagging on the same primary amine group. The G3P molecule contains an aldehyde group for reactions to occur and a phospho-group for TiO_2 isolation (164, 165). This tagging strategy represents an ideal solution for detergent removal, as detergents do not bind to TiO_2 nor influence binding of phospho-group containing peptides to TiO_2 (172).
We aimed to optimize the tagging reaction on the C18 micro-column resin, which would allow separation of peptides from protein contaminants before the reaction, and easy removal of residual reagents after completion of the reaction. Modified peptides were finally eluted from the micro-column and directly subjected to TiO_2 pull down. With this workflow, the number of drying and resuspending cycles, and consequently sample losses, was minimized.

As a first step, the chemical yields of an ordinary dimethyl labeling reaction on the resin of the C18 micro-column (176) were examined. The results confirmed that the reaction takes place dominantly on the N-terminal primary amine at lower pH (2.8), whereas at pH 5.5, the reaction occurred on the N-terminal and partially also on the Lys side chain primary amine group. At pH 8.2, the reaction resulted in nearly 100% conversion of the peptides on both types of primary amines. This experiment showed that it is possible to apply the dimethyl labeling chemical principles to peptides bound on the C18 micro-column.

We then proceeded to G3P labeling. First experiments at pH 3.5, 5.5 or 7.2 gave singly G3P tagged products and products with a G3P tag carrying an additional 14 Da mass increase (Figure 26). The amounts of the latter product were increased when the reaction was conducted at a pH higher than 3.5. Amounts were highest at pH 5.5. At pH 7.2, another product with a mass increase of 28 Da appeared. Doubly tagged peptides were not observed in high amounts in MS analysis after the reaction, due to the phospho-group hindering efficient ionization of peptides (180, 182, 207). The doubly tagged peptides were observed only when G3P tagged peptides were subjected to enzymatic dephosphorylation for phospho-group removal. This also led to overall higher signal intensities of modified dephosphorylated peptides in the MS analysis than those of their phospho-group containing counterparts (Figure 27).
Dephosphorylation of the G3P tagged peptides modified at pH 5.5 resulted in the expected products, which had mass increases of 74 Da, 88 Da or 148 Da (Figure 28 A, D, G). Manual inspection of spectra for the $E6_{9\text{-}19}$ FQDPQERPIKL peptide revealed other side reaction products with mass increases of 44 Da, 58 Da or 118 Da (Figure 28 B, C, H). These unexpected products corresponded to the expected tag masses, but with a 30 Da mass loss. The +88 Da tag had the form of the expected 14 Da+74 Da tag, but could also result from a 2x44 Da tag on the same or separate primary amine groups (Figure 28 D, F, I). Similarly, the 148 Da mass increase corresponded to 2x the 74 Da tag (Figure 28 G, Figure 43 B). Reaction products at pH 5.5 contained tags either only on the N-terminal amine group or on both N-terminal and Lys side chain primary amines in different combinations, such that a distribution of tags among both amine groups was observed. This is exemplified in the spectrum

in Figure 28 H for a +118 Da modified peptide, which was modified on the N-terminal amine group with either +44 Da, +74 Da or +118 Da and on Lys it was unmodified or modified with +44 Da or +74 Da.
As many side products were observed, the optimal pH condition for the highest product signal was examined with several pH reaction buffers. Peptides were isolated by TiO_2 pull down and subjected to enzymatic dephosphorylation. Besides the above described products, additional low abundant modifications of the peptide $E6_{9\text{-}19}$ FQDPQERPIKL were observed: +162 Da (74+88 Da), +192 Da (44+148 Da), +132 Da (44+88=58+74 Da), and +222 Da (3x74 Da).
Quantification of the most abundant reaction products showed that the highest signals were those of singly G3P modified peptides (+154 Da before dephosphorylation, and +74 Da after dephosphorylation), when the reaction took place at pH~2 for the selected peptide $E6_{9\text{-}19}$ FQDPQERPIKL. At higher pH, more side products appeared and thus the signal was distributed among them. Therefore, the most suitable pH for peptides with a Lys in their sequence would be at a pH between 1.5 – 2 (Figure 29 A).
The pH was of less importance for $E7_{11\text{-}19}$ YMLDLQPET conversion, as the product intensities did not change significantly when increasing the pH up to 3.5 (Figure 29 B). However, it was also observed that at pH 1.5 no complete modification was achieved for the $E7_{11\text{-}19}$ peptide (Figure 29 B). Furthermore, several other peptides were left unmodified at pH 1.5 as well, meaning that pH 1.5 was not suitable for their conversion. In the end, pH 2 – 2.2 was chosen as a compromise for the best reaction conditions for the majority of peptides. The determined pH was significantly lower than the pH required for ordinary dimethyl labeling reactions.

The overall experimental pipeline for chemical labeling is schematically represented in Figure 30. Briefly, the IP sample is subjected to acetic treatment for MHC I-peptide complex disassembly. The sample is then transferred onto a Zorbax micro-column, where the majority of proteins is removed, whereas peptide and detergent bind to the Zorbax material. On the column, G3P tagging of primary amines is performed. Subsequently, the remaining reagents are removed from the micro-column and modified peptides are eluted. TiO_2 pull-down is performed for separation of modified peptides from the detergent CHAPS. Afterwards, peptides are subjected to enzymatic dephosphorylation, which improves peptide signal intensity during MS analysis. Finally, samples are desalted and subjected to LC-MS^2/MS^3 analysis.

Manual inspection of the N-terminal phenylalanine ion (F) of the $E6_{9\text{-}19}$ FQDPQERPIKL peptide in either the unmodified (mass 120.08 m/z in Figure 28 E) or the +74 Da form (mass 194.12 m/z in Figure 28 A) with the elemental composition calculator of the Xcalibur program assigned the right elemental composition for both of them, $C_8H_{10}N$ or $C_{11}H_{16}O_2N$, respectively This demonstrated the calculator's correctness and its applicablity to other modifications on the N-terminal phenylalanine ion. The difference between +74 Da and +88 Da tagged N-terminal phenylalanine ions, with the elemental compositions $C_{11}H_{16}O_2N$ and $C_{12}H_{18}O_2N$, respectively, was CH_2, which resulted in a 14 Da mass shift. This mass shift could correspond to the elemental composition of a methyl group, which could be

formed after reaction of formaldehyde with primary amine groups, as also seen in dimethyl labeling reactions.

Molecular formulas assigned to the N-terminal phenylalanine ion with +44 Da or +58 Da tags were $C_{10}H_{14}ON$ or $C_{11}H_{16}ON$, respectively. The comparison of the chemical compositions between +74 Da and +44 Da as well as between the +88 Da and +58 Da tagged N-terminal phenylalanine ions revealed that the difference between both pairs was CH_2O, corresponding to a mass difference of 30 Da and the elemental composition of formaldehyde.

After G3P dephosphorylation, the remaining part of the tag molecule with a mass of 74 Da contains vicinal diols, whose C-C bond (Figure 37 A) can undergo cleavage in oxidative conditions. The vicinal-diol cleavage usually takes place with sodium periodate ($NaIO_4$) or lead tetraacetate ($Pb(OAc)_4$), resulting in the formation of aldehydes and ketones (208). Presumably, the vicinal-diol bond cleavage happened within the G3P tag after dephosphorylation without addition of any of the above mentioned reagents, resulting in the observed mass reduction of 30 Da, which corresponds to formaldehyde (Figure 37 A).

Presumably, a similar reaction as described above also happened as a side reaction of the G3P modification, which resulted in a G3P+14 Da tag and the presence of a +28 Da side product after G3P tagging under increasing pH. Spontaneous G3P dephosphorylation was the prerequisite reaction for vicinal-diol formation, which lead to bond cleavage in the residual G3P molecule and formaldehyde formation. Most likely, this formaldehyde was then competing with unmodified G3P to bind to the primary amine, resulting in the formation of G3P+14 Da, where +14 Da corresponds to a methyl group on the primary amine. Along these lines, the +28 Da product corresponds to dimethyl labeled primary amines. The suggested molecular formulas of the most frequently detected products are depicted in Figure 37 B.

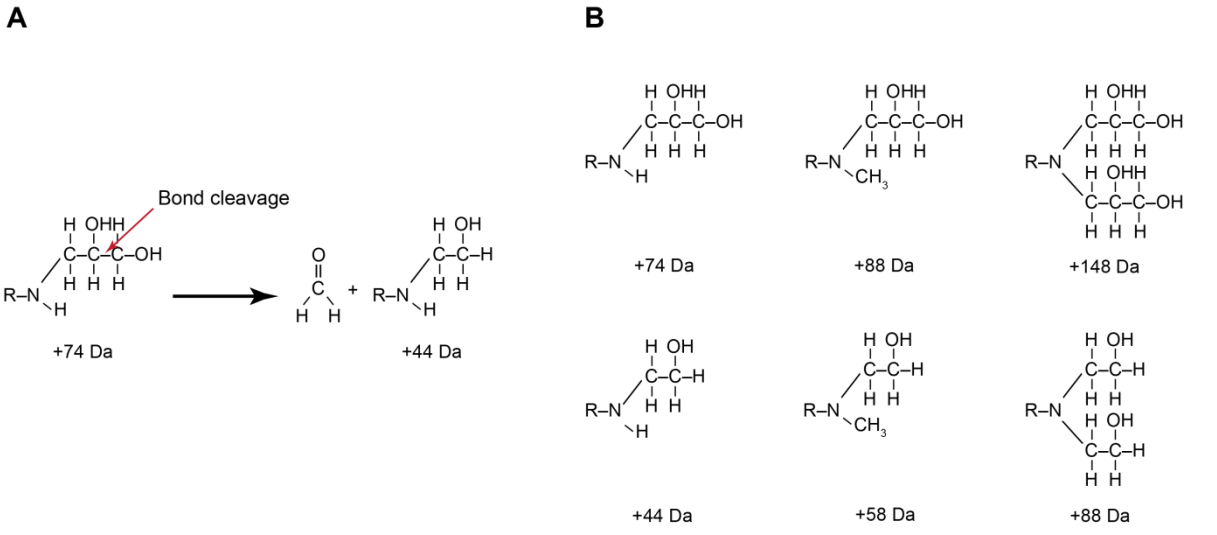
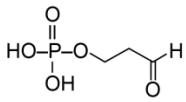


Figure 37. Vicinal-diol bond cleavage and molecular formulas of the most frequent modifications after G3P modification.
A) Vicinal-diol bond cleavage, which occurred to a singly dephosphorylated G3P modified peptide and B) molecular formulas of the most frequent modifications on primary amines after G3P reaction and enzymatic dephosphorylation.

One way to minimize formation of side products in the dephosphorylation step, is to exchange alkaline phosphatase with acid phosphatase, active at a lower pH. Judging from the G3P reaction taking place at pH 5.5, yielding a high amount of +14 Da products (Figure 29), one could try the enzymatic dephosphorylation at pH 4, which is the lowest pH optimum of commercially available enzymes. Presumably, the side reaction of vicinal-diol cleavage would be minimized, but not fully prevented. To achieve less vicinal-diol cleavage, the employed enzyme should be fully active at a pH below 3.5. One could also try chemical dephosphorylation with 70% hydrogen fluoride (HF)-pyridine (207, 209). However, possible new side reactions on the tag would need to be examined first.

In order to overcome all side reactions due to vicinal-diol cleavage, one could investigate the reaction with a different tagging molecule, e.g. 3-(phosphonooxy)-propanal (Figure 38). However, this molecule is not commercially available and would need to be synthesized custom-made.



3-(phosphonooxy)-propanal

Figure 38. An alternative tagging molecule.
Chemical formula of 3-(phosphonooxy)-propanal, which could be used as an alternative tagging molecule.

In addition to being present in low abundance, the majority of target HLA-A2 HPV16 E6 and E7 peptides contain methionine, which can undergo oxidation during sample preparation, resulting in two distinct chromatographic peaks, with the unmodified peptide eluting later. As seen in Figure 23, oxidation took place already at the peptide extraction level and was also observed after chemical modification of peptides.
The oxidation level increased with every additional processing step. This was seen by the relative amount of the G3P singly modified $E7_{11-19}$ YMLDLQPET +154 Da product with or without oxidation (Figure 29 B, +154 MetOx Z and +154 Z) being more in favor of the unoxidized peptide. In contrast, the amount of oxidized peptide after dephosphorylation (Figure 29 B +74 MetOx T and +74 T) was increased, yielding signals with comparable intensities for oxidized and unoxidized peptide.
It would be desirable that either the oxidized or the unoxidized form would dominate to avoid spreading of total peptide amount in two distinctly eluting species. As methionine oxidation is unavoidable, a reaction which allows complete methionine oxidation might be considered. One could examine an additional mild performic acid oxidation reaction to fully convert methionines into methionine sulfoxide as described in (210). In this reference, the reaction was performed on the protein level and resulted in complete conversion of methionine into methionine sulfoxide and of cysteine into cysteic acid without reactions on other amino acids. Besides free cysteines, also those forming disulfide bonds were converted, making this reaction an attractive way to solve two challenges at once, full methionine oxidation and avoidance of reduction and alkylation reaction of cysteines. Ideally, this reaction could be performed on the micro-column before G3P tagging to avoid any other side reactions on the tag.

Due to the above described side reactions of vicinal-diol cleavage and methionine oxidation, the intensity of the most dominant dephosphorylated singly G3P tagged peptides was 20-30% of those of

their unmodified counterparts before the reaction (Figure 29 C). One exception is the peptide $E6_{9-19}$ with 50-60% reaction yield at pH 2. Low reaction yields are the limiting factor of this experimental setup. However, it allows for detergent removal, which critically influences the LC separation and MS detection. Presumably, the suggested improvements of the tagging pipeline could increase the reaction yield to over 50% for most peptides.

5.3. Identification of viral epitopes

5.3.1. Identification of HPV16 epitopes

The target peptide $E7_{11-19}$ YMLDLQPET was not detected with any of the assessed IP sample preparations and detection strategies, although it was reported to be detected in cervical cancer cell lines and patient tumor biopsies with MS^3 analysis before (146, 147). However, endogenous HLA-A2-restricted peptides were detected in every measured sample. To be able to detect HPV16 E6 and E7-derived low abundant epitopes, we employed highly sensitive and specific LC-MS^3 analysis, where fragmentation energies need to be manually optimized for every transition (pair of peptide precursor and its MS^2 fragment) for optimal MS^3 spectrum generation.

The first reason for unsuccessful identification of the $E7_{11-19}$ peptide could be that our examined workflows resulted in high losses of the peptide, so that its amount dropped below the LC-MS^3 detection limit, which was estimated to be 0.3 fmol per analysis on our Qtrap6500 instrument (Figure 40). The estimated amount of $E7_{11-19}$ peptide in (146) was 25 peptide copies per CaSki cell, resulting in 2 – 3 fmol per one IP sample with the assumption of no sample preparation losses. Ten IP samples, even if processed with ultrafiltration with high sample losses, should thus have contained enough $E7_{11-19}$ peptide to detect the oxidized or unoxidized peptide form. This was however not the case. Moreover, 18 chemically tagged CaSki IP samples, containing no detergent which could interfere with column binding, elution and peptide ionization, were analyzed on the highly sensitive HRAM Q-Exactive instrument and also did not result in $E7_{11-19}$ peptide identification. However, after external pulsing on the surface of HLA-A2+ HPV16- BSM cells, the $E7_{11-19}$ peptide was successfully detected with the optimized IP protocol and ultrafiltration extraction, indicating that the experimental pipeline is capable of successfully isolating the peptide (Figure 16).

In contrast to our experimental workflow, after IP isolation of HLA-A2-peptide complexes and acetic dissociation, the methodology described in (146) used epitope extraction by C18 material pipette tips as the only purification strategy. Moreover, MS^3 measurement of the eluate was performed with direct injection into the instrument without prior chromatographic separation. Results were then analyzed with a specially developed probabilistic Poisson-transformation analysis (152), which calculates the possibility of a target peptide being present in the sample (146). The instrument used for measurements was a QTrap4000, which is the first generation of QTraps. This instrument is a low resolution instrument and does not have high mass accuracy. The authors overcame these problems by introducing a second fragmentation step resulting in a MS^3 spectrum, which reduces interferences and increases specificity and sensitivity of the analysis. The instrument that was used for most of the targeted LC-MS^3 analysis in this study is a QTrap6500, which is the third and most recently improved generation of QTrap instruments. According to the producer's specifications, sensitivity of the

QTrap6500 for targeted MS^2 and MS^3 scanning modes is improved by two orders of magnitude compared to the first QTrap4000 model. Furthermore, the linear ion trap (LIT) scanning speed in the MS^3 mode in the QTrap6500 is fast enough to allow measurements in the on-line LC configuration, which is not possible with the Qtrap4000. The advantage of LC-MS^3 is the possibility of separation and concentration of analytes, permitting the identification of less abundant compounds. It can be speculated that the identified $E7_{11-19}$ peptide in (146, 147) was a spectrum of interfering species with similar masses as the target peptide, generating a cumulative MS^3 spectrum with $E7_{11-19}$ characteristic peaks, which originally did not belong to the target peptide, but still generated a positive result after statistical data processing. This theory is supported by the results observed in the experiment where we compared HLA-A2 IP with direct elution of epitopes from the cell surface. The directly eluted sample was so complex that many measured transitions resulted in rich MS^3 spectra (Figure 44). This was observed despite 2 kDa ultrafiltration, which removed many contaminants. These results can be explained by the limited scope of used methodology dealing with highly complex samples and the measuring principles of the QTrap instruments. The instruments isolate precursor ions with m/z within 0.7 Da windows, which is a rather broad range in modern mass spectrometry. All ions corresponding to this criterion will pass to the collision cell for the first fragmentation. The same holds true in the LIT, where ions with m/z within a pre-set window are subjected to the second fragmentation. When a sample is as complex as one from direct elution or an IP sample that is not separated on the LC system prior to MS^3 analysis, the probability is high that other precursors and their fragment ions correspond to the set m/z windows and programed energies. This will result in fragmentation and complex MS^3 spectra generation for multiple ions simultaneously, which cannot be assigned to any particular analyte anymore.

Beside the differences between our and the previously reported (146, 147) methodology, another reason for unsuccessful $E7_{11-19}$ peptide identification could be that the peptide is simply not presented on CaSki and any other analyzed cells that were used during this project. Also in the Prof. Reinherz lab at the Dana-Farber Cancer Institute, Harvard Medical School (Boston, MA, USA), where the studies reported in (146, 147) were conducted, it was later observed that the detection of $E7_{11-19}$ on CaSki cells in later passages of cell culture was not successful (personal communication with Dr. Keskin, Harvard Medical School). As our CaSki cells were a kind gift from Prof. Hoppe-Seyler, DKFZ, Heidelberg, who could not provide us with the passage number of the cells, we assume that they were in late passages and had lost surface presentation of the $E7_{11-19}$ peptide. Only experiments with cells at a low passage number, and using a workflow that removes all interfering contaminants, will reveal if the reason for as yet unsuccessful identification of the target peptide $E7_{11-19}$ was due to cell biology or experimental reasons.

5.3.2. Identification of mCMV epitopes

In collaboration with the group of Prof. Čičin-Šain at the Helmholtz Centre for Infection Research (Braunschweig, Germany), we investigated the presence of the epitope HGIRNASFI in complex with the murine MHC I molecule $H\text{-}2D^b$ on the surface of cells infected with one of three murine

cytomegalovirus (mCMV) clones; mCMV wild type ($mCMV^{WT}$), $mCMV^{M45I\rightarrow A}$ or $mCMV^{M45\ C\text{-}term}$. Monitored H-$2D^b$ endogenous peptides AALENTHLL, FGPVNHEEL and KALINADEL were detected in all measured samples. Based on previous biological results, the peptide HGIRNASFI was expected to be present in small amounts on the wild type virus and in higher amounts in the $mCMV^{M45\ C\text{-}term}$ clone. However, the peptide HGIRNASFI was not expected to be presented on the cells transfected with the $mCMV^{M45I\rightarrow A}$ clone.

The HGIRNASFI peptide was only detected once at the very limit of detection in cells infected with the wild type virus, despite observed T cell activation in the biological system and thus indirect proof of its presence on cells. This shows that T cells require only low copy numbers of epitopes per cell for effective immune responses (manuscript submitted[1]).

We successfully identified the peptide HGIRNASFI on the $mCMV^{M45\ C\text{-}term}$ clone (manuscript submitted[1]). It is important to note that only 10^7 cells were used for IP sample generation, which is the lowest number of cells used for direct LC-MS^2/MS^3 identification of epitopes that we are aware of. Judging from the signal intensity of the target peptide HGIRNASFI on the cells infected with the $mCMV^{M45\ C\text{-}term}$ clone, the input could even be reduced to $5x10^5$ or $1x10^6$ cells per IP sample for reliable epitope identification with the described workflow. Presumably, we could also reduce the input cell number further when exchanging the ultrafiltration step with Zorbax micro-column epitope purification.

5.3.3. Identification of HIV epitopes

In collaboration with the group of Dr. LeGall at the Ragon Institute of MGH, MIT and Harvard (Cambridge, MA, USA), we aimed to identify naturally presented HIV epitopes on 293T HIV-transfected cells. The above mentioned group has successfully established an untargeted MS identification workflow of epitopes from directly eluted cells (manuscript submitted[2]). The scope of this approach is however limited, as untargeted MS^2 measurement is less sensitive than a targeted approach. In the collaboration project, we aimed to identify previously reported epitopes, that were identified through biological assays (185), which the LeGall group was not able to detect with their experimental system.

Following our optimal HLA-A2 IP protocol and Zorbax purification of epitopes, we were able to identify the endogenous HLA-A2-restricted peptides AIVDKVPSV and YLLPAIVHI in all samples. Furthermore, three low abundant HLA-A2-restricted HIV-derived epitopes, EPFRDYVDRFY, FLGKIWPSYK and VLEWRFDSRL, were detected on the surface of HIV transfected cells. The p24-derived epitope EPFRDYVDRFY has concurrent binding motives for HLA-A2 and -B7. In the above mentioned experiment, we were able to show that it was presented on HLA-A2 molecules. Theoretically, it could be presented by HLA-B7 as well, but this needs to be investigated in a separate HLA-B7 IP experiment. The VLEWRFDSRL peptide is the first Nef-derived epitope that was directly identified on the cell surface by LC-MS^2/MS^3 analysis.

[1]Dekhtiarenko I, Fischer S, **Blatnik R**, Holzki JK, Bokner L, Marandu TF, Hoppe S, Lisnić B, May T, Lemmermann NAW, Holtappels R, Reddehase MJ, Riemer AB, Cicin-Sain L. C-terminal epitope localization facilitates antigen processing, direct presentation and T-cell memory inflation.

[2]Rucevic M, Kourjian G, Boucau J, Garcia Bertran W, Berberich MJ, Walker BD and LeGall S. MHC-bound HIV peptides identified from various cell types reveal common nested peptides and novel T cell responses.

6. Conclusions and Future Perspectives

In this thesis, a methodology for isolation, purification and enrichment of cell-surface displayed T cell epitopes for MS detection was optimized. Epitopes were isolated from cells by affinity purification of MHC I-epitope complexes or by mild acetic treatment of life cells for direct elution from the cell surface. Epitope-containing eluates were subjected to various enrichment, purification and fractionation strategies, including ultrafiltration, normal and reverse phase chromatography, isoelectric focusing, and a newly established chemical tagging strategy for epitope isolation by TiO_2 pull down. The methodology was developed first for detection of HLA-A2-restricted HPV16 E6 and E7 epitopes, and then applied to detect mCMV- and HIV-derived epitopes. We were not able to identify the HPV16 $E7_{11-19}$ YMLDLQPET peptide on the surface of HPV16-transformed cells, although it was reported to be detected on cell lines and tumor samples before (146, 147). This can be explained by peptide losses during sample preparation, the presence of detergent influencing LC separation, or by the fact that the peptide was not presented by our HPV16-transformed cells, as they were analyzed at a very high passage number. To be able to identify the target peptide on the surface of CaSki and other HPV16-transforemd cells, various experiments will need to be performed with low passage cells and complete removal of sample contaminants to minimize their influence on the LC separation and MS^2/MS^3 detection.

However, the H-$2D^b$-restricted mCMV epitope HGIRNASFI was detected in high abundance on the surface of $1x10^7$ cells. The cell number needed for these experiments was the lowest reported so far for MS-based epitope detection, and it could be further reduced to $1x10^6$ for still reliable identification. Furthermore, three low abundant HLA-A2-restricted HIV-derived epitopes, EPFRDYVDRFY, FLGKIWPSYK and VLEWRFDSRL, were successfully detected. The epitope VLEWRFDSRL is the first directly MS-identified Nef-derived epitope reported. Taken together, this confirms the efficiency of the developed methodological pipeline and its broad applicability to various MHC I types and virus infected target cells.

Also other researchers in the field have struggled with identification of low abundant (mainly tumor-derived) epitopes, which was due to older, less sensitive, LC-MS instrumentation. In addition, they only measured in an untargeted manner, which is less sensitive than the targeted approach (114, 131-133, 188, 189). Furthermore, difficulties could also originate from insufficient sample purification influencing LC-MS^2 detection. Usually, all studies were performed with bigger sample inputs using on-column immunoaffinity purification strategies for MHC I-peptide complexes. Samples were then subjected to purification with ultrafiltration and reverse phase extraction (138, 198). None of the respective papers reported an influence of residual detergent in the sample on epitope identification. This could be due to either sufficient detergent removal or use of other LC-MS systems, which were not affected by detergents.

In the past years, LC-MS technology and sample preparation methodologies in proteomics advanced in a way that complex samples need less fractionation steps, and LC-MS^2 analyses result in a high

number of identified peptides and proteins. The same holds true for the identification of HLA peptidomes. A recent publication reported the identification of over 22,000 unique HLA peptides in seven cancer cell lines (139), demonstrating that high numbers of HLA peptides can be measured in a single study. However, this untargeted approach is less sensitive than targeted approaches; therefore many low abundant peptides remain undiscovered.

To overcome the challenge of low epitope abundance, just recently genome sequencing data of tumor samples to identify mutated proteins were used for *in silico* predictions and targeted identification of predicted epitopes on the tumor cell surfaces. This workflow resulted in the identification of several novel tumor-mutation-derived neoantigens (90, 91). Thus, these studies represent a milestone in tumor-derived neoantigen detection methodology.

On the other hand, new data-independent measuring methodologies, such as MS^E (211) or the most widely used SWATH-MS (212), have been developed. SWATH analysis is performed with HRAM quadrupole-quadrupole time-of-flight (Qq-TOF) instruments and combines targeted selected reaction monitoring and untargeted shotgun MS^2 analysis. The instrument measures and fragments all ions that enter the mass spectrometer, resulting in highly complex spectra. To process these complex data sets, SWATH-MS results are analyzed by targeted extraction data mining, using spectral libraries. SWATH-MS analysis allows identification and quantification of acquired data at the same time. Furthermore, SWATH-MS measurement permits the detection of low abundant species and the possibility of reanalysis of obtained data at a later time point with improved spectral libraries (212).
One of the main efforts in the epitope mass spectrometry community is to establish HLA peptidome MS spectral databases using HRAM MS instruments from normal, infected, damaged and malignant cells for comparative studies and tumor epitope identification. When these spectral libraries are generated SWATH-MS analysis can be performed and results used for future development of therapies against infectious diseases, autoimmune disorders, and cancer (137). Based on the presented approach, one could prepare a spectral library of all possible HPV E6 and E7 epitopes with synthetic peptides and then measure samples from HPV-transformed cells for identification and quantification of target low abundant peptides.

Another alternative option for IP eluate analysis would be capillary electrophoresis-electrospray ionization mass spectrometry (CESI-MS), which uses capillary electrophoresis for peptide separation instead of reverse phase liquid chromatography. The capillary electrophoreses device is on-line connected to a MS instrument. It maintains low nanoL/min flows, allowing concentration of analytes in smaller volumes, thereby increasing sensitivity. Furthermore, it has no sample carry-over, is not influenced by detergents and can separate peptides from proteins. CESI can be coupled to a triple quadrupole-linear ion trap instrument for targeted analysis, as well as to a HRAM Qq-TOF instrument to perform SWATH analysis (personal communication with Dr. Müller, Sciex GmbH, Darmstadt, Germany and Prof. Imhof, Ludwig-Maximilians-University Munich, Medical Faculty, Department of Molecular Biology, Martinsried, Germany).

In conclusion, the identification of low abundant epitopes will remain a challenging task, as long as the advanced MS technologies described above are not in standard use. However, our targeted MS approach can identify low abundant epitopes, as demonstrated in this work for HIV- and mCMV-derived epitopes. In combination with genome sequencing and identification of mutated proteins in tumors, our methodology can also be applied to the targeted identification of mutated tumor neoantigens, which are in general presented in low amounts on the cell surface and thus difficult to detect. The developed methodology is also very useful in the case of restricted starting sample amounts. In conclusion, the presented methodology can contribute to identification of epitopes for future immunotherapy design and/or therapeutic cancer vaccine development.

7. References

1. Knipe DM (2013) *Fields Virology* (Lippincott Williams & Wilkins, Philadelphia, PA, USA) 6 Ed.
2. Breitbart M & Rohwer F (2005) Here a virus, there a virus, everywhere the same virus? *Trends in Microbiology* 13(6):278-284.
3. zur Hausen H (2002) Papillomaviruses and cancer: from basic studies to clinical application. *Nature reviews. Cancer* 2(5):342-350.
4. Perz JF, Armstrong GL, Farrington LA, Hutin YJ, & Bell BP (2006) The contributions of hepatitis B virus and hepatitis C virus infections to cirrhosis and primary liver cancer worldwide. *Journal of hepatology* 45(4):529-538.
5. Beral V, Newton R, & Sitas F (1999) Human Herpesvirus 8 and Cancer. *Journal of the National Cancer Institute* 91(17):1440-1441.
6. Pattle SB & Farrell PJ (2006) The role of Epstein–Barr virus in cancer. *Expert Opinion on Biological Therapy* 6(11):1193-1205.
7. Ravanfar P, Satyaprakash A, Creed R, & Mendoza N (2009) Existing antiviral vaccines. *Dermatologic Therapy* 22(2):110-128.
8. Heydari DMaS (2011) B cell immunotherapy to resolve a persistent viral infection. *The Journal of Immunology* 186(53.10).
9. Hegde NR, Rao PP, Bayry J, & Kaveri SV (2009) Immunotherapy of viral infections. *Immunotherapy* 1(4):691-711.
10. Woodman CB, Collins SI, & Young LS (2007) The natural history of cervical HPV infection: unresolved issues. *Nature reviews. Cancer* 7(1):11-22.
11. Evander M*, et al.* (1995) Human papillomavirus infection is transient in young women: a population-based cohort study. *The Journal of infectious diseases* 171(4):1026-1030.
12. zur Hausen H (1977) Human Papillomaviruses and Their Possible Role in Squamous Cell Carcinomas. *Current topics in microbiology and immunology*, eds Arber W, Henle W, Hofschneider PH, Humphrey JH, Klein J, Koldovský P, Koprowski H, Maaløe O, Melchers F, Rott R*, et al.* (Springer Berlin Heidelberg), Vol 78, pp 1-30.
13. zur Hausen H (2009) Papillomaviruses in the causation of human cancers - a brief historical account. *Virology* 384(2):260-265.
14. Forman D*, et al.* (2012) Global Burden of Human Papillomavirus and Related Diseases. *Vaccine* 30, Supplement 5:F12-F23.
15. Moscicki A-B*, et al.* (2012) Updating the Natural History of Human Papillomavirus and Anogenital Cancers. *Vaccine* 30, Supplement 5:F24-F33.
16. Gillison ML*, et al.* (2012) Human Papillomavirus and Diseases of the Upper Airway: Head and Neck Cancer and Respiratory Papillomatosis. *Vaccine* 30, Supplement 5:F34-F54.
17. Doorbar J (2006) Molecular biology of human papillomavirus infection and cervical cancer. *Clin Sci (Lond)* 110(5):525-541.
18. Doorbar J*, et al.* (2012) The Biology and Life-Cycle of Human Papillomaviruses. *Vaccine* 30, Supplement 5:F55-F70.
19. Frazer IH (2004) Prevention of cervical cancer through papillomavirus vaccination. *Nature reviews. Immunology* 4(1):46-54.
20. Horvath CA, Boulet GA, Renoux VM, Delvenne PO, & Bogers JP (2010) Mechanisms of cell entry by human papillomaviruses: an overview. *Virology journal* 7:11.
21. Schiller JT, Day PM, & Kines RC (2010) Current understanding of the mechanism of HPV infection. *Gynecologic oncology* 118(1 Suppl):S12-17.
22. Stanley MA, Pett MR, & Coleman N (2007) HPV: from infection to cancer. *Biochemical Society transactions* 35(Pt 6):1456-1460.
23. Lowy DR, Solomon D, Hildesheim A, Schiller JT, & Schiffman M (2008) Human papillomavirus infection and the primary and secondary prevention of cervical cancer. *Cancer* 113(7 Suppl):1980-1993.
24. Joura EA*, et al.* (2015) A 9-valent HPV vaccine against infection and intraepithelial neoplasia in women. *The New England journal of medicine* 372(8):711-723.
25. Frazer IH, Leggatt GR, & Mattarollo SR (2011) Prevention and treatment of papillomavirus-related cancers through immunization. *Annual review of immunology* 29:111-138.
26. De Vincenzo R, Conte C, Ricci C, Scambia G, & Capelli G (2014) Long-term efficacy and safety of human papillomavirus vaccination. *International Journal of Women's Health* 6:999-1010.

27. Hildesheim A*, et al.* (2007) Effect of human papillomavirus 16/18 L1 viruslike particle vaccine among young women with preexisting infection: a randomized trial. *Jama* 298(7):743-753.
28. Stanley M (2006) Immune responses to human papillomavirus. *Vaccine* 24 Suppl 1:S16-22.
29. Khallouf H, Grabowska AK, & Riemer AB (2014) Therapeutic Vaccine Strategies against Human Papillomavirus. *Vaccines* 2(2):422-462.
30. Melief CJ & van der Burg SH (2008) Immunotherapy of established (pre)malignant disease by synthetic long peptide vaccines. *Nature reviews. Cancer* 8(5):351-360.
31. Kenter GG*, et al.* (2009) Vaccination against HPV-16 oncoproteins for vulvar intraepithelial neoplasia. *The New England journal of medicine* 361(19):1838-1847.
32. van Poelgeest MI*, et al.* (2013) HPV16 synthetic long peptide (HPV16-SLP) vaccination therapy of patients with advanced or recurrent HPV16-induced gynecological carcinoma, a phase II trial. *Journal of translational medicine* 11:88.
33. Britt W (2008) Manifestations of human cytomegalovirus infection: proposed mechanisms of acute and chronic disease. *Current topics in microbiology and immunology* 325:417-470.
34. Reddehase MJ, Simon CO, Seckert CK, Lemmermann N, & Grzimek NK (2008) Murine model of cytomegalovirus latency and reactivation. *Current topics in microbiology and immunology* 325:315-331.
35. Reddehase MJ, Podlech J, & Grzimek NK (2002) Mouse models of cytomegalovirus latency: overview. *Journal of clinical virology : the official publication of the Pan American Society for Clinical Virology* 25 Suppl 2:S23-36.
36. Landolfo S, Gariglio M, Gribaudo G, & Lembo D (2003) The human cytomegalovirus. *Pharmacology & therapeutics* 98(3):269-297.
37. Murphy E*, et al.* (2003) Coding potential of laboratory and clinical strains of human cytomegalovirus. *Proceedings of the National Academy of Sciences of the United States of America* 100(25):14976-14981.
38. Dunn W*, et al.* (2003) Functional profiling of a human cytomegalovirus genome. *Proceedings of the National Academy of Sciences of the United States of America* 100(24):14223-14228.
39. Griffiths PD & Grundy JE (1987) Molecular biology and immunology of cytomegalovirus. *The Biochemical journal* 241(2):313-324.
40. Wathen MW & Stinski MF (1982) Temporal patterns of human cytomegalovirus transcription: mapping the viral RNAs synthesized at immediate early, early, and late times after infection. *Journal of virology* 41(2):462-477.
41. Keil GM, Ebeling-Keil A, & Koszinowski UH (1984) Temporal regulation of murine cytomegalovirus transcription and mapping of viral RNA synthesized at immediate early times after infection. *Journal of virology* 50(3):784-795.
42. Messerle M, Buhler B, Keil GM, & Koszinowski UH (1992) Structural organization, expression, and functional characterization of the murine cytomegalovirus immediate-early gene 3. *Journal of virology* 66(1):27-36.
43. Lancini D, Faddy HM, Flower R, & Hogan C (2014) Cytomegalovirus disease in immunocompetent adults. *The Medical journal of Australia* 201(10):578-580.
44. Karrer U*, et al.* (2004) Expansion of Protective CD8(+) T-Cell Responses Driven by Recombinant Cytomegaloviruses. *Journal of virology* 78(5):2255-2264.
45. Snyder C (2011) Buffered memory: a hypothesis for the maintenance of functional, virus-specific CD8+ T cells during cytomegalovirus infection. *Immunol Res* 51(2-3):195-204.
46. Dekhtiarenko I, Jarvis MA, Ruzsics Z, & Cicin-Sain L (2013) The context of gene expression defines the immunodominance hierarchy of cytomegalovirus antigens. *J Immunol* 190(7):3399-3409.
47. WHO (2013 (retrived august 2015)) HIV global health observatory data; http://www.who.int/gho/hiv/.
48. Engelman A & Cherepanov P (2012) The structural biology of HIV-1: mechanistic and therapeutic insights. *Nature reviews. Microbiology* 10(4):279-290.
49. Ho DD & Bieniasz PD (2008) HIV-1 at 25. *Cell* 133(4):561-565.
50. Carla Kuiken TL, Brian Foley, Beatrice Hahn, Preston Marx, Francince McCutchan, Steven Wolinsky, and Bette Korber (2008) *HIV Sequence Compendium 2008* (HIV Sequence Compendium 2008., Los Alamos, New Mexico, USA).
51. Goulder PJR & Watkins DI (2004) HIV and SIV CTL escape: implications for vaccine design. *Nature reviews. Immunology* 4(8):630-640.
52. Missa Sanou JC, Jay Levy, and Janet Yamamoto (2011) Selection of conserved HIV-1 vaccine epitopes based on cross-reactivity to feline immunodeficiency virus. *The Journal of Immunology* 186:53.17.

53. Murphy K (2008) *Janeway's Immunobiology, 7th ed.* (Garland Science, Taylor & Francis Group, LLC, Abingdon, UK).
54. Joyce JA & Fearon DT (2015) T cell exclusion, immune privilege, and the tumor microenvironment. *Science* 348(6230):74-80.
55. Mellman I, Coukos G, & Dranoff G (2011) Cancer immunotherapy comes of age. *Nature* 480(7378):480-489.
56. Yewdell JW & Bennink JR (2001) Cut and trim: generating MHC class I peptide ligands. *Current opinion in immunology* 13(1):13-18.
57. Tanaka K (2009) The proteasome: overview of structure and functions. *Proceedings of the Japan Academy. Series B, Physical and biological sciences* 85(1):12-36.
58. Kloetzel PM (2001) Antigen processing by the proteasome. *Nature reviews. Molecular cell biology* 2(3):179-187.
59. Kloetzel PM & Ossendorp F (2004) Proteasome and peptidase function in MHC-class-I-mediated antigen presentation. *Current opinion in immunology* 16(1):76-81.
60. Stoltze L*, et al.* (2000) Two new proteases in the MHC class I processing pathway. *Nature immunology* 1(5):413-418.
61. Reits E*, et al.* (2004) A major role for TPPII in trimming proteasomal degradation products for MHC class I antigen presentation. *Immunity* 20(4):495-506.
62. Momburg F, Roelse J, Hammerling GJ, & Neefjes JJ (1994) Peptide size selection by the major histocompatibility complex-encoded peptide transporter. *The Journal of experimental medicine* 179(5):1613-1623.
63. Koopmann JO, Post M, Neefjes JJ, Hammerling GJ, & Momburg F (1996) Translocation of long peptides by transporters associated with antigen processing (TAP). *European journal of immunology* 26(8):1720-1728.
64. Wearsch PA & Cresswell P (2008) The quality control of MHC class I peptide loading. *Current opinion in cell biology* 20(6):624-631.
65. Neefjes J, Jongsma ML, Paul P, & Bakke O (2011) Towards a systems understanding of MHC class I and MHC class II antigen presentation. *Nature reviews. Immunology* 11(12):823-836.
66. http://hla.alleles.org/nomenclature/stats.html (July 2015).
67. Falk K, Rotzschke O, & Rammensee HG (1990) Cellular peptide composition governed by major histocompatibility complex class I molecules. *Nature* 348(6298):248-251.
68. Falk K, Rotzschke O, Stevanovic S, Jung G, & Rammensee HG (1991) Allele-specific motifs revealed by sequencing of self-peptides eluted from MHC molecules. *Nature* 351(6324):290-296.
69. Rotzschke O, Falk K, Wallny HJ, Faath S, & Rammensee HG (1990) Characterization of naturally occurring minor histocompatibility peptides including H-4 and H-Y. *Science* 249(4966):283-287.
70. Hunt DF*, et al.* (1992) Characterization of peptides bound to the class I MHC molecule HLA-A2.1 by mass spectrometry. *Science* 255(5049):1261-1263.
71. Calis JJ*, et al.* (2013) Properties of MHC class I presented peptides that enhance immunogenicity. *PLoS computational biology* 9(10):e1003266.
72. Sidney J, Peters B, Frahm N, Brander C, & Sette A (2008) HLA class I supertypes: a revised and updated classification. *BMC immunology* 9:1.
73. Sette A & Sidney J (1999) Nine major HLA class I supertypes account for the vast preponderance of HLA-A and -B polymorphism. *Immunogenetics* 50(3-4):201-212.
74. Engelhard VH (2007) The contributions of mass spectrometry to understanding of immune recognition by T lymphocytes. *International journal of mass spectrometry* 259(1-3):32-39.
75. Purcell AW & Gorman JJ (2004) Immunoproteomics: Mass spectrometry-based methods to study the targets of the immune response. *Molecular & cellular proteomics : MCP* 3(3):193-208.
76. Rausch S, Kruck S, Stenzl A, & Bedke J (2014) IMA901 for metastatic renal cell carcinoma in the context of new approaches to immunotherapy. *Future Oncology* 10(6):937-948.
77. Stevanovic S (2002) Identification of tumour-associated T-cell epitopes for vaccine development. *Nature reviews. Cancer* 2(7):514-520.
78. Singh-Jasuja H, Emmerich NP, & Rammensee HG (2004) The Tubingen approach: identification, selection, and validation of tumor-associated HLA peptides for cancer therapy. *Cancer immunology, immunotherapy : CII* 53(3):187-195.
79. Rammensee H, Bachmann J, Emmerich NP, Bachor OA, & Stevanovic S (1999) SYFPEITHI: database for MHC ligands and peptide motifs. *Immunogenetics* 50(3-4):213-219.

80. Bui HH*, et al.* (2005) Automated generation and evaluation of specific MHC binding predictive tools: ARB matrix applications. *Immunogenetics* 57(5):304-314.
81. Sidney J*, et al.* (2008) Quantitative peptide binding motifs for 19 human and mouse MHC class I molecules derived using positional scanning combinatorial peptide libraries. *Immunome research* 4:2.
82. Peters B & Sette A (2005) Generating quantitative models describing the sequence specificity of biological processes with the stabilized matrix method. *BMC bioinformatics* 6:132.
83. Nielsen M*, et al.* (2007) NetMHCpan, a method for quantitative predictions of peptide binding to any HLA-A and -B locus protein of known sequence. *PloS one* 2(8):e796.
84. Nielsen M*, et al.* (2003) Reliable prediction of T-cell epitopes using neural networks with novel sequence representations. *Protein science : a publication of the Protein Society* 12(5):1007-1017.
85. Mora M, Donati C, Medini D, Covacci A, & Rappuoli R (2006) Microbial genomes and vaccine design: refinements to the classical reverse vaccinology approach. *Current opinion in microbiology* 9(5):532-536.
86. Larsen MV*, et al.* (2010) Identification of CD8+ T cell epitopes in the West Nile virus polyprotein by reverse-immunology using NetCTL. *PloS one* 5(9):e12697.
87. Celis E*, et al.* (1994) Identification of potential CTL epitopes of tumor-associated antigen MAGE-1 for five common HLA-A alleles. *Molecular immunology* 31(18):1423-1430.
88. De Groot AS*, et al.* (2001) From genome to vaccine: in silico predictions, ex vivo verification. *Vaccine* 19(31):4385-4395.
89. Schumacher TN & Schreiber RD (2015) Neoantigens in cancer immunotherapy. *Science* 348(6230):69-74.
90. Yadav M*, et al.* (2014) Predicting immunogenic tumour mutations by combining mass spectrometry and exome sequencing. *Nature* 515(7528):572-576.
91. Gubin MM*, et al.* (2014) Checkpoint blockade cancer immunotherapy targets tumour-specific mutant antigens. *Nature* 515(7528):577-581.
92. Zhang Y, Fonslow BR, Shan B, Baek MC, & Yates JR, 3rd (2013) Protein analysis by shotgun/bottom-up proteomics. *Chemical reviews* 113(4):2343-2394.
93. Han X, Aslanian A, & Yates Iii JR (2008) Mass spectrometry for proteomics. *Current Opinion in Chemical Biology* 12(5):483-490.
94. Marcus K (2012) *Quantitative methods in proteomics*.
95. Huang HZ, Nichols A, & Liu D (2009) Direct Identification and Quantification of Aspartyl Succinimide in an IgG2 mAb by RapiGest Assisted Digestion. *Analytical chemistry* 81(4):1686-1692.
96. Leon IR, Schwammle V, Jensen ON, & Sprenger RR (2013) Quantitative assessment of in-solution digestion efficiency identifies optimal protocols for unbiased protein analysis. *Molecular & cellular proteomics : MCP* 12(10):2992-3005.
97. Di Palma S, Boersema PJ, Heck AJ, & Mohammed S (2011) Zwitterionic hydrophilic interaction liquid chromatography (ZIC-HILIC and ZIC-cHILIC) provide high resolution separation and increase sensitivity in proteome analysis. *Analytical chemistry* 83(9):3440-3447.
98. Nogueira FC*, et al.* (2013) Isotope labeling-based quantitative proteomics of developing seeds of castor oil seed (Ricinus communis L.). *Journal of proteome research* 12(11):5012-5024.
99. Rozenbrand J & van Bennekom WP (2011) Silica-based and organic monolithic capillary columns for LC: Recent trends in proteomics. *Journal of separation science* 34(16-17):1934-1944.
100. Luo Q*, et al.* (2005) Preparation of 20-microm-i.d. silica-based monolithic columns and their performance for proteomics analyses. *Analytical chemistry* 77(15):5028-5035.
101. Molnar I & Horvath C (1976) Reverse-phase chromatography of polar biological substances: separation of catechol compounds by high-performance liquid chromatography. *Clinical chemistry* 22(9):1497-1502.
102. Buszewski B & Noga S (2012) Hydrophilic interaction liquid chromatography (HILIC)--a powerful separation technique. *Analytical and bioanalytical chemistry* 402(1):231-247.
103. Aebersold R & Mann M (2003) Mass spectrometry-based proteomics. *Nature* 422(6928):198-207.
104. Tanaka K*, et al.* (1988) Protein and polymer analyses up to m/z 100 000 by laser ionization time-of-flight mass spectrometry. *Rapid Communications in Mass Spectrometry* 2(8):151-153.
105. Karas M & Hillenkamp F (1988) Laser desorption ionization of proteins with molecular masses exceeding 10,000 daltons. *Analytical chemistry* 60(20):2299-2301.

106. Fenn JB, Mann M, Meng CK, Wong SF, & Whitehouse CM (1989) Electrospray ionization for mass spectrometry of large biomolecules. *Science* 246(4926):64-71.
107. Hillenkamp F, Karas M, Holtkamp D, & Klüsener P (1986) Energy deposition in ultraviolet laser desorption mass spectrometry of biomolecules. *International Journal of Mass Spectrometry and Ion Processes* 69(3):265-276.
108. Dreisewerd K (2003) The desorption process in MALDI. *Chemical reviews* 103(2):395-426.
109. Calligaris D*, et al.* (2010) MALDI in-source decay of high mass protein isoforms: application to alpha- and beta-tubulin variants. *Analytical chemistry* 82(14):6176-6184.
110. Cramer R (2009) MALDI MS. *Proteomics,* Methods in Molecular Biology™, eds Reinders J & Sickmann A (Humana Press), Vol 564, pp 85-103.
111. Kebarle P (2000) A brief overview of the present status of the mechanisms involved in electrospray mass spectrometry. *Journal of Mass Spectrometry* 35(7):804-817.
112. Kebarle P & Verkerk UH (2009) Electrospray: From ions in solution to ions in the gas phase, what we know now. *Mass spectrometry reviews* 28(6):898-917.
113. Wilm M (2011) Principles of electrospray ionization. *Molecular & cellular proteomics : MCP* 10(7):M111 009407.
114. Domon B & Aebersold R (2006) Mass spectrometry and protein analysis. *Science* 312(5771):212-217.
115. Ahmed FE (2008) Utility of mass spectrometry for proteome analysis: part I. Conceptual and experimental approaches. *Expert review of proteomics* 5(6):841-864.
116. Schwartz J, Senko M, & Syka JP (2002) A two-dimensional quadrupole ion trap mass spectrometer. *Journal of the American Society for Mass Spectrometry* 13(6):659-669.
117. Douglas DJ, Frank AJ, & Mao D (2005) Linear ion traps in mass spectrometry. *Mass spectrometry reviews* 24(1):1-29.
118. Makarov A (2000) Electrostatic Axially Harmonic Orbital Trapping: A High-Performance Technique of Mass Analysis. *Analytical chemistry* 72(6):1156-1162.
119. Scigelova M, Hornshaw M, Giannakopulos A, & Makarov A (2011) Fourier transform mass spectrometry. *Molecular & cellular proteomics : MCP* 10(7):M111 009431.
120. Comisarow MB & Marshall AG (1974) Fourier transform ion cyclotron resonance spectroscopy. *Chemical Physics Letters* 25(2):282-283.
121. Marshall AG & Hendrickson CL (2002) Fourier transform ion cyclotron resonance detection: principles and experimental configurations. *International journal of mass spectrometry* 215(1–3):59-75.
122. Hillenkamp F, Karas M, Beavis RC, & Chait BT (1991) Matrix-assisted laser desorption/ionization mass spectrometry of biopolymers. *Analytical chemistry* 63(24):1193A-1203A.
123. Roepstorff P & Fohlman J (1984) Proposal for a common nomenclature for sequence ions in mass spectra of peptides. *Biomedical mass spectrometry* 11(11):601.
124. Biemann K (1988) Contributions of mass spectrometry to peptide and protein structure. *Biomedical & environmental mass spectrometry* 16(1-12):99-111.
125. Eng JK, McCormack AL, & Yates JR (1994) An approach to correlate tandem mass spectral data of peptides with amino acid sequences in a protein database. *Journal of the American Society for Mass Spectrometry* 5(11):976-989.
126. Yates JR (1998) Mass spectrometry and the age of the proteome. *Journal of Mass Spectrometry* 33(1):1-19.
127. Yates JR, Ruse CI, & Nakorchevsky A (2009) Proteomics by Mass Spectrometry: Approaches, Advances, and Applications. *Annual Review of Biomedical Engineering* 11(1):49-79.
128. Cui W, Rohrs HW, & Gross ML (2011) Top-down mass spectrometry: recent developments, applications and perspectives. *The Analyst* 136(19):3854-3864.
129. Pirmoradian M*, et al.* (2013) Rapid and Deep Human Proteome Analysis by Single-dimension Shotgun Proteomics. *Molecular & Cellular Proteomics* 12(11):3330-3338.
130. Imamura H, Wakabayashi M, & Ishihama Y (2012) Analytical strategies for shotgun phosphoproteomics: Status and prospects. *Seminars in Cell & Developmental Biology* 23(8):836-842.
131. Picotti P & Aebersold R (2012) Selected reaction monitoring-based proteomics: workflows, potential, pitfalls and future directions. *Nature methods* 9(6):555-566.
132. Picotti P*, et al.* (2010) High-throughput generation of selected reaction-monitoring assays for proteins and proteomes. *Nature methods* 7(1):43-46.
133. Gallien S, Duriez E, & Domon B (2011) Selected reaction monitoring applied to proteomics. *Journal of mass spectrometry : JMS* 46(3):298-312.

134. Fortin T*, et al.* (2009) Multiple reaction monitoring cubed for protein quantification at the low nanogram/milliliter level in nondepleted human serum. *Analytical chemistry* 81(22):9343-9352.
135. Olsen JV & Mann M (2004) Improved peptide identification in proteomics by two consecutive stages of mass spectrometric fragmentation. *Proceedings of the National Academy of Sciences of the United States of America* 101(37):13417-13422.
136. Ting L, Rad R, Gygi SP, & Haas W (2011) MS3 eliminates ratio distortion in isobaric multiplexed quantitative proteomics. *Nature methods* 8(11):937-940.
137. Caron E*, et al.* (2015) An open-source computational and data resource to analyze digital maps of immunopeptidomes. *eLife* 4.
138. Kowalewski DJ & Stevanovic S (2013) Biochemical large-scale identification of MHC class I ligands. *Methods Mol Biol* 960:145-157.
139. Bassani-Sternberg M, Pletscher-Frankild S, Jensen LJ, & Mann M (2015) Mass spectrometry of human leukocyte antigen class I peptidomes reveals strong effects of protein abundance and turnover on antigen presentation. *Molecular & cellular proteomics : MCP* 14(3):658-673.
140. Lee KJ*, et al.* (2003) Modulation of nonspecific binding in ultrafiltration protein binding studies. *Pharmaceutical research* 20(7):1015-1021.
141. Cunningham R, Wang J, Wellner D, & Li L (2012) Investigation and reduction of sub-microgram peptide loss using molecular weight cut-off fractionation prior to mass spectrometric analysis. *Journal of mass spectrometry : JMS* 47(10):1327-1332.
142. Alvarez-Navarro C, Martin-Esteban A, Barnea E, Admon A, & Lopez de Castro JA (2015) Endoplasmic Reticulum Aminopeptidase 1 (ERAP1) Polymorphism Relevant to Inflammatory Disease Shapes the Peptidome of the Birdshot Chorioretinopathy-Associated HLA-A*29:02 Antigen. *Molecular & cellular proteomics : MCP* 14(7):1770-1780.
143. Tan CT, Croft NP, Dudek NL, Williamson NA, & Purcell AW (2011) Direct quantitation of MHC-bound peptide epitopes by selected reaction monitoring. *Proteomics* 11(11):2336-2340.
144. Stanley MA (2012) Epithelial cell responses to infection with human papillomavirus. *Clinical microbiology reviews* 25(2):215-222.
145. Gonzalez-Galarza FF*, et al.* (2015) Allele frequency net 2015 update: new features for HLA epitopes, KIR and disease and HLA adverse drug reaction associations. *Nucleic acids research* 43(Database issue):D784-788.
146. Riemer AB*, et al.* (2010) A conserved E7-derived cytotoxic T lymphocyte epitope expressed on human papillomavirus 16-transformed HLA-A2+ epithelial cancers. *The Journal of biological chemistry* 285(38):29608-29622.
147. Keskin DB*, et al.* (2011) Direct identification of an HPV-16 tumor antigen from cervical cancer biopsy specimens. *Frontiers in immunology* 2:75.
148. Kessler JH*, et al.* (2003) Competition-based cellular peptide binding assays for 13 prevalent HLA class I alleles using fluorescein-labeled synthetic peptides. *Human immunology* 64(2):245-255.
149. Kessler JH*, et al.* (2004) Competition-based cellular peptide binding assay for HLA class I. *Current protocols in immunology / edited by John E. Coligan ... [et al.]* Chapter 18:Unit 18 12.
150. O'Brien PM & Saveria Campo M (2002) Evasion of host immunity directed by papillomavirus-encoded proteins. *Virus research* 88(1-2):103-117.
151. Grabowska AK & Riemer AB (2012) The invisible enemy - how human papillomaviruses avoid recognition and clearance by the host immune system. *The open virology journal* 6:249-256.
152. Reinhold B, Keskin DB, & Reinherz EL (2010) Molecular detection of targeted major histocompatibility complex I-bound peptides using a probabilistic measure and nanospray MS3 on a hybrid quadrupole-linear ion trap. *Analytical chemistry* 82(21):9090-9099.
153. Pattillo RA*, et al.* (1977) Tumor antigen and human chorionic gonadotropin in CaSki cells: a new epidermoid cervical cancer cell line. *Science* 196(4297):1456-1458.
154. Ku JL, Kim WH, Park HS, Kang SB, & Park JG (1997) Establishment and characterization of 12 uterine cervical-carcinoma cell lines: common sequence variation in the E7 gene of HPV-16-positive cell lines. *International journal of cancer. Journal international du cancer* 72(2):313-320.
155. Cox J & Mann M (2008) MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nature biotechnology* 26(12):1367-1372.
156. Cox J*, et al.* (2009) A practical guide to the MaxQuant computational platform for SILAC-based quantitative proteomics. *Nature protocols* 4(5):698-705.
157. MacLean B*, et al.* (2010) Skyline: an open source document editor for creating and analyzing targeted proteomics experiments. *Bioinformatics* 26(7):966-968.

158. Schilling B*, et al.* (2012) Platform-independent and label-free quantitation of proteomic data using MS1 extracted ion chromatograms in skyline: application to protein acetylation and phosphorylation. *Molecular & cellular proteomics : MCP* 11(5):202-214.
159. Ishihama Y, Rappsilber J, & Mann M (2006) Modular stop and go extraction tips with stacked disks for parallel and multidimensional Peptide fractionation in proteomics. *Journal of proteome research* 5(4):988-994.
160. Gobom J, Nordhoff E, Mirgorodskaya E, Ekman R, & Roepstorff P (1999) Sample purification and preparation technique based on nano-scale reversed-phase columns for the sensitive analysis of complex peptide mixtures by matrix-assisted laser desorption/ionization mass spectrometry. *Journal of mass spectrometry : JMS* 34(2):105-116.
161. Rappsilber J, Mann M, & Ishihama Y (2007) Protocol for micro-purification, enrichment, pre-fractionation and storage of peptides for proteomics using StageTips. *Nature protocols* 2(8):1896-1906.
162. Crowell AM, Wall MJ, & Doucette AA (2013) Maximizing recovery of water-soluble proteins through acetone precipitation. *Analytica chimica acta* 796:48-54.
163. Masuda T, Tomita M, & Ishihama Y (2008) Phase transfer surfactant-aided trypsin digestion for membrane proteome analysis. *Journal of proteome research* 7(2):731-740.
164. Larsen MR, Thingholm TE, Jensen ON, Roepstorff P, & Jorgensen TJ (2005) Highly selective enrichment of phosphorylated peptides from peptide mixtures using titanium dioxide microcolumns. *Molecular & cellular proteomics : MCP* 4(7):873-886.
165. Thingholm TE, Jorgensen TJ, Jensen ON, & Larsen MR (2006) Highly selective enrichment of phosphorylated peptides using titanium dioxide. *Nature protocols* 1(4):1929-1935.
166. Falkenby LG*, et al.* (2014) Integrated solid-phase extraction-capillary liquid chromatography (speLC) interfaced to ESI-MS/MS for fast characterization and quantification of protein and proteomes. *Journal of proteome research* 13(12):6169-6175.
167. Schirle M*, et al.* (2000) Identification of tumor-associated MHC class I ligands by a novel T cell-independent approach. *European journal of immunology* 30(8):2216-2225.
168. Pascolo S*, et al.* (2001) A MAGE-A1 HLA-A A*0201 epitope identified by mass spectrometry. *Cancer research* 61(10):4072-4077.
169. Nuwaysir LM & Stults JT (1993) Electrospray ionization mass spectrometry of phosphopeptides isolated by on-line immobilized metal-ion affinity chromatography. *Journal of the American Society for Mass Spectrometry* 4(8):662-669.
170. Neville DC*, et al.* (1997) Evidence for phosphorylation of serine 753 in CFTR using a novel metal-ion affinity resin and matrix-assisted laser desorption mass spectrometry. *Protein science : a publication of the Protein Society* 6(11):2436-2445.
171. Gruhler A*, et al.* (2005) Quantitative phosphoproteomics applied to the yeast pheromone signaling pathway. *Molecular & cellular proteomics : MCP* 4(3):310-327.
172. Jensen SS & Larsen MR (2007) Evaluation of the impact of some experimental procedures on different phosphopeptide enrichment techniques. *Rapid communications in mass spectrometry : RCM* 21(22):3635-3645.
173. Hsu JL, Huang SY, Chow NH, & Chen SH (2003) Stable-isotope dimethyl labeling for quantitative proteomics. *Analytical chemistry* 75(24):6843-6852.
174. Thompson A*, et al.* (2003) Tandem mass tags: a novel quantification strategy for comparative analysis of complex protein mixtures by MS/MS. *Analytical chemistry* 75(8):1895-1904.
175. Ross PL*, et al.* (2004) Multiplexed protein quantitation in Saccharomyces cerevisiae using amine-reactive isobaric tagging reagents. *Molecular & cellular proteomics : MCP* 3(12):1154-1169.
176. Boersema PJ, Raijmakers R, Lemeer S, Mohammed S, & Heck AJ (2009) Multiplex peptide stable isotope dimethyl labeling for quantitative proteomics. *Nature protocols* 4(4):484-494.
177. Lundblad RL (1984) *Chemical Reagents for Protein Modification* (CRC Press, Boca Raton, FL).
178. Hermanson GT (1996) *Bioconjugate Techniques* (Academic Press, San Diego, CA).
179. Baxter EW & Reitz AB (2004) Reductive Aminations of Carbonyl Compounds with Borohydride and Borane Reducing Agents. *Organic Reactions*, (John Wiley & Sons, Inc.).
180. Liao PC, Leykam J, Andrews PC, Gage DA, & Allison J (1994) An approach to locate phosphorylation sites in a phosphoprotein: mass mapping by combining specific enzymatic degradation with matrix-assisted laser desorption/ionization mass spectrometry. *Analytical biochemistry* 219(1):9-20.
181. Larsen MR, Sorensen GL, Fey SJ, Larsen PM, & Roepstorff P (2001) Phospho-proteomics: evaluation of the use of enzymatic de-phosphorylation and differential mass spectrometric

peptide mass mapping for site specific phosphorylation assignment in proteins separated by gel electrophoresis. *Proteomics* 1(2):223-238.
182. Ishihama Y*, et al.* (2007) Enhancement of the efficiency of phosphoproteomic identification by removing phosphates after phosphopeptide enrichment. *Journal of proteome research* 6(3):1139-1144.
183. Millioni R*, et al.* (2013) Pros and cons of peptide isolectric focusing in shotgun proteomics. *Journal of chromatography. A* 1293:1-9.
184. Righetti PG, Sebastiano R, & Citterio A (2013) Capillary electrophoresis and isoelectric focusing in peptide and protein analysis. *Proteomics* 13(2):325-340.
185. http://www.hiv.lanl.gov/content/immunology
186. Hu X*, et al.* (1996) Enhancement of cytolytic T lymphocyte precursor frequency in melanoma patients following immunization with the MAGE-1 peptide loaded antigen presenting cell-based vaccine. *Cancer research* 56(11):2479-2483.
187. Henderson RA*, et al.* (1992) HLA-A2.1-associated peptides from a mutant cell line: a second pathway of antigen presentation. *Science* 255(5049):1264-1266.
188. Anderson L & Hunter CL (2006) Quantitative mass spectrometric multiple reaction monitoring assays for major plasma proteins. *Molecular & cellular proteomics : MCP* 5(4):573-588.
189. Kim YJ, Zaidi-Ainouch Z, Gallien S, & Domon B (2012) Mass spectrometry-based detection and quantification of plasma glycoproteins using selective reaction monitoring. *Nature protocols* 7(5):859-871.
190. Munoz N, Castellsague X, de Gonzalez AB, & Gissmann L (2006) Chapter 1: HPV in the etiology of human cancer. *Vaccine* 24 Suppl 3:S3/1-10.
191. de Sanjose S*, et al.* (2010) Human papillomavirus genotype attribution in invasive cervical cancer: a retrospective cross-sectional worldwide study. *The Lancet Oncology* 11(11):1048-1056.
192. Schiffman M, Castle PE, Jeronimo J, Rodriguez AC, & Wacholder S (2007) Human papillomavirus and cervical cancer. *Lancet* 370(9590):890-907.
193. Scarinci IC*, et al.* (2010) Cervical cancer prevention: new tools and old barriers. *Cancer* 116(11):2531-2542.
194. Bruni L, Barrionuevo-Rosas L, Albero G, Aldea M, Serrano B, Valencia S, Brotons M, Mena M, Cosano R, Muñoz J, Bosch FX, de Sanjosé S, Castellsagué X. ICO Information Centre on HPV and Cancer (HPV Information Centre). Human Papillomavirus and Related Diseases in the World. Summary Report 2015-04-08. http://www.hpvcentre.net/statistics/reports/XWX.pdf.
195. Gillison ML*, et al.* (2000) Evidence for a causal association between human papillomavirus and a subset of head and neck cancers. *Journal of the National Cancer Institute* 92(9):709-720.
196. Michaud DS*, et al.* (2014) High-risk HPV types and head and neck cancer. *International journal of cancer. Journal international du cancer* 135(7):1653-1661.
197. Kanodia S, Da Silva DM, & Kast WM (2008) Recent advances in strategies for immunotherapy of human papillomavirus-induced lesions. *International journal of cancer. Journal international du cancer* 122(2):247-259.
198. Carralot JP, Lemmel C, Stevanovic S, & Pascolo S (2008) Mass spectrometric identification of an HLA-A*0201 epitope from Plasmodium falciparum MSP-1. *International immunology* 20(11):1451-1456.
199. Yeung YG, Nieves E, Angeletti RH, & Stanley ER (2008) Removal of detergents from protein digests for mass spectrometry analysis. *Analytical biochemistry* 382(2):135-137.
200. Yeung YG & Stanley ER (2010) Rapid detergent removal from peptide samples with ethyl acetate for mass spectrometry analysis. *Current protocols in protein science / editorial board, John E. Coligan ... [et al.]* Chapter 16:Unit 16 12.
201. McNulty DE & Annan RS (2008) Hydrophilic interaction chromatography reduces the complexity of the phosphoproteome and improves global phosphopeptide isolation and detection. *Molecular & cellular proteomics : MCP* 7(5):971-980.
202. Cool DR & DeBrosse D (2003) Extraction of oxytocin and arginine-vasopressin from serum and plasma for radioimmunoassay and surface-enhanced laser desorption-ionization time-of-flight mass spectrometry. *Journal of chromatography. B, Analytical technologies in the biomedical and life sciences* 792(2):375-380.
203. Bonifacino JS, Dell'Angelica EC, & Springer TA (2001) Immunoprecipitation. *Current Protocols in Immunology*, (John Wiley & Sons, Inc.).
204. Hodge K, Have ST, Hutton L, & Lamond AI (2013) Cleaning up the masses: exclusion lists to reduce contamination with HPLC-MS/MS. *Journal of proteomics* 88:92-103.

205. Trindade F*, et al.* (2015) Salivary peptidomic as a tool to disclose new potential antimicrobial peptides. *Journal of proteomics* 115:49-57.
206. Stewart, II, Thomson T, & Figeys D (2001) 18O labeling: a tool for proteomics. *Rapid communications in mass spectrometry : RCM* 15(24):2456-2465.
207. Woo EM, Fenyo D, Kwok BH, Funabiki H, & Chait BT (2008) Efficient identification of phosphorylation by mass spectrometric phosphopeptide fingerprinting. *Analytical chemistry* 80(7):2419-2425.
208. Wade L (2010) *Organic chemistry* (Pearson Prentice Hall).
209. Kuyama H, Toda C, Watanabe M, Tanaka K, & Nishimura O (2003) An efficient chemical method for dephosphorylation of phosphopeptides. *Rapid communications in mass spectrometry : RCM* 17(13):1493-1496.
210. Pesavento JJ, Garcia BA, Streeky JA, Kelleher NL, & Mizzen CA (2007) Mild performic acid oxidation enhances chromatographic and top down mass spectrometric analyses of histones. *Molecular & cellular proteomics : MCP* 6(9):1510-1526.
211. Silva JC, Gorenstein MV, Li GZ, Vissers JP, & Geromanos SJ (2006) Absolute quantification of proteins by LCMSE: a virtue of parallel MS acquisition. *Molecular & cellular proteomics : MCP* 5(1):144-156.
212. Gillet LC*, et al.* (2012) Targeted data extraction of the MS/MS spectra generated by data-independent acquisition: a new concept for consistent and accurate proteome analysis. *Molecular & cellular proteomics : MCP* 11(6):O111 016717.
213. Armbruster DA & Pry T (2008) Limit of Blank, Limit of Detection and Limit of Quantitation. *The Clinical Biochemist Reviews* 29(Suppl 1):S49-S52.
214. Mani DR, Abbatiello SE, & Carr SA (2012) Statistical characterization of multiple-reaction monitoring mass spectrometry (MRM-MS) assays for quantitative proteomics. *BMC bioinformatics* 13 Suppl 16:S9.

8. Appendix

8.1. Comparison of elution buffers

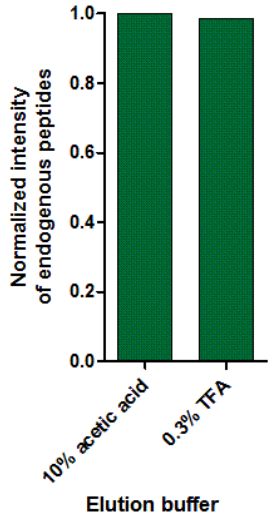


Figure 39. Comparison of the IP elution buffers 10% acetic acid and 0.3% TFA.
Two CaSki HLA-A2 IP samples from $6x10^7$ cells each were treated either with 70 µL 10% acetic acid in water or 70 µL 0.3% TFA in water containing murine MHC I restricted peptides FGPVNHEEL and SSIEFARL. Samples were ultrafiltered, desalted with the OMIX tips and analyzed with the nanoAcqutiy-QTrap5500 platform. Total intensities of the HLA-A2 endogenous peptides AIVDKVPSV and YLLPAIVHI were normalized to the total intensities of the added synthetic murine MHC I peptides. The results from one experimental replicate are shown.

8.2. Determining the limit of specific peptide detection

To determine the limit of detection (LOD) for the $E7_{11-19}$ YMLDLQPET peptide, a synthetic peptide was added to an HLA-A3 IP eluate from HPV16 negative EA cells to ensure similar sample composition as in the HLA-A2 IP from HPV16-transformed cells.

An HLA-A3 IP was performed from $1x10^8$ EA cells and subjected to acidic elution. The IP eluate was ultrafiltered, desalted with OMIX tips and divided in equal aliquots. The synthetic peptide was spiked into the samples to be injected at amounts corresponding to 0.01 fmol, 0.1 fmol, 1 fmol, 10 fmol or 100 fmol. Every sample was analyzed in three technical replicates. The following three MS^3 transitions were measured to confirm peptide identity; b_5 (555.26/636.30), b_6 (555.26/764.36) and b_8 (555.26/990.46). MS^3 spectra were reliably detected for all transitions with characteristic fragmentation patterns when 1 fmol or higher concentrations of the peptide were used for analysis (Figure 40 A – F). No signal was detected when 0.1 fmol or less of the peptide was analyzed.

The limit of detection (LOD) was determined with a calibration curve (213, 214) (Figure 40 G), based on the formula LOD=LOB+10xSD, where the LOB is the signal of a blank sample and SD its standard deviation, to result in a conservative LOD value (213). The LOD for the $E7_{11-19}$ YMLDLQPET was estimated to be 0.3 fmol.

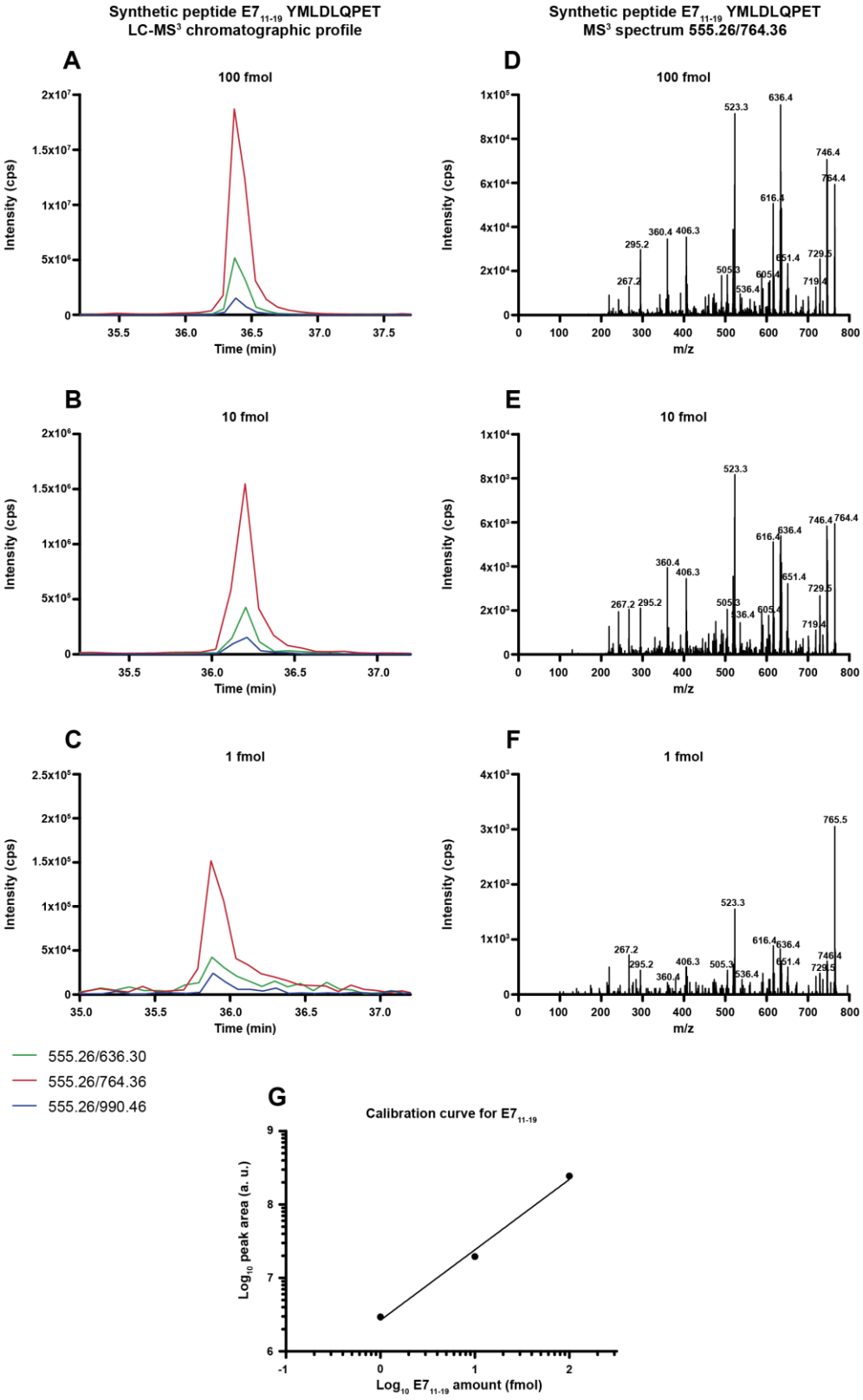


Figure 40. Determination of the limit of detection (LOD) for the peptide E7$_{11-19}$ YMLDLQPET.
An EA HLA-A3 IP eluate was subjected to ultrafiltration and desalting with OMIX tips prior to the addition of the synthetic peptide E7$_{11-19}$ YMLDLQPET. The peptide was added to result in 0.01 fmol, 0.1 fmol, 1 fmol, 10 fmol and 100 fmol per LC-MS^3 analysis. LC-MS^3 analysis was performed with the nanoAcquity UPLC-QTrap6500 system. A) – C) Extracted ion chromatograms for transitions b_5 (555.26/636.30), b_6 (555.26/764.36) and b_8 (555.26/990.46) of synthetic peptide E7$_{11-19}$, representative of one technical replicate; D) – F) MS^3 spectrum for transition b_6 (555.26/764.36) for the E7 $_{11-19}$ peptide, representative of one technical replicate; G) Calibration curve from three technical replicates for the peptide E7$_{11-19}$ at 1 fmol, 10 fmol and 100 fmol per analysis.

8.3. Acetone – ethyl acetate precipitation

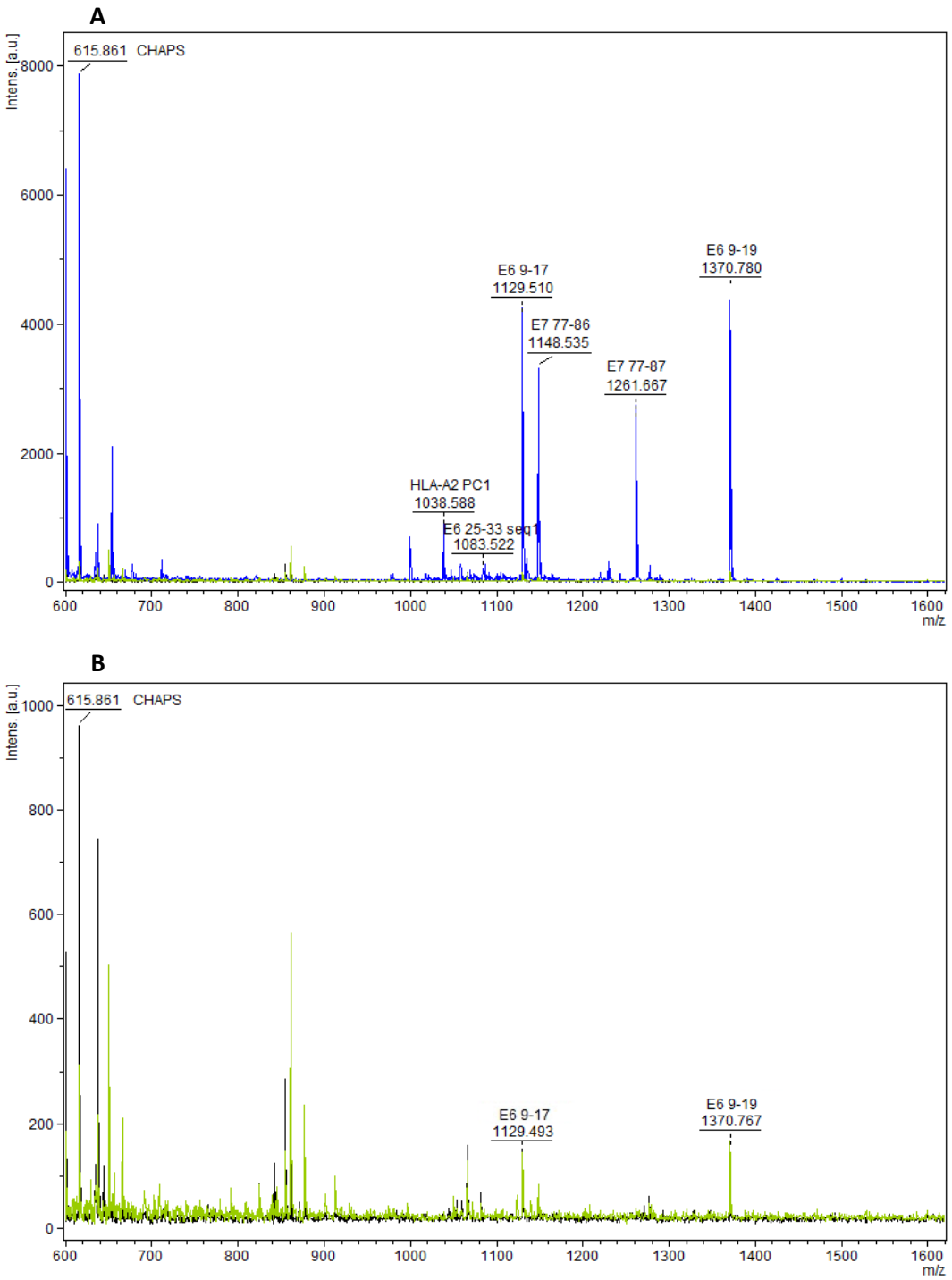


Figure 41. MALDI TOF MS peptide profile of a BSM IP sample with added peptides, after fractionation with acetone – ethyl acetate precipitation.
The IP eluate was subjected to acetone precipitation overnight and subsequently centrifuged to separate the protein pellet from peptides in the supernatant. The supernatant was subsequently vacuum dried to reduce its volume and precipitated with ethyl acetate to separate detergent and peptides. The starting unprocessed sample, acetone precipitation pellet, organic ethyl acetate top layer and aqueous ethyl acetate bottom layer were subjected to vacuum drying, resuspending in 50% ACN/0.1% TFA and MALDI TOF analysis with the UltrafleXtreme instrument. A) All fractions; B) all fractions except the unprocessed IP sample for better presentation of less intense signals. The unprocessed IP sample is represented by the blue line, the acetone pellet by the black line, the aqueous ethyl acetate bottom layer by the red lineand the organic ethyl acetate top layer by the green line. Both ethyl acetone fractions are directly overlaid, therefore the red line is not visible. Peptide peaks and the CHAPS peak are marked with their masses and names. Peptides and CHAPS were detected in all fractions.

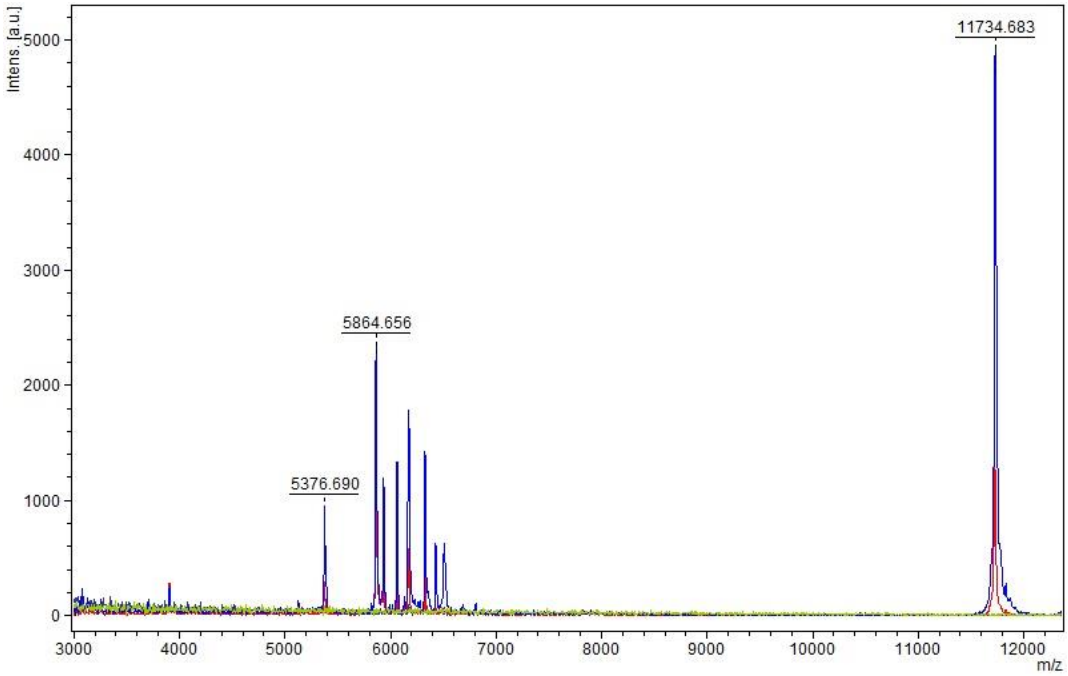


Figure 42. MALDI TOF MS protein profile of a BSM IP sample with added peptides after fractionation with acetone – ethyl acetate precipitation.
The IP eluate was subjected to acetone precipitation overnight and subsequently centrifugation to separate protein pellet from peptides in supernatant. The supernatant was subsequently vacuum dried to reduce its volume and precipitated with the ethyl acetate to separate detergent and peptides. The starting unprocessed sample, acetone precipitation pellet, organic ethyl acetate top layer and aqueous ethyl acetate bottom layer were subjected to vacuum drying, resuspending in 50% ACN/0.1% TFA and MALDI TOF analysis with the UltrafleXtreme instrument. The unprocessed IP sample is represented by the blue line, the acetone pellet by black line, the aqueous ethyl acetate bottom layer by the red line and the organic ethyl acetate top layer by the green line. Proteins were detected in the unprocessed IP sample and the aqueous ethyl acetate bottom layer fraction.

8.4. G3P tagging product

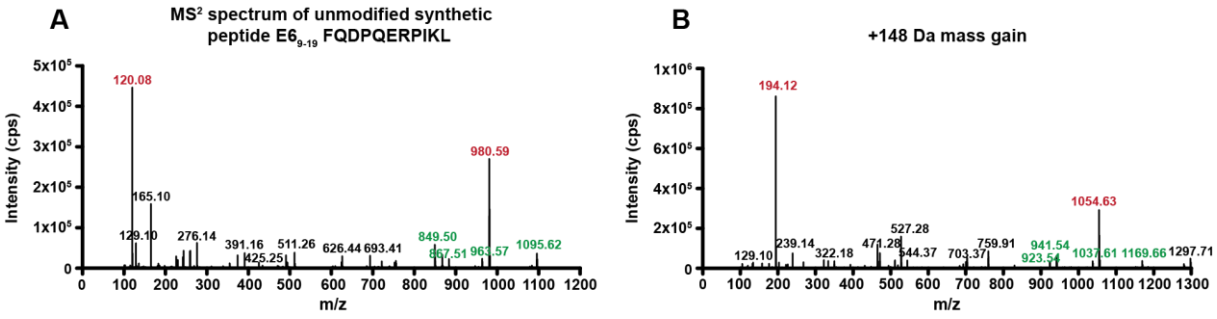


Figure 43. MS² spectra of the peptide E6₉₋₁₉ FQDPQERPIKL and its +148 Da product after G3P modification and dephosphorylation.
G3P modified peptides were treated with alkaline phosphatase before they were analyzed with the speLC-Q-Exactive instrument. Characteristic peaks for the N-terminal phenylalanine (102.08 m/z) and y_8 (980.59 m/z) in (A) together with their +74 Da modified counterparts (194.12 and 1054.63 m/z) in (B) are marked in red. The 74 Da mass shifts of other characteristic peaks at 849.50, 867.51, 963.57, 1095.62 in (A) to 923.54, 941.54, 1037.61 and 1169.66 in (B), are marked in green.

8.5. MS^3 spectra in a complex sample

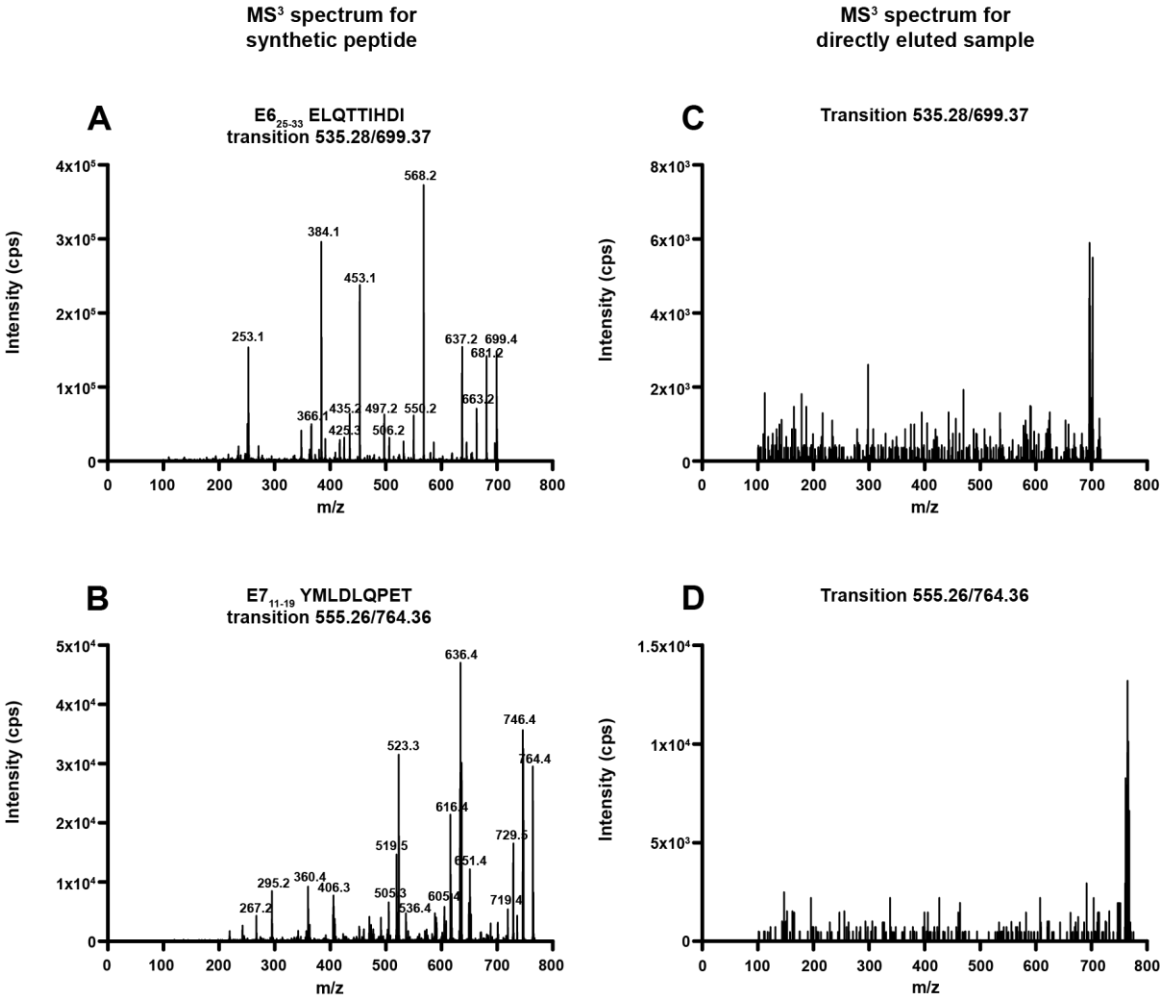


Figure 44. Detection of cumulative MS^3 spectra in a complex sample.
1x10^8 CaSki cells were subjected to mild acetic treatment for direct elution of epitopes. The eluate was subjected to 2 kDa ultrafiltration and purification with a Seppak cartridge prior to LC-MS^3 analysis with the nanoAcquity UPLC-QTrap6500 system. A) MS^3 spectrum for transition y_6 (535.28/699.37) for the synthetic peptide E6$_{25-33}$ ELQTTIHDI; B) MS^3 spectrum for transition b_6 (555.26/764.36) for the synthetic peptide E7$_{11-19}$ YMLDLQPET; C) MS^3 spectrum for transition 535.28/699.37 detected in the directly eluted sample and D) MS^3 spectrum for transition 555.26/764.36 detected in the directly eluted sample. MS^3 spectra in the directly eluted sample were generated within 1 minute of the expected elution time for the synthetic peptide. MS^3 spectra of the directly eluted sample contain numerous unspecific signals.

8.6. Isoelectric focusing

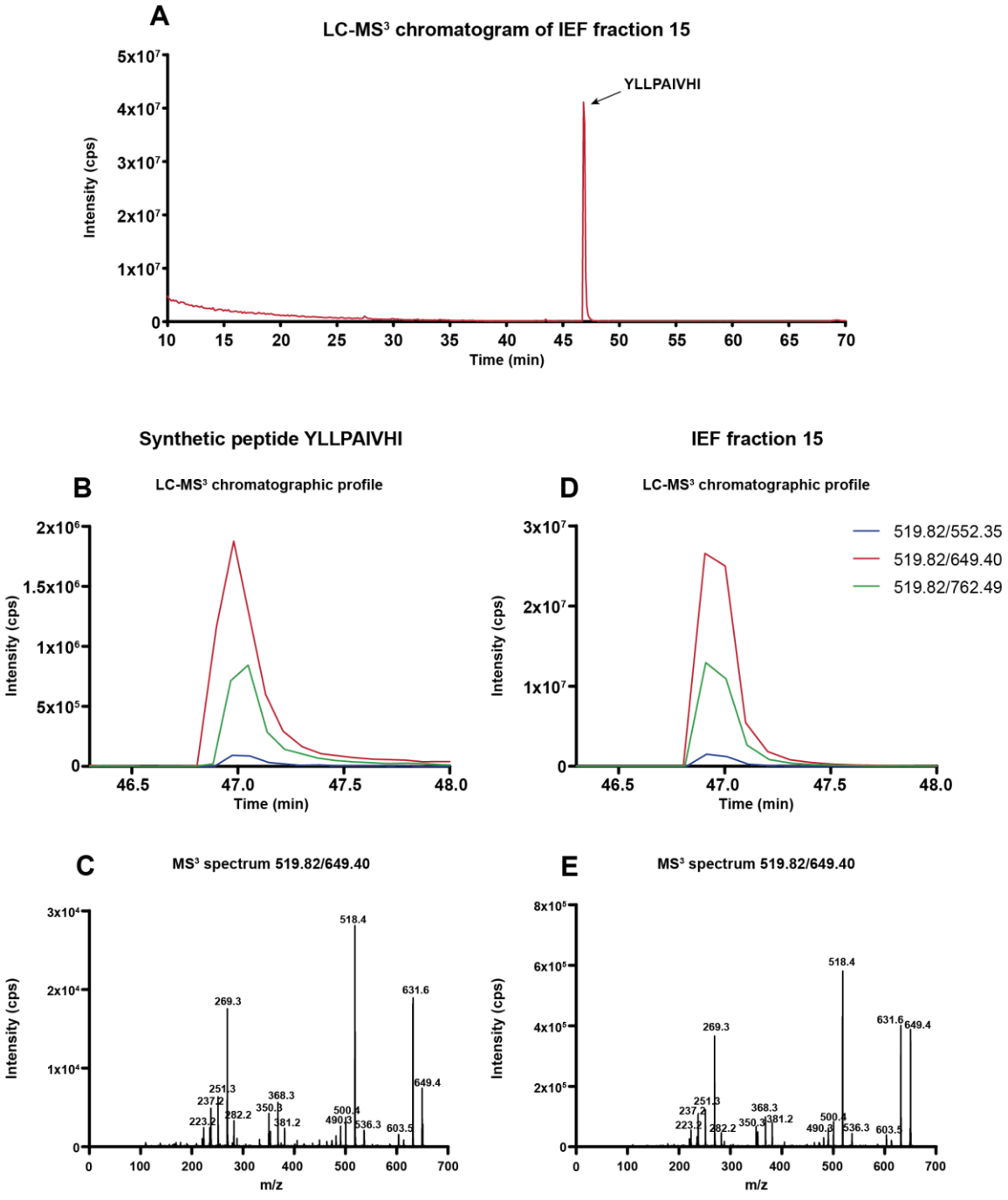


Figure 45. Isoelectric focusing (IEF) of directly eluted epitopes resulted in focusing of the endogenous peptide YLLPAIVHI in one fraction.
$1x10^8$ SNU17 cells were subjected to mild acetic treatment for direct elution of epitopes. The eluate was subjected to purification with a Seppak cartridge, vacuum dried and then fractionated by IEF. Fractions were desalted with a Zorbax micro-column prior to LC-MS^3 analysis with the nanoAcquity UPLC-QTrap6500 system. A) LC-MS^3 chromatographic profile for the endogenous peptide YLLPAIVHI detected in high abundance in fraction 15; B) Extracted ion chromatogram for transitions y_5 (519.82/552.35), y_6 (519.82/649.40) and y_7 (519.82/762.49) of the synthetic peptide YLLPAIVHI; C) MS^3 spectrum for the transition y_6 (519.82/649.40) for the synthetic peptide YLLPAIVHI; D) extracted ion chromatogram for the transitions 519.82/552.35, 519.82/649.40 and 519.82/762.49 detected in IEF fraction 15; E) MS^3 spectrum for the transition 519.82/649.40 detected in IEF fraction 15.

8.7. Detection of proteins in IP samples after fractionation

Lists of identified peptides and proteins after IP sample fractionations with a Seppak cartridge or a Zorbax micro-column are presented in digital format in Microsoft Office Excel files on the compact disc provided at the back of the thesis.