UNIVERSITY OF LIVERPOOL

FLUCLOXACILLIN PROTEIN BINDING LEADS TO THE PRESENTATION OF NOVEL MHC PEPTIDES

This thesis is submitted in accordance with the requirements of the University of Liverpool

for the degree of Doctor of Philosophy by

James C Waddington

August 2019

DECLARATION

I declare that the work presented in this thesis is all my own work and has not been submitted for any other degree.

_____

James C Waddington, MBiolSci (Microbiology).

“Covalent binding without function is confession without sin”

Professor B. Kevin Park

ACKNOWLEDGEMENTS

First and foremost, thank you to everybody who has been involved in my education, work and personal life up until now. Teachers, lecturers, supervisors, friends, family, colleagues; you have all played a part in the generation of this thesis, and for that I am eternally grateful.

Since starting my PhD, I have been in awe of the supervision I have received from Kevin, Xiaoli and Dean. From the overarching philosophical concepts provided by Kevin, to the technical experimental aspects taught by Xiaoli, to the making sense of it all from Dean. You have all been instrumental in the data generated for this thesis, and I can only hope that I have managed to take in a small fraction of what you have taught me. A special mention must be given to Xiaoli here. You have been so much more than a supervisor to me, and I could not have been luckier… 非常感谢你. My only regret is that I didn't make you sign your name sooner..!

One of my biggest achievements must be an invitation to 2 pm coffee within the first year of my PhD. Roz, you used to terrify me, but now I know just how much of a softie you are. Thank you for the countless hours in the lab teaching me how the mass specs work. Thank you for answering all my questions out of hours. Thank you for not shouting at me (too much) when I broke the instruments. I know that I have learnt a lot from you, and I'm so glad that now I can fix the machines when you break them. Jane, you too used to terrify me (even more than Roz). Aside from your help in the lab, and the countless gels you have run, thank you for being able to talk about something other than science. Breakfast club, trips to Lidl and defrosting the freezers will all be memories that I hold very fondly. I can only apologise for going to breakfast when you were busy. To my proteomics sidekick Bhav, thank you so much for being you. You kept my spirits high and convinced me that Jane and Roz weren't as scary as I thought. You are one of the most genuine people I have met, and I am so grateful to have joined the group while you were there. Gina, being from Yorkshire, we were always destined to be friends. You brought sanity to my time in the lab, just as much as you brought insanity to the group. You're one in a million, and I miss working with you every day, even if you do love the Lancashire rose. To Ana, thank you for making me laugh and for all your help over the years. Keep an eye on them, and make sure you keep throwing things away when they are not looking. Finally, thanks to the extended proteomics family, you know who you are!

Of course, being part of Dean's research group, I was initially placed in the immunology student office. I apologise for never cloning T cells to a novel compound. Thank you to all of those who have been around during my time here. You have all contributed to my project in one way or another. Thanks to Lee for teaching me all I needed to know in the lab, thanks to Monday for all your expertise and advice, and thanks to John for your help throughout my project. It's really daunting joining such a large research group, but I was never made to feel unwelcome. A special mention must go to Arun. I honestly don't know how I would have done this PhD without you. Both as a mentor and a friend, you have been there for me when I needed you, and I cannot thank you enough. You always become close to the people in your own year group, and I am so happy that Joel started his PhD at the same time as me. From our chats in the office, to living together, to the late-night whiskeys on the balcony, I have enjoyed every minute of working with you. Thank you for being such a good friend to me! Paul, you have certainly kept me entertained. When you're not talking about abacavir or an ELISpot, your positive outlook on life has been a breath of fresh air. Thanks for all your help, and for sharing my frustrations over the years. Of course, I couldn't not give a special mention

to the Andy's. Gibson, without sounding too soppy, you know how much you mean to me, and you know how much you are missed. I'm not sure I should put too much on paper here, and I don't think I need to explain why. Sully, your work ethic was truly inspiring, and I am so glad that I had the opportunity to work with you at the beginning of my time in the group. Finally, to the newer members of the team; Serat-E and Kareena (did we settle for Kerat or Serena?) - You have both been like my kids over the last year or so, and I cannot wait to see you fly the nest. Look after each other and start writing your thesis! Just make sure you don't break anything else.

Agnès, thank you for trying to make me go to the gym and be healthy; I'm sorry that I stopped going and stuck to wine instead. Harriet, you are *literally* the coolest person I know, please come back to the UK soon. Amy, Sammie and Lucy, thank you for everything, especially for joining the postgrad committee with me (I'm still sorry). Sophie, thank you for all your help over the last year or so, you have made my thesis much more colourful. I also want to thank my collaborators Patricia and Tony. My time in your lab set the foundation of my work, and I cannot thank you enough. And finally, to Neil French, for always finding the funding from somewhere to send me to all the workshops and conferences that I have been so lucky to attend.

Away from the University, the support I have received from my family and friends has been immeasurable. Mum and Dad, you have always taught me to put my education first, and I am only the person that I am today because of you. Alex, without you, I honestly don't think I would have done this PhD. Mainly because I refuse to let you be the only Dr in the house, but also because you've been there for me. Since we all started university together back in 2011, Beth, Sally, Conor and Lydia, I cannot put into words how much you have all supported me. I know that without you I would not be here writing this now. And of course, to James, thank you for being so patient over the last 12 months. I'm not sure I will know what to do with myself now that I have finished writing this thesis?! This support extends to my family, friends and loved ones from across the globe; thank you!

Finally, I would like to end with a very special mention to two mentors from before my PhD. Heather Allison, you sparked my initial interest for research, and the two years spent in your lab made me want to pursue a career in science. Thank you for everything. And to Pam Kennedy; you will never know how much your support helped me in getting here. Without you taking the extra time explaining mass spectrometry to me, I likely wouldn't have got through my A-level chemistry, yet alone a PhD. It's nothing short of ironic that I ended up writing a thesis on the use of mass spectrometry in the characterization of drug-protein adduct formation. I hope that you have finally made it to that Caribbean island.

TABLE OF CONTENTS

ABBREVIATIONS

5'OH-flucloxacillin	5'-hydroxymethy flucloxacillin
ABC	ATP-binding cassette
ABTS	2,2'-azino-bis(3-ethylbenzothiazoline-6-sulphonic acid
AC	Amoxicillin-clavulanate
ACN	Acetonitrile
ADR	Adverse drug reaction
AHS	Abacavir hypersensitivity syndrome
AKT	Protein B kinase
ALP	Alkaline phosphatase
ALT	Alanine aminotransaminase
AMS	Accelerated mass spectrometry
ANN	Artificial neural network
APAP	Acetaminophen
APS	Ammonium persulfate
AUC	Area under the curve
β2M	Beta-2-microglobulin
BiP	Binding protein
BP	Benzyl penicillin
BSA	Bovine serum albumin
BSEP	Bile salt export pump
CCL	CC-chemokine ligand
CD	Cluster of differentiation
CE	Collision energy
CLA	Cutaneous lymphocyte antigen
CMFDA	5-chloromethylfluorescein diacetate
CMV	Cytomegalovirus
CNX	Calnexin
CO_3	Carbonate
CRT	Calreticulin
CTLA4	Cytotoxic T-lymphocyte-associated protein 4
CV	Column volume
CXP	Collision cell exit potential
CYP	Cytochrome
DAG	Diacylglycerol
DAMPs	Damage associated molecular patterns
DC	Dendritic cell
DDA	Data dependent acquisition
DHR	Drug hypersensitivity reaction
DIA	Data independent acquisition
DILI	Drug induced liver injury
DILIN	DILI network
DNCB	Dinitrochlorobenzene
DNP	Dinitrophenol

DP	Declustering potential
DRESS	Drug reaction with eosinophilia and systemic symptoms
DTT	Dithiothreitol
EBV	Epstein-Barr virus
ELISA	Enzyme linked immunosorbent assay
EP	Entrance potential
ER	Endoplasmic reticulum
ERAP	Endoplasmic reticulum aminopeptidases
F	Phenylalanine
FA	Formic acid
Fab	Antigen binding fragments
FBS	Foetal bovine serum
Fc	Fragment crystallizable
FDA	Food and Drug Administration
FDR	False discovery rate
GM-CSF	Granulocyte-macrophage colony-stimulating factor
GSH	Glutathione
GSTP	Glutathione S-transferase pi
GWAS	Genome wide association studies
HBV	Hepatitis B virus
HC	Healthy volunteer
HCV	Hepatitis C virus
HEV	Hepatitis E virus
HHV	Human herpes virus
HLA	Human leukocyte antigen
HPLC	High performance liquid chromotography
H-Pt	Hypersensitive patient
HSA	Human serum albumin
HSP	Heat shock protein
I	Isoleucine
IAA	Iodoacetamide
IDA	Information dependent acquisition
iDILI	Idiosyncratic DILI
IFN	Interferon
IP3	Inositol triphosphase
ITAM	Immunoreceptor tyrosine-based activation motif
iTRAQ	Isobaric tag for relative and absolute quantification
KLH	Keyhole limpet hemocyanin
L	Leucine
LFTs	Liver function tests
LMP	Latent membrane protein
LSC	Liquid scintillation counting
LSEC	Liver sinusoidal endothelial
m/z	Mass-to-charge ratio
MAPK	Mitogen-activated protein kinase

MDR	Multidrug resistance
MGF	Mascot generic file
MHC	Major histocompatibility complex
MIP	Maximum intensity projection
MPE	Maculopapular exanthema
MRM	Multiple reaction monitoring
MRP	Multidrug resistance protein
MS	Mass spectrometry
MSA	Mouse serum albumin
MW	Molecular weight
NAPQI	*N*-acetyl-*p*-benzo-quinone imine
NB	Non binder
NCBI	National Center for Biotechnology Information
NET	Neutrophil extracellular trap
NK	Natural killer
NRTI	Nucleoside reverse transcriptase inhibitor
OATP	Organic anion transporting polypeptide
OR	Odds ratio
OVA	Ovalbumin
P#	Position/Pool #
PAMPs	Pathogen associated molecular patterns
PAS	Protein A sepharose
PBP	Penicillin binding protein
PBS	Phosphate buffer saline
PD-1	Programmed death-1
PE	Peptide elution
PI	Pharmacological interaction
PI3K	Phosphoinositide 3-kinase
pMHC	Peptide MHC complex
PO_4	Phosphate
PRRs	Pattern recognition receptors
PSM	Peptide-to-spectrum match
Pt	Patient
PTM	Post translational modification
PXR	Pregnane X receptor
PΩ	C-terminal amino acid residue
Q	Quadrupole
QTOF	Quadrupole time-of-flight
RM	Reactive metabolite
ROCK/MLCK	Rho/Myosin light chain kinase
RT	Room temperature
SB	Strong binder
SCX	Strong cation exchange
SDS-PAGE	SDS-polyacrylamide gel electrophoresis
SHP	Small heterodimer partner
SJS	Stevens-Johnson syndrome

SMX	Sulfamethoxazole
SMX-NO	SMX-nitroso
SNPs	Single nucleotide polymorphisms
TAP	Transporter associated with antigen processing
TCR	T cell receptor
TEMED	N,N,N′,N′-Tetramethylethylenediamine
TEN	Toxic epidermal necrolysis
TFA	Trifluoroacetic acid
T_{fh}	Follicular helper T cells
T_h	Helper T cell
T_h0	Naïve T_h cells
TIC	Total ion count
TIM-3	T cell immunoglobulin and mucin-domain containing-3
TLRs	Toll like receptors
TMT	Tandem mass tagging
TNF-α	Tissue necrosis factor α
TNP	Trinitrophenol
Tpn	Tapasin
TPP	Trans-Proteomics Pipeline
Treg	Regulatory T cells
TST	Tris/saline/Tween
V	Valine
W	Tryptophan
WB	Weak binder
WHO	World Health Organisation
Y	Tyrosine
ZAP-70	Zeta-chain-associated protein kinase 70

PUBLICATIONS

The following manuscripts in preparation are associated with this thesis:

Waddington, J.C., Penman, S., Ali, S., Jenkins, R.E., Chadwick, A., Sharma, P., Naisbitt, D.J., Meng, X. & Park, B.K. Bile canalicular localization of flucloxacillin disrupts membrane transporter activity, providing new insights for the development of drug induced liver injury

Waddington, J.C., Meng, X., Illing, P., Tailor A, Adair, K, Jenkins, R., Hamlett, J., Farrell, J., Purcell A.W., Naisbitt, D.J. & Park, B.K. Antigen profiling reveals natural presentation of haptenated flucloxacillin HLA-B*57:01 binding peptides

The following manuscripts are associated with this thesis:

Ali, S-E.*, **Waddington, J.C.***, Park, B.K. & Meng, X. (2019). Definition of the Chemical and Immunological Signals Involved in Drug-induced Liver Injury. Chemical Research in Toxicology. *Joint first author

Sullivan, A., Watkinson, J., **Waddington, J.,** Park, B. K. & Naisbitt, D. J. (2018). Implications of HLA-allele associations for the study of type IV drug hypersensitivity reactions. Expert Opinion on Drug Metabolism & Toxicology

Waddington, J. C., Meng, X., Naisbitt, D. J. & Park, B. K. (2017). Immune drug-induced liver disease and drugs. Current Opinions in Toxicology.

Meng, X., Ariza, A., **Waddington, J.** & Naisbitt, D. J. (2016). Immunological Mechanisms of Drug Hypersensitivity. Current Pharmaceutical Design.

Tailor*, A., **Waddington*, J. C.**, Meng, X. & Park, B. K. (2016). Mass Spectrometric and Functional Aspects of Drug–Protein Conjugation, Chemical Research in Toxicology. **Joint first author*

Additional contributions were made to the following publications:

Tailor, A., **Waddington, J.C.,** Hamlett, J., Maggs, J., Kafu, L., Farrell, J., Dear, G.J., Whitaker, P., Naisbitt, D.J., Park, K. & Meng X. (2019). Definition of haptens derived from sulfamethoxazole: in vitro and in vivo. Chemical Research in Toxicology

Alzahrani, A., Ogese, M., Meng, X., **Waddington, J. C.**, Tailor, A., Farrell, J., Maggs, J. L., Betts, C., Park, B. K. & Naisbitt, D. J. (2017). Dapsone and Nitroso Dapsone Activation of Naïve T-Cells from Healthy Donors. Chemical Research in Toxicology.

Meng, X., Earnshaw, C. J., Tailor, A., Jenkins, R. E., **Waddington, J. C.**, Whitaker, P., French, N. S., Naisbitt, D. J. & Park, B. K. (2016). Amoxicillin and Clavulanate Forms Chemically and Immunologically Distinct Multiple Haptenic Structures in Patients, Chemical Research in Toxicology.

CONFERENCE & WORKSHOP ATTENDANCE

Invited Speaker

December 2018- 49th LBMSDG meeting, Francis Crick Institute, London, UK.
'Characterization of the involvement of HLA peptides in drug hypersensitivity reactions'

Oral Abstracts

September 2019- HIPP Satellite Meeting, HUPO World Congress, Perth, Australia.
'β-lactam protein binding leads to the presentation of novel MHC peptides, generating a potential immunogen for drug-hypersensitive reactions'

December 2018- DMG Early Career Meeting, Imperial College, London, UK.
'Definition and localization of flucloxacillin-modified proteins that may have involvement in drug-induced liver injury'

July 2018- BSPR Annual Scientific Meeting, University of Bradford, UK.
*'Characterisation of the HLA-B*57:01 immunopeptidome in flucloxacillin-related drug-induced liver injury'*

Poster Abstracts

September 2019- HUPO World Congress, Perth, Australia.
'Cell membrane transporters facilitate the accumulation of drug-protein antigens, generating a localized pool of MHC epitopes associated with drug-induced liver injury'

July 2019- BSPR Annual Scientific Meeting, Southampton, UK.
*'Detection of flucloxacillin modified peptides naturally presented by HLA- B*57:01, a risk allele for flucloxacillin-induced liver injury'*

December 2018- Pharmacology 2018 (British Pharmacology Society), London, UK.
'Beta-lactam protein binding leads to the presentation of novel MHC peptides, generating a potential immunogen for drug-hypersensitivity reactions'

September 2018- HUPO World Congress, Orlando, Florida, USA.
'Characterisation of flucloxacillin-modified proteins leading to the presentation of flucloxacillin-modified MHC peptides, and their importance in iDILI'

April 2018- 8th Drug Hypersensitivity Meeting (EAACI), Amsterdam, The Netherlands.
*'Characterization of the HLA-B*57:01 immunopeptidome in the presence of flucloxacillin'*

January 2018- 16th EACCI Winter School, Saas-Fee, Switzerland.
'Characterization of drug-modified peptides that act as T- cell antigens'

September 2017- HUPO World Congress, Dublin, Ireland.
'Characterisation of naturally processed peptides that act as T- cell antigens in patients with drug hypersensitivity

Workshops

May 2019- Introduction to R, a statistical software package, University of Liverpool.

September 2018- Proteomics Informatics Course, University of Central Florida, Orlando, Florida, USA.
Run by the Seattle Proteome Centre from the Institute of Systems Biology. A 5-day intensive course in the use of the open suite software the Trans-Proteomic Pipeline (TPP).

ABSTRACT

Flucloxacillin is a β-lactam antibiotic associated with a high incidence of idiosyncratic drug-induced liver reactions. Although expression of HLA-B*57:01 increases susceptibility, little is known of the pathological mechanisms involved in the induction of the clinical phenotype. Off-target protein modification is suspected to drive the reaction, either through non-immune mediated pathways, through the modification of peptides that are presented to T cells by the risk allele, or both. In this thesis, the characterization of proteins haptenated by flucloxacillin was performed using proteomic techniques. As protein haptenation, followed by antigen processing and presentation of the drug-derived antigenic determinants may drive the adverse event, the immunopeptidome of HLA-B*57:01 in the presence and absence of flucloxacillin was determined.

In order to detect flucloxacillin modified proteins, an antibody specific for flucloxacillin was generated. Characterization of the antibody to determine cross reactivity and selectivity was determined. Anti-flucloxacillin antibody was used to identify flucloxacillin modified proteins in a number of immortalised cell lines and primary human hepatocytes. Western blot, immunocytochemistry and mass spectrometry were used for the detection, localisation and characterization, respectively, of drug modified proteins. C1R-B*57:01, B-lymphoblast cells transfected with HLA-B*57:01, were incubated with flucloxacillin for 48h. HLA peptide complexes were subsequently eluted and processed for mass spectrometric analysis. Finally, bioinformatic pipelines were generated to assist in the characterization of flucloxacillin-modified MHC peptides to allow a more high-throughput approach to immunopeptidomic data analysis.

The generation of a high titre flucloxacillin specific antibody was successful, with no cross reactivity with proteins. Isoxazole ring containing β-lactams did cross react, indicating this was the site of recognition. Intracellular protein modification was identified in all the cell lines examined, including primary human hepatocytes. Using the liver cell line HepaRG, localization within the bile canaliculi was observed. The function of the major hepatocellular efflux transporters MRP2 and P-gp was increased in the presence of flucloxacillin in a time-dependant manner. Modification of master regulators of MAP Kinase signalling molecules was detected. Flucloxacillin also modified HLA-B*57:01 protein directly, which could lead to neo-antigens being presented. In depth analysis of the global repertoire of peptides was interrogated. Flucloxacillin was found to alter the C-terminal amino acid on the majority of peptides, where an increase in phenylalanine and a decrease in tryptophan was observed. Peptides unique to flucloxacillin treatment were theoretically weaker binders to HLA-B*57:01, indicating flucloxacillin may assist in their stabilization. Of the peptides eluted from flucloxacillin treated C1R-B*57:01 cells, 7 were fully annotated to show flucloxacillin-lysine covalent binding, with other partially annotated peptides indicating modifications.

In this thesis a wide range of off-target protein modification has been determined, including proteins involved in regulatory signalling pathways. The localization of flucloxacillin was identified to occur in the site of clinical disease during flucloxacillin-induced liver injury. It was also demonstrated that drug-modified peptides are presented by HLA-B*57:01 and that global repertoires are altered by flucloxacillin. Further investigation into the immunogenicity of haptenated proteins in the onset of iDILI is required to determine the role of these peptides in drug hypersensitivity.

CHAPTER 1 - GENERAL INTRODUCTION

Adverse or unwarranted responses to drugs are a major impediment to the drug development process and an economic burden to healthcare services across the globe. The financial impact of adverse drug reactions (ADRs) varies, however in the UK a study of 18,820 patients revealed a cost of £450 million to the NHS, contributing to 6.5% of hospital admissions (Pirmohamed *et al.*, 2004). While often resolved with cessation of the treatment, in rare cases, ADRs can lead to severe conditions including gastrointestinal bleeding, renal failure, liver damage, and in some cases, mortality. Of course, strict regulations imposed on pharmaceutical companies have significantly reduced the incidence of ADRs. However, when new pharmaceuticals reach the market, exposure to large populations results in the presentation of side effects not observed during initial screening. These more severe reactions to drugs more commonly involve immune activation. Delayed reactions to drugs, including after cessation of treatment, are indicative of an adaptive immune response. In these cases, the adverse event cannot be predicted due to the idiosyncratic nature of the onset of disease (Uetrecht and Naisbitt, 2013). Indeed, as technology and our understanding of genetics has evolved it has become more apparent that certain genes can be risk factors for the onset of disease (Alfirevic and Pirmohamed, 2010).

Here, a thorough literature review was performed to provide a framework and overview of the areas that will feed into the experimental chapters, with a view to further our understanding of the immune triggers in idiosyncratic ADRs. The general introduction has therefore covered the following;

- ADRs; including drug induced liver injury (DILI).
- Cellular mediators of ADRs and mechanisms of immune activation.
- Proteomic methods to detect drug protein adducts.
- B-lactam antibiotics and flucloxacillin induced DILI.
- Thesis aims.

1.1 ADVERSE DRUG REACTIONS

ADRs are defined by the World Health Organization (WHO) as "a response to a medicine which is noxious and unintended, and which occurs at doses normally used in man for the prophylaxis, diagnosis or therapy of disease, or for modification of physiological function" (WHO, 1972). In more recent years this definition has been met with disagreement though the use of 'noxious'. Typically, noxious relates to something that is 'injurious, harmful or hurtful' as defined by the Oxford English Dictionary, therefore, can minor reactions be defined as an ADR under the WHO's 1972 definition? There needs to be some separation between mild and more severe ADRs otherwise the current surveillance system employed to monitor ADRs would not operate as intended. In response, Laurence's definition aimed to reduce the categorization of minor ADRs; "a harmful or significantly unpleasant effect caused by a drug at doses intended for therapeutic effect (or prophylaxis or diagnosis) which warrants reduction of dose or withdrawal of the drug and/or foretells hazard from future administration" (Laurence and Carpenter, 1998). While this was generally accepted the exclusion of non-drug triggers such as contaminants, inactive compounds and error made it less favourable by many in the scientific community. It wasn't until 2000, almost 30 years after the first definition, that Edwards and Aronson defined an ADR as "an appreciably harmful or unpleasant reaction, resulting from an intervention related to the use of a medicinal product, which predicts hazard from future administration and warrants prevention or specific treatment, or alteration of the dosage regimen, or withdrawal of the product" (Edwards and Aronson, 2000). This definition is more generally accepted as it more broadly describes ADRs while still allowing for a wider range of symptoms to be included.

1.1.1 DRUG HYPERSENSITIVITY REACTIONS

1.1.1.1 DEFINITION

Drug hypersensitivity reactions (DHRs) are a subset of ADR based on their immunological involvement. Typically, DHRs are idiosyncratic in nature, meaning that they cannot be

predicted by the pharmacology of the drug or their dose due to unknown off target toxicity. Individual susceptibility to a DHR is associated with a whole host of contributions including genetics, diet, immune status, multiple medications, treatment duration, age, gender and the chemical structure of the drug (Pichler, 2003). Patients with HIV and mononucleosis have a marked increase of an ADR to suflonamides and ampicillin respectively. Less specific factors such as undergoing open heart surgery or the influenza vaccine can generally increase the risk of an ADR (Ju and Reilly, 2012).

1.1.1.2 CLASSIFICATIONS

Immune mediated DHRs are categorized into four sub-groups depending on the onset of symptoms and the cellular components and mechanisms responsible for their action (Figure 1.1). Although defined in 1963 by Gell and Coombs, it is still generally widely accepted and termed types I, II, III and IV hypersensitivity (Gell and Coombs, 1963). The definitions for these reactions are explored below.

1.1.1.2.1 TYPE I

Type I reactions are immediate reactions that occur within minutes to hours of contact with the drug. Several symptoms are observed including urticaria, asthma, conjunctivitis, rhinitis and in severe cases cardiorespiratory (anaphylactic) shock. With type I hypersensitivity respiratory symptoms are generally rare as few drugs are administered though inhalation. Activation occurs when IgE bind to high affinity receptors on mast cells and basophils and are cross linked by a drug allergen. Once bridged IgE molecules trigger the degranulation of basophils leading to the release of preformed mediators such as histamine. Subsequently newly synthesized mediators such a thromboxanes, prostaglandins and leukotrienes are released (Atkinson and Kaliner, 1992; DeJarnatt and Grant, 1992; Descotes and Choquet-Kastylevsky, 2001).

1.1.1.2.2 TYPE II

Type II reactions are mediated by cytotoxic IgG and IgM antibodies and can cause cell damage via two mechanisms. Immunoglobulin covered target cells can be recognized by immune cells through the conserved (Fc) region of the antibodies by macrophages, neutrophils and/or eosinophils. The second mechanism of action is through the activation of the classical complement cascade resulting in cell lysis. Typically type II reactions are semi-delayed and will using manifest within a few days. Agranulocytosis, thrombocytopenia and immune-allergenic hemolytic anemia are all examples of type II hypersensitivity reactions (DeShazo, 1997; Descotes and Choquet-Kastylevsky, 2001).

1.1.1.2.3 TYPE III

Type III hypersensitivity is defined as semi-delayed immune-complex reactions mediated through IgG/IgM and soluble antigens. Soluble antigens react with immunoglobulins in the tissue space and form micro-precipitates in the endothelial lining of blood vessels leading to the activation of complement cascades thus tissue injury. The resulting inflammation recruits immune cells such as macrophages, neutrophils and platelets to the deposition site subsequently causing further damage. Typically deposits are found in the lungs, kidneys, the skin and joints (Descotes and Choquet-Kastylevsky, 2001).

1.1.1.2.4 TYPE IV

Type IV hypersensitivity is defined as delayed type with a clinical onset of disease up to 45 days after drug administration and is primarily T cell mediated. Importantly Type IV drug hypersensitivity is not mediated by immunoglobulins, even if present in the patient sera. Drug antigens presented on major histocompatibility complexes (MHC) by antigen presenting cells activate T lymphocytes resulting in the release of cytokines and cytolytic molecules leading to tissue damage. Skin eruptions are typically observed for type IV reactions and can manifest as relatively mild urticaria to fatal Stevens-Johnson Syndrome (SJS) and toxic epidermal necrolysis (TEN).

Type IV hypersensitivity is mediated by several cell types which are associated with different lymphokine secretory molecules. The term delayed type (IV) hypersensitivity was originally coined to describe T cell reactions to tuberculin, however it has since been used as an umbrella term for different cellular involvement and cytokine releases which manifest as differential clinical disorders. With this in mind Pichler set out to further define type IV hypersensitivity using classifications based on the effector cells and lymphokine secretions involved (Table 1.1) (Pichler, 2003).

Table 1.1. Type IV hypersensitivity reaction classifications. Recommendations set by Pichler (2003) have allowed for further categorization of type IV DHRs (a-d) based on the cellular components and the secretory molecules involved.

	Effector Cells	**Lymphokine secretion**
Type IV a	Monocytes	IFNγ
Type IV b	Eosinophils	IL-4, IL-5
Type IV c	Keratinocytes	Perforin, Granzyme B, Fas Ligand
Type IV d	Neutrophils	IL-8

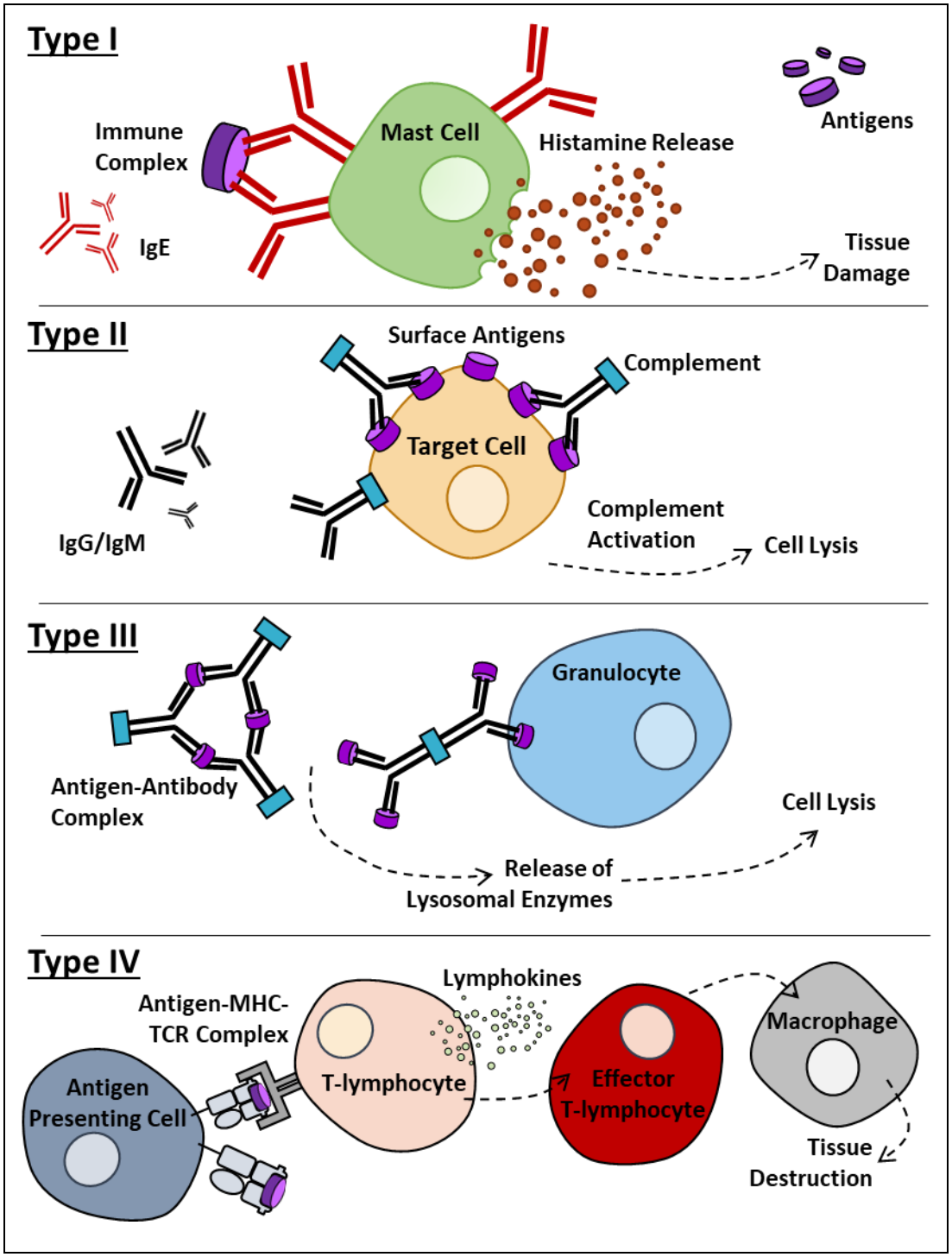


Figure 1.1. Type B (immune mediated) ADRs are further categorized into types I-IV. Type I occurs immediately following sensitization and is primarily mediated by IgE. Cross linking of antigens on mast cells results in the release of histamine leading to tissue damage. Type II hypersensitivity are semi-delayed cytotoxic reactions mediated through the activation of complement by IgG and IgM. Type III reactions are triggered through immune complex reactions formed by soluble antigens. Type IV, delayed type, hypersensitivity is mediated by lymphocytes. T cell receptors are activated by pMHC complexes resulting in the release of lymphokines. Effector T lymphocytes and recruited immune cells such as macrophages are responsible for the clinical response through the release of cytokines and cytolytic molecules.

1.1.1.3 CLINICAL MANIFESTATION OF DHRS

DHRs can affect several organs in the body, however skin is the one that is more commonly and noticeably described. The skin is the largest organ of the human body and is comprised on tiny blood vessels throughout its structure. It is therefore unsurprising that T cell mediated reactions occur within the tissue. Cutaneous reactions can be wide ranging, from mild rashes to severe reactions such as SJS and TEN. In this section we will briefly cover the different manifestations of DHRs associated with the skin, including the potential mechanisms for clinically relevant examples.

1.1.1.3.1 MACULOPAPULAR EXANTHEMA

Maculopapular exanthema (MPE) is the most common delayed type reaction, attributed to 95% of all cutaneous eruptions, characterized by the infiltration of CD4+ T cells. The clinical symptoms of MPE begin with erythematous macular or papular eruptions on the trunk of the body spreading to the extremities in a symmetrical fashion. Symptoms generally occur within 7-14 days post exposure to the drug with re-exposure resulting in a much faster response. In severe cases MPE can progress to more severe reactions like those seen in SJS and TEN (Fernandez *et al.*, 2008). A high prevalence of the cutaneous lymphocyte antigen (CLA) and the HLA-DR allele is associated with CD4+ T cells in MPE, with the active secretion of cytolytic molecules identified. CLA is a chemokine responsible for the homing of T cells to the skin, upon arrival a T_h1 phenotype (interferon (IFN)-γ release) has been observed (Fernandez *et al.*, 2008). It is these cytolytic molecules that result in the observed clinical maculopapular rash.

A number of drugs have been implicated in the progression of MPE, ranging from β-lactam antibiotics, quinolones and allopurinol (Romano *et al.*, 1995). Amoxicillin induced exanthema is associated with the viral infection, mononucleosis, caused by the Epstein-Barr virus (EBV). Patients with EBV infection are at higher risk of developing MPE triggered by an immune response to the β-lactam antibiotic amoxicillin. The incidence of MPE with EBV when

administering phenoxymethylpenicillin or tetracycline is 14-23%, however with amoxicillin this rises to 28-69%, with the incidence rising to 100% in children. It is understood that viral infection increases the incidence of drug mediated T cell activation due to cytokine/regulatory environment within the host. Alternatively, it is possible that altered drug metabolism could be as a result of alteration of the drug metabolism enzymes through viral infection. Either way, it is well documented that the incidence of MPE is increased during viral infection. The generally low severity of MPE means it is often treated in a fairly standardized way with cessation of the drug. As re-challenge with delayed-type hypersensitivity results in a more severe reaction, re-attempts to offer the same drug is often avoided (Renn *et al.*, 2002).

1.1.1.3.2 DRUG REACTIONS WITH EOSINOPHILIA AND SYSTEMIC SYMPTOMS

Drug reaction with eosinophilia and systemic symptoms (DRESS) is often referred to as hypersensitivity syndrome. Although it is mainly associated with damage of the liver, skin eruptions still occur with symptoms including rash, fever and eosinophilia (Cacoub *et al.*, 2011). DRESS is characterized by a long latency period despite discontinuation of the culprit drug, with periods of relapse (Niu *et al.*, 2015). IL-5 is the main mediator in disease progression which leads to the active recruitment of eosinophils. As with MPE, viral involvement is implicated. In 76% of patients with DRESS, viral reactivation of latent herpes virus (EBV, human herpes virus (HHV) -6, HHV-7 and cytomegalovirus (CMV)) occurs. Cross reactivity of EBV-driven CD8+ T cells have been observed resulting in damage to multiple organs in DRESS. This is believed to assist in the development of the condition and the development of autoimmune sequelae (Picard *et al.*, 2010; Niu *et al.*, 2015). DRESS is primarily mediated with CD8+ T cells. It is believed that the continuous priming of naïve CD8+ T cells to herpes virus is responsible for the prolonged pathogenesis of DRESS (Niu *et al.*, 2015). Multiple drugs have been associated with the onset of DRESS, including allopurinol, carbamazepine and SMX. Due to the overactivation of the immune system treatment often

includes the use of steroids. However, due to the severity of the disease 8% of patients do not survive (Peyrière *et al.*, 2006).

1.1.1.3.3 Stevens-Johnson Syndrome/Toxic epidermal necrolysis

SJS is one of the more serious forms of drug hypersensitivity estimated to affect 1 in 6 million people annually (Saito *et al.*, 2013). The clinical manifestation of SJS is a severe loss of skin attachment with a diagnosis made when detachment accounts for 10-30% of the total skin surface area. The mechanisms underlying SJS are not fully understood however natural killer (NK) lymphocytes and CD8+ cytotoxic lymphocytes are heavily implicated leading to the death of keratinocytes. The cytokine release includes granzyme B and perforin mediated pathways and include Fas/Fas-ligand involvement (Abe *et al.*, 2003; Chung *et al.*, 2008). Once skin detachment increases beyond the upper threshold of 30% in SJS, TEN is diagnosed. The incidence for TEN decreases from 1 in 6 to 1 in 2 million people affected (Schwartz, McDonough and Lee, 2013). Similar mechanisms are observed in TEN as SJS with a large-scale destruction of keratinocytes. It is not surprising that in both SJS and TEN multi-organ failure can occur with a mortality rate between 10-50% in all cases (Gomes and Demoly, 2005; Schneck *et al.*, 2008). One of the challenges in understanding SJS/TEN is the lack of *in vitro* tests to identify the offending agents. Although drugs are implicated in 95% of SJS/TEN reactions it is not ethically possible to re-challenge the patients to confirm the causative drug (Tripathi *et al.*, 2011).

As is often the case with DHRs, co-morbidities are a factor in disease progression. For example, HIV is known to be a risk factor for the development of SJS/TEN due to the disruption of the immune system. In many cases the genetic predisposition of a patient can make them more susceptible to DHRs from certain drugs, discussed in subsequent sections.

1.1.2 DRUG INDUCED LIVER INJURY

So far the nature of DHRs have been focussed on reactions relating to injury to the skin. In several cases ADRs result in the development of liver damage, which can be fatal. In this section we will discuss the incidence of liver reactions to drugs, as well as the mechanistic involvement and key compounds that have been associated.

DILI remains one of the leading causes of drug attrition, with around 33% of new therapeutic agents resulting in hepatotoxicity (Daly *et al.*, 2009). The incidence rates of DILI are often unknown in many populations due to the qualitative nature of clinical reporting. While traditionally pharmaceutical products were thought to be the sole initiators in DILI, it's now known that herbal remedies and dietary supplements can lead to adverse events (Navarro and Seeff, 2013). A lack of relevant biomarkers, guidelines to reporting and the complex nature of clinical presentation contribute to inconsistencies between clinicians (Davern *et al.*, 2011). The main inconsistencies lie with diagnoses made upon exclusion, i.e. after elimination of common causes of liver disease such as alcoholic hepatitis, and a high level of suspicion (Davern *et al.*, 2011; Fernández-Murga *et al.*, 2018). In addition case reporting in medical journals is often met with limited requirements for publishing adverse events, if any at all (Agarwal *et al.*, 2010). A thorough review in 2002 into a French cohort found that 8 in 100,000 patients presented with an ADR related to DILI over a 13 year period (Sgro *et al.*, 2002). While relatively rare, DILI accounts for 50% of all hospital admissions in the USA relating to acute liver failure, with 39% of these stemming from acetaminophen (APAP) overdose causing hepatotoxicity and 13% from idiosyncratic DILI (iDILI) triggered by other drugs (Holt and Ju, 2006). Strikingly, patients presenting with acute liver failure as a result of an ADR have a 25% chance of recovering their own liver function (Ostapowicz *et al.*, 2002), with transplantation almost certainly being required. In the clinic DILI refers to a broad term that can symptomatically mimic many forms of acute and chronic liver aetiologies (Stephens, Andrade and Lucena, 2014). This further contributes to the challenges faced with accurately

reporting genuine adverse events. In this section we will discuss common forms of DILI, and the immune mechanisms by which they are modulated.

1.1.2.1 LIVER PHYSIOLOGY

The liver (Figure 1.2) is the largest solid organ in the body and is critical to metabolic processes and immune function. Anatomically the liver splits into two lobes which subsequently form 16 segments (8 in each lobe). In every segment an estimated 1,000 lobules exist, each comprising of canaliculi flowing towards the common bile duct and small sinusoids or capillaries flowing from branches of the hepatic portal vein and hepatic artery to a 'central' vein, a branch of the hepatic vein. As the major organ in the body for metabolism, the liver is also a major site of biotransformation and for the detoxification and excretion of alcohol, drugs, other xenobiotics, and environmental toxins (Fernández-Murga *et al.*, 2018). These compounds are delivered to the liver through the portal vein in the blood derived from the digestive organs, where they are filtered by hepatocytes. Hepatocytes are also responsible for determining the uptake and storage of nutrients as well as any compounds that should be returned to the blood or eliminated through the transport of bile into the duodenum. In addition to detoxifying small molecules the liver is capable of breaking down proteins into ammonia. As ammonia is toxic to the body in high concentrations this is further converted into urea which is released into the blood and filtered by the kidneys where it is excreted in urine.

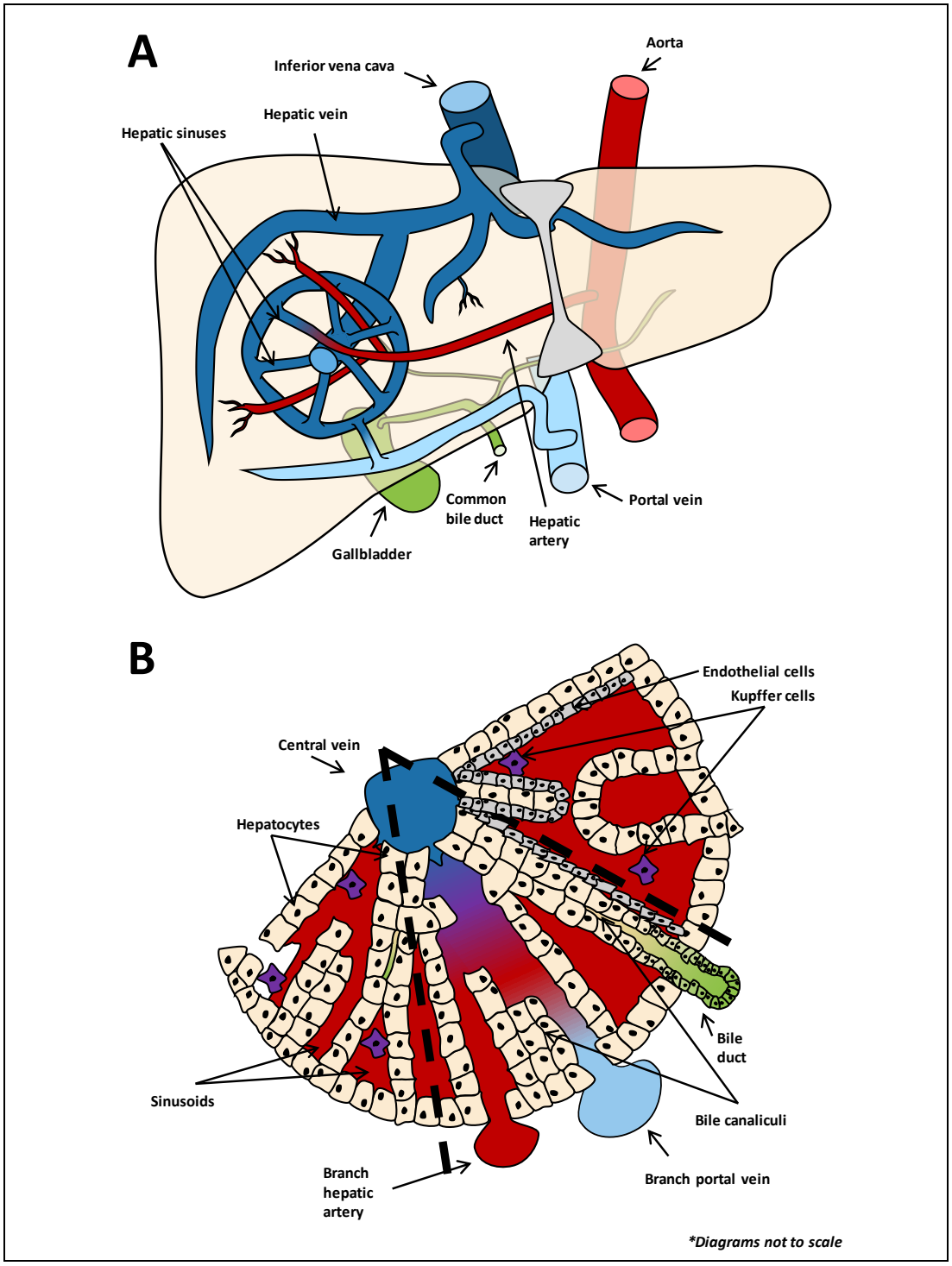


Figure 1.2. Structure of the liver. A) the liver is a large organ supplied by multiple ateries and veins. Oxygenated blood is supplied to the liver via a branch of the aorta. It exits via the hepatic vein to the inferior vena cava. The portal vein supplies blood from the digestive system for the liver to undertake its primary functions of metabolism, detoxification and others. B) Close up schematic of the lobules of the liver. Blood enters the lobules through branches of the portal vein and the hepatic artery where they are mixed in the sinusoids. Excretion products are transported into the bile canaliculi where they enter the bile duct and are subsequently extreted into the duodenum. Resident immune cells, such as Kupffer cells, are present in the sinusoids of the liver.

While generally perceived as an organ primarily for the metabolism and detoxification of molecules, it is known that the liver is deeply diverse in cells of the immune system. The constant bombardment of molecules within the liver makes it a site of continuously regulated inflammation in order to encourage tissue remodelling and repair. This makes for a complex balance of regulation as, although the resident immune cells must tolerate all of the commensal and dietary microorganisms it becomes exposed to, it must also be ready to respond to danger (Holt and Ju, 2006; Robinson, Harmon and O'Farrelly, 2016).

The innate immune system is the first line of defence towards pathogens and other foreign invaders. The liver has local immune mechanisms to cope with a variety of potential antagonists including pathogens, toxins, tumour cells and harmless dietary antigens. Kupffer cells account for the largest population of tissue macrophages in the liver, with NK and NKT cells also contributing to resident innate immunity (Li and Diehl, 2003). Phagocytosis of pathogens entering the liver through the portal venous blood is mediated largely by Kupffer cells (Holt and Ju, 2006). In addition the release of cytokines, nitric acid, prostanoids and reactive oxygen intermediates help to regulate hepatic inflammation as well as modulate NK and NKT cell phenotype (Hashimoto *et al.*, 1995; Tsutsui *et al.*, 1997). Kupffer cells also play an essential role in the presentation of antigens to the adaptive immune system's T lymphocytes. The liver is also home to a unique combination of resident lymphocytes. As well as the conventional CD4+ and CD8+ T lymphocytes the liver contains a high percentage of γδ (15-25% of hepatic T cells) and CD4- CD8- T cells (Exley and Koziel, 2004; Gao, Jeong and Tian, 2008). Functionally γδ T cells possess characteristics of both the innate and adaptive immune system acting as a bridge between the two. The high incidence of this subset of T cells in the liver is through their ability to have a protective as well as pathogenic responses, and often accumulate in inflamed liver tissue (Hammerich and Tacke, 2014).

One characteristic of the liver is its preference for innate cells to favour tolerance rather than activated immunity. A number of studies show this, with the tolerance of dietary antigens a

leading example. More strikingly the liver has been shown to be accepted in cases of allogenic organ transplantation. It has also been shown that pre-exposure of both soluble antigens and donor cells through the portal vein increases in the occurrence of systemic and solid tissue immune tolerance, respectively. Active suppression, immune deviation and the apoptosis of activated T cells are all thought to play a part in this immune tolerance (Holt and Ju, 2006). To confirm these hypotheses several studies have been performed. Transgenic T cell receptor (TCR) models have shown T cells from the spleen and lymph nodes that have become activated accumulate in the liver before undergoing apoptosis (Huang *et al.*, 1994; Bertolino *et al.*, 1995). This may demonstrate a clearance pathway whereby T cells that are programmed to undergo apoptosis (through increased B220 expression and a loss of recognition receptors) are trapped in the liver, dubbing the liver as the "T cell graveyard" (Crispe *et al.*, 2000). Another hypothesis for the induction of immune tolerance within the liver is through the priming of naïve T cells within the sinusoids where the blood flow is slowed down. Naïve T cell priming by liver sinusoidal endothelial (LSEC), hepatic dendritic, and Kupffer cells results in tolerance rather than activation, partly by the failure to differentiate into Th1 or cytotoxic cells. In addition, Kupffer and hepatic dendritic cells (DCs) have been shown to be less effective at antigen presentation when compared to immune cells in other systems (Rubinstein, Roska and Lipsky, 1986; Callery, Kamei and Flye, 1989; Knolle *et al.*, 1999; Limmer *et al.*, 2000). These unique immune responses allow the liver to efficiently remove pathogens, clear particles and soluble molecules from the circulation and delete activated T cells, all while maintaining a tolerance to food antigens (Holt and Ju, 2006).

1.1.2.2 CLINICAL MANIFESTATIONS OF DILI

The mechanisms by which DILI is triggered are still not fully understood, however they are widely accepted to follow two pathways; direct hepatotoxicity and adverse immune reactions. Intrinsic reactions based on the pharmacology of the drug are largely responsible for direct hepatotoxicity, with known thresholds needing to be reached before damage

occurs. APAP overdose is a well characterized example of this, with the intrinsic nature making it easier to develop animal models to aid in scientific understanding (McGill *et al.*, 2012). On the other hand, adverse immune reactions are idiosyncratic, making the pharmacology or dose correlate less well with predicting negative clinical outcomes. Of course, the latter makes it difficult for medical professionals to monitor the progression and susceptibility of disease. Due to this idiosyncratic nature reliable animal models are not available, limiting our understanding to pharmacogenetic studies and findings from intrinsic animal models (Stephens, Andrade and Lucena, 2014).

1.1.2.2.1 HEPATITIS

DILI is estimated to result in 10% of all hospital admissions through acute hepatitis, with hepatitis representing the most common presentation of DILI (Bleibel *et al.*, 2007; Fernández-Murga *et al.*, 2018). Severity can range from slight elevations in liver function tests (LFTs), transaminase elevation, hyperbilirubinemia, and diminished liver function (e.g. coagulation), to severe conditions ultimately resulting in death. Biochemical patterns can be used to predict the severity and type of DILI being observed. In the case of hepatitis an alanine aminotransaminase (ALT) result 2 times above the normal upper limit and alkaline phosphatase (ALP) within normal limits is a warning sign. Upon identifying elevated ALTs between 5 to 90 days after the initiation of the treatment clinical monitoring is performed in order to lead to a clinical diagnosis. If LFTs reduce >50% over a period of 8 days subsequent to cessation of the drug DILI is often reported. This is less likely should LFTs reduce to this value of a 30-day period. The challenge often faced is previous history of the offending drug. Drugs that cause DILI are quickly modified/improved/(withdrawn) and clinicians are often reluctant to report DILI for new compounds until clinical and/or scientific reports/papers have been published (Bleibel *et al.*, 2007).

Hepatitis can be the result of an intrinsic drug reaction, or through immune mediated (allergic) reactions. Mechanisms relating to both inductions of DILI will be described in the

next section. Patterns relating to hepatitis are from the onset of hepatocellular injury and have been associated with a wide range of pharmaceuticals through both immune mediated and non-immune mediated pathways. Although elevation in ALT is used to predict hepatocellular toxicity it is still difficult to determine the severity of the adverse event, with histology being one of the few robust methods available (Verma and Kaplowitz, 2009). The practise of diagnosing DILI reactions based on exclusion criteria can generate a bias towards only considering other common aetiologies. In the Western world the incidence of hepatitis E virus (HEV) infection is relatively low, so therefore is not usually considered as a differential diagnosis. This can ultimately result in the generation of incomplete statistics which further hampers the ability to carefully monitor DILI as a whole. In fact, of patient serum samples recruited from the DILI network (DILIN), a prospective study on patients with DILI, 16% tested positive for HEV IgG and 3% for HEV IgM. Of the 9 patients that tested positive for IgM, the most likely cause for their DILI diagnosis was HEV (Davern *et al.*, 2011). Interestingly this is not the only reported case of HEV infection resulting in the diagnoses of DILI. A retrospective study in the UK found that up to 13% of patients diagnosed with DILI tested positive for HEV, recommending HEV screening as standard in cases of suspected DILI (Dalton *et al.*, 2007). While these cases do only represent a small proportion of patients, and retrospective studies cannot prove the diagnoses for a past event, it is important to consider other differential diagnoses that may exist. This further highlights the need for robust biomarkers for genuine DILI along with tight guidelines for clinical reporting.

1.1.2.2.2 CHOLESTASIS

Presentation of choleostatic liver injury is variable with symptoms of jaundice, fever and pruritus often observed. Due to symptoms being common between other liver aetiologies clinical LFTs resulting in an elevation two times above the upper normal limit of ALP are often indicative of a choleostatic event. Intrahepatic build-up of bile acids is a characteristic phenomenon often only seen in cholestatic DILI and results from the ineffective transport of

bilirubin or bile salts and the ineffective flow of bile from the common bile duct into the duodenum (Sundaram and Björnsson, 2017; Fernández-Murga *et al.*, 2018). The build-up of these compounds makes them a target for early diagnostic tests to detect the onset of choleostatic liver injury early on. Bilirubin is a degradation product of haemoglobin which is conjugated with glucuronic acid in the liver. Subsequently bilirubin is secreted in bile. Upon choleostatic liver injury the flow of bile is disrupted resulting in the elevation of conjugated bilirubin. Linking elevated bilirubin with increases in ALP is often found to be a good diagnostic marker for choleostatic liver injury (Fernández-Murga *et al.*, 2018).

As with hepatitis, cholestasis can present as a whole host of symptoms and has been linked to a range of anti-infectious, anti-diabetic, anti-inflammatory and psychotropic agents, to name a few (Fernández-Murga *et al.*, 2018). Of these, anti-infectious drugs are thought to represent 32% of idiosyncratic liver injury with amoxicillin-clavulanate (AC) contributing to a large percentage of reported cases (Andrade *et al.*, 2005). Several features may arise from cholestasis; idiosyncratic liver injury can be characterized by the injury of bile ducts and includes 'vanishing bile duct syndrome'. Diagnosis of vanishing bile duct syndrome is rare (0.5% of patients with cholestasis) and is reported when only 50% of bile ducts are seen upon biopsy. Although our understanding is limited it is believe that it is a T cell mediated reaction where antigens from the biliary epithelial cells are targeted (Sundaram and Björnsson, 2017). The involvement of the immune system in DILI will be explored in detail in subsequent sections. Histologically, cholestasis leading to the dilation of bile canaliculi can be referred to as 'bland' presentation due to the lack of tissue necrosis or inflammation (Verma and Kaplowitz, 2009; Sundaram and Björnsson, 2017; Chatterjee and Annaert, 2018). It is estimated that around 50% of DILI reactions result in cholestasis (Petrov *et al.*, 2018). In recent years choleostatic liver injury has become increasingly important to understand; part to its frequent manifestation in DILI and poor prognosis. Recent figures have put the mortality rate of choleostatic DILI as high as 10% (Sundaram and Björnsson, 2017). While it

is estimated that over 1,000 drugs can lead to the onset of DILI only a few have been proven causal (Fernández-Murga *et al.*, 2018). In a lot of cases LFT abnormalities will begin to reverse upon termination of taking the drug, however this is often likely to require a longer amount of time compared with hepatitis (Sundaram and Björnsson, 2017). Although it is not possible to categorise a particular drug to cause either hepatocellular damage or cholestasis, the chemical properties of certain drugs may make a patient more susceptible to one form of DILI. Drugs possessing a difluorinated side chain are relatively lipophilic, as in the case of the quinolones temafloxacin and trovafloxacin, making them more likely to lead to choleostatic injury (Sundaram and Björnsson, 2017). Although risk factors do exist, such as age, other co-morbidities and genetic determinants, idiosyncratic liver injury leading to cholestasis can still not be predicted (Sundaram and Björnsson, 2017).

1.1.2.2.3 MIXED PATTERN OF INJURY

DILI can be categorized into a third diagnosis, with a mixed pattern of injury being reported in certain cases. The inability to correctly define hepatitis or cholestasis using the resources available to clinicians results in the 'appearance' of both conditions. When this is the case, and evidence for hepatocellular damage and choleostatic injury is present, a diagnosis of mixed pattern of injury is made. Generally, an increase in both ALT (above 2 time the upper normal limit) and ALP would lead to a diagnosis of mixed pattern of injury. Challenges with this type of DILI can lie with treatments plans due to what is essentially two distinct forms of liver disease presenting alongside each other (Bleibel *et al.*, 2007; Sundaram and Björnsson, 2017). Interestingly the incidence of mixed pattern appears to be more common with an older patient. This is thought to be linked to age being a risk factor for cholestasis, which is hypothesised to be due to a reduced expression of hepatocellular transporters resulting in disruption to the flow of bile (Sundaram and Björnsson, 2017).

1.2 CELLULAR MEDIATORS OF ADRS

The immune system is vital in maintaining a state of health in the body. External barriers to infection such as skin and through a complex network of cellular interactions, the immune system protects us from disease, pathogens and other potentially harmful bodies. One of the most important features of the immune system is to be able to tell self from non-self to prevent the catastrophic destruction of healthy tissue. In the context of this thesis, we are interested in the unwanted immune activation that occurs following drug administration. Broadly speaking the immune system can be categorised into two arms, innate and adaptive. The cellular components of the innate and adaptive immune response work together, with innate immunity responding immediately thus recruiting adaptive immunity (Figure 1.3). This section will review the major components of the immune system, including the interactions resulting in its activation, and key causes of drug hypersensitivity. It is first important to determine the difference between antigen, immunogen and hapten. For the purpose of this thesis an antigen is defined as 'a molecule that binds with high affinity to an immunological receptor', irrespective of any subsequent response. An immunogen can be an antigen and is a substance that 'stimulates an immune response'. Finally a hapten is 'a low molecular weight (MW) chemical which covalently modifies a protein/macromolecule', for example a drug bound to protein (Uetrecht and Naisbitt, 2013).

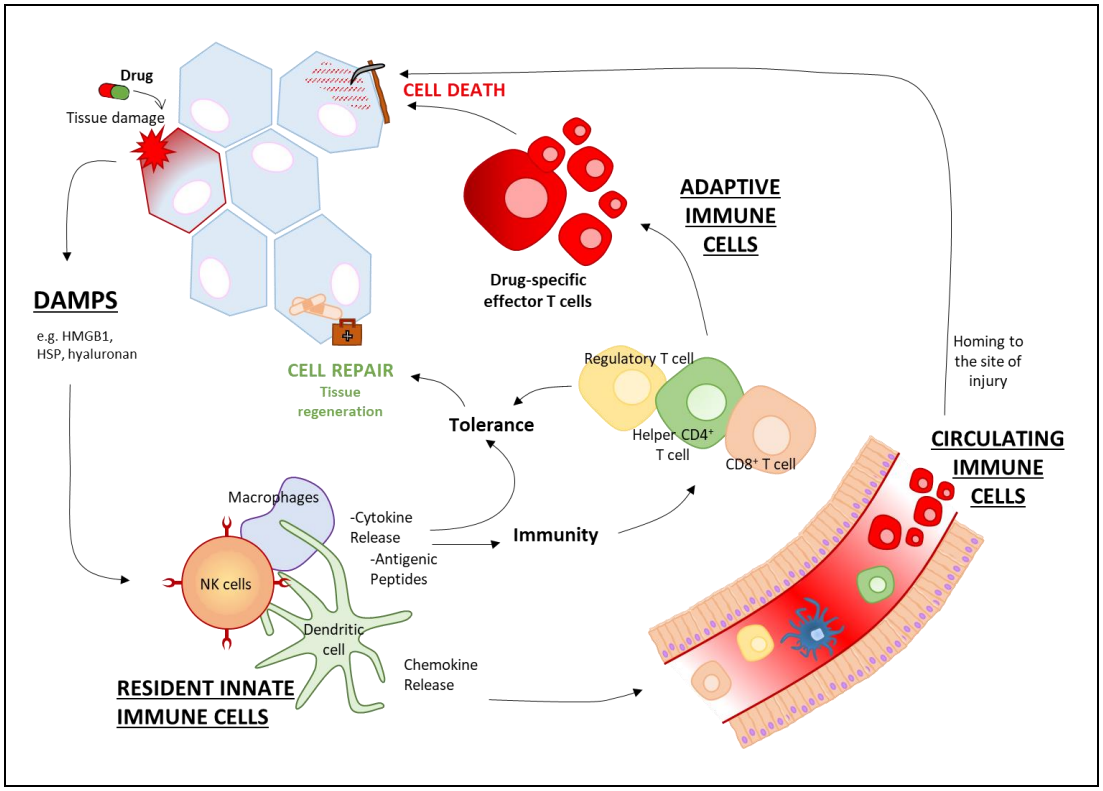


Figure 1.3. Immune-mediated drug-induced tissue injury. Innate immune cells (NK cell, macrophages and DCs) can be activated by damage-associated molecular patterns molecules released by apoptotic and necrotic cells. The innate immune cells can promote either tolerance/regeneration by the production of cytokines and antigenic peptides. Adaptive immune cells can be activated by a drug and its metabolites, drug-modified proteins, or drug altered self –peptides. Drug specific T cells that are either expanded at the site of injury, or that have migrated from blood, can lead to further tissue damage.

1.2.1 The Innate immune system

The innate immune system acts as a first response to pathogens/foreign bodies. Physical barriers, cellular components and the humoral response can all work independently or in tandem to clear the pathogen. Physical barriers are those such as skin, the gastrointestinal system, nasal cavity and tears, all assist in the prevention of pathogens from entering circulation/tissue. Cellular components of the innate immune system are important for when pathogens overcome the physical barriers to infection. Conserved molecular structures required for the life cycle of pathogens are recognized by cells of the innate immune system activating a response. Pathogen associated molecular patterns (PAMPs) are recognized by pattern recognition receptors (PRRs) encoded for in the germ line and require no 'learning' prior to initial exposure. Once activated, an array of cells are involved in the induction of the

inflammatory response mediated by secreted cytokines and chemokines, culminating in the eradication of the pathogen. Toll like receptors (TLRs) are perhaps the most extensively studied PRR and are considered to be the primary sensors of the innate immune system. In total 10 TLR family members have been identified in humans and are situated both on the cell surface (TLR1, 2, 4, 5 and 6) and in intracellular endocytic compartments (TLR3, 7, 8 and 9). More recently TLR10 has been shown to present in an inhibitory and regulatory role, with activation resulting in suppression of TLR2 responses. TLRs situated on the cell surface are responsible for the detection of PAMPs in the external environment including bacteria, fungi and protozoa whereas intracellular TLRs primarily recognize foreign nucleic acid PAMPs derived from bacteria and viruses (Kumar, Kawai and Akira, 2011; Oosting *et al.*, 2014). Once activated, TLRs initiate a series of signalling pathways resulting in increased phagocytosis and the release of pro-inflammatory cytokines leading to further immune recruitment. In addition to PAMPs, damage associated molecular patterns (DAMPs) released during necrotic tissue death and trauma can initiate an innate immune response (Medzhitov, Preston-Hurlburt and Janeway, 1997; Pétrilli *et al.*, 2007; Kumar, Kawai and Akira, 2011).

1.2.1.1 MONOCYTES

Monocytes are large white blood cells derived in the bone marrow from precursor cells. Circulating monocytes make up an estimated 10% of peripheral leukocytes and can further differentiate into tissue macrophages, dendritic cells (DCs) and foam cells. Their function within the innate immune response is primarily against pathogen defence, tumour-specific immune responses and inflammatory diseases such as atherosclerosis. Monocytes are known to play a critical role in a number of innate responses including phagocytosis, antigen presentation and cytokine production. Depending on their PRR expression, specifically CD14 and CD16, monocytes can be further categorised into various subsets. Classical monocytes expressing high levels of CD14 (CD14++) and low levels of CD16- (CD16-) are the most abundant monocyte and express CC-receptor 2 (CCR2). CCR2 is a chemokine receptor that

assists in the trafficking of monocytes. These classical DCs have high MHC II expression and characterisitcially long dendrite extensions. Intermediate (CD14++ CD16+) and non-classical monocytes (CD14+ CD16++) have a reduced expression of CCR2 and have a more functional role in migrating on the luminal surface of endothelial cells of small blood vessels, a process referred to as patrolling (Shi and Pamer, 2011). Recently it has been shown that intermediate monocytes express higher numbers of MHC class II molecules. These findings suggest that classical monocytes are recruited into an antigen presentation role during the inflammatory response (Shi and Pamer, 2011).

The differentiation of monocytes was first identified in seminal studies performed by Ebert and Florey in 1939 (Ebert and Florey, 1939). Terminal differentiation occurs upon tissue infiltration as a result of chemotaxis. During the inflammatory response nucleated cells increase CC-chemokine ligand 2 (CCL2) expression through pro-inflammatory cytokine stimulation or innate immune receptor activation by PAMPs/DAMPs. Monocytes are one of very few cells to express CCR2 and so are recruited to the site of inflammation (Nahrendorf, Pittet and Swirski, 2010; Shi and Pamer, 2011). This tissue micro-environment determines the differentiation of monocytes. For example, the presence of granulocyte-macrophage colony-stimulating factor (GM-CSF) and IL-4 are both involved in the differentiation to macrophages. Classical monocytes are more susceptible to differentiation into macrophages based on their surface expression of CCR2 recruiting them to the site of inflammation and so exposure to relevant cytokines (Ohradanova-Repic *et al.*, 2016).

1.2.1.2 DENDRITIC CELLS

DCs have a unique ability to stimulate primary immune responses through antigen presentation resulting in immunological memory. Their principal role is to capture antigens from the extracellular environment and present them on their cell surface, resulting in the activation of the adaptive immune response. DCs are also understood to be involved in

immunological tolerance, with a role in the thymus by deleting self-reactive T cells. Circulating DCs are relatively short-lived and are replaced with progenitor cells relatively quickly. Classical DCs differentiate from bone marrow myeloid progenitor cells, whereas monocyte derived DCs differentiate from monocytes, as previously described. Immature DCs have high phagocytic activity and actively engulf extracellular particles/proteins. Subsequently, immature DCs migrate to lymphoid tissue where maturation results in antigen presentation. Engulfed proteins are digested and peptides are displayed on MHC molecules for recognition by T lymphocytes. It is understood that classical DCs are more efficient in their presentation to T cells in comparison to other subtypes (Granot *et al.*, 2017). Subsequently activated T cells migrate to the site of damage and activate B cells for antibody production. This leads to the generation of plasma cells and so immunological memory. It is believed that upon T cell activation, DCs undergo apoptosis and are replaced by new progenitor cells in the circulatory system (Banchereau *et al.*, 2000; Steinman, Hawiger and Nussenzweig, 2003; Geissmann *et al.*, 2010).

The maturation of DCs occurs when they encounter PAMPs and DAMPs, resulting in phagocytosis and subsequent migration to lymphatic tissue. For this reason, immature DCs are often found residing in the extremities of the body where they encounter a higher number of antigens. In doing so immature DCs will detect danger signals resulting in the release of cytokines to stimulate other cells of the immune system, while making their way to the lymphatic organs to activate the adaptive immune response. Mature DC activation of T cells results in the clonal expansion of antigen specific T cells. When immature DCs activate T cells it is often antigen independent and so immune tolerance is promoted, resulting in the expansion of T regulatory cells (Steinman, 1991; Banchereau *et al.*, 2000; Kapsenberg, 2003; Steinman, Hawiger and Nussenzweig, 2003).

1.2.1.3 MACROPHAGES

Macrophages are fundamental in their role of identifying, ingesting and killing pathogens, as well as activating other cells of the immune system. Without them the modelling of organs, healing of wounds and clearance of pathogens would not be possible (Nathan, 1987). Upon tissue damage or infection, the differentiation of monocytes to macrophages occurs giving them their new primary role (Yang *et al.*, 2014). Their activation was established by Mackaness et al in the 1960s through bacterial proteins increasing antimicrobial activity of macrophages in a dose dependent, but antigen independent manner (Mackaness, 1962). Since then we have furthered our understanding of the factors involved in the activation of macrophages. IFN-γ, IL-12 and IL-18 are all known to activate macrophages. The secretion of cytokines during cell stress/inflammation results in recruiting macrophages to the site to assist with the clearing of cell debris. As with other cells of the innate immune system macrophages are involved in the presentation of peptides on the cell surface to T cells, resulting in the activation of the adaptive immune response (Gordon, 2003).

Tissue specific macrophages are distributed throughout the body ready to identify and respond to physiological changes and pathogen infiltration. Importantly local macrophages underpin the response to infection, either through the clearance and secretion of pro-inflammatory molecules, or through the secretion of signals to restore tissue homeostasis and repair (Gordon and Plüddemann, 2017). Kupffer cells are macrophages only found to reside within the sinusoids of the liver, and were named after C. von Kupffer, the pathologist to first identify this cell type (Decker, 1990). Accounting for 80-90% of the tissue macrophages found to reside within the body, Kupffer cells are the first macrophage population to come into contact with dietary microorganisms (Bilzer, Roggel and Gerbes, 2006). While their function in tissue is specialised, it is similar to that of circulating macrophages. In the case of Kupffer cells, differentiating dietary microorganisms to

pathogenic microorganisms makes them more susceptible to tolerogenic responses compared to circulatory macrophages.

1.2.1.4 GRANULOCYTES

Granulocytes are another cell of the innate immune system responsible for phagocytosis of extracellular pathogens. There are 4 different sub-types of granulocytes including basophils, eosinophils, neutrophils and mast cells. A defining feature of granulocytes is the presence of granules in the cytoplasm and a multi-lobed nucleus. Granulocytes are known to influence the function of DCs contributing to both innate and adaptive immune function (Breedveld *et al.*, 2017). The most abundant phagocytes in the blood are neutrophils, accounting for 65% of all white blood cells. They respond to activation by ingesting microorganisms followed by the release of soluble anti-microbials and the generation of neutrophil extracellular traps (NETs). NETs are made up of fibres composed of chromatin and serine proteases involved in the trapping of extracellular microbes, resulting in the neutralization of pathogens (Clark *et al.*, 2007). Neutrophils contain a number of anti-microbial compounds including defensins to kill bacteria, proteolytic enzymes to digest proteins and lysozyme to break down the lipopolysaccharide of bacterial cell walls, making them professional bacterial killers (Linderkamp *et al.*, 1998; Hickey and Kubes, 2009).

Conversely, basophils are one of the least abundant cells in the blood (~2%) and are pre-loaded with pro-inflammatory mediators. Like neutrophils, basophils are recruited to the site of infection where they will release histamine and leukocytes involved in the IgE response (Qi *et al.*, 2010). Eosinophils are responsible for the destruction of parasites such as helminths and are involved in allergic disease. Upon activation they can secrete a wide variety of cytokines including IL-2, IL-4, IL-5, IL-10, IL-12, IL-13, IL-16, IL-18 and TGFα/β. All these molecules have pro inflammatory effects, including the upregulation of cellular trafficking and the activation of the regulation of vascular permeability. In addition,

eosinophils are involved in antigen presentation, acting as APCs, as well as the destruction of tissue through the release of toxic granule proteins and lipid mediators (Rothenberg and Hogan, 2006).

1.2.1.5 NATURAL KILLER CELLS

NK cells are cytotoxic lymphocytes critical for the function of the innate immune system. Their mechanism of action is like that of CD8+ T cells (discussed in the next section) without antigenic specificity. Generally, NK cells will response to virally infected cells and cells expressing tumour phenotypes (Vivier *et al.*, 2011). NK cells are activated through the presence of ligands that appear in 'distress', such as the stress induced self-ligands recognized by NKG2D. TLRs are also involved in the activation of NK cells, inducing the production of IFN-γ. This been found to be more of an accessory activation, as the secretion of IFN-γ by NK cells increases when in the presence of other immune cells (Vivier *et al.*, 2011). The secretion of IFN-γ acts to recruit other cells to the site of infection/inflammation in addition to increasing the effector function of macrophages. Out of the 2 billion NK cells circulating in the body at any one time, 100% of those residing in the secondary lymphoid tissue (~10% in the blood) express CD56 in high abundance. This surface marker for NK cells can be used to sub-categorise NK cells into $CD56^{bright}$ and $CD56^{dim}$. $CD56^{bright}$ NK cells are those with a high surface marker abundance and can quickly produce immunoregulatory cytokines. In contrast $CD56^{dim}$ populations are terminally differentiated from the bright population with the primary function of cytolytic activity (Caligiuri, 2008; Chester, Fritsch and Kohrt, 2015).

1.2.2 THE ADAPTIVE IMMUNE SYSTEM

The adaptive immune response, sometimes termed acquired, is responsible for the targeting and neutralization of pathogens/foreign bodies that enter the host. As one of the two branches of immunity, the adaptive response works in conjunction with the innate response to achieve its desired effect. In the innate immune section we have discussed the relevance

of the cellular components of innate immunity in triggering the adaptive immune response through antigen presentation to T lymphocytes. T lymphocytes, or simply T cells, are a major component of the adaptive immune response along with B lymphocytes. The precise role of these cell types will be discussed in detail in the following section. Briefly, T cells are responsible for recognising antigens presented to them by other cell types. Should an immunogen be presented, subsequent activation of the T cell results in clonal expansion and the triggering of further signalling cascades. B lymphocytes, once activated, can become a source of immunological memory and actively secrete antibodies to the antigen by which they are exposed to. This results in rapid activation of the adaptive immune system on re-challenge from subsequent exposures.

1.2.2.1 T LYMPHOCYTES

T lymphocytes can be sub-categorised based on their activation and subsequent effector response into cytotoxic T cells, helper T cells and regulatory T cells. Both T cells and B cells (involved in antibody production and immunological memory; discussed in chapter 2) develop from the same pluripotent hemopoietic stem cells. Their main development stems from where they are matured into their respective cell type. Hemopoietic stem cells are generally located in the bone marrow and remains the site of B cell development. T cells are generated from the same precursor cell after migration to primary lymphoid organs, such as the thymus, through the blood. Autoreactive T cells are deleted in the thymus before they migrate further into tissue and secondary lymphoid organs (Alberts, Johnson and Lewis, 2002).

1.2.2.1.1 CYTOTOXIC T LYMPHOCYTES

Cytotoxic T cells are characterized by the presence of the CD8 cell surface marker. Their primary role is to identify infected cells and tumours resulting in the destruction of the target cell. Almost all nucleated cells present MHC class I molecules, responsible for the presentation of antigens derived from intracellular proteins, to the TCR protein on the

surface of CD8+ T cells. In doing so, activation of the T cell results in the release of cytotoxic molecules such as granzyme B, perforin and fas ligand resulting in the death of the antigen presenting cell.

Naïve T cells react to an enormous range of pathogens resulting in their expansion and differentiation into antigen specific killer cells, disseminating throughout the body to clear the infection. After initial priming in secondary lymphoid tissue and migration to the site of infection CD8+ T cells release their cytotoxic molecules neutralizing the target cell (Zhang and Bevan, 2011). Initial release of IFN-γ and tissue necrosis factor α (TNF-α) have anti-tumour and anti-microbial effects through direct interference with viral attachment or the induction of apoptosis (Zhang and Bevan, 2011). On release of perforin and granzyme B, the same as those found in NK cells of the innate immune system, cytotoxic effects occur. Perforin forms a pore in the membrane of the target cell, allowing granzymes to enter the infected cell, resulting in the cessation of viral reproduction and apoptosis. Astonishingly there is limited bystander effect from the release of these cytolytic molecules as they are directed along an immunological synapse towards the target cell. It is also known for CD8+ T cells to kill one cell, before moving onto a second and third before exhaustion, termed 'serial killing'. The production of Fas ligand upon activation of cytotoxic T cells allows for binding of the Fas receptor on the target cell. Activation of the Fas receptor results in a signalling cascade culminating in the apoptosis of the target cell. Towards the end of the infectious stage Fas ligand on activated CD8+ T cells can bind to Fas receptors on other CD8+ T cells resulting in the elimination of their immune effector response (Carter and Dutton, 1995; Zhang and Bevan, 2011).

1.2.2.1.2 HELPER T LYMPHOCYTES

Helper T cells (T_h) express the CD4 surface marker and have a primary role in the activation of further immune cells such as macrophages and phagocytic cells. They have a varied role in immune responses and can trigger a multitude of effector functions depending on their

maturation. T_h cells interact with antigen presenting cells which phagocytose extracellular proteins and display peptides on the cell surface through MHC class II mechanisms. Unlike CD8+ cytotoxic T cells, naïve T_h cells (T_h0) differentiate into a wide number of sub-sets upon antigen exposure and activation (Figure 1.4). The cytokines that are present during the time of activation is indicative of the differentiation that is undergone. Figure 1.4 is by no means an exhaustive list of the cytokines involved and the number of sub-sets that have been identified, however this well illustrates the components involved in T_h differentiation and effector function (Akdis and Akdis, 2009; Uetrecht and Naisbitt, 2013).

Initial studies of T_h sub sets revealed the differentiation into T_h1 and T_h2 cells (Mosmann and Coffman, 1989), differentiated by IL-2, IL-4 and IL-12. It is also recorded that exposure of IFN-γ can result in this differentiation showing the complex nature of these cytokine microenvironments. T_h1 cells primarily secrete IFN-γ resulting in the recruitment of a cytotoxic response, whereas T_h2 cells secrete cytokines known to be involved in class switching, recruiting B lymphocytes and mast cells associated with the development of antibody mediated responses and immunological memory. At the time of their discovery it was widely believed that the modulation of adaptive immunity and the recruitment of other immune cells and complexes was as a result of T_h1 and T_h2 subsets, however more recently many other sub-sets have been defined (Abul K. Abbas, Andrew H. Lichtman, 2012). The more recently identified sub-sets of T_h cells (T_h9, T_h17, T_h22 and T_{fh}) give greater details to the responses and definition of their specific roles in the adaptive immune response.

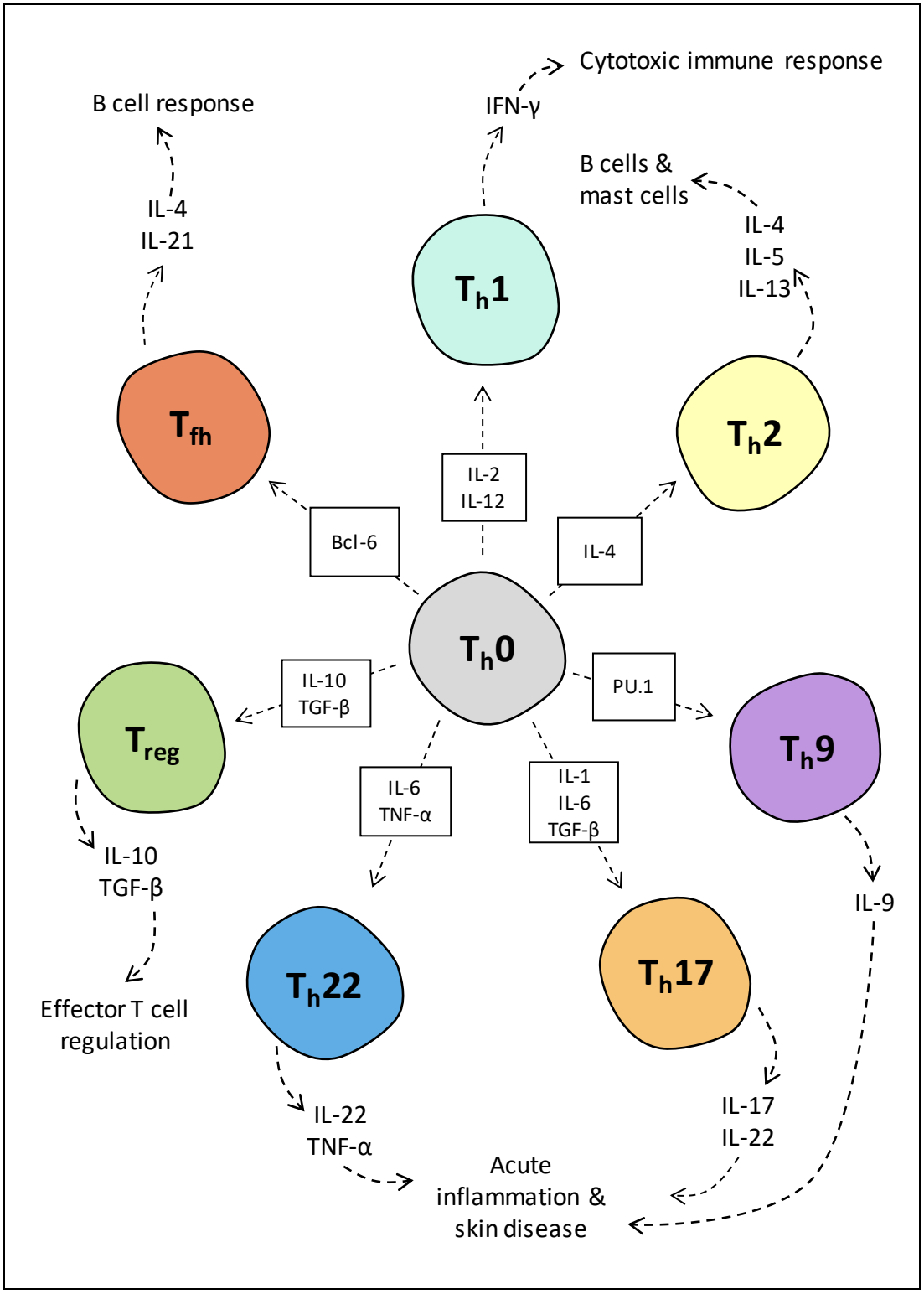


Figure 1.4. CD4+ Helper T cell differentiation. Upon activation the cytokine microenvironment determines the differentiation of T_h0 cells. Once activated and matured, each different T_h subset has a different effector function within the immune response. The main cytokines involved in the differentiation of T_h0 cells, and their effector cytokine secretions are displayed.

The recent identification of T_h9 cells is attributed to the secretion of IL-9, a cytokine associated with asthma, anaphylaxis and helminth immunity. IL-9 is also known to stimulate the survival of T cells, and assists in class switching to IgE production from B cells (Veldhoen *et al.*, 2008; Licona-Limón *et al.*, 2013). Their new discovery means that their primary role is still under scrutiny, however, they have been implicated in a number of conditions and are understood to recruit of a number of effector cells. It has also been discovered that T_h9 cells can directly cause tissue damage through the release of TNFα and granzyme B (Schlapbach *et al.*, 2014). T_h17 and T_h22 have too been implicated acute inflammation and tissue damage, in particular the skin. T_h17 and T_h22 are both differentiated in the presence of IL6, with T_h22 requiring the additional signal from TGF-β to fully differentiate. Both subsets have the capacity to secrete IL-22, involved in the inflammatory response, with T_h17 cells also capable of secreting IL-17, associated with immune regulation (Pennino *et al.*, 2010; Uetrecht and Naisbitt, 2013).

Follicular helper T cells (T_{fh}) have a primary role in the priming of B lymphocytes to a specific antigen. This is a similar role to T_h2 cells in promoting immunological memory and antibody production. Regulatory T cells, or T_{reg} cells are anti-inflammatory and are differentiated by TGF-β and IL-10. TGF-β plays a predominant role in immune tolerance and is the primary cytokine secreted by T_{reg} cells. Regulatory cells are an important component in the onset of disease and are becoming better understood. As they are heavily involved in antigenic tolerance their dysregulation can result in an unwanted immunological response. While T_h9 cells have also been implicated in the secretion of TGF-β it is becoming more apparent that they are more important in promoting inflammation rather than tolerance (Li and Flavell, 2008; Kaplan, 2013; Uetrecht and Naisbitt, 2013; Vinuesa *et al.*, 2016; Kobayashi *et al.*, 2017).

1.2.2.2 T LYMPHOCYTE ACTIVATION

The stimulation of T lymphocytes can be achieved through 3 signalling mechanisms; defined as signals 1, 2 and 3. Signal 1 is provided through the interaction of the peptide MHC complex (pMHC) and the TCR, whether that be a CD4+ or CD8+ T cell. This interaction alone is not enough for the full activation of the T cell and subsequent signalling cascades and will only usually result in toleration. For full activation of the T cell signal 2 must also be present. These relate to the interaction between signalling molecules that are non-specific to the antigen. Signalling molecules can be both co-stimulatory and co-inhibitory with the outcome dependant on the interplay between the two. Finally, signal 3 refers to the signalling molecules, such as cytokines, released from APCs following antigen uptake. In all, these signals are imperative in defining the nature of the full cellular immune response (Curtsinger *et al.*, 1999).

The interaction between the TCR and pMHC is crucial for T cell clonal expansion and the activation of the adaptive immune response. Several hypotheses exist for how the interaction occurs, however the basis of the interaction remains the same. The αβ TCR receptor recognizes the pMHC as a single unit, resulting in subsequent intracellular signalling and an effector response. Studies have been performed to look at the importance of co-stimulatory signals in T cell activation. It is apparent that in many cases a single pMHC would be unable to stimulate a T cell on its own, with studies characterizing short half-lives for the interaction between the two molecules. The interaction between the pMHC and the TCR is likened more to that of an antibody-antigen, rather than the peptide-MHC interaction. This results in an often-rapid dissociation between the TCR and pMHC, likely to result in failed activation. It is therefore apparent that multiple interactions are involved leading to the activation of T cells (Corr *et al.*, 1994). This interaction between the pMHC and the TCR has been eloquently referred to as the 'immunological synapse', consisting of a cluster of TCRs surrounded by a ring of integrin family adhesion molecules. The formation of this synapse is

likely to assist in the formation of sustained pMHC TCR engagement. There are a number of factors attributing to the often-short half-life of the pMHC TCR complex. Large glycoproteins on the cell surface of both the APC and the T cell, relative to the size of the TCR/MHC, may act as a physical barrier preventing the interaction. It is also important to think about the number of antigenic peptides that are being presented at any one time. If the abundance is low, it is less likely to activate the T cell above a required threshold. This is also in addition to the movement of cells reliant on the contact between a TCR and the pMHC complex. It may be the case that the antigenic peptide, although has high MHC binding affinity, may not have a high affinity to the TCR. Taking all of these factors into account it is clear that co-stimulatory molecules are important in sustained TCR engagement with the pMHC resulting in the initiation of the tyrosine phosphorylation cascades and further signalling mechanisms (Grakoui *et al.*, 1999).

Broadly, MHC I and II activate cytotoxic CD8+ and helper CD4+ T cells respectively, resulting in the cellular immune response. In TCR activation of CD8+ T cells the CD8 surface molecules act as a co-receptor assisting the TCR binding to pMHC I (Figure 1.5) (CD4 in CD4+ T cells). Once the TCR/CD8 and pMHC I association is formed Lck, a tyrosine kinase, interacts with the cytoplasmic region of the CD8 surface marker leading to the phosphorylation of the CD3 immunoreceptor tyrosine-based activation motifs (ITAMs). The recruitment and subsequent activation of ZAP-70 follows as a result of CD3 phosphorylation, in turn activating LAT. LAT kinase links with the TCR and facilitates signalling during CD8 T cell activation.

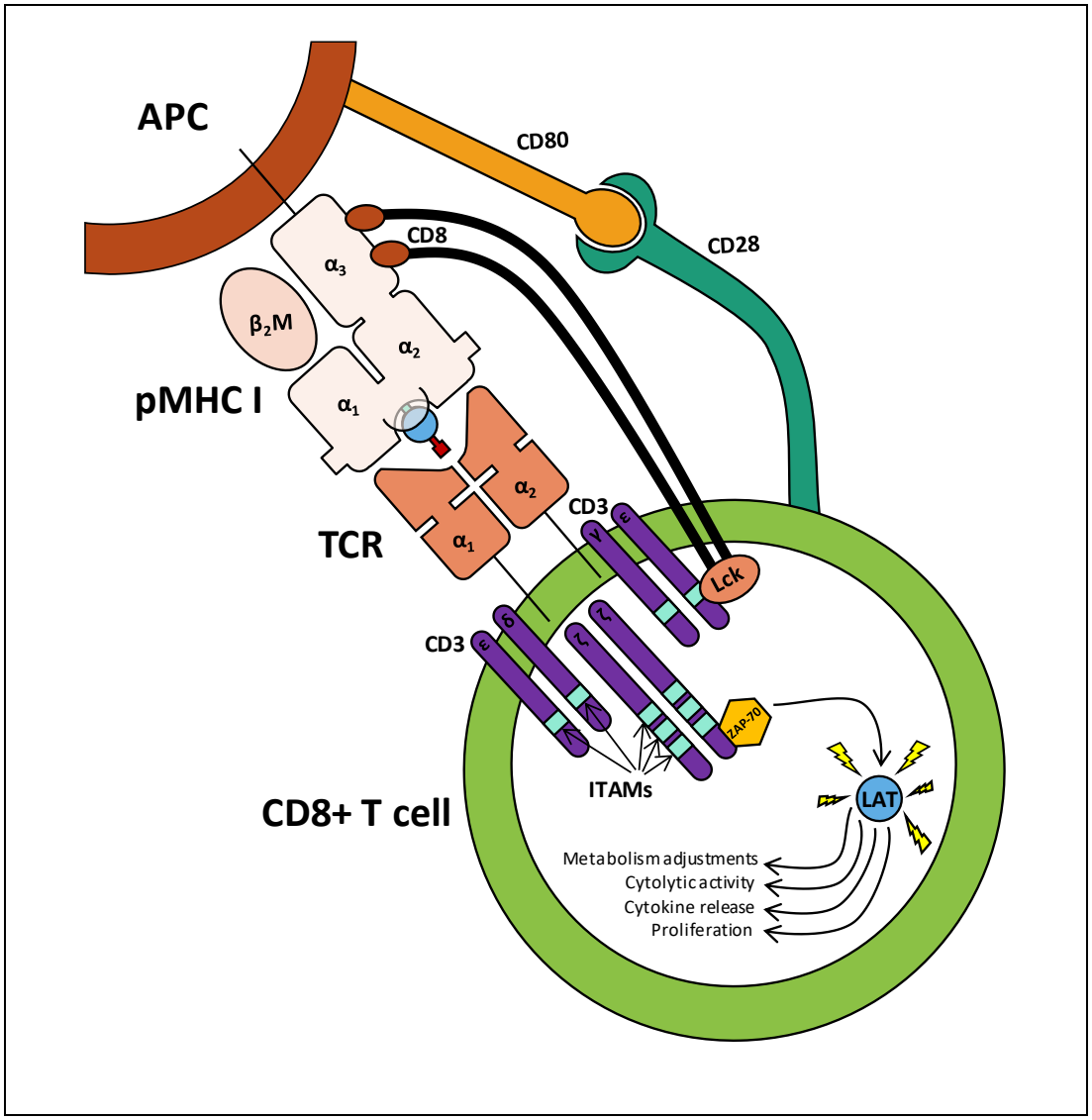


Figure 1.5. CD8+ TCR activation by MHC I. Interaction between the pMHC complex of class I molecules and the αβ chain of the CD8+ TCR initiated the adaptive immune response (signal 1). The CD8 co-receptor interacts with the α3 chain of the MHC I protein further amplifying the response through the activation of Lck. Subsequent phosphorylation of CDS results in the recruitment and activation of ZAP-70. LAT activation through ZAP-7 results in the effector response, determined by the co-stimulatory molecules present. Co-receptor signalling between CD28 and CD80 results in further activation leading to T cell proliferation (signal 2).

Described as a 'signalosome' it can result in a host of cellular responses including cytokine release, cytolytic activity, proliferation and metabolic adjustments. For example, the release of diacylglycerol (DAG) results in the activation of MEK/ERK signalling leading to the upregulation of CD69 surface expression, representing full cell activation. Inositol triphosphase (IP3) release results in the release of stored calcium ions from the ER, critical to the production of IL-2 and lymphocyte survival (D'Ambrosio *et al.*, 1994; Liu *et al.*, 1998).

While signal 1, the pMHC/TCR interaction, is central to the activation of T cells from specific antigens, signal 2 is required for their full activation. Several co-stimulatory (CD28, CD40 and CD54) and co-inhibitory (CTLA-4, PD-1, CD45, KIR, TNFR1) molecules on the surface of T cells are important in determining the balance between activation or tolerance (Wang and Reinherz, 2012; Rosenberg and Huang, 2018). Signal 2 has also been implicated in the modulation of T lymphocyte differentiation into Th1 and Th17. CD28 is a well characterized signal 2 pathway involving the cell surface expression of CD28 on the T cell and CD80/CD86 on the APC (Figure 1.5). Through CD28-CD80/CD86 signalling, the activation of phosphokinase C triggers the release of IL-2 promoting T lymphocyte survival. Inhibition of CD28 has been shown to prevent T cell activation, making it an important checkpoint for autoimmunity (Liu *et al.*, 1998; Acuto and Michel, 2003; Sharpe and Abbas, 2006; Mizui *et al.*, 2008). Other signal 2 pathways focus on a co-inhibitory effect and include programmed death-1 (PD-1), cytotoxic T lymphocyte-associated protein 4 (CTLA4) and T cell immunoglobulin and mucin-domain containing-3 (TIM-3). All these cell surface molecules are a regulator of T cell inhibition and have been shown to regulate autoimmunity on a wide number of cell types. T cell activation leads to the expression of PD-1 on the cell surface which in turn can be bound by is ligand, PD-L1. The ligand is only present when induced by other pro-inflammatory mediators such as IFN-γ. Activation of PD-1 leads to the dephosphorylation of ZAP-70 through SHP-1 and SHP-2, resulting in the deactivation of the T cell. Other co-inhibitory pathways work in tandem with PD-1 giving the overall desired effect (Nishimura *et al.*, 1999; Sheppard *et al.*, 2004; Keir, Freeman and Sharpe, 2007; Chen and Flies, 2013).

T cell surface receptors also include those that respond to chemokines, cytokines and antibodies (Fc receptors). Activation of these receptors represents the signal 3 pathway of activation. It has been shown that the absence of signal 3 results in poor viability of naïve T cells with low effector function (Mescher *et al.*, 2006). Signal 3 is important in the activation

of CD8+ T cells through IL-12 and/or IFNα/β cytokines to have both a productive response and to avoid the induction of death and/or tolerance to the antigen. It is understood that cytoskeleton remodelling can be controlled with signal 3 molecules through the regulation of >350 genes. CD4 cells are also believed to need a third signal to result in a productive response with IL-1 thought to provide this signal (Curtsinger and Mescher, 2010). Although T cells can respond to antigens presented by MHC molecules, signal 1 alone is unlikely to result in an active immune response. Signal 2 molecules act to regulate immune responses to prevent over activation of the immune system while signal 3 allows T cells to assess the cellular microenvironment. The absence of any of these signals will usually result in a limited response of T lymphocytes.

1.2.2.2.1 MECHANISMS INVOLVED IN DILI

It is generally accepted that a number of individual patient susceptibility traits contribute to the onset of disease. Notably genes that encode for HLA and cytochrome (CYP) P450 enzymes play a role in the small percentage of patients that will develop the adverse reaction. Using the ATL/ATP ratios is a relatively easy way to assess the potential for the onset of DILI, however, it is common for several drugs to elevate LFTs without causing significant injury. The liver is an unusual organ in that the immune system promotes tolerance and regeneration over activation, meaning drug exposure may initially result in LFT increases which subsequently return to normal levels. While the mechanisms underpinning the reduced LFT over time are still unknown, this adaptation to injury is seen in ~3% of patients taking statins (Rashid, Goldin and Wright, 2004; Bleibel *et al.*, 2007).

Chemically reactive drugs and/or reactive metabolites (RMs) are generally implicated in the induction of DILI through the altering of normal cellular processes. Binding of drugs (or reactive metabolites) to cellular macromolecules can result in a whole host of disruption including protein dysfunction, DNA damage, oxidative stress and lipid peroxidation. Additionally cellular energy production can be reduced through mitochondrial dysfunction owed to disruption of ionic gradients or changes to intracellular calcium stores (Holt and Ju, 2006). Drug metabolism to RMs is in a many cases necessary in order to achieve the desired pharmacological effect, however, when not controlled it can lead to the disruption mentioned above (Stephens, Andrade and Lucena, 2014). Drug metabolism can be categorized into three stages; I, II and III. Phase I drug metabolism is mostly mediated by the family of CYP P450 enzymes usually through oxidation and reduction. The cells within the liver largely express CYPs making the liver a prime target for RMs as it is the central organ for drug metabolism. Working in coordination with phase I metabolism is phase II, or detoxification (Yuan and Kaplowitz, 2013). Often, it is the fine balance between phases I and II that is critical in determining the incidence of DILI. Small molecules and proteins are formed

with the purpose of quenching/inactivating RMs resulting in their detoxification. Phase II metabolism acts to change the physical properties of RMs in order to assist with their exit from the cell. Conjugation with glucuronate, acetyl and glutathione (GSH) all enhance the hydrophobicity of RMs resulting in their excretion into the bile or urine through phase III metabolism. Phase III is largely mediated by ATP-binding cassette (ABC) transporters, culminating in the cellular excretion of detoxified metabolism products. (Holt and Ju, 2006; Yuan and Kaplowitz, 2013; Stephens, Andrade and Lucena, 2014). Importantly, membrane transporters are thought to underlie some forms of DILI. The role of phase III metabolism will be discussed in further detail in chapter 3.

Although mechanisms underpinning DILI, in particular cholestasis, have been investigated since the 1980's, they are still somewhat debatable (Fernández-Murga *et al.*, 2018). Understanding mechanistic pathways resulting in DILI is largely through the study of APAP metabolism, perhaps the most well characterized form of hepatotoxicity. APAP is largely metabolised through glucuronidation and sulphation, however, phase I metabolism by CYP2E1 to *N*-acetyl-*p*-benzo-quinone imine (NAPQI) contributes to hepatotoxicity. NAPQI binds the free thiols on intracellular GSH (phase II metabolism) resulting in depletion of cytosolic and mitochondrial stores. Once cell defences are depleted NAPQI can bind to free thiols on off-target proteins resulting in their dysfunction. Disruption in mitochondrial proteins results in the generation of mitochondrial reactive oxygen species, leading to necrosis of the cell (Yuan and Kaplowitz, 2013). APAP is a form of intrinsic DILI resulting from an increase in APAP concentration that is directly related to NAPQI production and so hepatotoxicity. There is still however some element of patient susceptibility through expression of CYP enzymes making them a target for genetic DILI studies. Variable expression and the catalytic activity of the CYP enzymes may result in an increased abundance of RMs which subsequently reduces intracellular levels of phase II quenching molecules. To date

however, compelling evidence is yet to be discovered (Holt and Ju, 2006; Bleibel *et al.*, 2007; Stephens, Andrade and Lucena, 2014).

1.2.3 MHC AND ANTIGEN PROCESSING

The immune system acts as a constant surveillance system identifying abnormalities in the proteins present and/or expressed within individual cells and tissues. The MHC is a protein encoded by human leukocyte antigen (HLA) genes in humans. The HLA coding site is often described as the most complex and polymorphic region of the genome, with in excess of 220 genes encoding up to 10,000 different allelic forms (Robinson *et al.*, 2015). The HLA system is located within the 6p21.3 region of the short arm of chromosome six comprising of 21 polymorphic genes. The immune responses to pathogens and tumours are reliant on the product of these gene clusters and can be further categorised into three regions (Caron *et al.*, 2015; Robinson *et al.*, 2015). HLA class I and HLA class II genes code for the MHC class I and MHC class II protein products respectively. These are located within the telomeric ends (class I) and the centromeric end (class II) of the MHC. MHC can be further classified into HLA-A, -B, -C, -E, -F & -G (MHC class I) and HLA-DP, -DM, -DOB, -DQ & -DR (MHC class II), based on their protein structure and function (Pritchard *et al.*, 2015). The third region in the cluster encodes for a number of non-HLA genes that carry immune function (Robinson *et al.*, 2015). The complex nature of this system makes for a complex nomenclature. Briefly, HLA alleles are named based on their gene, allelic group and HLA protein, with subsequent coding for other polymorphisms (Figure 1.6).

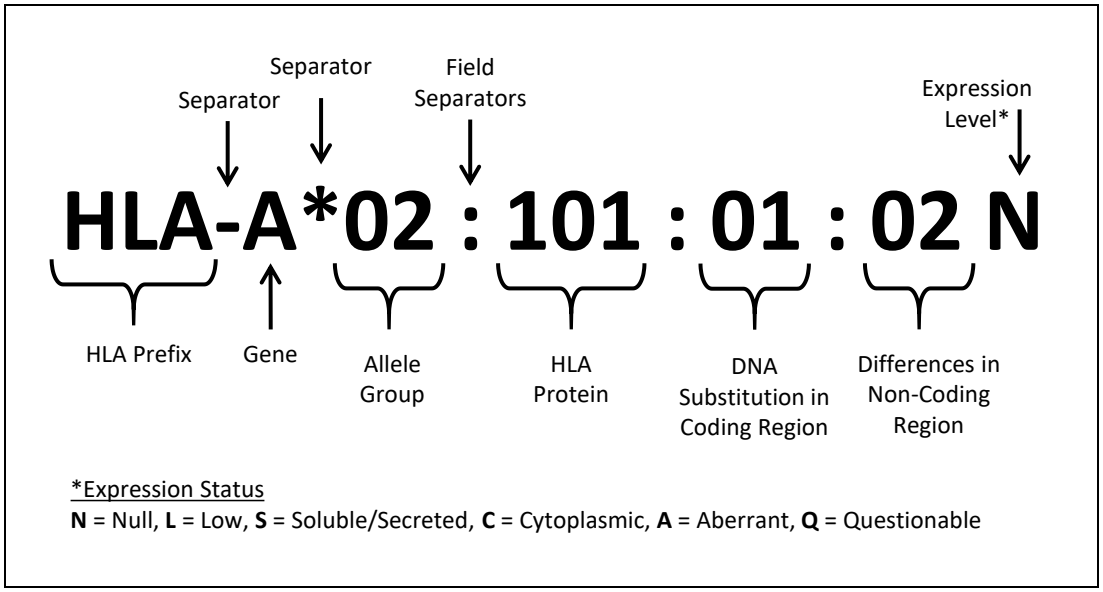


Figure 1.6. HLA allele nomenclature. The diverse range of HLA alleles results in a complicated system in order to take into account all of the known polymorphisms. The HLA prefix is always followed by a hyphen separator and the gene name. An asterisk is followed by the allelic group, a further separator and the HLA protein. For genes that require further annotation DNA substitutions in the coding and non-coding regions may follow, and in some cases includes the expression level.

MHC proteins are responsible for presenting peptides to circulating and tissue resident T cells. Normally, self-peptides derived from endogenous proteins are presented on the cell surface of nucleated cells. When neo-antigens appear, from pathogenic proteins or tumour cells, immune activation follows. This explains the need for a huge diversity of genes that encode for HLAs. Typically HLA polymorphisms occur in the genomic regions encoding the HLA binding cleft. This gives a large diversity in the different peptides that can be loaded and subsequently presented to the immune system. It is for this reason that certain populations have better immunity to local pathogens, due to positive selection for the genes that better present certain peptides over an evolutionary period. While the mechanisms regulating these interactions are tightly controlled, breakdown in molecular communication can lead to severe clinical conditions (Fortier *et al.*, 2008; Caron *et al.*, 2015).

MHC class I molecules appear on all nucleated cells and are responsible for presenting short peptides to CD8+ T cells. Peptides of 8-12 amino acids can typically be loaded into the MHC binding cleft. Classical MHC I alleles include HLA-A, -B and -C, with each estimated to present

1,000 to 10,000 peptides. As investigations into the MHC peptide repertoire expands the number of known MHC ligands increases. If we consider 10,000 peptides can be presented by a single MHC class I allele, that number represents <0.1% of the number of potential 9-mer ligands that could be derived from the human protein coding genes. This further highlights the specificity of individual HLA alleles, with the collection of peptides presented by a specific allele termed the immunopeptidome. There are several functions of the MHC I immunopeptidome, these include shaping the repertoire of developing thymocytes, immune surveillance and amplification of responses to intracellular pathogens and tumours. The peptides presented by MHC I are largely derived from proteasomal degradation of cytosolic proteins. The source protein can be from a pathogen (culminating in the initiation of T cell activation), functional endogenous proteins or defective ribosomal products that arise from defective protein synthesis. If a host cell becomes infected by a pathogen MHC class I present peptides to cytotoxic T cells resulting in neutralization of the cell (Caron *et al.*, 2015; Robinson *et al.*, 2015; Abelin *et al.*, 2017; Reeves and James, 2017).

Structurally, MHC class I molecules are formed of two heterodimers consisting of one membrane-anchored heavy chain with extracellular α1, α2 and α3 regions, and a soluble β2-microglobulin. The peptide binding grove is located between the α1 and α2 sites of the MHC heavy chain (Figure 1.7) (Madden, 1995). The MHC I binding site consists of conserved polar tyrosine residues at each end, forming hydrogen bonds at the N- and C-terminal positions (P1 and PΩ) restricting the length of the peptides that can be accommodated. This has been confirmed using X-ray crystallography and is conserved both across all MHC class I alleles (in humans and mice) as well as peptides of different length, with long peptides accommodation a bulging phenotype in the middle. Exception to this rule does exist; peptides that are too short to accommodate the entire binding groove may still be presented with only one termini stabilized with a hydrogen bond. Additionally, more recent studies have identified N-terminal extension of MHC class I peptides out of the HLA binding groove. These interactions

at the peptide termini are however separate to the anchoring residues and are sequence independent (Bouvier and Wiley, 1994; Madden, 1995; Rammensee, 1995; Pymm *et al.*, 2017; Wieczorek *et al.*, 2017). The specific binding motif is different for each allele, making the peptides that can be presented by each allele different. Typically for MHC class I proteins, peptides are anchored at positions 2 (P2) or P5/P6 and the C-terminal position (PΩ) (Wieczorek *et al.*, 2017).

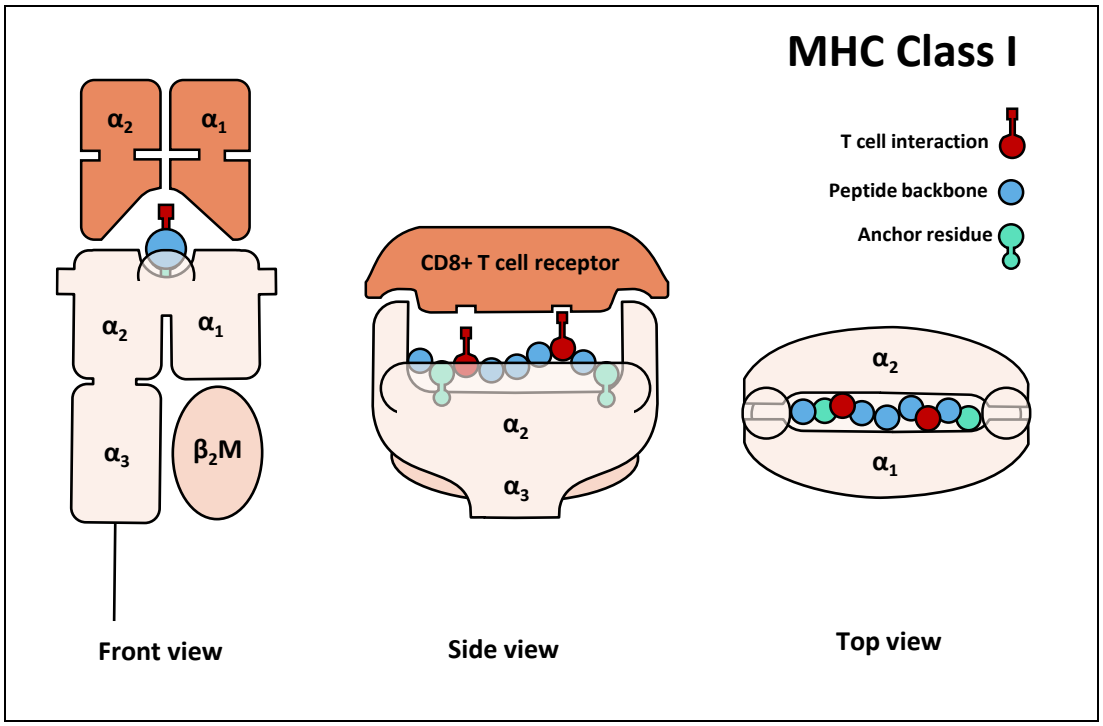


Figure 1.7. MHC class I structure from the front (left) side (middle) and top (right). MHC class I molecules are formed of two heterodimers consisting of one membrane bound heavy chain (α1, α2 and α3) and one soluble β2-microglobulin. MHC class I peptides are recognized by CD8+ TCRs where interactions between the TCR and MHC protein, and TCR and peptide results in T cell activation. MHC class I peptides are anchored into the binding groove by anchor residues, often found at P2 and PΩ (C-terminal amino acid). The binding groove of MHC class I peptides is constrained to fit amino acids of (typically) 8-12mers.

The digestion of intracellular proteins by the proteasome and subsequent presentation on MHC class I molecules is a complex mechanism mediated by a number of enzymatic events (Figure 1.8). The preparation of antigens and subsequent loading occurs in two main events. Firstly, proteins are digested in the cytosol by the proteasome. While the proteasome has a defined function in MHC peptide processing it is a complex multi-catalytic enzyme which is

also involved in cell homeostasis ensuring misfolded proteins are eliminated, preventing aggregation. Within the proteasome there are three catalytic sub-units latent membrane protein (LMP) 2, LMP7 and LMP10) which are responsible for the activity of the core 20S subunit. In a normal cellular state IFN-γ (and other inflammatory cytokines) promote the up-regulation of these sub-units increasing the generation of antigenic peptides. Often the cleavage of proteins into smaller peptides will result in the final C-terminal amino acid being prepared. LMP2, LMP7 and LMP10 all favour hydrophobic C-terminal residues promoting binding to the PΩ anchor (Reeves and James, 2017; Wieczorek *et al.*, 2017).

Following digestion within the cytosol, peptides enter the endoplasmic reticulum (ER) where further processing follows. Transporter associated with antigen processing (TAP) (consisting of two ATP-hydrolysing subunits, TAP1 and TAP2) transports peptides from the cytosol into the lumen of the ER, and a major component in the second event in MHC class I presentation. TAP preferentially transports peptides of 11-14 amino acids long, longer than the preferred 8 to 12-mer MHC class I associated ligands. In order for peptides to be loaded on to the MHC, the immature α-chain and the β-2-microglobulin (β_2M) must be associated along with TAP, tapasin (Tpn) and ERp57 to form the mature peptide loading complex. During this mechanism Tpn is instrumental in the correct association of the peptide loading complex and acts as a bridge between the other molecules. This association occurs within the ER through the chaperone proteins calnexin (CNX), binding protein (BiP) and calreticulin (CRT).

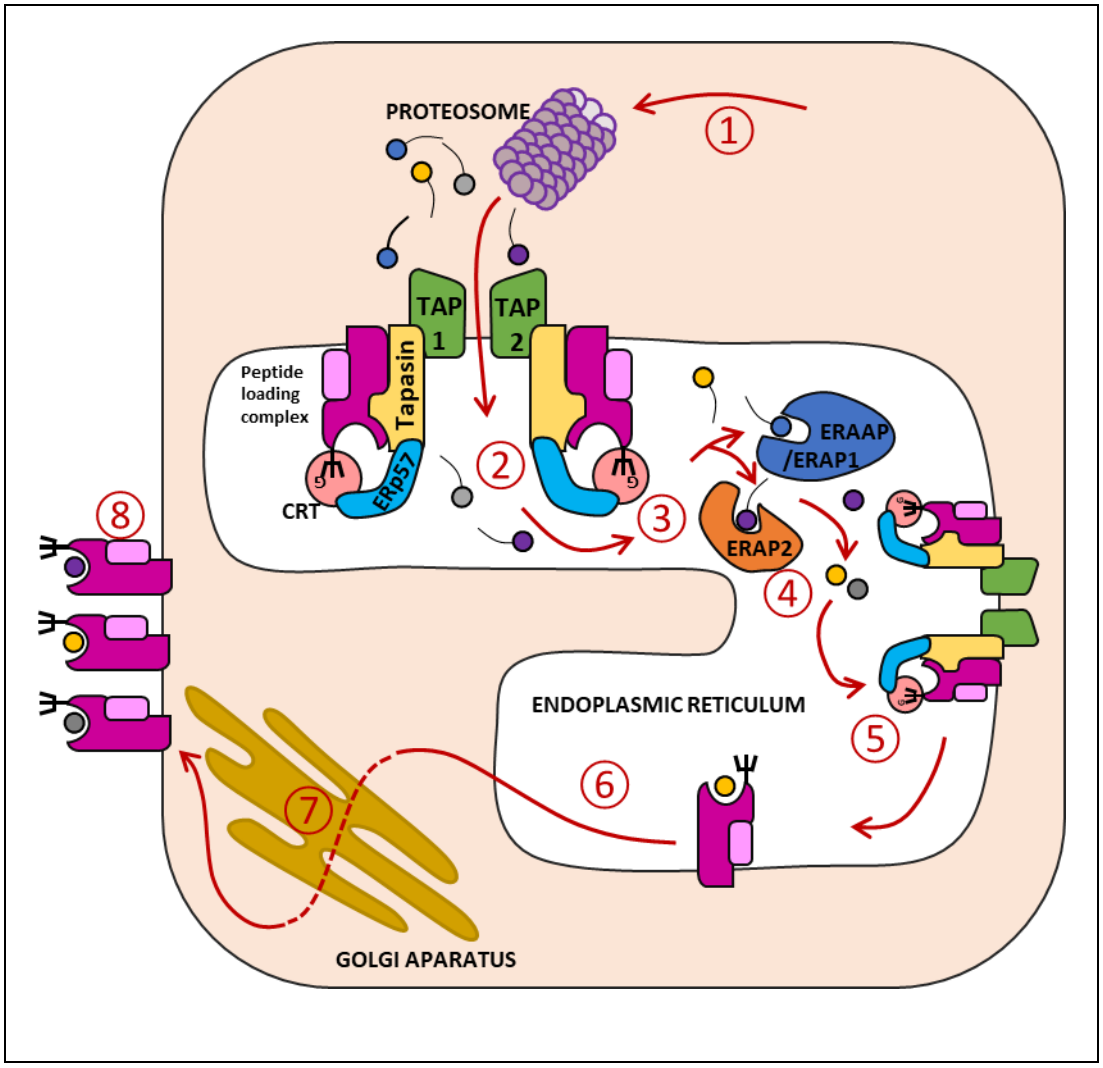


Figure 1.8. TAP antigen processing mechanism. (1) Endogenous proteins are digested by the proteasome in the cytosol. (2) Peptides (mostly 8-16aa) are transported into the ER via the peptide binding pocket formed by TAP1 and TAP2. TAP1 and TAP2 has preference for hydrophobic C-terminal aa's at this point. Tapasin acts as the mediator for the interaction between the MHC and TAP. (3) ERAAP, ERAP1 and ERAP2 trim peptides within the ER to allow them to be loaded onto MHC. (4) Peptides are transported towards (5) ERp57, a thiol oxidoreductase, which catalyses the formation and breakage of disulphide bonds and assists with the loading of the peptide into the MHC binding groove. (6) MHC peptide complexes are translocated out of the ER where they are (7) transported through the golgi and (8) are presented on the cell surface.

The correct assembly of the peptide loading complex is crucial in the ability to bind MHC peptides. As peptides of 11-14 amino acids long would not preferentially bind to MHC class I binding groove in a stable manner, trimming of the N-terminal amino acid residue(s) is performed. Endoplasmic reticulum aminopeptidases (ERAP) 1 and 2 are both involved in N-terminal trimming, generating a pool of antigens with high affinity to the MHC class I binding cleft. Following N-terminal trimming of the peptides ERp57, a thiol oxidoreductase, catalyses

the formation and breakage of disulphide bonds. This mechanisms contributes to the loading of peptides into the MHC binding groove (Reeves and James, 2017; Wieczorek *et al.*, 2017). Once peptides are loaded into the HLA-binding groove the interaction between Tpn and the MHC peptide loading complex is dissociated. This results in the transport of the mature peptide MHC class I complex from the ER to the cell surface, passing through the golgi apparatus within the cytosol.

Although the digestion and subsequent processing in the ER happen independently of one another, impairments in either mechanisms can result in the ineffective presentation of MHC class I peptides. Mechanisms to prevent the processes involved in antigen processing appear in a number of diseases and is extensively studied in cancer progression. Investigations have revealed that a number of cancers have developed to disrupt pathways in all aspects of antigen processing (Mehta *et al.*, 2007). In malignant cells the three subunits contributing to the 20S core of the proteasome are replaced. LMP2, LMP7 and LMP10 reduction significantly reduces the number of peptides that are suitable for MHC binding. Instead the incorporation of β1, β2 and β5 proteins are responsible for the catalytic activity of the 20S core. Disruption in LMP2 and LMP7 have also been reported in the case of cervical carcinoma. Single nucleotide polymorphisms (SNPs) in the genes coding for these proteins results in a worse prognosis for survival rates, thought to be due to reduced protein function. As IFN-γ is known to promote the expression of LMP2, LMP7 and LMP10 in non-malignant cells, therapeutic uses in cancer therapy are being evaluated. Unsurprisingly other mechanisms have identified disruption in the processes occurring within the ER, for example a loss of ERAP1 results in a 50% loss of peptide expression. (Mehta *et al.*, 2007; Hasim *et al.*, 2012; Wehenkel *et al.*, 2012; Reeves and James, 2017). This highlights the challenges faced when trying to understand the mechanisms that are underlying the progression of disease when it comes to MHC class I involvement.

MHC class II differ in terms of both structure and function. Structurally MHC class II alleles consist of two membrane bound heterodimers comprising of one α (α1 and α2) and one β (β1 and β2) chain. The α1 and β1 chains form the peptide binding cleft (Figure 1.9) (Madden, 1995). MHC class II proteins are only presented on professional antigen presenting cells including DCs, macrophages and B lymphocytes. This is in comparison to MHC class I, which are presented on almost all nucleated cells. Conversely to MHC class I, class II molecules are not as restricted in their ability to present longer peptides of typically 10-25 amino acids. This is due to open ends of the binding cleft without the stabilizing of N- and C-terminal amino acids through hydrogen bonds. This results in 'nested sets' of peptides derived from the same protein being presented. Here a stretch of peptide is continually anchored in the binding groove but is flanked on either side by extended sequences protruding the ends of the binding cleft (Rammensee, 1995). Anchoring positions in MHC class II are more commonly associated with P1, P4, P6 and P9. As the N-terminal amino acid on the peptide chain may be protruding, the peptide binding cleft P1 refers to the first anchored amino acid, and is often regarded as the most important anchoring site (Wieczorek *et al.*, 2017).

The peptides presented on MHC class II alleles are the result of proteosomal degradation of endocytosed proteins of extracellular origin. As these peptides are not indicative of an intracellular infection/pathogen cytotoxic CD8+ T cell activation would often result in a detrimental neutralization of the presenting cell. Therefore, peptides presented on MHC class II proteins are detected by CD4+ T cells. The primary response of helper CD4+ T cells is to stimulate B cells for antibody generation or activation of macrophages to enhance their phagocytic functions (Caron *et al.*, 2015).

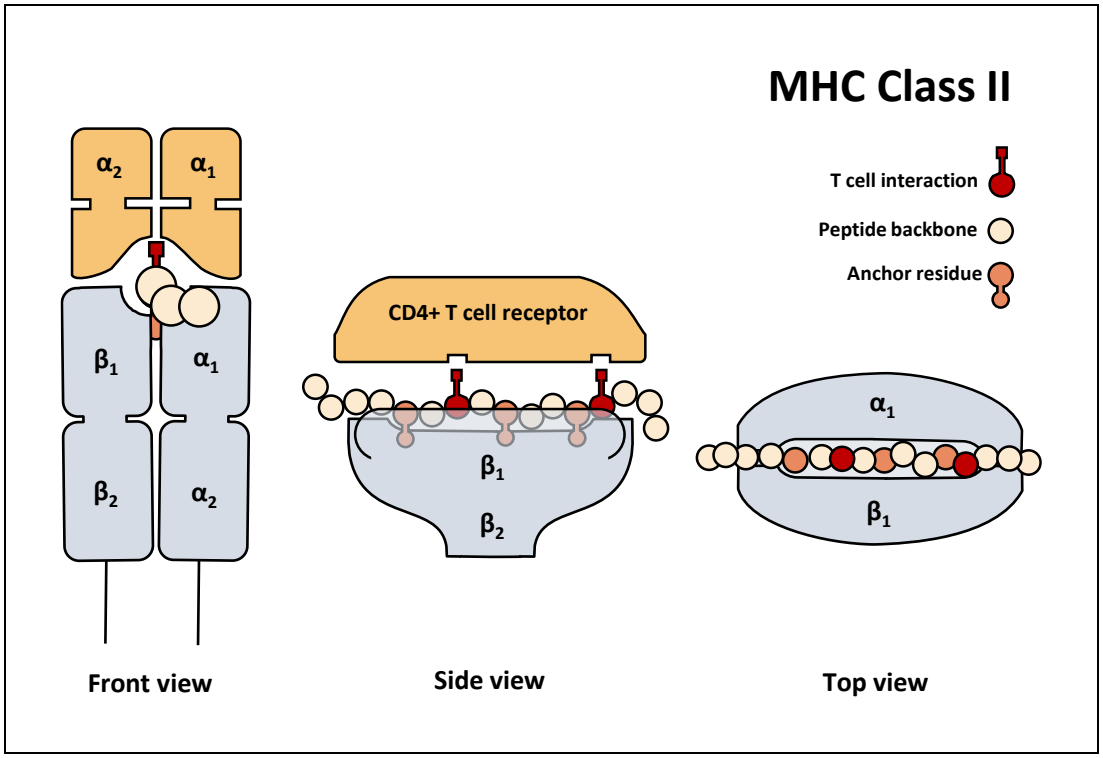


Figure 1.9. MHC class II structure from the front (left) side (middle) and top (right). MHC class II molecules are formed of two heterodimers consisting of two membrane bound heavy chains, one α (α1 and α2) and β (β1 and β2). MHC class II peptides are recognized by CD4+ TCRs and are more frequently derived from extracellular proteins. The open edges of the MHC class II binding groove make it possible for longer peptides to be accommodated, typically 10-25 amino acids long.

In recent years it has become apparent that MHC class II presentation can occur on non-hematopoietic cells. Generally, MHC class II expression on professional APCs results in the migration to the lymphatic system, activating and maturing naïve CD4+ T cells. This atypical presentation of MHC class II on tissue resident cells may be involved in immune tolerance mechanisms, although this is currently not fully understood. We do however know that naïve CD4+ cells can be activated in the absence of DCs, meaning this atypical presentation may have some immunogenic effect (Kambayashi and Laufer, 2014). As with class I, the ability of MHC class II alleles to bind a large range of peptides is based on the diversity of the structure of the binding cleft. Humans present slightly more different allelic forms of class II alleles, with 8 in total, again one from each parent. The different alleles include HLA-DRA, DRB, DQ and DP (Caron *et al.*, 2015).

The mechanisms for antigen processing of the MHC class II pathway are similar to that of MHC class I. Briefly, proteolytically processed peptides/degraded proteins are loaded onto MHC class II proteins though a series of enzymatic reactions comparable to that of MHC class I. While MHC class I antigen processing occurs in the ER, MHC class II antigen processing occurs within late endosomes. Once peptides with high affinity are loaded the MHC class II peptide complex is transported to the cell surface where it is displayed to CD4+ helper T cells (Wieczorek *et al.*, 2017).

1.2.4 Interactions between the pMHC and TCRs in type IV hypersensitivity

As previously described, autoreactive T cells are deleted in the thymus prior to release into the circulatory system (Alberts, Johnson and Lewis, 2002). For T cells to be involved in the onset of DHRs, the interaction between the MHC and TCR must either be disrupted or directly involve drug interactions. The mechanisms by which T cells are activated by drugs and/or RMs through antigen presentation follow three distinct hypotheses (Figure 1.10). These are commonly known as the pharmacological interaction (PI) (Figure 1.10A), hapten (Figure 1.10B) and altered peptide repertoire (Figure 1.10C) hypotheses. The details of these interactions are described in further detail in the next section. Briefly, the drug (and/or metabolite) can interact with the peptide MHC (pMHC) complex and the TCR in three different ways resulting in T cell activation and lymphokine secretion.

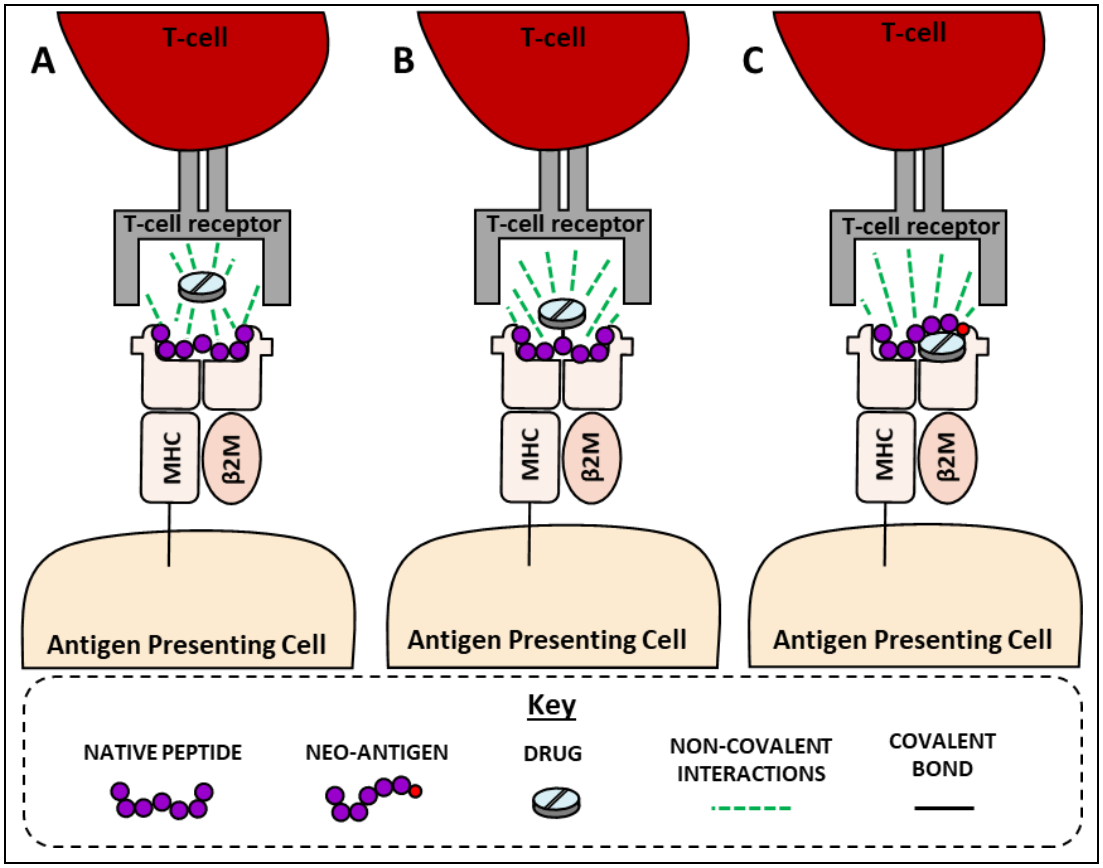


Figure 1.10. Type IV hypersensitivity T cell activation hypotheses. Interactions between the peptide, MHC, drug and T cell receptor (TCR) follow three different hypotheses. (A) The drug interacts non-covalently with the peptide, MHC and the TCR. (B) The drug covalently binds to the peptide presenting a neo-antigen to the TCR. (C) The drug interacts with the MHC altering the conformation of the peptide binding groove so able to accommodate non-drug-modified neo-antigens. In all cases these lead to T cell activation and subsequent lymphokine release.

1.2.4.1 PHARMACOLOGICAL INTERACTION

The PI hypothesis is the result of an independent interaction between small molecules, the pMHC complex and the TCR through non-covalent interactions. The rapid activation of T-lymphocytes from chemically non-reactive drugs is well documented through investigations with sulfamethoxazole (SMX), lamotrigine and carbamazepine. T cell responses to SMX from hypersensitive patients found that clones were highly specific for SMX, and could be activated in the absence of metabolism and antigen processing (Schnyder *et al.*, 1997; Zanni *et al.*, 1998; Watkins W., 2013). As SMX is a chemically inactive compound, and antigen processing was blocked, the interaction between SMX, the TCR and pMHC complex must have occurred in a series of non-covalent interactions. The activation of T cells activated to

SMX has been confirmed to be rapid through the identification of calcium release in < 10 minutes, too short for antigen processing to have occurred. It has also been shown that SMX is not restricted to HLA alleles, but instead stabilizes a number of pMHC complexes (von Greyerz *et al.*, 2001). Cross reactivity is implicated in the activation of clones to compounds such as SMX. Usually the adaptive immune system first recruits T cells to the site of inflammation, resulting in an antigen driven response. It has been found that certain drugs are able to bypass the innate immune response by activating memory T cell responses that are similar to that of other compounds (Pichler, 2005).

While SMX does not have a known HLA restriction, genome wide association studies (GWAS) have identified associations between carbamazepine and HLA-B*15:02. This allele is more commonly found in the Taiwanese population expressing a TCR vβ-11-ISGSY clonotype, making HLA-B*15:02 the only component in the clinical manifestation. In some patients with both the HLA allele and the TCR clonotype, hypersensitivity to carbamazepine was found, making the specific interaction between drug, the MHC and the TCR dependant on the presence of these genes (Chen *et al.*, 2011; Ko *et al.*, 2011; Wei *et al.*, 2012). These specific patient factors make the discovery of DHRs a challenge during drug development as they often happen in a small number of patients that would only be detected when the drug is available to the market.

1.2.4.2 HAPTEN

The oldest model of drug hypersensitivity is the hapten hypothesis, first described in 1935 by Landsteiner and Jacobs. In this study dinitrochlorobenzene (DNCB) was used to sensitize guinea pigs which were later re-challenged with much lower concentrations. Sensitized animals showed a strong response in comparison to the control group suggesting that an immune response to DNCB was responsible. Furthermore, these responses could be seen by administering low doses of DNCB directly onto the skin. Interestingly the authors noted that

immunization would only be likely if there was a combination of these 'simple substances' or drugs with proteins due to previous studies noting the sensitization of rabbits to formaldehyde through exposure to formalinized proteins. Perhaps more convincing in the case of DNCB previous attempts to use bacterial carbohydrates, which are known to have antigenic activity, to produce skin sensitization had failed. These findings resulted in the hypothesis that small molecules bind to proteins resulting in immunogenicity (Landsteiner and Jacobs, 1935).

SMX-nitroso (SMX-NO), the RM of SMX, is known to irreversibly bind to protein and activate T cells *in vitro*. This compound has been used as a model antigen in DHR studies for its interaction both via the PI mechanism and the hapten. SMX-NO has been found to stimulate T cells from patients with drug hypersensitive reactions, indicating the formation of the RM *in vivo*. This interaction between SMX-NO and these activated T cell clones derived from patients could be blocked by inactivating the antigen uptake and presentation, indicating intracellular processing was required (Naisbitt *et al.*, 1999; Lavergne *et al.*, 2009; Castrejon *et al.*, 2010).

Since the initial report by Landsteiner and Jacobs, the binding of small molecules to proteins has been well defined. Jenkins et al characterized the irreversible binding of the β-lactam antibiotic flucloxacillin to human serum albumin (HSA) through nucleophilic attack of the β-lactam ring by primary amines on lysine residues. Importantly it was observed that flucloxacillin bound to lysine residues on HSA from both *in vitro* incubations and *in vivo* patient serum samples (Jenkins *et al.*, 2009). While drug-protein binding does not always lead to hypersensitivity it has to be noted that nucleophilic residues exist on proteins that are important in activating T lymphocytes. Using kinetic spectrophotometric measurements to develop a simple absorbance model Chipinda *et al* identified skin sensitization was mainly driven by electrophilic reactivity of small molecules, further confirming Landsteiner's initial hypothesis (Chipinda *et al.*, 2010).

The hapten hypothesis in drug hypersensitivity involves the irreversible binding of a small molecule, in this case drug, to the peptides presented by the MHC to the TCR. The hapten mechanism can occur through both processing dependant and independent pathways. Drug modification of endogenous proteins may undergo intracellular digestion leading to the presentation of drug modified peptides. Alternatively, the modification of already bound MHC peptides may occur through a processing independent mechanism. Typically, the latter occurs when reactive drugs, or inert compounds (pro-haptens) that undergo metabolism to become RMs, are involved (Naisbitt *et al.*, 2001; Posadas and Pichler, 2007). Whichever mechanism is responsible the drug modified peptides act as neo-antigens to the TCR resulting in T cell activation. Studies using cloned T lymphocytes *in vitro* has confirmed these two mechanisms. Processing independent activation of T cells is observed when antigen presenting cells exposed to reactive drugs/metabolites for a short period of time are washed and presented to the T lymphocyte. Other studies have confirmed processing dependence when T cell activation is lost through blocking antigenic processing pathways (Whitaker *et al.*, 2011; El-Ghaiesh *et al.*, 2012; Jenkins *et al.*, 2013).

1.2.4.3 ALTERED PEPTIDE REPERTOIRE

The most recent hypothesis to be identified is the altered peptide repertoire. In 2002 Mallal first reported the association between the nucleoside reverse transcriptase inhibitor (NRTI) abacavir, used in HIV-1 treatment, and HLA-B*57:01 (Mallal *et al.*, 2002). This discovery lead groups to understand the interactions between this HLA polymorphism and why it was required for abacavir hypersensitivity reactions to occur. It was not until 2012 that multiple groups reported a novel MHC self-repertoire appearing in the presence of abacavir, with neo-self-antigens accounting for 20% of the HLA-B*57:01 immunopeptidome. As abacavir alters the tertiary structure of the MHC and the peptides that can be presented, normally inert/irrelevant MHC peptides are thought to trigger T-lymphocyte, in particular CD8+, activation (Chessman *et al.*, 2008; Bharadwaj *et al.*, 2012; Illing *et al.*, 2012).

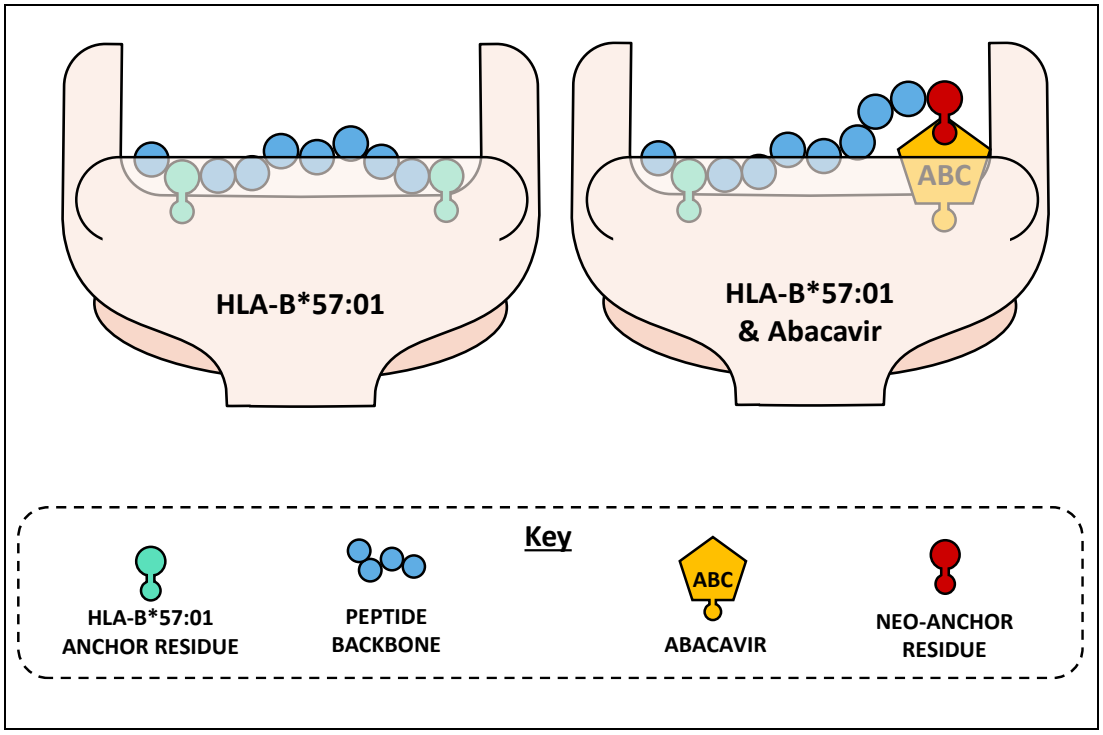


Figure 1.11. Abacavir binds to HLA-B*57:01 altering the PΩ anchor preference. Abacavir binds to the F-pocket of HLA-B*57:01 making C-terminal aromatic amino acids (tryptophan, tyrosine and phenylalanine) unable to bind. Subsequently peptides terminating in isoleucine and leucine accommodate the HLA binding groove giving way to the presentation of a subset of neo-antigens.

Importantly this was only observed with HLA-B*57:01, and not in other closely related HLA alleles (Illing *et al.*, 2012; Norcross *et al.*, 2012). Specifically, abacavir interacts with the F-pocket of the peptide binding cleft making it difficult for peptides terminating in a C-terminal phenylalanine, tryptophan or tyrosine to bind (the usual HLA-B*57:01 C-terminal anchor residues). Instead peptides terminating in leucine and isoleucine are favoured with the additional stabilization from abacavir (Figure 1.11) (Illing *et al.*, 2012). These novel peptide antigens are derived from endogenous proteins and they themselves are not modified by the drug. Thus far only abacavir is known to interact with the MHC in this way. Although covalent interactions between the drug and MHC could result in similar changes to the global peptide repertoire examples are yet to be identified.

1.2.4.4 GENETIC PREDISPOSITION

Several drug-related hypersensitive reactions have been linked to genetic associations. GWAS are now commonly used to identify the genetic similarities between patients suffering

from an ADR to the same drug. Table 1.2 represents drugs that have been associated with different clinical manifestations of DHRs along with their associated genetic predisposition. In most cases carrying the gene does not alone represent the only factor in disease progression. Often it is the culmination of several attributes that result in the onset of disease.

As previously described, abacavir, a retroviral treatment used in HIV infection, is strongly associated with the carriage of HLA-B*57:01. In 2-5% of patients an adverse reaction resulting in fever, rash, nausea and vomiting occurs. While the cessation of abacavir generally results in recovery, re-challenge results in a must more serious disease with mortality reported in a few cases. A number of studies identified the genetic association between abacavir and HLA-B*57:01 giving a positive predictive value of 48% and a negative predictive value of 100%. These numbers represent the effect of HLA-B*57:01 on the likelihood of developing the disease. Carriage of HLA-B*57:01 results in a 48% chance of developing AHS, while 100% of cases are from patients carrying the allele (Mallal *et al.*, 2002, 2008; Illing *et al.*, 2012; Norcross *et al.*, 2012; Martin and Kroetz, 2013).

Table 1.2. Genetic associations between drugs and hypersensitivity reactions. The genetic associations, predominantly HLA alleles, between drugs are described. Associations are categorised into drug reactions with eosinophilia and systemic symptoms (DRESS), Stevens-Johnson syndrome/toxic epidermal necrolysis (SJS/TEN), acute generalized exanthematous pustulosis (AGEP), drug induced liver injury (DILI) and abacavir hypersensitivity syndrome (AHS).

Type of Reaction	Drug	Genetic association	References
DRESS	Phenobarbital	CYP2C19*2	(Manuyakorn *et al.*, 2013; Polak *et al.*, 2014)
	Dapsone	HLA-B*13:01	(Zhang *et al.*, 2013; Polak *et al.*, 2014)
	Carbamazepine	HLA-A*31:01 HLA-B*51:01	(Beeler *et al.*, 2006; Hsiao *et al.*, 2014; Polak *et al.*, 2014)
	Allopurinol	HLA-B*58:01	(Gonçalo *et al.*, 2013; Polak *et al.*, 2014)
SJS/TEN	Lamotrigine	HLA-B*15:02	(Pichler, 2003; Beeler *et al.*, 2006; Phillips *et al.*, 2011; Zeng *et al.*, 2015)
	Carbamazepine	HLA-B*15:02	(Pichler, 2003; Phillips *et al.*, 2011; Hsiao *et al.*, 2014)
	Phenobarbital	CYP2C19*2 HLA-A*02:07 HLA-B*51:01	(Pichler, 2003; Phillips *et al.*, 2011; Kaniwa *et al.*, 2013; Manuyakorn *et al.*, 2013)

	Co-trimoxizole	HLA-B*15:02 HLA-C*06:02 HLA-C*08:01	(Pichler, 2003; Phillips *et al.*, 2011; Kongpan *et al.*, 2015)
	Allopurinol	HLA-B*58:01	(Pichler, 2003; Tassaneeyakul *et al.*, 2009; Phillips *et al.*, 2011; Gonçalo *et al.*, 2013)
	Nevirapine	HLA-B*35:05 HLA-DRB1*01:01 HLA-B*14:02	(Pichler, 2003; Phillips *et al.*, 2011)
	Methazolamide	HLA-B*59:01	(Pichler, 2003; Kim and Lee, 2010; Phillips *et al.*, 2011)
AGEP	Sulfonamides	HLA-DQ3 HLA-DR11 HLA-B*51	(Beylot, Doutre and Beylot-Barry, 1996; Pichler, 2003; Beeler *et al.*, 2006; Pavlos *et al.*, 2014)
	Ampicillin		(Saissi *et al.*, no date; Beylot, Doutre and Beylot-Barry, 1996; Matsumoto *et al.*, 2008; Nacaroglu *et al.*, 2014; Pavlos *et al.*, 2014)
	Amoxicillin		(Saissi *et al.*, no date; Beylot, Doutre and Beylot-Barry, 1996; Matsumoto *et al.*, 2008; Pavlos *et al.*, 2014)
	Terbinafine		(Saissi *et al.*, no date; Beylot, Doutre and Beylot-Barry, 1996; Rubegni *et al.*, 2008; Pavlos *et al.*, 2014)
	Corticosteroids		(Beylot, Doutre and Beylot-Barry, 1996; Buettiker *et al.*, 2006; Pavlos *et al.*, 2014)
	NSAIDs		(Beylot, Doutre and Beylot-Barry, 1996; Byerly *et al.*, 2005; Teixeira, Silva and Selores, 2006; Pavlos *et al.*, 2014)
	Pristinamycin		(Saissi *et al.*, no date; Beylot, Doutre and Beylot-Barry, 1996; Pavlos *et al.*, 2014)
DILI	Flucloxacillin	HLA-B*57:01	(Kaplowitz, 2004; Daly *et al.*, 2009; Daly, 2010)
	Amoxicillin-clavulanic acid	HLA-A*02:01-HLA-DRB1*15:01-HLA-DQB1*06:02 haplotype HLA-A*30:02 HLA-B*18:01	(Hautekeete *et al.*, 1999; Kaplowitz, 2004; Daly *et al.*, 2009; Daly, 2010)
	Lapatinib	HLA-DRB1*07:01 HLA-DQA1*02:01	(Kaplowitz, 2004; Spraggs *et al.*, 2011; Schaid *et al.*, 2014)
AHS	Abacavir	HLA-B*57:01	(Mallal *et al.*, 2008; Illing *et al.*, 2012; Norcross *et al.*, 2012)

The strong association between abacavir and HLA-B*57:01 is rare in that the positive and negative predictive values are so high. Take this in comparison to flucloxacillin (discussed in detail at the end of this chapter) with the same HLA association, however only 8 in 100,000 develop a hypersensitive response. While HLA-B*57:01 increases the incidence of

flucloxacillin mediated liver damage by 30-60 times, its predisposition is still not as strong (Daly *et al.*, 2009).

Both class I and class II MHC proteins present peptides to CD8+ and CD4+ T cells respectively, leading to their clonal expansion, activation and maturation (Soen *et al.*, 2003). The investigation and the characterization of the peptides presented by HLA molecules to T lymphocyte and other immune cells can give an understanding to the mechanisms underpinning disease associated with HLA polymorphisms, thus leading to better treatments. To characterize these precise peptide signatures, mass spectrometric techniques are often utilized. The details of the mass spectrometric analysis of MHC peptides will be explored in detail in chapters 4 and 5.

1.2.5 IMMUNE MEDIATED DRUG INDUCED LIVER INJURY

Stress or damage to hepatocytes results in the release of signals that stimulate the activation of both innate and adaptive immune responses. Kuppfer cells, NK cells and NKT cells are particularly active in cases of innate immune responses to DILI (Holt and Ju, 2006). Individual susceptibility to DILI is contributed to by the balance between cytokine release relating to DAMPs and hepatoprotective signals. The role of the immune system in DILI is mediated through the culmination of complex interactions between the innate and the adaptive response. A number of cells are involved in the progression of hepatotoxicity while at the same time have been shown to induce hepatoprotective mechanisms. These confounding results makes it difficult to accurately predict the precise onset of DILI.

Innate immunity plays an important role in responding to drug induced stress. Limited animal models exist for DILI. However, APAP overdose has furthered our knowledge of the underpinning mechanisms. NAPQI induced hepatocyte damage leads to the activation of the innate immune system leading to the inflammatory response and the infiltration of circulating immune cells into the liver. Once cells are activated the release of cytokines,

chemokines and reactive oxygen/nitrogen species progress the manifestation of disease (Holt and Ju, 2006). IFN-γ, Fas and Fas-ligand are all known to directly cause liver damage through knock-out mice becoming resistant to APAP toxicity (Ishida *et al.*, 2002; Liu, Govindarajan and Kaplowitz, 2004). Other studies have shown that the release of IL-10, IL-6 and COX-2 cytokines can have a protective role, with increases in mRNA expression being observed during APAP toxicity. Mutant mouse models deficient in IL-10 and IL-6 were found to have an increased susceptibility to APAP DILI, leading to their hepatoprotective role being confirmed (Reilly *et al.*, 2001; Bourdi *et al.*, 2002; Masubuchi *et al.*, 2003). Kupffer cells responding to hepatocellular damage have been shown to secrete TNF-α, resulting in direct tissue damage through Bim/Bmf caspase pathways leading to apoptosis (Bleibel *et al.*, 2007). Additionally, IL-12 and IL-18 are secreted activating NK and NKT cells. On the other hand, as with many immune cells involved in DILI, Kupffer cells can also play a protective role secreting IL-10 and IL-6 thus counteracting inflammation and promoting regeneration (Ju *et al.*, 2002; Holt and Ju, 2006).

NK, NKT, circulating macrophages and neutrophils are all involved in the innate immune response during DILI. Depletion of NK and NKT cells results in reduced interferon gamma (INF-γ) release leading to reduced direct hepatotoxicity. It has also been shown that aside from reducing other cytokine release the depletion of NK and NKT cells results in a decrease in neutrophil accumulation (Liu, Govindarajan and Kaplowitz, 2004; Holt and Ju, 2006), suggesting they have a role in the initial inflammatory response. The specific role of neutrophils in DILI remains to be fully elucidated, however, their accumulation has been characterized in a number of studies (Ishida *et al.*, 2002; Liu, Govindarajan and Kaplowitz, 2004; Holt and Ju, 2006). Studies in the rat have shown that anti-neutrophil antibodies can induce protective mechanisms, reducing ALT levels and histological improvements, however, whether their primary role is in the clearance of cellular debris or hepatotoxicity is still under investigation (Smith *et al.*, 1998).

The adaptive immune response in DILI is more closely related to idiosyncratic reactions. There is a multitude of clinical features which implicate the adaptive immune system in the onset of disease from certain drugs. The occurrence of a rash, fever and eosinophilia often with delayed onset are indicative of a learnt response arising from exposure to a drug. Often in these cases re-challenge with the offending drug results in a much more serious reaction with fast onset, showing immunological memory exists. Further evidence for adaptive involvement is through the detection of antibodies isolated from DILI patients specific to drug-modified hepatic proteins. There are a number of drugs that contribute to these mechanisms, including diclofenac, carbamazepine and halothane (Kenna, 1997; Zimmerman, 2000; Reilly *et al.*, 2001).

The adaptive response in the liver is unique in that it will usually promote tolerance rather than activation (Figure 1.3). This is in part why immune mediated liver toxicity is relatively rare, making patient susceptibility key. Unless these barriers for tolerance are crossed, an adaptive immune response is unlikely (Holt and Ju, 2006). In the case of halothane induced iDILI treatment of mice with polyI:C, a molecule of microbial mimicry, resulted in an increase level of toxicity (Reilly *et al.*, 2001). This has also been shown to be the case with a number of other drugs when administered with the addition of lipopolysaccharide (Ju and Reilly, 2012). In using pathogen associated molecules (adjuvants) it appears to be possible to overcome the toloregenic barriers of the adaptive immune system, resulting in an increased response to drugs that otherwise would be tolerated. However, the lack of genuine iDILI animal models makes this difficult to interrogate for a number of drugs.

Typically, the involvement of the adaptive immune system in DILI follows the DHR hypotheses; the hapten or PI hypothesis. Ultimately both are dependent on the interaction between the drug/RM, MHC and TCR (Holt and Ju, 2006). While generally regarded as dose independent it has become apparent in recent years that the onset of iDILI is often increased through pharmaceuticals taken at doses >50 mg/day. Whether this increased incidence of

iDILI among doses above 50 mg/day is due to a dose relationship or due to more frequent use of medications with doses above this threshold is yet to be fully understood (Stephens, Andrade and Lucena, 2014). That said, even if there is a dose dependent onset it is thought that this too varies among individuals based upon their specific exposures/circumstances (Han *et al.*, 2013).

1.2.5.1 DRUGS ASSOCIATED WITH DILI

DILI is one of the main reasons for the failure of drugs making it to market, or subsequently being removed from the market. The rare incidence of DILI makes adverse reactions difficult to pick up until available to larger number of people (Zimmerman, 2000; Pessayre and Larrey, 2008). The U.S. Food and Drug Administration (FDA) have removed several drugs due to liver toxicity including, but not limited to, bromfenac, troglitazone and ebrotidine. Other drugs suspected to cause DILI, such as nefazodone and trovafloxacin are available, however come with a black box warning (Holt and Ju, 2006; Stephens, Andrade and Lucena, 2014).

1.2.5.1.1 AMOXICILLIN-CLAVULANATE

One leading cause of DILI, after APAP, is the co-treatment of the β-lactam antibiotics amoxicillin and clavulanic acid. The presentation of amoxicillin-clavulanate (AC) DILI can be mixed, but is more commonly associated with cholestasis and has in rare cases resulted in vanishing bile duct syndrome. It has been reported that the incidence of AC-DILI can be as high as 43 in 100,000; representing a large number of potential patients when considering its widespread use. While most recover fully there are cases where liver transplant may be required otherwise there is a significant chance of mortality. AC-DILI is more common among men and those over the age of 55, as is commonly associated with the onset of cholestasis (deLemos *et al.*, 2016). The incidence of AC-DILI is increased by the addition of clavulanic acid due to reporting of amoxicillin single treatment resulting in DILI being low. Genetic predisposition to AC-DILI has been associated with carriage of the HLA-DRB1*15:01, HLA-DRB5*01:01, HLA-DQB1*06:02 haplotype. Interestingly the HLA-DRB1*15:01-DQB*06:02

HLA alleles have been linked to cholestatic injury whereas HLA-A*30:02 is linked to hepatocellular DILI (Kim *et al.*, 2015).

T lymphocytes are implicated in the onset of disease and have been the focus of several studies. It has been possible to generate both CD4+ and CD8+ T cell clones from amoxicillin and clavulanic acid independently. In the generation of CD4+ and CD8+ clones the panel of cytokines secreted were assessed. It was apparent that the clones responding to amoxicillin were able to secrete a panel of cytotoxic cytokines including IFN-γ and IL-22. The cytokine secretion from clavulanic acid clones however were more restricted. It was shown that the clones were likely activated by a hapten mechanism through several *in vitro* tests. Firstly, antigen presenting cells were required for optimal activation and responses were not abrogated when APCs were pulsed with amoxicillin, meaning internal modification of drug/drug modified proteins resulted in antigen processing and presentation. When antigen processing was knocked out, responses were not longer observed (Kim *et al.*, 2015).

1.2.5.1.2 NEVIRAPINE

HLA-B*35, HLA-C*04 and HLA-DRB1*01:01 are all implicated in the onset of nevirapine DILI in a number of populations. However, to date the association is not strong enough to accurately describe predicative values (Cornejo Castro *et al.*, 2015). Nevirapine DILI is associated with HIV patients, however, it is difficult to determine the attributing factors of HIV as nevirapine is a HIV treatment. Symptoms include rash, skin blistering and hepatotoxicity. In a recent study of 151 nevirapine-hypersensitive patients (and 182 tolerant controls) the incidence of HLA-C*04:01 resulted in an increased risk of developing SJS/TEN. While this does not explain the incidence resulting in hepatotoxicity it does further our understanding of immune involvement in nevirapine induced hypersensitivity. Other genes coding for the ERAP proteins involved in the loading of antigens into MHC have been identified to offer a protective phenotype in nevirapine DHRs, however the full mechanisms are yet to be understood (Carr *et al.*, 2017).

1.3 PROTEOMIC METHODS TO DETECT DRUG-PROTEIN ADDUCTS

Multiple methods are used to detect drug-protein adducts. The most suitable method is dependent on type of sample, the required specificity and sensitivity, and the depth of analysis needed. While mass spectromety (MS) based techniques now dominate drug-protein adduct analysis, traditional methods of protein purification are still essential.

1.3.1 SAMPLE ISOLATION

Biological samples such as blood serum contain an extremely complex mixture of in excess of 10,000 proteins with abundances spanning 9 orders of magnitude (Adkins, 2002). This makes it a challenge to detect proteins that have or may be targeted by drugs. Highly abundant proteins such as HSA account for 80% of the total protein concentration in serum (Steel *et al.*, 2003). HSA is often targeted by drugs (and/or their metabolites) both due to its high abundance and its ability to non-specifically and reversibly bind to many lipophilic organic compounds (Elsadek and Kratz, 2012; Zheng *et al.*, 2014). Additionally, HSA is particularly susceptible to modification by β-lactam antibiotics, which bind covalently to available lysine residues (Levine and Ovary, 1961; Batchelor, Dewdney and Gazzard, 1965; Meng *et al.*, 2011; Jenkins *et al.*, 2013; Garzon *et al.*, 2014). Isolating individual protein adducts for mass spectrometric analysis can be achieved using both high performance liquid chromatography (HPLC) methods (protein depletion and sample fractionation) and SDS-PAGE. Affinity antibody/ligand modified resin is a good tool to deplete target proteins and is particular useful for these complex biological samples (Fang and Zhang, 2008; Jenkins *et al.*, 2009). Other HPLC methods include the use of resins with certain chemical properties that can fractionate samples based on their hydrophobicity, for example. Again, this pre-fractionation is mainly required for complex samples.

The *in vitro* incubation of drugs (metabolites) with model proteins such as HSA and glutathione S-transferase pi (GSTP) has given insight into the chemical mechanisms of drug-

protein adduct formation. For example, the chemical nature, the level of modification and the stability of adducts can be investigated by altering incubation conditions such as buffer pH, drug concentrations and incubation conditions. Model compounds give the ability to generate sufficient quantities of drug-protein adducts to determine the structures of potential adducts and possible chemical pathways for the adducts which might be formed *in vivo*. Such *in vitro* studies are required for the development of methods with sufficient sensitivity for analysis of drug-protein adducts formed in low abundance from biological samples (Jenkins *et al.*, 2009, 2013; El-Ghaiesh *et al.*, 2012). While *in vitro* drug protein incubations enable us to identify binding sites it is important that any free drug is removed prior to further processing to prevent adduct formation on normally inaccessible protein locations.

1.3.2 DRUG-PROTEIN ADDUCT DETECTION USING ANTIBODIES

Antibodies specific to particular molecular targets can be used to detect drug-protein adducts through a variety of techniques, including immunoblots and ELISAs. While these are very coarse methods of hapten detection they represent a simple way to quickly detect proteins of interest. In addition, methods used to separate complex mixtures, through 1D and 2D SDS-PAGE and HPLC, can be used in conjunction with antibodies to identify drug-protein adducts. Immunoblots are highly qualitative and are best used to visually display a set of data. In contrast, absorbance readings from ELISA samples provide quantitative measurement that can be statistically analyzed and validated.

Drug specific antibodies are used for the detection of drug modified proteins. While this practice is widely accepted in the field, there are limitations. As immunoblotting and ELISAs rely on antibodies that recognize a specific epitope, cross reactivity can be problematic. Figure 1.12 demonstrates how chemical moieties are conserved within drug families, such as the thiazolidine ring in β-lactam antibiotics (McVey *et al.*, 2001; El-Ghaiesh *et al.*, 2012)

Although the thiazolidine ring is conserved across the family of β-lactam antibiotics, penicillin antibodies are not cross reactive with all members. Commercially available antibodies raised to benzyl penicillin (BP) that recognizes the thiazolidine ring cross react with other β-lactams such as amoxicillin, piperacillin and penicillin V but not with flucloxacillin, oxacillin, cloxacillin and dicloxacillin. Although these all possess a thiazolidine ring, molecular stereochemistry prevents recognition by BP specific antibodies.

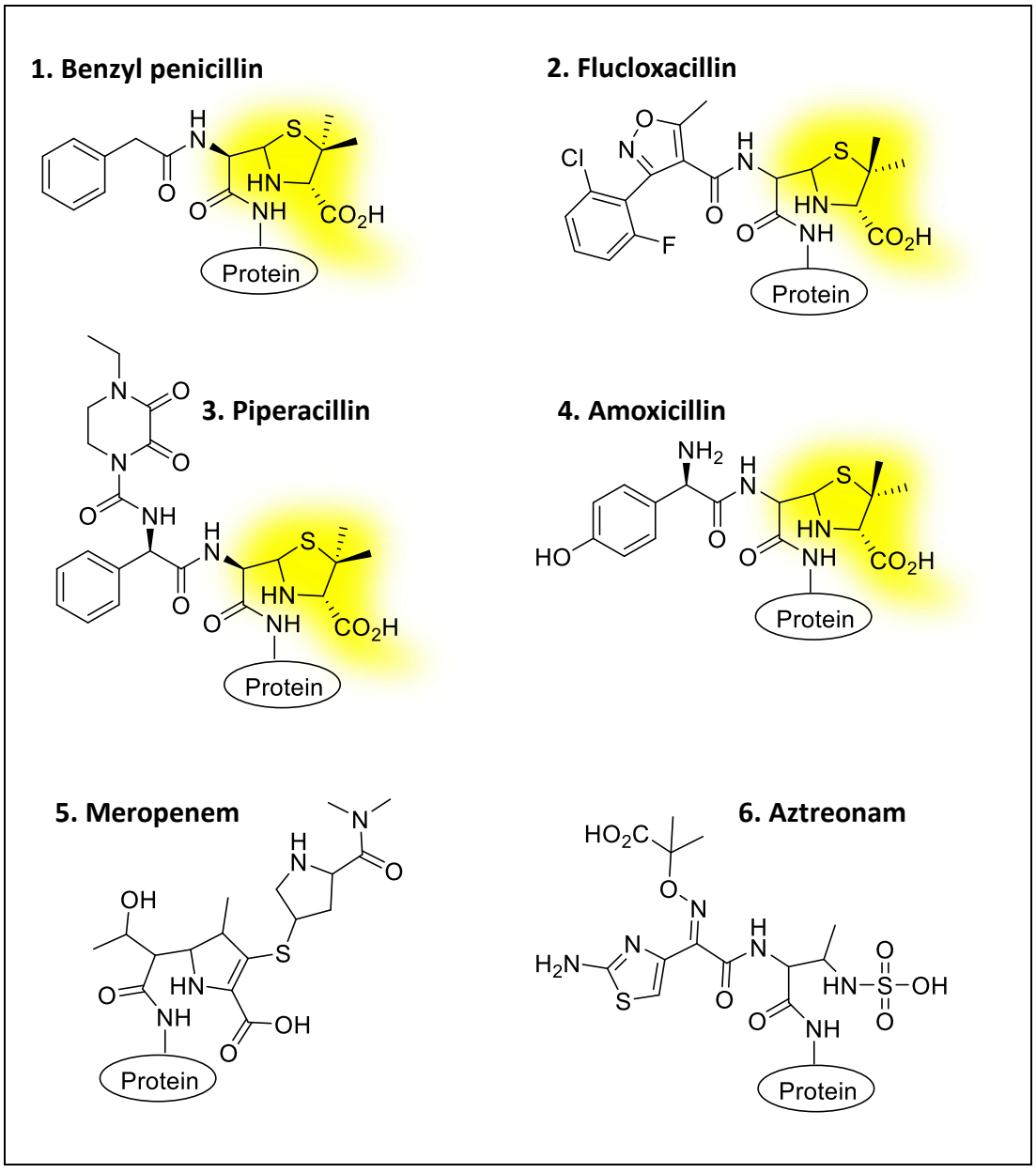


Figure 1.12. Haptenic structures of β-lactam-protein complex. Nucleophilic addition to the β-lactam ring of 1) BP, 2) flucloxacillin, 3) piperacillin, 4) amoxicillin, 5) meropenem & 6) aztreonam by lysine residues on proteins results in ring opening and formation of a stable amide adduct

1.3.3 MASS SPECTROMETRY BASED PROTEOMICS FOR THE ANALYSIS OF DRUG-PROTEIN ADDUCTS

While all the aforementioned techniques have the ability to detect drug-protein adducts, more sophisticated tools are required to characterize the precise chemical nature and for the identification of novel drug-protein adducts. MS is an exceptionally powerful tool used to elucidate structures of both known and unknown adducts (Domon and Aebersold, 2006). Through peptide sequencing with MS techniques it is possible to determine the precise amino acid modification in the protein structure. In the analysis of peptides using MS following ionization, mass-to-charge ratios (*m/z*) of peptides are calculated. Tandem MS utilizes a second fragmentation whereby nitrogen bombards the peptide inducing fragmentation. This process, termed collision induced dissociation, allows further characterization and is the tool of choice when sequencing peptides derived from proteins (Angel *et al.*, 2012). Due to this second fragmentation step occurring within tandem MS it is usually referred to as MS/MS. MS platforms have also been developed to perform differently based on sample type and expected results. A 'trade-off' between sensitivity and mass accuracy can be observed between different MS platforms (Domon and Aebersold, 2006) where depending on the sample mixture one may be favored more than the other. This section will outline the main platforms of MS analysis used for the detection and quantification of drug-protein adducts and their limitations for particular sample types.

1.3.3.1 TOP-DOWN MASS SPECTROMETRIC ANALYSIS

As protein targets for many drugs (metabolites) in complex biological samples (e.g. blood plasma) remain unknown, less targeted MS approaches are required for adduct discovery. Top-down MS works by introducing whole proteins into the mass spectrometer where the MW of an intact protein is first identified (Figure 1.13) (Bogdanov and Smith, 2005). Further fragmentation by tandem MS (the "down" part) allows identification of proteins with full sequence coverage. Since top-down MS analyses intact proteins without proteolytic

digestion, this strategy is extremely useful for mapping of labile drug-protein adducts that may be lost during digestion (Zhang and Ge, 2011).

Top-down MS is generally performed in high resolution instruments because of the need to resolve the high MW of intact proteins (Zhang and Ge, 2011). Historically, Fourier transform ion cyclotron resonance mass spectrometers have been mostly used for top-down MS analysis. More recently, new instruments with a high resolving power that facilitate tandem MS experiments have been developed (Siuti and Kelleher, 2007). In proteomics, thousands of highly accurate spectra can be produced relatively quickly. Consequently, data can be matched to protein databases in order to identify proteins present in the sample.

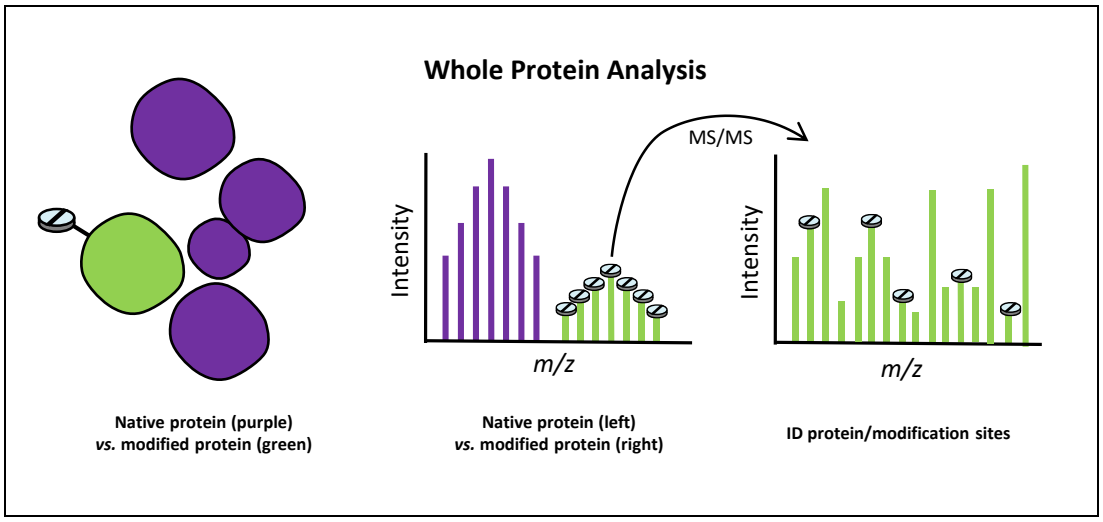


Figure 1.13. Top-down MS method for the characterisation of drug protein adducts. Whole protein is directly analyzed by MS without digestion. Further MS/MS fragmentation enables the localisation of drug-binding amino acid residues.

Top-down MS is used for the detection of drug modified adducts from complex sample mixtures (Figure 1.13). Initially, the protein of interest will be isolated from the sample. This can be achieved using HPLC, either through protein depletion, enrichment or fractionation. While protein enrichment using affinity capture is generally more preferred, this is limited by the availability of antibodies/ligands to proteins of interest. Where antibodies/ligands are unavailable, top-down mass spectrometric methods are not favoured due to the complexity of the data acquired. This is particularly important with the drug modification of proteins, as

they usually appear in lower abundance to the native protein. Where more abundant proteins exist, particularly in complex mixtures, repeated selection and fragmentation can result in a loss in dynamic range (Durbin *et al.*, 2016). Where enrichment can be achieved, on analysis drug-protein adducts will gain a mass addition, depending on the level of modification, resulting in a mass-shift that would be identifiable on the mass spectra produced. The incorporation of novel MS fragmentation techniques such as electron capture dissociation, has also allowed for the identification of modification sites (Zubarev *et al.*, 2000; Siuti and Kelleher, 2007). Although top-down MS has proven to be a powerful technique for the identification of the sites of protein modification and determination of the order of multiple modification, several technical challenges such as protein solubility and sensitivity are yet to be resolved (Zhang and Ge, 2011). An alternative bottom-up MS approach, involving proteolytic digestion of proteins into small peptides prior to MS analysis, has been widely used for identification and quantification of drug-protein adducts.

1.3.3.2 BOTTOM-UP MASS SPECTROMETRIC ANALYSIS

The digestion of proteins into polypeptides and analysis by MS is known as a 'bottom-up' approach. In this method samples undergo reduction and alkylation, using chemicals such as dithiothreitol (DTT) and iodoacetamide (IAA) respectively, to unfold the tertiary protein structure and prevent refolding. Samples are subsequently digested using proteolytic enzymes (Figure 1.14) (Kelleher *et al.*, 1999; Helm *et al.*, 2014). Ultra-high resolution instruments, such as the Thermo Orbitrap mass spectrometers, can offer a resolving power of 240,000 (Michalski *et al.*, 2012). Alternatively, Sciex QTOF (quadrupole time-of-flight) MS/MS platforms can offer faster data acquisition, however, their resolving power is limited to 40,000. This makes it important to choose the correct platform depending on the sample being analysed (Yamada *et al.*, 2002). Similar to top-down analysis, proteins must first be extracted from the sample mixture using the same methods. Alternatively, in bottom-up analysis proteins of interest can be isolated by SDS-PAGE followed by in-gel digestion of the

bands of interest. It is advantageous for samples to undergo various clean-up steps, such as desalting, concentration and purification in order to enhance the sensitivity of detection (Bagshaw, Callahan and Mahuran, 2000; Liang *et al.*, 2000). Samples are then loaded into the mass spectrometer as a mixture of digested peptides.

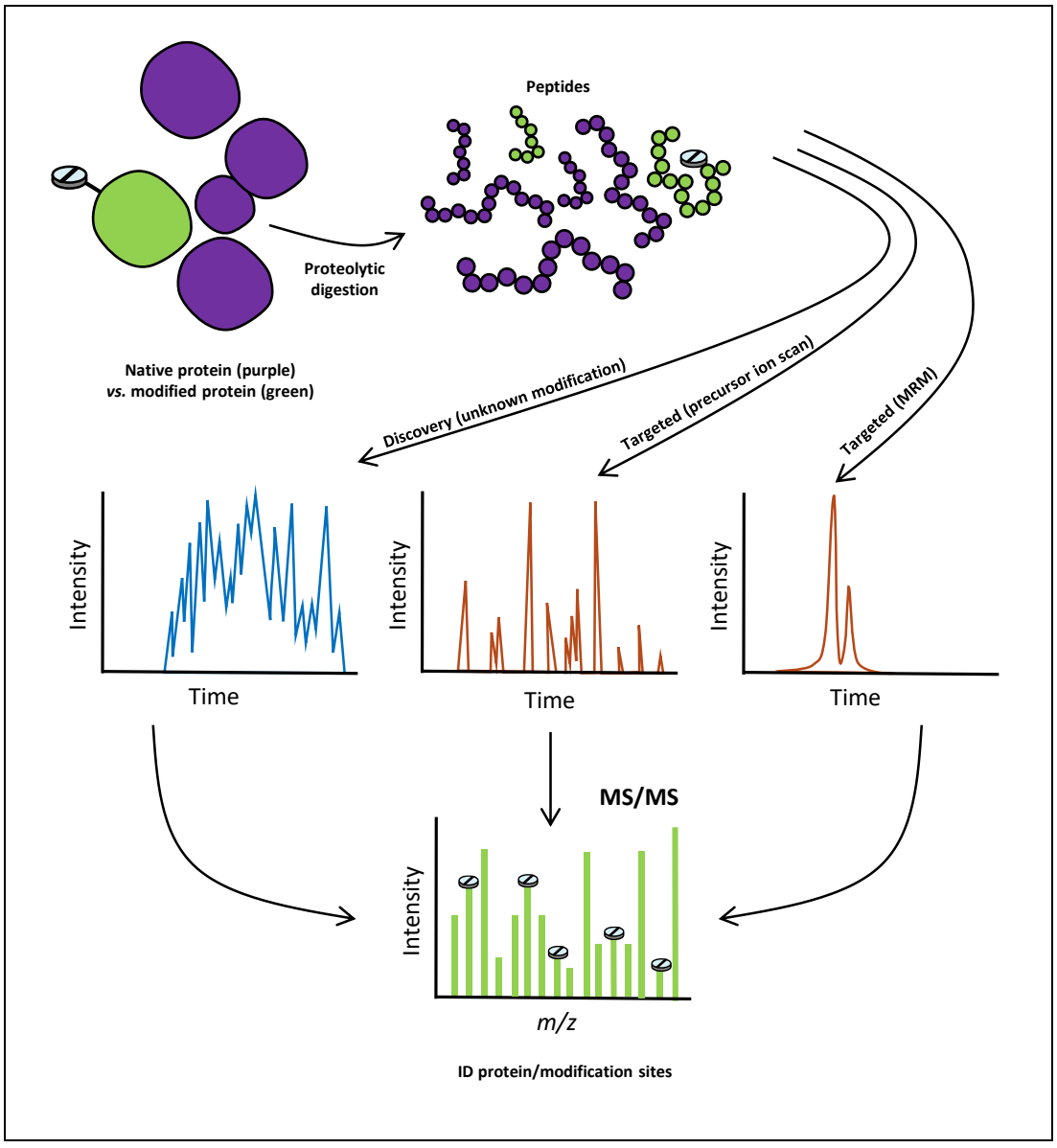


Figure 1.14. Bottom-up MS method for the characterisation of drug protein adducts. Proteins are enzymatically digested into peptides that can be analyzed by various MS techniques: Discovery MS/MS analysis allows for identification of unknown adducts; more targeted analyses (precursor ion scanning and multiple reaction monitoring scanning) enable the detection of known adducts with great sensitivity.

In the discovery stage, non-targetted LC-MS/MS is a powerful platform used for the detection and analysis of unknown drug-protein adducts. LC-MS/MS ion scanning is a non-

targeted approach where the initial ion (precursor ion) is detected in the first ionization stage, and then further fragmented to characterize the full peptide sequence. Although this is a non-specific scanning technique, modern QTOF platforms can achieve up to 100 MS/MS spectra per cycle, relaying masses of sample data for analysis. Similarly, precursor ion scanning can be used to identify suspected modifications in an unknown sample (Figure 1.14). Here fragmentation is induced and the instrument is set to scan for a specific fragment ion. When the fragment ion is detected the mass spectrometer is triggered to acquire all MS/MS spectra. For example, a characteristic peak at *m/z* 160 corresponding to the cleavage of the thiazolidine ring can be used for identifying adducts formed by β-lactam antibiotics (Jenkins *et al.*, 2009; Meng *et al.*, 2011). This type of experiment is extremely useful for monitoring unknown protein targets for known adducts within a complex mixture. While this is a partially targeted approach to drug-protein adduct detection, it can give rise to false positives. Ions that are not derived as a result of drug cleavage may correspond to the same mass, therefore it is important to manually characterize the amino acid sequence on any spectra where drug modification is suspected. Putative drug modified peptide sequences can be matched against peptide databases with the hope of identifying protein origins. Although in theory this is a promising method to identify novel drug-protein adducts, in practice it is laborious and takes a trained eye to calculate peptide sequences.

MRM is a more targeted MS tool enabling the user to scan for specific parent and fragment ions. MRM transitions specific for drug modified peptides are generally designed for this purpose. The parent ions can be calculated for all possible peptides derived from theoretical digests with a mass addition of proposed haptens. The parent ion masses are then paired with the proposed fragment masses derived from drug molecules. MRM-MS/MS actively searches for pre-defined *m/z* ratios for precursor and fragment ions and a full scan MS/MS is only triggered for the ions that match the search criteria (de Hoffmann, 1996). Ions which

do not correspond to pre-defined *m/z* will not be acquired, making analysis clearer and more specific (Figure 1.14) (Kitteringham *et al.*, 2009; Gillette and Carr, 2013).

MRM is particularly useful in the identification of proteins modified by β-lactam antibiotics. Protein digestion using known enzymes, such as trypsin, enables theoretical digests and expected *m/z* ratios of peptides to be calculated. β-lactam antibiotics covalently bind to exposed NH_2 groups on the side chains of lysine residues (Mauri-Hellweg *et al.*, 1996). When a drug is covalently bound, proteolytic digestion of trypsin cleaving the C-terminal side of lysine and arginine is interrupted (Olsen, Ong and Mann, 2004). This is due to the chemical morphology at a β-lactam antibiotic bound lysine residue blocking chemical interactions, thus incorporating missed cleavage sites (Thiede *et al.*, 2000; Jenkins *et al.*, 2009). Missed lysine cleavages are therefore included when performing theoretical digests. This enables the specific selection of peptides with the potential to be modified at a non-terminal lysine residue. To specifically search for drug modified peptides the *m/z* ratios are calculated with a mass addition corresponding to the MW of haptens. When analyzing the spectra, characteristic features of the drug molecule can be identified, for example, the cleavage of thiazolidine ring of penicillin leads to the fragment ion at m/z 160 (Jenkins *et al.*, 2009; Meng *et al.*, 2011).

Sample preparation is important in being able to successfully identify drug-protein adducts, especially when they are present in low abundance. Therefore, a three-dimensional approach to sample preparation can be performed. Proteins of interest can be first isolated from complex biological samples using affinity capture and subjected to proteolytic digestion. After digestion HPLC platforms are used for ion exchange, separating samples into multiple fractions depending on their ionic interactions with the column used. Each fraction will then be purified to remove the salt using C18 reverse phase HPLC, before MS analysis. Sample preparation in this way spreads the peptide mixture across two gradients, enhancing the sensitivity of detection.

1.4 BETA-LACTAMS ANTIBIOTICS

β-Lactam antibiotics are widely prescribed for bacterial infections and are a frequent cause of allergic reactions. β-Lactams are intrinsically reactive and can form drug-protein adducts with lysine residues on protein through the nucleophilic opening of the β-lactam ring. Recent studies investigating blood from patients using advanced MS methods have identified albumin as the major target for β-lactams. The precise haptenic structures and the exact location of modification on albumin have been defined for a number of β-lactams including flucloxacillin, amoxicillin and piperacillin (Jenkins *et al.*, 2009; Meng *et al.*, 2011; Whitaker *et al.*, 2011). At low concentrations, different β-lactams appeared to selectively target different lysine residues in HSA, with non-covalent interactions positioning the drugs in favorable orientations in the protein-binding pocket to facilitate covalent binding. Thus, the three-dimensional shape of the drug, as well as its inherent chemical reactivity, determines the selectivity of covalent binding.

Detecting drug-modified proteins in a complex biological sample is challenging, and often requires enrichment, HPLC separation and novel MS/MS techniques to be developed. A common enrichment method is immunoaffinity purification using either drug-specific or protein-specific antibodies. For example, isolation of HSA using anti-HSA antibodies has allowed identification of many drug-modified albumin present in patient plasma (Hammond *et al.*, 2014; Meng *et al.*, 2014, 2015, 2016; Yip *et al.*, 2017) . In 2009 Jenkins *et al* performed a comprehensive investigation into the *in vitro* binding of flucloxacillin and its metabolite to HSA (Jenkins *et al.*, 2009). Flucloxacillin modifies up to 10 lysine residues in HSA *in vitro*, with Lys190 and Lys212 being most readily observed. Analysis of HSA isolated from the serum of eight flucloxacillin-tolerant patients also showed that modification of Lys190 and Lys212 occurred in all the patients, with up to nine other modified lysine residues being detected. In addition, the detection of HSA adducts derived from its metabolite, 5'-hydroxymethly flucloxacillin (5'OH-flucloxacillin), provides the evidence of local metabolism within

hepatocytes (Jenkins *et al.*, 2009). Incubation of HSA with piperacillin results in the formation of two distinct haptens: a cyclized and a hydrolyzed form in which the dioxopiperazine ring had undergone hydrolysis (Whitaker *et al.*, 2011). Piperacillin was able to modify up to 13 out of 59 lysine residues at high concentrations, with Lys541 being the only modified residue detected at lower concentration. Four of these (Lys190, Lys195, Lys432 and Lys541) were also detected in plasma from 4 piperacillin treated, tolerant patients. Fourteen different modified lysine residues were identified on HSA isolated from patients' serum (Meng *et al.*, 2011). Apart from albumin adducts, amoxicillin-modified transferrin and immunoglobulin were also detected when human serum was incubated with amoxicillin *in vitro*. The hapten was identified as amoxicilloyl on Lys190, Lys199 and Lys541 with higher concentrations eliciting more extensive modification (Ariza *et al.*, 2012).

Unlike the penicillin class of β-lactams, meropenem and aztreonam contain a β-lactam ring fused to another ring that is different from the thiazolidine rings present in penicillins. Thus, meropenem and aztreonam formed complex and structurally distinct haptenic structures with lysine residues on HSA. Each drug modified Lys190, with less common modifications at Lys12, Lys199 and Lys 351 for meropenem and Lys137, Lys432 and Lys541 for aztreonam (Jenkins *et al.*, 2013).

1.4.1 FLUCLOXACILLIN DILI

As previously described in this introduction, flucloxacillin is a β-lactam antibiotic implicated in the onset of immune mediated cholestatic DILI. Flucloxacillin is a broad range synthetic, penicillinase-resistant, isoxazolyl antibiotic primarily used to treat infections of the skin, predominantly those caused by *Staphylococcus spp* (Jenkins *et al.*, 2009). Their clinical effect is attributed to the binding of drug to the penicillin binding protein, a key component in the synthesis of the bacterial cell wall. In binding to its target, flucloxacillin prevents cell wall synthesis and so Gram-positive bacteria are killed (Piddock and Wise, 1986). While *in vitro*

flucloxacillin has bacteriocidal activity against Gram-negative species, other classes of penicills are more frequently used due to higher efficacy and licencing (PubChem CID: 21319). Generally, flucloxacillin is well tolerated at doses up to 4 gram per day, however in a small number of cases, estimated in the UK at 8.8 in 100,000, patients have developed symptoms such as nephritis, rashes, fever, eosinophilia and hepatic cholestasis. Symptoms occur more commonly in elderly patients, females, and in those who are undergoing prolonged periods of treatment. The onset of these reactions is usually delayed type, manifesting between 1 and 45 days after initial exposure, meaning the adaptive immune system is likely involved in the onset of disease. While discontinuation of flucloxacillin assists in the recovery of patients, symptoms can remain for a number of weeks post cessation of the drug. After APAP, flucloxacillin is in some populations considered the second leading cause of drug induced liver failure (Daly *et al.*, 2009; Jenkins *et al.*, 2009).

Flucloxacillin is metabolised to 5'OH-flucloxacillin in the presence of CYP3A4 in the liver. While both flucloxacillin and its metabolite were not found to be toxic to hepatocytes, 5'-hydroxymethylflucloxacillin did induce damage to biliary epithelial cells. Interestingly in cases of flucloxacillin DILI, hepatic cholestasis is involved where disruption in the excretion of bile is observed, with little hepatic necrosis. While this may indicate direct toxicity of flucloxacillin on the biliary cells, skin rash can also present in patients, more common in T cell mediated immunity (Daly *et al.*, 2009; Jenkins *et al.*, 2009). The identification of flucloxacillin specific T cells isolated from the blood of patients with flucloxacillin induced nephritis, as well as the detection of flucloxacillin-specific IgE from allergic patients, confirms immune involvement (Jenkins *et al.*, 2009). As manifestation of flucloxacillin DILI is idiosyncratic in nature, the full mechanistic understanding behind its pathology is yet to be fully elucidated (Kim *et al.*, 2015).

In 2009 a comprehensive GWAS study was performed after reports that amoxicillin clavulanate and ximelagatran, both implicated in DILI, may be related to the expression of

the MHC class II alleles HLA-DRB1*15:01 and HLA-DRB1*07:01, respectively. This large study used 866,399 markers in 51 cases of flucloxacillin DILI with 282 controls, matched for both gender and ancestry. Upon analysis, it revealed links with the genes encoding for the MHC region had the strongest association (P = 8.7 x 10^{-33}) for rs2395029[G], a marker in linkage disequilibrium with HLA-B*57:01. This association was confirmed with MHC typing of tolerant controls, giving an odds ratio (OR) of 80.6 with a probability of 9 x 10^{-19}. In addition to the association with HLA-B*57:01, the gene ST6GAL1 was also present in the flucloxacillin DILI patients who were carriers of the associated HLA allele. ST6GAL1 has been found to be increased during hepatic inflammation, meaning it could determine susceptibility to immunoinflammatory responses (Daly *et al.*, 2009). Since the initial discovery, further GWAS studies have been performed by the same group to increase the numbers in the study cohort. In the second study 197 flucloxacillin patients were compared with 6,835 controls. The same association with HLA-B*57:01 was observed with an odds ratio of 36.62, P= 2.67 x 10^{-97}. While the increased risk of flucloxacillin DILI from the drug reduced from 80.6 times to 36.62 times, the statistical significance was much higher. In addition to HLA-B*57:01, HLA-B*57:03 was also shown to have an association (OR = 79.21, P= 1.2 x 10^{-6}). The effect of these HLA-B alleles on adverse flucloxacillin responses was shown to be down to the effect of val97 within the protein coding region (Nicoletti *et al.*, 2019). While these genetic associations do not fully explain the onset of disease, predisposition to immune related alleles makes it possible to interrogate the effects of these alleles using both proteomic and *in vitro* techniques.

In previous studies, CD8+ T lymphocytes have been isolated from patients with flucloxacillin DILI and have subsequently been activated by the drug *in vitro*. HLA-B*57:01+ healthy volunteers were recruited into the study, where it was possible to characterize CD4+ and CD8+ T cell clones. CD45RA+CD8+ T cells were activated when flucloxacillin was presented alongside DCs, suggesting antigen processing was required. Additionally CDR9+ T cell clones were found to secrete IFN-γ, T_h2 cytokines, perforin, granzyme B and Fas Ligand; all

associated with the cytotoxic immune response. Interestingly flucloxacillin-induced activation was indeed processing dependent, with cross reactivity identified with T cell clones from patients with the closely related HLA-B*58:01 allele. Clones were cross reactive against drugs with a similar chemical structures, such as oxacillin, cloxacillin and dicloxacillin (Monshi *et al.*, 2013). While cross reactivity in this study was observed, HLA-B*57:01 associations with oxacillin, cloxacillin and dicloxacillin were not observed. However, this was only tested on a very small study cohort so the results are to be taken with caution (Nicoletti *et al.*, 2019).

1.5 AIMS

The main aim of this thesis is to improve our understanding of the mechanisms that underpin flucloxacillin related DILI. Genetic susceptibility indicates that specific interactions between HLA-B*57:01 and TCRs lead to the activations of T-lymphocytes resulting in tissue damage. Both signal 1 and 2, the precise peptides invoking T-lymphocyte responses and the triggers recruiting the immune cells to the site of damage respectively, will be explored. In order to answer these questions specific aims, outlines below, were explored.

1. Generate a sensitive flucloxacillin-specific antibody to enable the detection of flucloxacillin protein binding and cellular localization.
2. Implement experimental and analytic workflows to enable the study of MHC derived peptides.
3. Investigate the binding of flucloxacillin to HLA-B*57:01 restricted MHC peptides and interrogate their involvement in functional T cell assays.
4. Characterize the global peptide repertoire of HLA-B*57:01 and identify any changes that occur as a result of flucloxacillin exposure.
5. Improve analysis pipelines to enable the high throughput interpretation of flucloxacillin modified peptides; both tryptic and non-tryptic.

CHAPTER 2 - FLUCLOXACILLIN ANTI-SERA PRODUCTION AND CHARACTERIZATION

2.1 INTRODUCTION

B lymphocytes play a pivotal role in the adaptive immune response through the development of immunological memory and rapid responses upon re-challenge to a specific pathogen. Antibodies, also known as immunoglobulins, are circulating proteins that are produced in response to specific foreign structures. Recognition of antigens by antibodies is a similar concept to that of the TCR and MHC in its specificity, however antibodies have a much wider diversity with stronger binding to their target epitope. Produced exclusively by B lymphocytes, antibodies can exist as both membrane bound and secreted proteins. Cell membrane bound antibodies localised at the surface of naïve B cells act as antigen receptors where upon stimulation leads to terminal differentiation into plasma cells. Plasma cells are responsible for the secretion of antibodies specific to the antigen that triggered cell differentiation. In addition to blood plasma, antibodies are present in mucosal secretions and interstitial fluid (Abbas, Lichtman and Shiv, 2018).

The general function of antibodies is to neutralize microbial toxins, prevent microbial cell entry and trigger effector responses for the elimination of pathogens. In humans antibodies exist as 5 different isoforms; IgA, IgD, IgE, IgG and IgM. Each isoform has a different structure and function, determined by the heavy chain fragment crystallizable (Fc) region, with all playing an important role in the humoral immune response (Figure 2.1). IgA is most frequently observed as a dimer, but can also exist as a monomer and trimer. Acting primarily in mucosal immunity, IgA has a plasma concentration of 3.5 mg/mL with a half-life of approximately 6 days. IgG is the most abundant isoform in plasma at 13.5 mg/mL with a prolonged half-life of 23 days. This monomeric proteins' key role is in opsonisation, complement activation, antibody dependent cell mediated toxicity, neonatal immunity and feedback inhibition of B cells. IgE and IgM are both relatively low in abundance with a plasma concentration of 0.05 and 1.5 mg/mL and a half-life of 2 and 5 days respectively. While IgM is an important component in naïve B cell antigen recognition, IgE is involved in defences

against parasites, particularly helminths, and immediate hypersensitivity reactions. Although IgD exists, this monomeric isomer appears at trace amounts and has the shortest half-life of 2 days, acting only as a B cell antigen receptor (Abbas, Lichtman and Shiv, 2018).

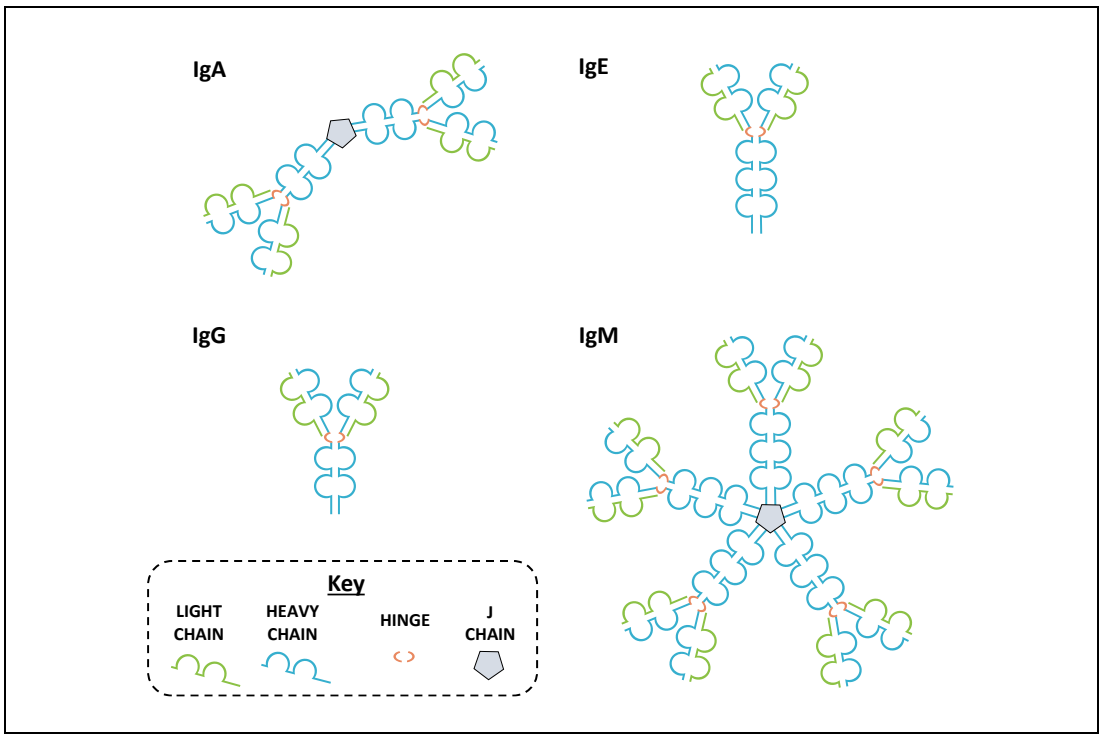


Figure 2.1. Structural differences between the different isoforms of human immunoglobulins. Human antibodies exist as 5 different isoforms, IgA, IgE, IgG, IgM and IgD (structure not shown). IgA primarily exist as dimers joined together by a J chain, however can exist as monomers and trimers. IgE and IgG both exist as monomers, with IgE having a longer Fc region (described in the next section). IgM are pentameric proteins, with 5 immunoglobulins joined with by a J chain.

Among all the different isoforms the general structure of antibodies remains conserved (Figure 2.2). The core structure of the antibody contains one symmetric Fc region and two antigen binding fragments (Fab). Heavy and light chains are made up of both constant and variable regions, the latter allowing for the remarkable degree of antigens that can be recognized. The variable region, also known as the antigen binding site, is specific to the millions of different B lymphocytes circulating in the body. The hinge connecting the Fc to the Fab region is flexible thus allowing the antibody to bind two antigens simultaneously. While the Fab region is responsible for the recognition of a specific epitope, the Fc region

interacts with cell surface Fc receptors enabling activation of immune responses (Abbas, Lichtman and Shiv, 2018).

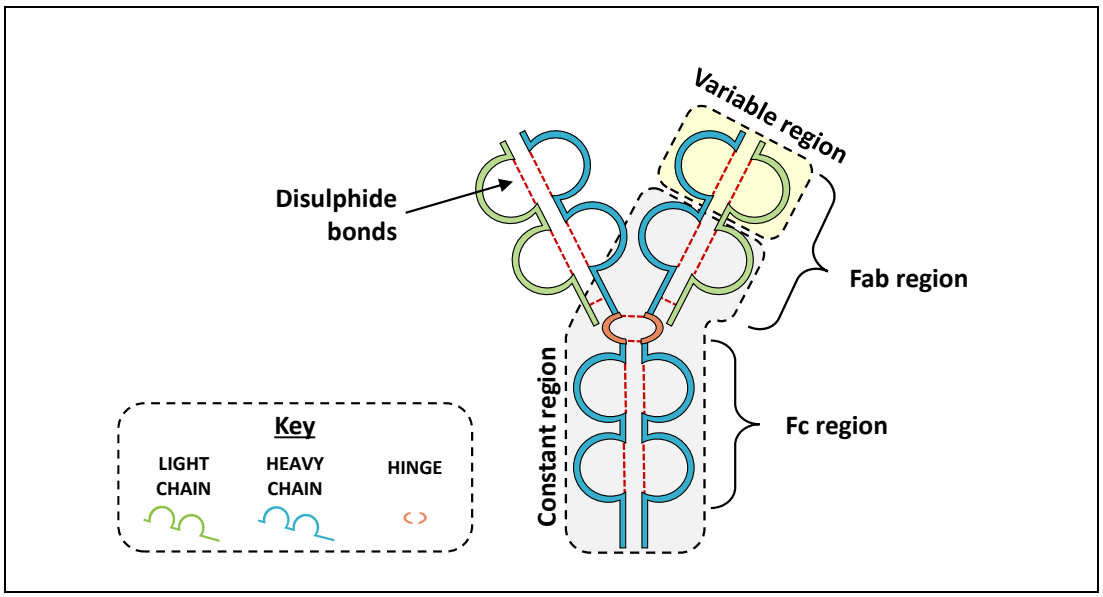


Figure 2.2. The structure of the different isoforms of human antibodies can generally be defined by conserved structures. All antibodies consist of two symmetrical heavy chains, spanning the length of the protein, and two light chains. The Fc region is constant between all of the antibodies within an isoform of a species. The Fab region is responsible for antigen binding through vast differences in the structure of the variable region. The hinge connecting the Fab to the Fc region is flexible to allow binding on multiple epitopes at a time, with disulphide bonds throughout the protein to constrict the structure.

Although antibodies are generally described in the context of pathogenic responses, a role in drug mediated hypersensitivity has often been identified. In addition to the detection of flucloxacillin specific CD8+ T cells from patients with flucloxacillin mediated DILI, the presence of IgE specific to flucloxacillin has been identified in the sera of allergic patients (Baldo, Pham and Weiner, 1995). In the general introduction to this thesis the use of proteomic techniques for the detection of drug modified proteins was briefly discussed. While antibodies are an exceptionally important component of the adaptive immune response, it is possible to utlize them for the discovery of drugs bound to proteins *in vitro* (Jenkins *et al.*, 2009). This offers a biological tool for the detection of antigens in both research and clinical laboratory settings. In order to detect drug haptens using techniques such as immunoblots, immunofluorescence microscopy and enzyme linked

immunoabsorbance assays, the availability of antibodies specific to drug epitopes are required. The general concept behind these methods of detection rely on an immobilized antigen, e.g. drug bound to protein, to be detected by a specific antibody. A secondary antibody conjugated to a chemical signal specific for the Fc region of the primary antibody (Fc portions are generally conserved between species) binds to its target. Upon development of the chemical signal, a detector (e.g. immunofluorescent microscope) records the absorbance of the signal (Figure 2.3). The methods used for the detection of drug modified proteins using antibodies are described in detail in this thesis.

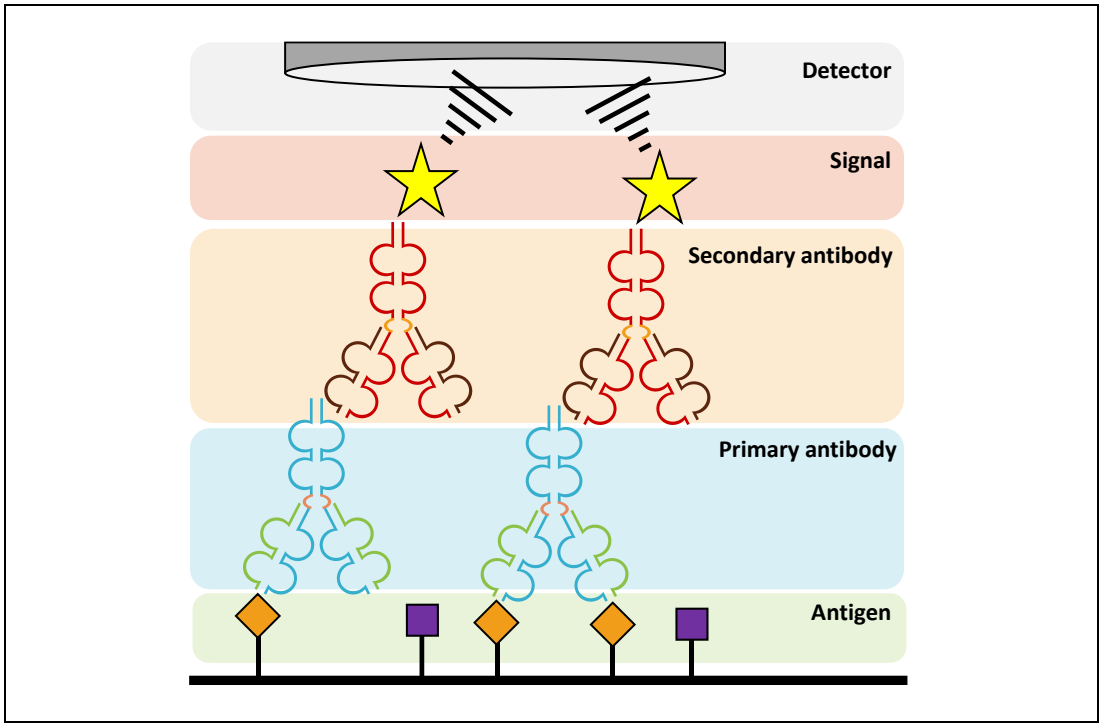


Figure 2.3. General overview of the use of antibodies for laboratory based assays. Antigens bound by primary antibodies are subsequently bound by secondary antibodies. Secondary antibodies are conjugated to a chemical signal that can be developed and detected using a number of techniques.

For several drugs commercially available anti-drug antibodies can be purchased, however it was not possible to find a specific antibody for flucloxacillin. Previous studies have successfully generated antibodies specific to flucloxacillin, and while flucloxacillin modified proteins could be detected there was a degree of cross reactivity between untreated controls (Carey and van Pelt, 2005). Drugs such as flucloxacillin are too small to elicit an immune

response that will result in the production of antibodies, therefore carrier proteins are used for the conjugation of the epitope of interest. A number of different carriers including gelatine, albumin and keyhole limpet hemocyanin (KLH) are commonly used to generate antigens. The choice of carrier protein is dependent on several factors and can vastly impact the quality of the antibody that is generated by the host. Physiochemical interactions between the hapten and carrier and the downstream application of the antibody have great implications on this decision (Fasciglione *et al.*, 1996). For example, the use of flucloxacillin bound to HSA as a carrier protein in a rabbit would elicit an immune response and, potentially, a high titre antibody. However, the polyclonal antibody (contained within the serum of the rabbit) would not be suitable for the detection of flucloxacillin in any human cell/tissue samples due to the cross reactivity with HSA.

2.2 AIMS

In order to characterize the cellular binding of flucloxacillin using proteomic techniques, a high titre antibody was required. Previous studies have used mass spectrometry (MS) to identify the presence of flucloxacillin modified albumin in patients. Due to the complexity and low abundance of flucloxacillin-modified proteins it was not possible to identify the depth and degree of low abundant cellular proteins that are potentially modified. Therefore, in this chapter the following aims were addressed.

1. Identify a suitable carrier protein and host for the generation of a flucloxacillin specific antibody.
2. Optimise the conditions required to achieve maximum hapten density on the carrier protein.
3. Upon receipt of the antibody, determine the specificity and cross reactivity of the flucloxacillin antibody.
4. Optimise the conditions required for use of the antibody using proteomics techniques.
5. Investigate the precise chemical epitope recognized by the antibody.

2.3 METHODS

2.3.1 PROTEIN CONJUGATION

HSA (>97% lyophilized powder, Sigma-Aldrich) and ovalbumin (OVA) (Imject™, Thermo Scientific) were used as protein carriers. Flucloxacillin (Wockharat) and amoxicillin (GSK) were gifted from collaborators. Penicillin V, BP, piperacillin, oxacillin, cloxacillin and dicloxacillin were purchased from Sigma-Aldrich. Drug-protein incubations were performed in phosphate (PO_4) buffer (13.08 mM KH_2PO_4, 62.27 mM K_2HPO_4), pH 7.4 unless specified. Carrier proteins were made up to 1 mM solutions and were incubated at molar ratios to the respective antibiotic. Drug-protein incubations were performed at 37°C for at least 16 hours. Protein purification was performed using methanol precipitation (for subsequent proteomic analysis) or spin filters with a 3 kDa molecular weight cut off (Amicon Ultra, Merck) (for antibody conjugate preparation). Ovalbumin-flucloxacillin incubations used for the production of the antibody immunogen were performed in carbonate or phosphate-carbonate buffer to determine the optimal conditions for drug-protein binding. Carbonate (CO_3) buffer (0.1M Na_2CO_3, 0.1M $NaHCO_3$), pH 11, and PO_4:CO_3 buffer (mixed at a 1:1 volume ratio) were used for drug protein incubations at 50:1, 100:1, 250:1 and 500:1 molar ratios.

2.3.2 PROTEIN QUANTIFICATION

Proteins and drug-modified proteins were quantified using the Bradford assay. A standard curve was produced using 5 µL aliquots of known concentrations of reference proteins (HSA or OVA) and unknown concentrations of samples were mixed with 200 µL of Bradford reagent (Sigma-Aldrich). After 10 minutes the absorbance was read at 405 nm (Dynex Technologies MRXe) using Revelation 4.25 software. Protein concentration was calculated using the formula generated by the line of best fit (R^2 value > 0.95) of the standard curve.

2.3.3 SDS-PAGE

SDS-polyacrylamide gel electrophoresis (SDS-PAGE) was performed for both protein visualisation (coomassie blue staining) and epitope detection (Western/dot blotting). Samples were prepared by addition of 4 x Laemmli sample buffer (25% glycerol (VWR Chemicals), 100 mM Tris (pH 6.8) (National Diagnostics), 2.6% SDS (Sigma-Aldrich), 1.3% bromophenol blue (w/v) (BDH/VWR), 5% mercaptoethanol (Sigma-Aldrich)) to a final concentration of 1 µg/µL and boiled at 100°C for 10 minutes. A 10% polyacrylamide gel was made following manufacturers protocol (32.6% ProtoGel (National Diagnostics), 25.7% ProtoGel resolving buffer (National Diagnostics), 0.1% N,N,N',N'-Tetramethylethylenediamine (TEMED) (Sigma-Aldrich) and 0.1% ammonium persulfate (APS) (Sigma-Aldrich) in deionised water). Unless otherwise stated, 10 µg of protein was loaded onto the gel and was subsequently subjected to electrophoresis at 30 mA until the sample had run the length of the gel. Seeblue plus 2 (Invitrogen, Thermo Fisher) protein marker was used as a MW ladder.

2.3.3.1 WESTERN BLOT ANALYSIS

SDS-polyacrylamide gels were transferred onto 0.45 µM nitrocellulose membrane (Amersham) by electroblotting at 250 mA for 1 hour. Nitrocellulose membrane was washed with deionised water and blocked with Tris/saline/Tween (TST) buffer (150 mM NaCl, 10 mM Tris-HCL, 0.05% Tween 20, pH 8.0) containing 10% nonfat dry milk (Bio-rad). The membrane was incubated with primary antibody (varying concentrations) in 10% nonfat dry milk in TST for a minimum of 16h at 4°C. Four 4 minute washes in TST were performed and horseradish peroxidase anti-rabbit secondary antibody (Dako) was added in 10% non-fat dry milk TST buffer at a 1:1,000 dilution. After 1 hour a further 4 x 4 minute TST washes was followed by signal detection by enhanced chemiluminescence (Western Lighting, Perkin Elmer) on autoradiography film (Amersham).

2.3.3.2 COOMASSIE BLUE STAINING

SDS-polyacrylamide gels were fixed for 1 hour in 40% methanol (Fisher Scientific) and 7 % acetic acid (Fisher Scientific). Coomassie stain was added for a minimum of 1 hour, consisting of 0.08% coomassie brilliant blue (VWR) in 1.6% phosphoric acid (Riedel-de Haen), 8% ammonium sulphate (Sigma-Aldrich) and 10% methanol. SDS-polyacrylamide gels were destained with 10% acetic acid in 25% methanol for 60 seconds before storage in 25% methanol.

2.3.4 MASS SPECTROMETRY

Prior to mass spectrometric analysis samples were reduced with 10 mM DTT for 20 minutes and alkylated with 55 mM IAA for a further 20 minutes. Digestion of drug-protein conjugates was performed using sequencing grade modified trypsin (Promega) overnight at 37°C. Samples were purified using C18 ZipTips (Millipore) and dried in a centrifugal concentrator (Eppendorf speedvac). Samples were reconstituted in 2% acetonitrile (ACN) 0.1% formic acid (FA) (v/v). Identification of flucloxacillin modified OVA peptides was performed by delivering samples onto a QTrap 6500 hybrid quadrupole-linear ion trap mass spectrometer (Sciex) by an Ultimate 3000 HPLC (Dionex) and auto sampler. Samples were injected onto a nanoACQUITY UPLC Symmetry C18 Trap Column, 100Å, 5 µm internal diameter, 180 µm x 20 mm (Waters) in 2% ACN/0.1% FA. Samples were subsequently fractionated by applying a mobile phase gradient (Table 2.1) in line with the trap column onto the analytical column (ACQUITY UPLC Peptide BEH C18 nanoACQUITY column, 130Å, 1.7 µm internal diameter, 75 µm x 100 mm (Waters)) at 300 nL/min. The analytical column was maintained at 45°C throughout the duration of the mass spectrometric acquisition.

Table 2.1. HPLC gradient table. Mobile phase A (0.1% FA in LC-MS H_2O) and mobile phase B (0.1% FA in ACN).

Time	Mobile phase A	Mobile phase B
00:30	95%	5%
05:00	95%	5%
45:00	50%	50%
45:06	1%	99%
65:00	1%	99%
65:06	95%	5%
80:00	95%	5%

Samples were introduced into the mass spectrometer via a 10-µm inner diameter PicoTip (New Objective) using the following conditions. MS 1 ions ranged from a charge state of 2+ to 5+ and included unknown charge states. The intensity threshold was set to 300 with an exclusion time window of 20 seconds for ions previously detected. A mass tolerance of 1,000 mDa was applied. MRM transitions specific for ovalbumin-flucloxacillin modified peptides (Table 2.2) were selected based on a number of criteria.

Table 2.2. MRM transitions of flucloxacillin modified OVA tryptic peptides.

Q1 (*m/z*)	Q3 (*m/z*)	Peptide ID and charge	DP (volts)	EP (volts)	CXP (volts)	CE (volts)
990.77	160.1	ADHPFLFCIKHIATNAVLFFGR 3+	70	10	10	53.54
743.33	160.1	ADHPFLFCIKHIATNAVLFFGR 4+	70	10	10	37.94
594.86	160.1	ADHPFLFCIKHIATNAVLFFGR 5+	70	10	10	30.96
670.24	160.1	AFKDEDTQAMPFR 3+	70	10	10	37.51
502.93	160.1	AFKDEDTQAMPFR 4+	70	10	10	26.64
402.54	160.1	AFKDEDTQAMPFR 5+	70	10	10	21.92
950.68	160.1	ELKVHHANENIFYCPIAIMS ALAMVYLGAK 4+	70	10	10	47.68
760.74	160.1	ELKVHHANENIFYCPIAIMS ALAMVYLGAK 5+	70	10	10	38.75
805.08	160.1	FDKLPGFGDSIEAQCGTSVN VHSSLR 4+	70	10	10	40.84
644.27	160.1	FDKLPGFGDSIEAQCGTSVN VHSSLR 5+	70	10	10	33.28
716.27	160.1	GLWEKAFK 2+	70	10	10	45.11
477.84	160.1	GLWEKAFK 3+	70	10	10	27.89
844.66	160.1	GSIGAASMEFCFDVFKELK 3+	70	10	10	46.23
633.75	160.1	GSIGAASMEFCFDVFKELK 4+	70	10	10	32.79
507.2	160.1	GSIGAASMEFCFDVFKELK 5+	70	10	10	26.84
671.28	160.1	IKVYLPR 2+	70	10	10	42.59
447.85	160.1	IKVYLPR 3+	70	10	10	26.39
853.57	160.1	MEEKYNLTSVLMAMGITDVF SSSANLSGISSAESLK 5+	70	10	10	43.12
624.68	160.1	MKMEEK 2+	70	10	10	39.98
416.79	160.1	MKMEEK 3+	70	10	10	24.84
882.41	160.1	NVLQPSSVDSQTAMVLVNAI VFKGLWEK 4+	70	10	10	44.47
706.13	160.1	NVLQPSSVDSQTAMVLVNAI VFKGLWEK 5+	70	10	10	36.19
705.79	160.1	TQINKVVR 2+	70	10	10	44.52
470.86	160.1	TQINKVVR 3+	70	10	10	27.54
767.77	160.1	VASMASEKMK 2+	70	10	10	47.99
512.18	160.1	VASMASEKMK 3+	70	10	10	29.61
972.92	160.1	VHHANENIFYCPIAIMSALA MVYLGAKDSTR 4+	70	10	10	48.73
778.54	160.1	VHHANENIFYCPIAIMSALA MVYLGAKDSTR 5+	70	10	10	39.59
827.35	160.1	YPILPEYLQCVKELYR 3+	70	10	10	45.37
620.76	160.1	YPILPEYLQCVKELYR 4+	70	10	10	32.18
496.81	160.1	YPILPEYLQCVKELYR 5+	70	10	10	26.35
913.05	160.1	VTEQESKPVQMMYQIGLFR 3+	70	10	10	49.65
685.04	160.1	VTEQESKPVQMMYQIGLFR 4+	70	10	10	35.2
548.23	160.1	VTEQESKPVQMMYQIGLFR 5+	70	10	10	28.77
912.06	160.1	DILNQITKPNDVYSFSLASR 3+	70	10	10	49.6
684.29	160.1	DILNQITKPNDVYSFSLASR 4+	70	10	10	35.16
547.64	160.1	DILNQITKPNDVYSFSLASR 5+	70	10	10	28.74
735.33	160.1	KIKVYLPR 2+	70	10	10	46.18
490.55	160.1	KIKVYLPR 3+	70	10	10	28.53
421.14	160.1	KIK 2+	70	10	10	28.58

The mass to charge ratio (*m/z*) was first calculated for all possible peptides with a missed cleavage at lysine or arginine residues, taking into account the inability for trypsin to cleave at these residues where flucloxacillin modification is present. The mass addition of flucloxacillin (453 Da) was added to the theoretical mass of each potential drug modified peptide. The precursor ion (Q1) paired with the 160.1 Da fragmentation ion of flucloxacillin (Q3) were used for MRM transitions. MRM transitions were acquired at 1 unit resolution in Q1 and Q3 to maximise specificity in positive ion mode. Each experiment contained 1,983 cycles with a cycle time of 3,906 ms. Collision energy was calculated using the following formula; $\text{slope} \times m/z + \text{intercept}$ (Table 2.3).

Table 2.3. Information dependent acquisition (IDA) collision energy parameters

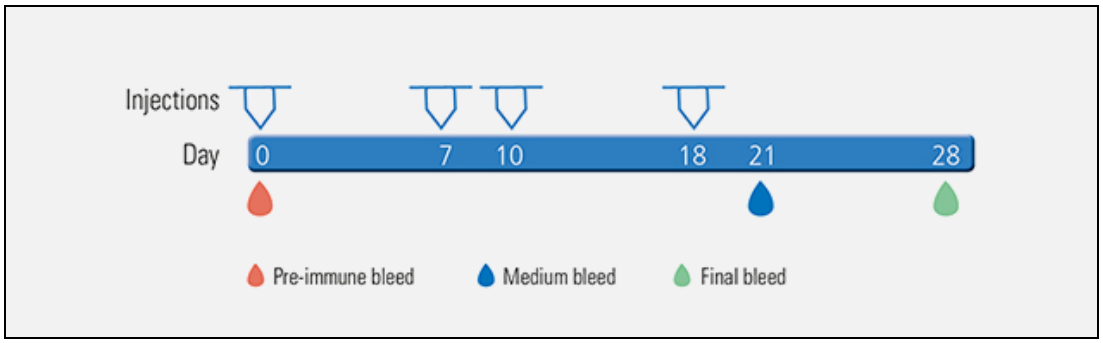
Charge state	Slope	Intercept
Unknown	0.044	5.000
1	0.058	9.000
2	0.044	5.000
3	0.050	4.000
4	0.050	3.000
5	0.050	3.000

Sample information was acquired using Analyst 1.6.2 software (Sciex) where ion intensity for each MRM transition was calculated. The area under the curve (AUC) for MRM transitions was normalised by the total ion count (TIC), allowing comparison between sample conditions. The MRM method developed identifies flucloxacillin modified tryptic OVA peptides based on the drug mass addition and flucloxacillin fragmentation ions. Manual interpretation of spectra was therefore performed to confirm flucloxacillin modification of these peptides.

2.3.5 ANTIBODY PRODUCTION

Antibody production was performed externally by Kaneka Eurogentec S. A. (Belgium). The speedy 28-polyclonal package was used for the production of a high antibody titre with high antibody affinity. Injections (100 µg/injection) are administered to two rabbits at days 0, 7,

10 and 18. A pre-immune bleed is taken at day 0, a medium bleed at day 21 and a final bleed at day 28 (Box 1). The bleeds are supplied as serum containing polyclonal antibody. Subsequent ELISA is performed by Eurogentec and in house (in addition to Western blot analysis).



Box 1. Kaneka Eurogentec S.A. Speedy 28-day program immunization and bleed schedule protocol. Taken from https://secure.eurogentec.com/speedy.html (date accessed 19/02/2019).

2.3.6 ELISA

Enzyme linked immunosorbent assays (ELISA) were performed to enable quantification of drug-hapten antibody binding at different hapten densities and antibody concentrations. High protein binding 96 well plates (Immunlon 4 HBX, Thermo Fisher) were coated with 100 ng of protein (positive & negative controls) or drug-protein conjugates overnight at 4°C. Wells were washed 5 times with phosphate buffer saline (PBS)/Tween (1 x PBS, 0.1% Tween-20) and blocked with blocking buffer (1 mg/mL bovine serum albumin (Sigma) (BSA) in PBS). After a 2 hour incubation at room temperature the blocking buffer was removed and wells were incubated with primary anti-drug antibodies at varying concentrations in blocking buffer for 2 hours at room temperature. Subsequent washes (5 x PBS/Tween) were followed by incubation of horseradish peroxidase secondary anti-rabbit antibody (Dako) at a 1 in 2,000 dilution in blocking buffer. After a further 2 hour incubation at room temperature 5 washes with PBS/Tween was followed by 2 washes with PBS. Signal was developed using 2,2'-azino-bis(3-ethylbenzothiazoline-6-sulphonic acid) (ABTS) (Sigma Aldrich) and detected at 405 nm (Dynex Technologies MRXe) using Revelation 4.25 software.

2.3.7 DOT-BLOT

In order to detect antibody responses to haptens, drug-protein was spotted onto nitrocellulose membrane and left to dry. After several spots were applied to the same site (small volumes were dotted at a time to prevent running) the nitrocellulose membrane was blocked and probed with antibody as described in section 2.3.3.1.

2.3.8 FLUCLOXACILLIN ANTIBODY INHIBITION

Flucloxacillin N-acetyl lysine conjugate was synthesized by NewChem Technologies (Durham, UK). A final concentration of 100 uM of the conjugate was used for hapten inhibition, and was incubated with the anti-flucloxacillin antibodies for 1 hour prior to use.

2.4 RESULTS

2.4.1 ANTI-DRUG ANTIBODY CROSS REACTIVITY

The aim of this chapter was to develop a high titre antibody for the detection of flucloxacillin conjugated to protein. BP is a beta lactam antibiotic with conserved structures with flucloxacillin. Commercially available anti-BP antibody was purchased (Serotec) to define cross reactivity with other structurally related compounds. B-lactam antibiotics including BP, piperacillin, amoxicillin and penicillin V, all containing the conserved thiazolidine ring (Figure 2.4, highlighted yellow), were conjugated to HSA at 1:1 and 100:1 molar ratios (drug:protein). The drug-protein conjugates were fully characterized using western blot and mass spectrometric analysis before being used for cross-reactivity assessment (Figure 2.4). Cross reactivity was observed between BP, piperacillin, amoxicillin and penicillin V, however, it was not possible to detect flucloxacillin haptenated HSA with anti-BP antibody. Although the thiazolidine ring is present on flucloxacillin (Figure 2.4, highlighted red) the epitope cannot be detected. This data confirmed the requirement for the custom generation of anti-flucloxacillin antibodies to enable detection of flucloxacillin using proteomics techniques.

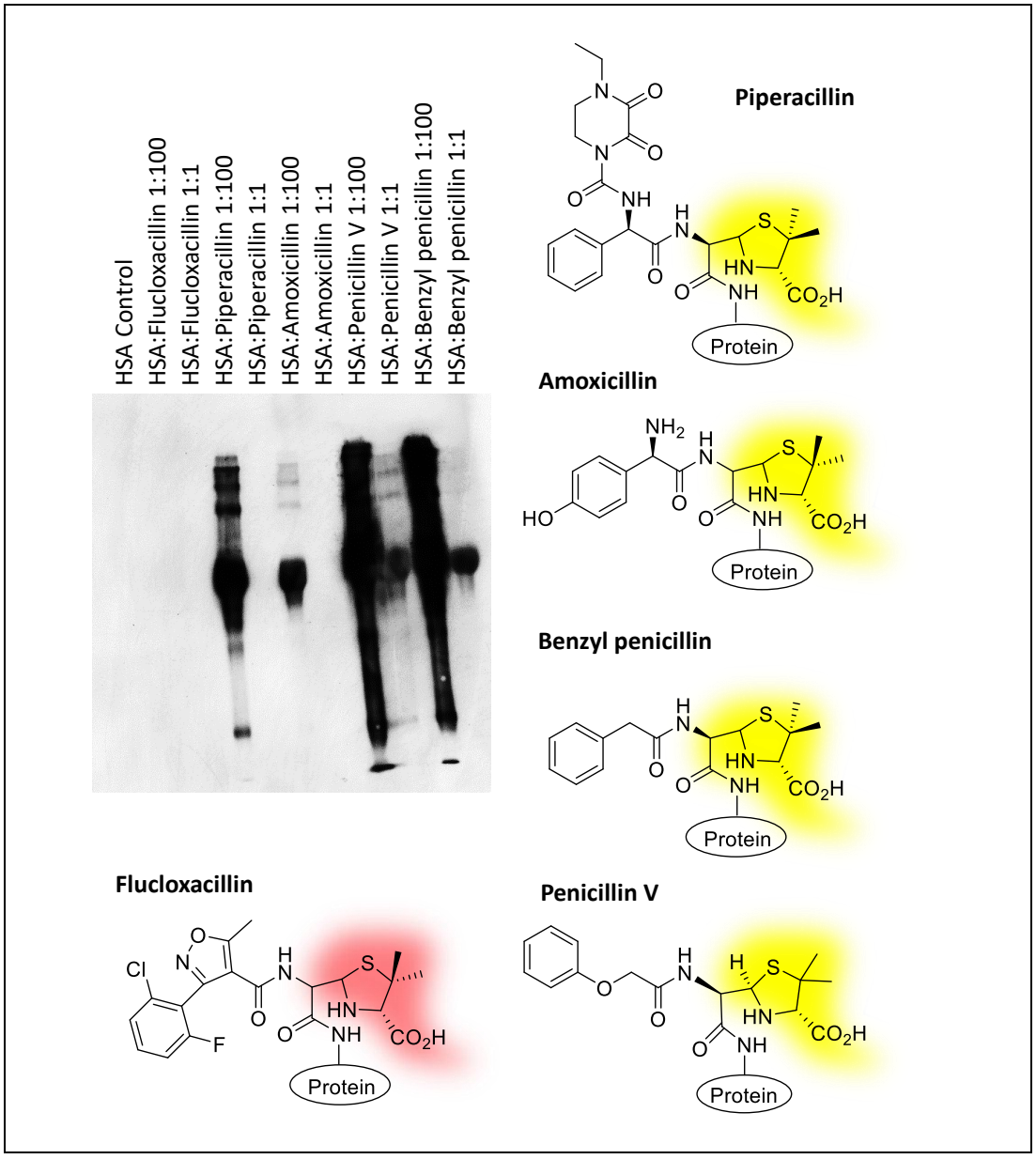


Figure 2.4. Anti-BP antibody cross reactivity. HSA was conjugated to flucloxacillin, piperacillin, amoxicillin, penicillin V and BP at 1:10 and 1:100 molar ratios. Western blot analysis was used to identify anti-BP cross reactivity amongst these β-lactam antibiotics. Cross reactivity was observed between piperacillin, amoxicillin, BP and penicillin V through recognition of the conserved thiazolidine ring (yellow). No signal was detected from flucloxacillin conjugated HSA, even though the thiazolidine ring is conserved (red).

2.4.2 HAPTEN CARRIER PROTEIN

A suitable hapten carrier protein was chosen based upon amino acid sequence homology to HSA. Limited sequence homology was required to prevent cross reactivity of polyclonal antibodies. Antibodies reactive to HSA would limit the ability to detect flucloxacillin bound to cellular human proteins where HSA was present. HSA (P02768) and BSA (P02769) were

aligned using NCBI BLAST suite to compare amino acid sequences. Unsurprisingly sequence homology was found to be as high as 88%, making cross reactivity highly likely. OVA (P01012), the main protein in hen egg white, was another alternative due to its availability and accessible lysine residues for flucloxacillin modification. NCBI BLAST found limited sequence homology between the proteins.

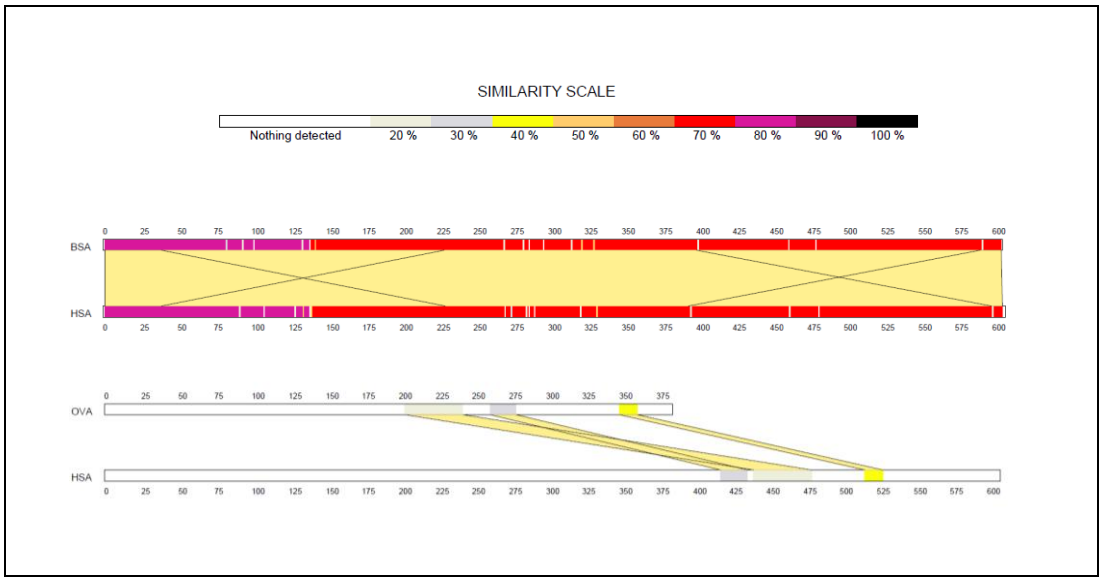


Figure 2.5. Protein sequence homology. BSA (top) and OVA (bottom) amino acid sequences were compared with HSA.

To fully interrogate sequence homology between HSA, BSA and OVA the ExPASy Sim alignment tool (Swiss Institute of Systems Biology) was used. Amino acid alignment of HSA to BSA resulted in 76.6% sequence identity with an overlap in 607 residues (Figure 2.5, top). In contrast alignment of HSA with OVA results in 22.0% sequence homology with an overlap of only 41 residues (Figure 2.5, bottom). Due to its low sequence homology to HSA, flucloxacillin binding to OVA was explored for use as a hapten carrier protein.

2.4.3 FLUCLOXACILLIN MODIFICATION OF OVALBUMIN FOR ANTIBODY GENERATION

The quantity of the antigens for immunization is vital for antibody production. Optimization of the flucloxacillin-OVA reaction conditions was therefore performed to generate antigens with a high epitope density. Preparation of the antigen required optimization of the

flucloxacillin-OVA reaction conditions. In this section the importance of buffer composition and molar drug protein ratios for maximal drug binding will be discussed.

2.4.3.1 OPTIMAL CONDITIONS FOR MAXIMAL HAPTEN DENSITY

Mass spectrometry was used to define the precise nature of flucloxacillin modified ovalbumin. Multiple reaction conditions were set up to include both buffer composition and drug protein molar ratio. Flucloxacillin OVA incubations that were performed in PO_4 CO_3 buffer resulted in low levels of modification which were comparable across all molar ratios. Flucloxacillin was found to bind to OVA more readily in carbonate buffer due to the increased pH (Figure 2.6).

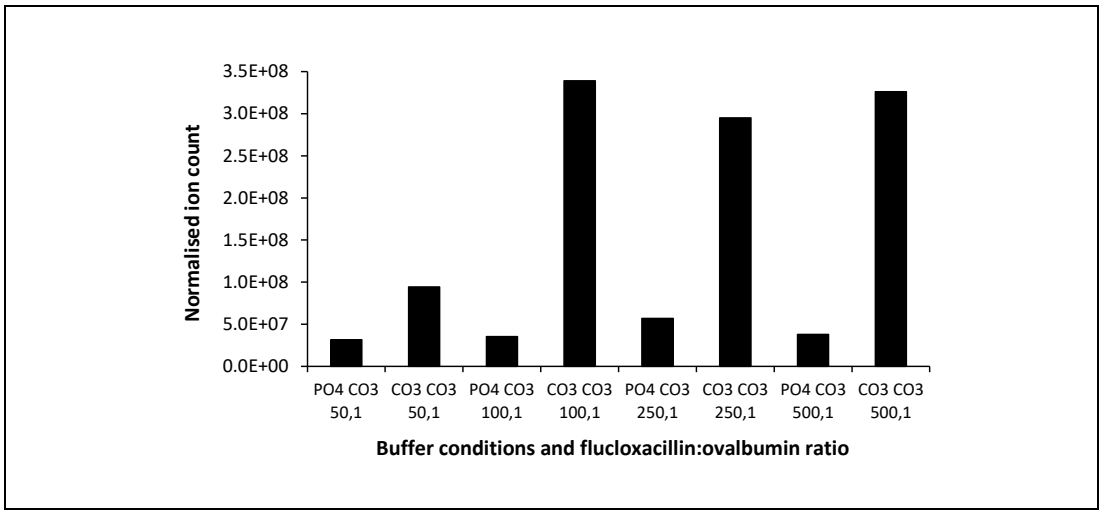


Figure 2.6. Relative quantification of fucloxacillin binding to OVA under different experiemental conditions. The overall level of flucloxacillin binding was determined by the addition of the AUC of relevant MRM transitions, normalised across samples using the TIC for each data set. Conjugation buffers were either phosphate:carbonate (PO_4 CO_3) or carbonate (CO_3 CO_3). Different molar ratios of flucloxacillin to OVA were prepared in each buffer. PO_4 CO_3 buffer resulted in lower hapten density compared to CO_3 CO_3, with 100:1 resuling in maximal binding.

Flucloxacillin modification of OVA was optimal in CO_3 buffer at a 1:100 molar ratio. While extensive binding was observed at 250 and 500:1 molar ratios this was comparable to 100:1 (Figure 2.6). Therefore, the additional drug did not increase the level of binding beyond 100:1. Of the 20 lysine residues in OVA (385 amino acids), only 6 lysine residues were found to be readily modified by flucloxacillin (Figure 2.7).

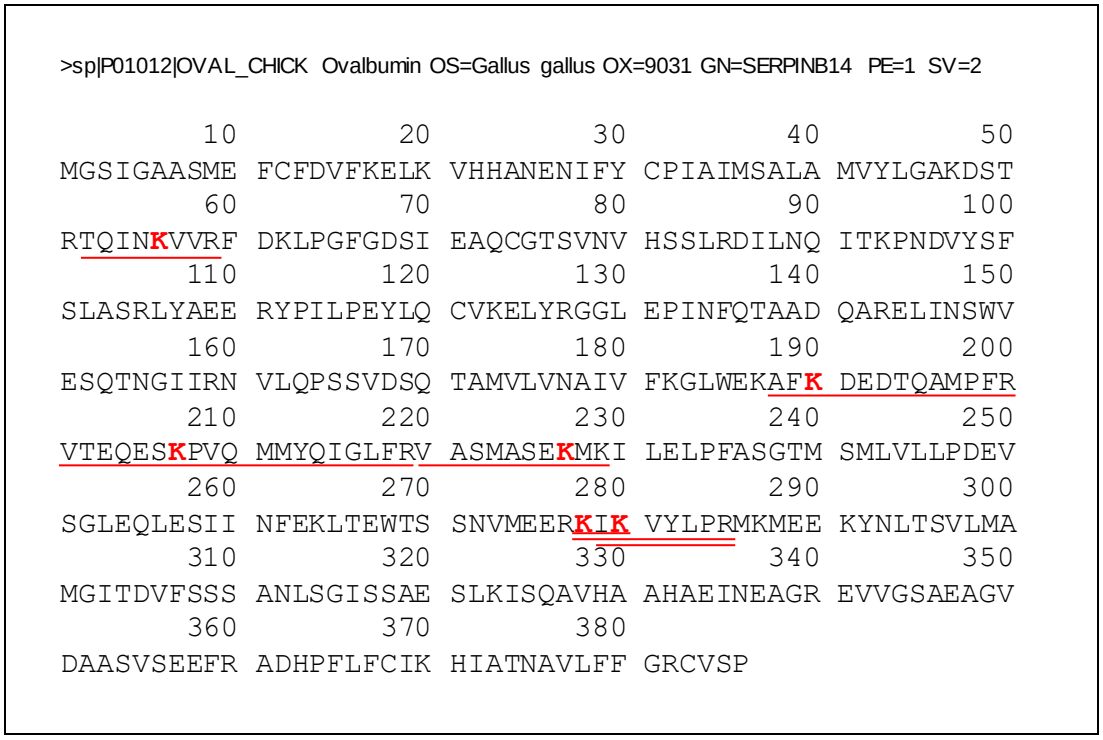


Figure 2.7. OVA amino acid sequence. Of the 20 lysine residues present in the amino acid sequcne of ovalbumin, 6 (highlighted in red) were found to be modified by flucloxacillin. Modified peptide sequences are shown as underlined. Modifications on lysine 278 and 280 were observed as K*IK, K*IKVYLPR and KIK*VYLPR, with * indiciating flucloxacillin.

Characterization of modified lysine residues revealed that under optimal conditions (CO_3 buffer at a 100:1 molar ratio) flucloxacillin binds to lysine 278 and lysine 280 most readily. Interestingly, this binding pattern appears irrespective of the buffer composition or the molar ratio of drug to protein (Figure 2.8).

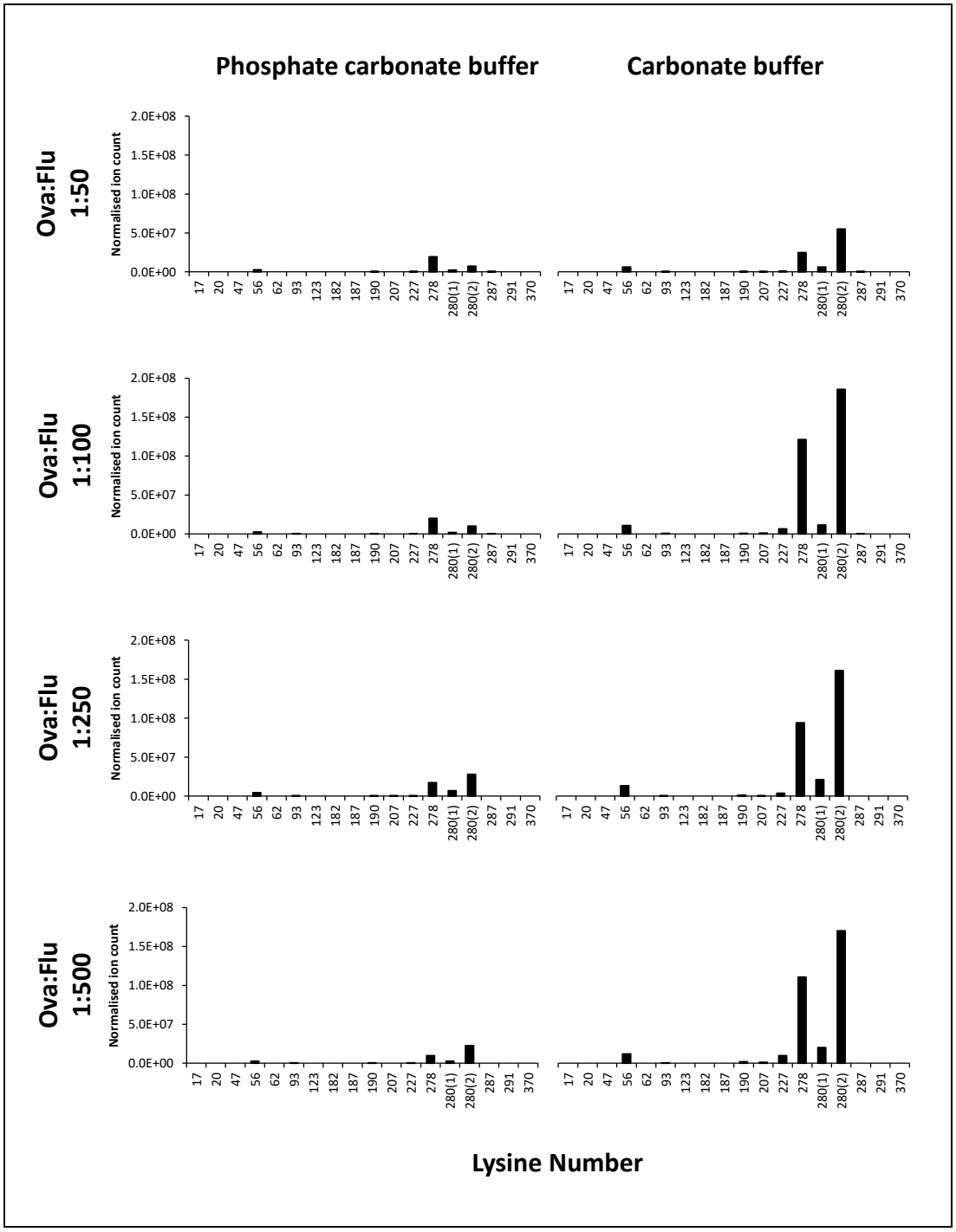


Figure 2.8. Optimal buffer composition and protein to drug molar ratio for maximal hapten density. Drug modified lyine residues were quantified when flucloxacillin was incubated with a 50 to 1 (top), 100 to 1, 250 to 1 and a 500 to 1 (bottom) molar ratio of drug to protein. Phosphate carbonate buffer (left) or carbonate buffer (right) were used to assess the impact of pH on flucloxacillin binding. Two peptides were found to be modified containint Lys 280; KIK[Flucloxacillin]VYLPR and IK[Flucloxacillin]VYLPR, and are denoted as 280(1) and 280(2) respectively.

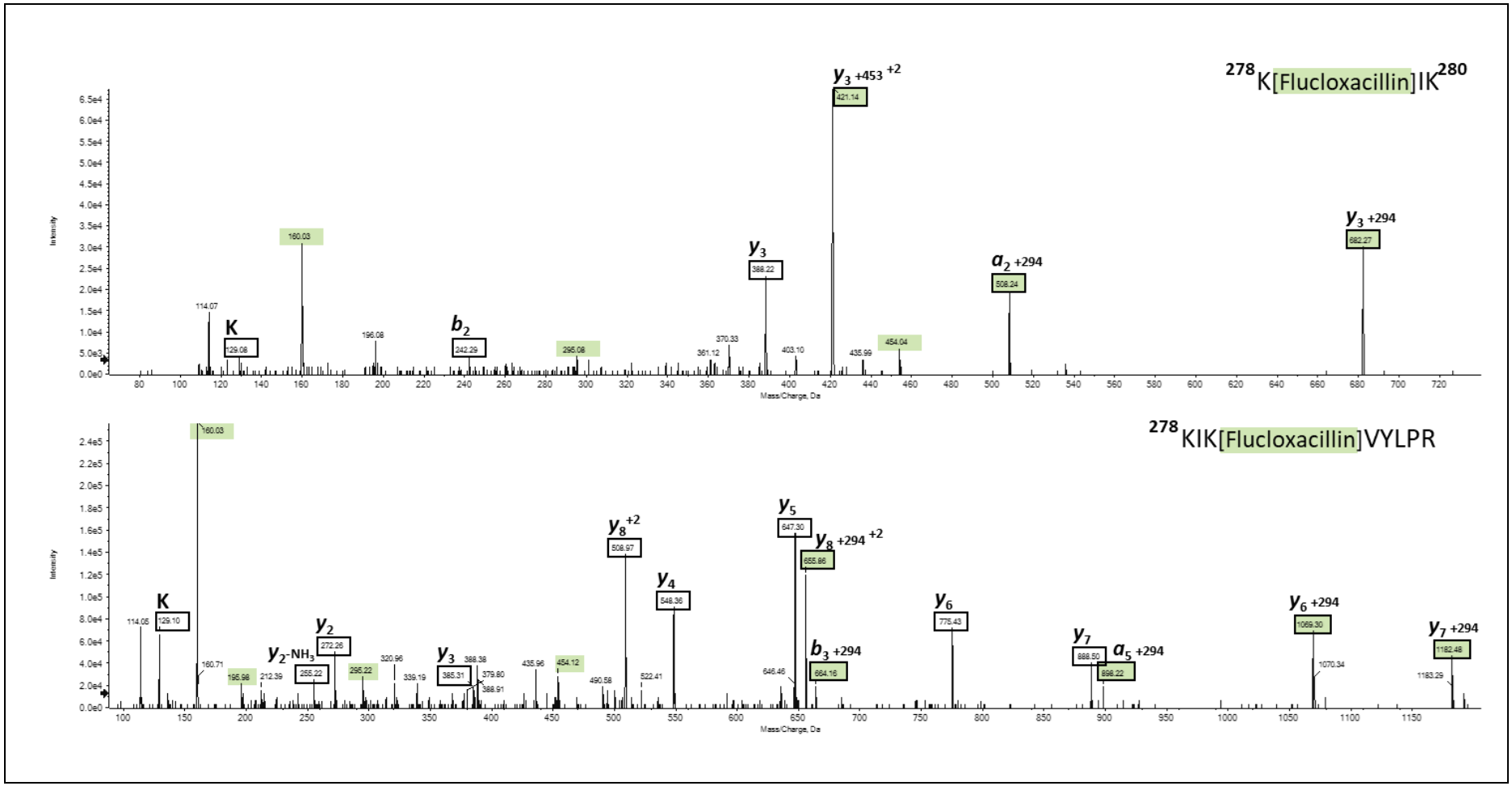


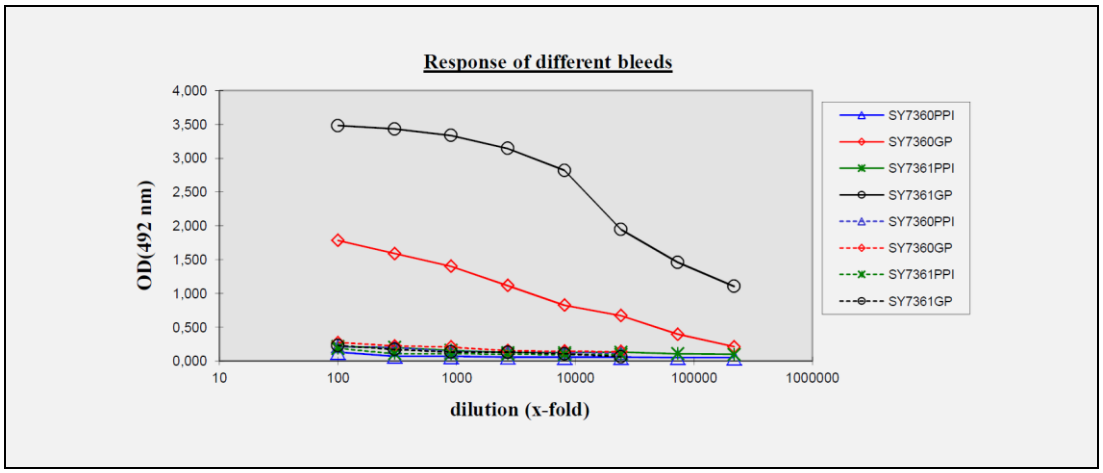
Figure 2.9. Mass spectrometric analysis of flucloxacillin modified ovalbumin. Modification of lysine 278 (top) and lysine 280 (bottom) were characterized using mass spectrometry.

2.4.3.2 MASS SPECTROMETRIC CHARACTERIZATION OF FLUCLOXACILLIN MODIFIED OVALBUMIN

MS/MS spectra corresponding to peptides containing lysine 278 and 280 were the most abundant across all incubation conditions. Modification of peptides was confirmed using several pieces of evidence; 1) flucloxacillin mass addition on the parent ion, 2) peptide sequences corresponding to *b* and *y* ions, 3) characteristic flucloxacillin fragment ions and 4) drug modified *b* and *y* ions. **K**(flucloxacillin)IK and KI**K**(flucloxacillin)VYLPR were both annotated to show both the presence and location of flucloxacillin (Figure 2.9). For both peptides all major peaks could be confidently assigned; leading to preparation of the flucloxacillin-OVA protein conjugate for immunization.

2.4.4 ANTI-FLUCLOXACILLIN ANTIBODY REACTIVITY

As part of the antibody generation, Eurogentec performed an ELISA to confirm responses to the immunogen used (OVA-flucloxacillin); the results are displayed in box 2.



Box 2. Antibody response to OVA-flucloxacillin conjugate used as immunogen. SY7360 and SY7361 correspond to the two different rabbits used to raise the antibody. PPI = preimmune serum, GP = large bleed. Straight line = antigen, dashed line = negative control.

Two rabbits (SY7360 and SY7361) were used in the immunization protocol. From the results provided by Eurogentec, both SY7360 and SY7361 gave positive results to the immunogen, with SY7361 giving a higher absorbance reading at 492 nm. It is important to note that these

are polyclonal mixtures of antibody in antisera, however from herein will be referred to as antibodies.

2.4.4.1 FLUCLOXACILLIN SPECIFICITY

In order to fully interrogate the specificity of the antibody provided, further assessment was performed. In house ELISAs were used to determine both the specificity of the anti-flucloxacillin anti-sera returned and the concentration required for optimal signal. In addition to the immunogen (OVA-flucloxacillin), responses to ovalbumin, HSA-flucloxacillin (1:100) and HSA were initially carried out (Figure 2.10).

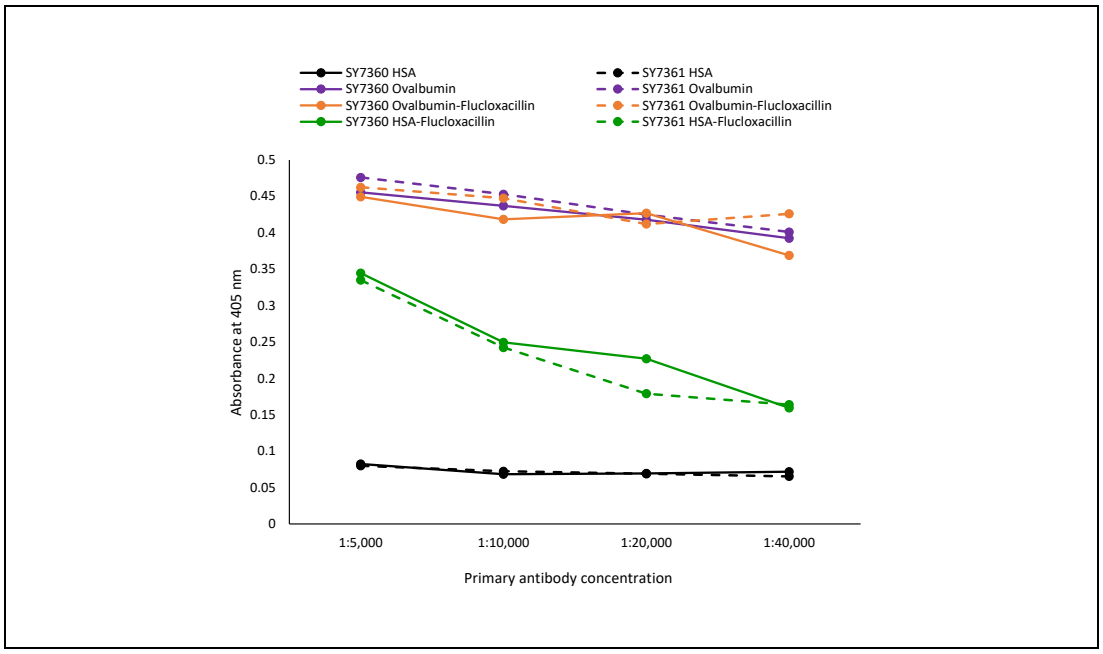


Figure 2.10. SY7360 and SY7361 (anti-flucloxacillin antibody/sera) specificity to flucloxacillin. An ELISA was performed to determine the specificity to flucloxacillin and optimal concentration of the sera provided from Eurogentec.

In contrast to the ELISA performed by Eurogentec, both SY7360 and SY7361 performed comparably to all of the test epitopes. As expected, highest responses were observed to OVA-flucloxacillin and ovalbumin. Both antibodies recognized HSA-flucloxacillin in a concentration dependant manner, but no response to HSA itself is observed, indicating they are both highly specific to flucloxacillin. From this data it is apparent that a dilution of 1:5,000

is optimal for both antibodies. In addition to ELISAs, dot blotting was used to identify the optimal conditions and any cross reactivity for use in Western blot analysis.

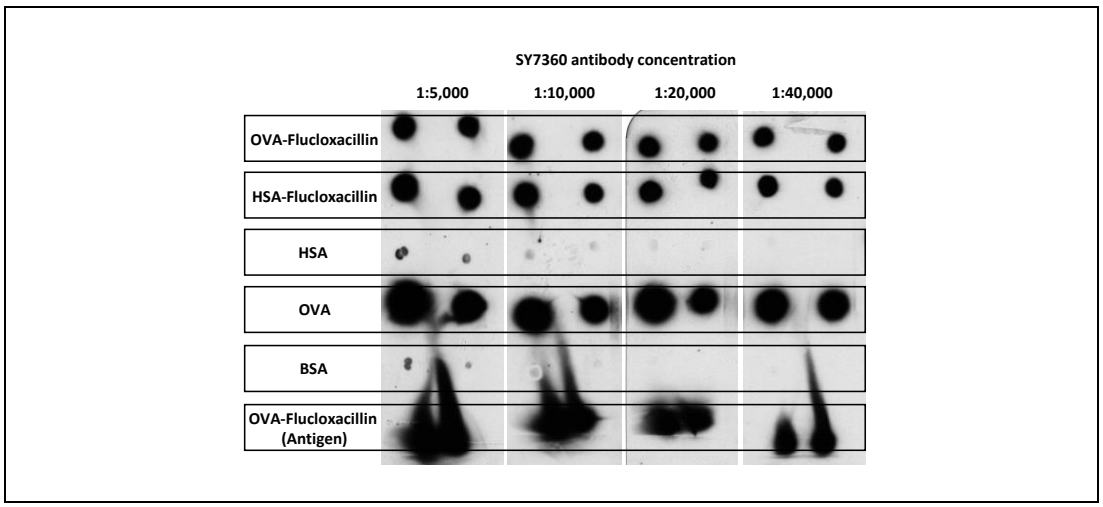


Figure 2.11. Dot immuno-blot to detect cross reactivity and optimal SY7360 antibody concentration for Western blot analysis. At all concentrations a detection of flucloxacillin modified proteins (OVA and HSA) were observed, in addition to the carrier protein (OVA). Limited cross reactivity was observed to HSA and BSA at high antibody concentrations (1:5,000).

The results from the dot blot mirrored that of the ELISA in terms of specificity and cross reactivity (Figure 2.11). Strong recognition to OVA-flucloxacillin, OVA and HSA-flucloxacillin conjugates were observed. Both HSA and BSA showed a limited signal. High concentrations of the antibody (e.g. Figure 2.11, 1 in 5,000) resulted in a signal being detected from HSA and BSA, while the signal to flucloxacillin was unchanged. As high antibody concentrations may result in cross reactivity/non-specific binding of HSA in cells derived from human sources, optimal antibody concentrations for immunoblot assays were determined to be 1 in 20,000.

2.4.4.2 HAPTEN DENSITY AND INHIBITION

Up to this point detection of flucloxacillin bound to HSA was only performed at a 1:100 HSA to flucloxacillin molar ratio. Molar ratios of 100:1, 1:1 and 0.1:1 drug to protein were prepared to determine the limit of detection using ELISA. In addition, hapten inhibition is important to confidently determine that the antibody is recognising the specific hapten. Incubation of the anti-flucloxacillin antibodies with flucloxacillin conjugated to N-acetyl lysine prior to use in assays should reduce the signal in the ELISA. Using ELISA it was only

possible to detect HSA with bound flucloxacillin at a 1:100 molar ratio (Figure 2.12). Signal from 1:1 and 1:0.1 molar ratios could not be distinguished from the background, indicating either the antibody and/or assay were not sensitive enough. Importantly, prior incubation of both antibodies with N-acetyl lysine-flucloxacillin resulted in the complete abrogation of signal when probing for HSA-flucloxacillin at previously detected molar ratios (100:1), indicating both antibodies are highly hapten specific.

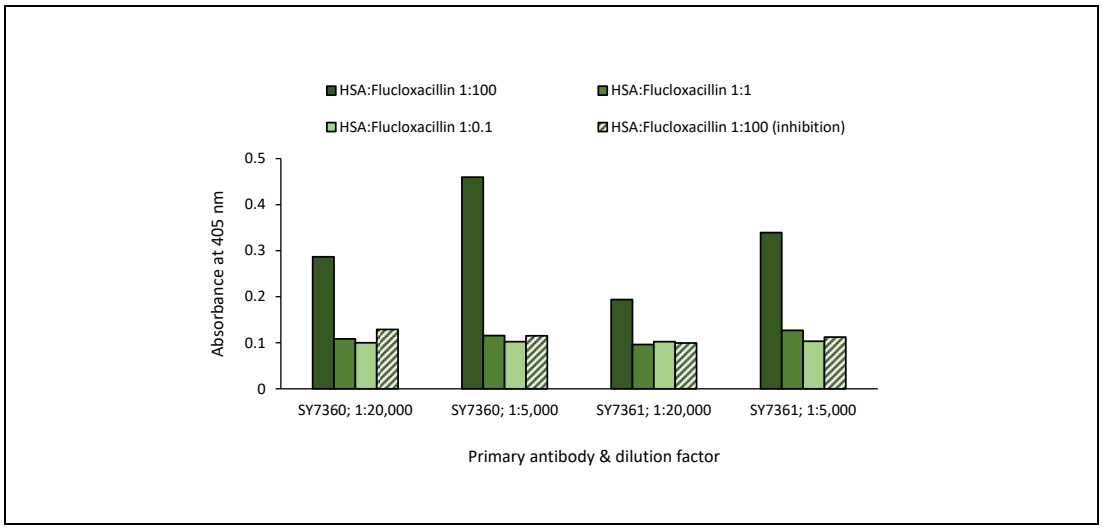


Figure 2.12. Detection of low hapten density using anti-flucloxacillin antibody. ELISAs were used to identify the limit of flucloxacillin at varying hapten density. Both antibodies were able to detect flucloxaillin conjugated to HSA at 1:100 HSA to drug ratios. HSA to flucloxacillin ratios of 1:1 and 1:0.1 could not be detected using ELISA. Full antibody inhibition was observed when pre incubated with N-acetyl lysine.

2.4.4.3 CROSS REACTIVITY WITH BETA LACTAM ANTIBIOTICS

At the beginning of this chapter cross reactivity of the commercially available anti-BP antibody was demonstrated using Western blot. This experiment was repeated using ELISA using β-lactam-HSA conjugates (flucloxacillin, amoxicillin, piperacillin, BP and penicillin V) at molar ratios of 1:100 (Figure 2.13).

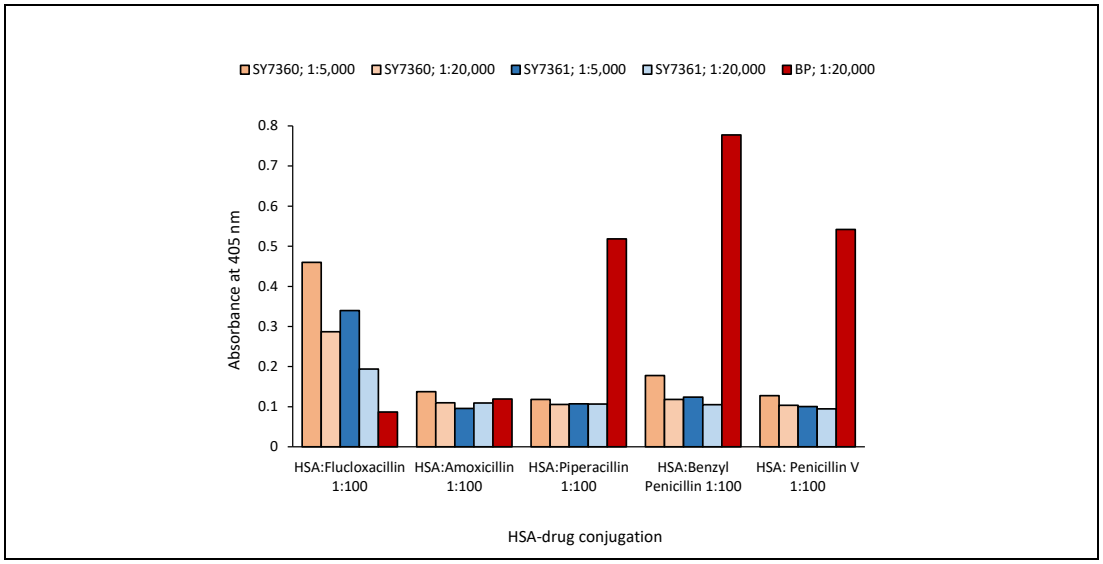


Figure 2.13. Cross reactivity of flucloxacillin and BP (commercially available) antibodies using ELISA. Fluxlocaxillin specific antibody was only able to detect HSA modified by flucloxacillin at both 1:5,000 and 1:20,000 concentrations. BP specific antibody was able to detect BP, piperacillin and penicillin V conjugated to HSA. Amoxacillin HSA could not be detected by either antibody using ELISA at these concentrations.

In addition to anti-BP, each anti-flucloxacillin antibody was used to detect the HSA drug conjugates at a 1 in 5,000 and 1 in 20,000 dilution. As shown before, cross reactivity of the anti-BP antibody with piperacillin and penicillin V were observed. Surprisingly, no cross reactivity with amoxicillin could be detected. Anti-flucloxacillin antibodies were only specific to HSA modified by flucloxacillin at both concentrations.

It was shown that the ELISA was unable to detect cross reactivity of the anti-BP antibody with amoxicillin. Previously when using Western blotting the cross reactivity for amoxicillin HSA was lower than that of the other β-lactams (Figure 2.4). Therefore, we hypothesised that ELISAs had a reduced level of sensitivity compared with Western blotting. Subsequent Western blots were performed to further interrogate the cross reactivity of the anti-flucloxacillin antibody with both lower molar ratios and β-lactam haptens as attempted by ELISA (Figure 2.14).

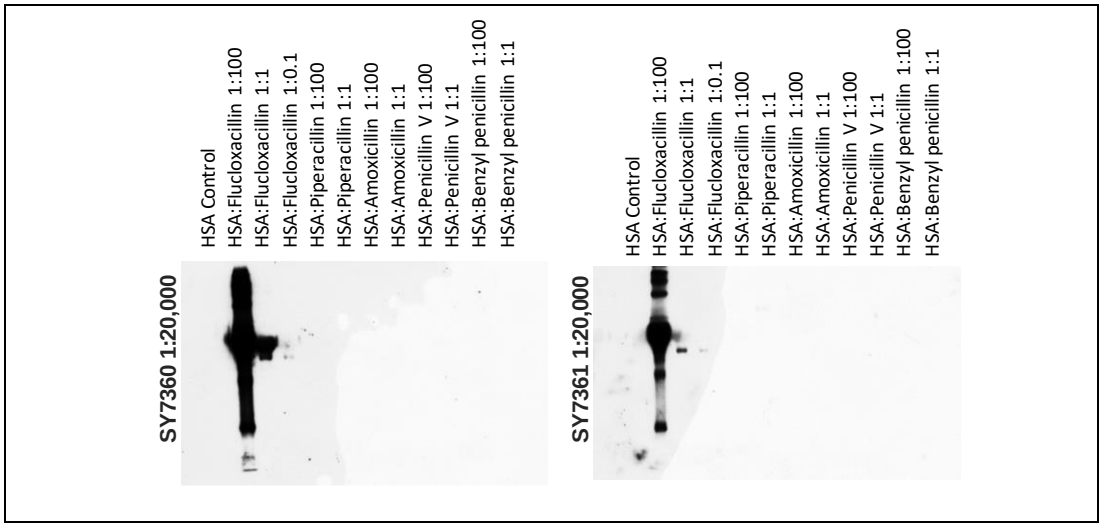


Figure 2.14. Detection of flucloxacillin HSA at different hapten densities and cross reactivity to other β-lactam antibiotics using Western blot analysis. Using Western blot, HSA flucloxacillin could be identified at 100:1, 1:1 and 0.1:1 molar ratios. No cross reactivity to piperacillin, amoxacillin, penicillin V and BP was obsereved.

Using Western blot it was possible to detect HSA flucloxacillin conjugates at molar ratios of 1:0.1 using both antibodies, with no cross reactivity to HSA or any other of the β-lactam haptens. This confirmed that both SY7360 and SY7361 were sensitive and specific to flucloxacillin. While flucloxacillin does contain a thiazolidine ring this was not believed to be the epitope recognized by the antibody due to this lack of cross reactivity. When compared to amoxicillin, piperacillin, BP and penicillin V, flucloxacillin is the only β-lactam to contain an isoxazole ring. To determine whether this was the epitope recognition site further HSA conjugates to β-lactam antibiotics, each containing the isoxazole ring, was performed.

HSA was conjugated to oxacillin and cloxacillin at 1:100 molar ratios. It was not possible to determine the molar ratio of HSA to dicloxacillin due to remaining stocks being in solution at an unknown concentration, however as binding of flucloxacillin reached a maximal level we believed it would be modified enough for detection using Western blot. Indeed, cross reactivity of the anti-flucloxacillin antibody was observed between flucloxacillin, oxacillin, cloxacillin and dicloxacillin using Western blot (Figure 2.15). Importantly all signals could be diminished using hapten inhibition, showing that the epitope for antibody detection is conserved between these structurally related compounds.

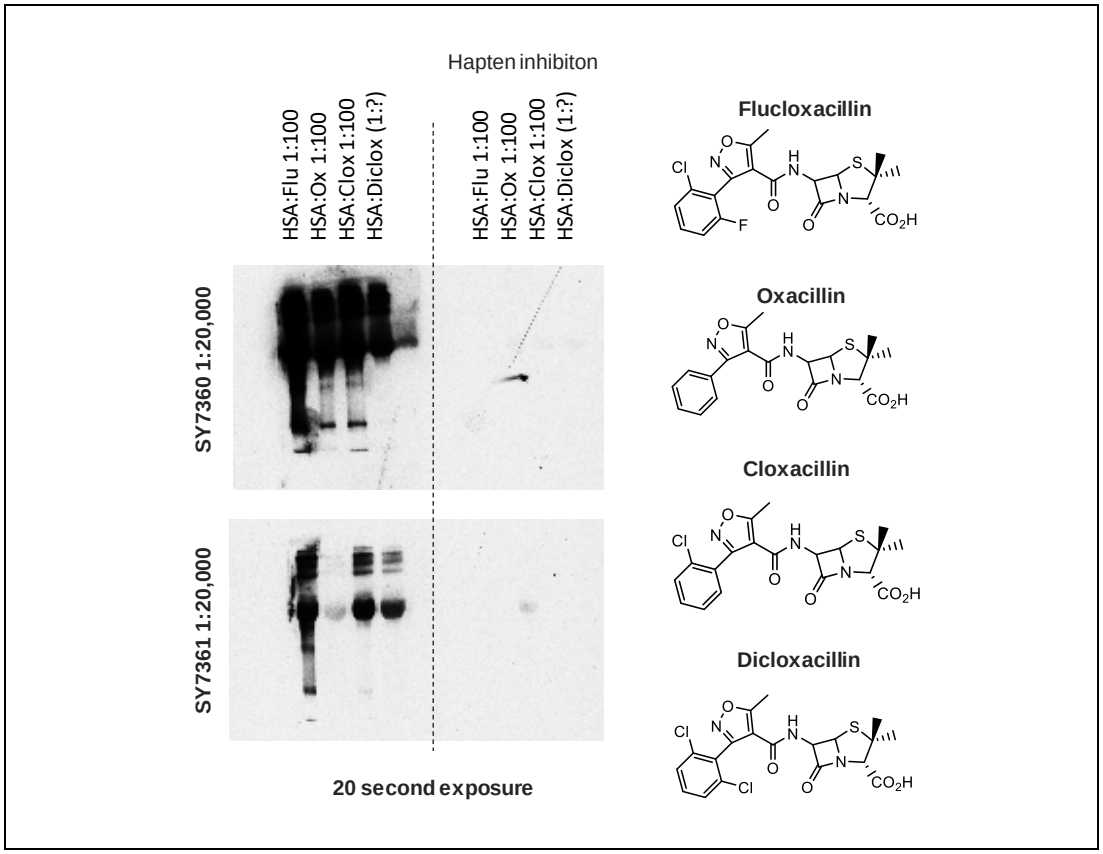


Figure 2.15. Cross reactivity of the flucloxacillin specific antibody with structurally related β-lactam antibiotics. Both flucloxacillin antibodies were able to cross react with oxacillin, cloxacillin and dicloxacillin (left). Hapten inhibition was oberved with both antibodies when pre-incubated with N-acetyl lysine flucloxacillin (right).

2.5 DISCUSSION

Anti-drug antibodies can be purchased commercially for use in hapten detection. BP antibody is cross reactive with certain β-lactam antibiotics that contain a thiazolidine ring (amoxicillin, piperacillin and penicillin V), but not flucloxacillin. Therefore, the aim of this chapter was to develop an antibody to detect flucloxacillin, with limited cross reactivity to human proteins, to enable use in the detection of flucloxacillin modified proteins in cellular systems. The decisions made for the choice of hapten carrier, incubation conditions, characterization and quality control were instrumental in the successful generation of the antibody.

Low MW compounds do not generally trigger an immune response, and so they must be conjugated to a suitable carrier (Gefen *et al.*, 2015). The secretion of antibodies by differentiated plasma cells is reliant on the activation of membrane bound immunoglobulins on the cell surface. The carrier protein/polypeptide that is used for antibody generation must, most importantly, be capable of eliciting an immune response in the host. Another important factor in the selection of a carrier protein depends on the applications of the antibody undergoing development. As described, the use of human proteins as a carrier would indeed most likely result in an immunogenic response to proteins when administered in a different species, however antibodies specific to the carrier would be secreted into the blood plasma. If HSA were to be used it is likely that any sample originating from human cells would contain HSA, resulting in an anti-HSA response. With this said, it is also important to select a protein that has poor sequence homology to, in this example, HSA. BSA was found to have a high sequence homology to HSA and for this reason was immediately rejected for use as a carrier protein.

Several carrier proteins are often used for the development of antibodies, including both OVA and KLH. Although, OVA is weakly immunogenic in comparison to KLH (Fasciglione *et*

al., 1996). As we were interested in the hapten bound to the carrier protein, rather than the carrier itself, it was important to consider the availability of lysine residues for drug binding. OVA is 45 kDa in size, while KLH is a mega protein at 390 kDa. Although it is the most commonly used carrier protein to elicit an immune response of a hapten, the ability to characterize KLH adducts with mass spectrometry is quite challenging. While reactive lysine residues are present on KLH its size makes it difficult to analyze multiple proteins at one time using mass spectrometry. On the other hand, although less immunogenic, OVA contains free lysine residues that could potentially be modified by flucloxacillin. Sequence homology between OVA and HSA was found to be limited to 22% making it an ideal candidate as a carrier protein, therefore the modification of OVA by flucloxacillin was investigated.

Previous experiments performed in our group have indicated that flucloxacillin modification on HSA increases at a higher pH. For this reason, flucloxacillin was incubated in phosphate buffer (pH 7.4) or carbonate buffer (pH 11). At high molar ratios large amounts of flucloxacillin, that contains a free carboxylic acid, reduced the pH in the phosphate buffer to more acidic conditions. Generally under acidic conditions covalent binding is not favourable. Indeed, we found that the level of binding to OVA was lower than in the carbonate buffer. Interestingly the binding of flucloxacillin to OVA plateaued at a ratio of 100:1 flucloxacillin to OVA, meaning there was no added benefit in increasing the molar ratio above this limit. Once the optimal conditions were defined, Imject OVA (Thermo Scientific), formulated specifically for use as a carrier protein, was used for immunogen preparation.

Mass spectrometry is a powerful tool enabling the precise characterisation of flucloxacillin modified OVA peptides. MRM was used due to its increased specificity compared with QTOF MS/MS. QTOF MS/MS is reliant on the top *n* most abundant ions present at a point in chromatographic time, triggering fragmentation. Mass spectrometry revealed modification on 6 of the 20 available lysine residues, with lysine 278 and 280 consistently undergoing haptenation. The presence of flucloxacillin can be determined by characteristic

fragmentation ions derived from flucloxacillin. The presence of 160, 195, 295 and 454 Da ions in MS/MS spectra all related to the MS fragmentation of flucloxacillin, indicating covalent binding of flucloxacillin to peptides. The most abundant fragmentation ion is at 160 Da, corresponding to the cleavage of the thiazolidine ring. In conjunction with other fragmentation ions, especially 454 Da (whole flucloxacillin mass), it is usually indicative of a flucloxacillin modified amino acid within the sequence. Of course, full peptide sequencing was performed to ensure that the *B* ions (starting from the N terminus) and the *Y* ions (starting from the C terminus) could be confidently annotated.

ELISAs performed by Eurogentec showed that the antibodies present in the sera of the immunized rabbits bound to the immunogen. Importantly, the antibodies also bound to flucloxacillin conjugated to HSA while no signal was detected from HSA alone. This indicated the antibody was specific for the hapten (flucloxacillin), rather than the protein backbone. Indeed, the antibodies responded to flucloxacillin conjugated to HSA at a range of concentrations. It was found that a concentration of 1 in 5,000 gave the best signal using an ELISA while maintaining the same background absorbance. In addition to an ELISA a dot immuno blot was performed to ensure flucloxacillin could be detected using Western blotting protocols. However, at high concentrations cross reactivity to HSA and BSA were observed; this is understood to be due to experimental technique. Prior to dotting samples on the nitrocellulose, circles were drawn on the nitrocellulose membrane to allow multiple spots to be placed in the same area. This may have resulted in a slight depression of the membrane, making a physical 'groove' for the antibody to sit in. When washing, antibodies within this groove may not have been removed, resulting in apparent background. As the signal for flucloxacillin HSA did not diminish when the antibody concentration was reduced from 1 in 5,000 to 1 in 20,000, the latter was believed to be optimal for this technique. Indeed, this condition was primarily used due to the lack of cross reactivity with HSA/BSA.

Both ELISAs and Western blots were performed in order to determine the level of drug modification required to enable detection. Using ELISA only high levels of modification could be detected, whereas using Western blotting signals from both 1:1 and 0.1:1 flucloxacillin to HSA ratios were identified, making it clear that this was a far more sensitive technique. In both cases, hapten inhibition using flucloxacillin modified N-acetyl lysine resulted in the complete abrogation of the signal, indication the anti-flucloxacillin antibodies are highly specific. Prior incubation of the antibodies with N-acetyl lysine flucloxacillin results in binding to the antigen binding site of the Fab region, thus preventing interaction with flucloxacillin immobilized on the nitrocellulose membrane. While from this inhibition it is not possible to categorically define the epitope recognition, it is likely to be flucloxacillin itself rather than flucloxacillin bound to lysine. Previous studies investigating antibody epitope sizes range from ~20 to 400 amino acid contact residues, with the most frequently observed between 50 and 79 amino acids (Stave and Lindpaintner, 2013). Other anti-β-lactam penicillin antibodies subsequently produced in the lab can be inhibited by the drug alone, without the need for conjugation to N-acetyl lysine, further supporting this hypothesis.

Cross reactivity to other β-lactam antibiotics confirmed BP specific antibodies were able to detect piperacillin-, penicillin V- and indeed BP- modified proteins using ELISA. Neither flucloxacillin- nor amoxicillin- modified proteins could be detected, even when high molar ratios of drug to protein (100:1) were used. Similarly, the flucloxacillin antibody could only detect flucloxacillin- modified proteins. As Western blots were found to be a more sensitive, cross reactivity was assessed using this technique. Although flucloxacillin antibodies can detect very low levels of modification (drug to protein ratio 0.1:1) it cannot cross react to proteins modified to other β-lactam antibiotics, even when high levels of modification were formed (100:1). This study indicated that the anti-flucloxacillin antibodies are highly specific and may bind to the side chain rather than the thiazolidine ring.

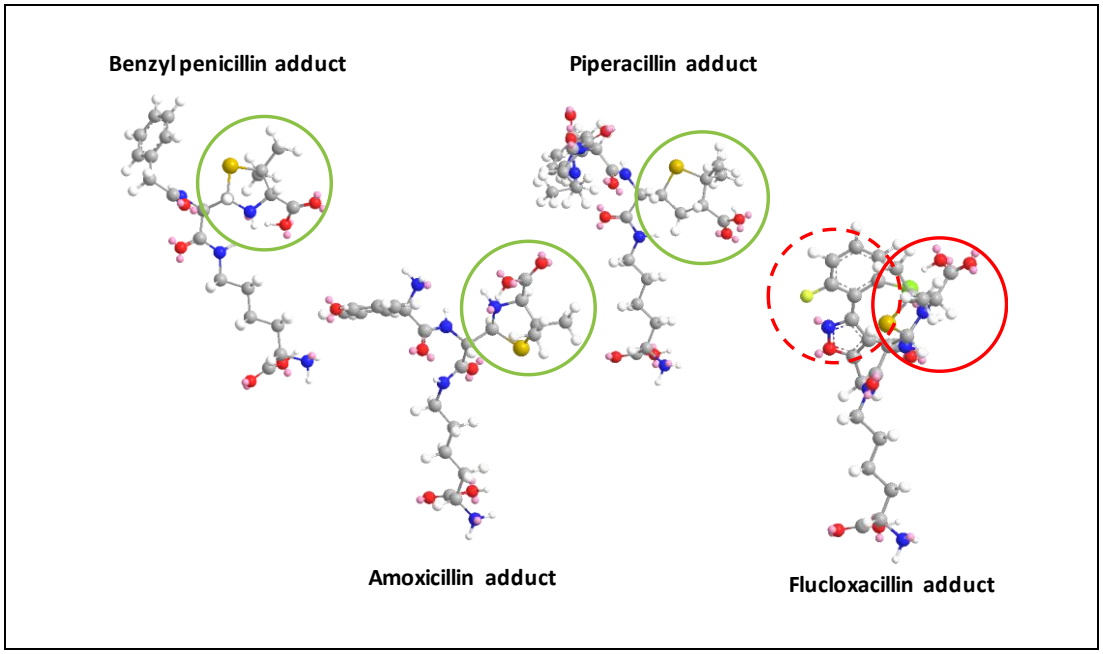


Figure 2.16. Chemical structures of BP, piperacillin, amoxicillin and flucloxacillin adducts. The conserved thiazolidine ring protrudes from the core of the molecule in BP, piperacillin and amoxicllin, leading to antibody recognition and cross reactivity (green circle). Flucloxacillin, although contains a thiazolidine ring (red solid circle), it does not extrude from the main structure. Instead, the isoxazole ring (red dashed circle) is the most likely antibody epitope.

In contrast, cloxacillin, dicloxacillin and oxacillin are all β-lactam antibiotics with further structural similarity to flucloxacillin. In addition to the thiazolidine ring, they all contain an isoxazole ring in their structure. For this reason, cross reactivity was assessed using Western blot analysis. Interestingly, flucloxacillin antibodies can detect HSA modified by all four compounds. These signals could all be diminished with the use of the antibody inhibition. This study further confirmed that the antibody is specific to the isoxazole ring. Upon investigating the 3D structure of BP, piperacillin and amoxicillin it was clear that the thiazolidine ring stuck out from the main core structure for antibody recognition (Figure 2.16, green circles). With flucloxacillin, the stereochemistry results in the inversion of the thiazolidine ring towards the core structure, thus preventing interactions with antibodies (Figure 2.16, red solid circle). Therefore, as the only side chain present with a conserved structure between flucloxacillin, cloxacillin, dicloxacillin and oxacillin, the isoxazole ring (Figure 2.16, red dashed circle) was identified as the most likely epitope for antigenic recognition and discrimination.

Previous studies have identified anti-drug-antibodies in the sera of patients with DHRs, however their role in delayed type hypersensitivity is not fully understood. In order for drugs to elicit an immune response they must first bind to macromolecules acting as carrier proteins. The aim of this chapter was to covalently bind flucloxacillin to a suitable carrier for immunization of rabbits for the generation of drug specific antibodies. Indeed, a high titre antibody specific to flucloxacillin was successfully generated. Cross reactivity was determined between other relevant proteins, with no cross reactivity to HSA observed, even at high antibody concentrations. Furthermore, cross reactivity to other β-lactams gave us an understanding of the highly specific nature of the flucloxacillin antibody. This specificity was further confirmed through inhibiting antibody detection of flucloxacillin-modified proteins through pre-incubation with flucloxacillin. In this chapter the specific nature of the flucloxacillin antibodies generated were interrogated. However, the detection of flucloxacillin was limited to *in vitro* drug incubations on model proteins. In order to understand the pathophysiology of flucloxacillin in DILI it is important to investigate flucloxacillin binding and localization in a more physiological context, including liver derived cell lines. Here we have shown the development of an important tool to further interrogate flucloxacillin in a more relevant context, performed in the next chapter using a range of different proteomic techniques.

CHAPTER 3 - DETECTION OF FLUCLOXACILLIN BINDING IN CELL CULTURE SYSTEMS

3.1 INTRODUCTION

Previous findings show flucloxacillin binds to albumin in patients, however many cellular protein targets or localisation of covalent binding remains unknown. Previous studies tried to understand flucloxacillin binding in animal models, where rats had been treated with varying doses of flucloxacillin over specific time periods before being euthanized to harvest the liver (Carey and van Pelt, 2005). While in this study the generation of a flucloxacillin specific antibody was successful, using Western blot analysis on liver lysates resulted in a high degree of cross reactivity to proteins from untreated rats. Nevertheless, Carey *et al* were able to hypothesise that the localization in which flucloxacillin modified proteins existed by separating subcellular fractions using centrifugation. Nuclear/membrane, mitochondrial, microsomal and cytosolic preparations were individually probed using SDS-PAGE and Western blotting. In total 6 adducts were consistently identified in different subcellular fractions, however it was difficult to positively identify the protein sources. Interestingly within the microsomal fraction adduct formation was detected at the MW of CYP P450 enzymes. This 52 kDa adduct was present at all doses administered to the rats; the only adduct detected in the low dose treatment. As CYP3A4 is involved in flucloxacillin metabolism, and is a part of the P450 enzyme family, it could indeed be a target for adduct formation upon activation (Carey and van Pelt, 2005).

Membrane transporters are important for the influx and efflux of drugs and their metabolites between different cell types. The progression to disease from a range of drugs has been linked to the activity of transporter proteins resulting in cholestatic effects.

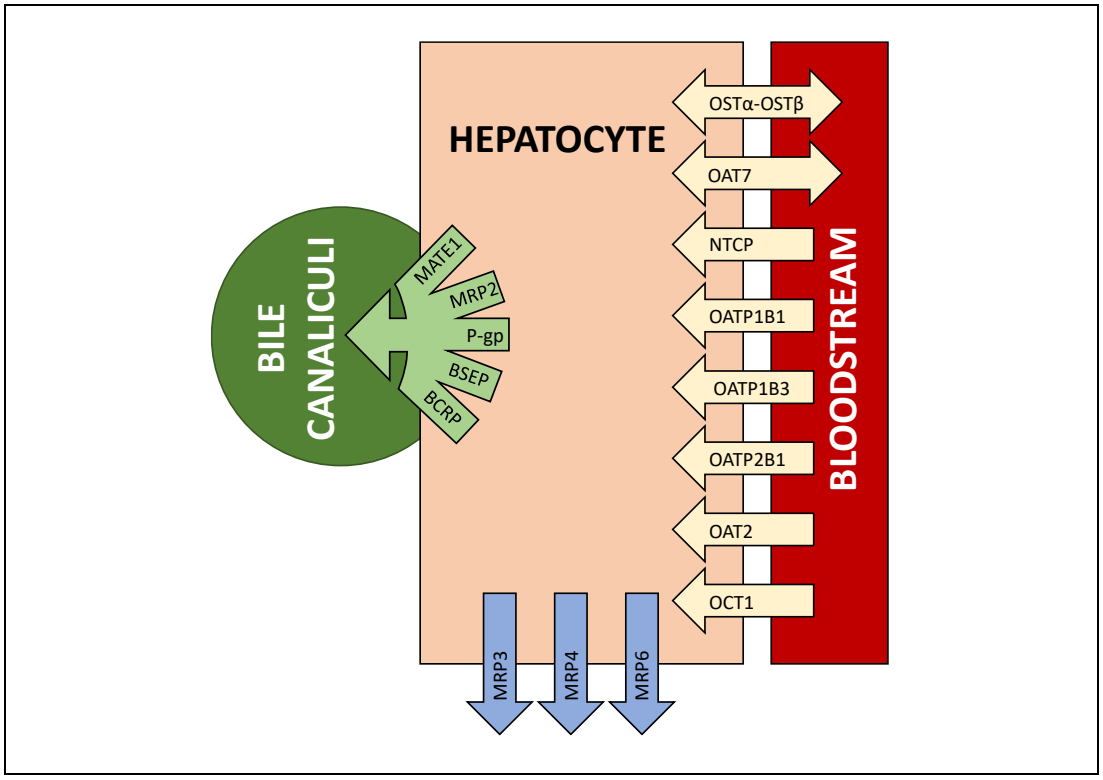


Figure 3.1. Common membrane transporters of the liver. OATPs contribute to a large proportion of the influx transporters on the apical membrane of hepatocytes. MDR proteins are largely responsible for the transport of products out of hepatocytes into adjoining cells, in particular the bile canaliculi.

Influx transport from the apical membrane of the liver is largely reliant on the organic anion transporting polypeptide (OATP) proteins, however others do exist (Figure 3.1) (Sundaram and Björnsson, 2017). A number of different OATPs are present depending on the type of cell, for example OATP1B3 and -1B1 are specifically expressed in the liver (Smith, Figg and Sparreboom, 2005). In addition, the movement of drugs and their metabolites is often specific to certain transporters, for example, benzyl penicillin is most actively transported by OATP2B1 and -1B1 (Tamai *et al.*, 2000). This may explain why certain drugs contribute to disease types in specific organs. Currently, little is known as to the transporters associated with flucloxacillin. Efflux transporters are responsible for the movement of drugs into the bile canaliculi from hepatocytes. These glycoproteins, collectively part of the multidrug resistance (MDR) protein (MRP) family include MDR1 (P-gp), MRP2 and the bile salt export pump (BSEP). A range of other transporter proteins are known to be present in different cell types (Figure 3.1). While all efflux transporter proteins may have involvement in drug

clearance, BSEP is largely implicated in the transport of drug metabolites out of hepatocytes (Sundaram and Björnsson, 2017). Flucloxacillin efflux transporters are largely unknown, however patients with mutations in the genes that encode MDR3 and BSEP have a 3-fold increase in the risk of developing DILI with certain other antibiotics (Lang *et al.*, 2007). Protein detection based methods, such as immunofluorescence imaging, have been developed to interrogate the activity of transporters under certain conditions, i.e. the presence of drugs. For example, transporter activity of MRP2 and P-gp can be assessed using fluorescently labelled 5-chloromethylfluorescein diacetate (CMFDA). While CMFDA can passively enter the cell, it can only be effluxed out of the cell via active transport through MRP2 and P-gp. Accumulation within the cell cytoplasm represents an absence or inactivity of these membrane transporters (Gaskell *et al.*, 2016).

In order to detect flucloxacillin protein targets in relevant human material, the liver-like cell lines HepG2 and HepaRG are particularly useful tools. While neither can completely replicate the physiology and microenvironment of primary human hepatocytes, their availability and ability to survive in cell culture make them an ideal candidate for the study of liver toxicity. HepG2 cells, first reported in 1980 by Knowles *et al*, were derived from the liver biopsy of an adolescent child (Caucasian male, 15 years old) with primary hepatocellular carcinoma. One of the key features of HepG2 cells is their ability to synthesize and secrete 17 of the major human plasma proteins into the cell culture medium (Knowles, Howe and Aden, 1980). The HepaRG cell line was too isolated from a patient tumour, however in this case Hepatitis C virus (HCV) was present. HepaRG cells were the first cell line to be successfully infected by Hepatitis B virus (HBV), with infection only previously achieved in primary human hepatocytes. This was particularly important as HBV has a very narrow cell specificity and is restricted to infecting differentiated cells that can support its full replication cycle. This highlights one of the key features of HepaRG cells in their ability to maintain efficient proliferentiation and differentiation during cell culture. The addition of DMSO in the last 15

days of cell culture leads to this differentiated state, where increases in the metabolising enzymes CYP1A and CYP3A4 are observed (Gripon *et al.*, 2002). The HepaRG differentiation cell cycle begins after one-week post seeding where cells commit to either hepatocyte or biliary pathways. After 2 weeks hepatocyte-like colonies are surrounded by epithelial cells (primitive biliary cells). These colonies can be detected through the formation of numerous bile canaliculi, characteristic of polarized hepatocytes (Cerec *et al.*, 2007).

The choice of human hepatic tumour derived cell lines depends on the nature of the experiment conducted. Compared to primary human hepatocytes, HepG2 CYP 450 expression is reduced by 90%, much lower than a 60% reduction in HepaRGs. That said, HepaRGs overexpress CYP3A4 compared to primary cells. Membrane transporter expression is another important factor when selecting the appropriate cell line to use for a specific purpose. For example, the efflux transporter MRP2 is expressed equivalently in HepaRG and primary hepatocytes, however, is not expressed in HepG2 monolayer cells. Another efflux transporter, P-gp (MDR1), was found to be expressed equivalently in HepG2s while over expressed in HepaRGs; again, compared to primary cells. In general terms however, it is accepted that HepaRG cells diverge less from primary human hepatocytes in terms of protein expression (Sison-Young *et al.*, 2015).

The C1R cell line (Storkus *et al.*, 1987; Zemmour *et al.*, 1992) is a B lymphoblastoid cell line deficient in MHC-class I. C1R-B*57:01 is transfected to express HLA-B*57:01. As flucloxacillin DILI is associated with the carriage of HLA-B*57:01, it is important to identify any protein binding that may occur within cells carrying the allele. The use of antigen presenting cells for this investigation will allow for the determination of the fate of modified proteins in subsequent chapters. Although in the context of this thesis the major aim is to elucidate the immune involvement in flucloxacillin induced DILI, non-immune mediated cholestatic liver injury is also believed to be triggered by flucloxacillin (Burban *et al.*, 2017). Non-immune cholestatic effects were described by Burnan *et al*, and were found to be mediated by the

activity of P38α. P38α (MAPK14) is a MAPK (mitogen-activated protein kinase) responsible for the induction of a number of signalling pathways through phosphorylation activity. Heat shock protein (HSP) 27 has known function in the Rho/Myosin light chain kinase (ROCK/MLCK) regulation of actin polymerization. Disruption of ROCK/MLCK signalling pathways is heavily implicated in cholestatic insults through cytoskeleton rearrangement and bile canaliculi deformations (Sharanek *et al.*, 2016). From the study performed by Burban *et al*, the activation of P38 is believed to result in the phosphorylation of HSP27. Once activated HSP27 acts as a chaperone, interacting with protein B kinase (AKT). Subsequent formation of a P38/HSP27/AKT protein kinase complex scaffold allows for the phosphorylation of AKT by phosphoinositide 3-kinase (PI3K), blocking apoptosis. Generally, this mechanism is characteristic of a protective phenotype allowing cells to survive periods of stress. However, in the case of flucloxacillin the blocking of apoptosis is related to the onset of cholestatic effects (Burban *et al.*, 2017).

3.2 AIMS

The purpose of this chapter was to further characterize the cellular protein targets for flucloxacillin modification, and their potential role in the progression of flucloxacillin induced liver injury. A range of proteomic techniques, utilising the antibody generated in chapter one, were used in relevant cell lines to assess both the distribution and localization of haptenated proteins. In order to investigate this, the following aims were addressed;

1. Using the anti-flucloxacillin antibody detect the level of flucloxacillin protein binding in C1R-B*57:01, HepG2, HepaRG cell lines and primary human hepatocytes using SDS-PAGE and Western blotting.
2. Investigate the localization of flucloxacillin modified proteins in C1R-B*57:01, HepG2 and HepaRG cells using immunofluorescence microscopy.
3. Assess the impact of flucloxacillin treatment on membrane transporter proteins using fluorescent substrates in fully differentiated HepaRG cells.
4. Using mass spectrometry, positively identify protein sources for modification from C1R-B*57:01 cells.
5. Identify cellular signalling pathways that may be responsible for non-immune induced liver cholestasis.

3.3 METHODS

3.3.1 PREVIOUSLY DESCRIBED METHODS

- Protein conjugation - 2.3.1, p85
- Protein quantification - 2.3.2, p85
- SDS-PAGE - 2.3.3, p86
    - Western blot analysis - 2.3.3.1, p86
    - Coomassie blue staining - 2.3.3.2, p87
- Mass spectrometry - 2.3.4, p87

3.3.2 TISSUE CELL CULTURE

All cells were grown in pre-treated (proprietary Nunclon™ Delta surface treatment) sterile cell culture flasks and multi-well plates purchased from Thermo Fisher (Nunc). Cell viability and quantification was performed using trypan blue (Sigma Aldrich) staining and light microscopy.

3.3.2.1 HEPG2 CELLS

HepG2 cells were cultured and maintained in DMEM supplemented with 10% FBS (foetal bovine serum) (v/v), 4 mM L-glutamine, 25 mM glucose, 1 mM sodium pyruvate and 1 mM HEPES (37°C, 5% CO_2). Cells were treated with flucloxacillin supplemented media. For growth on glass cover slips, rat tail collagen (Invitrogen) was used for pre-treatment.

3.3.2.2 HEPARG CELLS

HepaRG cells were cultured in growth media (Williams E (Sigma Aldrich) supplemented with 10% FBS (Invitrogen, Paisley, UK), 2mM L-glutamine, 100 µg/mL penicillin, 100 U/mL streptomycin, 5 µg/mL insulin and 50 µM hydrocortisone) (37°C, 5% CO_2) for 1.5 weeks. Cells were differentiated in 50:50 growth:differentiation media (growth media supplemented with 1.7% DMSO) for 0.5 weeks, and differentiation media for a further 4 weeks. Cells were

treated with flucloxacillin supplemented into differentiation media. For growth on glass cover slips, rat tail collagen (Invitrogen) was used for pre-treatment.

3.3.2.3 C1R-B*57:01 CELLS

C1R-B*57:01 cells were maintained in F1 media (RPMI 1640 supplemented with 10% FBS (Invitrogen, Paisley, UK), 100 mM L-glutamine, 1mM HEPES, 100 µg/mL penicillin, 100 U/mL streptomycin) and 50 µg/mL geneticin (37°C, 5% CO_2). Treatment with flucloxacillin was for 48h unless otherwise stated.

3.3.2.4 PRIMARY HUMAN HEPATOCYTES

Liver biopsies were obtained from Aintree Hospital, Liverpool, Merseyside from donors undergoing resections for varying aetiologies. All donors provided written informed consent to partake in the study which has received approval from the appropriate research ethics committees. Liver resections were perfused using 1x Hepes buffer to remove residual blood. Collagenase type IV (Sigma Aldrich) was used to digest the tissue releasing hepatocytes. Hepatocytes were cultured in Williams E supplemented by 2 mM L-glutamine, 100 µg/mL penicillin, 100 U/ml streptomycin, 100x insulin-transferin-selenium, and 1 µM/ml dexamethasone. After 3 hours cells were washed to remove unattached cells. After a further 24 hours (37°C, 5% CO2) media was removed and replaced with flucloxacillin supplemented media. Cells were harvested after 16-24 hours of drug incubation.

3.3.3 PROTEIN EXTRACTION FROM CELL CULTURE

Cells were harvested and washed prior to being pelleted and snap frozen. Cell pellets were lysed (7.0M urea, 2.0M thiourea, 4% CHAPS, 40mM Tris base and 1% DTT) and cell supernatants were collected for protein quantification.

3.3.4 2D SDS-PAGE

One hundred micrograms of protein from C1R-B*57:01 cell lysates were separated in two dimensions. The first dimension was performed by rehydrating IPG strips (Immobiline™

DryStrip, Amersham Pharmacia Biotech) with sample in rehydration solution and separating based on pH using the Multiphor Electrophoresis System (GE Healthcare, MA, USA). The second dimension and Western blot was performed as previously described in Chapter 2.

3.3.5 PATIENT SERUM

Patients receiving flucloxacillin treatment (i.v. and/or oral) were recruited. Ethical approval was obtained from Liverpool local research ethics committee and each patient gave informed consent to participate in the study.

3.3.5.1 ISOLATION

Venepuncture samples were extracted into heparinised tubes at least 8 h post treatment. Samples were centrifuged at 2,000 x *g* at 4°C for 15 min and stored at -80°C.

3.3.5.2 HSA DEPLETION

Serum samples were depleted by HSA affinity chromatography (#5188-6562, Agilent Technologies) on the Agilent 1200 HPLC. PBS was flushed through the column to remove unbound serum proteins before elution in 12 mM HCl and neutralization using Tris (pH 9). Buffer exchange into PBS was performed using a 3 kDa MWCO filter (Merck Millipore) for subsequent SDS-PAGE analysis.

3.3.6 IMMUNOFLUORESCENCE MICROSCOPY

Adherent cells were grown on glass cover slips for use with immunofluorescence microscopy. C1R-B*57:01 cells were cultured in the presence of flucloxacillin and adhered to glass cover slips using Cell-Tak (Corning). Cells were washed with PBS (pH 8.0) and fixed using 4% paraformaldehyde for 30 minutes at 4°C. Cells were permeabilized (0.004% Tween 20, 0.025% Triton-X-100, PBS) for 30 minutes and blocked with BSA (Sigma Aldrich) (5% in permeablilizaton buffer) for 1 hour at room temperature. Subsequently blocking buffer containing polyclonal rabbit-anti-flucloxacillin antibody (1 in 2,000 dilution) was added overnight at 4°C. After washing with permeablilizaton buffer, goat anti-Rabbit IgG secondary

antibody (Alexa Fluor 488, ThermoScientific, 1 in 1,000 dilution) was applied for one hour in the dark. After further washes, the cells were incubated in hoescht (DAPI, ThermoScientific, 1 in 5,000 dilution) and phalloidin (Alexa Fluor 568, ThermoScientific, 1 in 250 dilution) for nuclear and f-actin staining respectively. Cover slips were mounted onto glass microscope slides with Pro-Long Gold (ThermoScientific, MA, USA) and sealed. Images were taken using a Carl Zeiss Axio Observer microscope with Apoptome using a 5x and a 40x oil objective. For detection of MRP2 and P-gp, 1 in 200 dilutions of primary antibody were used (ProteinTech, MRP2; 24893-1-AP, P-gp; 22336-1-AP).

3.3.6.1 CMFDA TRANSPORTER FUNCTION ASSAY

HepaRG cells were cultured to differentiation as described (3.3.2.2) in a Lumox (Sarstedt) 24 well plate. Media was removed and replaced with uptake media (136 mM NaCl, 5.2 mM KCl, 1.1 mM KH_2PO_4, 0.7 mM $MgSO_4$, 2.3 mM $CaCl_2$, 10 mM Hepes and 11 mM glucose) (pH 7.4) containing 3 µM CMFDA solution (Invitrogen). CMDFA was incubated for 30 minutes at 37°C before removal and washing with uptake media. Cells were imaged as previously described directly through the plate using a 5x objective. For images taken using the 40x objective, HepaRG cells were cultured on collagen coated glass cover slips and processed as previously described for transporter function and immunofluorescence microscopy.

3.3.6.2 MRP2/P-GP MEMBRANE TRANSPORTER INHIBITION

HepaRG cells were cultured to differentiation as described. Prior to the addition of flucloxacillin, wells were incubated with 30 µM MK571 and 12.5 µM valspodar for a minimum of 1 hour to block MRP2 and P-gp activity respectively. MK571 and valspodar were also added to media supplemented with flucloxacillin. Images were acquired as previously described.

3.3.7 COOMASIE STAINED IN-GEL PROTEIN DIGESTION

Gel bands were excised and placed into clean protein lo-bind Eppendorf tubes. To the gel pieces, 200 µL 25 mM NH_4HCO_3 in H_2O/ACN (1:1, v/v) was added. Gel pieces were incubated at RT for 10 minutes while shaking. NH_4HCO_3 was removed and replaced with 200 µL ACN to dehydrate the gel pieces. After 5 minutes (RT, shaking) this step was repeated. After removing the ACN, 50 µL 0.1 M NH_4HCO_3 was added and gel pieces were incubated for 30 minutes at 37°C. Gel pieces were resuspended in 50 µL 10 mM ammonium bicarbonate and incubated overnight with sequencing grade modified trypsin at 37°C (Promega). Supernatant was collected and transferred to a clean protein lo-bind Eppendorf tube. To the gel piece, 200 µL of 3% acetic acid in H_2O/ACN (1:1, v/v) was added and sonicated at 37°C in a water bath for 30 minutes. Supernatant was combined and samples concentrated by vacuum centrifugation at 37°C. Samples were resuspended in 0.1% trifluoroacetic acid (TFA) and purified using C18 ZipTips (Millipore) following manufacturer's instructions. Samples were subsequently resuspended in 11 µL 0.1% FA and 5 µL was injected into the mass spectrometer (Sciex 6600 triple TOF) using previously described parameters (Chapter 2, 3.4).

3.4 RESULTS

3.4.1 DIVERSITY OF FLUCLOXACILLIN PROTEIN BINDING

3.4.1.1 HEPARG AND HEPG2 CELL LINES

Investigation into the diversity of flucloxacillin protein binding was performed in the liver-like cell lines HepG2 and HepaRG. Both cell lines were incubated in the presence of different concentrations of flucloxacillin at non-toxic doses. After cell lysis, protein isolation and quantification were performed and proteins were visualised using coomassie gel staining (Figure 3.2, left). Across all cell culture conditions, the protein quantity and diversity were comparable. Using the same cell lysates, Western blot was performed to detect the level of flucloxacillin protein binding (Figure 3.2, right). Proteins from both HepG2 and HepaRG cells underwent modification at 1 mM flucloxacillin concentrations. Proteins contained within HepaRG cells were more sensitive to modification by flucloxacillin, as binding was observed qualitatively in larger quantities and a wider diversity at lower drug concentrations.

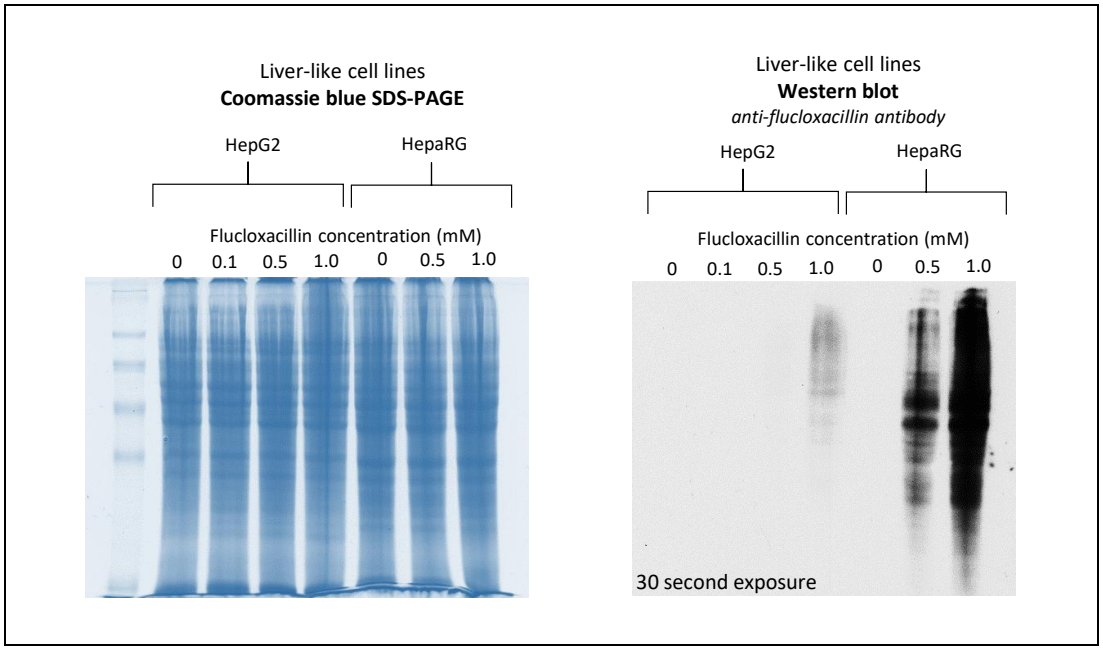


Figure 3.2. SDS-PAGE coomassie blue and Western blot of liver like cell lines HepG2 and HepaRG. Coomassie blue SDS-PAGE shows no alteration in protein abundance and equal loading across different flucloxacillin treatments in HepG2 and HepaRG cell lines. Western blot analysis using anti-flucloxacillin antibody reveals flucloxacillin protein binding in both cell lines, in a dose dependant manner. Proteins extracted from HepaRG cells are modified by flucloxacillin at lower concentrations compared with HepG2 cells.

3.4.1.2 PRIMARY HUMAN HEPATOCYTES

In addition to liver like cell lines, primary human hepatocytes from a patient (pt) donor (pt 196) was cultured in the presence of flucloxacillin for 16 hours. In-keeping with liver like cell lines, the protein abundance in primary human hepatocytes did not change in the presence of flucloxacillin, as determined by coomassie SDS-PAGE (Figure 3.3, top left). Western blot analysis of patient hepatocytes (Figure 3.3, top right) showed the modification of multiple proteins in a dose dependent manner.

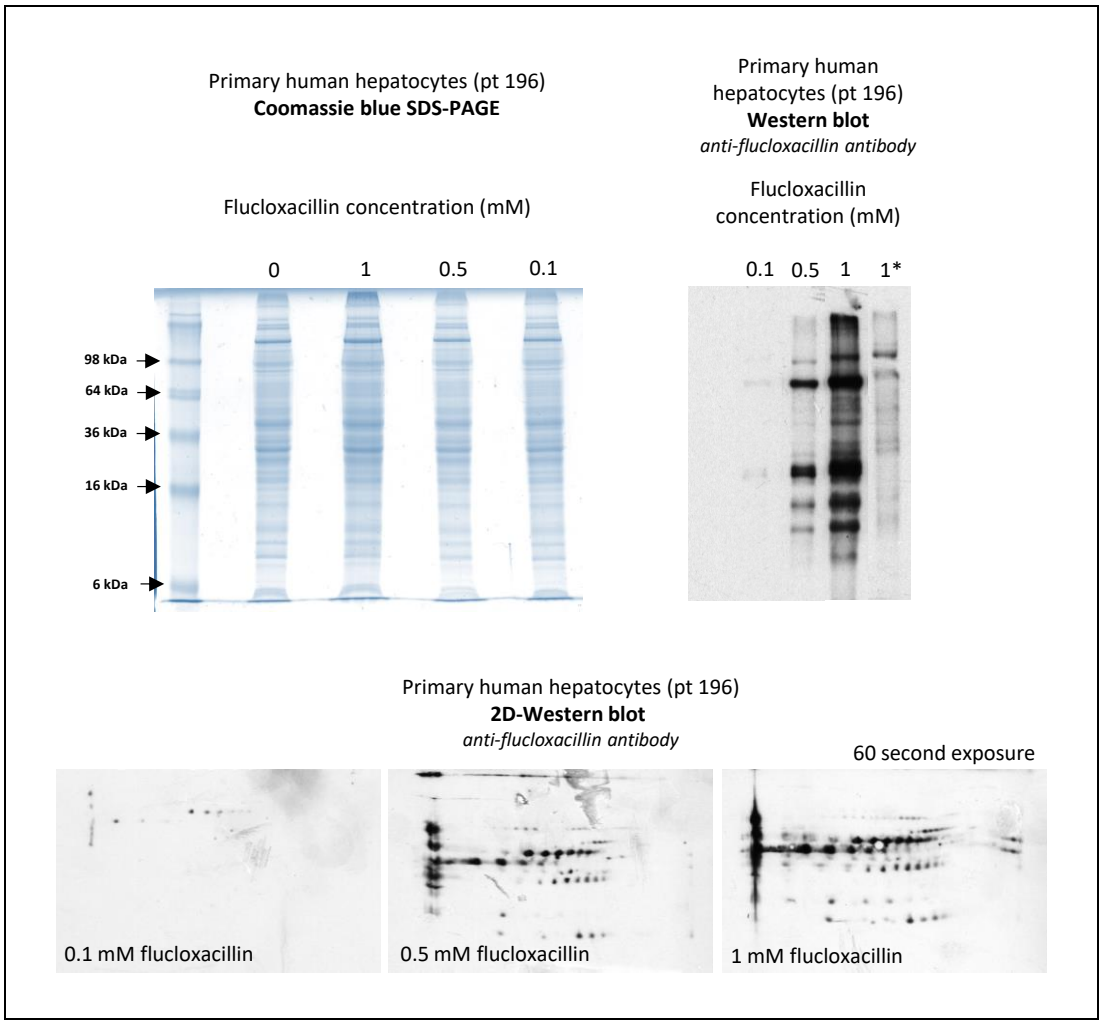


Figure 3.3. SDS-PAGE coomassie blue and Western blot of primary human hepatocytes. Primary human hepatocytes were incubated with different concentrations of flucloxacillin for 16 hours. Coomassie blue staining shows no change in protein abundances. A dose dependent increase in flucloxacillin protein binding was detected using anti-flucloxacillin antibody in primary human hepatocytes. Two-dimensional Western blot analysis reveals flucloxacillin modification of multiple proteins in a dose dependent manner. *cells were incubated with flucloxacillin for 1 hour prior to cell lysis.

Modification of proteins extracted from primary hepatocytes was observed after a 1-hour incubation with flucloxacillin in cells isolated from pt 196. This highlighted the speed at which modification occurs. No cross reactivity was observed between the anti-flucloxacillin antibody produced in house and non-modified cellular proteins. In addition to single dimension Western blot, 2D SDS-PAGE and Western blot was performed (Figure 3.3, bottom) to further highlight the range of proteins that can be modified by flucloxacillin.

3.4.1.3 C1R-B*57:01 CELLS

Two-dimensional SDS-PAGE was performed on C1R-B*57:01 cell lysates cultured in the presence of flucloxacillin. Coomassie blue staining (Figure 4.1, top) indicated no difference in the proteome after a 24-hour incubation with flucloxacillin at non-toxic concentrations.

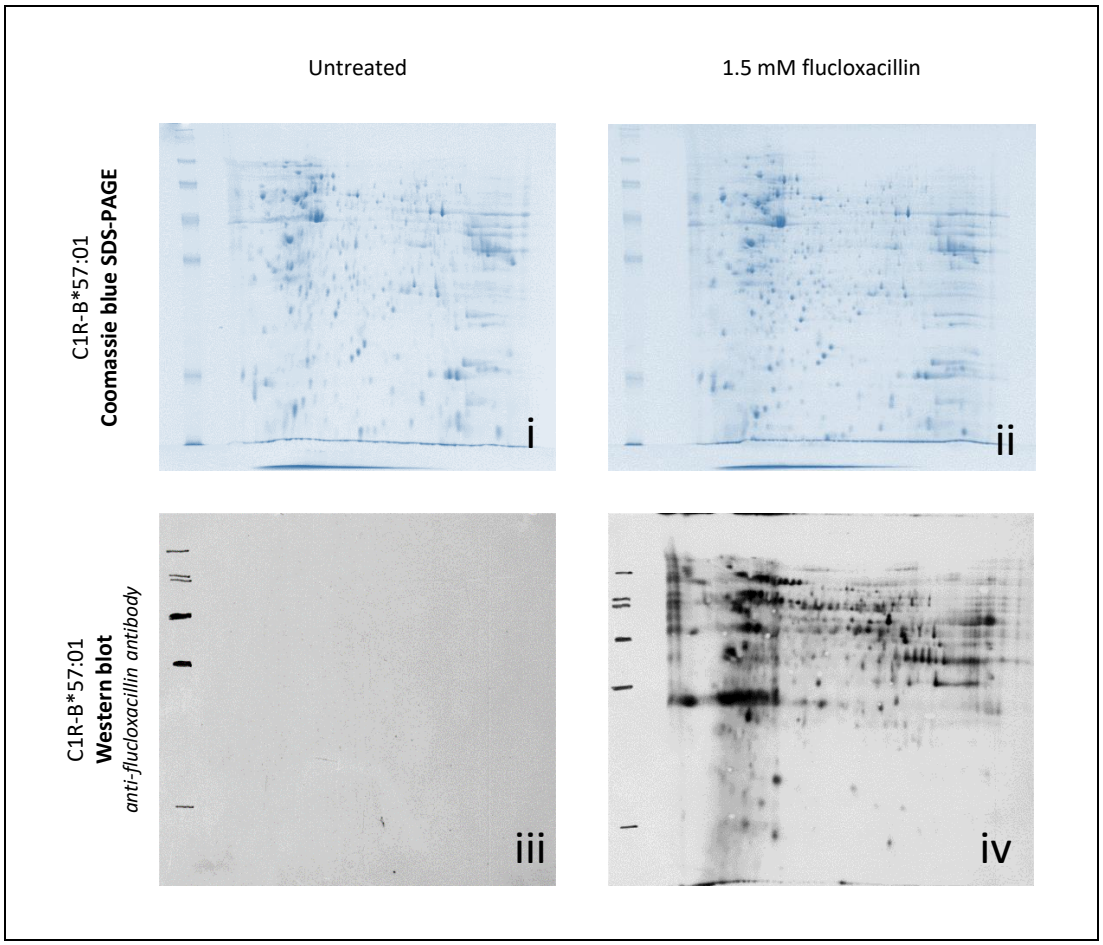


Figure 3.4. Two-dimensional SDS-PAGE coomassie blue and Western blot of C1R-B*57:01 cells. Two-dimensional coomassie blue staining revels no change in the proteome of C1R-B*57:01 cells when incubated in the presence of flucloxacillin (i vs ii). Western blot analysis revels the modification of a wide range of proteins in the treated sample (iv), with no cross reactivity in the untreated control (iii).

Western blot analysis was performed to characterize the diversity of flucloxacillin protein binding in these antigen presenting cells. Modification was found on a wide range of proteins with the antibody maintaining a high specificity for flucloxacillin (Figure 3.4, bottom).

3.4.1.4 PATIENT SERUM

Serum isolated from patients undergoing treatment with flucloxacillin was prepared for SDS-PAGE. An aliquot from each patient (pt 50, 51 and 64) along with control serum was taken and the HSA captured, removing all other serum proteins. Patients 50 and 51 had taken 2 g of flucloxacillin 3 and 4 times a day, respectively, intravenously for 10 days. Dosing information for patient 64 was not available.

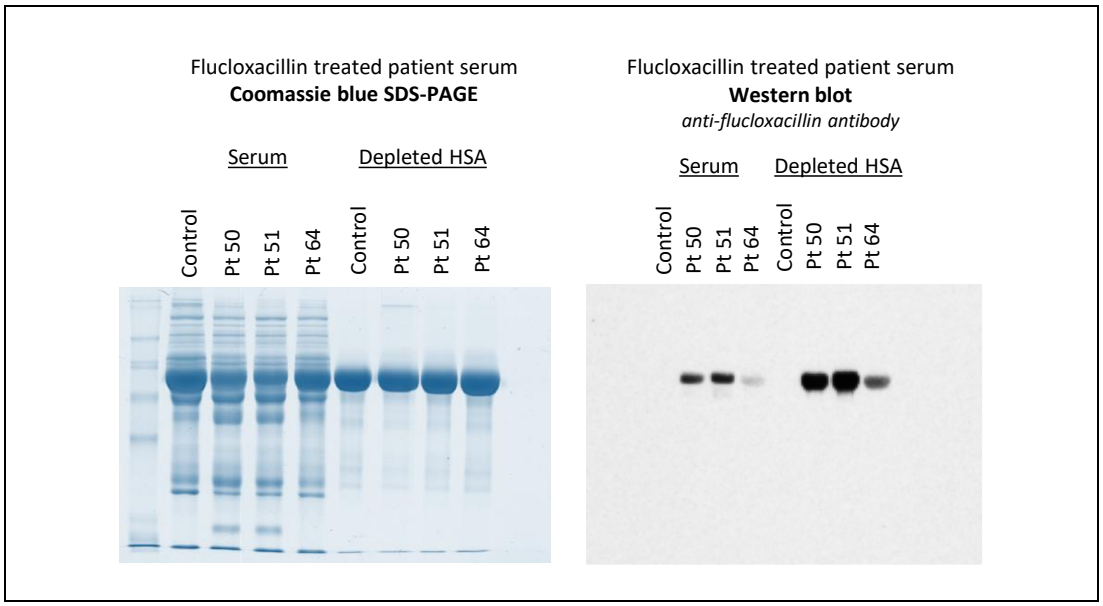


Figure 3.5. SDS-PAGE coomassie blue and Western blot of flucloxacillin patient serum and depleted HSA. Serum from patients taking flucloxacillin were HSA depleted. Coomassie blue was used to visualize the successful depletion of HSA from the serum protein. Western blot analysis reveals only HSA is modified in patient serum, with increased signal from the depleted HSA. Control patient serum shows no cross reactivity.

Whole serum and depleted HSA from each sample was separated using SDS-PAGE and stained using coomassie blue (Figure 3.5, left). Western blot analysis using the anti-flucloxacillin antibody (Figure 3.5, right) detected only HSA to be modified in all serum samples isolated from patients undergoing flucloxacillin treatment.

3.4.2 CHARACTERIZATION OF FLUCLOXACILLIN PROTEIN BINDING

A wide range of flucloxacillin modified proteins were detected using anti-flucloxacillin antibody from a range of cell types. Therefore, C1R-B*57:01 cells incubated with flucloxacillin were prepared for mass spectrometric analysis to allow the characterization of modified proteins. Due to the limitations with software in identifying flucloxacillin modified peptides (discussed in chapter 6) manual characterization was performed. One protein, 14-3-3 gamma, was found to be modified at lys-50 (NLLSVAYK[Flucloxacillin]NVVGAR) (in relatively high abundance (Figure 3.7A). The 14-3-3 family consists of a number of isoforms, including gamma, epsilon and theta, each involved in regulatory processes (Zerr *et al.*, 1998). Compared with 14-3-3 gamma, epsilon contains a valine to isoleucine substitution (NLLSVAYKNV**I**GAR) while theta contains an arginine to glycine substitution (NLLSVAYKNVVG**G**R) (Figure 3.7B). Indeed, modification was also found on 14-3-3 epsilon (Figure 3.8A) and theta (Figure 3.8B). Interestingly, modification of the lysine residue contained within this peptide sequence was observed for all three isoforms, indicating preferential binding to a site of the protein (Figure 3.6).

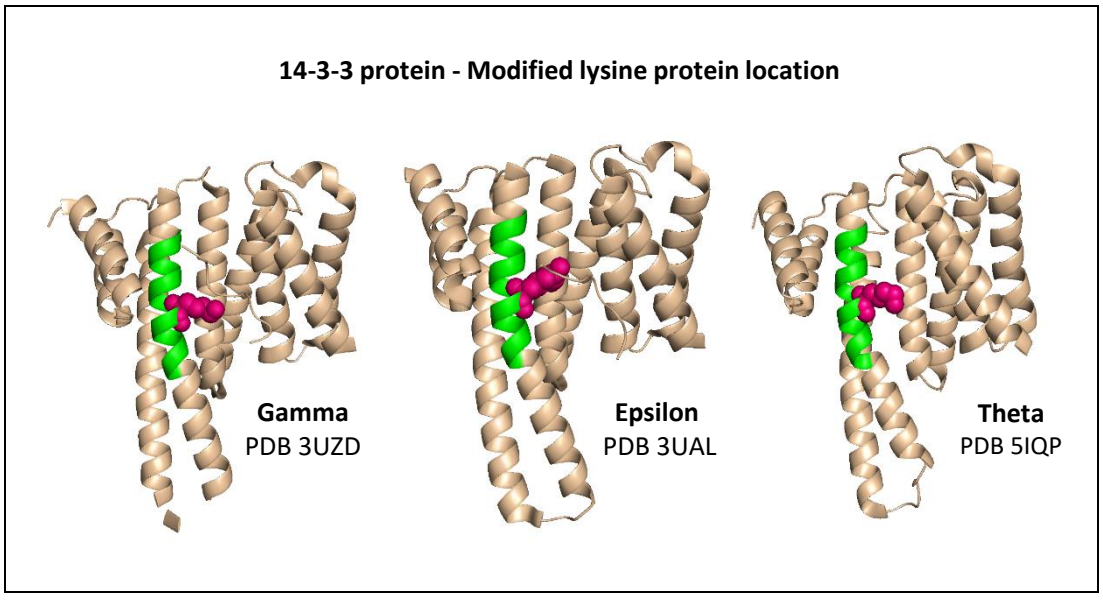


Figure 3.6. Flucloxacillin modification of 14-3-3 protein. The same region of 14-3-3 gamma, epsilon and theta (green) were modified at the lysine residue (pink).

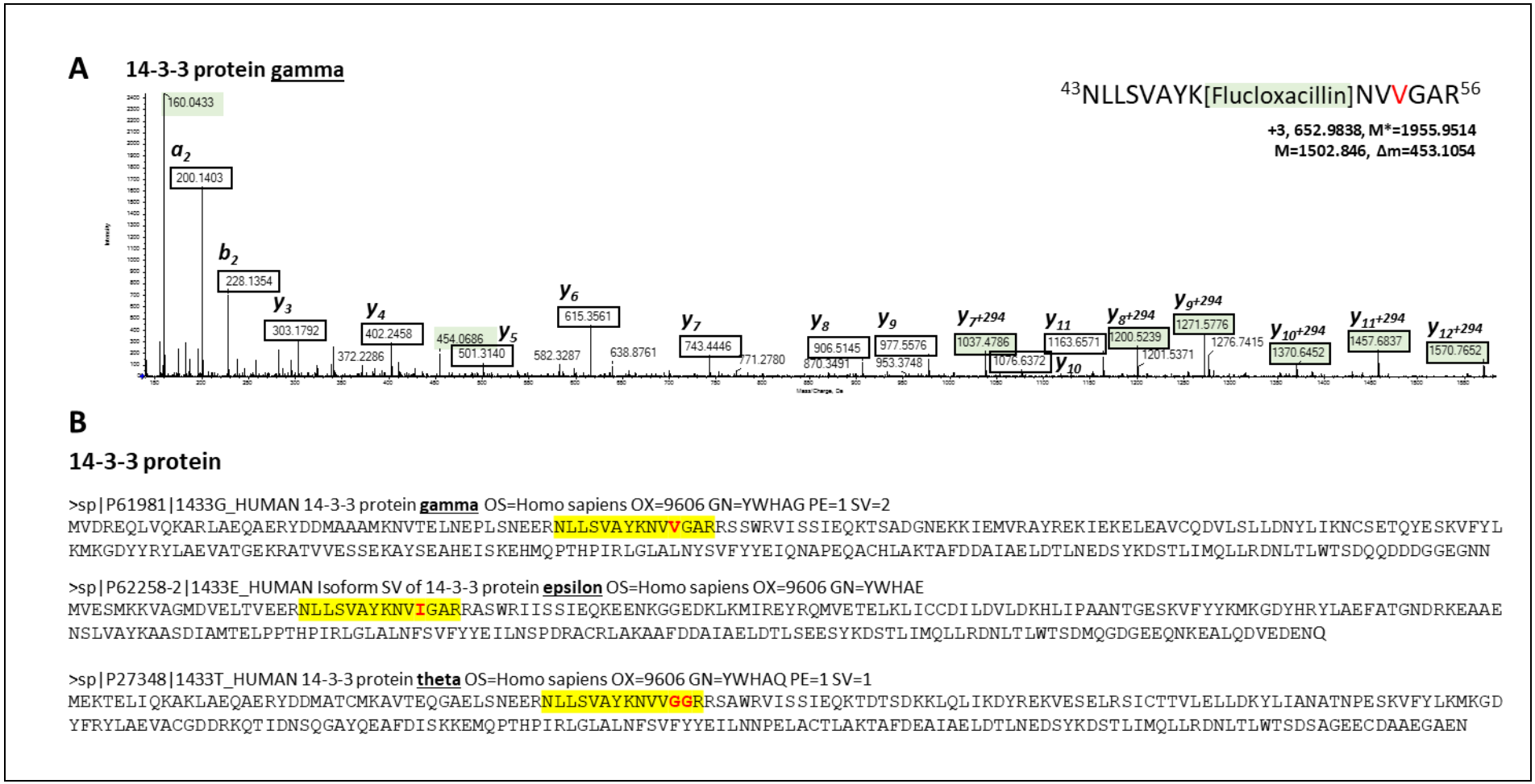


Figure 3.7. Flucloxacillin modification of 14-3-3 protein. C1R-B*57:01 cells incubated with flucloxacillin were lysed and prepared for mass spectrometric analysis. Manual characterization reveals adduct formation on (A) 14-3-3 gamma (NLLSVAYK[Flucloxacillin]NVVGAR). Multiple isoforms of 14-3-3 exist, including (B) gamma, epsilon and theta.

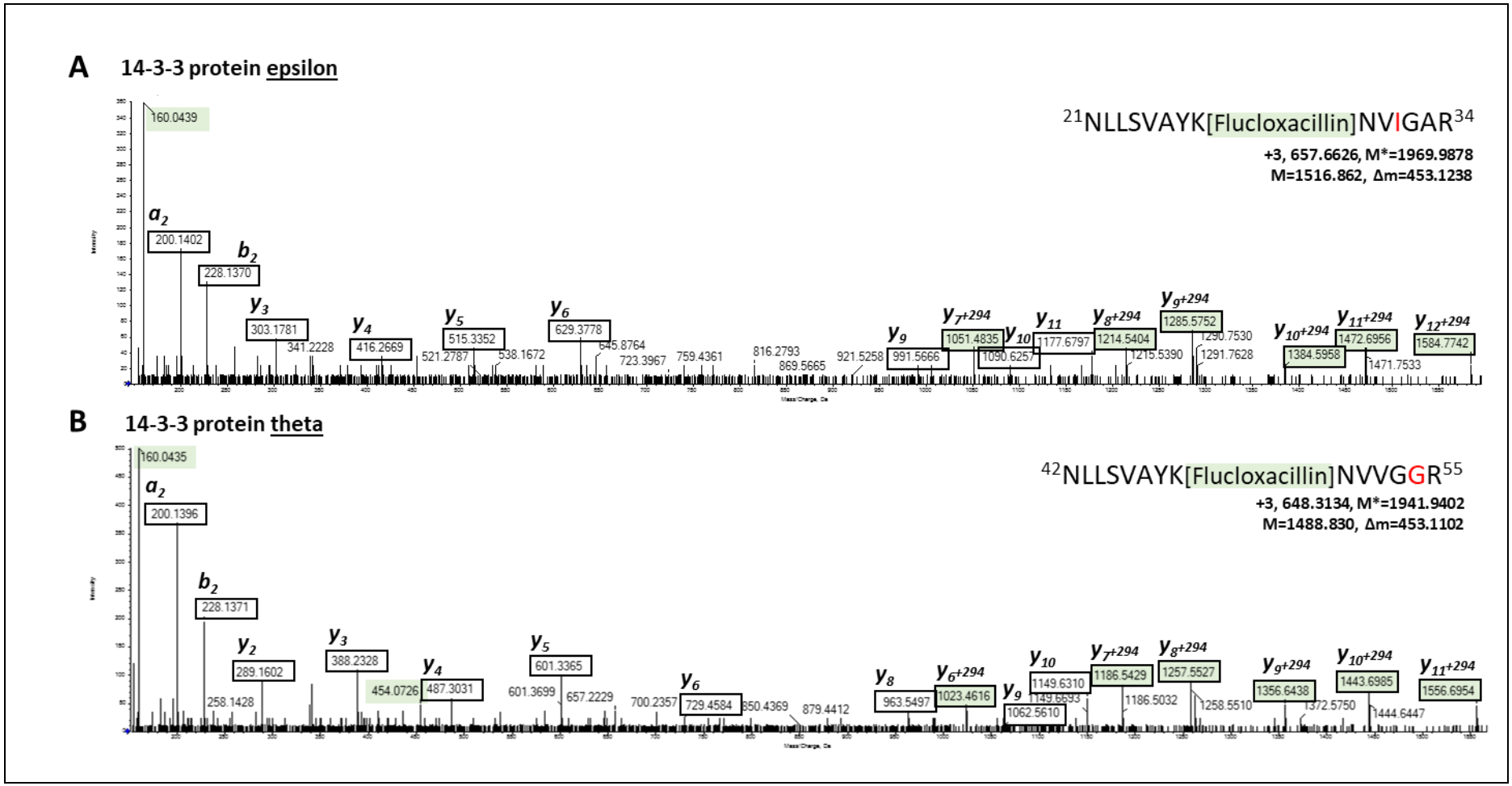


Figure 3.8. Flucloxacillin modification of 14-3-3 protein. C1R-B*57:01 cells incubated with flucloxacillin were lysed and prepared for mass spectrometric analysis. Manual characterization reveals adduct formation on (A) 14-3-3 epsilon (NLLSVAYK[Flucloxacillin]NVIGAR) and (B) theta (NLLSVAYK[Flucloxacillin]NVVGGR).

3.4.3 LOCALIZATION OF FLUCLOXACILLIN PROTEIN BINDING

Immunocytochemistry was used to visualize flucloxacillin binding within cells. Importantly, images comparing different conditions were acquired using the same exposure times throughout this chapter.

3.4.3.1 C1R-B*57:01 CELLS

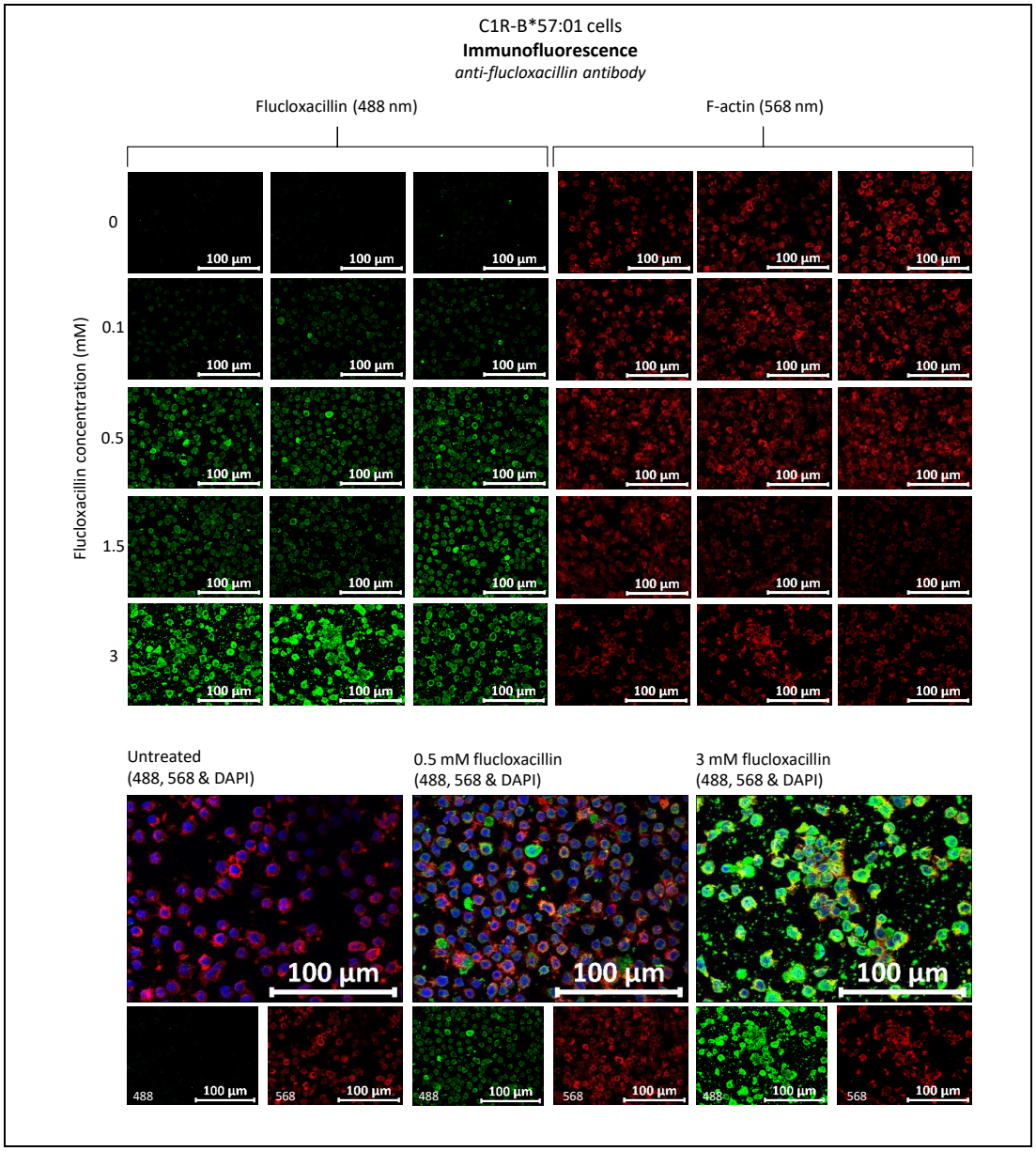


Figure 3.9. Immunocytochemistry of C1R-B*57:01 cells treated with flucloxacillin. Anti-flucloxacillin antibody was used to visualize flucloxacillin protein modification in C1R-B*57:01 cells in a dose dependent manner. Multiple images represent technical replicates from the same slide. Flucloxacillin (488 nm, green) is observed at all concentrations (0.1 - 3 mM) while untreated cells showed no signal. F-actin (568 nm, red) is consistent between all concentrations. Overlaid images show the localisation of flucloxacillin in relation to F-actin and the nucleus (DAPI, blue). A dose response can be observed up to 3 mM, where drug modified cell debris can be observed at a toxic concentration.

Modification of proteins by flucloxacillin was observed in a dose dependent manner in C1R-B*57:01 cells treated at a range of non-toxic (0.1, 0.5 and 1.5 mM) and toxic (3 mM) concentrations (Figure 3.9, top, green). F-actin (red) and nuclear staining (blue) appear relatively consistent at all concentrations, while no modification was observed in untreated cells. At toxic concentrations (3 mM), what is thought to be flucloxacillin modified cell debris is also observed (Figure 3.9, bottom, right). Immuno-fluorescence was also used to identify the localisation of flucloxacillin binding. In C1R-B*57:01 cells the green signal is maintained throughout the entire cell indicating the modification of intracellular proteins (Figure 3.10). In a few cases a higher intensity of flucloxacillin is observed at the cell surface of C1R-B*57:01 cells (Figure 3.10, white arrows). This is attributed to the modification of surface/membrane proteins or peptides presented by HLA-B*57:01.

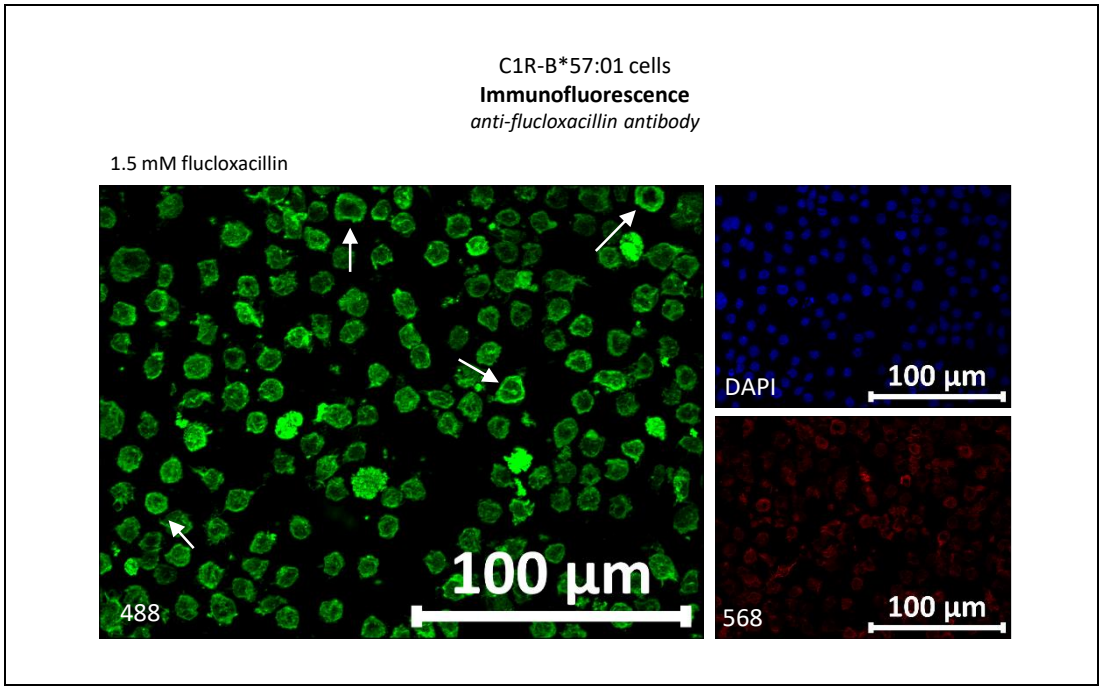


Figure 3.10. Localization of flucloxacillin in treated C1R-B*57:01 cells. Flucloxacillin appears to modify intracellular proteins. Cell membrane proteins are also believed to be modified due to increased signal on the cell surface at 488 nM (green, white arrows).

3.4.3.2 HepG2 and HepaRG cells

The diversity and localisation of flucloxacillin protein binding was also assessed in the liver like cell lines HepG2 and HepaRG. The modification of HepG2 cells is largely on the cell

surface (Figure 3.11) with limited flucloxacillin detected internally within the borders of the F-actin cytoskeleton. This surface modification is observed throughout all the Z planes of the image (Figure 3.11, 488 channel).

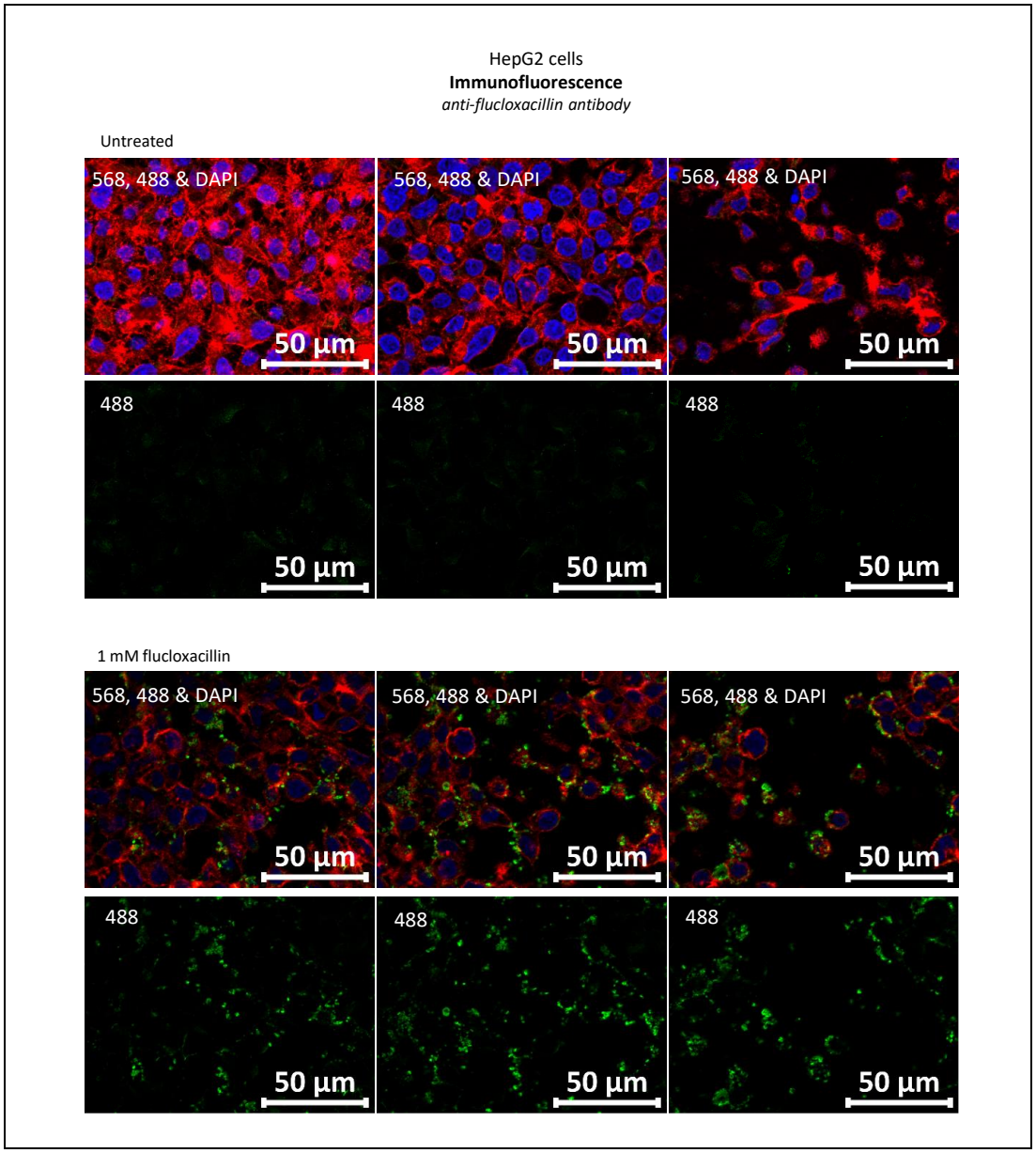


Figure 3.11. Flucloxacillin binding and localization in HepG2 cells using immunocytochemistry. Flucloxacillin mainly modifies HepG2 membrane proteins as limited signal is observed within the cell cytosol. Extracellular modification is observed in all layers of the Z-stacks. No flucloxacillin signal (488 nm, green) is observed in untreated cells.

In HepaRG cells, flucloxacillin modification of intracellular proteins is observed in a dose dependant manner (Figure 3.12).

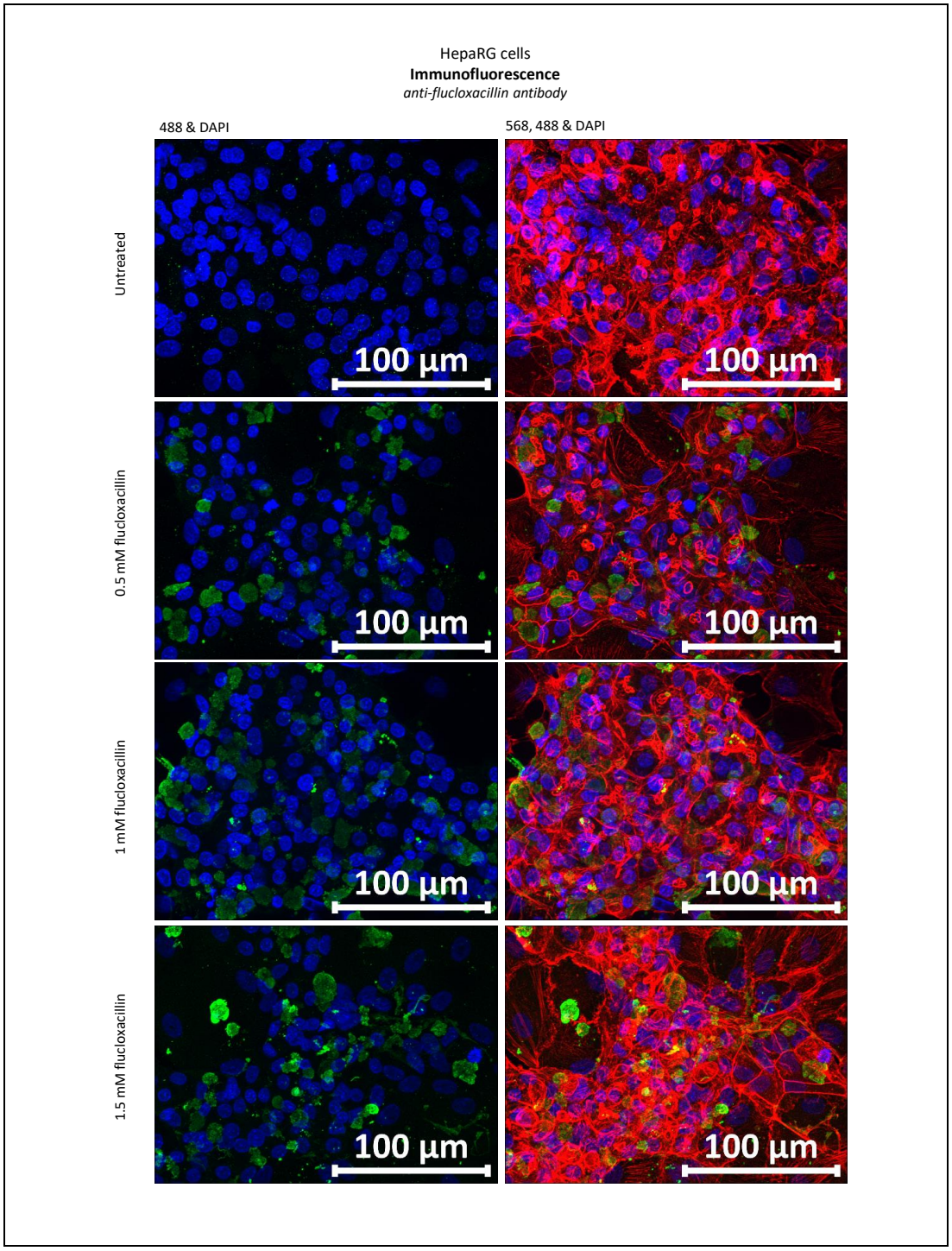


Figure 3.12. Flucloxacillin binding in HepaRG cells using immunocytochemistry. Flucloxacillin binding is observed at all concentrations in a dose dependant manner. Nuclear staining (DAPI, blue) and flucloxacillin (488 nm, green) are shown in the absence (left) and presence (right) of F-actin stain (568 nm, red).

Upon further examination, high intensities of flucloxacillin signal appear within tight actin bundles (Figure 3.13, white arrows); indicative of the bile canaliculi formed between multiple hepatocyte like cells. Localisation of flucloxacillin in the bile canaliculi provides some

evidence for the active transport of flucloxacillin out of the hepatocyte like cells by membrane transporters.

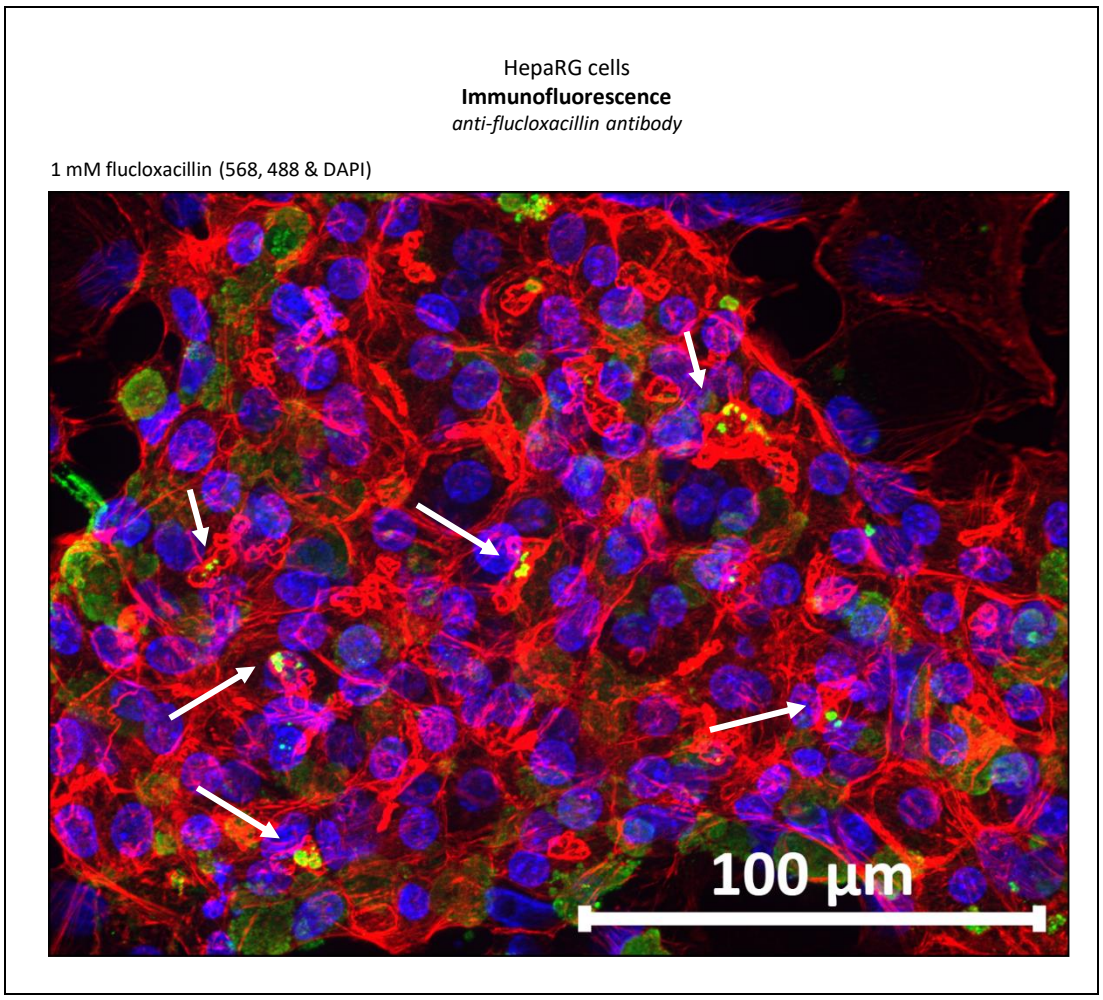


Figure 3.13. Localization of flucloxacillin binding in HepaRG cells. Flucloxacillin appears to be localized within tight actin bundles (white arrows) between groups of hepatocyte-like cells.

3.4.4 CELL MEMBRANE TRANSPORTERS

3.4.4.1 MRP2 AND P-GP EXPRESSION

In order to confirm the presence of functional biliary canalicul, immunofluorescence was performed using anti-MRP2 and anti-P-gp antibody. MRP2 and P-gp were found to be present at comparable levels in untreated cells (Figure 3.14). As anticipated, MRP2 and P-gp both localised within the tight actin bundles, as observed with flucloxacillin (Figure 3.13).

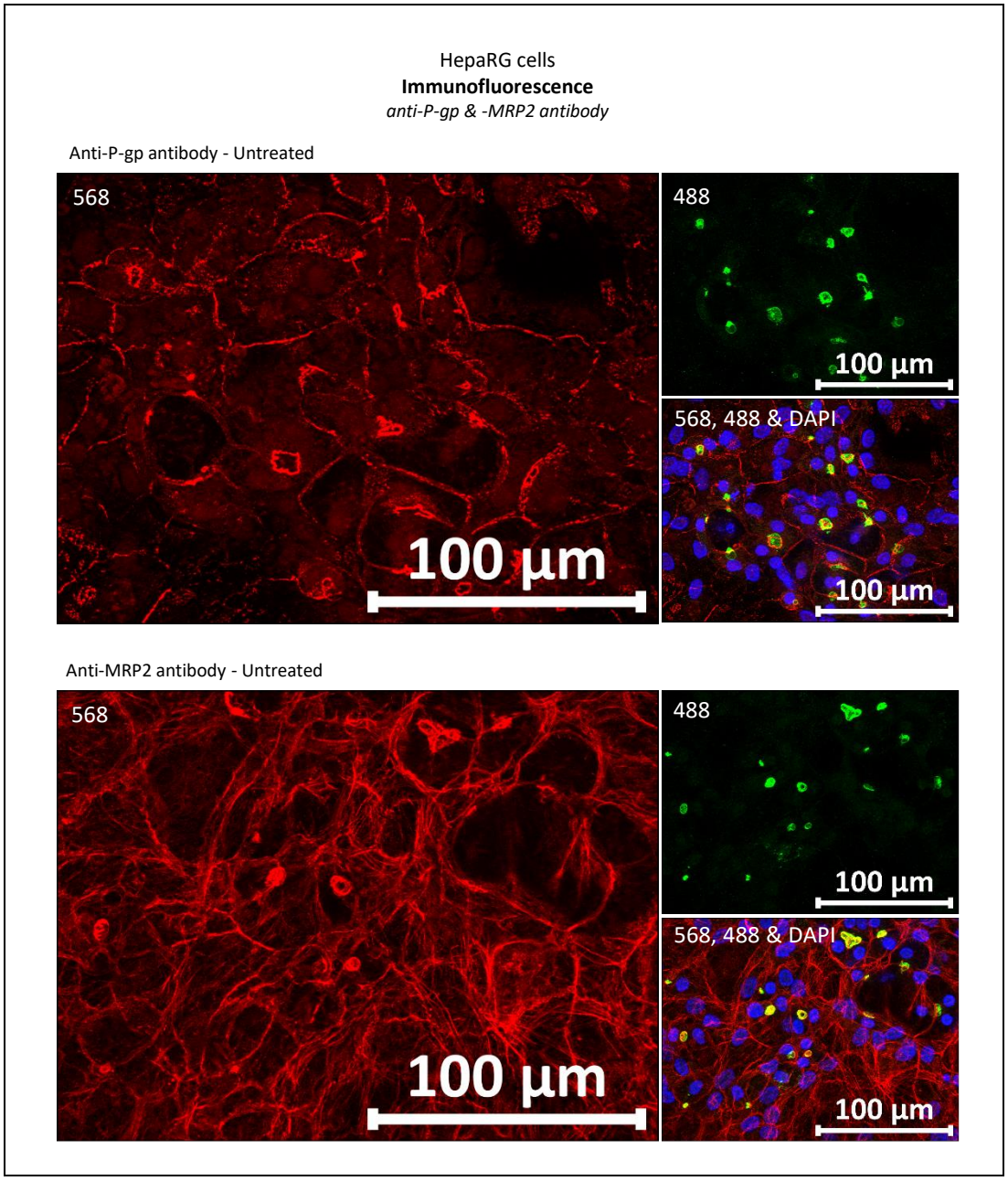


Figure 3.14. Detection of P-gp and MRP2 membrane transporters in HepaRG cells using immunocytochemistry. Anti-P-gp and anti-MRP2 antibodies were used to detect the expression of P-gp and MRP2 in HepaRG cells. P-gp and MRP2 expression (488 nm, green) was localised within the tight actin bundles (568nm, red).

MRP2 expression was investigated when HepaRG cells were incubated with flucloxacillin for an increasing period of time, with fresh flucloxacillin supplemented media replaced each day (Figure 3.15). Untreated cells express MRP2 at the tight actin bundles as shown in the previous experiment. After 24h 1.5 mM treatment, bile canaliculi become dilated, with further dilation seen after 7 days treatment. Interestingly, after 24h treatment MRP2

appears to be present in the cell cytoplasm (Figure 3.15, white arrows), while at 7 days this intensity decreases. This may be indicative of MRP2 being expressed intracellularly prior to localisation to biliary epithelia. Alternatively, flucloxacillin may be causing relacalisation of MRP2, which may lead to cholestasis.

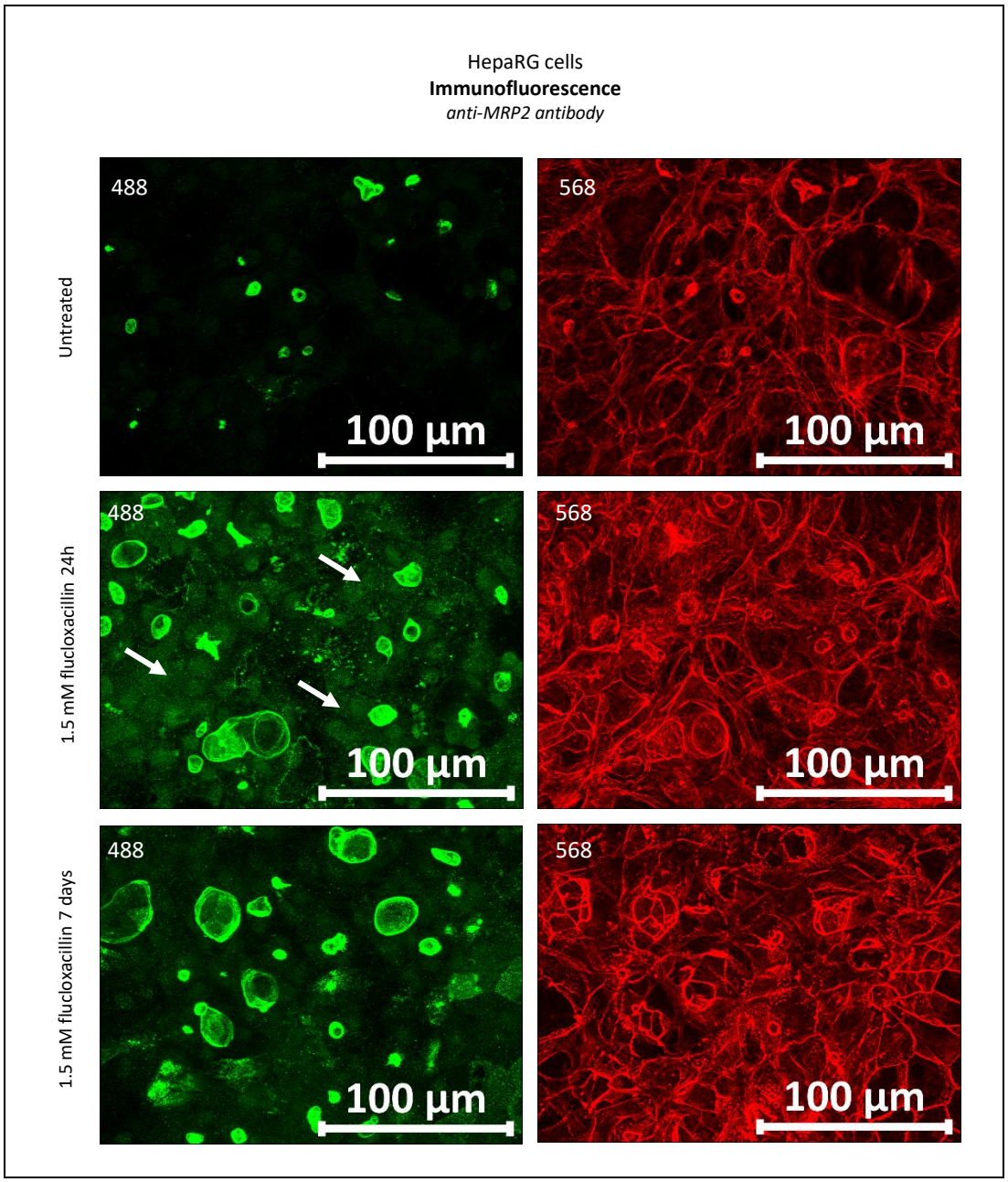


Figure 3.15. Biliary canalicular dilation in response to flucloxacillin treatment. Anti-MRP2 antibody was used to detect the expression MRP2 in HepaRG cells to identify biliary cells. Bile canaliculi were shown to become dilated after prolonged flucloxacillin exposure. MRP2 appears to be within the cytoplasm of hepatocyte like cells (488 nm, green) after 24h flucloxacillin treatment, likely due to MRP2 transcription prior to localisation.

3.4.4.2 MRP2 AND PGP FUNCTIONAL ACTIVITY

As the distribution of MRP2 changed across untreated and flucloxacillin treated HepaRG cells, the function of the MRP2 membrane transporter was assessed. CMFDA was applied to the cells for 30 minutes, where upon passive diffusion into the cell is transformed into an impermeable fluorescent product. MRP2 is responsible for organic anion secretion into the bile canaliculi, and can transport the impermeable fluorescent dye. Accumulation of CMFDA in the cell cytoplasm indicates little or no MRP2 expression. Fluorescence was retained within the epithelial cells of the HepaRG culture (Figure 3.16, top, left), confirming no activity of MRP2 in these cells.

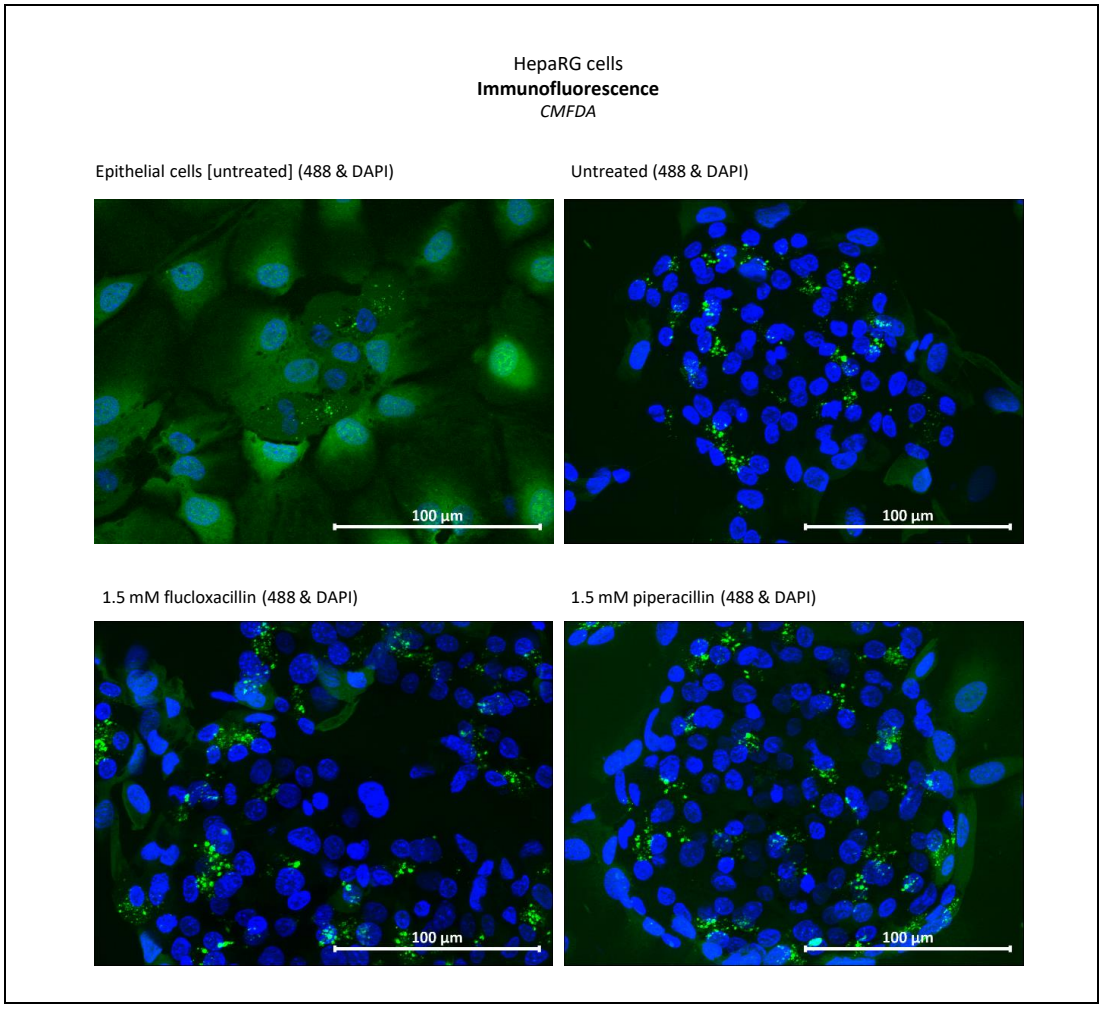


Figure 3.16. Assessment of membrane transporter function in HepaRG cells. CMFDA was used to identify active transport out of HepaRG cells. Epithelial cells (top, left) do not express MRP2, therefore CMFDA (488, green) is maintained within the cell cytosol. HepaRG cells actively transport CMFDA out of the cell cytoplasm into bile canaliculi (top, right). Flucloxacillin and piperacillin did not alter membrane MRP2 activity after 24 h treatment.

In the HepaRG hepatocellular-like cells CMFDA was transported out of the cell cytoplasm, localising in the bile canaliculi. HepaRG cells were treated with flucloxacillin and piperacillin for 24 hours prior to the addition of CMFDA. Piperacillin was used to assess the impact of a different β-lactam antibiotic. Between both drug treatments and the untreated control, hepatocellular-like cells did not display any changes in CMFDA transport (Figure 3.16).

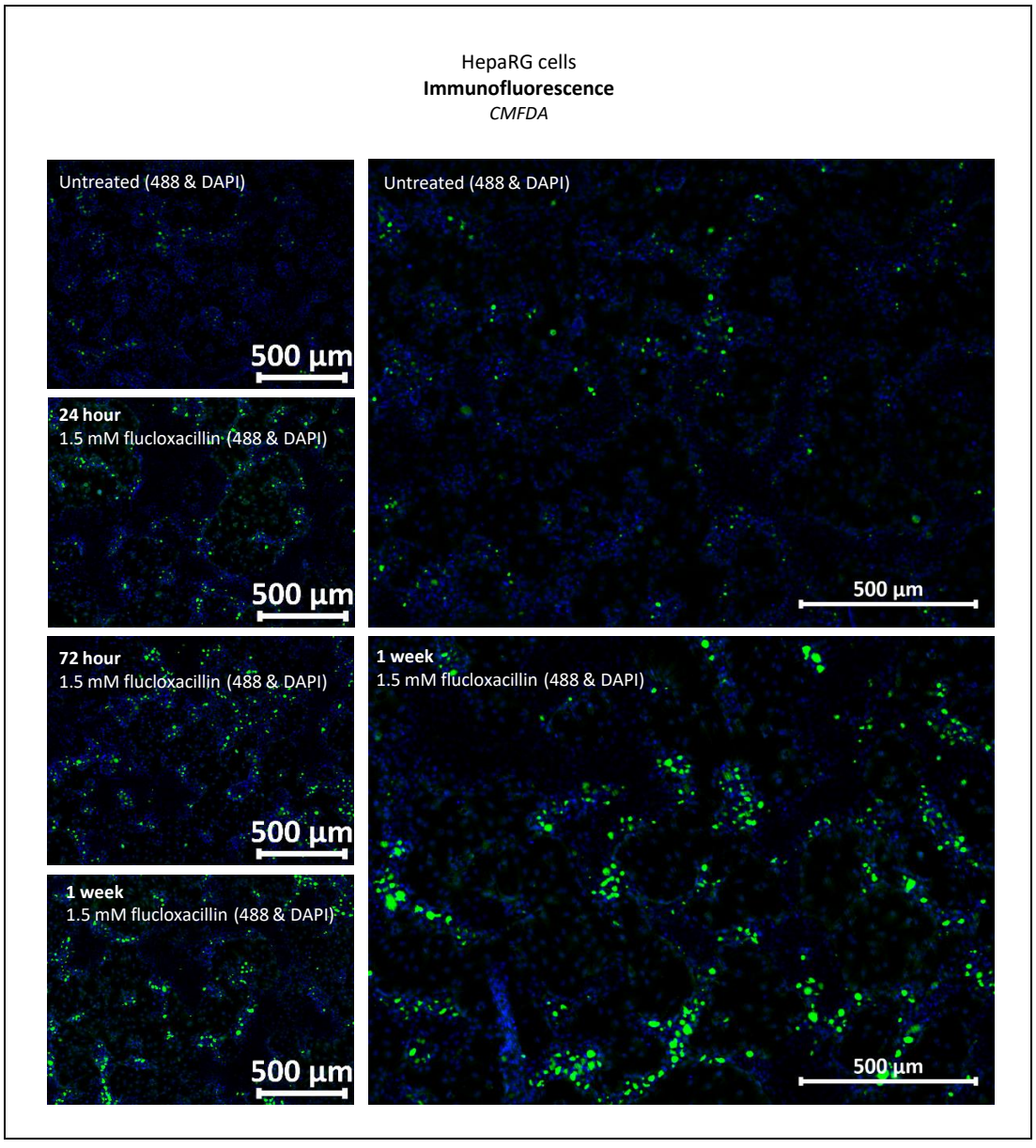


Figure 3.17. Assessment of membrane transporter function in HepaRG cells after prolonged flucloxacillin exposure. CMFDA localization in the bile canaliculi region of HepaRG cells increases after prolonged exposure to flucloxacillin (left, top to bottom). The intensity of CMFDA (488 nm, green) increases greatly when HepaRG cells are incubated with flucloxacillin for 1 week. Cells were observed using a 5x objective to look at overall CMFDA accumulation across a large number of cells.

The onset of flucloxacillin induced DILI is delayed, therefore the effect of continuous dosing was assessed. HepaRG cells were incubated in the presence of flucloxacillin after 24 hours, 72 hours and 1 week of dosing with flucloxacillin. Media supplemented with flucloxacillin was made fresh for each dose to prevent quenching of the drug by serum proteins. The accumulation of CMFDA in the biliary epithelial of HepaRG cells increased as the duration of flucloxacillin incubation became longer (Figure 3.17, left). In addition to the increased accumulation of the fluorescent marker within the bile canaliculi, the morphology of the canaliculi appears to change with dilation occurring. This is particularly apparent after 1 week continual dosing (Figure 3.17, right). This accumulation in the biliary epithelia is indicative of active transport of CMFDA from the hepatocyte like cells via MRP2 and P-gp.

In order to quantify the increase of CMFDA accumulation in the presence of flucloxacillin, the mean intensity of green fluorescence (488 nm) was normalised to the nuclear staining intensity (DAPI). This was performed to give an estimate of 'green fluorescence per cell'.

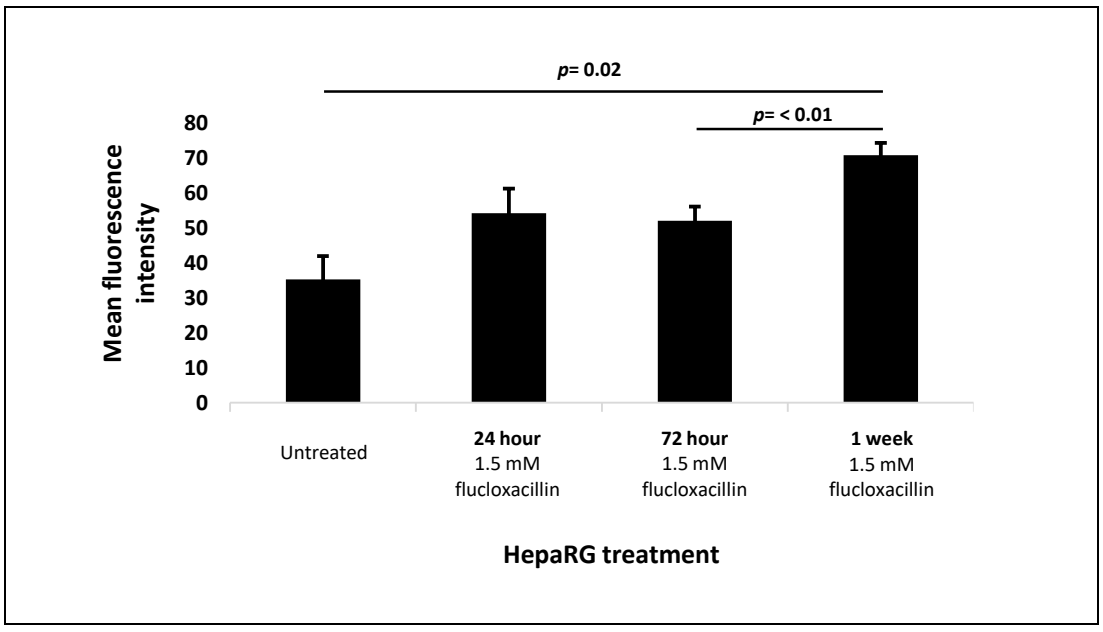


Figure 3.18. Quantification of membrane transporter function in HepaRG cells after prolonged flucloxacillin exposure. CMFDA (488, green) intensity was normalised to nuclear (blue, DAPI) intensity to calculate a 'CMFDA signal per nucleus' value. After prolonged flucloxacillin exposure the normalised CMFDA signal intensity significantly increases. All images were taken using the same exposure times.

Images were taken at 3 random points of each slide at a 5 x magnification in an attempt to remove technical error and bias towards particular clusters of cells. In keeping with the qualitative interpretation, it appears that the mean fluorescence of CMFDA does increase with the prolonged culture with flucloxacillin (Figure 3.17). Through performing a paired two-tailed t-test between the different conditions it was revealed that at 1 week flucloxacillin exposure CMFDA intensity significantly increased from the untreated and 72 hour treated cells. While a significant increase from 24 hours to 1 week was not observed, the *p*-value was still relatively low (0.075).

MK571 and valspodar were used to block the activity of MRP2 and P-gp respectively. Cells were cultured in the presence of 1.5 mM flucloxacillin overnight with the addition of one or both of the transporter blockers. Prior to incubation with flucloxacillin, cells were pre-treated with their respective blockers for 1 hour. In the control (no membrane transporter inhibitors) flucloxacillin can easily be seen localising within the bile canaliculi. This is observed as microtubule like structures on the maximum intensity projection image (MIP) (Figure 3.19, top left). The localisation of flucloxacillin within the F-actin bundles, representing bile canaliculi regions, are observed throughout the different Z-planes contributing to the final MIP (Figure 3.19, top right). Upon adding MRP2 block, covalent binding of flucloxacillin localisation within bile canaliculi reduced. Instead, covalent binding of flucloxacillin appears to be contained within the cell cytoplasm appeared to be increased (Figure 3.19, middle left). The same can be seen when adding P-gp block (Figure 3.19, middle right). Interestingly, when both MK571 and valspodar are added flucloxacillin modification can still be observed within bile canalicular regions (Figure 3.19, bottom left) as confirmed by Z-stack slices (Figure 3.19, bottom right). While this is the case, flucloxacillin does appear to be distributed within the cell cytoplasmic regions, however it is not conclusive whether this is significantly increased when compared to the control.

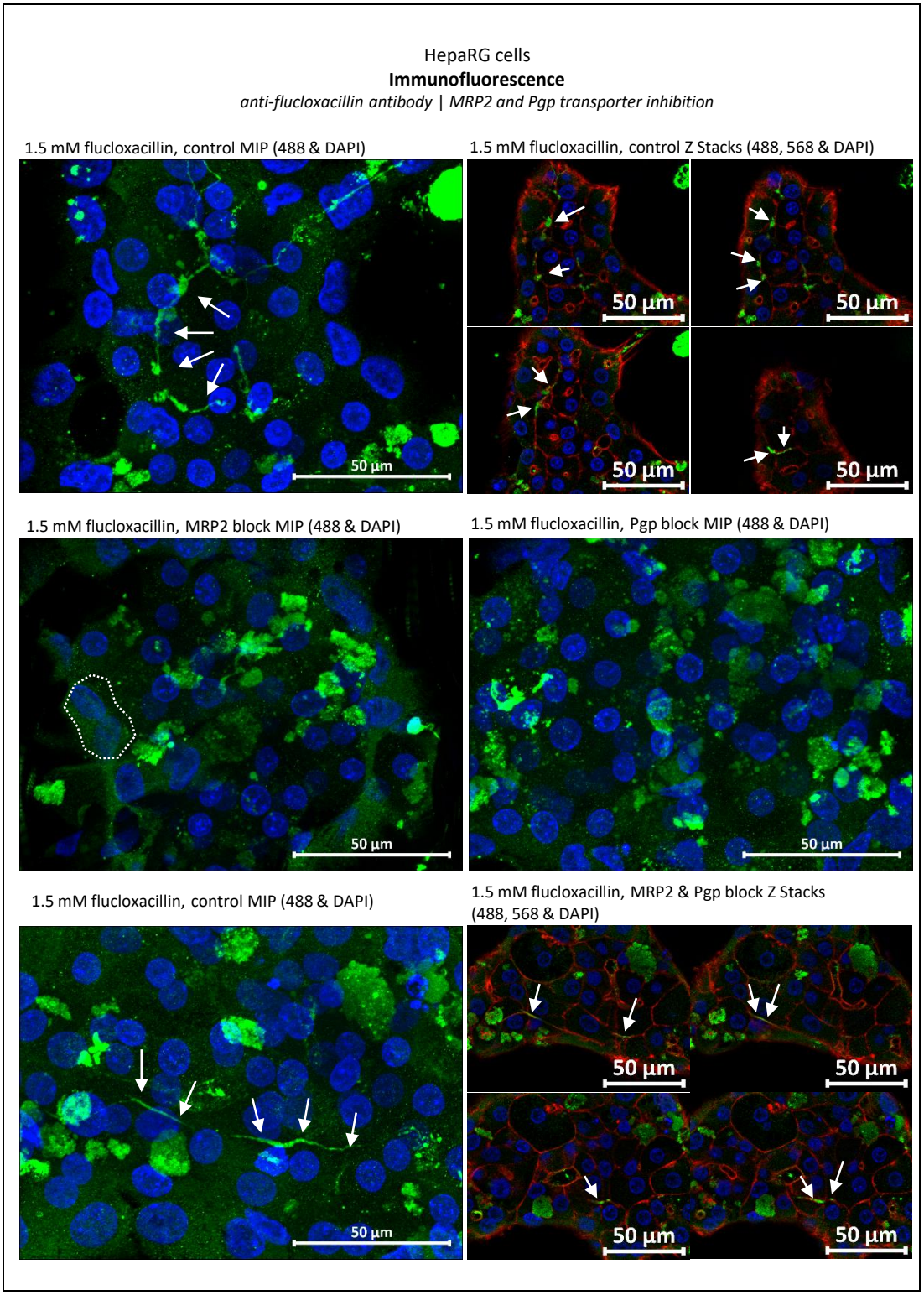


Figure 3.19. Assessment of the localization of flucloxacillin within bile canaliculi with the addition of transport inhibitors. MK571 and valspodar were used to block the activity of MRP2 and P-gp membrane transporters respectively. Flucloxacillin treatment at 1.5 mM overnight resulted in localization of drug within the bile canaliculi, as overserved in the MIP (top left) and Z-stack images (top right) (white arrows). MRP2 (middle left) and P-gp (middle right) inhibition resulted in less bile canalicular and more cytoplasmic modification (dotted line). When both blocks were added simultaneously (bottom) flucloxacillin could still be observed localizing within the bile canaliculi (white arrows, MIP (left) and Z-stacks (right)).

3.4.5 P38/MAPK14 MODIFICATION BY FLUCLOXACILLIN

P38α has been implicated in the activation of map kinase signalling pathways in the presence of flucloxacillin. Furthermore, previous studies have identified acetylation of key lysine residues augments p38α phosphorylation activity (K-53 & K-152) (Pillai *et al.*, 2011) (Figure 3.20A). Therefore, flucloxacillin binding to p38α was assessed. At high molar ratios of flucloxacillin to protein (50, 100, 150 and 200 to 1), flucloxacillin modification of p38α was observed using Western blot (Figure 3.20B). Although modification was observed, the sites of modification cannot be determined using this technique.

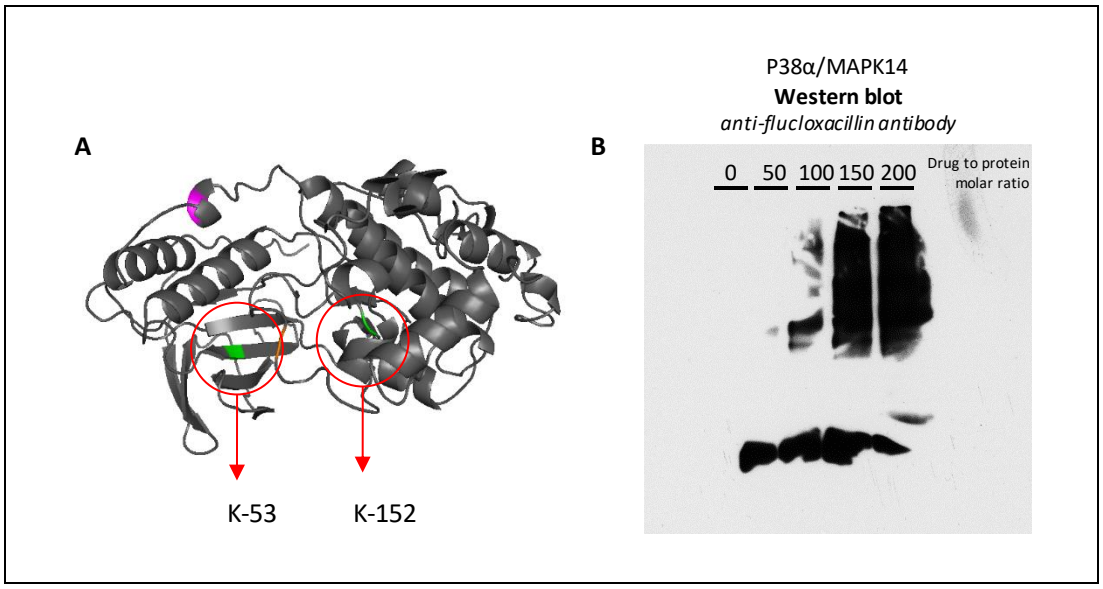


Figure 3.20. Modification of P38α/MAPK14 by flucloxacillin. P38α/MAPK14 acetylation of lysine 53 and 152 (A, left) augments phosphorylation activity. Modification of P38α/MAPK14 by flucloxacillin was detected using Western blot (B, right). Molar concentrations of 0, 50, 100, 150 and 200 to 1 (left to right) of flucloxacillin to protein were used.

Coomassie staining of SDS-PAGE was used to excise bands corresponding to p38α protein to be characterized using mass spectrometry. At all conditions a single protein was consistently observed, with limited protein degradation occurring as drug concentration increases.

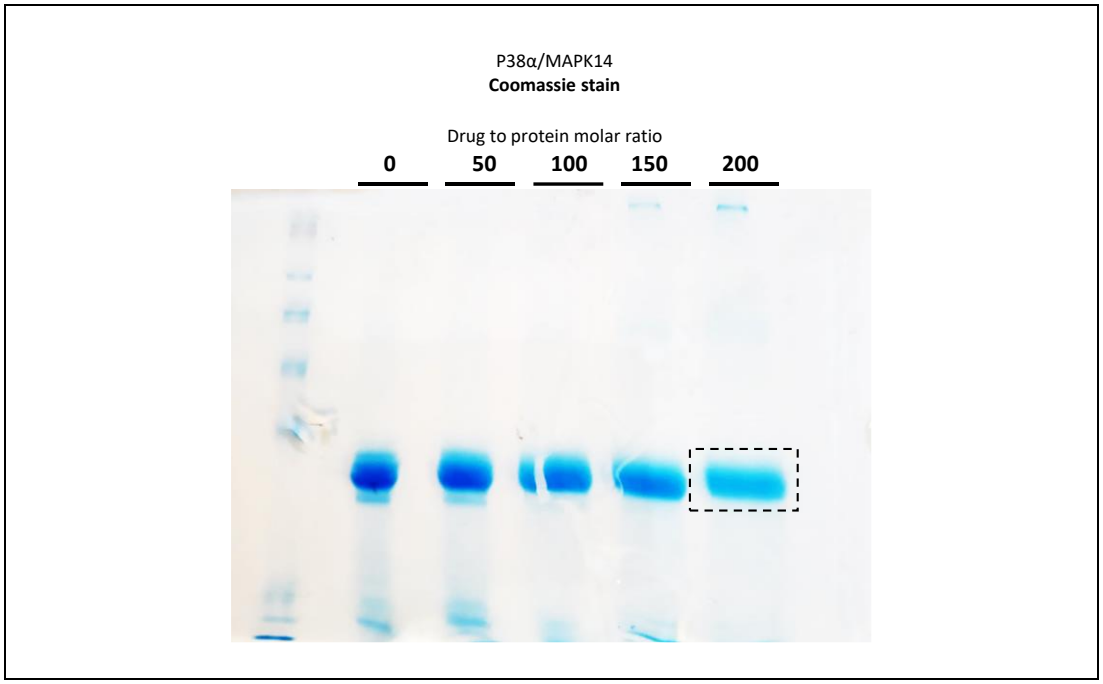


Figure 3.21. Coomassie stained SDS-PAGE of P38α/MAPK14 after flucloxacillin modification. P38α/MAPK14 flucloxacillin was separated using SDS-PAGE. Molar concentrations of 0, 50, 100, 150 and 200 to 1 (left to right) of flucloxacillin to protein were used. The band corresponding to P38α incubated with flucloxacillin at a 1:200 molar ratio was excised for further analysis (dotted line).

P38α incubated with flucloxacillin at a 1 to 200 molar ratio was excised from the coomassie stained SDS-PAGE gel (Figure 3.21, dotted line) and prepared for mass spectrometric analysis. Discovery proteomics was performed to enable characterization of all adducts formed, eliminating bias towards only K-53 and K-152. Indeed, the most common flucloxacillin adduct was formed on K-53 (Figure 5.4A), with modification apparent in a number of peptides. Further characterization revealed simultaneous modification of K-53 and K-54 (Figure 5.4B) along with adduct formation on K-45 (Figure 3.23A), K-76 (Figure 3.23B) and K-139. Flucloxacillin binding to K-152 was not identified, although the native versions of the peptide were detected at relatively high intensity.

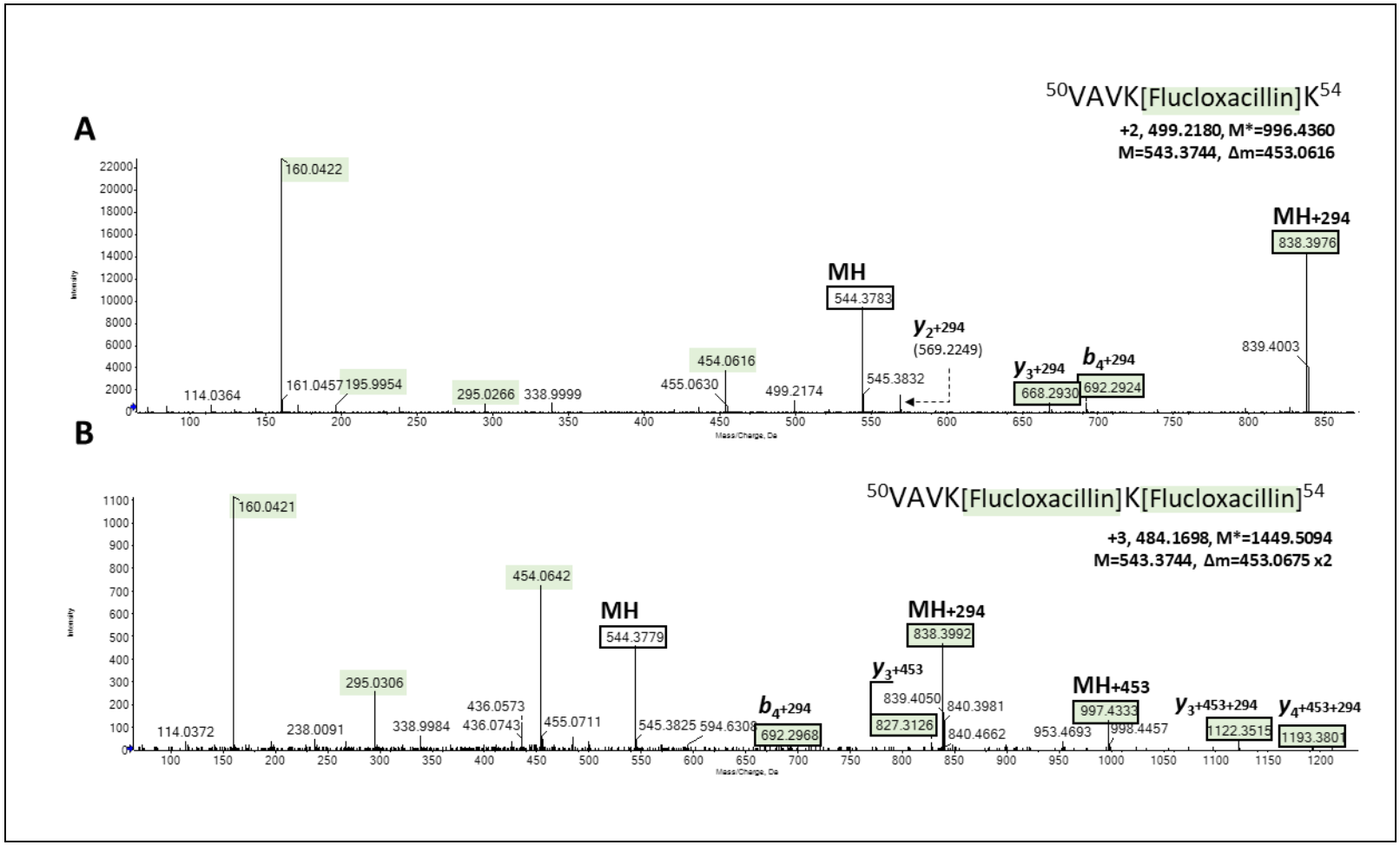


Figure 3.22. Flucloxacillin modification of P38α/MAPK14 protein. P38α was incubated with flucloxacillin at a 1:200 molar ratio and subjected to SDS-PAGE. Protein was excised from a coomassie stained gel and prepared for mass spectrometric analysis. Manual characterization reveals adduct formation on (A) K-53 [VAVK*K] and (B) K-53 & K-54 simultaneously [VAVK*K*].

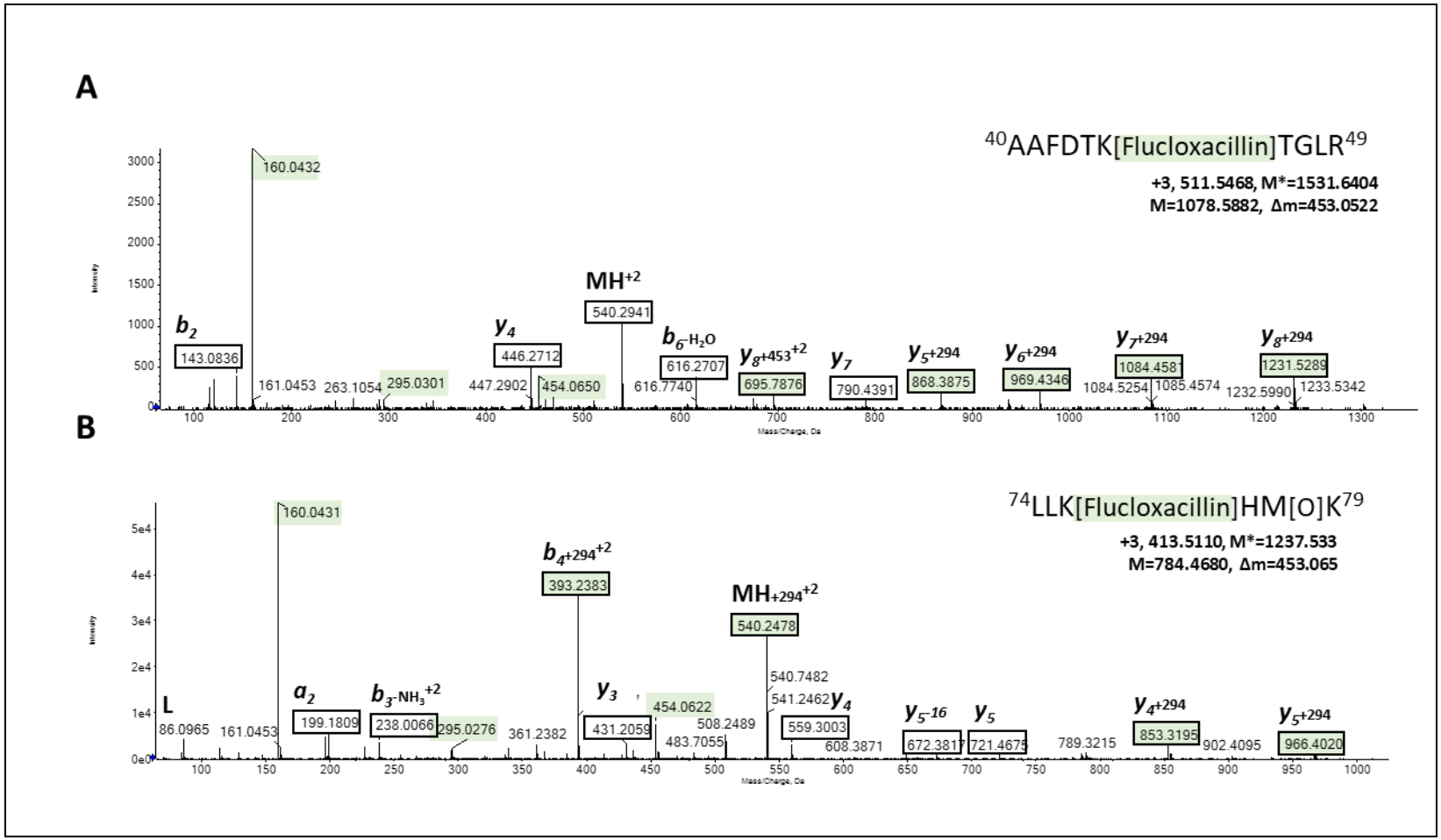


Figure 3.23. Flucloxacillin modification of P38α/MAPK14 protein. P38α was incubated with flucloxacillin at a 1:200 molar ratio and subjected to SDS-PAGE. Protein was excised from a coomassie stained gel and prepared for mass spectrometric analysis. Manual characterization reveals adduct formation on (A) K-45 [AAFDTK*TGLR] and (B) K-76 [LLK*HMK].

3.5 DISCUSSION

The aim of this chapter was to understand the diversity and localisation of flucloxacillin modified proteins using immuno-based assays. In order to meet these aims SDS-PAGE, Western blot and immunofluorescence microscopy were performed utilising the anti-flucloxacillin antibody developed in chapter 1. For the characterization of flucloxacillin modified proteins relevant cell lines were investigated. The proteins extracted from liver like cell lines, HepG2 and HepaRG cells, in addition to primary human hepatocytes were characterized. As flucloxacillin is related to DILI it was important to investigate the cells considered to be most closely related to the human liver. Although the study of primary hepatocytes may be more physiologically relevant, limitations in cell culture techniques exist. For example, primary hepatocytes are already terminally differentiated and cannot be cultured for long durations; compared with cell lines. In addition, flucloxacillin induced liver injury is linked with the incidence of HLA-B*57:01, therefore protein binding in C1R-B*57:01 antigen presenting cells solely expressing HLA-B*57:01 was also investigated.

HepaRG and HepG2 cells were used for the initial characterization of flucloxacillin modified proteins. Both cell lines were cultured as described in the presence and absence of flucloxacillin. Upon cell lysis and protein extraction, there was no observable difference in total protein diversity as defined using SDS-PAGE with Coomassie staining. Western blot analysis shows that proteins extracted from both cell lines are modified by flucloxacillin in a time and concentration dependent manner, which is consistent with previous publications (Jenkins *et al.*, 2009; Meng *et al.*, 2011). Protein diversity in C1R-B*57:01 cells was also found to be unchanged, with multiple proteins modified by flucloxacillin. In all these cell lines it is important to note that flucloxacillin antibody is highly specific and does not react with any proteins within untreated cell lysates. While previous studies have identified a handful of proteins (~6) undergoing modification in rat liver (Carey and van Pelt, 2005), here there is strong evidence for a much larger diversity. This is either since the high titre antibody

developed in this study was more sensitive than that used in previous studies, or that relatively high drug concentrations were used in cell monolayers cultures. To remove some of the limitations associated with liver like cell culture systems, primary human hepatocytes were also cultured in the presence of flucloxacillin. Again, a dose dependent increase in flucloxacillin modification of a wide range of proteins was detected, indicating flucloxacillin modification can occur in a more physiologically relevant system. Moving forward, we are investigating flucloxacillin covalent binding in more physiologically relevant hepatic models such as the 3D spheroids recently developed for drug safety screening (Gaskell *et al.*, 2016).

A previous study using mass spectrometric analysis identified flucloxacillin modified HSA from patient serum samples (Jenkins *et al.*, 2009). In this chapter serum samples were probed with the anti-flucloxacillin antibody to determine if any other serum proteins were targets for flucloxacillin modification. HSA was the only protein found to be targeted by flucloxacillin. As HSA is by far the most abundant serum protein, along with its evolutionary function as a small molecule carrier protein, it is perhaps not surprising that it was exclusively modified by flucloxacillin within blood derived from patients receiving flucloxacillin therapy. Reversible drug binding to albumin is an important factor in drug dosage calculations. Usually the main consideration in the determination of drug dose is the 'free' unbound quantity for several factors. Protein binding can prevent drugs to be both metabolized into reactive compounds or to reach toxic levels. Furthermore, irreversibly bound fractions can be released slowly to maintain an equilibrium. This binding and release of drugs changes depending on both physiological and chemical conditions, making it a complex process to understand (Tatlidil, Ucuncu and Akdogan, 2015). While the function of albumin in reversible drug binding is widely understood, less is known of the fate of covalent interactions.

Immunofluorescence imaging was used to detect the localisation of flucloxacillin binding in the previously mentioned cell culture systems. C1R-B*57:01 cells show both intracellular and extracellular modification. While flucloxacillin covalent binding is mainly located in the

cytoplasmic region of these C1R-B*57:01 cells it is important to note that the method associated with image acquisition could contribute to a false positive signal. Multiple images are taken through different planes of focus (Z-stacks) before being merged into one image (MIP). This could, in theory, make purely extracellular modification appear to localise within the cell (Figure 3.24). As intracellular modification was found on multiple planes of focus, and proteins extracted from C1R-B*57:01 cells were modified as detected by Western blot, this cytoplasmic protein modification appears genuine.

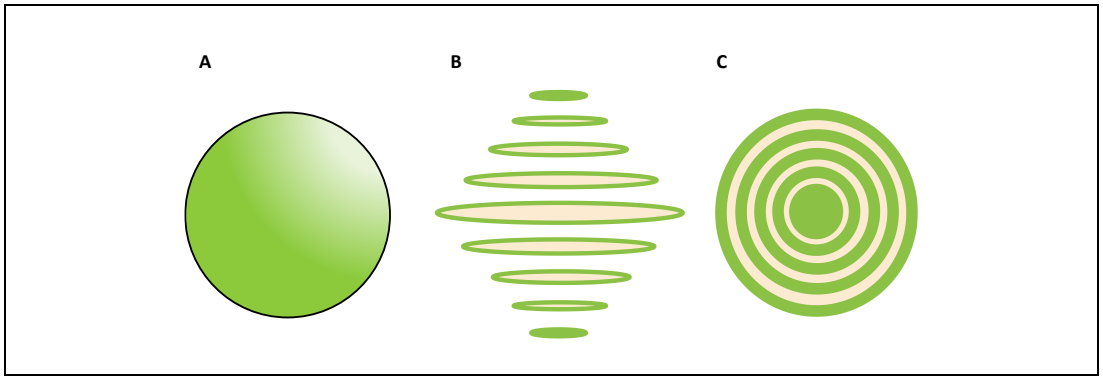


Figure 3.24. Assessment of intracellular protein modification using fluorescence microscopy. As cells are spheroid (A) and not flat, Z-stack images are taken through different focus points/planes of the cell (B). Once combined, extracellular modification could be mistaken for intracellular modification (C).

Interestingly, proteins extracted from HepaRG cells are more readily modified than those from HepG2 cells as detected using Western blot analysis. While the underlying mechanism for this is still largely unknown, protein modification on HepG2 cells is largely extracellular, with very limited flucloxacillin detected within the cell cytoplasm. This was identified in both SDS-PAGE Western blots and immunofluorescence microscopy, and can be attributed to a couple of different factors. Firstly, HepG2 cells may not have the required influx transporters to facilitate flucloxacillin entering the cell. While the proteins responsible for the active transport of flucloxacillin are largely unknown, it is interesting to note that HepG2 cells have low transporter abundance compared with primary cells (Sison-Young *et al.*, 2015). Secondly, HepG2 cells are naturally high in the abundance of the nuclear pregnane X receptor (PXR), responsible for maintaining bile acid homeostasis. During cholestasis bile acid concentrations

accumulate to toxic concentrations resulting in the increase of efflux activity in a PXR mediated fashion. Flucloxacillin has been identified as a potential PXR ligand, resulting in the increase of PXR-related gene expression. This in turn is believed to result in the increased activity of efflux transporters within HepG2 cells when exposed to flucloxacillin (Andrews *et al.*, 2010). In this case, limited intracellular flucloxacillin binding may be attributed to fast clearance upon entering the cell.

In contrast to the extracellular binding to HepG2 cells, intracellular modification of proteins within HepaRG cells treated with flucloxacillin was observed. Perhaps more interestingly, flucloxacillin treatment resulted in the localization of drug within the condensed F-actin regions. This is indicative of the bile canaliculi, as confirmed by the presence of MRP2 and P-gp in a similar positioning to flucloxacillin. Once metabolised to 5-hydroxymethly-flucloxacillin (5'OH-flucloxacillin), as generated by CYP3A4, flucloxacillin becomes toxic to biliary epithelial cells while remaining non-toxic to hepatocytes (Lakehal *et al.*, 2001; Andrews *et al.*, 2010). Hepatic cholestasis is characterized by the obstruction of bile flow out of the liver, therefore covalent binding at this location within the bile canaliculi could be associated with the observed toxicity. It is important to note that 5'OH-flucloxacillin can also form protein adducts which could be detected by the flucloxacillin antibody. Therefore, as HepaRG cells do over express CYP3A4 (Sison-Young *et al.*, 2015) is it likely that the metabolite is also residing within the bile canaliculi.

CMFDA was used to detect the efflux activity of the membrane transporters MRP2 and P-gp in HepaRG cells. As cells were pre-incubated with flucloxacillin for longer time courses, the active transport of CMFDA from hepatocytes into biliary epithelia increased. It was also noticed that bile canaliculi dilated when cells were incubated with flucloxacillin for prolonged periods of time, as observed in previous publications (Burban *et al.*, 2017). This is indicative of flucloxacillin mediating increased efflux activity over time, perhaps due to interactions with PXR. Studies characterizing the role of transporters in liver disease have shown that

polymorphisms in the bile salt efflux pump (BSEP) membrane transporter are associated with different forms of liver injury (Lang *et al.*, 2007), making the role of transporter activity relevant to DILI. However, whether these transporters are involved in the induction of flucloxacillin induced cholestatic liver injury are still not fully understood. To investigate efflux transporters on flucloxacillin translocation, MK571 and valspodar were used to block the activity of MRP2 and P-gp, respectively. HepaRG cultures containing no transport inhibition resulted in flucloxacillin localizing within bile canaliculi. Localization of flucloxacillin within bile canaliculi was not as defined when MRP2 or P-gp were blocked independently. However, when both transporters were inhibited simultaneously, localization within the bile canaliculi was still observed. Secondly, MRP2 and P-gp may not directly be involved in flucloxacillin transport, and are generally upregulated during a prolonged presence of flucloxacillin through a cellular stress response. Other important efflux transporters such as BSEP were identified to be downregulated in the presence of 6 different cholestatic inducing compounds in one study (Burbank *et al.*, 2015). Although contradictory to our results with MRP2 and P-gp, this highlights the complexity of the interplay between all of these membrane transporters during periods of cell stress. Further studies using a range of transport inhibitors must be performed in order confirm the conclusions for this section of the work.

Although a wide range of proteins were found to be modified by flucloxacillin in different cell types, the identity of the vast majority remains unknown. Mass spectrometry was used to characterize modified proteins, however due to MS/MS fragmentation of flucloxacillin it is not possible to confidently identify the peptide sequence, and so protein source, using commercially available software. That said, manual interpretation of peptide spectra revealed modification of the of 14-3-3 protein. This family of proteins are a group of conserved regulatory molecules expressed in all eukaryotic cells. Their main function is involved in the regulation of protein phosphorylation in MAPK pathways. Multiple isoforms

exist, each with differential localisation and function. For example the gamma isoform, originally thought to be specific to nervous tissue, has been shown to be a potential marker in the diagnosis of Creutzfeldt-Jakob disease (Zerr *et al.*, 1998). However, it has also been implicated in vascular smooth muscle cell activation and metabolism in responses to vessel damage (Autieri *et al.*, 1996). Further studies have identified the role of 14-3-3 epsilon in gastric cancer (Leal *et al.*, 2012) while 14-3-3 theta has a protective role in Parkinson's disease (Slone, Lesort and Yacoubian, 2011). The modification of other peptides was identified through the presence of flucloxacillin fragmentation ions, however manual characterization was not possible. Characterization of proteins modified by flucloxacillin were attempted in HepaRG cells, however due to the low relative abundance of modified proteins, resulting in low quality MS/MS spectra containing characteristic drug fragment ions, peptide annotations could not be manually assigned.

While this thesis is primarily focussed on the immune involvement in DILI, non-immune mediated cholestatic liver injury has been observed with flucloxacillin *in vitro*. P38α map kinase activity results in the observed cholestasis in the study performed by Burban *et al.* However, previous publications have observed P38α as the trigger for a number of pathways resulting in both immune activation and cell death (Arrighi *et al.*, 2001; Aiba *et al.*, 2003; Burban *et al.*, 2017). One theory for the enhanced phosphorylation activity of P38 in the presence of flucloxacillin is direct binding to key amino acid residues involved in its activity. It is understood that acetylation of K-152, located in the substrate binding domain, and K-53, in the ATP-binding pocket, results in increased enzymatic activation (Pillai *et al.*, 2011). Flucloxacillin was found to modify P38α by Western blot and mass spectrometric analysis at high molar ratios of drug to protein. However, the concentrations used unlikely reflect physiologically relevant conditions. While flucloxacillin modification of K-53 was found, this does not implicate a role in the augmentation of P38α phosphorylation activity. In fact, as K-53 is involved with ATP-binding, it's more likely that flucloxacillin would inhibit the activity of

P38. The relevance of the other amino acids modified by flucloxacillin is currently not understood. In order to test the effect of flucloxacillin modified P38α, kinase activity would need to be assessed. Challenges would appear in trying to purify flucloxacillin modified P38α while keeping the protein enzymatically active. For these reasons alone, it was not feasible in the scope of this project to perform such assays. Future studies would need to address both limitations in order to draw conclusions from the effect of flucloxacillin on P38α activity, and its relevance in non-immune mediated cholestatic liver injury.

In summary, a diverse range of intracellular and extracellular proteins derived from both immune and hepatic cells were found to be targets of flucloxacillin. These flucloxacillin modified proteins in HepaRG cells mainly localized within bile canaliculi, potentially providing a pool of localised antigens involved in the induction of local immune reactions. Furthermore, covalent binding of flucloxacillin to proteins involved in cellular signalling pathways may alter the functions of kinase activity, leading to non-immune mediated cholestatic liver injury. Regardless of the protein targets and localisation, flucloxacillin modified proteins could be processed intracellularly and presented by antigen presenting cells, leading to the activation of immune responses.

CHAPTER 4 - THE HLA-B*57:01 IMMUNOPEPTIDOME AND FLUCLOXACILLIN

4.1 INTRODUCTION

The ability for the immune system to detect foreign antigens leading to subsequent cell death of infected cells/tissue is an evolutionary response to infections. In addition to microbial infections, cancer accumulates in 1000s of genetic mutations resulting in the presentation of neoantigens, culminating in the destruction of malignant cells by T cells (Diken *et al.*, 2017). Understanding this interaction between antigens and immune cells was crucial in understanding tissue rejection is cases of organ transplantation. Early studies found that skin graft rejection occurred within 10-14 days of surgery. In animal models, upon re-challenge with skin from the same donor, the recipient rejected the tissue within a few days; indicating immunological memory (Abbas *et al.*, 2014). This led to the discovery of the involvement of the adaptive immune system in transplant rejection, in particular MHC types. Interestingly, MHC was termed from '*histo*, tissue' compatibility with the host immune system. While other mechanisms of transplant rejection exist, such as polymorphic antigens other than MHC molecules, these are termed minor histocompatibility antigens. MHC are responsible for the observed strong and rapid immune reactions mediated through T cell activation. This occurs through several molecular mechanisms. Firstly, recipient CD8+ T cells can recognize self MHC molecules presenting donor derived peptides. Secondly, CD8+ self T cells may interact with donor MHC resulting in allogenic activation, due to thymic deletion not accounting for foreign MHC proteins. This may or may not include an interaction with the MHC peptide itself, and all three are termed direct recognition. Indirect recognition states that allogenic donor peptide-MHC complexes are taken up by the recipient APCs, processed intracellularly, and the donor MHC peptide is presented on the recipient MHC. This results in the activation of CD4+ T cells, and immune recruitment (Abbas, Lichtman and Shiv, 2018).

The idiosyncratic nature of drug-induced type B hypersensitivity reactions makes the involvement of the adaptive immune system a major component in the development and progression of disease. At present, three models are proposed for the interaction between drugs and the immune system through MHC and TCRs. Non-covalent interactions between the drug, peptide-MHC complex and TCR, termed the PI hypothesis (Schnyder *et al.*, 1997), is believed to result in the activation of T cells. Alternatively, the hapten/pro-hapten hypothesis states that reactive drugs or drug metabolites (haptens) bind covalently to proteins to initiate immune reactions (Landsteiner and Jacobs, 1935). Both hypotheses are reliant on intracellular processing of protein antigens to peptides followed by presentation of peptide-HLA complexes by antigen presenting cells. Although processing is required to generate the peptide antigen, the PI mechanism occurs independently of the processing event. However, there remains controversy as to whether drugs interact preferentially with HLA proteins or HLA binding peptides. Functional studies using parent drugs, metabolites and synthetic drug-modified proteins or designer HLA binding peptides, and T cell lines and clones from patients with hypersensitivity, support both hapten and pharmacological interactions hypotheses (Schnyder *et al.*, 2000; Manchanda *et al.*, 2002; Naisbitt *et al.*, 2007; El-Ghaiesh *et al.*, 2012; Ogese *et al.*, 2014; Yaseen *et al.*, 2015; Meng *et al.*, 2016, 2017). However, knowledge of structures generated naturally and displayed by HLA proteins remain largely ill-defined.

In recent years, mass spectrometric analysis of HLA-peptide complexes has identified thousands of peptides naturally presented on the cell surface by HLA molecules. The peptide sequence information derived from this method, in conjunction with X-ray crystallographic analysis of HLA peptide complexes, has provided new insights into the mechanisms of immune-mediated disease. Mass spectrometry was first utilised for the characterization of MHC peptides in 1992 (Hunt *et al.*, 1992). These earlier studies characterizing the cell surface presentation of MHC peptides identified low sequence IDs in comparison to more recent

studies. Investigating the peptides presented by HLA-B*51:01 in a 2006 publication resulted in 64 peptides identified from an initial starting material of $4x10^9$ cells (Gebreselassie, Spiegel and Vukmanovic, 2006). More recently studies focussing on MHC I peptide presentation on breast cancer cell lines identified 3,186 MHC peptides from a total of $2x10^9$ cells ($1x10^8$ per experiment) (Rozanov *et al.*, 2018). The improvement in peptide identification is largely due to the availability of technology with high sensitivity enabling deeper profiling of MHC peptides. Typically, MHC peptide identification using mass spectrometry is performed using data-dependant acquisition (DDA), meaning a certain ion intensity threshold, relative to other ions in the sample, needs to be reached in order to induce fragmentation. Therefore, extensive sample pre-fractionation must be performed to facilitate the capture of low abundant peptides. HPLC is often used in the preparation of complex samples. Here, fractionation is based on the chemical properties of the analytes, for example, C18 will separate samples based on peptide/protein hydrophobicity.

The identification of antigenic MHC peptides has not only allowed for developing personalised immunotherapy for cancer and auto-immune diseases, but has also shed light on the mechanisms of diseases. In terms of drug hypersensitivity, research from three teams found that the antiretroviral drug abacavir interacts with the peptide binding cleft of HLA-B*57:01 altering the shape and chemistry of the HLA molecule and the peptides that bind (Illing *et al.*, 2012; Norcross *et al.*, 2012; Ostrov *et al.*, 2012). This gave rise to the third hypothesis of MHC TCR interactions. Abacavir, a NRTI used in the treatment of HIV infection, causes abacavir hypersensitivity syndrome (AHS) in 55% of patients taking the drug who are HLA-B*57:01 positive (Mallal *et al.*, 2002; Lucas *et al.*, 2015). While the genetic association, and subsequent genetic screening prior to treatment, between AHS and HLA-B*57:01 was identified in the early 2000's it was not until 2012 when the molecular mechanisms of disease were defined (Illing *et al.*, 2012; Norcross *et al.*, 2012; Ostrov *et al.*, 2012). Due to the genetic predisposition of AHS and HLA-B*57:01 the overall repertoire of MHC peptides presented by

the allele in the presence of abacavir was investigated using mass spectrometry. Illing *et al* were among the first to discover that in the presence of abacavir the peptides that are presented by HLA-B*57:01 are altered with up to 25% of the eluted peptides being novel self-peptides. MHC I peptides are generally 9-11 amino acids long and have specific residues (position 2 and Ω; the C-terminal amino acid residue) where they anchor to the HLA molecule. While the overall length distribution of peptides was found to remain the same, the abundance of leucine and isoleucine at the PΩ residue was increased, while tyrosine, tryptophan and phenylalanine all decreased (Illing *et al.*, 2012). Further structural elucidation of the drug peptide-HLA complex revealed that abacavir bound non-covalently in the vicinity of the F-pocket of the HLA binding groove. Up to a thousand abacavir unique HLA binding peptides were identified, however, their contribution to the CD8+ T cell response seen in abacavir hypersensitive patients is still to be defined. Importantly, the interaction between abacavir and HLA-B*57:01 with peptide repertoire change has not been observed with other drugs.

The immune involvement in flucloxacillin DILI, in particular the genetic predisposition to HLA-B*57:01, makes the investigation of the immunopeptidome key to this study. In the previous chapter flucloxacillin was found to bind extensively to intracellular and extracellular proteins in a range of cell types, including C1R-B*57:01 B lymphoblastoids. As the mechanisms for immune activation in drug hypersensitivity are dependent on antigen processing, the immunopeptidome of C1R-B*57:01 cells cultured in the presence of flucloxacillin was investigated. Although typically flucloxacillin is believed to activate T cells through the hapten hypothesis, the shared genetic predisposition between abacavir and flucloxacillin meant it was important to first identify any changes in the peptide repertoire in the presence of drug treatment.

4.2 AIMS

Direct binding of flucloxacillin or its degradation products to HLA-B*57:01 through non-covalent bonds may alter the peptide binding repertoire similar to abacavir. Thus, the objective of this chapter was to profile the flucloxacillin HLA-B*57:01 antigenic repertoire, with the intention of discovering whether altered self-peptides can be displayed on the surface of antigen presenting cells for presentation to CD8+ T cells. In order to achieve this, the following aims were set;

1. Development of workflows to enable immunoaffinity capture of MHC I peptide complexes.
2. Implementation of HPLC fractionation and mass spectrometric methods to study the MHC I repertoire of low abundant analytes.
3. Characterize the antigenic profile of MHC I peptides eluted from control, abacavir and flucloxacillin treated C1R-B*57:01 cells.

4.3 METHODS

4.3.1 CELL CULTURE

The lymphoblastoid C1R cell line (Storkus *et al.*, 1987; Zemmour *et al.*, 1992) deficient in MHC-class I, transfected to express HLA-B*57:01, were used as a source of antigen presenting cells. C1R-B*57:01 cells were maintained in F1 media (RPMI 1640 supplemented with 10% FBS (Invitrogen, Paisley, UK), 100 mM L-glutamine, 100 µg/mL penicillin, 100 U/mL streptomycin), 1mM HEPES and 50 µg/mL geneticin (37°C, 5% CO_2). C1R-B*57:01 cells were grown in multiple T175 culture flasks (Nunc) until a cell number of $1x10^9$ cells was achieved. Treatment with 35 µM abacavir or 1.5 mM flucloxacillin sodium salt was for 10 days or 48 hours, respectively, unless otherwise stated. Cells were subsequently harvested by centrifugation (RT, 453 x g), snap frozen in liquid nitrogen and stored at -80°C until required. W6/32 (Sigma Aldrich, Dorset, UK) mouse hybridoma cells lines, used for the production of anti-HLA-A,B,C IgG2a, were maintained in F1 media (37°C, 5% CO_2). W6/32 hybridomas were grown to confluence before the supernatant was recovered after centrifugation and stored at 4°C until required.

4.3.2 ASSESSMENT OF FLUCLOXACILLIN TOXICITY ON C1R-B*57:01 CELLS

Flucloxacillin toxicity was assessed by culturing C1R-B*57:01 cells in the presence of drug at varying concentrations. Cell proliferation was measured by the addition of [3^H]thymidine (0.5 µCi/well, 5 Ci/mmol, Morovek Biochemicals Ltd, Brea, CA, USA) for 16h of culture followed by scintillation counting.

4.3.3 PAN-MHC-CLASS I ANTIBODY PURIFICATION AND VALIDATION

W6/32 anti-HLA-A,B,C IgG2a antibody was purified through affinity capture through a 10 mL bed volume of protein A sepharose (PAS) (fast flow, Sigma Aldrich, Dorset, UK). Supernatant was pooled, pH adjusted to 8.0, and filtered through 0.45 µM pore size nitrocellulose (Sarstedt, Leicester, UK) prior to affinity capture. After washing with PBS W6/32 antibody

was eluted in 0.1M citric acid and pH adjusted with 1M Tris. W6/32 antibody was quantified using the Nanodrop 1200 spectrophotometer (ThermoScientific, MA, USA) at A280. Flow cytometry was used to identify W6/32 binding to MHC I on multiple B cell lines through adding 10 µg (unless otherwise described) W6/32 antibody to $1x10^6$ cells for 1 hour, followed by PE-A conjugated anti mouse secondary antibody (Sigma-Aldrich). The FACS Canto II flow cytometer was used to acquire data that was subsequently analyzed using associated FACS DIVA software.

4.3.4 PURIFICATION OF MHC-B-57:01 PEPTIDE COMPLEXES

Cell pellets at a cell density of $1x10^9$ were lysed in 0.5% IGEPAL, 50 mM Tris pH 8.0, 150 mM NaCl and protease inhibitors (complete protease inhibitor cocktail, Roche). Pellets were manually disrupted with pipetting until a single pellet was no longer visible. Lyzed cells were incubated while rocking at 4°C for 1h. Lysates were cleared by ultracentrifugation (180,000 x *g*) and passed through unconjugated PAS (Repligen, USA). MHC class I complexes were captured through anti-MHC 1 (W6/32) antibody conjugated to PAS, prepared as previously described (Dudek *et al.*, 2016). For every 10^9 cells used, 10 mg of W6/32 conjugated to 1 mL of PAS, termed column volume (CV) was used. Proteins and peptides captured by the unconjugated PAS pre-column and W6/32 affinity column were washed with 10 CV of cold wash buffer 1 (0.005% (w/v) IGEPAL, 50 mM Tris-HCl pH 8.0, 150 mM NaCl, 5 mM EDTA, L pepstatin A), followed by 10 CV of cold wash buffer 2 (50 mM Tris-HCl pH 8.0, 150 mM NaCl), then 10 CV of cold wash buffer 3 (50 mM Tris-HCl pH 8.0, 450 mM NaCl) and finally 10 CV of cold wash buffer 4 (50 mM Tris-HCl pH 8.0).

MHC-bound complexes were eluted in 5 CV 10% acetic acid and manually loaded onto a monolithic C18 column (100 x 4.6 mm Onyx, Phenomonex) in 0.1% trifluoroacetic at 2 mL/min connected to an Agilent 1260 HPLC. Sample loading was performed for 6 minutes with a wavelength absorbance set to 254 nm. The mixture of MHC-bound peptides, MHC

heavy chain and β2-microglobulin were subsequently fractionated using mobile phases containing 0.1% TFA(A) and 80% ACN/0.1% TFA (B) with a wavelength detection of 215 nm (Table 4.1).

Table 4.1. MHC-bound peptide HPLC method parameters. MHC peptides and protein components of eluted peptide-MHC complexes were separated using mobile phases containing 0.1% TFA (A) and 80% ACN/0.1% TFA (B). Fractions were taken across the gradient at 0.25 min and 1 minute intervals for hydrophilic and hydrophobic constituents, respectively.

Time	%A	%B	Flow (ml/min)	Wavelength	Fractionation Trigger
0:00	98	02	2	215	Off
0:15	85	15	2	215	Time-based with 15 second timeslices
4:15	70	30	2	215	
12:15	60	40	2	215	
22:15	55	45	2	215	
23.00					Time-based with 1 min timeslices
24:15	01	99	2	215	
26:15	00	100	2	215	
31.00					Off
32:15	98	02	2	215	
35.00	98	02	2	215	

4.3.5 MASS SPECTROMETRIC ANALYSIS OF HLA-B*57:01 PEPTIDE-MHC

Peptide containing HPLC fractions were pooled (as described later in this chapter) and concentrated to 10 µL in a vacuum centrifuge (Speedvac, Eppendorf) at 30°C. Peptide pools were analyzed using a Triple TOF 5600 mass spectrometer (Sciex) and were delivered into the instrument using a Eksigent NanoLC Ultra HPLC system. Samples were injected onto a nanoACQUITY UPLC Symmetry C18 Trap Column (P/N 186007496, Waters, MA, USA) and washed for 10 min at 2 µL/min with 0.1% FA. A gradient from 1.7% ACN/0.1% FA to 64% ACN/0.1% FA was applied over 79 min at a flow rate of 300 nL/min through a Peptide BEH C18 nanoACQUITY Column (P/N 186003815, Waters, MA, USA). MS was operated in positive ion mode with survey scans of 200 ms, with an MS/MS accumulation time of 150 ms for the 20 most intense ions (total cycle time 3.2 seconds). A threshold for triggering MS/MS of 40 counts per second was used with an exclusion of former target ions for 30 seconds. Rolling collision energy was applied. Data dependant acquisition of ions in the mass range of 200-

1,800 amu (MS) and 60-1,800 amu (MS/MS) was collected using Analyst TF 1.6. The instrument was automatically calibrated using a β-galactosidase digest every 3rd sample. MHC heavy chain data was acquired in the same way with the following adjustments. A gradient of 5% ACN/0.1% FA to 50% ACN/0.1% FA was applied at 300 nL/min over 90 mins. MS survey scans of 250 ms with MS/MS accumulation times of 100 ms were applied to ions over 100 counts per second for the 25 most abundant ions (total cycle time 2.8 s). Former target ions were excluded for 12 seconds. Ions in the region of 300-1,650 amu were acquired for MS and 100-1,400 amu for MS/MS.

4.3.6 HLA PEPTIDE REPERTOIRE ANALYSIS

ProteinPilot (Sciex, version 5.0, revision 4,769, Paragon algorithm 5.0.0.0.0. 4767) was used to search the peak lists against the UniProt-Sprot database (all species 557,986 entries, July 2018) with Homo sapiens species restriction (20,386 entries) and decoy reverse database (total 40,772 entries searched against). Biological modifications were set as variable modifications and enzymatic restriction was removed. All MS/MS peptide pool datasets were processed in one batch using a false discovery rate (FDR) cut-off of 5%. Further validation of peptides within the 5% false discovery threshold were processed through the optimisation of previously described methods (Illing *et al.*, 2012) (section 4.4.5).

4.3.7 HLA-B* 57:01 MODELLING

The crystal structure of HLA-B*57:01 (PDB 3VRJ) (Illing *et al.*, 2012) was used to generate models by removal of abacavir and the peptide using Pymol (2.0, Schrodinger). GOLD 5.2 (CCDC software) (Jones *et al.*, 1997) was used to dock flucloxacillin and penicilloic acid within the binding groove, with the binding site defined as 15 Å around the binding point. A generic algorithm with ChemPLP as the fitness function was used to generate 10 binding modes per ligand. Default settings were retained for the “ligand flexibility”, “fitness and search options”, and “GA settings”.

4.4 RESULTS

4.4.1 MHC CLASS I SURFACE EXPRESSION ON RELEVANT CELL LINES

The assessment of MHC class I surface expression was important to determine the suitability of C1R-B*57:01 cells peptide elution studies. While C1R-B*57:01 cells were used in the 2012 abacavir paper (Illing *et al.*, 2012), the determination of their physiological relevance with regards to MHC I expression was compared to other cell types using flow cytometry (Figure 4.1).

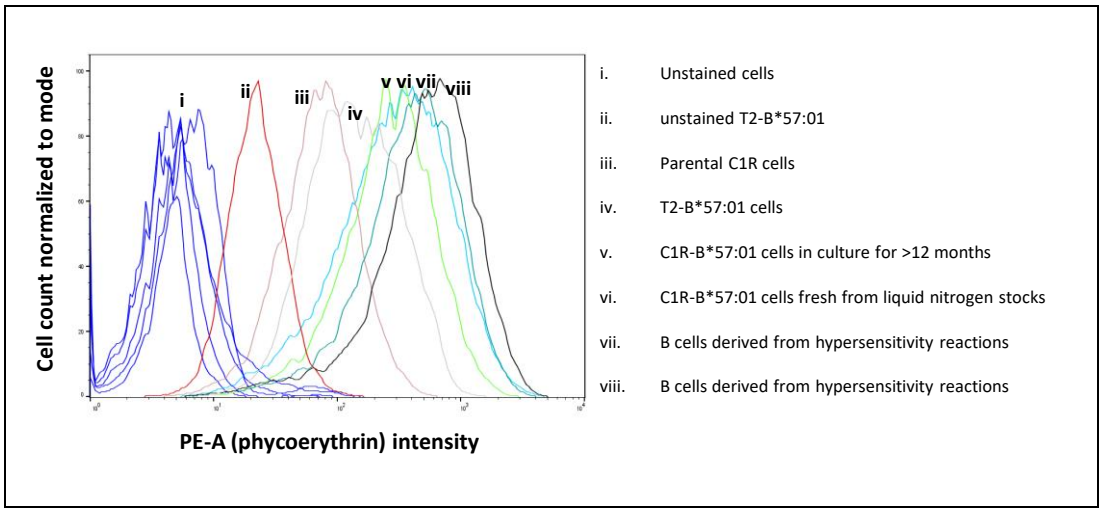


Figure 4.1. Quantification of MHC class I cell surface expression on B cells. C1R-B*57:01 cells fresh from liquid nitrogen stocks (vi) and C1R-B*57:01 cells in culture for >12 months (v) shows no reduction in MHC I expression. Parental C1R cells (iii) and T2-B*57:01 cells (iv) lacking in antigenic processing machinery express low MHC I as expected. B cells derived from patients with hypersensitivity reactions show the highest MHC I surface expression (vii & viii). Unstained cells are shown (i) with the exception of unstained T2-B*57:01 (ii) which reported a higher background autofluorescence.

As C1R-B*57:01 cells were kept in culture for a considerable period of time, it was important to confirm that expression of MHC I did not decline over time. C1R-B*57:01 cells cultured fresh from liquid nitrogen stocks (Figure 4.1, vi) showed no increase in HLA class I expression from C1R-B*57:01 cells kept in culture for >12 months (Figure 4.1, v). Parental C1R cells are known to have low HLA class I expression which was confirmed (Figure 4.1, iii). T2-B*57:01 cells carry the HLA-B*57:01 gene however do not possess TAP antigen processing machinery. Therefore, due to the lack of the ability to load antigens onto MHC through TAP processing pathways HLA class I surface expression is lower than C1R-B*57:01 cells (Figure 4.1, iv). T2-

B*57:01 MHC I expression remained higher than parental C1Rs as other processing pathways exist. Expression of MHC I was also similar to that of many immortalised cell lines derived from patients (Figure 4.1, vii & viii). Further analysis of MHC I expression was performed on patient derived B cells, from both healthy controls (HC) (Figure 4.2) and hypersensitivity patients (H-Pt) (Figure 4.2). This comparison also performed to understand HLA class I expression on C1R-B*57:01 cells compared with more physiologically relevant sources. Here we can conclude that C1R-B*57:01 cells and immortalised patient derived antigen presenting cells expressed comparable levels of MHC class I molecules.

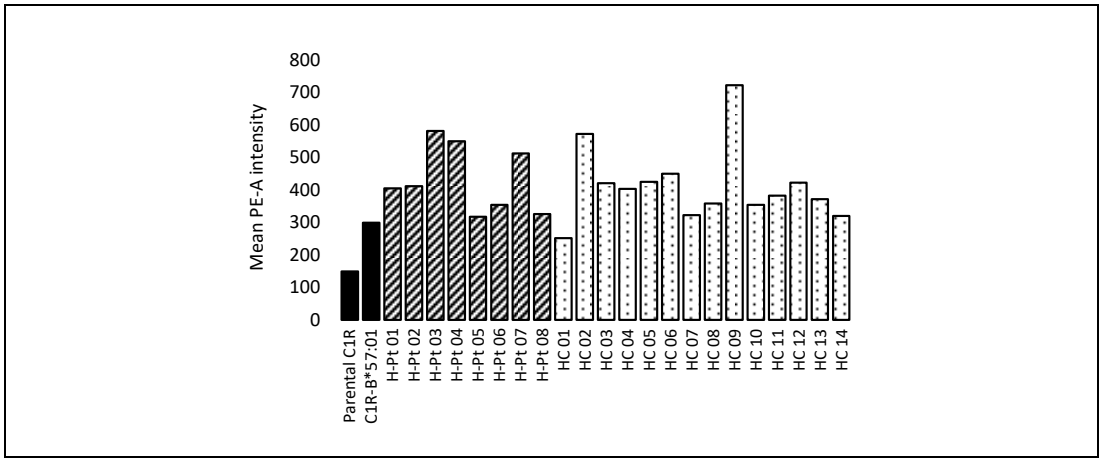


Figure 4.2. Further analysis of MHC I surface expression from patient derived B cells in comparison to C1R-B*57:01 and parental C1R cells. Proteins on the surface of C1R-B*57:01 cells and B cells from healthy volunteers (HC) and hypersensitive patients (H-Pt) shows that similar levels of MHC class I proteins are expressed.

4.4.2 FLUCLOXACILLIN TOXICITY ON C1R-B*57:01 CELLS

In order to detect any changes to the HLA-B*57:01 immunopeptidome it was important to use a concentration of flucloxacillin that would maximise the likelihood of any alterations to be observed, while maintaining cell viability. Therefore, the toxicity of flucloxacillin on C1R-B*57:01 cells was assessed. Cells incubated at a range of flucloxacillin concentrations found that up to 2 mM did not reduce cell viability (Figure 4.3). Beyond 2 mM, toxicity was quickly observed. Based on these results, a concentration of 1.5 mM flucloxacillin was used for subsequent MHC peptide elution studies.

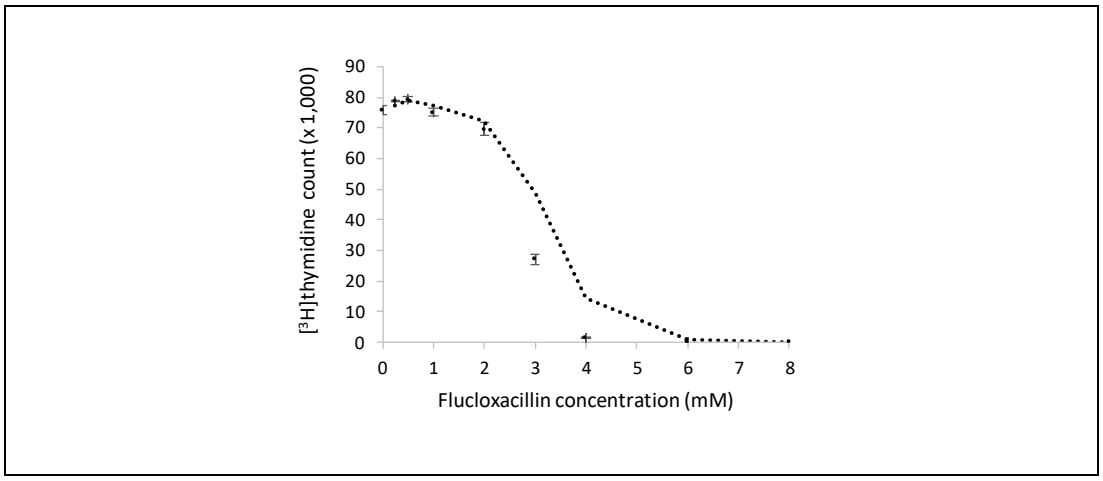


Figure 4.3. Flucloxacillin toxicity on C1R-B*57:01 cells. C1R-B*57:01 cells were incubated at different concentrations of flucloxacillin for 24 hours. Subsequent incubation with the incorporation of tritiated thymidine for 16 hours shows toxicity is limited up to 2 mM (n = 3).

4.4.3 Optimisation of C1R-B*57:01 cell culture conditions

To observe any changes in the immunopeptidome of C1R-B*57:01 cells, comparisons between treated and untreated (control) cells were required. It was therefore important to understand whether the cell culture condition could alter the levels of MHC I presentation. Firstly, C1R-B*57:01 cell culture medium was supplemented with FBS and grown in cell culture flasks (Figure 4.4A) or roller bottles (Figure 4.4B). After identifying no difference in MHC I abundance between the growth vessels used, the effect of media and flucloxacillin treatment was assessed. C1R-B*57:01 cells were cultured in media supplemented with human (AB) serum (Figure 4.4C), again no change in MHC I abundance was identified. Finally, the addition of flucloxacillin also resulted in no changes (Figure 4.4D). This enabled the conclusion to be drawn that MHC I abundance remains consistent irrespective of the cell culture conditions used.

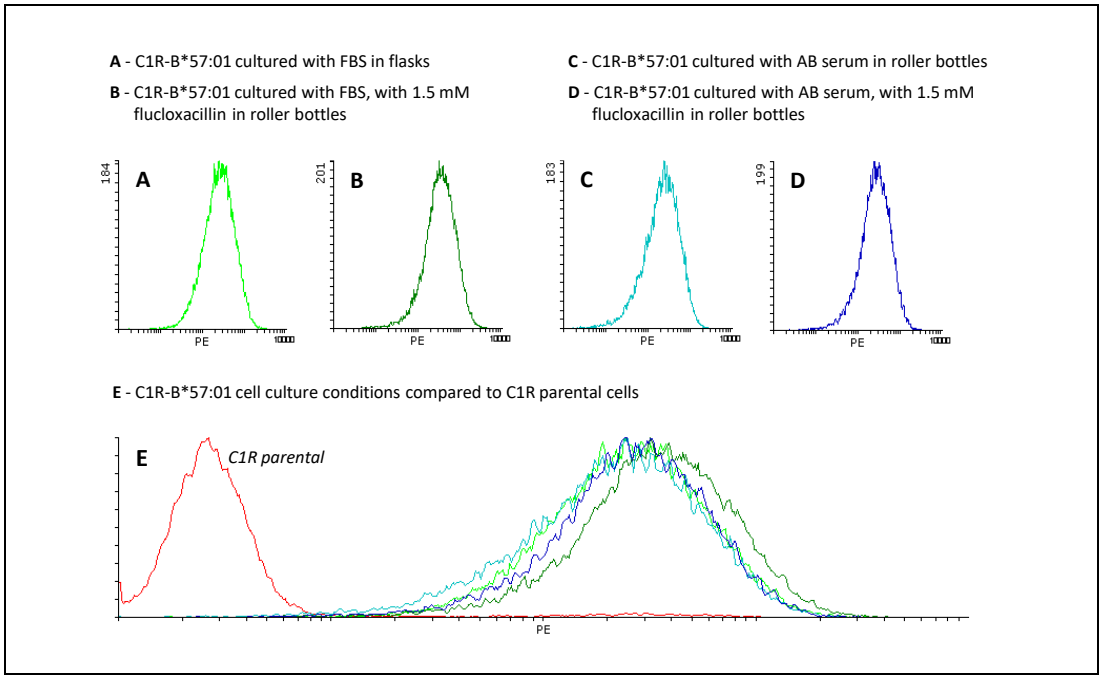


Figure 4.4. Effect of cell culture conditions on MHC I expression. (A) C1R-B*57:01 cells cultured in the presence of FBS and grown in T75 flasks. (B) C1R-B*57:01 cells cultured in the presence of FBS in roller bottles. (C) C1R-B*57:01 cells grown in the presence of human AB serum in roller bottles. (D) C1R-B*57:01 cells cultured in human AB serum, with 1.5 mM flucloxacillin in roller bottles. MHC I expression was comparable across all conditions.

4.4.4 MHC I IMMUNOAFFINITY CAPTURE METHOD OPTIMISATION

For the investigation of drug-related immunopeptidomics immunoaffinity capture of MHC I along with extensive peptide pre-fractionation prior to mass spectrometric analysis is required. One of the initial steps in the peptide elution protocol is to lyse antigen presenting cells to release MHC I into cell lysate. Cryogenic milling of frozen cell pellets is usually performed at sub - 150°C temperatures prior to cell lysis. Here a 10 mm stainless steel ball is used to manually smash the cell pellet in a RETCH mixer mill prior to the addition of lysis buffer. As this is not possible in Liverpool, manual lysis using a pipette tip was compared to cryogenic milling. This optimisation was performed within the Purcell laboratory, Monash University, Melbourne Australia. For this, the quantity of β_2M extracted from 2 identical cell pellets was assessed. As β_2M is a small sub-unit of MHC I, the retention time when fractionated using HPLC is separate from MHC heavy chain and protein contaminants; a peak at around 30 mins is indicative of β_2M. It was observed that the quantity of β_2M extracted

from both cell pellets was comparable (Figure 4.5). This removed the requirement of cryogenic milling for C1R-B*57:01 MHC I extraction, allowing for implementation using equipment available within Liverpool.

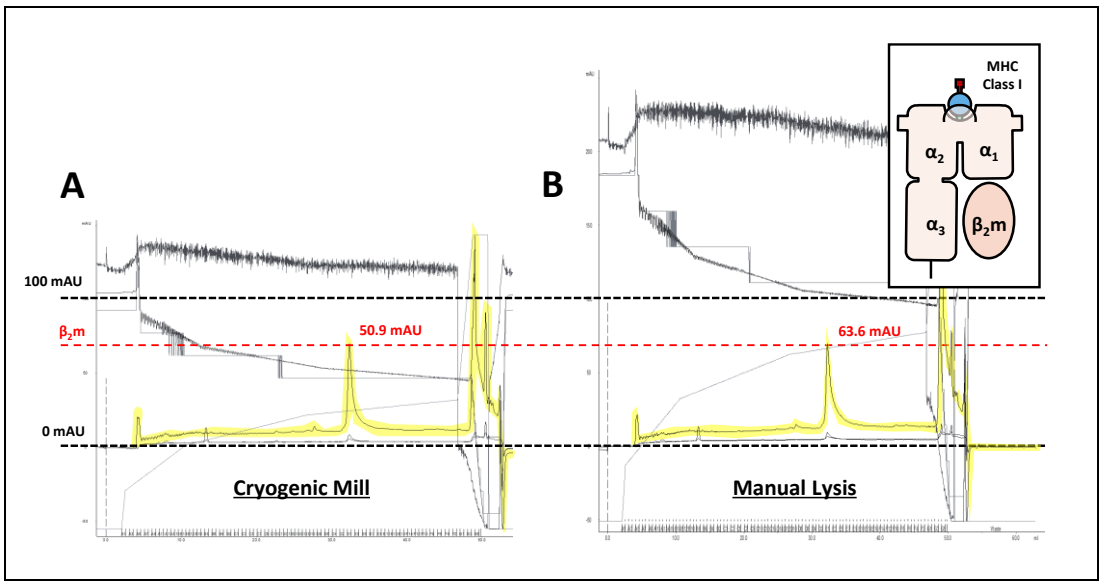


Figure 4.5. Assessment of the cell lysis method for the extraction of MHC I. C1R-B*57:01 (5 x 10^8) cells were lysed using cryogenic milling (A) or manual lysis by pipetting (B) prior to MHC I immunoaffinity capture. Comparable MHC I levels were isolated, as determined by β_{2M} intensity.

Further assessment of MHC I expression on C1R parental and C1R-B*57:01 cells was performed to compare those in Liverpool to those routinely used in the Purcell Lab.

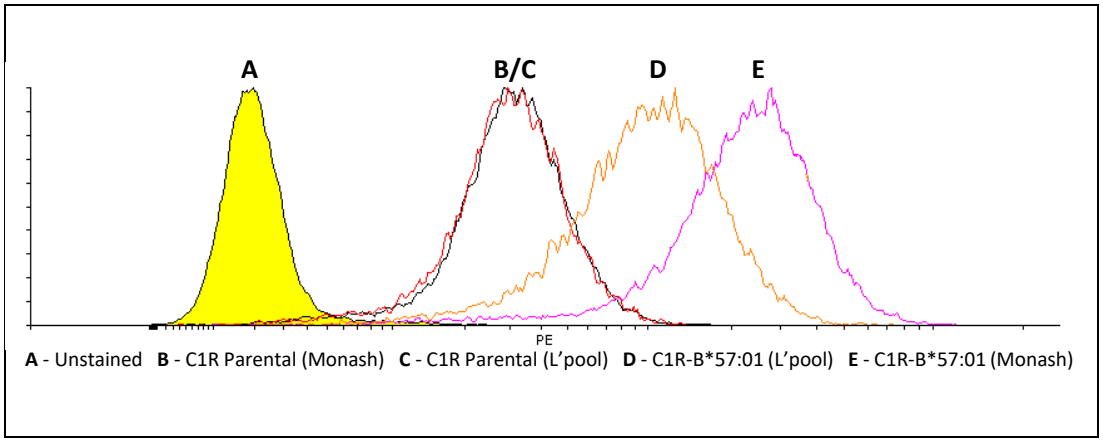


Figure 4.6. HLA-B*57:01 encoded MHC class I surface expression on C1R parental cells and C1R-B*57:01 cells from cultures in Liverpool and Monash University. Parental C1Rs (B & C) equally expressed HLA-B*57:01 encoded surface MHC I. C1R-B*57:01 cells cultured at Monash University (E) showed higher expression of MHC I compared with Liverpool (D), however both showed increased expression when compared to both parental C1Rs and the unstained cells (A).

C1R parental cells expressed similar levels of MHC I in both those cultured in Liverpool and those at Monash (Figure 4.6B & C). C1R-B*57:01 cells in Liverpool (Figure 4.6D) showed lower MHC I expression compared with those cultured in Monash (Figure 4.6E). That said, MHC I expression of C1R-B*57:01 cells from Liverpool were higher compared to parental cells.

To capture MHC class I from cell lysates, anti-HLA-A, B, C antibody (W6/32) was required. For each elution 10 mg of W6/32 antibody was used, therefore in house production was essential due to the costs associated with purchasing such large quantities of antibody. Cell culture of W6/32 lymphoblastoid hybridoma cells that secrete W6/32 antibody was scaled up in order to generate large volumes of cell supernatant. W6/32 antibody was subsequently captured by passing the supernatant through a PAS column. After washing with PBS, antibody was eluted from PAS using mild acid extraction (0.1 M citrate, pH 3.0) and buffer exchanged into PBS to allow for accurate quantification. Supernatant flow through was retained and antibody re-captured for a total of 4 purifications. Flow cytometry was used to assess the level of MHC I staining on C1R-B*57:01 cells by purified antibody and subsequent supernatant flow-through. From this analysis, W6/32 supernatant was found to be present in the supernatant after the first purification (Figure 4.6). Therefore, in order to maximise purification efficiency, supernatant was passed through equilibrated PAS twice prior to being discarded.

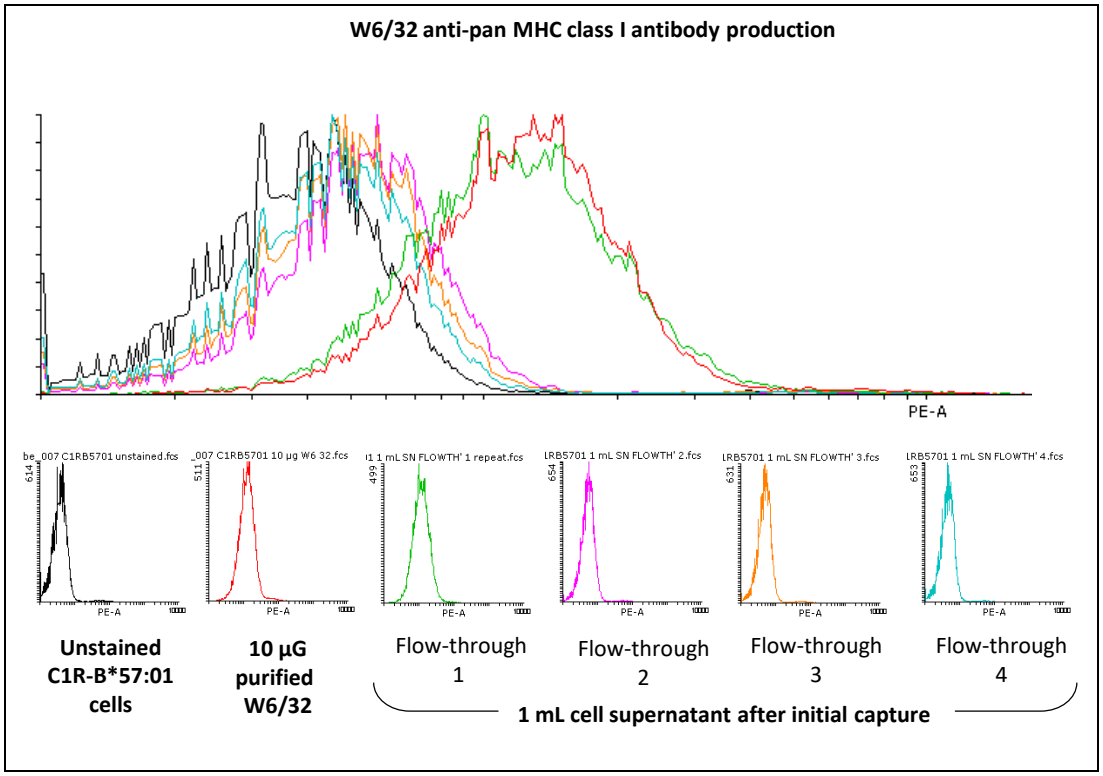


Figure 4.7. W6/32 anti-pan MHC class I antibody production at the University of Liverpool. W6/32 hybridoma cells were cultured to secrete anti-MHC class I antibody for use in the peptide elution protocol. Protein A sepharoase was used to capture W6/32. Flow cytometry was used to validate MHC class I binding on C1R-B*57:01 cells with 10 µG purified W6/32 antibody. Flow cytometry using W6/32 supernatant flow-through reveals antibody is still present after the first purification. After the 2nd purification antibody levels in the supernatant was found to be minimal and PE-A signal was comparable to the unstained. C1R-B*57:01 cells.

Next, peptide elutions were performed. HPLC absorbance traces indicative of MHC I fractionation were used to confirm successful implementation of the methods (Figure 4.8, A). As 10% acetic acid is used to elute peptide-MHC complexes from anti-MHC I antibody, both peptides and the β_2M subunit dissociate from the MHC heavy chain protein. Therefore, HPLC fractionation can be divided into 4 main sections. Firstly, hydrophilic peptides are fractionated, followed by the β_2M subunit. Hydrophobic peptides elute off the HPLC column next, followed by MHC I heavy chain and any other protein contaminants (Figure 4.8A).

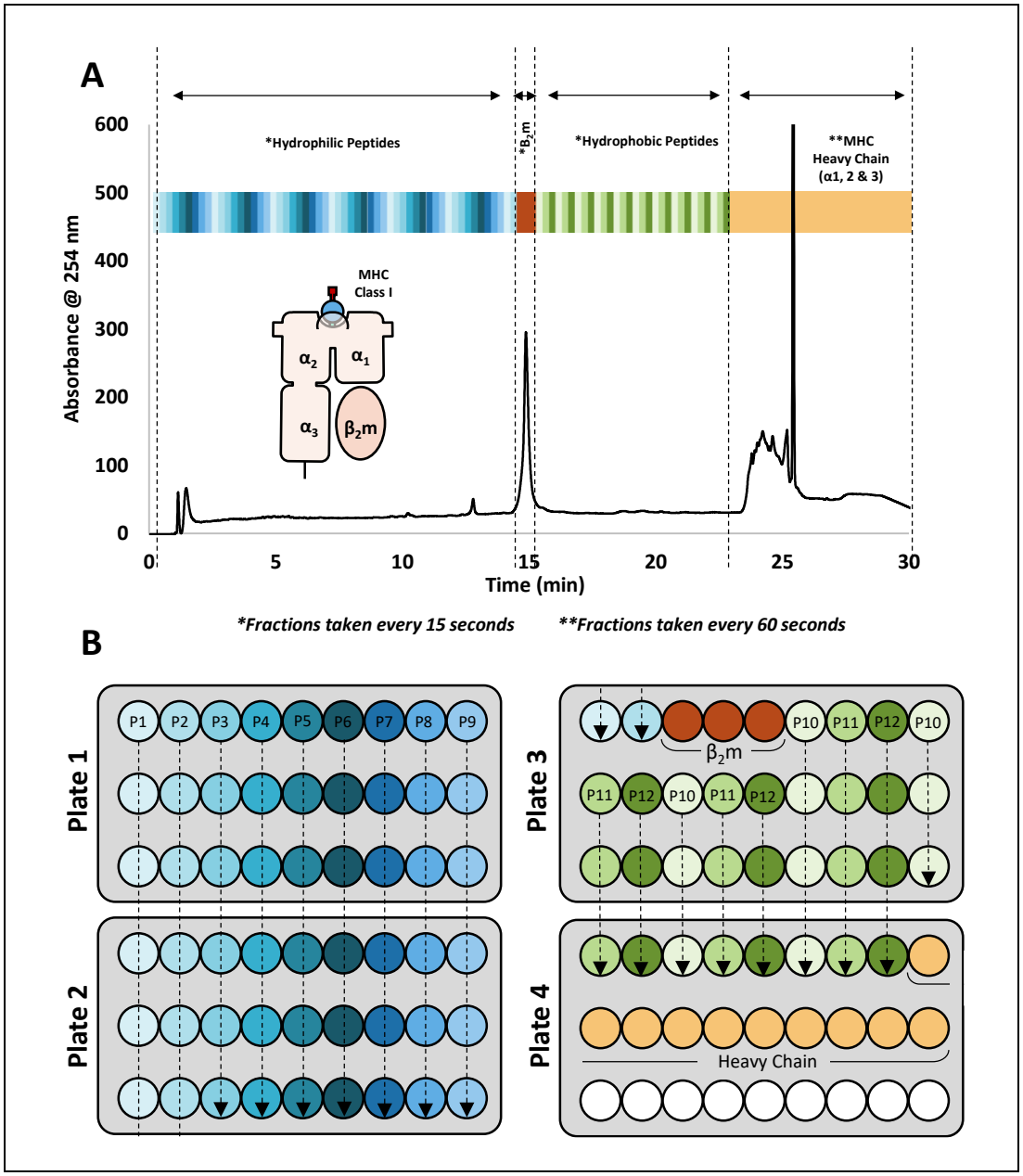


Figure 4.8. Peptide-MHC complex elution and fractionation performed at the University of Liverpool. (A) Fractionation trace with the presence of the β_2M peak and MHC heavy chain protein indicating successful MHC I isolation. (B) Peptide pools are generated by combining fractions across regions, i.e. hydrophilic peptides (blue), β_2M (brown), hydrophobic peptides (green) and heavy chain (yellow).

Sample pre-fractionation is important to enhance the sensitivity when peptides are detected using mass spectrometry. To do this, peptide pools are generated from 99 fractions collected across the HPLC gradient (Figure 4.8B). As a C18 gradient is used for both HPLC fractionation and mass spectrometry, hydrophilic peptide fractions that are collected 2:15 minutes apart are combined to generate a single pool (P1-9). As hydrophobic peptides are not as abundant,

fractions collected 45 seconds apart are pooled (P10-12). β_2M subunits and heavy chain protein were collected separately; fractionation of single protein was not required.

Three independent peptide elutions were performed from untreated C1R-B*57:01 cells (Figure 4.9). Samples were loaded onto the C18 HPLC column using 0.1% TFA in order to wash away the acetic acid they were eluted in. The absorbance of acetic acid at 254 nm allowed for correct sample loading to be determined (Figure 4.9, left). Once washed for 7 minutes at 2 mL/min, the peptide elution gradient was applied resulting in sample fractionation. For optimal peptide absorbance, a wavelength of 215 nm was used (Figure 4.9, right). Experiments were named by the prefix “PE” (peptide elution) followed by the experiment number to allow tracking of different assays within the group. PE_006, PE_008 and PE_027 are all representative of elutions from untreated controls. Through overlaying the HPLC traces (Figure 4.9, bottom) it subsequently observed that the retention time for the β_2M was later for PE_027 (Figure 4.9, bottom, blue box). This is likely due to a temperature variation during fractionation. Importantly, fraction pooling was altered to accommodate this shift. In addition to the retention time shift it was observed that PE_027 has a higher MHC I absorbance reading (Figure 4.9, bottom, orange box). The reason for this remains unclear.

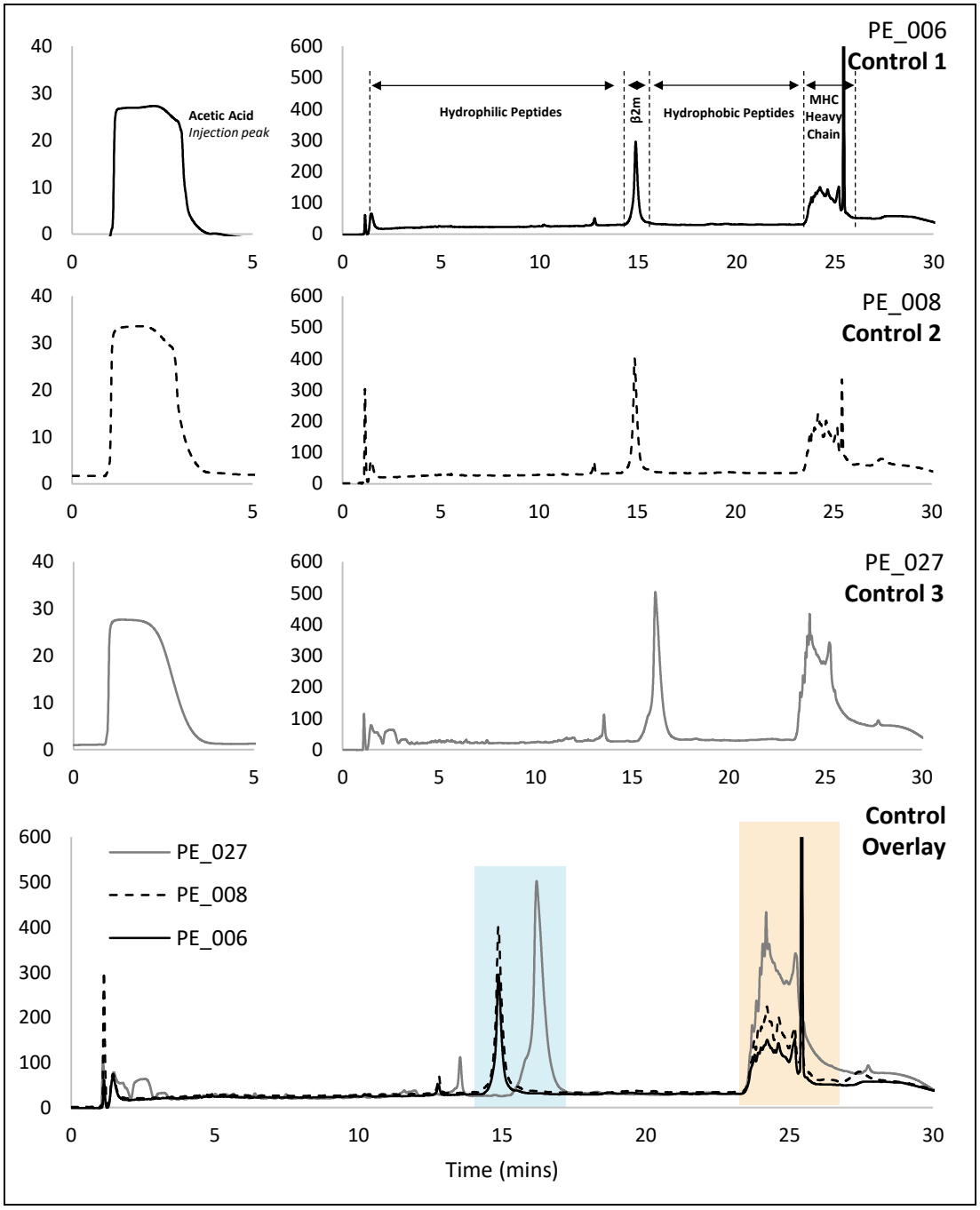


Figure 4.9. HPLC traces of successful untreated control C1R-B*57:01 MHC-peptide elution fractionations. Sample injection was performed separately to fractionation to ensure successful loading prior to the application of solvent gradient (left). An absorbance of 254 nm was used to detect the loading buffer (acetic acid). Subsequent fractionation was performed at 215 nm (right) to improve peptide/protein absorbance. Overlaying traces (bottom) reveals a retention time shift of PE_027 as determined by β_2M (blue box). PE_027 MHC abundance was also presumed higher through increases in both β_2M and heavy chain peaks (blue and orange box, respectively).

In order to confirm the successful implementation of the peptide elution method, C1R-B*57:01 cells were incubated with abacavir. This allowed for the reproduction of the data

observed in 2012 (Illing *et al.*, 2012). Peptide elutions from abacavir treated cells (PE_001, 003 and 004) were all similar in β_2M retention times and MHC quantity (Figure 4.10).

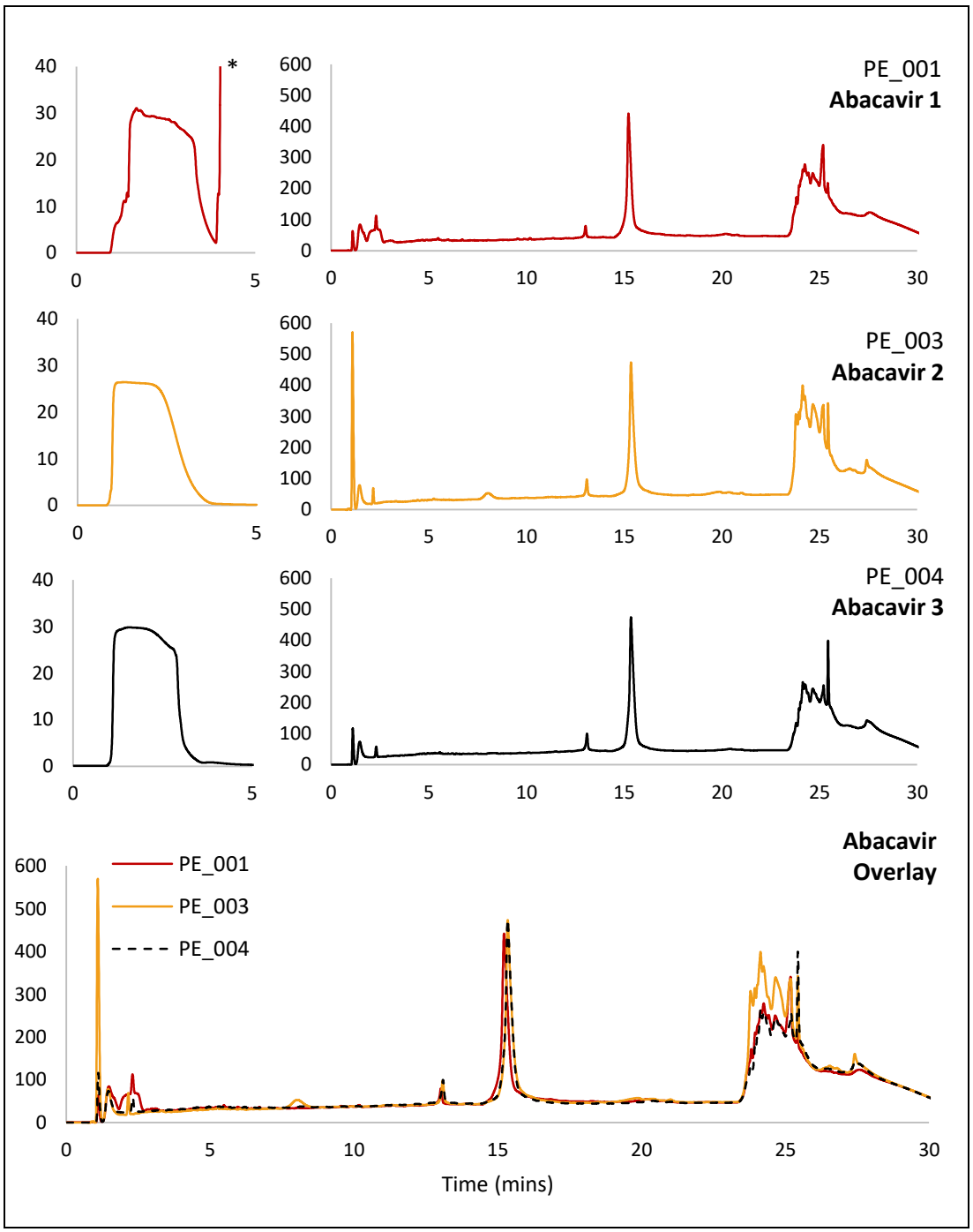


Figure 4.10. HPLC traces of successful abacavir treated C1R-B*57:01 MHC-peptide elution fractionations. Sample injection was performed separately to fractionation to ensure successful loading prior to the application of solvent gradient (left). An absorbance of 254 nm was used to detect the loading buffer (acetic acid). Subsequent fractionation was performed at 215 nm (right) to improve peptide/protein absorbance. Overlaying traces (bottom) reveals a comparable retention times and intensities. *absorbance reading increases due to suspected air bubble after injection.

As mentioned previously, the addition of abacavir to C1R-B*57:01 cells results in a shift in the C-terminal (PΩ) anchor residue. Native HLA-B*57:01 peptide ligands preferentially terminate in phenylalanine (F), tryptophan (W) or tyrosine (Y). Abacavir was found to occupy the F-pocket of the MHC binding grove, meaning these bulky aromatic amino acids can no longer accommodate this region. Therefore, an increase in the smaller aliphatic amino acids leucine (L), isoleucine (I) and valine (V) are reported at the PΩ position when abacavir is present (shown in 4.4.6.1). Post mass spectrometric acquisition of peptide pools, bioinformatic tools are required to translate MS/MS spectra into peptide sequences.

4.4.5 BIOINFORMATIC ANALYSIS OF MHC PEPTIDES

The process of getting from mass spectrometric output files to a list of peptides unique to a particular treatment required the development of a robust workflow (Figure 4.11). While there are a number of other search algorithms available (discussed in later sections), they all work on a similar principle. MS/MS spectra are matched against a protein (or peptide) database with the aim of identifying several fragmentation ions that correspond to a peptide sequence. Once a peptide sequence is assigned, statistical analysis internal to the algorithm will allocate a confidence score to the hit. While this is a good measure of accuracy, it is not particularly useful to compare different data sets. For example, a 95% confidence score in one output may not translate to 95% in different data set. This is due to the percentage ranking being based on the rest of the data in that set. Therefore, FDRs are calculated for each dataset. Here, a decoy database containing reversed peptide sequences (that do not exist in the proteome) are searched against. If a match is found with high confidence it is almost certain to be incorrect. Therefore, based on the number of incorrectly assigned matches an FDR can be calculated. For immunopeptidomic data the level of complexity is very high, hence a 5% FDR is often used. While this is not perfect, having a strict FDR requirement will significantly reduce the number of peptides that can be used for further analysis. The challenges surrounding bioinformatic tools for the assessment of MHC peptides

will be discussed in detail in chapter 6. Briefly, having an unknown protein source and the protease used, along with any biological modifications possible makes it a very large search space within the algorithm. This inherently reduces the number of confidently assigned peptide sequences.

For this study the uniprot/swissprot proteome (downloaded in July 2018), containing 557,896 proteins, was used. With the decoy database included, this results in a search space consisting of 1,115,972 possible proteins, before post translational modifications (PTMs) are considered. By limiting the species to *Homo sapiens*, the search space was reduced to 20,386 proteins, doubling to 40,772 with the addition of the decoy database. The reduction in possibilities saw an increase of identified peptides with >95% confidence from 1,310 to 2,223 in one experiment. Using a protein database solely comprising of human proteins did not result in any additional positive hits. The parameters used in the search were relatively relaxed. Cysteine alkylation was not performed, digestion enzymes were not known and any biological modification could exist. As there were a large number of possibilities, a thorough search effort was applied. For each experiment, 12 different pools were searched simultaneously resulting in a single output file.

The distinct peptide summary, listing all peptides identified, were extracted from each raw output file. For each spectra multiple hypotheses can be formed, with confidence scores in rank order. Therefore, the number one hypothesis was selected based on the first peptide locus. Next, any peptides that were assigned to keratins were removed as these are likely contaminants from the lab environment. Reverse hits, from the decoy database, were also removed. Peptides were sorted based on their confidence score. Only those meeting the 5% FDR for the dataset were retained. As C1R-B*57:01 cells contain low levels of HLA-C*04:01 expression, peptides matching the P2 binding motif (F, Y & W) were removed. The PΩ binding motif (L, I, G, A, V, P, F, M & W) was not used due to the wide range of amino acids and the overlap with HLA-B*57:01 anchors. Other known contaminants identified from previous

experiments within the Purcell group (total number of contaminants in library = 20,845) were also removed. These contained a mixture of peptides from parental C1Rs, class II ligands and other known contaminants. Finally, duplicated peptides were removed. While peptides of the same sequence should elute together, those with PTMs may be acquired at different times. For repertoire analysis, PTMs were not considered.

Next, data from three different experiments were combined to generate a list of peptides with a 5% FDR for each treatment group. At this point peptide length analysis and anchor residue abundance can be performed. For comparisons between treatment groups, unique peptides to each condition were identified. Peptides that were identified across two different sets of confident peptides were removed, leaving only peptides that appeared when the condition was present. These unique peptide lists were taken forward for further analysis. As the unique peptide lists were generated through the combination of replicates, statistical analysis was not possible. Therefore, where statistical calculations were performed, the databases of peptides unique to treatment conditions were used to identify the unique ligands from each replicate. Using this robust workflow (Figure 4.11), we were confident that the peptides we were analysing were genuine HLA-B*57:01 ligands presented under each different treatment condition.

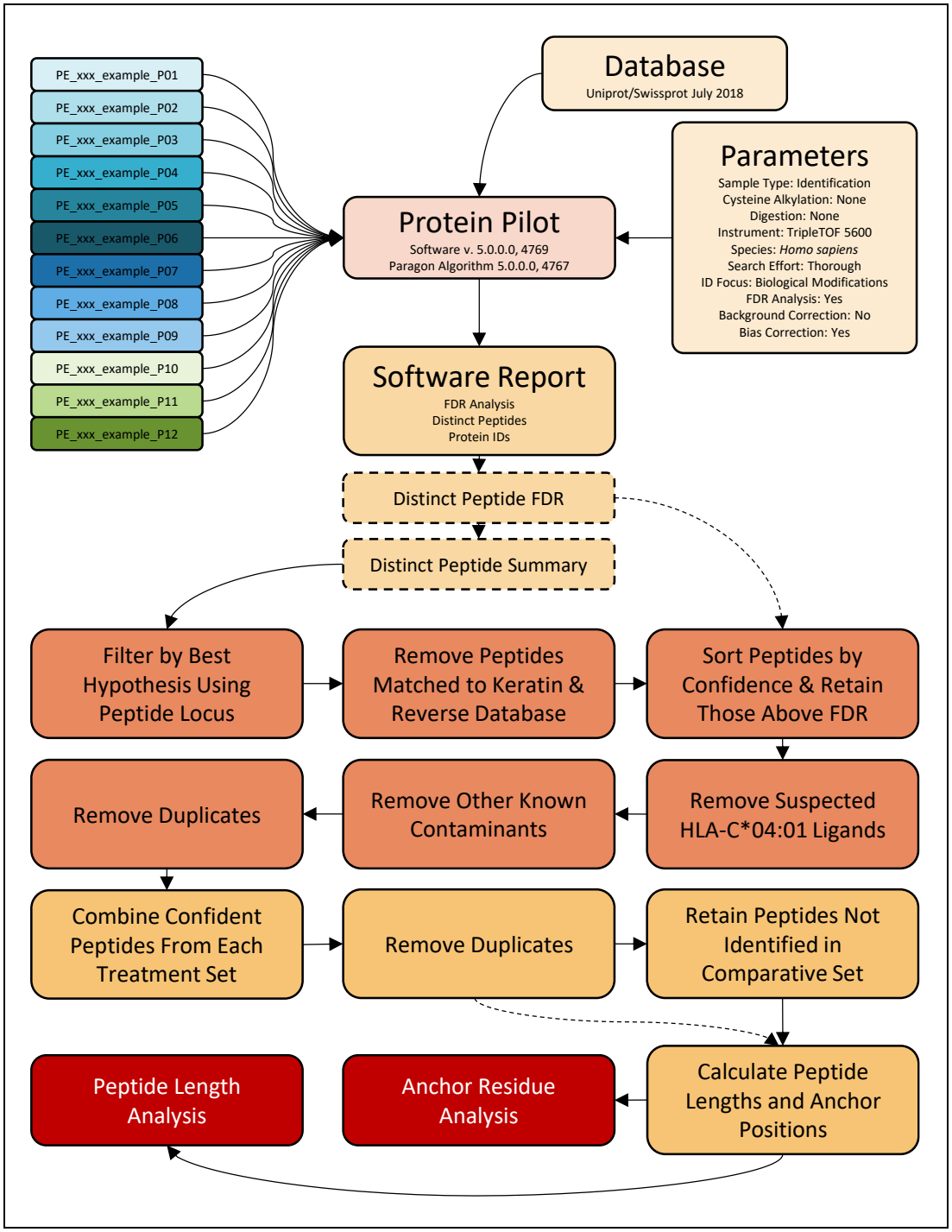


Figure 4.11. Post-acquisition bioinformatic analysis workflow of eluted MHC peptides. Raw mass spectrometry from each experiment are combined and searched against the *Homo sapien* proteome to identify peptides and assign statistical analysis. Relaxed parameters are used due to the complexity of immunopeptidomic analysis. Subsequent analysis is used to removes all suspected contaminants resulting in a list of confident MHC peptides from HLA-B*57:01. Further interpretation is performed to identify peptides unique to treatment groups, before anchor residue and peptide length analysis are investigated.

4.4.6 THE C1R-B*57:01 IMMUNOPEPTIDOME

4.4.6.1 ABACAVIR

Peptides at 5% FDR eluted from untreated C1R-B*57:01 cells (n= 2,189) were compared to those identified from abacavir treated cells (n= 1,013). Unique peptides were determined by comparing both sets of peptides at a 5% FDR. Of the peptides eluted, 1,753 were unique to the control, 437 were shared between treatment groups and 576 were unique to abacavir treatment (Figure 4.12A). The length of peptides eluted from each treatment group were first compared (Figure 4.12B). Although abacavir treatment resulted in slightly fewer 9-mers and slightly more 10-mers, overall the length of the peptides eluted was maintained. Next, comparisons between the P2 anchor residues were interrogated. In line with the previous publication, alanine, serine and threonine were found to be most abundant, with no significant changes observed between treatments (Figure 4.12C, P2). The PΩ anchor residue was the most important comparison to be made due to the repertoire shift previously identified. As anticipated, abacavir resulted in an increased abundance of isoleucine (2% to 28%) and leucine (4% to 16%), in addition to reductions in the abundance of phenylalanine (29% to 10%), tryptophan (45% to 21%) and tyrosine (12% to 6%) (Figure 4.12C, PΩ). In identifying this shift in the HLA-B*57:01 peptide repertoire in the presence of abacavir, the methods to extract, capture and acquire MHC peptide data were confirmed to be well implemented. Furthermore, it was important to see that the criteria used for post-acquisition analysis was robust enough to compare abacavir treated cells to the control.

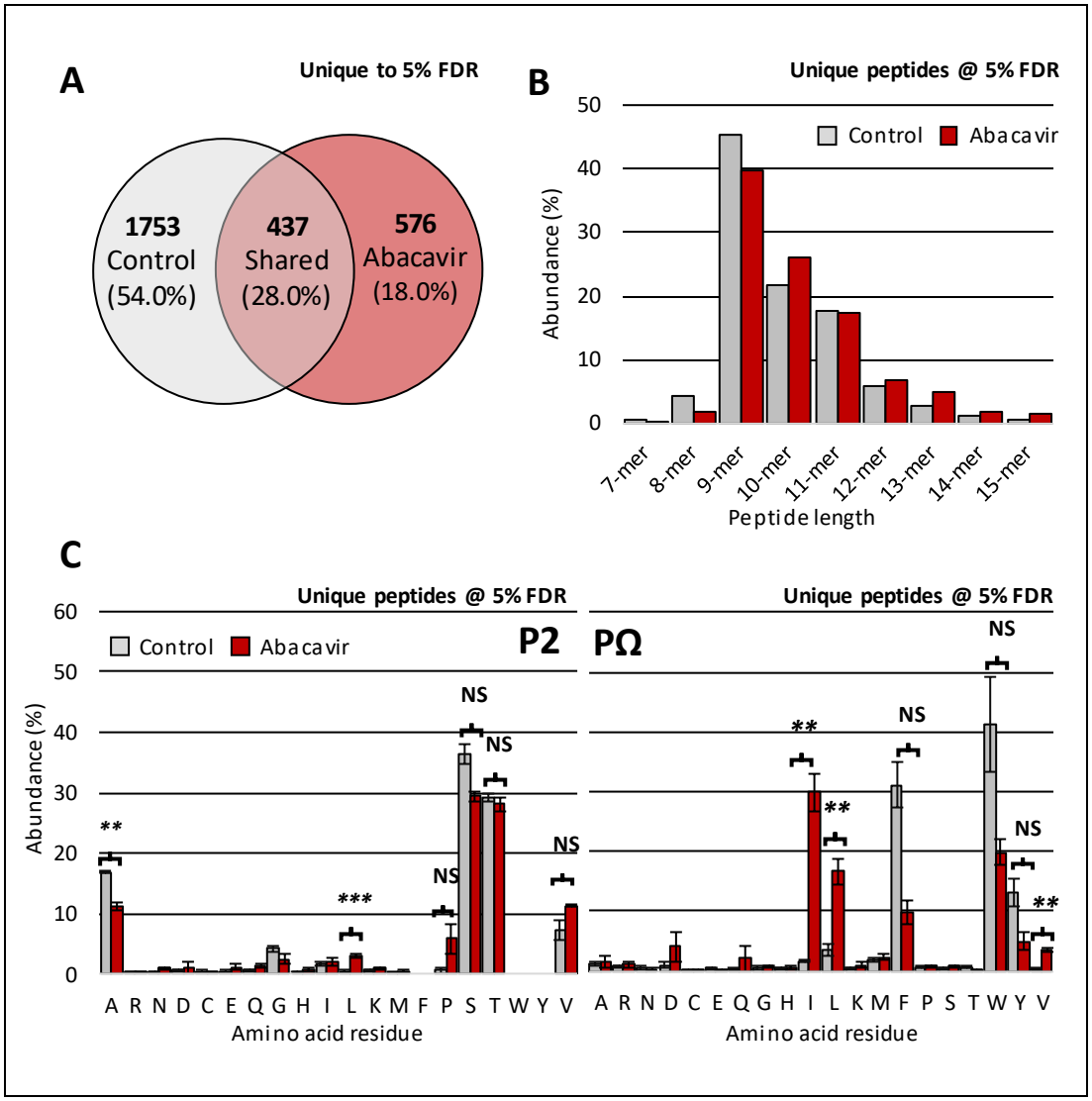


Figure 4.12. In depth analysis of abacavir-associated HLA-B*57:01 ligandome. (A) Total number of peptides identified with a 5% FDR (n= 2,766) from untreated and abacavir-treated cells. From the control cells 1,346 peptides were unique, while 576 peptides were uniquely presented in the presence of abacavir. The remaining 437 peptides were presented in both sets. (B) Length distributions of HLA-B*57:01 ligands across 7-15mers with abacavir treatment (red) and without (grey). (C) The abundance of specific amino acid residues at the P2 anchor residue showed no differences between the two data sets (left). At the PΩ anchor residue binders of HLA-B*57:01 with (red bars) or without abacavir (grey bars) a decrease in phenylalanine, tryptophan and tyrosine, and a significant increase in isoleucine and leucine was observed (right). Two-tailed paired T-test *p*-value within 0.05 (*), 0.01 (**) and <0.01 (***).

4.4.6.2 FLUCLOXACILLIN

After validation of the implementation of peptide elution at Liverpool, attention was turned to flucloxacillin. As flucloxacillin and abacavir hypersensitivity share a genetic predisposition to HLA-B*57:01, interrogation of the immunopeptidome in the presence of flucloxacillin was

an important question to address. Therefore, peptide elution of C1R-B*57:01 cells cultured in the presence of flucloxacillin was performed (Figure 4.13).

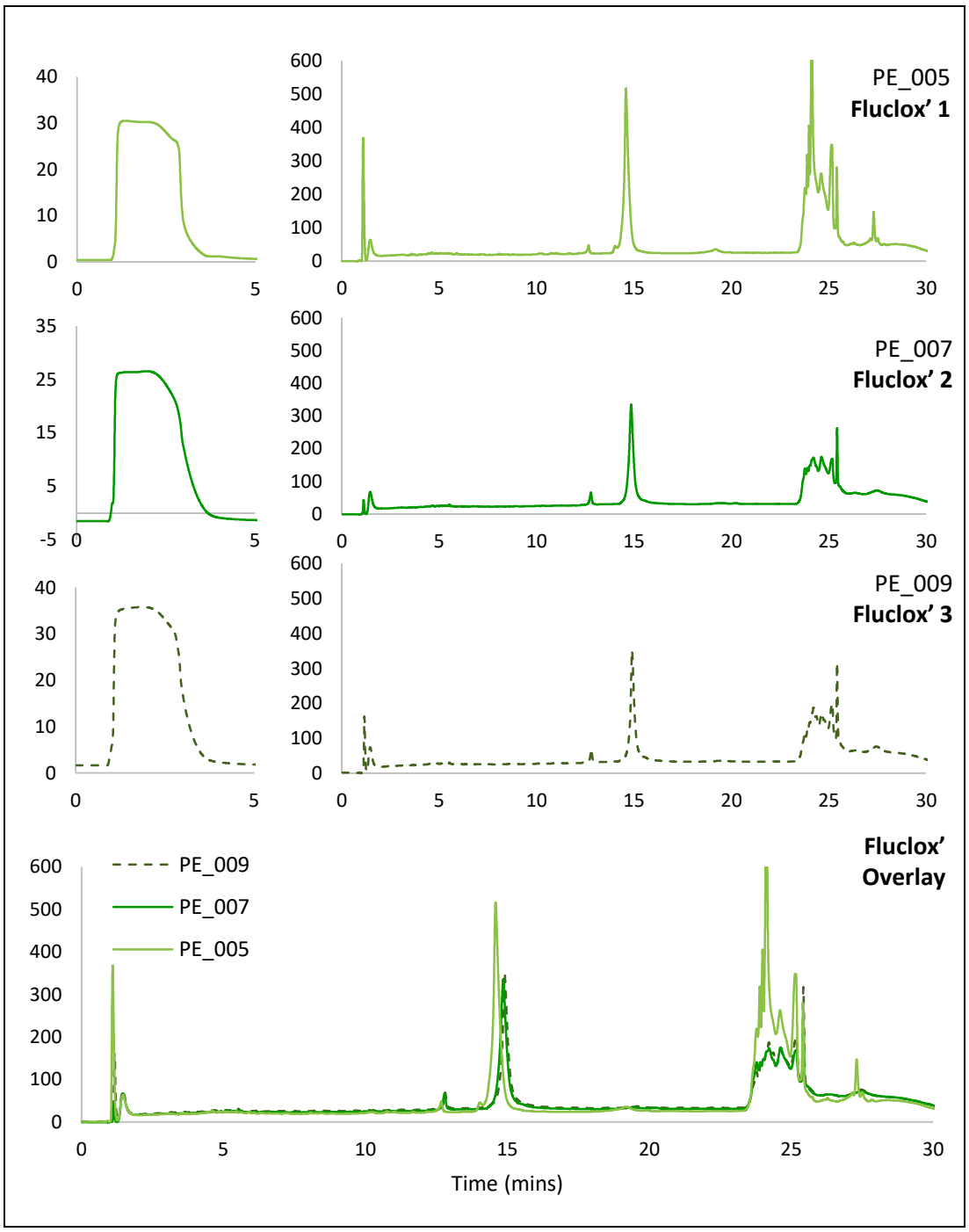


Figure 4.13. HPLC traces of successful flucloxacillin treated C1R-B*57:01 MHC-peptide elution fractionations Sample injection was performed separately to fractionation to ensure successful loading prior to the application of solvent gradient (left). An absorbance of 254 nm was used to detect the loading buffer (acetic acid). Subsequent fractionation was performed at 215 nm (right) to improve peptide/protein absorbance. Overlaying traces (bottom) reveals comparable retention times, with PE_005 showing higher absorbances for both β_2M and MHC heavy chain.

As with previous elutions, samples were injected onto the HPLC on a separate method to the fractionation gradient (Figure 4.13). All β_2M peaks eluted at the same retention time for flucloxacillin treated C1R-B*57:01 cells (Figure 4.13, bottom), with a similar MHC abundance recorded. While the β_2M and heavy chain appears to be slightly higher with the first flucloxacillin elution (Figure 4.13, PE_005), the number of confident peptides identified was not enhanced (PE_005 n= 803, PE_007 n= 755 and PE_009 n= 1,080). Comparative analysis was performed between combined peptides presented by flucloxacillin (n= 1,693) and untreated (n= 2,189) cells. Using a 5% FDR for the identification of unique peptides to each treatment, 1,346 were found exclusively in the control and 850 only identified in the flucloxacillin treated group. This results in 843 (43.5%) peptides being shared between sets (Figure 4.14A). Peptide length distribution revealed no change in the presence of flucloxacillin (Figure 4.14B), in correlation with abacavir. The amino acid abundance at positions across each peptide were examined further to identify any alterations in the presence of flucloxacillin. With flucloxacillin treatment there was no apparent change in the P2 anchor residue (Figure 4.14C, P2), with alanine, serine and threonine remaining abundant. Interestingly, changes were observed in the PΩ anchor residue. The abundance of phenylalanine rose from 21% to 43% in the presence of flucloxacillin, while tryptophan was reduced from 54% to 17%. A small increase in tyrosine was also observed, from 10% to 17% (Figure 4.14C, PΩ).

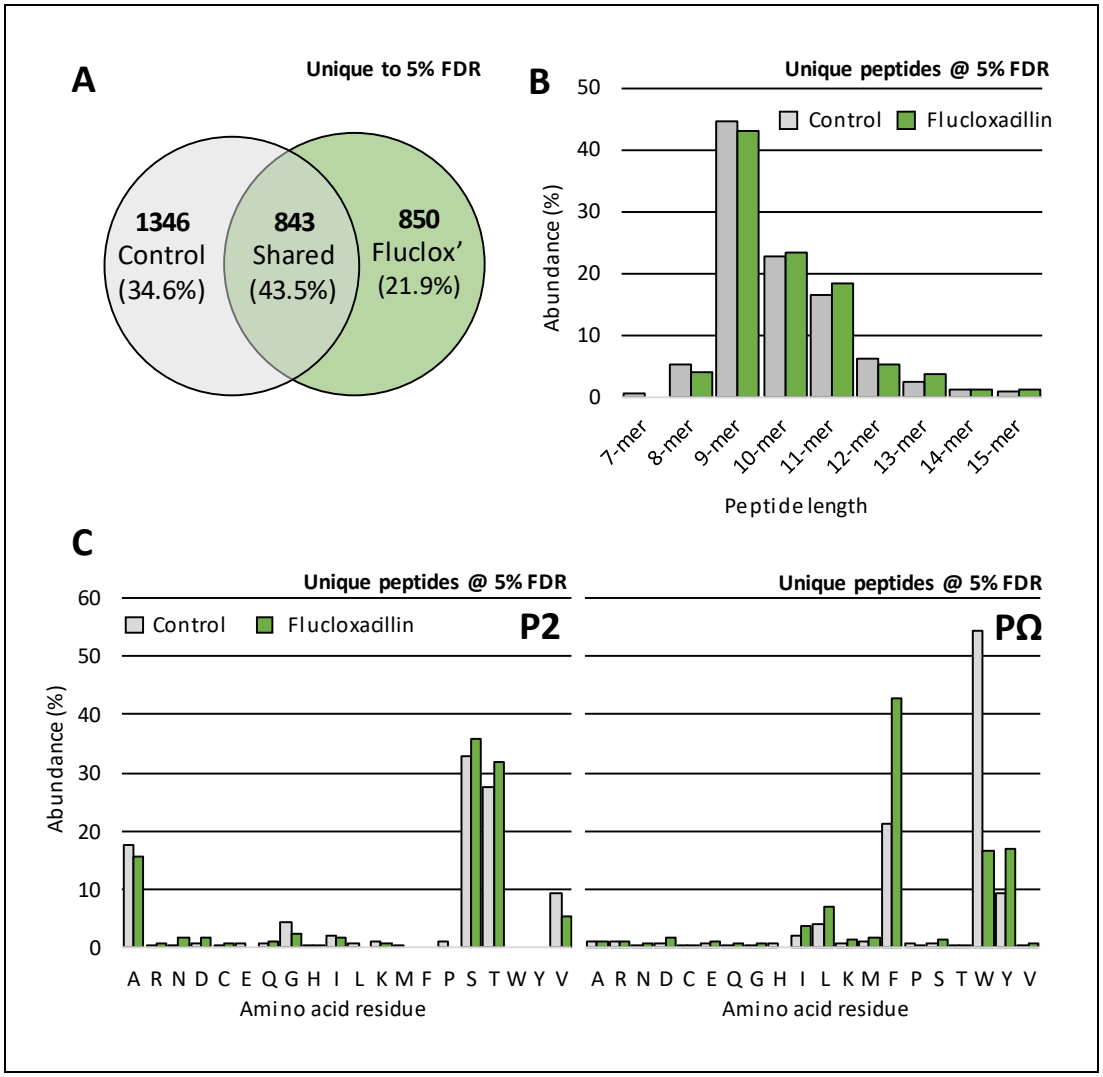


Figure 4.14. In depth analysis of flucloxacillin-associated HLA-B* 57:01 ligandome. (A) Total number of peptides identified with a 5% FDR (n= 3,887) from untreated and flucloxacillin-treated cells. From the control cells 1,346 peptides were unique, while 855 peptides were uniquely presented in the presence of flucloxacillin. The remaining 843 peptides were presented in both sets. (B) Length distributions of HLA-B*57:01 ligands across 7-15mers with flucloxacillin treatment (green) and without (grey). (C) The abundance of specific amino acid residues at the P2 anchor residue showed no differences between the two data sets (left). At the PΩ anchor residue binders of HLA-B*57:01 with (green bars) or without flucloxacillin (grey bars) a decrease in phenylalanine and an increase in tryptophan was observed (right).

As this result was largely unexpected, it was important to properly interrogate the post-acquisition methods to ensure the change in repertoire was not an artefact of the analysis method. Comparing two peptide sets at a 5% FDR directly may result in falsely identified 'unique' peptides. For example the peptide RVWDIVRHY appears in the flucloxacillin treated set within the 5% FDR cut off. In the control set, the same peptide is present, however, with a confidence score of 87.52% it was too low to be included within the 5% FDR. Therefore,

RVWDIVRHY was deemed to be unique to flucloxacillin treatment, when in reality it is likely not. In light of this, further analysis using additional criteria was performed to identify unique MHC peptides. Peptides at a 5% FDR from each treatment group were analyzed against peptides with a confidence score above 20% from the comparative data set. If the peptide at a 5% FDR in the test group appeared in the comparative set with a confidence above 20% it was no longer deemed unique. To make this analysis as robust as possible the retention time, mass difference and any PTMs were also taken into consideration. With the example of RVWDIVRHY, the mass difference between peptide identified from the flucloxacillin treatment was -0.005 Da when compared to that identified in the control set. In addition, the retention time was 0.28 minutes apart. Therefore, RVWDIVRHY was no longer accepted as a unique peptide to flucloxacillin. While in this example the confidence score was still relatively high (87.52%), SAAPLFFSW, previously assumed unique to flucloxacillin treatment, was identified in the control set with a 24.44% confidence. Using the same interrogation method, a mass difference of 0.038 Da and a retention time shift of 0.86 minutes makes it likely that it is not unique to flucloxacillin treatment. The assessment of unique peptides was therefore re-performed using both a 20% confidence in the comparative set; crucially all test peptides were still within the 5% FDR. This new peptide list of unique peptides to the control and flucloxacillin was then used to identify the unique ligands to each replicate, to enable statistical analysis to be performed (Figure 4.15A). By performing the additional analysis, it enabled more confidence to be had in the repertoire change at the PΩ position, with the previous findings being statistically significant, when C1R-B*57:01 cells were cultured in the presence of flucloxacillin.

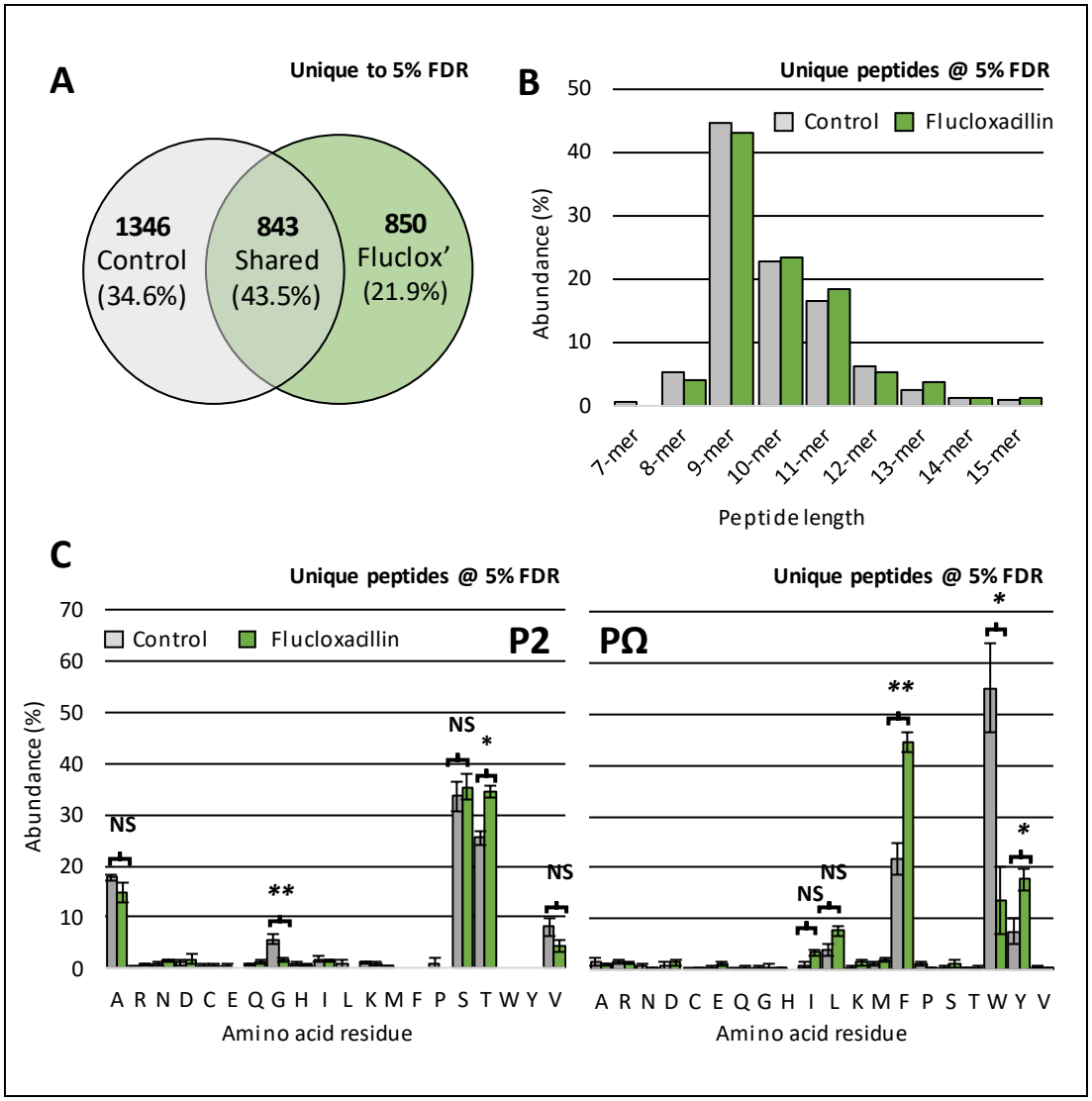


Figure 4.15. Further interrogation of the flucloxacillin-associated HLA-B* 57:01 ligandome. The criteria for defining unique peptides was enhanced to prevent peptides being falsely identified as unique. Test peptides at 5% FDR were compared to peptides from comparative data sets with 20% confidence scores. Retention times, mass differences and PTMs were also taken into consideration. Unique peptides from each replicate were identified. Further analysis revealed statistical separation between PΩ amino acid abundance when C1R-B*57:01 cells were cultured in the presence of flucloxacillin. Two-tailed paired T-test *p*-value within 0.05 (*), 0.01 (**) and <0.01 (***).

4.4.6.2.1 ANCHOR RESIDUE AMINO ACID ABUNDANCE AND PEPTIDE LENGTH

Ice Logos (Colaert *et al.*, 2009) were used to visualize amino acid abundance at each position, relative to natural abundance, for peptides of different lengths (Figure 4.16A-E). Across all 8-12mers tryptophan was abundant in the PΩ position in the untreated set (Figure 4.16, left) whereas compared to phenylalanine in flucloxacillin unique peptides (Figure 4.16, middle). This was the only major observable difference (Figure 4.16, right) across all sets.

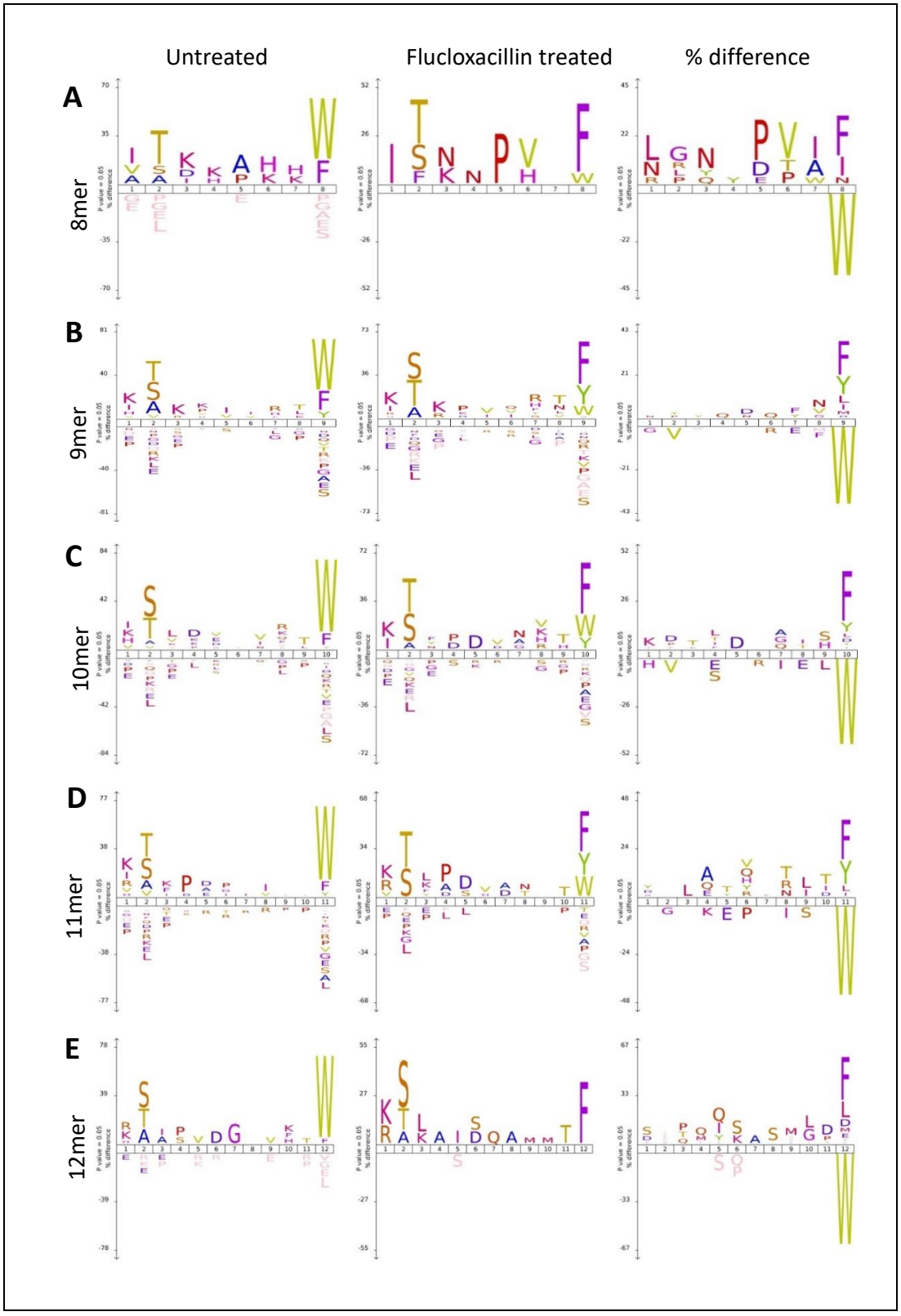


Figure 4.16. Sequence motif analysis of HLA-B*57:01 peptides presented in the presence of flucloxacillin. Peptides eluted from untreated (left) and flucloxacillin-treated (middle) cells further highlights the differences observed at PΩ (right). Peptide repertoire data is shown for 8-mer to 12-mers (A-E, respectively).

4.4.6.2.2 NON-COVALENT INTERACTIONS BETWEEN FLUCLOXACILLIN AND HLA-B*57:01

While the repertoire change appears genuine, the precise mechanism for this remains largely unknown. However, modelling of penicilloic acid and flucloxacillin with HLA-B*57:01 revealed that it is possible for both compounds to occupy the C-F binding pockets in HLA-B*57:01 (Figure 4.17). Flucloxacillin penicilloic acid is the result of the hydrolysis of the β-lactam ring. Close interactions between flucloxacillin penicilloic acid with the key amino acid residues at the binding groove, for example, Asn77, Asp114, Ser116, and Tyr123 may have altered the presentation of peptides through stabilising phenylalanine at the PΩ position, leading to an increased presentation of peptides terminating in phenylalanine.

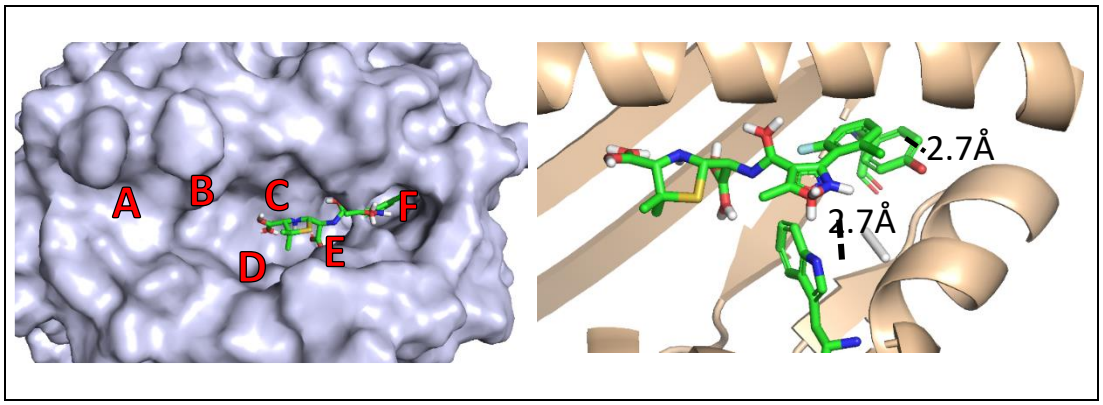


Figure 4.17. Interaction of flucloxacillin and the HLA-B*57:01 binding groove. Modelling studies show that flucloxacillin penicilloic acid can occupy C-F binding pockets in HLA-B*57:01, interacting closely with Tyr123 and Trp147.

If an interaction between penicilloic acid and/or flucloxacillin and HLA-B*57:01 is occurring in a non-covalent manner (as modelled), it would be anticipated that free drug would be detected during HPLC fractionation. Flucloxacillin was not previously identified as eluting from peptide-MHC complexes during fractionation (Figure 4.18), however this could be simply due to limits of detection.

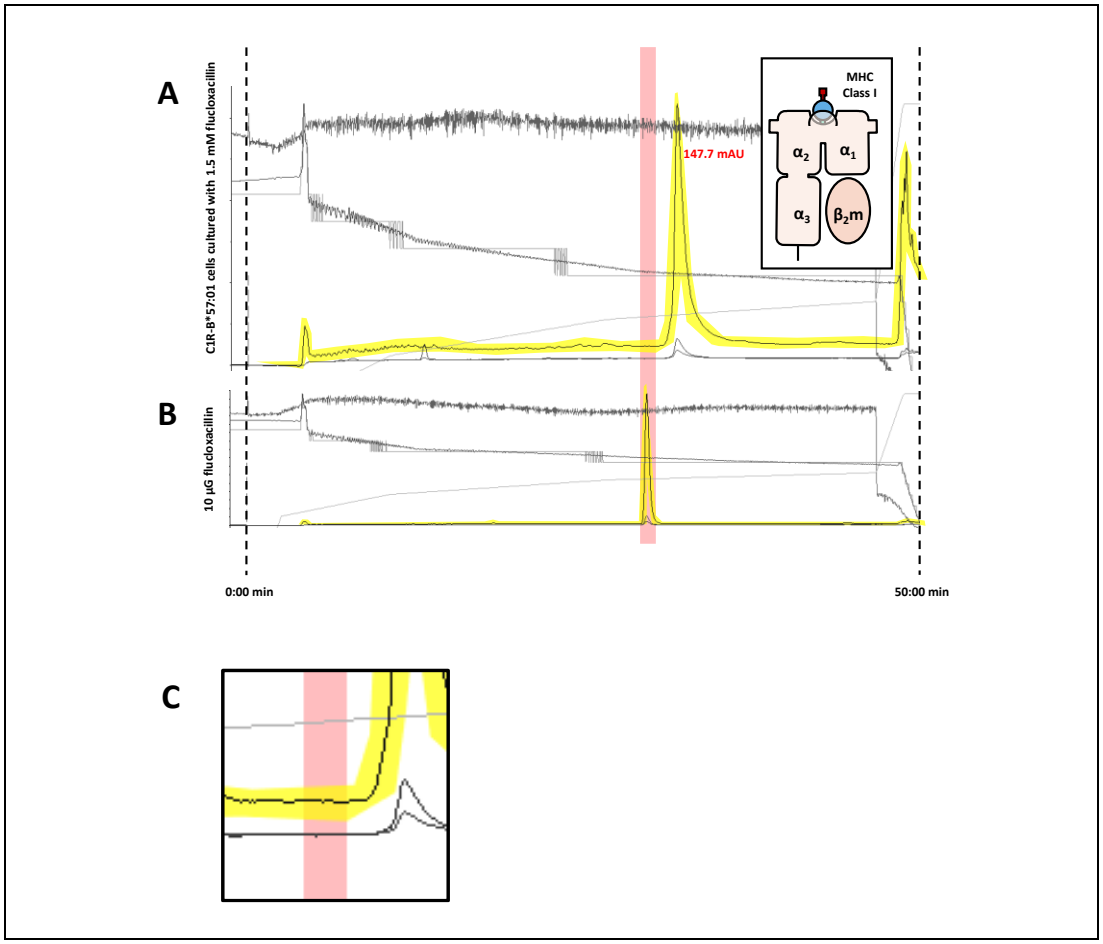


Figure 4.18. Determination of non-covalently bound flucloxacillin to MHC I in C1R-B*57:01 cells cultured in the presence of drug. (A) Peptide-MHC complexes were eluted from C1R-B*57:01 cells cultured in the presence of 1.5 mM flucloxacillin for 48h and fractionated using HPLC. (B) Flucloxacillin retention time was determined by loading 10 µG using the same parameters. (C) Unbound flucloxacillin was not detected in the peptide-MHC complex fractionation when C1R-B*57:01 were cultured in the presence of flucloxacillin.

The presence of free drug was subsequently investigated in the mass spectrometric data acquired from flucloxacillin treated cells. Indeed, both flucloxacillin (*m/z* 454.0696) (Figure 4.19A) and penicilloic acid (Figure 4.19B) could be identified. Reassuringly, penicilloic acid was found to be present in relatively large quantity, as demonstrated by the AUC of the parent ion (*m/z* 472.2989) (Figure 4.19C).

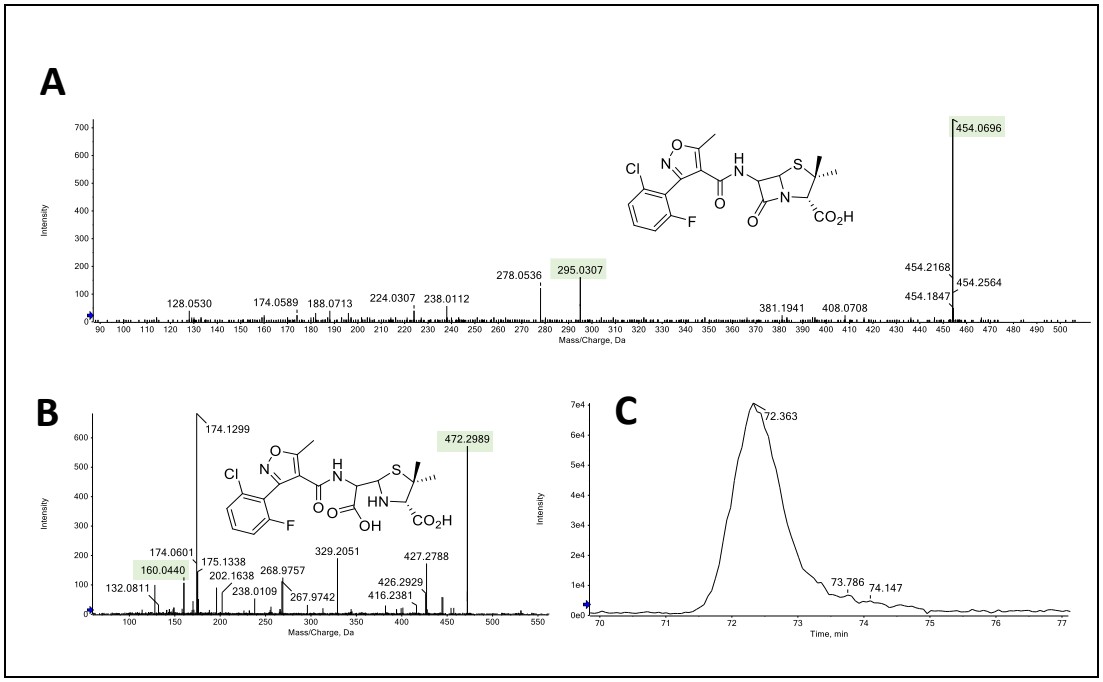


Figure 4.19. Flucloxacillin and degradation products identified in the peptide-MHC complex. (A) Flucloxacillin (m/z 454.0696) and (B) its degradation product, flucloxacillin penicilloic acid (m/z 472.2989) were detected in the peptide-MHC-Complex eluted from C1R-B*57:01 cells, indicating both compounds are involved in MHC peptide presentation. (C) The extracted ion count of the penicilloic acid shows it to be in relatively high abundance.

*4.4.6.2.3 PEPTIDE BINDING AFFINITY TO HLA-B*57:01*

To further assess the impact of flucloxacillin on the self-peptide presentation, the binding affinity of eluted peptides to HLA-B*57:01 was predicted using NetMHC 4.0. Analysing binding affinities of unique peptides presented from each treatment group at a 5% FDR (control 1,346 vs flucloxacillin-treated 850), we observed fewer peptides bound to HLA-B*57:01 with high affinity in the flucloxacillin dataset (control 42.3% vs flucloxacillin-treated 22.2%) (Figure 4.20A, i). The binding affinity of the eluted peptides to other HLA alleles was also predicted using the same peptide lists. As expected, the number of peptides that bind to the closely related HLA alleles, such as HLA-B*58:01 (Figure 4.20A, ii) stay relatively the same. As anticipated peptide binding affinity to the distantly related HLA-B*15:01 decreased in both sets of data (Figure 4.20A, iii). Interestingly, both the control and flucloxacillin unique peptides appear to have very similar affinity to HLA-B*15:17 (Figure 4.20A, iv).

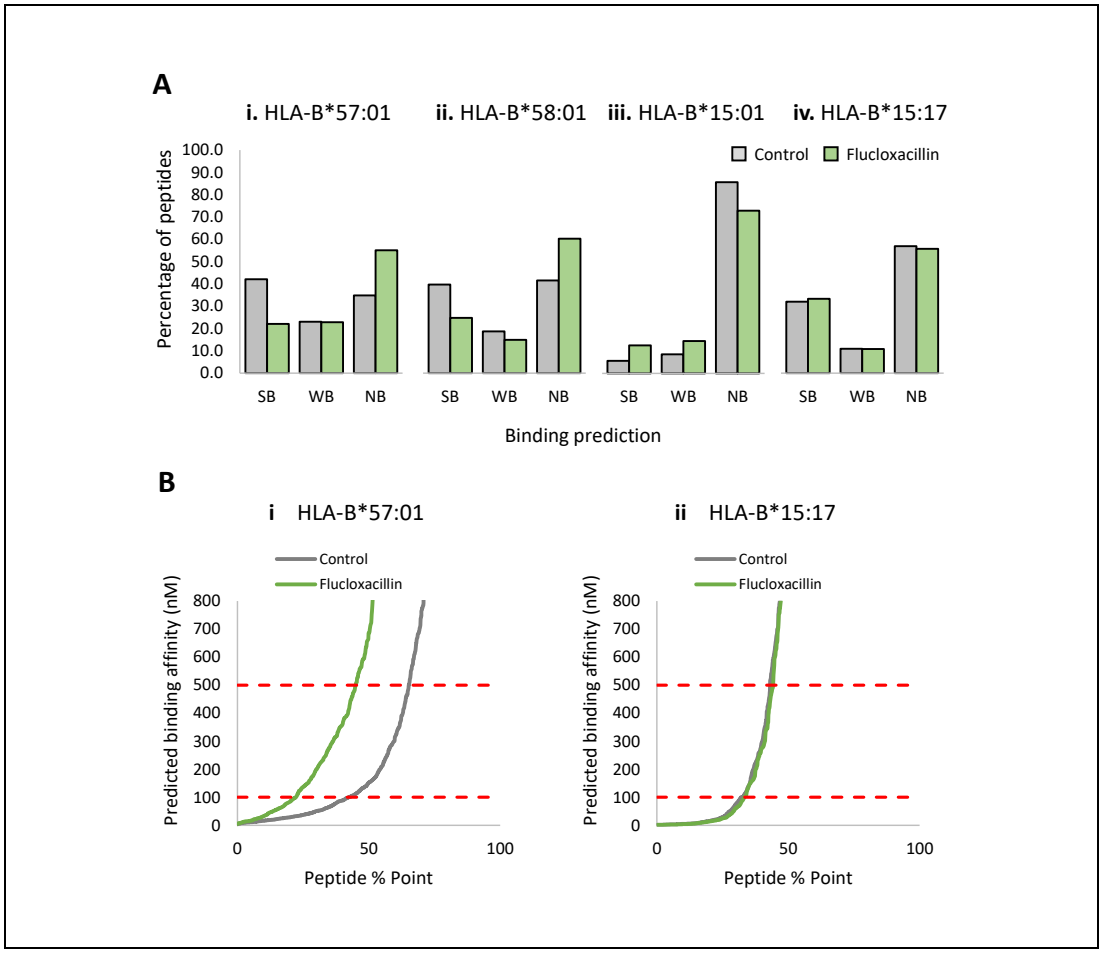


Figure 4.20. Prediction of binding affinity of eluted peptides to HLA-B alleles. (A) Predicted binding affinity of unique peptides eluted from untreated (grey bars, 1,346 peptides) and flucloxacillin-treated cells (green bars, 850 peptides) was performed using NetMHC 4.0. A decrease in the number of peptides with high binding affinity to HLA-B*57:01 was observed with flucloxacillin treated cells compared to untreated cells (i). Binding affinity was similar in the closely related allele HLA-B*58:01 (ii), however was lost in both treatment groups when binding affinity was assessed to the unrelated allele HLA-B*15:01 (iii). Peptides unique to flucloxacillin treatment appear to bind more favourably to the unrelated HLA-B*15:17 allele (iv). (B) The number of peptides with high (<500 nm) and very high (<100 nm) affinity for each treatment group. The number of peptides unique to flucloxacillin treatment with a very high binding affinity increases from 22.2% with HLA-B*57:01 (i) to 33.5% with HLA-B*15:17. SB-strong binder, WB-weak binder, NB-non binder

Analysis of the binding affinity of individual peptides from each treatment group to HLA-B*57:01 and HLA-B*15:17 was performed (Figure 4.20B). The percentage of strong binders (SB) (<100 nm) and weak binders (WB) (<500 nm) to HLA-B*57:01 is greater in the control group than in the flucloxacillin treated group (Figure 4.20B, i). Surprisingly, an increase in the number of SB to HLA-B*15:17 in the flucloxacillin treated group (33.5%) when compared to those to HLA-B*57:01 (22.2%) were observed (Figure 4.20B, ii). This may indicate that the peptides being presented by HLA-B*57:01 in the presence of flucloxacillin are more similar

to those being presented by HLA-B*15:17. It is possible that the presence of flucloxacillin (penicilloic acid) may stabilize the presentation of peptides which otherwise have low predicted binding affinities.

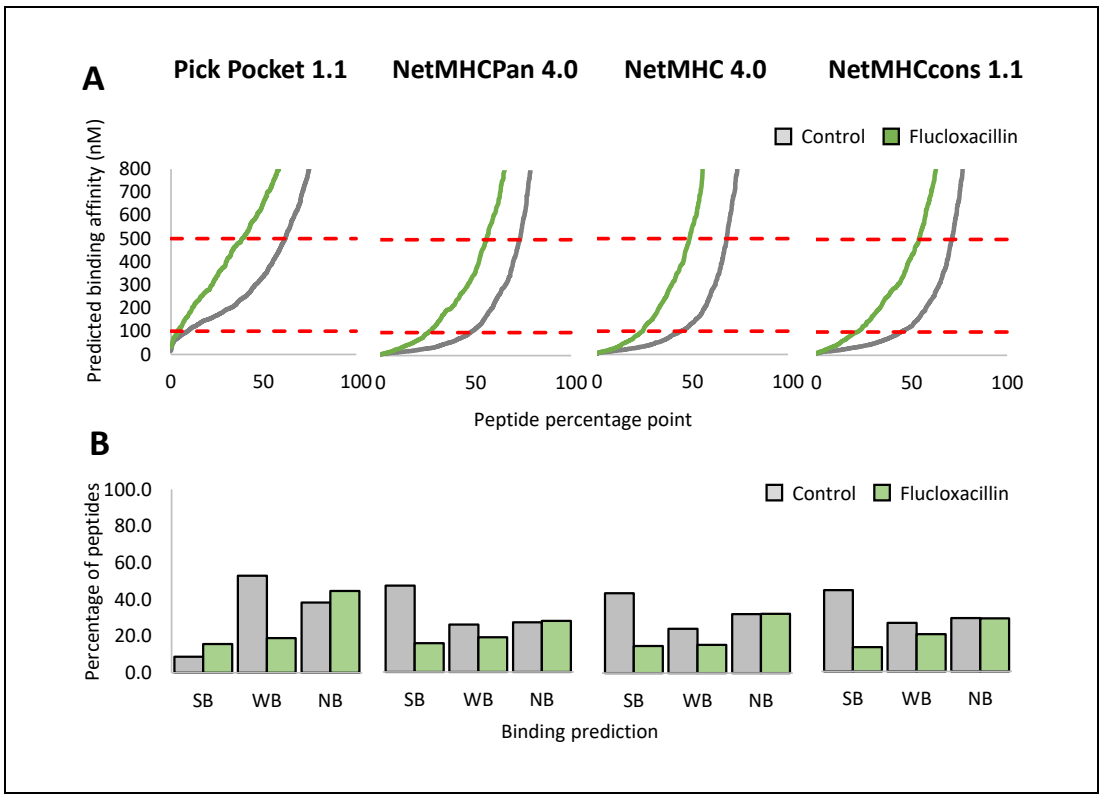


Figure 4.21. Comparison of available algorithms for the assessment of peptide binding affinity. Four different algorithms, Pick Pocket, NetMHCPan, NetMHC and NetMHCcons, were used to predict the binding affinity of unique peptides to HLA-B*57:01. (A) While small differences were observed with Pick Pocket compared to the other 3 algorithms, separation between flucloxacillin and untreated unique peptides remained. (B) Total MHC binding for each algorithm reveals that similar conclusions can be drawn from all 4. SB-strong binder, WB-weak binder, NB-non binder

When assessing peptide binding affinity using predictive tools the choice of modelling algorithm used will determine the conclusions that can be drawn. For the analysis already discussed NetMHC 4.0 was used, however a number of other algorithms including Pick pocket, NetMHCPan and NetMHCcons are available for predicting MHC peptide binding affinity to specific HLAs. NetMHC and NetMHCPan are both artificial neural network (ANN) based allele-specific methods trained on 94 class I alleles and 115,000 quantitative binding data covering 120 different MHC molecules, respectively. Pick Pocket produces results on a matrix-based method reliant on receptor-pocket similarities between MHC molecules,

having been trained on 94 different MHC alleles. NetMHCcons is a consensus method integrating results from all three of the other algorithms. By predicting the binding affinity of both individual peptides (Figure 4.21A) and the immunopeptidome as a whole (Figure 4.21B) to HLA-B*57:01 for unique peptides from each treatment, data separation was observed irrespective of the algorithm used. Pick Pocket did predict fewer high binders, likely due to the difference in the matrix-based method rather than ANN, however a similar separation of individual peptides can be observed. In all instances, peptides unique to flucloxacillin treatment interact less favourably with the HLA-B*57:01 binding pocket.

4.4.6.2.4 PROTEIN SOURCES OF UNIQUE PEPTIDES

So far, changes to the HLA-B*57:01 immunopeptidome in the presence of flucloxacillin has been entirely focussed at the peptide level. Up to this point it was unclear as to whether flucloxacillin changed the abundance of specific proteins in the cell cytoplasm. If this were the case, it is possible that changes in the proteome could result in altered peptide presentation. Therefore the protein sources of unique peptides from each treatment group were compared using the Panther 14.1 classification system (Mi, Muruganujan, Ebert, *et al.*, 2019; Mi, Muruganujan, Huang, *et al.*, 2019). Different parameters were used to interrogate any changes between treatment groups, including molecular function, biological process, cellular component, protein class and pathway involvement. Interestingly it was not possible to identify any change between treatment groups across all the parameters that were used (Figure 4.22).

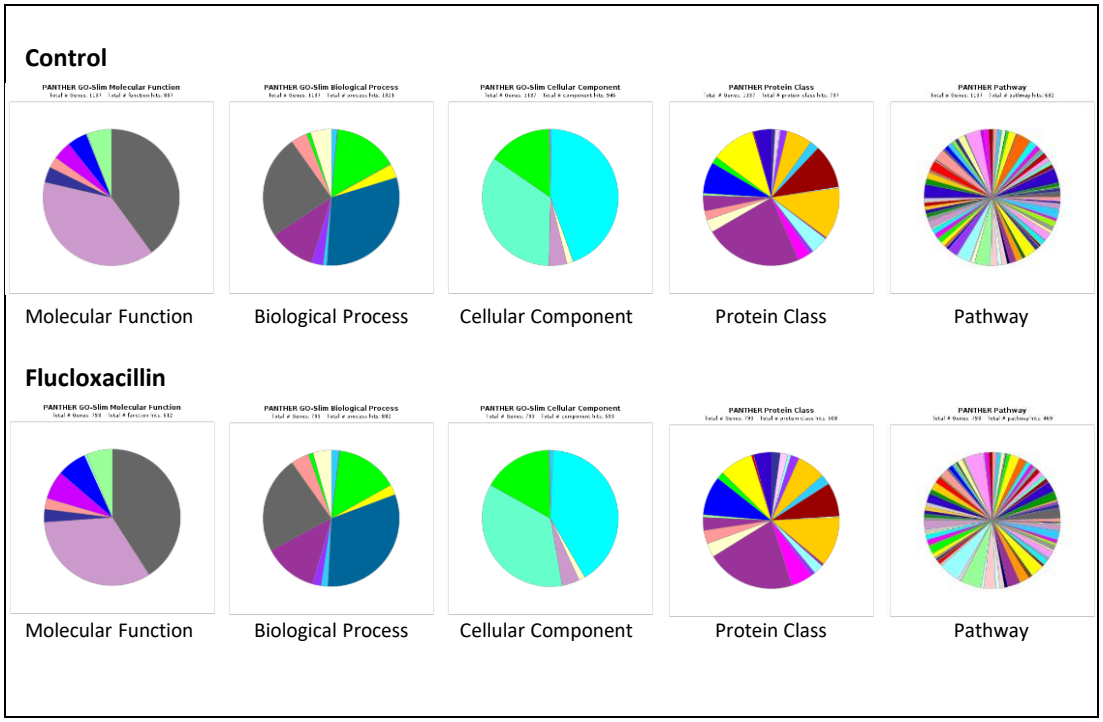


Figure 4.22. Protein sources of unique peptide identified from each treatment group. The molecular function, biological process, cellular component, protein class, and pathway involvement of the protein sources of unique peptides were compared between treatment groups using Panther 14.1. No changes between treatments were observed across all parameters.

4.5 DISCUSSION

Previous studies demonstrated alterations in the immunopeptidome of HLA-B*57:01 when C1R-B*57:01 cells were cultured in the presence of abacavir (Illing *et al.*, 2012; Norcross *et al.*, 2012; Ostrov *et al.*, 2012). To date, this is the only known example of the altered repertoire hypothesis in DHRs. In this chapter, the main aim was to develop the methods necessary to perform MHC class I elution studies in order to interrogate the immunopeptidome of C1R-B*57:01 cells in relation to flucloxacillin. Through a collaborative exchange, methods employed by Dr Illing under the supervision of Prof Purcell were optimised for transfer to the CDSS at the University of Liverpool. In this optimisation, the relative MHC class I expression of C1R-B*57:01 cells were comparable to patient derived APCs and was not affected by both drug treatment and culture conditions. A concentration of 1.5 mM flucloxacillin was used to identify changes to the immunopeptidome without toxicity towards C1R-B*57:01 occurring. The clinical relevance of this concentration is debatable; the maximal flucloxacillin concentration in patient plasma has previously been identified as 41.4 μM (Jenkins *et al.*, 2009), 36x lower than the concentration used in this study. That said, *in vitro* patient T cell responses to flucloxacillin are activated at 1.5 mM. Furthermore, when 2 mM piperacillin was used for T cell activation in another study, the levels of piperacillin adduct formation identified from the incubation media was comparable to the level of binding in patient plasma (Meng *et al.*, 2017). In the previous chapter localization of flucloxacillin was identified in the bile canaliculi of HepaRG cells, indicating at the site of injury the physiological concentration could be much higher. In fact, a recent study identified a plasma concentration of 209 mg/L (0.46 mM) upon repeated exposure of the drug, however it is well noted that patient variability exists (Maier-Salamon *et al.*, 2017). While this figure is still > 3x lower than the concentration used, free flucloxacillin will likely have been quenched by serum proteins resulting in a lower concentration ultimately interacting will the C1R-B*57:01 cells.

After successful implementation of the peptide elution methods was achieved, the analysis of the immunopeptidome of abacavir treated C1R-B*57:01 cells was performed to act as a positive control. Our data demonstrated that an increased abundance of leucine and isoleucine occurs at the C-terminal anchor residue, which are consistent with previous results (Illing *et al.*, 2012). Similar approaches were therefore used to investigate flucloxacillin associated HLA-B*57:01 immunopeptidomics. Flucloxacillin treatment was associated with an increase in phenylalanine and a decrease in tryptophan at the C terminus of the peptides presented by HLA-B*57:01, irrespective of the peptide length. Further evaluation of the bioinformatic workflow reveals the same overall change in repertoire, therefore hypotheses for this shift were postulated. Modelling of flucloxacillin penicilloic acid to HLA-B*57:01 revealed that it may reside within the C-F pockets of the antigen-binding cleft, with the thiazolidine ring and carboxylic acid positioned into the D and E pockets. The aromatic ring and isoxazole side chain pointed towards the F pocket, which would disfavour the presence of bulky aromatic side chain of tryptophan at the PΩ position but favour phenylalanine due to the close contacts between aromatic rings. Indeed, using mass spectrometry, relatively high quantities of flucloxacillin penicilloic acid were detected. The interaction of flucloxacillin penicilloic acid with HLA-B*57:01 provided a basis for understanding the observed preferences of phenylalanine at PΩ. Furthermore, the presence of flucloxacillin and penicilloic acid at the binding cleft could change the chemistry of peptides binding to HLA-B*57:01.

Using HLA peptide binding affinity prediction tools, peptides unique to flucloxacillin treatment were found to bind with lower affinity to HLA-B*57:01. While different prediction tools exist, four different algorithms presented the same separation in the two unique peptide data sets. As anticipated, the binding affinity to closely related alleles such as HLA-B*58:01 were much the same, while distant related alleles such as HLA-B*15:01 showed weak binding affinity. Interestingly, peptides unique to flucloxacillin treatment had

comparable binding affinity with control peptides to HLA-B*15:17. This indicates that the peptides presented in the presence of flucloxacillin are similar to peptides that are usually accommodated by other HLA alleles. As such, the co-presentation of drugs and peptides that would not usually accommodate the native HLA binding repertoire could create a novel T cell epitope that may be implicated in the development of autoimmune-like reactions. The concept of transplant rejections is important when considering the impact of drugs on the MHC peptide repertoire. Direct recognition of allogenic MHC complexes can occur in the absence of peptides, however foreign peptides are still implicated in the subsequent T cell response.

Adoptive T cell therapy is a powerful technique whereby patient T cells primed *ex vivo* to targets are used to neutralize cells expressing a particular antigen. Its success has been well documented in disease treatment (Magalhaes *et al.*, 2019) however in a small number of studies catastrophic consequences have occurred. In 2013, Cameron *et al* investigated the mechanisms responsible for off target toxicity resulting in the mortality of patients undergoing adoptive T cell therapy. Affinity enhanced T cells were directed to HLA restricted MAGE A3 antigens (EVDPIGHLY), a common antigen expressed across several tumours with limited expression in healthy tissue. Pre-clinical screening showed no cross reactivity or cross target antigen recognition. Unexpectedly, these engineered T cells cross reacted with a peptide derived from the muscle protein Titin (ESDPIVAQY), the most likely cause of the observed toxicity (Cameron *et al.*, 2013). While adoptive T cell therapy is not necessarily linked to this study, it does show the power of the immune system when it comes to antigen recognition. Here, patient derived T cells were responding to antigens presented by their own MHC molecules, highlighting the importance of the peptide itself. In the current study, novel peptides were present when C1R-B*57:01 cells were incubated in the presence of flucloxacillin. Therefore, it is feasible that alterations in the peptides alone, as a result of interactions with flucloxacillin, could result in drastic downstream effects.

Unique peptides were identified from each treatment group, with changes in the amino acid abundance at the C terminal anchor residue observed. Flucloxacillin was identified to bind to a range of proteins within multiple cellular systems (Chapter 3). Crucially, MRP2/P-gp transporter function was increased when HepaRG cells were treated with flucloxacillin for a prolonged period. Therefore, it is possible that flucloxacillin could alter the overall proteome present in the cell cytoplasm. While this could not be detected with immunoblot assays, the protein source of unique peptides to each treatment could be compared. Strikingly, based on protein molecular function, biological process, cellular component, protein class and pathway involvement, there were not observable differences in the proteome. This further highlights the observed change in the C terminal anchor residue abundance antigen is likely processing dependant. Two hypotheses for this are apparent, either protein haptenation is resulting in missed cleavages (Jenkins *et al.*, 2009) in the proteasome, or flucloxacillin (penicilloic acid) is assisting with the binding of peptides which not otherwise have high affinity to HLA-B*57:01.

Previous studies, into the HLA-B*57:01 repertoire in response to abacavir were successfully reproduced. Furthermore, similar repertoire changes with flucloxacillin in relation to HLA-B*57:01 were identified. While crystallographic data is not available to prove interactions between flucloxacillin and the binding cleft, it is thought that flucloxacillin penicilloic acid can interact with C-F pockets, resulting in the anchoring of otherwise low affinity peptides. However, protein haptenation is known to readily occur though the nucleophilic attack of the β lactam ring (Jenkins *et al.*, 2009). Therefore, further analysis into covalent binding of flucloxacillin to proteins, resulting in the potential presentation on HLA-B*57:01, is necessary.

CHAPTER 5 - DETECTION OF FLUCLOXACILLIN MODIFIED MHC PEPTIDES

5.1 INTRODUCTION

Previous analysis found covalent, irreversible, binding of flucloxacillin to proteins. This was observed in both *in vitro* protein and cell incubations, as well as serum proteins derived from patients. In the previous chapter the changes in the HLA-B*57:01 immunopeptidome of C1R-B*57:01 cells were investigated when cultured in the presence of flucloxacillin. While repertoire shifts were identified, the bioinformatic workflow did not consider covalent modification of peptides by flucloxacillin. Therefore, covalent binding to intracellular proteins, peptide-HLA complexes, and the HLA molecule itself are all possible. A wide diversity of proteins from C1R-B*57:01 cells were previously shown to be modified, raising the question, what is their fate?

To date, a limited number of studies have been performed to define the molecular initiating events in hapten mediated DHRs. Pulsing of drugs and/or their metabolites with antigen presenting cells prior to exposure to, and subsequent activation, of T cells gives strong evidence for the hapten hypothesis. With this method, cells are incubated with drug for a short period of time before being washed. In this way, T cells are never directly exposed to free drug, yet are still activated. For example, incubation of piperacillin with patient derived APCs showed a T cell proliferative response after a 4 hour pulse (El-Ghaiesh *et al.*, 2012). While there is general understanding of these mechanisms at the protein level, experiments performed at the peptide level are sparse. This is a particularly challenging method, due in part to variations with genetic susceptibility. As HLA proteins are highly polymorphic, patients with different alleles will present different peptides. Therefore, defining a single peptide responsible for immune activation is nigh-on impossible. For the study of flucloxacillin the genetic predisposition to HLA-B*57:01 makes this a ‘simpler’ task, but piperacillin, for example, does not have any known genetic association; although the chemistry of covalent binding is very similar to that of other penicillins.

Early studies by the Weltzien group utilised nitrophenol compounds, namely trinitrophenol (TNP) and dinitrophenol (DNP), for the investigation of protein haptens in T cell activation. Tryptic digests of TNP modified BSA activated T cells, with glutaraldehyde fixation of antigen presenting cells revealing intracellular processing of peptides was not required. Chromatographic separation of tryptic peptides revealed TNP modification of BSA at position 227 was the antigenic determinant. Modification of a similar peptide from mouse serum albumin (MSA) resulted in cross reactivity, leading to the conclusion that the peptide sequence was only partially involved in T cell recognition. Furthermore, tryptic digests of both OVA and KLH modified by TNP resulted in some cross reactivity. It was concluded that the peptide sequence is important in the anchoring of peptide within the MHC binding groove while a hapten is required for subsequent T cell activation (Ortmann *et al.*, 1992). Another study by Martin *et al* used synthetic hapten-peptide conjugates to investigate their impact on T cell specificity in a haplotype-specific way. In this study, the presence of amino acids at specific positions complementing the MHC binding repertoire under investigation enhanced the antigenicity of TNP synthetic peptides. While this was focussed on the mouse H-2K^{b} repertoire, it was one of the first models to investigate hapten-T cell models in a MHC restricted fashion (Martin *et al.*, 1992). Indeed, further studies were required for the translation of the results into a model more relevant to human disease.

From the knowledge generated in their previous studies, Padovan *et al* used designer synthetic peptides for the characterization of T cell responses to drug modified peptides. As penicillins are a major contributor to hapten mediated drug allergy, BP was selected as the model compound in this study. BP specific T cell clones were successfully generated from BP haptented peptides fitting the MHC class II allele, HLA-DRB1*04:01. These three peptides (EAYAAAASKAAA, EAYAAAKSAAAA and EAYAKAASAAAA) all contained a lysine at different positions to alter the site of BP modification. In the T cell clones generated, optimal activation was achieved through the positioning of BP across the peptide backbone. For the first time,

this study not only provided evidence into the role of drug-modified peptides in T cell activation, but that the structural positioning of the modification is important in specific T cell recognition. Strikingly, this study also provided evidence for T cell activation by BP modified HLA-DRB1*04:01 self-peptides, showing the hapten is a major determinant in T cell activation through otherwise tolerated peptide sequences (Padovan *et al.*, 1997). At the time of this study, links between HLA associations and drug hypersensitivity predisposition were largely unknown, making these findings even more relevant in the present day.

A running theme throughout these studies is the carrier independence of peptides resulting in the stimulation of T cells. Padovan *et al* designed peptides with physical properties to enable binding to the HLA binding groove under investigation. HLA-DRB*01:04 has specificity for tyrosine in position 1 and 6, with position 1 being defined by the first anchor binding site. Along with glutamic acid for peptide solubility and lysine for hapten binding, the remaining amino acids were occupied by the relatively inert alanine. It seems that by having these basic properties T cell activation can occur in a very specific manner, even though *in vivo* these peptides would not exist (Padovan *et al.*, 1997). While this is an exciting step forward in the field of drug hypersensitivity, there is no indication as to the precise proteins and peptides naturally presented by MHC resulting in the clinical manifestation of disease. Since this study, further investigations into benzyl-penicillin (BP) modified peptides have been performed with more physiologically relevant peptide sources. Azoury *et al* recently described the priming of naïve T cells from healthy donors to BP peptides derived from HSA. In total 3 peptides, with BP bound at lysine 159, 212 and 525, were found to be recognized by naïve T cells. Interestingly, BP bound at lysine 159 and 525 were found to induce PBMC proliferation in patients with penicillin allergy (Azoury *et al.*, 2018). Here, the results showed that BP haptenation of HSA proteins could lead to the presentation of BP-HSA peptides, with the potential to prime naïve T cells.

HSA is a major target for modification by β lactam penicillins as described in a number of studies (Jenkins *et al.*, 2009; Meng *et al.*, 2011, 2016, 2017; Whitaker *et al.*, 2011). In previous chapters of this thesis, HSA was found to be the only protein modified by flucloxacillin that could be detected in patient serum samples. That said, the diversity of intracellular proteins modified by flucloxacillin was found to be extensive in multiple relevant cell lines. The fate of these proteins is not currently known; however, it is likely that they are digested by the proteasome and subsequently presented on the cell surface by MHC molecules. In the previous chapter repertoire changes were observed in unmodified peptides, presented when C1R-B*57:01 cells were cultured with flucloxacillin. Nevertheless, it is not yet clear whether drug-protein haptenation results in dysregulation of proteasomal processing. Therefore, in this chapter the methods developed for the investigation of the immunopeptidome of C1R-B*57:01 cells were utilised further in the identification of naturally presented flucloxacillin modified peptides.

5.2 Aims

Previously flucloxacillin DILI has been defined by the hapten mechanism of drug hypersensitivity. Here the methods developed in the previous chapter were used to characterize the role of flucloxacillin modified proteins in antigen presentation. Therefore, flucloxacillin modification of MHC heavy chain and MHC binding peptides presented by HLA-B*57:01 was investigated. In order to perform this study, the following aims were defined;

1. Identify and characterize flucloxacillin binding to MHC heavy chain proteins
2. Identify flucloxacillin HLA-B*57:01 binding peptides using characteristic drug fragmentation ions
3. Annotate flucloxacillin modified peptide sequences using *de novo* techniques
4. Identify protein sources of flucloxacillin modified peptides and their contribution to the identified immunopeptidome
5. Synthetically generate flucloxacillin modified HLA-B*57:01 binding peptides for use in T cell culture assays

5.3 METHODS

5.3.1 PREVIOUSLY DESCRIBED METHODS

- Dot Blot - 2.3.7, p92
- Immunoaffinity MHC purification - 4.3.4, p163
    - Mass spectrometry - 4.3.5164
    - Post-acquisition analysis - 4.3.6, p165
    - Theoretical docking analysis - 4.3.7, p165

5.3.2 SYNTHETIC PEPTIDE PRODUCTION

5.3.2.1 DRUG CONJUGATION

Fmoc-HSATQKEHGW (95% purity, Syn Peptide Co, China) was incubated with flucloxacillin at a 1:10 molar ratio (0.5 to 5 mM respectively) in 70% ACN/30% H_2O for 48 hours at 37°C. Reactions were performed in glass to reduce the loss of material. After 48 hours the reaction was stopped by freezing at -20°C.

5.3.2.2 HPLC FRACTIONATION

Fmoc-HSATQKEHGW flucloxacillin conjugate was thawed and diluted 1 in 50 into 70% ACN/30% H_2O prior to injection onto a monolithic C18 column (100 x 4.6 mm Onyx, Phenomonex) at 1 mL per minute using an Agilent 1260 HPLC system. Method optimisation for loading is discussed in section 5.4.7 of this chapter. For the separation and fractionation of the reaction mixture, unmodified peptide and flucloxacillin degradation products, a 30 minute gradient was applied with a wavelength detection of 215 nm (Table 5.1). Fractions were calculated based on the retention time of peaks of interest, as described in section 5.4.5. All fractions were collected into protein LoBind (Eppendorf) tubes and dried using vacuum centrifugation at 30°C until dry. Fractions were stored at -20°C prior to mass spectrometric analysis

Table 5.1. Synthetic flucloxacillin modified Fmoc-HSATQKEHGW HPLC method parameters. Modified peptide was separated from flucloxacillin degradation products and unmodified peptide using mobile phases containing 0.1% TFA (A) and 100% ACN/0.1% TFA (B).

Time	%A	%B	Flow (ml/min)	Wavelength
0:00	98	02	1	215
0:15	98	02	1	215
20:00	25	75	1	215
20:01	00	100	1	215
21:00	00	100	1	215
22:30	98	02	1	215
30:00	98	02	1	215

5.3.2.3 DEPROTECTION AND PURIFICATION

Piperidine was incubated with Fmoc-HSATQK*EHGW at a 10:1 ratio (90 µL 1 mM peptide, 10 µL 100 mM piperidine in 70% ACN/30% H_2O) for 2 hours at 37°C. For each run, 20 µL of the reaction mixture was loaded onto a Phenomenex C_{18} Kinetex 5µm column at 1 mL per minute using an Agilent 1200 HPLC system. A 30 minute gradient was applied with absorbance measured at a wavelength of 215 nm (Table 5.2). Fractions were collected and processed as described in section 5.3.2.2 prior to mass spectrometric analysis.

Table 5.2. Synthetic flucloxacillin modified deprotected HSATQKEHGW HPLC method parameters. Deprotected modified peptide was purified by fractionation using mobile phases containing 0.1% TFA (A) and 100% ACN/0.1% TFA (B).

Time	%A	%B	Flow (ml/min)	Wavelength
0:00	98	02	1	215
0:06	98	02	1	215
20:00	25	75	1	215
20:06	00	98	1	215
25:00	00	98	1	215
22:06	98	02	1	215
30:00	98	02	1	215

5.3.2.4 MASS SPECTROMETRIC ANALYSIS

Deprotected flucloxacillin modified HSATQK*EHGW was resuspended in 20% ACN/10% TFA and purified using strong cation exchange (SCX) ZipTips (Merck Millipore). ZipTips were equilibrated with 3 x 10 µL 0.1 TFA before samples were loaded onto the ZipTip by pipetting 15 - 20 x. After washing with 5 x 10 µL 0.1% TFA, peptides were eluted in 10 µL 5% ammonium

hydroxide/30% methanol/0.1% TFA. Samples were dried using nitrogen and resuspended in 30µl 0.1% FA/2% ACN prior to analysis using a Triple TOF 5600 mass spectrometer (Sciex) delivered into the instrument using a Eksigent NanoLC Ultra HPLC system. Samples (5 µL) were injected onto a nanoACQUITY UPLC Symmetry C18 Trap Column (P/N Waters, MA, USA) and washed for 10 min at 2 µL/min with 0.1% FA. A gradient from 1.6% ACN/0.1% FA to 95% ACN/0.1% FA was applied over 95 minutes (Table 5.3) at a flow rate of 300 nL/min through a Peptide BEH C18 nanoACQUITY Column (Waters, MA, USA). MS was operated in positive ion mode with survey scans of 200 ms, with an MS/MS accumulation time of 150 ms for the 20 most intense ions (total cycle time 3.2 s). A threshold for triggering MS/MS of 40 counts per second was used with an exclusion of former target ions for 30 seconds. Rolling collision energy was applied. Data dependent acquisition of ions in the mass range of 200-1,800 amu (MS) and 60-1,800 amu (MS/MS) was performed using Analyst TF 1.6.

Table 5.3. Synthetic flucloxacillin modified deprotected HSATQKEHGW MS HPLC method parameters. Deprotected modified peptide was analyzed using mass spectrometry through in line HPLC separation using mobile phases containing 0.1% TFA (A) and 100% ACN/0.1% TFA (B).

Time	%A	%B	Flow (nl/min)
00.00	98.4	1.6	300
75.00	72.0	28.0	300
79.00	36.0	64.0	300
80.00	5.0	95.0	300
85.00	5.0	95.0	300
86.00	98.4	1.6	300
95.00	98.4	1.6	300

5.3.2.5 STATISTICAL VALIDATION

Hypothetical calculations of *b* and *y* ions for flucloxacillin modified HSATQK*EHGW were generated using UCSF ProteinProspector MS-Product (v 5.22.1). Ions were calculated for the unmodified peptide, in addition to a mass addition of 294 Da and 453 Da on the lysine. Peak picking was performed on the eluted and synthetic fragment ions using Protein Pilot (v 5.000). Observed fragmentation ions were compared to hypothetical fragmentation ions, and those within a mass tolerance of 0.1 Da were accepted. A Pearson two-tailed correlation test was used to compare the two sets of accepted ions using SPSS Statistics (v 24).

5.4 RESULTS

5.4.1 COVALENT FLUCLOXACILLIN MODIFICATION OF MHC HEAVY CHAIN

5.4.1.1 IMMUNO-ABSORBANCE DETECTION

In the previous chapter, a novel set of HLA-B*57:01 restricted peptides presented in the presence of flucloxacillin were identified. Modelling experiments revealed flucloxacillin may interact with the peptide binding groove in a non-covalent manner to stabilise peptides which otherwise have low binding affinity to HLA-B*57:01. Previous experiments have focussed on flucloxacillin adduct formation with respect to irreversible binding. Therefore, covalent binding of flucloxacillin to HLA-B*57:01 was investigated. After HPLC fractionation of eluted peptide-MHC complexes from C1R-B*57:01 cells cultured in the presence of flucloxacillin, individual fractions were dotted onto nitrocellulose membrane in a grid format. Subsequent detection using anti-flucloxacillin antibody identified flucloxacillin present in the last three fractions (Figure 5.1A), indicative of the MHC heavy chain (Figure 5.1B).

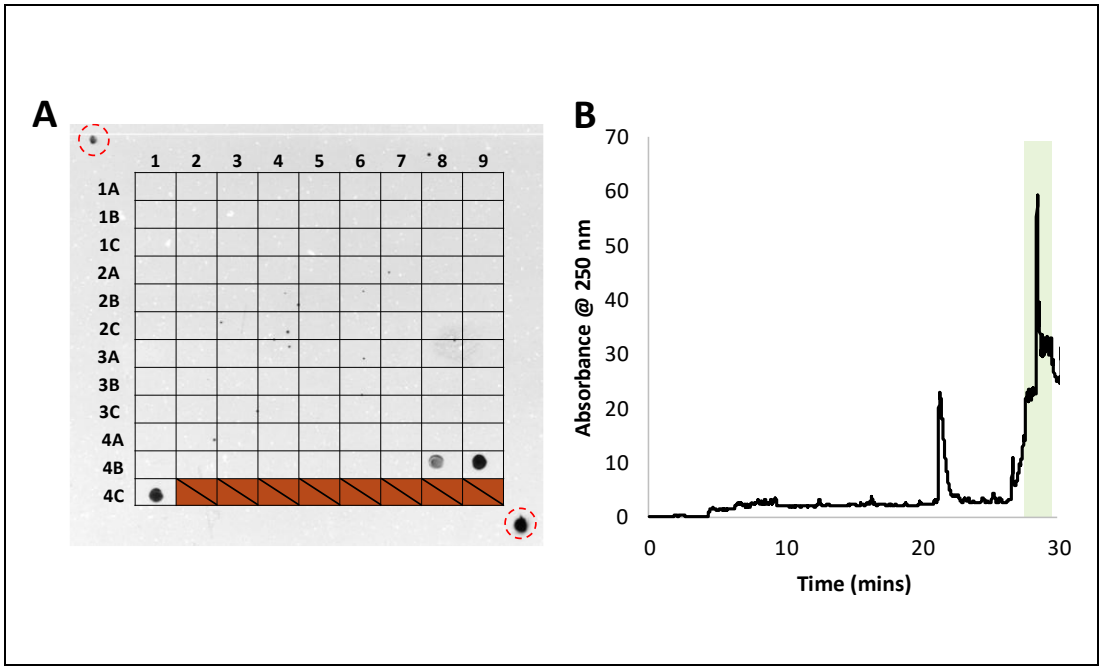


Figure 5.1. Immnodetection of flucloxacillin bound to MHC heavy chain. (A) Peptide-MHC complexes from flucloxacillin treated C1R-B*57:01 cells were fractionated using HPLC and dotted onto nitrocellulose membrane. Immunodetection using anti-flucloxacillin antobody revels modification in the last three fractions (B) indicative of MHC heavy chain protein. Positive controls (OVA) used for alignment of grid on nitrocellulose membrane with the developed film (red circles).

5.4.1.2 MASS SPECTROMETRIC CHARACTERIZATION

Mass spectrometry was used to characterize the site of MHC heavy chain adduct formation. MHC heavy chain was therefore pooled, reduced and alkylated, and digested using trypsin. Subsequent mass spectrometry revealed the presence of flucloxacillin bound to peptides. Commercially available software is unable to detect flucloxacillin modified peptides due to the partial loss of the flucloxacillin structure during MS/MS fragmentation. Characteristic flucloxacillin fragmentation ions of *m/z* 160, 295 and *m/z* 454 (full flucloxacillin mass) (Figure 5.2) indicate modifications are present on peptides. Loss of the thiazolidine ring when flucloxacillin is bound to peptides results in a mass residue of 294 Da on the target lysine.

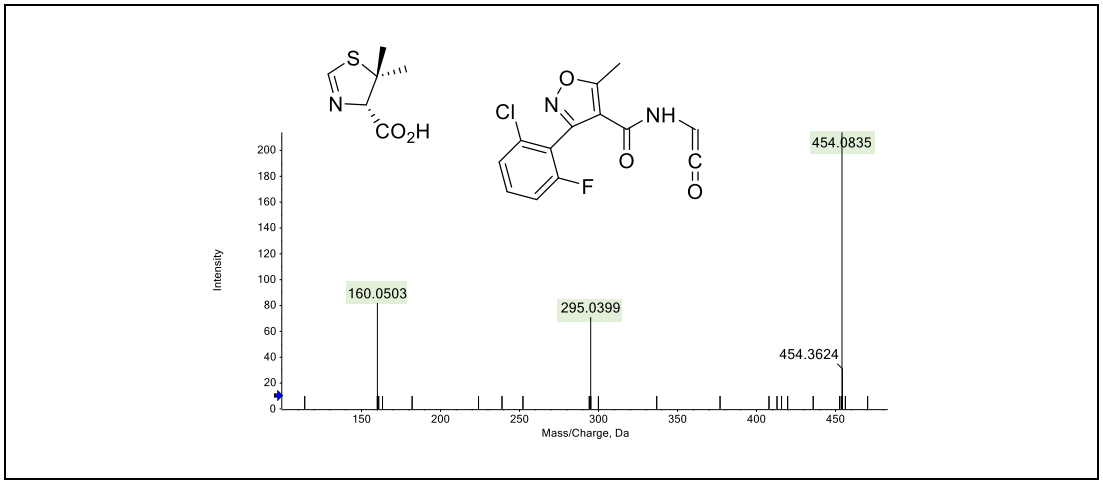


Figure 5.2. Flucloxacillin partial loss ions after MS/MS fragmentation. Flucloxacillin (*m/z* 454.0835) undergoes partial loss during MS/MS fragmentation resulting in loss of the thiazolidine ring (*m/z* 160.0503). The resultant part of flucloxacillin can also be identified (*m/z* 295.0399).

Analysis of the HLA alpha chain fractions from eluted complexes revealed flucloxacillin bound to the protein at the N-terminal glycine ([Flucloxacillin]-GSHSMR) (Figure 5.3). NH_2-GSHSMR is part of the extracellular region of the HLA-B*57 alpha chain, making it one of the most susceptible regions of the protein to undergo direct modification from the external environment. As signal peptides are typically involved in the translocation of newly synthesised proteins to the cell surface, it is unlikely that intracellular modification occurred prior to transport. With this is mind, it is perhaps less likely that this modification is involved in the loading of novel peptides into the binding cleft.

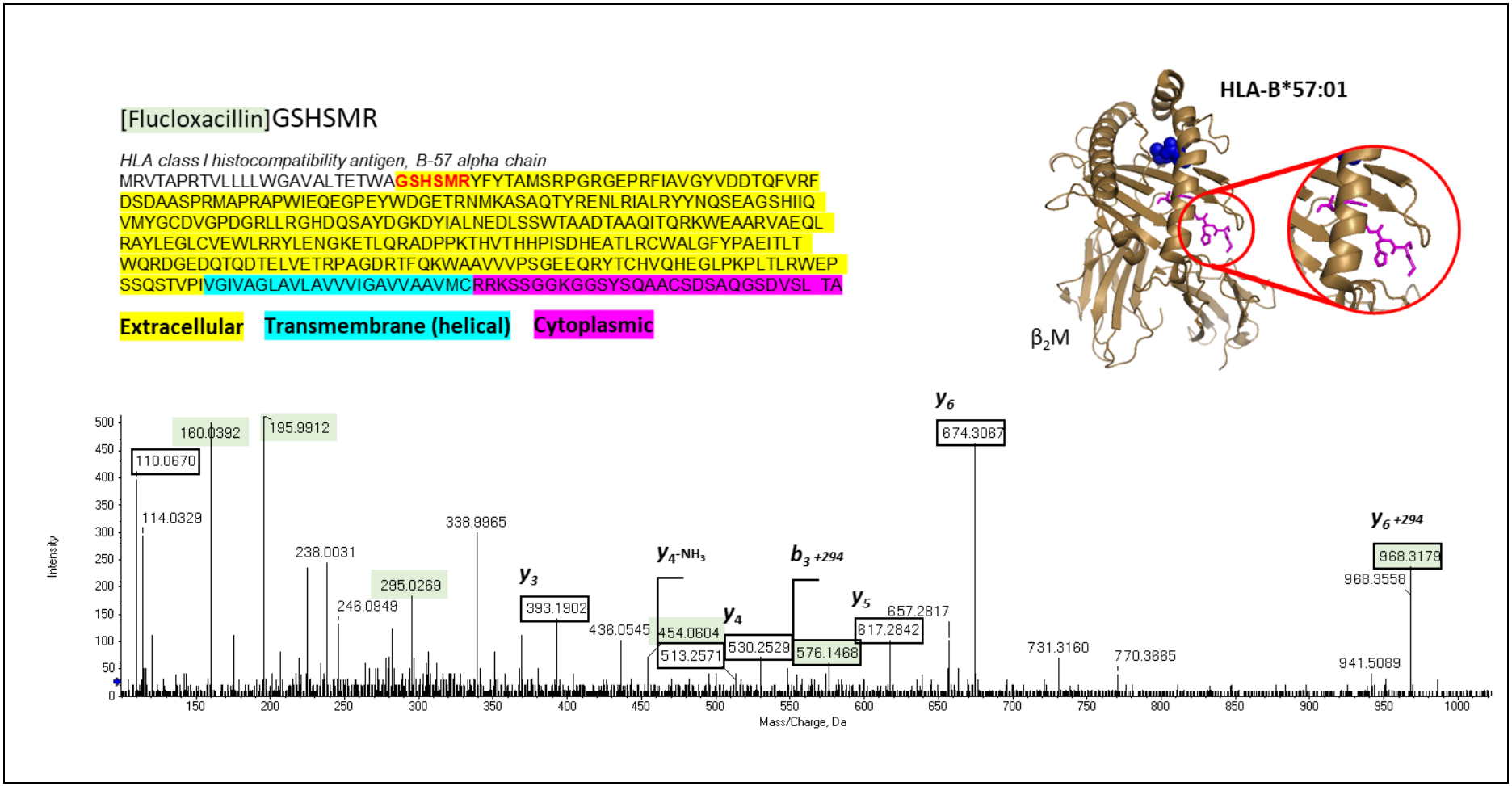


Figure 5.3. Mass spectrometric analysis of flucloxacilin modified MHC alpha chain. Mass spectrometry reveals covalently bound flucloxacillin to the N-terminal amino acid on the extracellular region of HLA-B*57:01, once cleavage of the signal peptide has been performed. The characteristic fragmentation ions 160.0392, 195.9912 and 295.0269 and 454.0604 are all indicative of the presence of flucloxacillin.

Importantly, the GSHSMR peptide sequence is conserved across many HLA proteins including HLA-C*04:01 (P30504). HLA-C*04:01 is expressed on C1R-B*57:01 cells, therefore, it is unlikely that the observed modification contributes to only the HLA-B*57:01 genetic susceptibility to flucloxacillin liver injury. While N-terminal modification of GSHSMR was the only one to be found in this study, it is possible that other sites of adduct formation could contribute to the loading of neo-epitopes in an HLA restricted manner.

5.4.2 CHARACTERIZATION OF FLUCLOXACILLIN MODIFIED MHC BINDING PEPTIDES

Peptides eluted from C1R-B*57:01 cells cultured in the presence of flucloxacillin were assessed for flucloxacillin modification. Manual interpretation of the spectra containing characteristic flucloxacillin fragmentation ions was performed. Of all the peptides eluted from flucloxacillin-treated C1R-B*57:01 cells, over 30 peptides were identified with an indicative modification, with 7 having been fully annotated (Table 5.4).

Table 5.4. Flucloxacillin-haptenated peptides eluted from C1R cells expressing HLA-B* 57:01

Figure	m/z	Charge state	Peptide mass (Da)	[1]Sequence	Protein source	[2]Predicted affinity (IC50, nM)
Figure 5.4A	533.8612	3	1145.5836	HTAHIAC(O_3)K*FA	human elongation factor 1-alpha 1 (EF1A1_HUMAN, P68104)	1210.6 (WB)
Figure 5.4B	630.2267	3	1434.624	VSDHEATLR*C(O_3)WA	HLA class I histocompatibility antigen, (1C04_HUMAN, P30504) (1C15_HUMAN, Q07000), (1B82_HUMAN, Q29718	2658.7 (NB)
Figure 5.5A	634.8969	3	1448.64	ISDHEATLR*C(O_3)WA		1017.23 (WB)
Figure 5.5B	634.8969	3	1448.64	ISDHEATLR*C(O_3)WA		1017.23 (WB)
Figure 5.6A	545.1719	3	1179.5157	HSATQK*EHGW	Not known	23.4 (SB)
Figure 5.6B	567.326	3	1245.5406	LFDPTNC(SO_2SH)K*MN	Not known	30810.6 (NB)
Figure 5.7A	561.8524	3	1229.503	LFDPTNC(O_3)K*MN	Not known	30810.6 (NB)

[1] Peptide sequences were determined manually by de novo sequencing, *indicates the flucloxacillin binding sites; [2] Predicted affinity of unmodified peptides to HLA-B*57:01 was determined using NetMHC4.0. SB-strong binder, WB-weak binder, NB-non binder.

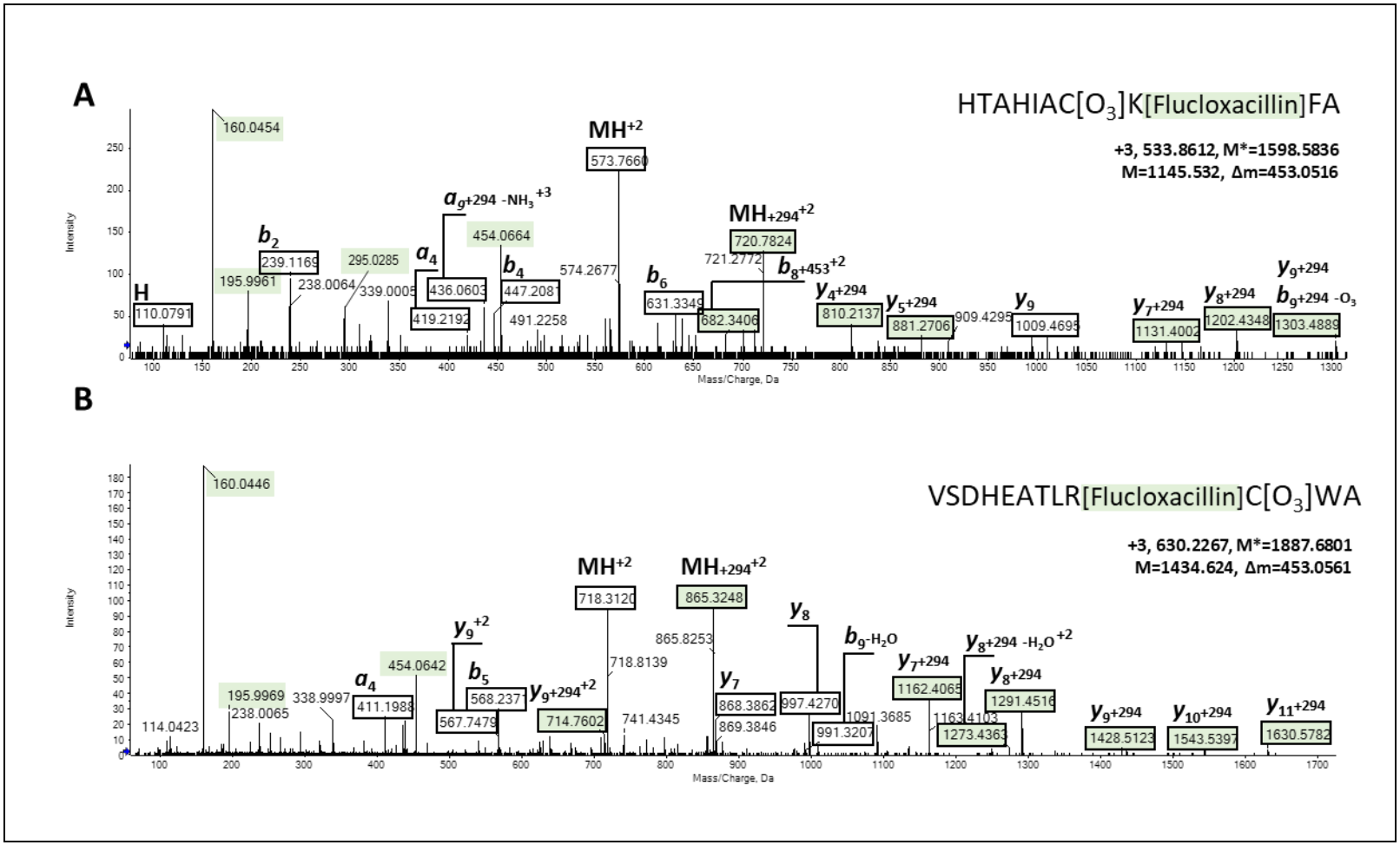


Figure 5.4. Naturally presented flucloxacillin-modified MHC class I peptides. MS/MS spectra showing flucloxacillin modification of peptides eluted from C1R-B*57:01 cells cultured in the presence of drug for 48h. Characteristic fragmentation ions indicate flucloxacillin modification (*m/z* 160, 195, 259 & 454). Manual interpretation reveals amino acid sequence; (A) HTAHIACK[Flucloxacillin]FA & (B) VSDHEATLR[Flucloxacillin]C[O_3]WA.

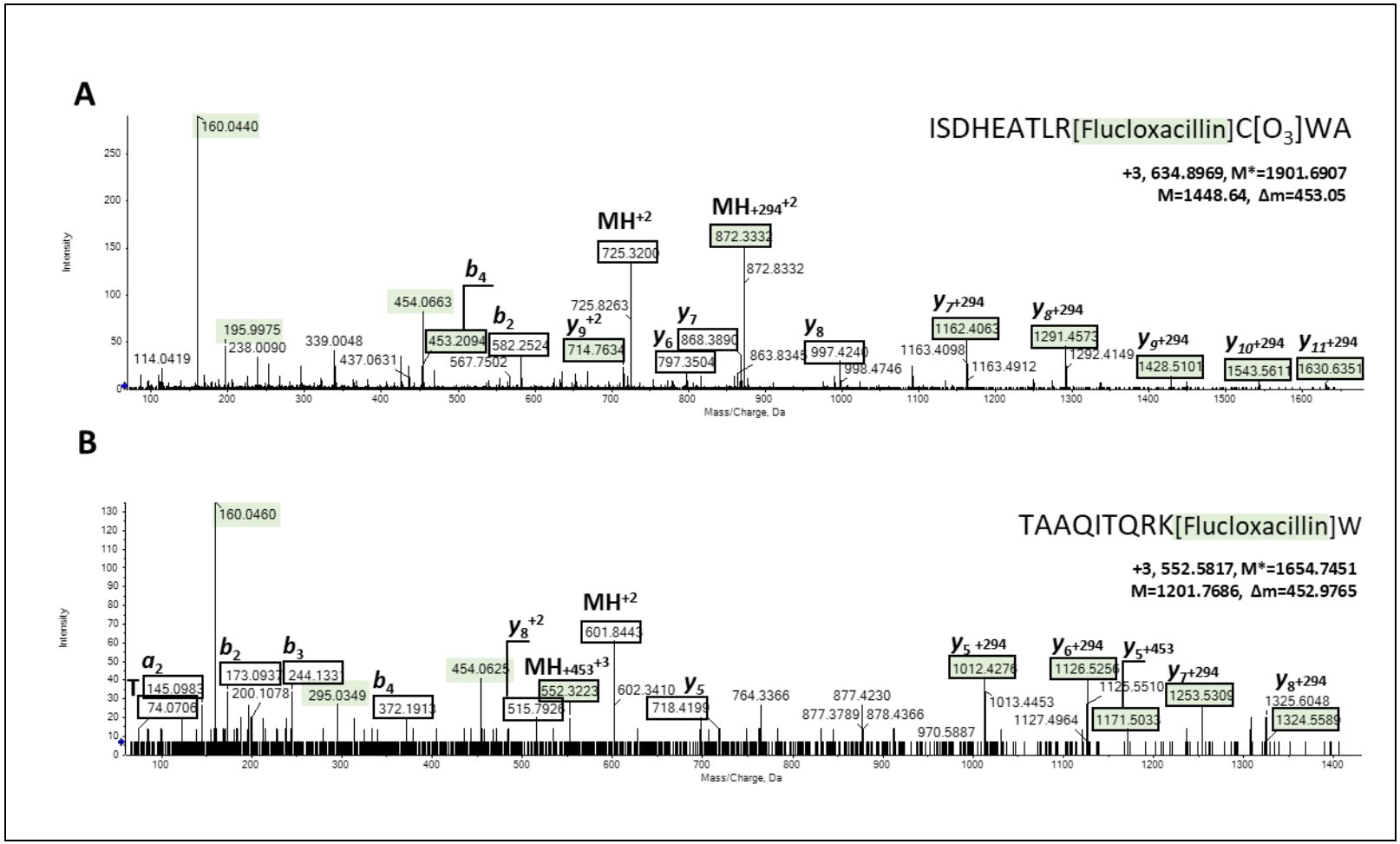


Figure 5.5. Naturally presented flucloxacillin-modified MHC class I peptides. MS/MS spectra showing flucloxacillin modification of peptides eluted from C1R-B*57:01 cells cultured in the presence of drug for 48h. Characteristic fragmentation ions indicate flucloxacillin modification (*m/z* 160, 195, 259 & 454). Manual interpretation reveals amino acid sequence; (A) ISDHEATLR[Flucloxacillin]C[O_3]WA & TAAQITQRK[Flucloxacillin]W.

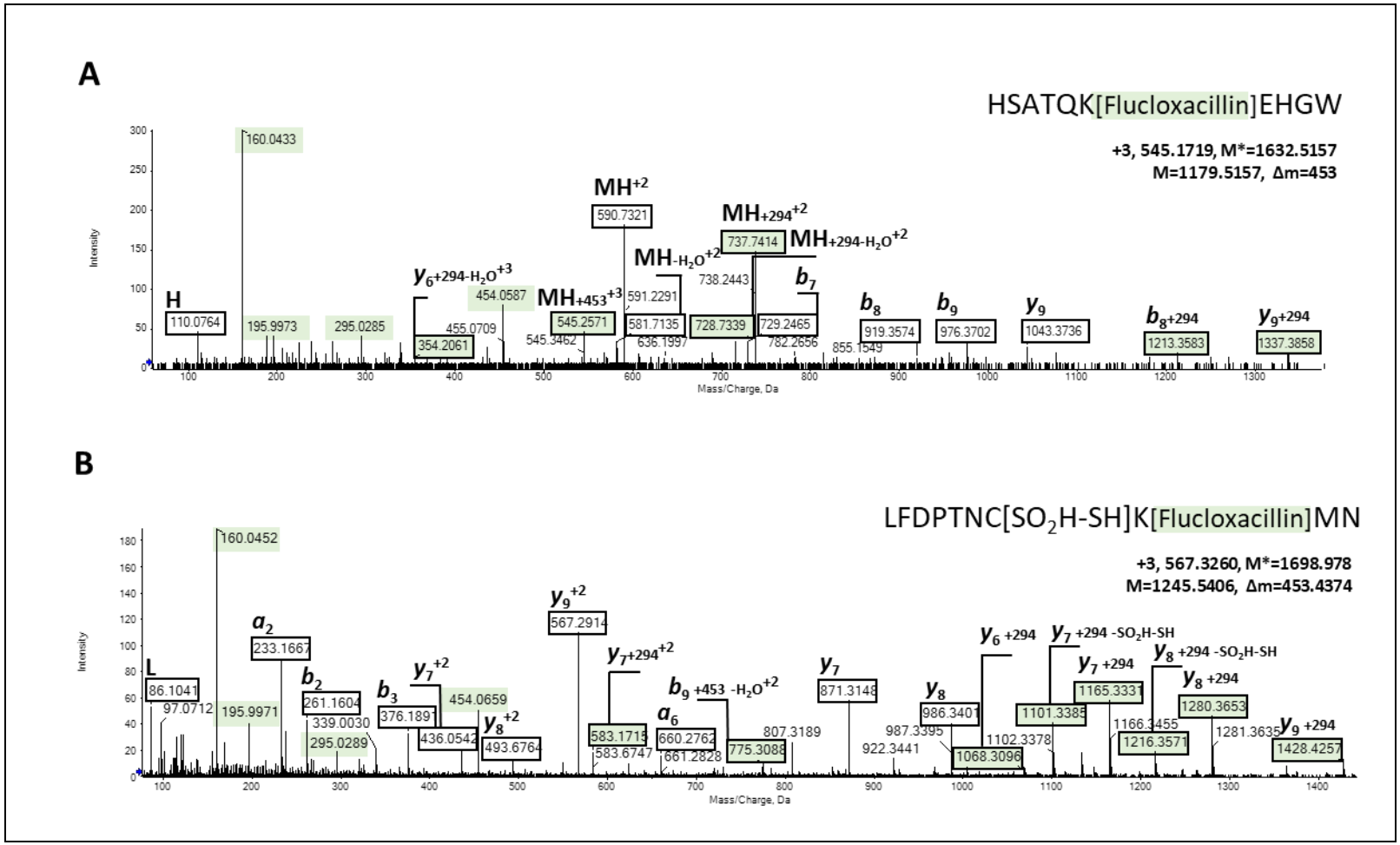


Figure 5.6. Naturally presented flucloxacillin-modified MHC class I peptides. MS/MS spectra showing flucloxacillin modification of peptides eluted from C1R-B*57:01 cells cultured in the presence of drug for 48h. Characteristic fragmentation ions indicate flucloxacillin modification (*m/z* 160, 195, 259 & 454). Manual interpretation reveals amino acid sequence; (A) HSATQK[Flucloxacillin]EHGWS & (B) LFDPTNC[SO_2H-SH]K[Flucloxacillin]MN.

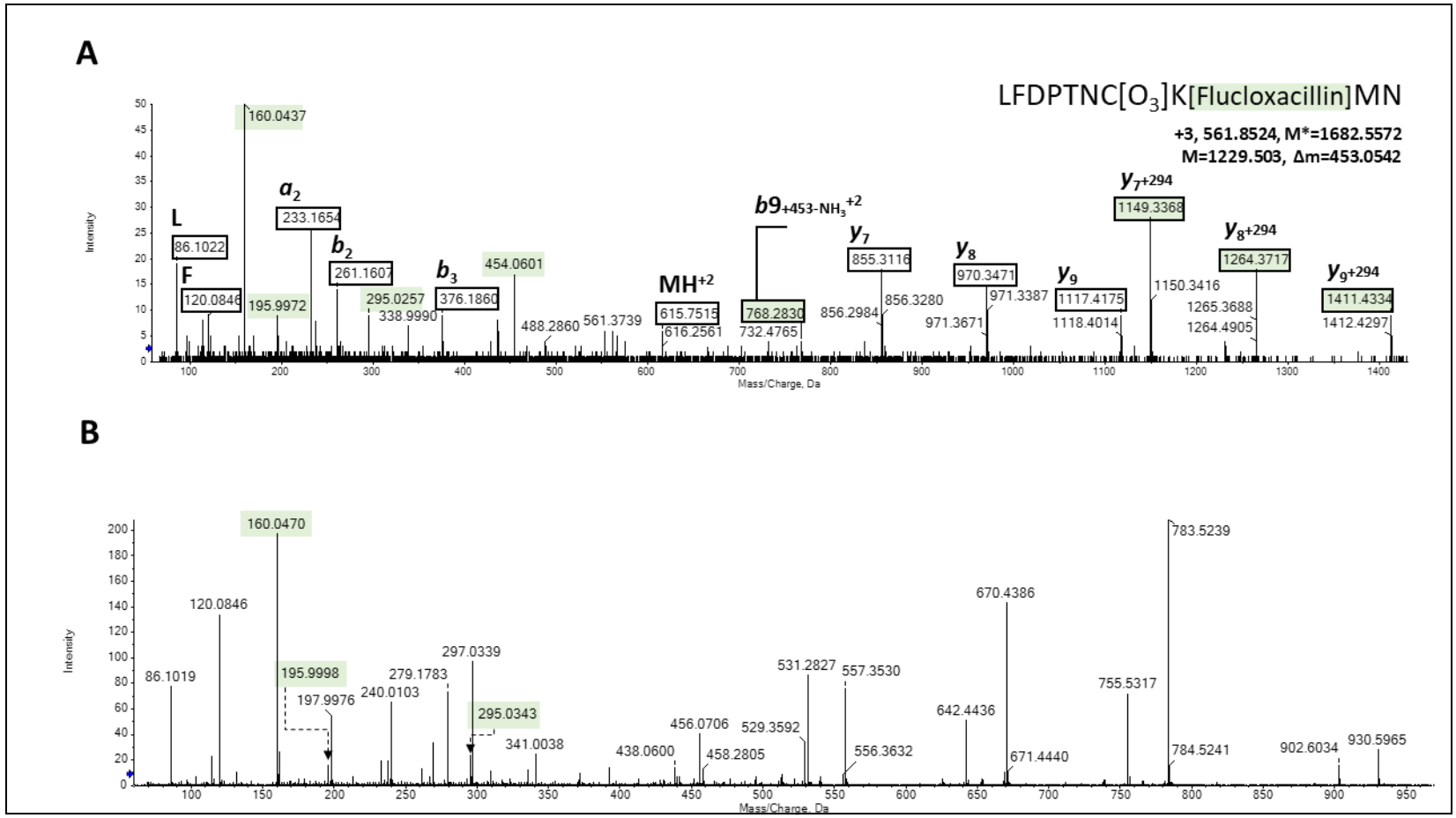


Figure 5.7. Naturally presented flucloxacillin-modified MHC class I peptides. MS/MS spectra showing flucloxacillin modification of peptides eluted from C1R-B*57:01 cells cultured in the presence of drug for (A) 48h & (B) 10 minutes. Characteristic fragmentation ions indicate flucloxacillin modifiication (*m/z* 160, 195, 259 & 454). Manual interpretation reveals amino acid sequence; (A) LFDPTN[CO_3]K[Flucloxacillin]MN. (B) Antigen processing requires >10 minutes, therefore direct modification of presented MHC I peptides is expected. Sequence annotation has not yet been possible.

Although only 7 MHC peptide sequences haptenated by flucloxacillin could be manually interpreted, it was possible to identify a range of protein sources. Peptide HTAHIACK*FA (Figure 5.4A) was derived from human elongation factor 1-alpha 1 (EF1A1_HUMAN, P68104), an abundant and multifunctional protein. Of these 7 peptides, 3 appear to come from the MHC alpha chain (VSDHEATLR*CWA (Figure 5.4B), ISDHEATLR*CWA (Figure 5.5A), and TAAQITQRK*W (Figure 5.5B). These were likely modified, processed by the proteasome and presented by a second HLA molecule. Interestingly, the protein sources for the other two peptides, HSATQK*EHGW (Figure 5.6A) and LFDPTNCK*MN (Figure 5.6B & Figure 5.7A) could not be determined using NCBI Blast searches, indicating they were potentially spliced peptides or originated from proteins not present in the human proteome database. Although these did not have a known protein source, spliced peptides are believed to account for a large proportion of the immunopeptidome (Faridi *et al.*, 2018). Using NetMHC 4.0 binding prediction tools LFDPTNCKMN is considered a non-binder to HLA-B*57:01 (30,810.6 nM), however it may be that flucloxacillin assists with the binding to HLA-B*57:01. While it is possible that the peptide was derived from HLA-C*04:01, the binding affinity prediction score is even lower (31,074.7 nM).

The determination of whether MHC peptides are directly modified, or if haptenated proteins are subsequently digested in the proteasome, was challenging. Therefore, flucloxacillin was incubated with C1R-B*57:01 cells for only 10 minutes. This would not give enough time for flucloxacillin to pass through proteasomal digestion and subsequent MHC presentation pathways. Strikingly, flucloxacillin-haptenated peptides were still detected (Figure 5.7B). While it was not possible to positively interpret the peptide sequence, the presence of characteristic fragment ions (*m/z* 160, 195 and 295) suggests flucloxacillin can bind to peptides already accommodated in the peptide binding groove.

5.4.3 ANTIGEN PROCESSING DISRUPTION

In vitro studies have shown that the binding of flucloxacillin to lysine results in missed cleavages when proteins are digested using trypsin (Jenkins *et al.*, 2009). It is therefore likely that modification may result in missed cleavages within the proteasome. To determine whether covalent binding of flucloxacillin interferes with natural protein processing, hotspots of peptide presentation across proteins were determined. Eluted peptides from human elongation factor 1-alpha 1 (EF1A1_HUMAN, P68104), where flucloxacillin was found bound to Lys371 (HTAHIACK*FA), were separately aligned to the source protein sequences (Figure 5.8A). Indeed, the presence of flucloxacillin generated a unique set of peptides. Specifically, covalent binding of flucloxacillin to Lys371 in EF1A1 resulted in a hotspot for ligand presentation (Figure 5.8A). Distinct hotspots associated with flucloxacillin treatment were also detected with other proteins (Figure 5.8C).

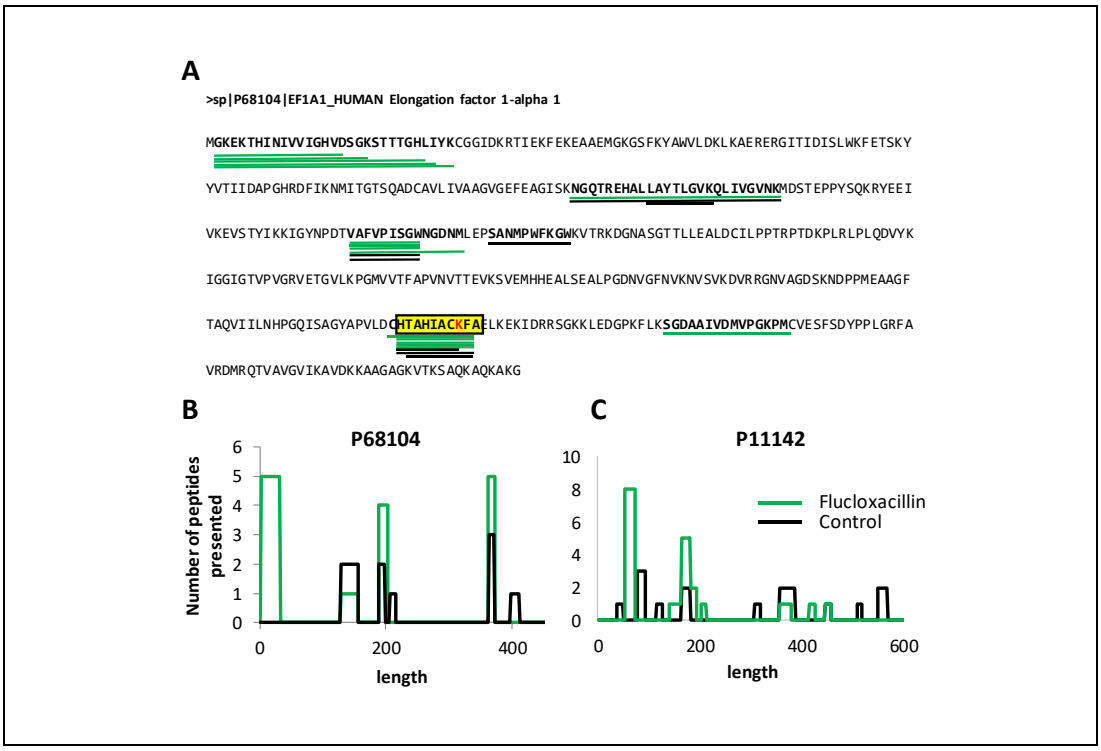


Figure 5.8. Flucloxacillin affects overall proteasomal cleavage sites. (A) The presence of flucloxacillin may affect the proteasome cleavage sites, leading to the generation of unique set of peptide pool. Peptides generated from the same protein (P68104, human elongation factor 1-alpha 1) presented by flucloxacillin-treated (green) and untreated cells (black) display a distinct digestion pattern. (B) Covalent binding of flucloxacillin to Lys371 (A, yellow box) results in hotspots of presentation. (C) Similar hotspots were observed in the human heat shock cognate 71K Da protein (P11142).

As evidence for disruption of antigen processing was observed, enzyme activity within the proteasome was investigated. In the previous chapter, changes to the MHC peptide PΩ amino acid abundance was observed in the presence of flucloxacillin. This suggests antigen processing may too be altered in the presence of flucloxacillin. Therefore, the amino acids flanking the termini of unique peptides from their source protein were identified for each treatment (Figure 5.9). The results of the comparison showed that at either end of the unique peptides, the relative abundance of the subsequent amino acid remained the same. This indicated that, although the two data sets were previously distinct based on PΩ amino acid abundance and binding affinity, the activity of specific enzymes was not altered by the presence of flucloxacillin.

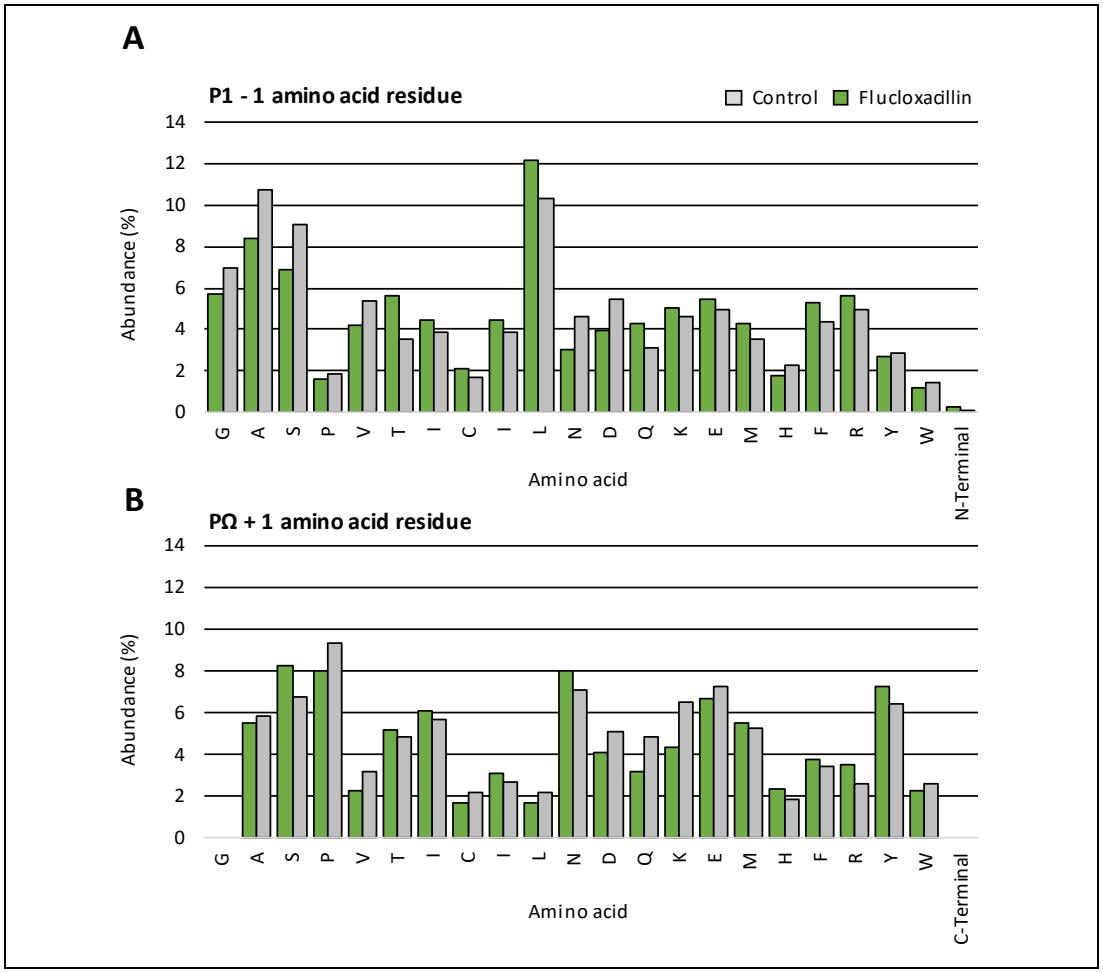


Figure 5.9. Amino acid abundance flanking N and C terminal ends of unique peptides to each treatment group. The abundance of amino acids present at (A) P1 minus 1 and (B) PΩ plus 1 was assessed. In all cases, the abundance of amino acids flanking each peptide termini did not change between treatment groups.

5.4.4 THEORETICAL DOCKING OF FLUCLOXACILLIN MODIFIED PEPTIDES TO HLA-B*57:01

To further assess the binding of haptenated peptides to HLA-B*57:01, theoretical modelling and docking analysis was performed (HLA-B*57:01 crystal structure PDB 3VRJ) with the assistance of Dr Xiaoli Meng. One peptide of particular interest was HSATQK*EHGW (Figure 5.6A) due to its unknown protein source and high binding affinity. In addition, no PTMs were observed; cysteine containing peptides are often oxidised, however it is challenging to determine whether oxidation took place *in vivo* or *ex vivo* during sample processing. Therefore, both the unmodified and modified HSATQK*EHGW peptide were docked into the HLA-B*57:01 binding cleft using Gold Suite v.5.2.1 (Figure 5.10).

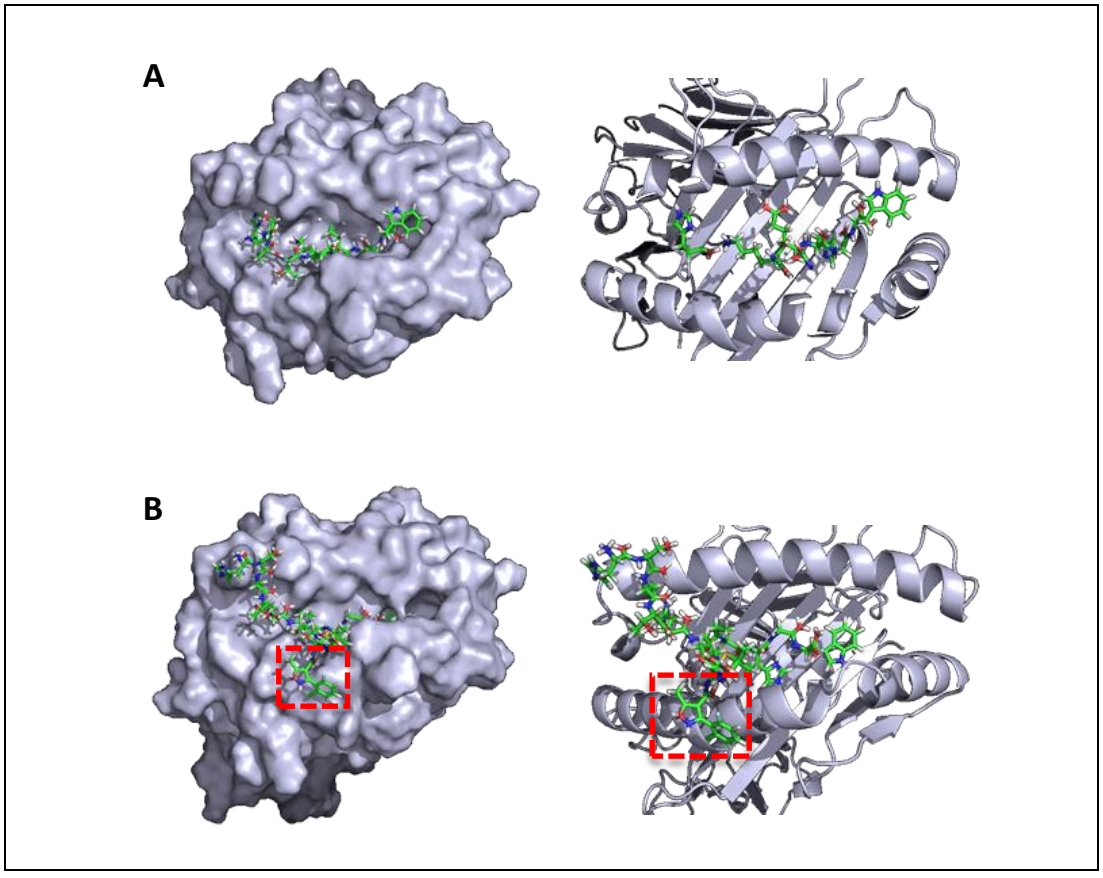


Figure 5.10. Theoretical docking of HSATQKEHGW into the HLA-B*57:01 binding cleft. Gold suite v.5.2.1 was used to dock (A) unmodified and (B) flucloxacillin modified HSATQKEHGW into the binding cleft of HLA-B*57:01. Both peptides bind with high affinity (Gold fitness scores of (A) 95.78 and (B) 79.31), with the flucloxacillin modification protruding outwards in its best orientation hypothesis.

Unmodified HSATQKEHGW (Figure 5.10A) was found to bind into the HLA-B*57:01 binding cleft with the C-terminal tryptophan residing within the F-pocket as anticipated. In a similar

style, HSATQK*EHGW (modified by flucloxacillin) also accommodated the binding groove in a canonical fashion (Figure 5.10B). Interestingly, the best hypothesis for binding of modified HSATQK*EHGW results in flucloxacillin protruding from the binding groove. While modelling is purely based on theoretical simulations, in this orientation flucloxacillin would most likely interact with the TCR. In turn, the presence of a neo-antigen is believed to be the most likely mechanism resulting in T cell activation, leading to an immune response. Of course, the anticipation that flucloxacillin modified HSATQK*EHGW would result in T cell activation is based on a several factors. In order to address this, synthetic purification of flucloxacillin modified HSATQK*EHGW was performed for *in vitro* T cell priming studies.

5.4.5 ISOLATION OF FLUCLOXACILLIN MODIFIED HSATQKEGHW

The N-terminus of HSATQKEHGW was protected with an Fmoc group during synthesis by Synpeptide Co., Ltd (China). As flucloxacillin is attacked by nucleophilic residues, it was important to protect the free amine on the N-terminal histidine. Fmoc-HSATQKEHGW, flucloxacillin and Fmoc-HSATQKEHGW/flucloxacillin incubations were all incubated for 48 hours in 70% ACN/30% H_2O. All three samples were subsequently loaded onto the HPLC in equal quantities (22.4 µg peptide & 74.4 µg flucloxacillin) (Figure 5.11A). HPLC analysis revealed separation of the three samples as hoped. In order to identify the peak representing modified peptide, all three analytes were run on the mass spectrometer. Unfortunately, although analyte differences were observed, these did not perfectly align with the purification HPLC traces (Figure 5.11B). Both the unmodified peptide, and the degradation products of flucloxacillin after a 48h incubation, were used to identify peaks which where exclusively present when Fmoc-HSATQKEHGW and flucloxacillin were co-incubated for 48 hours. In order to determine which peak was indicative of Fmoc-HSATQK*EHGW, manual fractionation was performed.

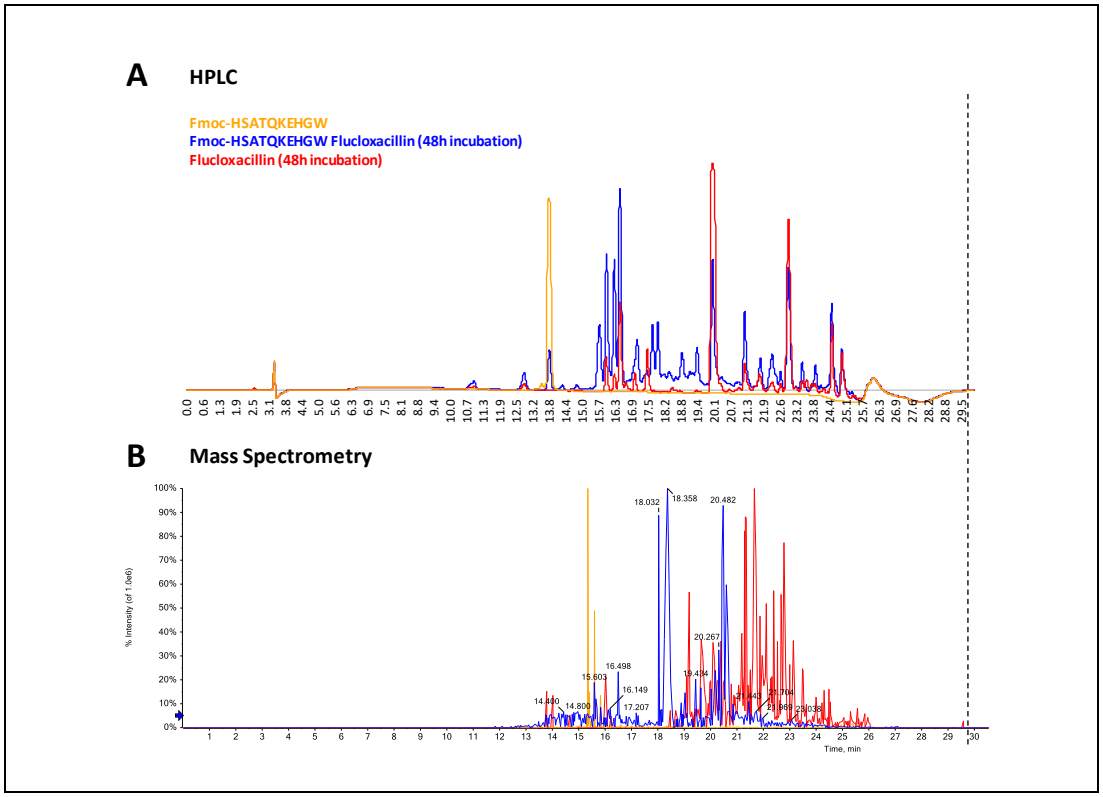


Figure 5.11. Fractionation of Fmoc-HSATQKEHGW, flucloxacilln and the product of the Fmoc-HSATQKEHGW flucloxacillin reaction. (A) HPLC and (B) mass spectrometry were used to identify analytes relating to the modification of Fmoc-HSATQKEHGW. Although separation was achieved using both HPLC and mass spectrometry, they were not comparable.

Unmodified Fmoc-HSATQKEHGW was fractionated (Figure 5.12A, i) and collected (13.6 to 13.9 min) for mass spectrometric analysis. MS/MS revealed singly, doubly and triply charged *m/z* ions corresponding to the unmodified peptide (Figure 5.12B). It was also observed that the absorbance reading for the unmodified peak was reduced in the incubation with flucloxacillin, indicating modification and retention time shift. Immediately after the unmodified peptide peak on the HPLC, an analyte only present in the incubation of flucloxacillin with Fmoc-HSATQKEHGW appears (15.5 to 15.85 min) (Figure 5.12A, ii). After collection of the fraction, mass spectrometry was performed, revealing both doubly and triply charged ions relating to the mass of Fmoc-HSATQK*EHGW (Figure 5.12B). All other peaks present exclusively in the incubation of drug alone and unmodified peptide (Figure 5.12A, iii) could not be determined, however were not indicative of modified peptide.

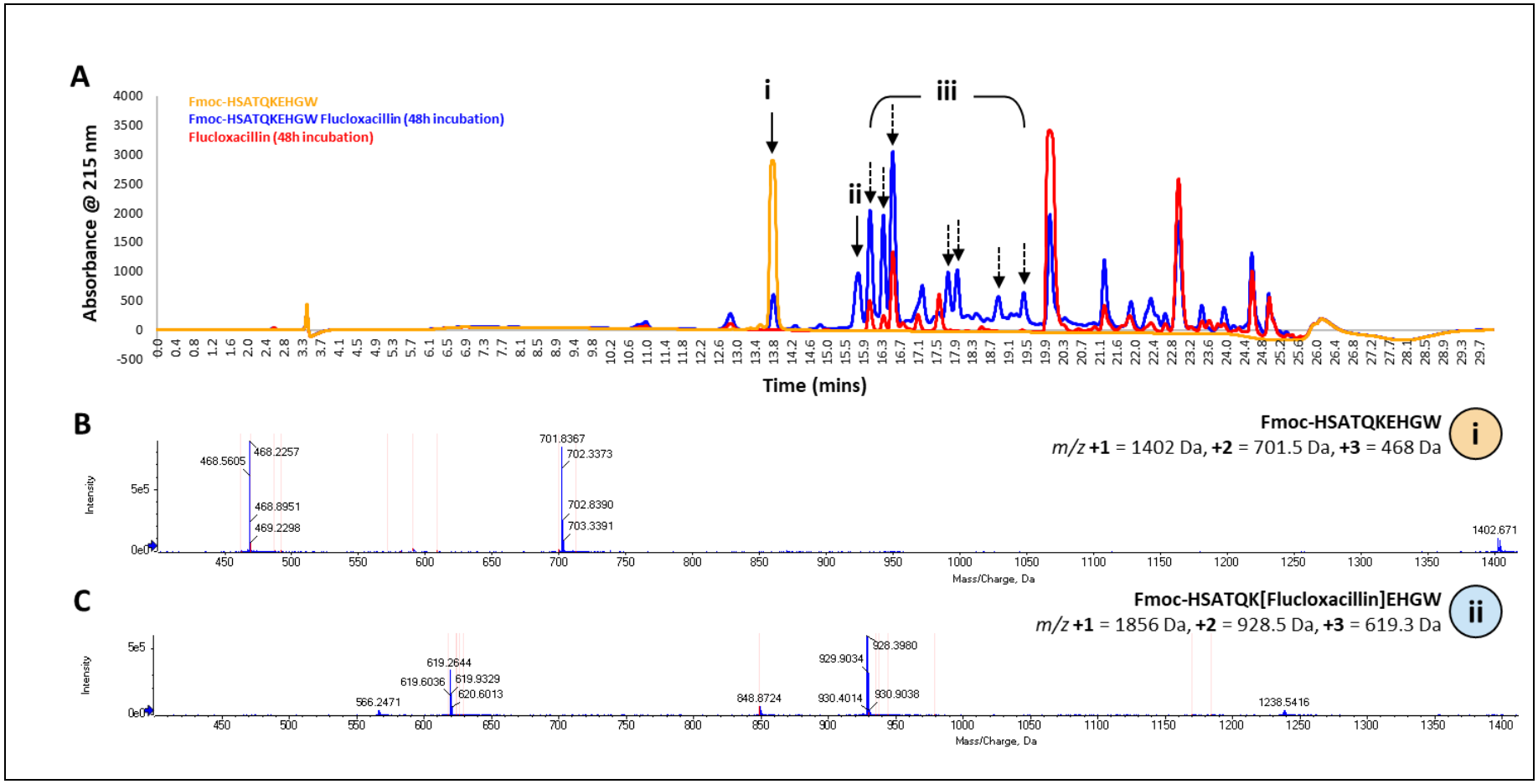


Figure 5.12. HPLC fractionation and mass spectrometric characterization of modified Fmoc-HSATQK*EHGW. (A) HPLC was used to determine the peak corresponding to the modifed peptide. Flucloxacillin incubated for 48 hours, and unmodified peptide, were used to determine absorbance readings only present in the reaction mixture. (B) Unmodified Fmoc-HSATQKEHGW was collected and confirmed to be of the correct mass (A, i). (C) Modified Fmoc-HSATQK*EHGW was collected (A, ii) and also confirmed using mass spectrometry. (A, iii) while other analytes were present, they did not correspond to the modification of Fmoc-HSATQKEHGW.

Piperidine was used for removal of the Fmoc from the N-terminal histidine of modified HSATQK*EHGW. HPLC was used to assess the retention time and purity of the deprotected modified peptide. As anticipated, the modified peptide with the Fmoc removed appeared as a single peak (Figure 5.13A, iii). As the modified peptide was generated for use in functional T cell culture assays, the purity of the collected fraction was investigated further using mass spectrometry. Modified HSATQK*EHGW fractions were prepared for mass spectrometric analysis using SCX prior to loading. Samples were deemed to be pure through the instance of 2 distinct peaks on the MS trace (Figure 5.13B). Both peaks corresponded to the modified peptide, with the separation due to different isomers of the modified peptide interacting with the analytical column differently.

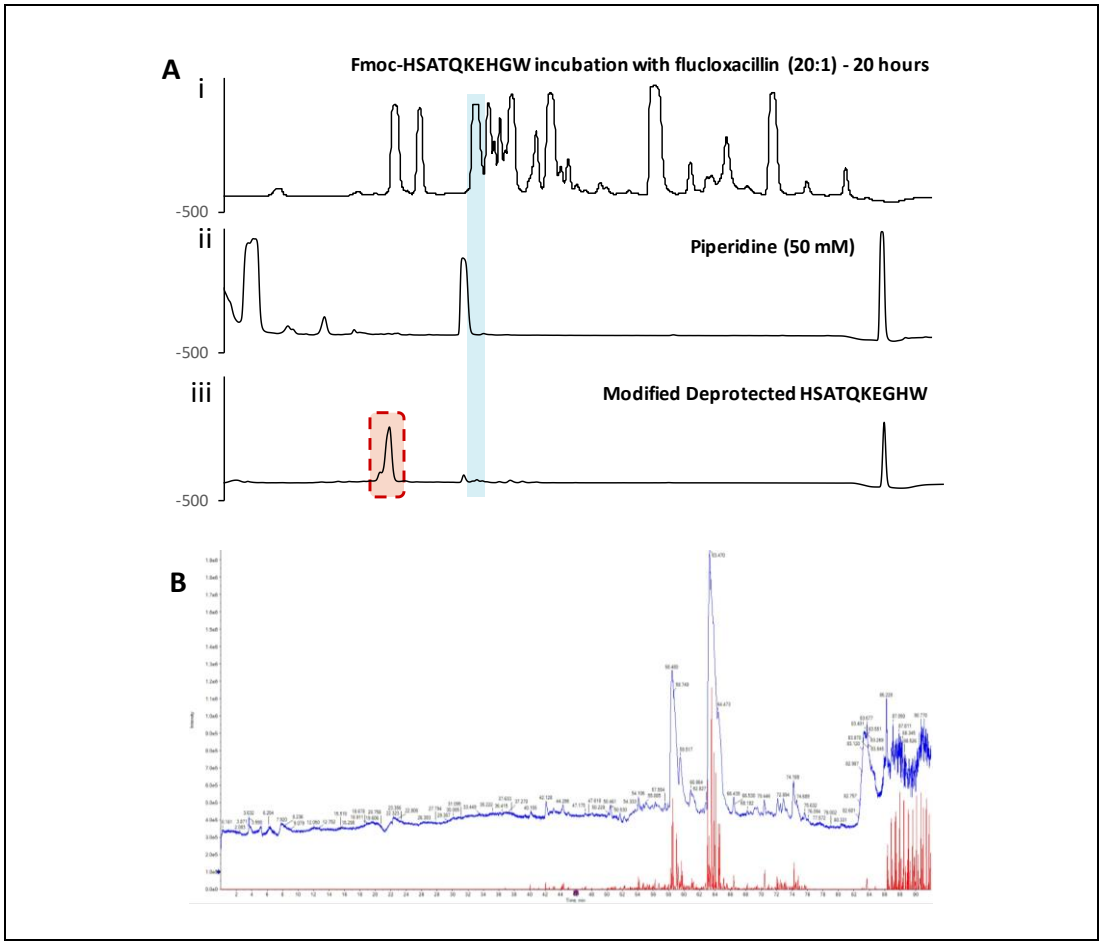


Figure 5.13. Deptotection of Fmoc-HSATQK*EHGW using piperidine. (A) HPLC separation of the (i) Fmoc-HSATQKEHGW flucloxacillin reaction mixture (ii) piperidine and (iii) the product of Fmoc-HSATQK*EHGW (red box) after incubation with piperidine at a 1:10 molar ratio. (B) Mass spectrometric analysis of the purity of HSATQK*EHGW after SCX preparation.

5.4.6 COMPARISON OF ELUTED VS SYNTHETIC MODIFIED HSATQK*EHGW

Synthetic HSATQK*EHGW was generated to confirm the sequence annotation of the eluted peptide; for which 3 different MS/MS spectra were acquired.

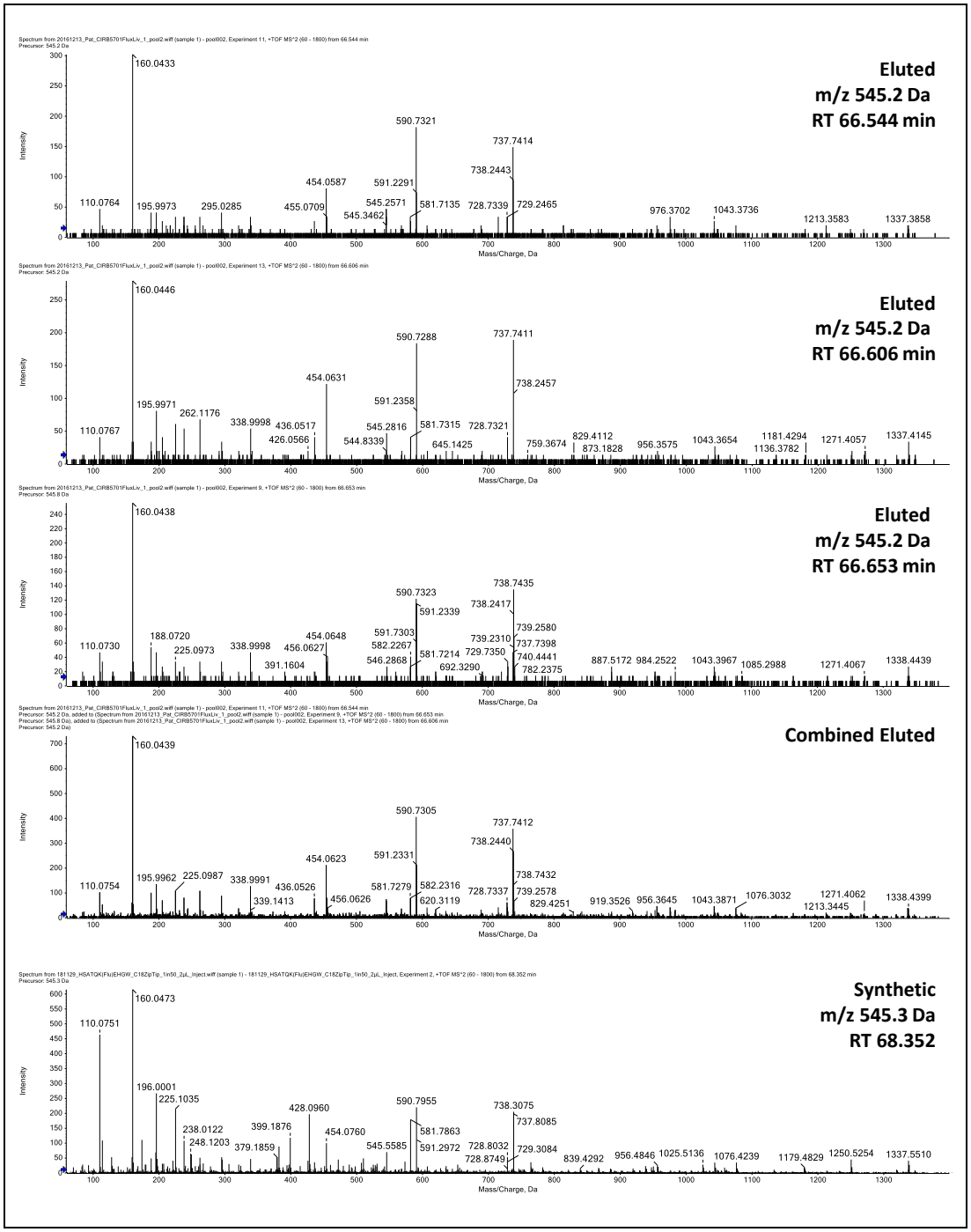


Figure 5.14. Alignment of eluted and synthetically modified HSATQK*EHGW. MS/MS spectra of eluted HSATQKEHGW were combined and compared with the fragment ion seried of the synthetic peptide A number of matches with similar ion intensities were identified between both peptides.

As the eluted peptides were in low abundance, combination of multiple spectra was performed to confirm true positive fragment ions from background noise. Upon alignment, it was apparent that many of the same fragment ions were found in both the synthetic and the eluted HSATQK*EHGW peptide (Figure 5.14). Relative intensities of fragment ions were normalised to the TIC from both the synthetic and combined eluted peptide. Fragment ions from both spectra were subsequently compared to theoretically expected ions for HSATQK*EHGW. Using a mass tolerance of 0.5 Da, to incorporate average isotopic masses, accepted fragmentation ions from each peptide were compared using Pearson correlation. Correlation was found to be signification at the 0.01 level using a 2-tailed test, with a correlation value of 0.76 (Figure 5.15).

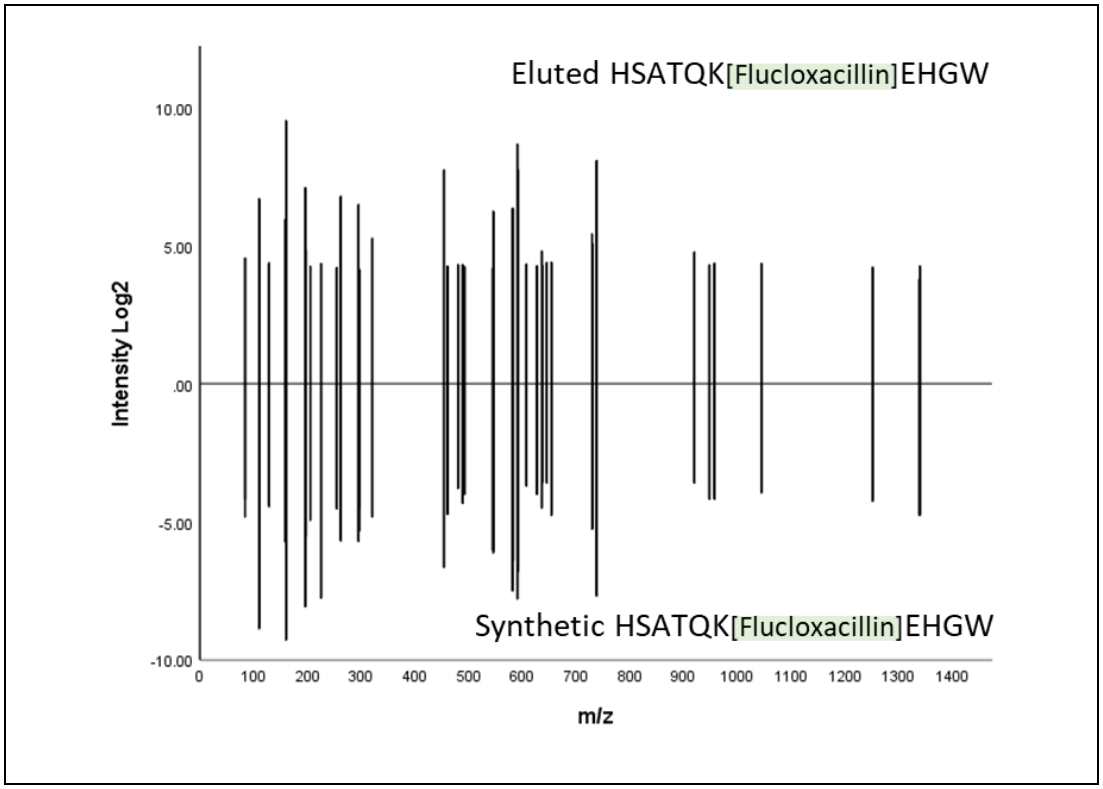


Figure 5.15. Comparative analysis of fragment ions from the eluted and synthetically modified peptide HSATQK*EHGW. The Log2 fragment ion intensities of theoretically correct ions shared between synthetic and eluted were compared and analyzed using Pearsons correlation.

5.4.7 SYNTHETIC PEPTIDE YIELD OPTIMISATION

In order to perform functional T cell studies with the modified peptide, peptide stability during purification as well as a high yield were necessary. This was both to minimise batch to batch variation in addition to practicality. HPLC fractionation of the Fmoc peptide incubated with flucloxacillin was performed immediately after 48 hours, 7 days storage at -20°C and overnight at room temperature in the HPLC autosampler (Figure 5.16A). All three absorbance traces appeared identical, with crucially, the modified fraction remaining. To maximise the yield of the modified fraction, increasing volumes of the incubation mixture (0.5 mM peptide, 5 mM flucloxacillin) were loaded onto the HPLC (Figure 5.16B). Separation of the modified peptide from the rest of the reaction mixture, i.e. unmodified peptide and flucloxacillin degradation products, was only possible up to a 40 µL injection volume (44.8 µg peptide, 148.8 µg flucloxacillin).

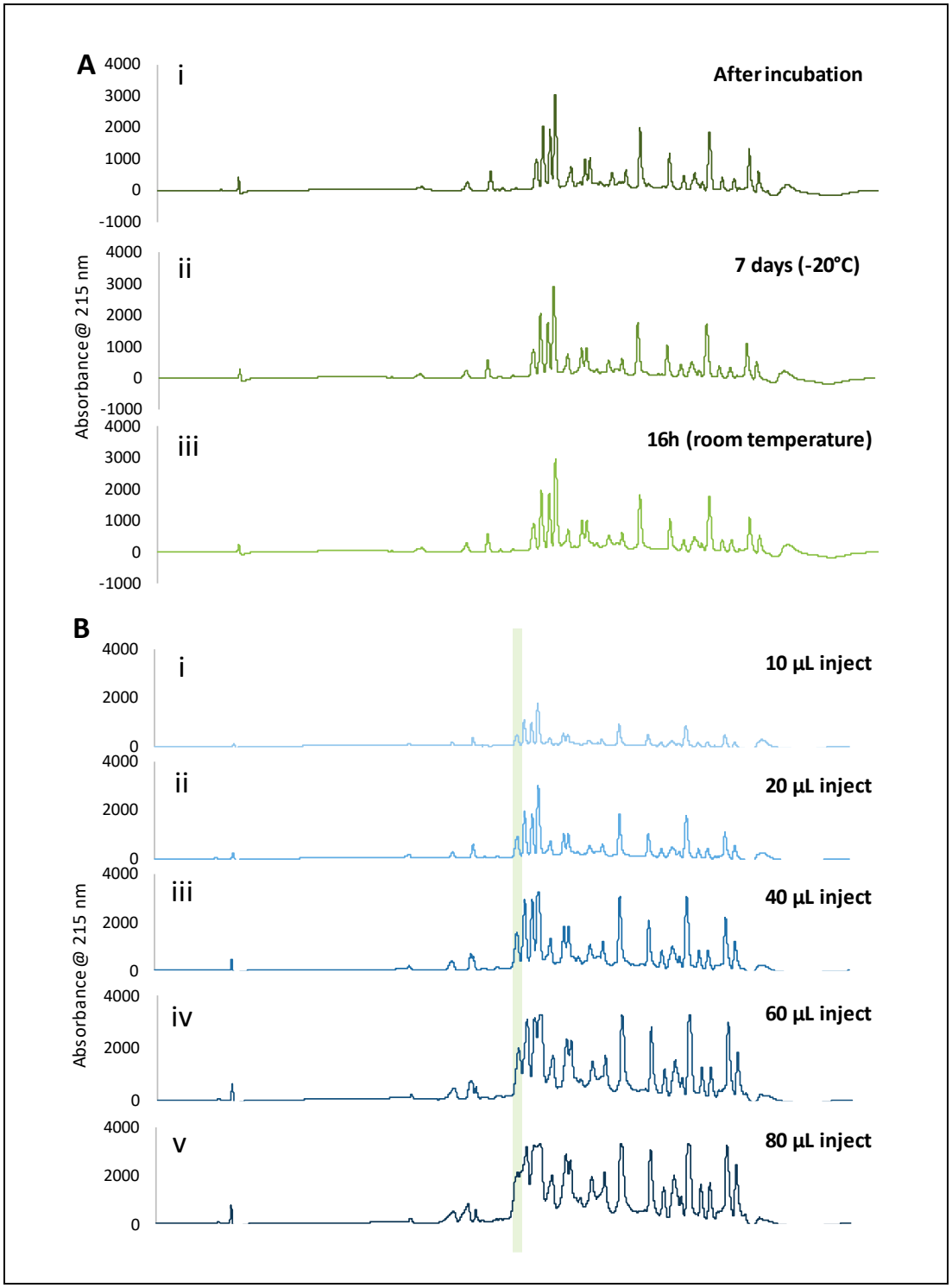


Figure 5.16. Assessment of the stability of the Fmoc-HSATQKEHGW flucloxacillin reaction mixture and maximum loading calculation. (A) The reaction mixture was loaded onto the HPLC after (i) 48 hours, (ii) overnight in the autosampler and (iii) for 1 week at -20°C. All traces were comparable, indicating the reaction mixture was stable. (B) Maximum loading of the reaction mixture while maintaining separation of the modified peptide fraction was assessed. A 40 µL injection of the reaction mixture (iii) was identified as optimal, to prevent contamination from other analytes.

5.5 DISCUSSION

Although advanced proteomics technologies have allowed global, systems-wide investigations of peptides presented to the immune system, less is known about the antigen processing and presentation in the presence of drugs, in particular reactive drugs that can covalently bind to proteins/peptides. It has been postulated that peptides derived from haptenated protein are likely to trigger drug-specific immune responses, in some cases, leading to tissue injury. Indeed, the screening of new drugs for bioactivation to chemically reactive intermediates is based on the theoretical concerns that such metabolites may induce an immune response in patients and cause serious idiosyncratic reactions that are only detected late in drug development (Park *et al.*, 2011). Much progress has been made to characterize the phenotype and function of the drug-specific immune cells that are involved in ADRs, but the chemistry of haptenated peptides naturally presented by specific HLA alleles that can initiate an immune response remain to be defined. In this chapter, a streamlined workflow was developed that has enabled the detection of natural HLA class I peptides presented on the surface of antigen presenting cells treated with flucloxacillin. For the first time, it has been demonstrated that flucloxacillin-haptenated peptides can be naturally presented by HLA-B*57:01.

Previous studies have demonstrated a strong association between HLA-B*57:01 and susceptibility to flucloxacillin-induced liver injury. Furthermore, activation of T cells from patients with liver injury by flucloxacillin has been detected and shown to be HLA-B*57:01-restricted. Thus, the primary objective of this chapter was to identify the flucloxacillin-associated structures displayed by HLA-B*57:01 on the surface of antigen presenting cells using an immortalized HLA-B*57:01 expressing B cell line. Close to 3,000 HLA-B*57:01 binding peptides were identified from flucloxacillin treated cells, among which more than 30 flucloxacillin-haptenated peptides were detected. Because of the complex partial loss of flucloxacillin in the mass spectrometer, identification of peptide sequences is not possible

using commercial software. During initial MS detection, flucloxacillin modified peptides were detected with the whole drug intact, giving a mass addition of 453 Da. Upon MS/MS, flucloxacillin is itself fragmented, resulting in mass additions of 453 Da, 294 Da (loss of the thiazolidine ring) and unmodified product ions. In addition to the unknown protein source, unknown protease involved in proteosomal digestion and any unknown biological modifications, fragmentation of flucloxacillin makes the database search space too large to confidently assign any peptide sequences. In addition to the complexity described, the presence of spliced peptides in the immunopeptidome adds another layer of complexity in the use of bioinformatics for the annotation of MHC peptides (discussed in detail in the next chapter). Therefore, manual *de novo* sequencing allowed us to annotate 7 of the flucloxacillin-haptenated peptides.

Flucloxacillin-haptenated peptides were anticipated to be presented through processing-dependent and -independent pathways. Through the processing-dependent pathway, flucloxacillin-haptenated proteins likely undergo intracellular processing and HLA peptide loading processes prior to presentation on the cell surface. Although a huge diversity of flucloxacillin-haptenated intracellular proteins were detected, only a few modified peptides were found to be naturally processed and presented by HLA-B*5701 for T cell recognition. Many of the flucloxacillin-modified proteins may have an impact on other biological functions such as cellular signalling. For example, flucloxacillin-haptenated eEF1A protein is the second most abundant protein (1-3% of total protein content) expressed in many cells including immune cells and liver cells (Abbas, Kumar and Herbein, 2015). The site of flucloxacillin covalent binding, K371 (HTAHIACK*FA), is not one of the sites for epigenetic modification that is involved in its canonical role of translation elongation (Andersen, Nissen and Nyborg, 2003; Hamey and Wilkins, 2018); however, this modification may play a role in non-canonical functions including cell regulation of the cytoskeleton through AKT and PI3K signalling pathways (Amiri *et al.*, 2007; Abbas, Kumar and Herbein, 2015).

Flucloxacillin was found to directly bind to peptides already presented on the cell surface within HLA peptide binding cleft. In this respect, a flucloxacillin-haptenated peptide (TAAQITQRK*W) and an unmodified counterpart, derived from HLA alpha chain, were detected, indicating direct peptide modification is possible. Furthermore, flucloxacillin-modified peptides were detected after C1R-B*57:01 cells were treated with flucloxacillin for only 10 minutes, providing firm evidence of direct haptenation. This direct haptenation could lead to immediate activation of T cells upon flucloxacillin treatment without the need for antigen processing. Indeed, Wuillemin *et al* demonstrated that stimulation of flucloxacillin-specific T cells could occur independent of proteasomal processing (Wuillemin *et al.*, 2013). Why flucloxacillin selectively modified HLA-B*57:01 binding peptides is not clear. One possible explanation is that flucloxacillin binds directly to peptide-HLA-B57:01 complexes with high affinity, and this selective non-covalent interaction could position the drug in favourable orientations to facilitate covalent binding. Our preliminary data with piperacillin, another drug with a similar core structure to flucloxacillin, indicates that the display of haptenated HLA-B*57:01 peptides is not a common feature of all β-lactam antibiotics.

Apart from altering antigen presentation, reactive drugs may interfere with antigen processing through multiple pathways. It is feasible that covalent binding, if it is close to the proteasomal cleavage sites, may alter enzymatic cleavage leading to the generation of *de novo* peptides. There already exists evidence that proteins with post-translational modifications are more resistant to proteolysis and the cleavage patterns are influenced by the sites of modification (Ninkovic and Hanisch, 2007; Purcell, van Driel and Gleeson, 2008). Indeed, covalent binding of flucloxacillin to proteins led to the presentation of peptides with a C-terminal alanine; VSDHEATLR*CWA, ISDHEATLR*CWA, HTAHIACK*FA. All three of these flucloxacillin-haptenated peptides were derived from intracellular proteins that were subsequently displayed by HLA-B*57:01. Importantly, a C-terminal alanine is strongly disfavoured by proteasomal cleavage and TAP transport (Tenzer *et al.*, 2005), providing

strong evidence of altered antigen processing. While this appears to be true of modified peptides terminating with a C-terminal alanine, global analysis of proteasomal activity showed potential disruption to be inconclusive. That said, the bioinformatic tools used to generate unique peptide lists cannot consider flucloxacillin modified peptides. Therefore, the results of the analysis must be taken with caution. The molecular features of flucloxacillin may assist the binding of proteins to proteasome catalytic sites, which strongly favour hydrophobic residues, leading to the generation of peptides from unique parts of proteins. Indeed, ligands derived from control- and flucloxacillin-treated cells were found to display strikingly distinct pools of peptides derived from the same protein.

In addition to haptenation, drugs such as flucloxacillin could change the shape and chemistry of the antigen-binding cleft through direct covalent binding to HLA molecules, leading to an alteration in the repertoire of endogenous HLA peptides. Immuno-detection and mass spectrometric analysis of the alpha chain of HLA-B*57:01 revealed that flucloxacillin can indeed irreversibly bind to the N-terminal glycine. Through analysis of the alpha chain protein sequence the glycine modification appears at the first amino acid upon cleavage of the signal peptide. This is of interest, as the signal peptide is responsible for transport of the MHC complex to the cell surface (Rapoport, 2007). Therefore, it is likely that flucloxacillin modification occurred extracellularly and did not result in the presentation of neo-antigens. That said, it is still possible that other binding sites within the HLA-B*57:01 protein are targeted by flucloxacillin.

Whether the novel flucloxacillin-associated peptides (haptenated and altered self-peptides, chapter 4) identified in this study play a role in the onset of flucloxacillin-induced liver injury remains to be determined. Evaluation of the immunogenicity of these peptides with T cells from patients with flucloxacillin-induced liver injury is extremely challenging due to the difficulty of generating synthetic flucloxacillin-haptenated peptides. We have therefore attempted to predict the immunogenicity of haptenated peptides to reduce the numbers of

candidate peptides to be evaluated in future studies. Since there is no algorithm available for haptenated peptides, HLA binding algorithms for predicting the binding affinity of the corresponding-unmodified peptides were used. Several of these peptides were predicted to be weak HLA-B*57:01 binders. This is not surprising as a flucloxacillin modification will alter the 3D molecular arrangements within the peptide structure and hence its ability to ligate HLA proteins. The flucloxacillin molecule will likely orientate towards the solvent surface and thus has the potential to be displayed to the T cell receptor on specific T cells. Upon theoretical docking, the conformation of the flucloxacillin haptenated HSATQK*EHGW peptide is indeed different from the native peptide when it bound to HLA-B*57:01, with the flucloxacillin side chain reaching towards the solvent surface for T cell recognition.

Collectively, this data shows how the β-lactam antibiotic flucloxacillin interacts covalently with HLA-B*57:01. For the first time, drug-haptenated MHC peptides have been characterized though binding to a specific HLA, B*57:01, resulting in cell surface presentation. These peptides originate from either direct haptenation of peptides already displayed in the HLA peptide binding groove or following processing of drug-modified intracellular proteins. These data indicate that diverse drug-associated antigens have the potential to trigger drug-specific T cell responses and the diversity may drive the complex clinical picture of drug-induced liver injury.

CHAPTER 6 - OPTIMISING FLUCLOXACILLIN MHC PEPTIDE CHARACTERIZATION

6.1 INTRODUCTION

The complexity of the data generated when investigating MHC peptides using mass spectrometry poses a real challenge to the analysis pipelines developed, and make manual interpretation of entire data files impossible (Codrea and Nahnsen, 2016). Two main challenges exist; firstly, the accurate identification of MHC peptides and secondly, the characterization of drug modified MHC peptides. Typically, the database used for peptide-to-spectrum matches (PSMs) can be reduced in size by adding assay specific details into the search parameters. This could be proteins/organisms of interest, known digestive enzymes or other known modifications, for example. In reducing the number of theoretical sequences contained within the database, confidence scores for PSMs are subsequently boosted. The challenge with MHC peptides comes from the inability to positively hypothesise the protein source, the presence of biological modifications or the proteolytic enzyme involved in protein digestion prior to loading into the MHC. Consequently this results in a very large database, and as a result, weaker confidence in correctly determining the peptide sequences of the spectra acquired. Therefore, bioinformatics pipelines are used to assign peptide identifications with the additional of confidence scores and statistical validation.

The volume of data that is now generated from state-of-the-art mass spectrometers can be in excess of 1TB (Kirkwood *et al.*, 2013; Kim *et al.*, 2014) and therefore analysis is becoming a challenge for the computational methods that were previously designed (Shteynberg *et al.*, 2011). It is generally accepted that three types of search engines exist; those that match with theoretical spectra, spectral library search engines and *de novo* search engines which attempt to derive sequence information from MS/MS fragmentation ion patterns alone (Shteynberg *et al.*, 2013). Perhaps the most common approach to assigning MS/MS spectra to peptide sequences comes from the *in silico* digestion of proteins contained within a database. By reducing this search space the confidence of PSMs can be boosted. In order to assign confidence scores to PSMs, decoy databases containing non-existent peptide

sequences are used. If a spectrum is assigned with high confidence to a decoy entry, FDRs can be calculated. While this can improve the reliability of peptide annotations used for subsequent analyses, the technique in which different search engines make their assignments vary, producing different outputs.

There are many pieces of software available for the bioinformatic analysis of mass spectrometric data; both open source and commercially available. Advantages and disadvantages to using a particular piece of software depends on the type of data being interrogated and the question that is being probed. Several search engines exist, each with different advantages and disadvantages. For example, MS Amanda is particularly suitable for high mass accuracy data, assigning more PSMs than commercially available search engines such as Mascot. Mascot, however, is particularly useful at identifying complex amino acid modifications, but is limited to peptides generated from a known protease and is not free to use (Dorfer *et al.*, 2014). While the statistical validation methods of open source algorithms can be interrogated, commercially available platforms do not usually disclose how PSM scoring is assigned. The combining of results from different search platforms has been used to improve accurate PSMs. For this approach the Institute for Systems Biology have been instrumental in their contributions to the development of novel proteomic tools through the Trans-Proteomics Pipeline (TPP) (Codrea and Nahnsen, 2016). More recently the TPP has introduced iProphet (Shteynberg *et al.*, 2011) which incorporates multi-level integrative analysis of data dependant acquisition (DDA) data. The purpose of iProphet is to compare the results from a number of shot-gun DDA searches and statistically validate the results into a single output, thus improving the overall confidence of a PSM and improve source protein identification (Shteynberg *et al.*, 2011).

A comprehensive review into the improvement in PSM assignment with iProphet resulted in an increase of ~40% when combining six different search engines (InSpecT, MyriMatch, OMSSA, Mascot, SEQUEST and X!Tandem). This could be improved by a further ~26% when

a seventh search engine, SpectraST (Lam *et al.*, 2007), was incorporated into the analysis (Shteynberg *et al.*, 2013). SpectraST is different in its PSM assignment due to fragment ions being searched against a spectral library rather than protein sequence databases, making it a more precise and efficient method (Lam *et al.*, 2007). While SpectraST offers increased precision and efficiency only spectra that have previously been annotated will give a positive match, making it impossible for searching novel peptides. A few studies have now utilised the power of combination searches using iProphet in identifying MHC peptides. The murine MHC class I immunopeptidome was interrogated using the search engines Comet (Eng, Jahan and Hoopmann, 2013), MSGF (Kim and Pevzner, 2014) and X!Tandem (Craig, Cortens and Beavis, 2004) through the TPP PeptideProphet pipeline, combining the results in iProphet yielding 28,448 high confidence peptides (Schuster *et al.*, 2018). While the number of positive PSMs from each individual search are not available, the addition of analysis using combination searches undoubtedly improved the number of confident PSMs.

The use of open source software is useful for experienced data scientists as users can understand PSM scoring and edit scripts to suit their own needs. That said, these algorithms are all developed based on bottom-up proteomic techniques where peptides are generated from known proteases. Indeed, this incorporates a level of bias as PSMs are assigned based on a database-driven approach of possible theoretical peptides. While the PSM assignment algorithms are unknown, commercially available software such as Protein Pilot v.5 (used in this study) and PEAKS (v.8.5) incorporate *de novo* sequencing steps in combination with a database-driven approach. While limitations exist with relation to the modification of internal algorithms for personalised approaches, the possibility to identify novel MHC peptides not included in a pre-compiled database is far greater (Purcell, Ramarathinam and Ternette, 2019). Spliced MHC peptides were first identified by Hanada *et al* where two FGF-5 peptide sequences that originated from different regions of the protein were recognized by cytotoxic T lymphocyte clones (Hanada, Yewdell and Yang, 2004). This phenomenon was

later described by Vigneron et al where spliced gp100 resulted in two non-continuous fragments of the protein being presented. The presentation of this peptide was restricted when protease inhibitors were introduced indicating that protein digestion and MHC presentation pathways were resulting in the splicing of peptides on the cell surface (Vigneron *et al.*, 2004). Since these early studies many other groups have reported the incidence of spliced peptides originating from two or more protein sources (Hanada, Yewdell and Yang, 2004; Vigneron *et al.*, 2004; Ebstein *et al.*, 2016; Liepe *et al.*, 2016; Mishto and Liepe, 2017; Platteel *et al.*, 2017; Mannering *et al.*, 2018). A comprehensive study into novel analysis workflows to positively identify spliced peptides was recently published by Faridi *et al*. It's thought that cis-spliced-peptides, where peptide segment are derived from the same protein, may account to up to 30% of the immunopeptidome (Liepe *et al.*, 2016). While trans-spliced-peptides consisting of peptide sequences derived from different proteins do too exist, the function of all spliced peptides is not yet fully understood (Faridi *et al.*, 2018). In the present study Faridi *et al* presents well described bioinformatic workflows that can positively discriminate between linear and spliced peptides using *de novo* sequencing methods (Faridi *et al.*, 2018). Therefore, in the case of spliced peptides matching fragment ions to *in silico* digests would not result in positive PSMs. The *de novo* feature of ProteinPilot and PEAKS is instrumental in positively identifying genuine spliced ligands; however due to the commercial nature of these products they are extremely costly and do not integrate with other pieces of software well.

It goes without saying that the bioinformatic analysis of MHC peptides is inherently challenging. An additional complication arises when drug involvement is added the search parameters. In the case of abacavir, a global change in MHC peptides was identified, therefore Protein Pilot was applicable for this study (Illing *et al.*, 2012). In the case of flucloxacillin, a global change was too identified in the C-terminal amino acid abundance with the use of Protein Pilot v.5 (Chapter 4), validating its use for immunopeptidomic analysis.

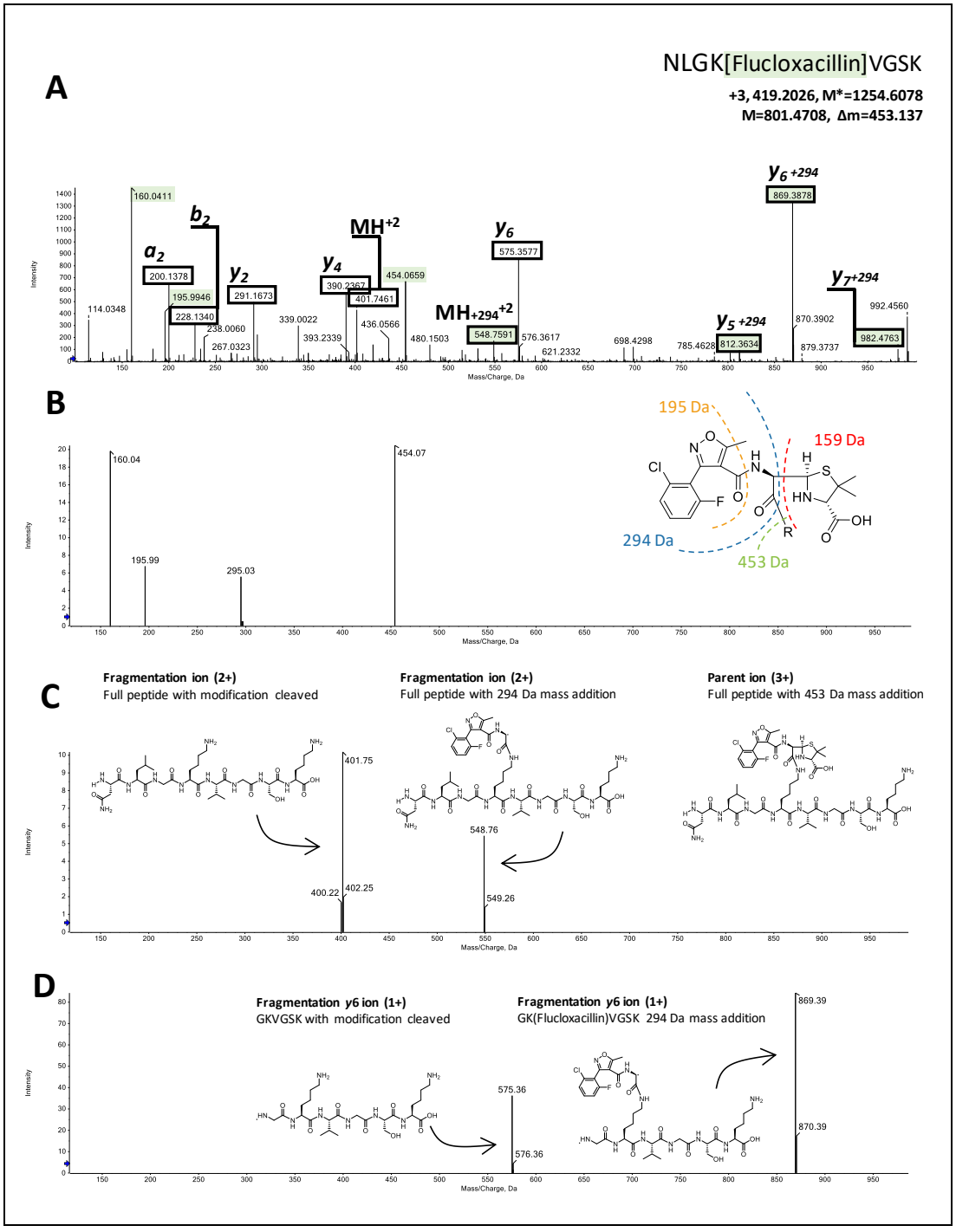


Figure 6.1. Flucloxacillin modification of HSA tryptic peptide NLGKVGSK. Mass spectrometric analysis of tryptic peptides derrived from flucloxacillin modified HSA. (A) Flucloxacillin modified NLGKVGSK was identified through the theoretical mass addition of flucloxacillin (453 Da), with characteristic fragmentation ions annotated. (B) Fragmentation ions relating to flucloxacillin. (C) Doubly charged fragmentation ions correspoinding to unmodified NLGKVGSK and partially modified peptide (+294 Da). (D) Singly charged fragmentation ions (*y*6 - GKVGSK) with and without modification.

Modification of peptides by chemicals can be incorporated into user specific 'special factors'

in Protein Pilot, where additional search parameters can be included in the algorithm. This

allows target amino acids, the probability of a modification arising and the chemical formula of drug to be considered when PSMs are performed. The challenge with flucloxacillin, however, is that the drug breaks apart during MS/MS fragmentation. Take the representative example NLGKVGSK, a tryptic peptide derived from HSA. Flucloxacillin modified NLGKVGSK can be manually identified through the theoretical addition of the flucloxacillin MW (453 Da) to that of the peptide (801.47 Da). Through searching the peptide spectra acquired, flucloxacillin modified NLGKVGSK can be identified as a triply charged ion ((1254.6078+3)/3 = 419.2026 Da) (Figure 5.1A). Fragmentation ions corresponding to flucloxacillin (160, 195 & 454 Da) confirm the presence of drug on the peptide (Figure 5.1A & B). In addition to the triply charged parent ion with a full flucloxacillin mass addition of 453 Da, a doubly charged unmodified fragment ion (401.75 Da) and a doubly charged fragment ion with a partial flucloxacillin modification (548.76 Da), are present (Figure 5.1C). This partial modification is indicative of the loss of the thiazolidine ring (159 Da) (Figure 5.1B). Additional fragment ions are also observed, both with the presence and absence of the flucloxacillin modification (Figure 5.1D). This contributes to a diverse array of possible fragment ions, with multiple ions corresponding to the same peptide sequence (+/- partial or full modification) (Figure 6.2).

NLGKVGSK					NLGK(Flucloxacillin)VGSK								
***b* ions**				***y* ions**	***b* ions +453**	***b* ions +294**	***b* ions**				***y* ions**	***y* ions +294**	***y* ions +453**
-	1	**N**	8	-	-	-	-	1	**N**	8	-	-	-
228.1343	2	**L**	7	688.4352	-	-	228.1343	2	**L**	7	688.4352	982.4352	1141.435
285.1557	3	**G**	6	575.3511	-	-	285.1557	3	**G**	6	575.3511	869.3511	1028.351
413.2507	4	**K**	5	518.3297	866.2507	707.2507	413.2507	4	**K**	5	518.3297	812.3297	971.3297
512.3191	5	**V**	4	390.2347	965.3191	806.3191	512.3191	5	**V**	4	390.2347	-	-
569.3406	6	**G**	3	291.1663	1022.341	863.3406	569.3406	6	**G**	3	291.1663	-	-
656.3726	7	**S**	2	234.1448	1109.373	950.3726	656.3726	7	**S**	2	234.1448	-	-
-	8	**K**	1	147.1128	-	-	-	8	**K**	1	147.1128	-	-

Figure 6.2. Theoretical *b* and and *y* fragment ions from NLGKVGSK. Fragment ions from unmodified NLGKVGSK result in two ion series. Due to drug fragmentation, an increased in up to 6 theoretical fragment ions series is observed when peprides are modified with by flucloxacillin.

Although not necessarily related to the drug, other PTMs will further amplify the number of potential fragment ions. In the case of flucloxacillin, fragmentation results in additional ions (relatively) unique to the drug. These will too be used as peptide derived ions by the algorithm while attempting to make PSMs. With these complexities, in addition to those previously described, the bioinformatics analysis of flucloxacillin-modified MHC peptides is limited using established search methods.

Once confident PSMs have been assigned, further bioinformatics analysis is often required for the interrogation of the MHC peptide repertoire. Upon analysis of the immunopeptidomics data in Chapter 5, several different Excel spreadsheets were generated to produce the final peptide lists unique to each treatment. While for this thesis the total number of elutions was 9 (3 x untreated, 3 x abacavir treated and 3 x flucloxacillin treated) this manual approach would not be appropriate for a high throughput data analysis. Programming languages are becoming more and more popular for researchers to overcome the issue of 'big data'. One particularly useful language and environment is 'R' (www.r-project.org) and R studio, respectively. R is primarily used for statistical computing and the generation, manipulation and display of data for publication quality graphics. Another attractive feature with R is that it is free to use and therefore has developed a large community base where users share script packages and offer peer support. Ultimately, 'R' can speed up the analysis of MHC peptide data which makes it instrumental in the development of high-throughput analysis platforms.

6.2 AIMS

The 'black-box' nature of Protein Pilot's scoring and PSM assignment makes the incorporation of novel bioinformatics tools into its search impossible. In order to overcome these challenges, software engineers in the Computational Biology Facility at the University of Liverpool were involved in generating novel bioinformatic tools that could be used for the analysis of MHC peptide data acquired for this thesis. Using their expertise, the primary aim of this project was to develop a pipeline for the automatic identification and characterization of flucloxacillin modified MHC peptides. A second aim of this chapter was to use 'R' within R-Studio to write scripts that would speed up data analysis. Using these novel pipelines will enable a more streamlined approach to the study of immunopeptidomic data.

N.B. A project report, prepared by data scientists at the Computational Biology Facility, can be found at the end of this chapter (Appendix 1).

6.3 METHODS & RESULTS

Due to the nature of this chapter, the methods and results will be combined into one section. The complexity of the MHC peptide mass spectrometric data makes it challenging to develop a single method which fits all the requirements; therefore, the progression of this chapter will be presented in a chronological format.

6.3.1 BIOINFORMATIC ANALYSIS OF DRUG MODIFIED TRYPTIC PEPTIDES

Protein Pilot was used for the analysis of MHC peptide data due to its availability (commercial licence required) and *de novo* sequencing capabilities. However, it was not possible to positively identify any flucloxacillin modified MHC peptides. While a lot of the difficulties can be attributed to the nature and complexity of MHC peptides themselves, Protein Pilot was still unable to confidently assign flucloxacillin modification to tryptic peptides. In this example the tryptic HSA peptide (ASSAKQR) can be manually annotated, with almost all the fragmentation ions accounted for (Figure 6.3A). When a search was set up using Protein Pilot, with the algorithm set to identify a highly probable mass addition of 453 Da on lysine residues, an incorrect annotation (GSYMEVEDN[Deamidated]R) was assigned (Figure 6.3B). As the sequence assignment confidence was low (<1%) it would not be considered in further analysis. In this case, it is understandable due to the user specific parameters stating a mass addition of 453 Da on lysine would occur. While this is true of the parent ion, product ions have a mass addition of 294, due to the cleavage of the thiazolidine ring. Subsequent parameters were input into the search algorithm, accounting for a partial loss of 160 Da, however the results were limited to multiple hypotheses with a confidence score <1%. This simple experiment highlights the incapability of Protein Pilot to correctly assign drug modification to tryptic peptides where a partial flucloxacillin loss occurs, therefore it would be impossible to use this method for the detection of flucloxacillin modified MHC peptides.

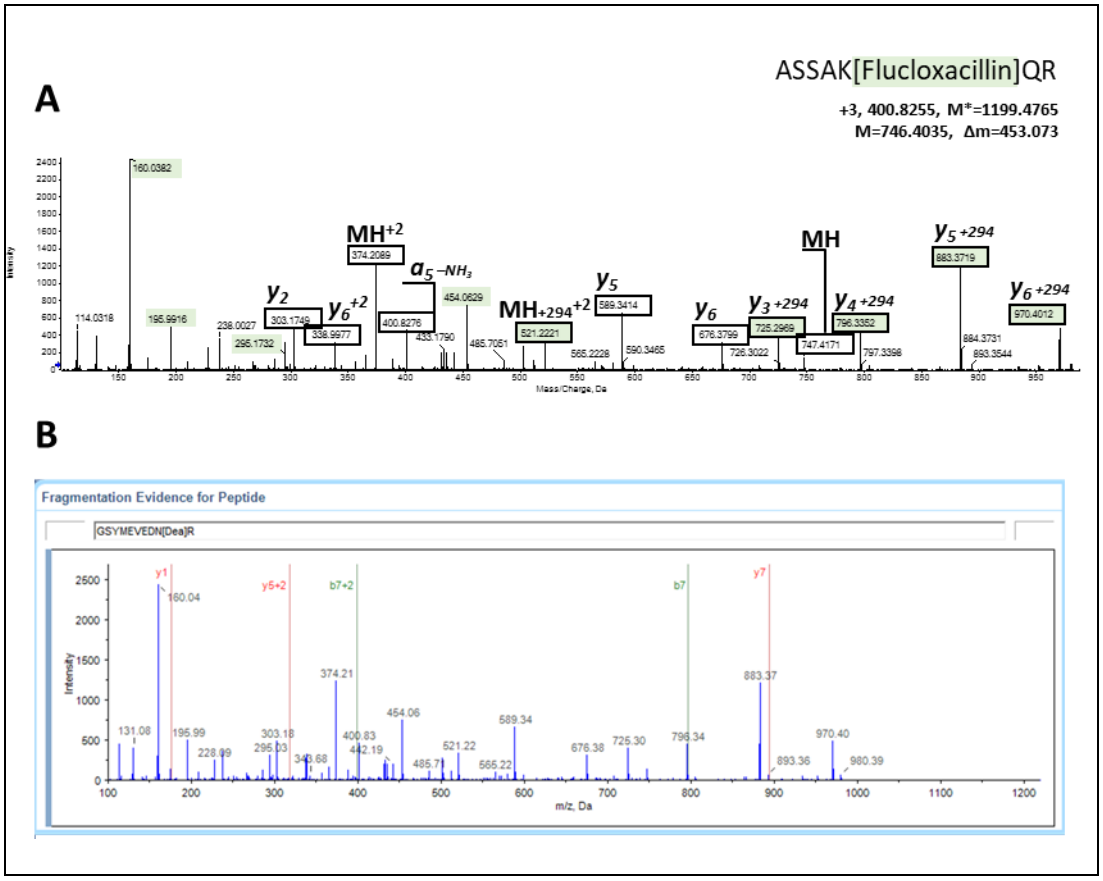


Figure 6.3. Identification of flucloxacillin modified tryptic HSA peptides using Protein Pilot. (A) A representivive example of a manually annotated tryptic HSA spectrum (ASSAK[Flucloxacillin]QR). Characteristic fragmentation ions (160, 195, 295 and 454 Da) confirm the presence of flucloxacillin. (B) The same specrtrum is annotated incorrectly when Protein Pilot investigates the fragment ion series, with fucloxacillin modification incorportated into the search algorith.

A pipeline was therefore initially generated for the identification of tryptic HSA peptides likely modified by flucloxacillin. For this, the search engine MASCOT (Matrix Science) was used due to its availability to our collaborators (licenced software) and previous successes with correctly annotating complex mass additions. To search through Protein Pilot, the original mass spectrometry output files (.wiff) can be used, as both data acquisition software (Analyst) and Protein Pilot are developed by Sciex. For MASCOT, .mgf (mascot generic format) files are required. Therefore, initial .wiff file searches were performed using Protein Pilot, and the peaks generated were subsequently exported as .mgf files.

Due to the complexities with flucloxacillin partial losses, initial searches were set up in MASCOT with a mass addition of 294, accounting for the loss of the thiazolidine ring. The

partial loss of the thiazolidine ring of flucloxacillin occurs during MS/MS fragmentation, even when relatively low collision energy (CE) is applied. Through diffusing flucloxacillin directly into the mass spectrometer (Sciex QTOF 5600, used for MHC peptide elution experiments), the loss of the thiazolidine ring was observed when CE was set to 10 (Figure 6.4A). Naturally, fragmentation of flucloxacillin is amplified when the CE is increased (Figure 6.4B).

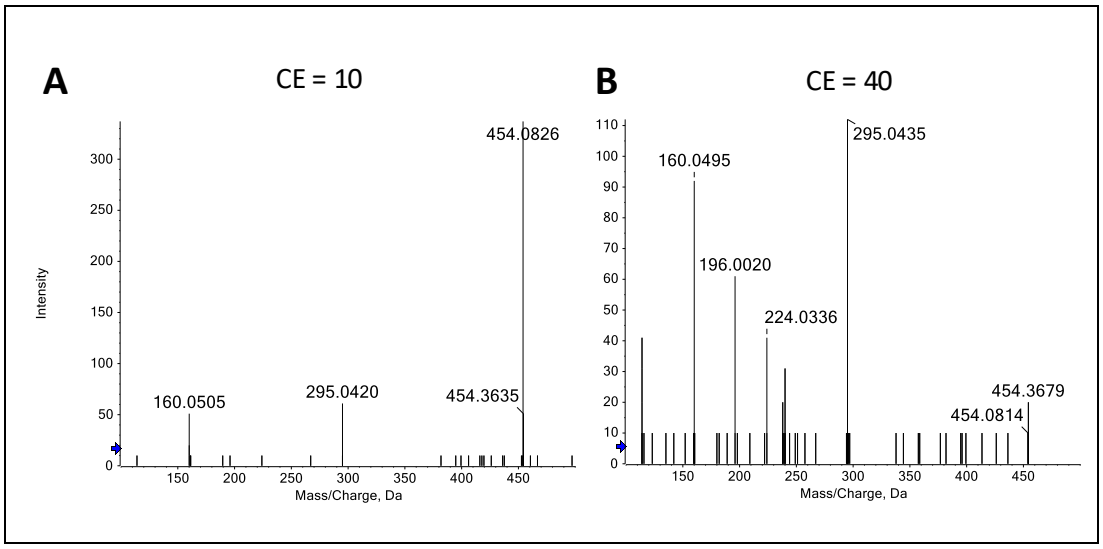


Figure 6.4. Flucloxacillin fragmentation ions. Flucloxacillin was manually injected into a TripleTof 5600 mass spectrometer (Sciex) with a collision energy of (A) 10 and (B) 40.

As the thiazolidine ring is particularly labile, a flucloxacillin mass addition of 294 Da can be observed on some parent ions, prior to MS/MS. MASCOT searches were set up to identify tryptic peptides (mass tolerance 15 ppm (MS1) and 0.05 Da (MS2)) with fixed carbamidomethyl (iodoacetic acid) modification of cysteine and a variable modification of 294 Da on lysine and arginine. The UniProt *Homo Sapiens* data base from February 2015 was set as the search database, with a decoy database incorporated. Indeed, when MASCOT (v.2.3.02) searches were performed on the .mgf files exported from Protein Pilot, some confident PSM were assigned (Table 6.1). Upon manual curation of the results, the sequence annotations presented were correct.

Table 6.1. Flucloxacillin modified HSA peptides using MASCOT search engine. MASCOT search parameters were adapted to accommodate a +294 Da mass addition at lysine and arginine residues.

Protein	*m/z*	Mass (Da)	Charge	Score	Sequence
sp|ALBU_HUMAN|	438.5377	1312.5913	3	63.27	AFKAWAVAR
sp|ALBU_HUMAN|	645.3244	1932.9514	3	56.75	KVPQVSTPTLVEVSR
sp|ALBU_HUMAN|	712.3169	2133.9289	3	55.76	EQLKAVMDDFAAFVEK
sp|ALBU_HUMAN|	679.6448	2035.9126	3	50.91	HPYFYAPELLFFAK
sp|ALBU_HUMAN|	588.0372	2348.1199	4	50.36	RHPYFYAPELLFFAKR
sp|ALBU_HUMAN|	474.9124	1421.7154	3	50.32	KQTALVELVK
sp|ALBU_HUMAN|	604.9386	1811.794	3	44.92	LDELRDEGKASSAK
sp|ALBU_HUMAN|	795.8611	1589.7076	2	42.5	LAKTYETTLEK
sp|ALBU_HUMAN|	450.5402	1348.5988	3	40.52	KYLYEIAR
sp|ALBU_HUMAN|	652.9246	1955.752	3	39.82	YKAAFTECCQAADK
sp|ALBU_HUMAN|	731.6786	2192.014	3	33.7	HPYFYAPELLFFAKR
sp|ALBU_HUMAN|	780.3782	2338.1128	3	27.44	VFDEFKPLVEEPQNLIK

While these initial results were promising, MASCOT is not a powerful tool for the analysis of non-tryptic peptides due to its lack of ability to perform *de novo* sequencing. Therefore, method optimisation was turned to PEAKS studio. While Protein Pilot has a *de novo* sequencing aspect (Purcell, Ramarathinam and Ternette, 2019), PEAKS was chosen due to its availability to the data scientists collaborating on this project. Peaks was set up to identify the full mass addition of 453 Da, with an expected partial loss of 294 Da. The remaining search parameters were comparable to those used in MASCOT. Positive results from the search were obtained (Table 6.2) with the localisation of the modification annotated in the peptide sequence. From the results (Table 6.2) those highlighted contain the same sequence information as MASCOT. In total, every single peptide identified in MASCOT was identified using Peaks. Peaks, however, was able to provide additional peptide sequences with relatively high scores (a Peaks score (-10logP) of 20 is generally regarded as highly confident). Therefore, the decision to utilise Peaks appeared to be justified.

Table 6.2. Flucloxacillin modified HSA peptides detected using Peaks studio. Peaks search engine was used to identifiy a mass addition of +453 Da at lysine and arginine residues, with a partial loss of 159 Da applied to accommodate for the cleavade of the thiazolidine ring. Highlighted entries were previsouly identified using MASCOT (Table 6.1).

Accession	m/z	Mass	Score	Peptide
P02768|ALBU_HUMAN	588.7678	2351.044	80.62	HPYFYAPELLFFAK(+453.06)R
P02768|ALBU_HUMAN	765.3287	2292.964	78.12	EQLK(+453.06)AVMDDFAAFVEK
P02768|ALBU_HUMAN	613.2571	2448.978	73.94	NEC(+57.02)FLQHK(+453.06)DDNPNLPR
P02768|ALBU_HUMAN	527.9226	1580.748	68.3	K(+453.06)QTALVELVK
P02768|ALBU_HUMAN	833.39	2497.144	67.98	VFDEFK(+453.06)PLVEEPQNLIK
P02768|ALBU_HUMAN	698.336	2091.987	66.97	K(+453.06)VPQVSTPTLVEVSR
P02768|ALBU_HUMAN	784.6937	2351.044	64.82	R(+453.06)HPYFYAPELLFFAK
P02768|ALBU_HUMAN	732.6566	2194.943	63.72	HPYFYAPELLFFAK(+453.06)
P02768|ALBU_HUMAN	627.7961	2507.145	59.54	RHPYFYAPELLFFAK(+453.06)R
P02768|ALBU_HUMAN	795.7043	2384.086	58.66	SLHTLFGDK(+453.06)LC(+57.02)TVATLR
P02768|ALBU_HUMAN	652.2982	1953.869	54.56	FYAPELLFFAK(+453.06)R
P02768|ALBU_HUMAN	583.9282	1748.753	53.68	LAK(+453.06)TYETTLEK
P02768|ALBU_HUMAN	657.9504	1970.825	48.67	LDELRDEGKASSAK(+453.06)
P02768|ALBU_HUMAN	840.3361	1678.654	48.46	FK(+453.06)DLGEENFK
P02768|ALBU_HUMAN	493.7144	1970.825	48.4	LDELRDEGK(+453.06)ASSAK
P02768|ALBU_HUMAN	503.552	1507.637	48.34	K(+453.06)YLYEIAR
P02768|ALBU_HUMAN	491.5496	1471.627	48.22	AFK(+453.06)AWAVAR
P02768|ALBU_HUMAN	918.3912	2752.154	45.7	NYAEAK(+453.06)DVFLGMFLYEYAR
P02768|ALBU_HUMAN	701.9899	2102.944	43.62	AEFAEVSK(+453.06)LVTDLTK
P02768|ALBU_HUMAN	546.2316	1635.669	41.85	AK(+453.06)TYETTLEK
P02768|ALBU_HUMAN	705.9364	2114.774	39.51	YK(+453.06)AAFTEC(+57.02)C(+57.02)QAADK
P02768|ALBU_HUMAN	467.537	1399.583	39.42	LK(+453.06)C(+57.02)ASLQK
P02768|ALBU_HUMAN	493.7156	1970.825	37.84	LDELR(+453.06)DEGKASSAK
P02768|ALBU_HUMAN	628.2706	1254.527	37.07	NLGK(+453.06)VGSK
P02768|ALBU_HUMAN	698.3344	2091.987	36.76	KVPQVSTPTLVEVSR(+453.06)
P02768|ALBU_HUMAN	449.1718	1344.486	36.58	DEGK(+453.06)ASSAK
P02768|ALBU_HUMAN	467.8701	1400.59	33.25	FK(+453.06)AWAVAR
P02768|ALBU_HUMAN	678.0175	2708.041	32.37	LDELRDEGKASSAK(+453.06)QR(+453.06)
P02768|ALBU_HUMAN	424.1848	1269.527	31.27	ATK(+453.06)EQLK
P02768|ALBU_HUMAN	706.6526	2116.932	31.24	YFYAPELLFFAKR(+453.06)
P02768|ALBU_HUMAN	678.0175	2708.041	28.43	LDELRDEGK(+453.06)ASSAK(+453.06)QR
P02768|ALBU_HUMAN	727.3338	1452.653	28.25	K(+453.06)QTALVELV
P02768|ALBU_HUMAN	784.694	2351.044	26.75	HPYFYAPELLFFAKR(+453.06)
P02768|ALBU_HUMAN	797.8787	1593.743	26.75	K(+453.06)LVAASQAALGL
P02768|ALBU_HUMAN	678.0163	2708.041	26.03	LDELR(+453.06)DEGK(+453.06)ASSAKQR
P02768|ALBU_HUMAN	537.5734	1609.703	24.22	AWAVAR(+453.06)LSQR

6.3.2 BIOINFORMATIC ANALYSIS OF DRUG MODIFIED MHC PEPTIDES

6.3.2.1 PRELIMINARY ANALYSIS

The success of Peaks in correctly annotating flucloxacillin modified tryptic HSA peptides made it an excellent approach for the analysis of MHC peptides. As described throughout this thesis, MHC peptides are inherently difficult to characterize without the further complication of drug modification. For the analysis of MHC peptides, .mgf files were generated using Protein Pilot from individual peptide pools that contained the manually annotated spectra in the previous chapter (Chapter 5). Peaks was able to assign 226 PSMs containing a flucloxacillin modification from a total of 9,154 annotations across the data set. Unfortunately, of the 226 PSMs with flucloxacillin, 91 were assigned the decoy database, making the confidence of these annotations low. The -10logP value for these PSMs ranged from 16.39 to 5.07, with a mean value of 7.11, well below the accepted score of 20. The highest confidence scoring peptide was annotated as ISR(+453.06)VTF and was derived from a protein database (Q9UIQ6|LCAP_HUMAN).

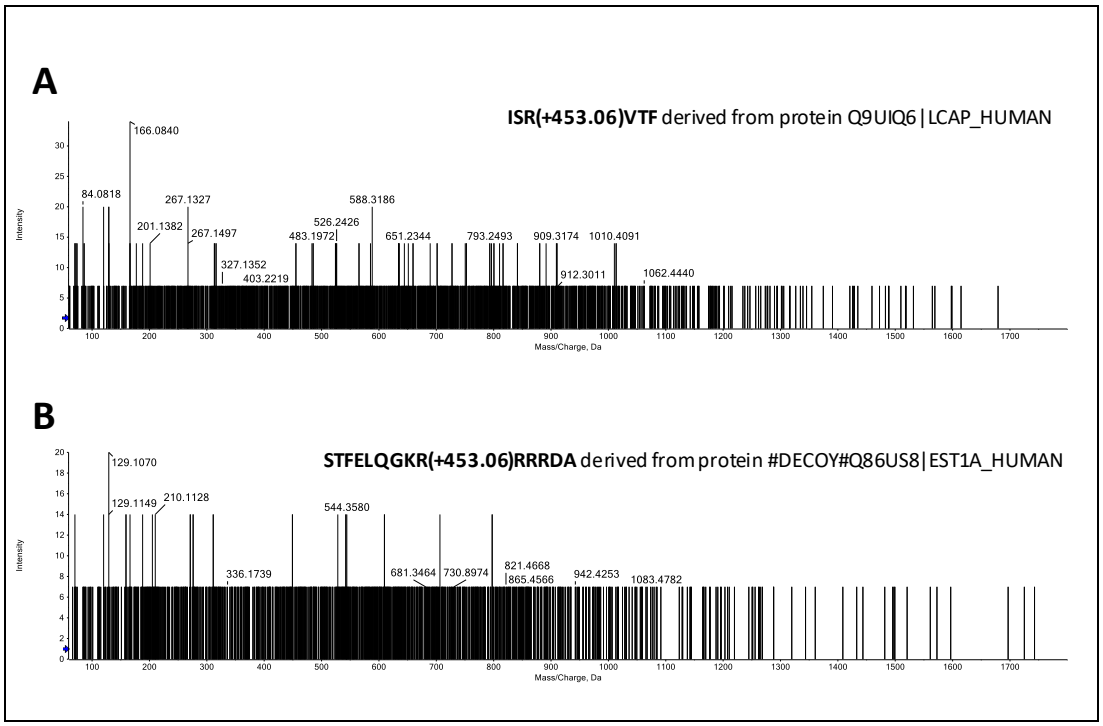


Figure 6.5. Peaks annotation of flucloxacillin modified MHC peptides. (A) ISR(Flucloxacillin)VTF and (B) STFELQGK(Flucloxacillin)RRRRDA were both annotated as potential drug modified MHC peptides.

Interestingly, both the P2 (S) and PΩ (F) amino acids matched those of the HLA-B*57:01 binding motif. Therefore, the original spectrum was identified for manual confirmation of the peptide sequence (Figure 6.5A). Immediately it was possible to tell that the spectrum quality was bad, with a limited number of low intensity fragmentation ions. Perhaps more importantly, there was a complete absence of diagnostic flucloxacillin fragmentation ions. An annotation from the decoy database was also examined further (Figure 6.5B). As expected, the PSM was highly inaccurate. From this data, it was clear that a more robust approach was required for the correct sequence annotation of flucloxacillin modified MHC peptides.

6.3.2.2 PROFILING FLUCLOXACILLIN PARTIAL LOSS & MANIPULATING MS/MS FRAGMENTATION IONS

It was first important to confirm that the dataset used for the flucloxacillin MHC peptide search did in fact contain drug modified peptides. Although in a qualitative sense drug fragment ions could be observed, a more quantitative approach was required. Therefore, Python scripts were developed to identify partial losses between parent ions and dominant fragment ions across all spectra (both MHC and tryptic HSA peptides) containing a 160 Da fragment ion (Table 6.3).

Table 6.3. Partial loss values identified from flucloxacillin modified peptide spectra. Spectra containing the characteristic ion peaks were investigated to identifiy partial loss values from the precursor ion and the base peak (most intense peak in the spectrum). The most common partial losses identified were within 0.1 Da to 453 and 159 Da as anticipated. Other frequent losses require further characterization

Loss Window	Intensity	% of Base Peak	% of Precursor (Naked)	Count
453.0_453.1	317055.06	88.45782212	239.5267395	290
159.0_159.1	412247.48	93.05516632	296.7914107	222
177.0_177.1	81417.30	8.889597465	35.21976414	102
396.0_396.1	22485.20	4.129842479	20.82850262	98
230.0_230.1	50348.71	6.459489601	29.08917993	78
305.1_305.2	10573.54	3.094847451	11.95872658	75
377.1_377.2	119386.09	13.82230818	166.5079222	67
404.1_404.2	15183.07	1.80774128	7.603212835	52
272.1_272.2	23284.49	4.581997833	29.98729171	51

Loss windows were produced from to 0.1 to 454 Da in increments of 0.1 Da. The intensity column is the summed intensity of all the fragment ions that appeared within a specific mass range. To assist with visualisation of the remaining values, the NLGK(flucloxacillin)VGSK tryptic HSA peptide has been used (Figure 6.6). For any given spectra, the base peak corresponds to the highest ion in the spectra (ignoring flucloxacillin fragmentation ions). The loss peak refers to an ion which is a specific mass lower than the BP. For example, in NLGKVGSK the base peak is 869.39 Da. All loss windows are assessed, and an intense fragment ion is present with a mass of 575.36 Da. This mass shift is equal to 294 Da; this loss peak ion intensity is used to record a loss window of 159 Da (453 minus the mass difference) (Figure 6.6, blue). The percantage of BP column refers to the average intensity of the loss peak compared to the BP; 575 Da is roughly 90% of the intensity of the BP. The naked precursor corresponds to the mass of the parent ion, minus 453 Da (flucloxacillin). Here, this equates to 801.47 Da (Figure 6.6, red). The same approach is applied, where the intensity of ions corresponding to loss windows are recorded. A mass addition of 294 Da, 548.76 Da, is recorded with lower intensity. This matches the overall percentage of precursor (naked), as the naked precursor ion is roughly 3x more intense than the modified precursor.

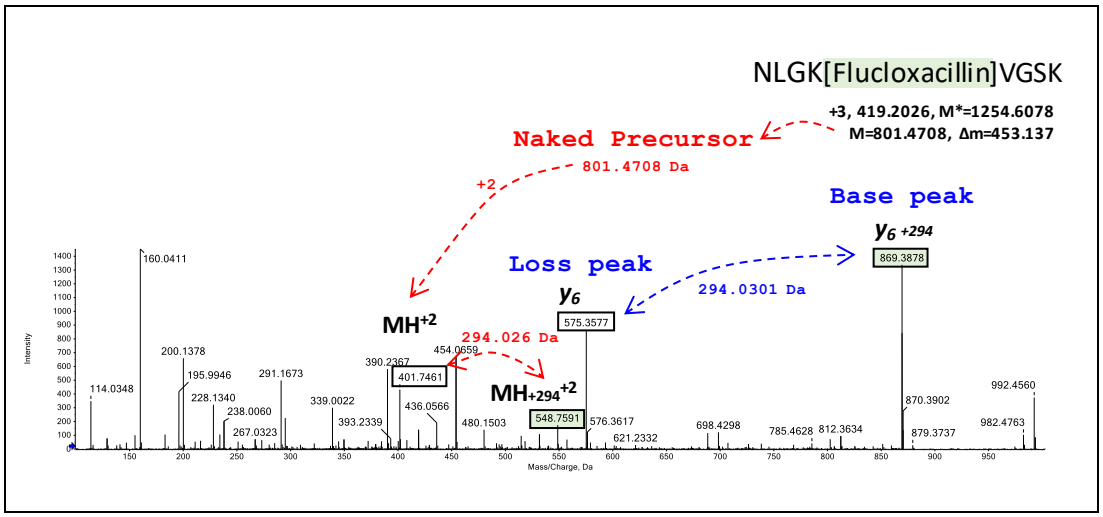


Figure 6.6. Diagramatic explanation of the method employed for the identification of spectral partial losses. The naked precursor is the parent ion MW with the full flucloxacillin mass addition removed (-453 Da). The base peak is the most intense ion in the fragmentation series, with the loss peak corresponding to the ion within the partial loss window being interrogated.

From this data we were confident that the partial loss of 159 Da and the whole drug loss (453 Da) were a true representation of flucloxacillin modification. While other losses were also highlighted in this experiment, they were not as abundant. For example, the loss window of 177.0-177.1 Da was identified, with a mass difference of 276 Da compared with 453 Da. The likely explanation for this is the loss of water from the ion with a 294 Da mass addition. This only further highlights the complexities involved in the identification of flucloxacillin modified MHC peptides. As previously shown, the number of possible fragment ions increased drastically when flucloxacillin is present (Figure 6.2). Considering each of these ions could too have further losses, such as H_2O and NH_3, the number of unknowns and different possible combinations made it unlikely to produce high quality PSMs from a typical database search.

Due to the complexity of the fragment ion series, the next strategy was to simplify the fragment ions though data manipulation. Again, the example NLGK(flucloxacillin)VGSK tryptic HSA peptide has been used to assist with the visualisation of this methodology (Figure 6.7). Contained within the fragment ion series the diagnostic ions 160, 196 and 454 Da can be observed (Figure 6.7A). The first step was to remove these from the spectra as the drug fragment ions should not be used for PSMs (Figure 6.7B). Next, ions pairs that are 294 and 453 Da apart were identified (Figure 6.7B, dotted lines), and the 'heavier' fragment ion intensity combined to the 'lighter' fragment ion intensity (Figure 6.7C). Finally, the precursor mass was reduced by 453 Da. Essentially, the idea was to remove the modification from the fragment ion series, so that it could be run through Peaks as an 'unmodified' MHC peptide. Any positive PSMs would then be annotated to the original spectra with the modification present. Unfortunately, this again did not yield any positive results that could be trusted after manual interpretation was performed.

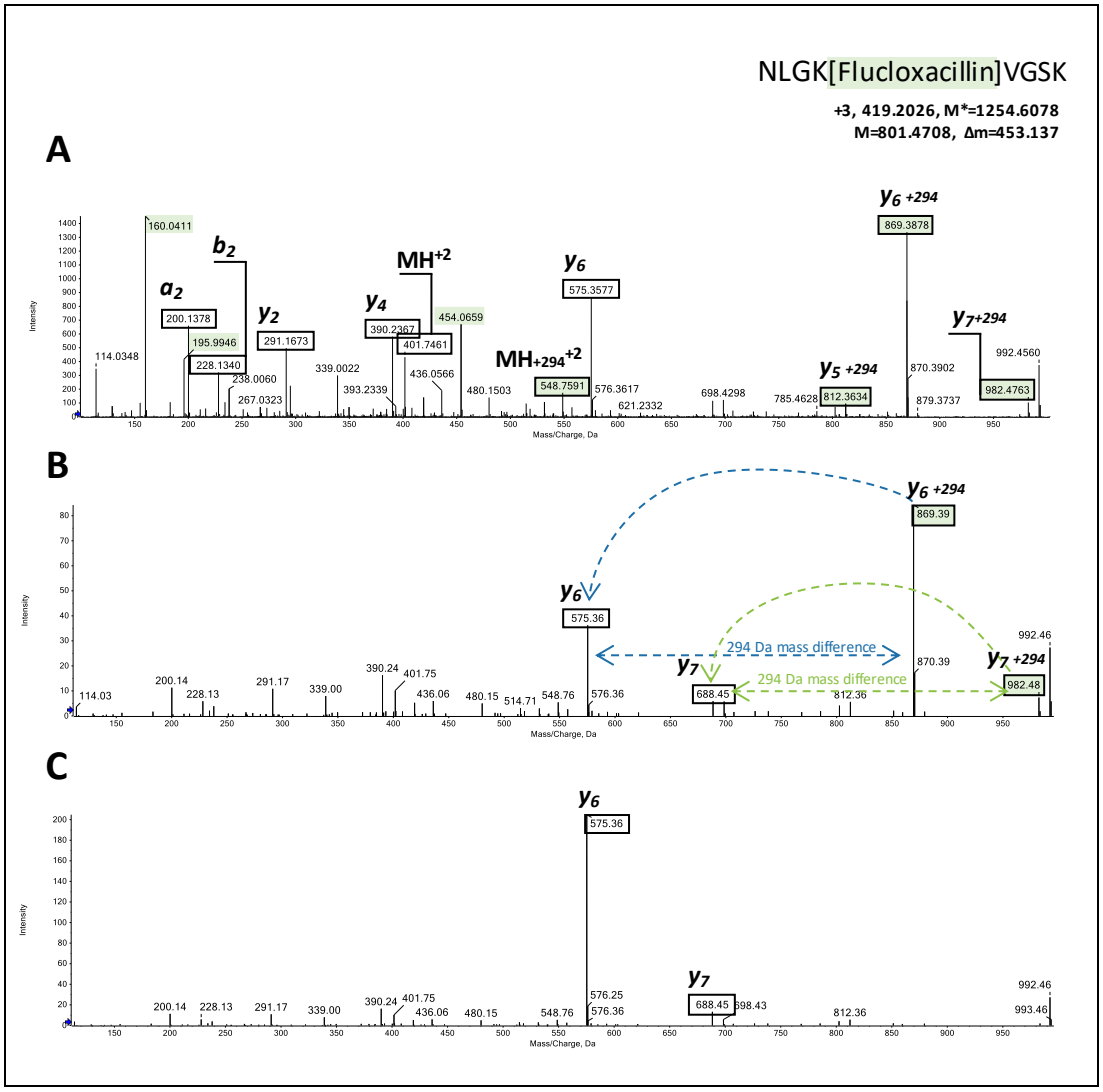


Figure 6.7. Diagramatic representation of spectral maniupluation to remove the flucloxacillin modification from the ion series using HSA tryptic peptide NLGK[Flucloxacillin]VGSK. (A) Manual annotation of drug modified NLGKVGSK. (B) Characteristic flucloxacillin fragment ions have been removed, and ion pairs with a partial loss of 294 Da are identified. (C) Ion pairs are combined to the value of the lower *m/z* (i.e. unmodified ion).

6.3.2.3 CUSTOM PEPTIDE DATABASE

It was clear that the combination of complex fragmentation ion series and a database of almost infinite possible theoretical peptides was not conducive to gaining positive sequence annotation. Therefore, attention was turned to the database used for PSM assignments. In this study the anchor residues of HLA-B*57:01 were used to assist with spectral matching. Instead of using a protein database for PSMs, the use of a pre-compiled peptide database was used to limit the search space. Indeed, this could be easily generated from a list of HLA-

B*57:01 peptides eluted from untreated samples. However, this would not enable any peptides not previously seen to be incorporated into the database.

As shown in previous chapters, the binding motif of HLA-B*57:01 is relatively conserved. However, changes are observed at the PΩ anchor when flucloxacillin is applied. Luckily, this alteration does not introduce novel amino acid anchor residues on unmodified peptides, therefore the amino acids contributing to the binding motif remain the same. Secondly, flucloxacillin undergoes nucleophilic attack, leading to the opening of the β-lactam ring and subsequent covalent binding. Therefore, for a peptide to be modified it must contain either a lysine or arginine residue. Finally, from previous elution data it is well documented that MHC I peptides derived from HLA-B*57:01 are by majority 9-11 amino acids in length.

Using these parameters, a new database was generated from the UniProt *Homo sapiens* proteome. Peptides were generated that were 9-11 amino acids in length, had relevant amino acids at the P2 (S, T & A) and PΩ (F, W & Y) anchor and contained a lysine or arginine in P3 to PΩ(-1). While PSMs would not be made to peptides that did not meet these criteria, it would assist with generating a smaller search space to examine covalent binding of flucloxacillin to HLA-B*57:01 peptides. As the algorithm behind scoring in Peaks could not be determined, MS Amanda was used as the search engine. One of the main benefits of using MS Amanda is due to all possible fragment ion matches being reported for each spectrum, regardless of PSM scoring. During the search, only spectra which contained the diagnostic fragmentation ions of flucloxacillin were included.

Multiple PSMs were made to the peptide database, with some peptides seeming plausible upon manual assessment. Unfortunately, in the current release, MS Amanda does not indicate which fragmentation ions were used for PSM annotation. One such example is YTK(flucloxacillin)VSATLNW (Figure 6.8A, i). The majority of the high intensity ions do correspond to the annotated peptide sequence, however during acquisition this was a low

abundant peptide (Figure 6.8A, ii) with a questionable charge state (Figure 6.8A, iii). One of the highest scoring PSMs related to TSNSGDPHK(flucloxacillin)GF (Figure 6.8B).

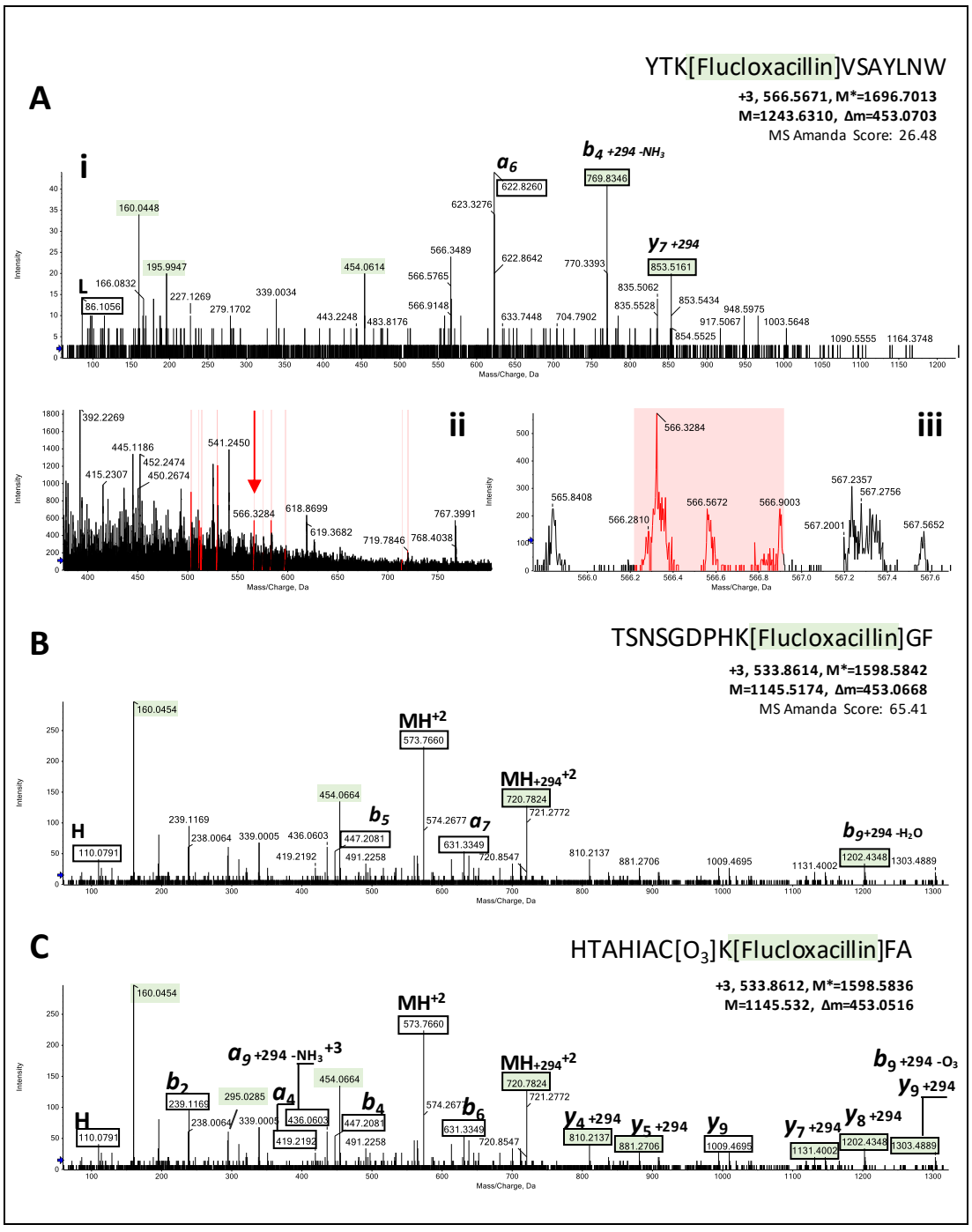


Figure 6.8. MS Amanda annotation of flucloxacillin modified MHC peptides using a peptide database library. (Ai) MS Amanda scored relatively high for the annotation of the fragment ion series as (i) YTK[Flucloxacillin]VSAYLNW, however, (ii) the precursor ion abundance was low and (iii) the charge state questionable. (B) A second peptide annotation with the higest score from the search, however many ions are not accounted for. (C) The same spectrum as B was previously manually annotated as a completely different peptide sequence, however HTAHIACKFA would not have been in the peptide library generated.

The spectrum used in the PSM assignment is of much higher quality, with a large proportion of the ions annotated. However, this is the same spectrum manually annotated as HTAHIACK(flucloxacillin)FA (Figure 6.8C) (Chapter 5). From comparing the two annotations, it's clear that the manual interpretation accounts for a much larger proportion of the fragment ions. Due to the peptide database, HTAHIACKFA would not exist as the PΩ amino acid is outside of the specified criteria. This is important when considering a repertoire change was observed with flucloxacillin and abacavir, as peptides with novel anchor residues would not be contained in the search database.

6.3.3 Using R-script to calculate confidently assigned HLA-B*57:01 peptides

Processing of immunopeptidomic data post acquisition is challenging due to the volume of information generated. In this thesis, 12 different mass spectrometry (.wiff) files were simultaneously run in Protein Pilot in order to identify PSMs. As described in this chapter, the ability to detect flucloxacillin modified peptides was severely limited. However, unmodified peptides can still be confidently annotated using this method. Within each Protein Pilot output file, a FDR is generated. Subsequently, this was used to apply a cut-off for peptide confidence scores. As described in chapter 4, a series of further steps are taken to ensure the authenticity of the data. This results in the generation of multiple spreadsheets in Microsoft Excel, where data is copied, pasted and further interrogated. This can unintentionally lead to errors due to the sheer volume of data being analyzed at any one time. In order to increase both the efficiency and robustness of this method, scripting using R language can be advantageous. Here a Protein Pilot output file was loaded into the script, leading to automatic data analysis being performed through a series of steps (Figure 6.9). One of the many benefits of using R language is the community support (https://community.rstudio.com/). As is it an open source language, R is based largely on peer to peer assistance, with several forums available for the more novice users to seek

support. Experienced users can develop packages that can be installed on other users' machines. This ability to share code packages is particularly useful in scientific research, as they act as a grounding basis for many forms of analysis.

Understandably, the time taken to generate the script is dependent on the user, however in all cases once written can be saved and reused. The steps in the script (Figure 6.9) are the same as those outlined in chapter 4. The first part of the code is to load all the relevant packages already written by other users. Once loaded, the data file generated from Protein Pilot is loaded (highlighted black). Next, the FDR value is extracted from the Protein Pilot data file, to enable the correct cut off to be used for different outputs. As with the previous analysis using Excel, known contaminants are removed through loading a pre-compiled list (highlighted blue) as well as the binding motif for HLA-C*04:01 (highlighted green). The remaining lines of code perform the analysis leaving a list of confidently assigned PSMs, which is exported as an Excel compatible file (.csv).

```
# LOAD PACKAGES ------------------------------------------------------------------------------
library(plyr)
library(tidyverse)
library(gplots)
library(ggseqlogo)
library(dplyr)
library(stringr)
library(tidyr)
library(stringi)
library(readxl)
# IMPORT DATA --------------------------------------------------------------------------------
# import data using readr package using txt file from proteinpilot delimited tab separated data
fdrfile <- "PE_027_JCW_C1RB5701_Control_HomoSapien__FDR"
elution_dat <-
  readxl::read_xlsx(paste(fdrfile, "xlsx", sep = "."), sheet = "Distinct Peptide Summary")
fdr_report <-
  readxl::read_xlsx(paste(fdrfile, "xlsx", sep = "."), sheet = "Single Column Summary", range =
"D28:E46")
fdr_report$confidence <- (fdr_report[,2]*100)
FDR5 <- as.character(fdr_report[8,3])
contaminants dat <-
  readr::read csv("C1RB5701 contaminants.csv")
cw4anchor_dat <-
  readr::read_csv("HLAC0401 motif.csv")
# Look at the data to ensure it has loaded correctly
glimpse(elution_dat)
# MUTATE DATA & ADD RELEVANT VARIABLES -------------------------------------------------------
# Mutate the data to make a new column of the lengths of each peptide
elution_dat <- elution_dat %>%
  mutate(Lengths = str_count(Sequence))
# Merge sequence and modifications variables
elution_dat$Unique <- paste(elution_dat$Sequence, elution_dat$Modifications)
# Create new column of sequences which are reversed for easy analysis
elution dat$seqreversed <- stri reverse(elution dat$Sequence)
# Create new column to assess contaminants in the sample and compare using the match function
(in%in) to the contaminants list
elution_dat["Contaminants"]<-NA
elution_dat$Contaminants <- elution_dat$Sequence %in% contaminants_dat$Sequence
# Create new column for p2 and p-omega anchors and columns for their match to Cw4 binding motif
elution_dat["P2"]<-NA
elution dat["PO"]<-NA
elution_dat["Cw4P2Match"]<-NA
elution_dat["Cw4POMatch"]<-NA
# Uses the substring function to select the P2 and Pomega positions of each peptide
elution_dat$P2<-substr(elution_dat$Sequence, 2, 2)
elution dat$PO<-substr(elution dat$seqreversed, 1, 1)
# Uses the substring function to select just the first locus for each peptide
elution dat["Locus"]<-NA
elution_dat["Locreversed"]<-NA
as.character(elution_dat$`Peptide Locus`)
elution_dat$Locreversed <- stri_reverse(elution_dat$`Peptide Locus`)
elution_dat$Locus<- substr(elution_dat$Locreversed, 1, 1)
# Uses the match function to asess matches to the Cw4 anchors
elution_dat$Cw4P2Match <- elution_dat$P2 %in% cw4anchor_dat$P2
elution_dat$Cw4POMatch <- elution_dat$PO %in% cw4anchor_dat$PO
# FILTERING THE DATA -------------------------------------------------------------------------
# Begin by filtering for peptides of FDR confidence interval
elution_dat_processed <-
  elution dat %>% filter(`Best Hypoth Conf`>=FDR5)
# Remove decoy database results using stringr_detect function ! means not
# https://cran.r-project.org/web/packages/stringr/vignettes/stringr.html
elution_dat_processed <-
  elution_dat_processed %>% filter(!str_detect(Names,'REVERSED'))
elution_dat_processed <-
  elution_dat_processed %>% filter(!str_detect(Names,'Keratin'))
# Remove any duplicate peptides using distinct function and keep all other variables using
.keep_all = TRUE
# https://www.rdocumentation.org/packages/dplyr/versions/0.7.7/topics/distinct
elution_dat_processed <-
  elution_dat_processed %>% distinct(Sequence, .keep_all = TRUE)
# Remove other loci
elution dat processed <-
  elution dat processed %>% filter(str_detect(Locus,"1"))
# Remove contaminants
elution_dat_processed <-
  elution_dat_processed %>% filter(!str_detect(Contaminants,'TRUE'))
# Remove matches to the Cw4 binding motif
elution_dat_processed <-
  elution_dat_processed %>% filter(!str_detect(Cw4P2Match, 'TRUE'))
# Create a file from the processed data
               write.csv(elution_dat_processed, paste(fdrfile, "csv", sep = "."))
```

Figure 6.9. R-Script used for the automatic assessment of HLA-B*57:01 peptides within a 5% FDR. Protein Pilot v.5 FDR output files are loaded into R-Studio. Known contaminants and the HLA-C*04:01 binding motif are loaded as a separate file. Script annotatations are denoted by '#'.

6.4 DISCUSSION

For this chapter, the aims were to use bioinformatic tools to enable the analysis of MHC peptide data to be managed in a more robust and higher throughput capacity. While manual interpretation of fragmentation ion series has enabled the characterization of drug modified peptides, this is time consuming and requires a high level of skill (and patience!). The challenges associated with MHC peptide annotations using software are further exasperated when complex drug binding is added into the mix. Flucloxacillin fragmentation leads to a different mass addition on the parent peptide compared with the fragmentation ion series. While this is helpful for the identification of flucloxacillin modification on a peptide, it severely hampers the ability to sequence spectra with any confidence. While there are a whole host of software packages available for the annotation of spectra, it was that none of those tested were capable in producing the required outputs. When selecting the algorithm to use, it's important to consider the pros and cons to each, and the questions you are asking. For example, it was found that MASCOT and Peaks are excellent at identifying flucloxacillin modification using tryptic peptide searches. Ultimately, applying trypsin significantly reduces the number of theoretical peptides that could be present in a particular set of proteins, therefore confidence in PSM assignment is much higher.

The bioinformatics analysis methods generated tried to approach this challenge from both directions. Firstly, the spectra were manipulated in order to 'remove' the modification. Due to the complex nature of the drug fragmentation, it was not always possible to identify which ions were paired in terms of 'modified' and 'unmodified' siblings. Through assessing the difference in mass between fragment ions, sibling pairs were combined to give a single 'unmodified' version of the fragmentation ions. While in principle this approach seemed sensible, in practice very few results were yielded with low confidence scores. Furthermore, upon manual interpretation, the results were found to be meaningless. Several factors may have impeded its success. Firstly, Peaks software is very intelligent and is based largely upon

machine learning of real spectral information. Manipulation of real spectra could have been identified as 'strange looking' to the algorithm, and so PSMs would not be scored with high confidence. As the algorithm behind Peaks is unknown (proprietary information) this is purely based on assumption. A more likely explanation is the mass differences (294 and 453 Da) between ion pairs were not specific to flucloxacillin fragmentation. In fact, 294 Da can also correspond to M(+16)M(+16), LT(+80), MY, FM(+16) and FF. Indeed, the combinations that add up to 453 will be higher. If this is the case, unrelated fragment ions may be selected as pairs and subsequently merged together.

The second approach was focussed on refining the database used for the PSM annotations, in order to reduce the search space. By having fewer possible options, confidence scores increase. A peptide database was generated using a set of rules, which meant the peptides must fit the HLA-B*57:01 binding motif (anchor residues and length) while containing a lysine or arginine. While this approach does seem to give some more reliable PSMs than previous methods, it is still limited by the number of peptides in the database. As we believe flucloxacillin may interrupt antigen processing, this approach too has its flaws. Interestingly, the peptide spectra manually annotated as HSATQKEHGW in chapter 5 did not return any possible hits. While it is hard to definitively hypothesise why, the manual annotation cannot be derived from any protein in the human database, therefore it was determined to be a spliced peptide. As this database is reliant on peptides derived from human proteins, any spliced variants would too be missed.

Finally, R studio was used for the generation of scripts to automatically analyze MHC peptide data for unmodified peptides. While the analysis method underpinning the script did the same as described in chapter 4, there were clear advantages. Using this method results in a reduction of time spent analysing the data from hours to seconds, with confidence that the analysis is performed correctly. While this script only generates a list of confidently assigned MHC peptide sequence, others performing peptide length and anchor residue abundance

are further incorporated into the analysis pipeline. While for this thesis the data interpretation using Excel was manageable, considerations in using languages such are R are essential for a higher throughput approach.

Overall, the collaboration with data scientists from the Centre for Computational Biology further confirmed the challenges associated with the detection of drug modified MHC peptides. Although progress was made, further work needs to be done in order to truly believe any of the results obtained from the searches. The script using R studio was beneficial to all members involved in MHC peptide research within the group and is now being routinely used for the initial analysis of MHC mass spectrometric data. Further scripts are now being developed to further the data interrogation using R studio.

6.5 APPENDIX 1

Simon Perkins[1], Andy Jones[1]*

(* corresponding author: andrew.jones@liverpool.ac.uk)

[1] Computational Biology Facility, University of Liverpool L69 7ZB

CBF

COMPUTATIONAL BIOLOGY FACILITY

6.5.1 CONTEXT

The project involved identifying peptide sequences that are covalently bound to the antibiotic drug Flucloxacillin.

The short peptide sequences are as a result of larger proteins that are non-selectively cleaved by the immune system apparatus. It is believed that Flucloxacillin 'interacts' with host proteins, which causes an errant recognition and immune reaction. Identifying sequences that this drug, and other drugs, may bind to is of potential significance for future research studies and clinical decision making.

Using a mass spectrometry (MS) approach, peptides can be identified that may have a drug (or any modification) bound to them. This is made easier if the drug behaves predictably under peptide ionisation and fragmentation, however this is not always the case. Additionally identifying a small subset of peptides modified with a drug is easier if the traits of potential peptide sequences is known beforehand.

6.5.1.1 THE DRUG

In the case of Flucloxacillin, intact peptides with the drug bound are thought to frequently undergo neutral losses under experimental conditions. A neutral loss involves the partial or complete fragmentation of the chemical groups attached to the peptide, meaning that peptide fragments identified which would be expected to include the mass of the modification (drug in this case), have an altered mass. Preliminary analysis of spectra generated from Flucloxacillin- modified peptides showed that the drug can undergo several different types of neutral loss (different mass losses), in a somewhat unpredictable manner. This means that it is very difficult to identify peptides with the drug bound as the full mass of the drug addition is very infrequently seen. There are some 'known' losses which are thought to occur but their detection is difficult, so identifying peptides based on an MS1 m/z value is difficult, as is identifying fragment ions which may be linked to a neutral loss.

If a peptide can be identified with the drug bound, then localisation is canonically the next relevant issue, and it is very important to understand the behaviour of the drug under fragmentation conditions. If this is not fully understood, then it can be very difficult to localise a modification to an amino acid residue – this also applies to identifying a peptide from the MS2 fragments.

6.5.1.2 THE PEPTIDE

The pattern of sequences of peptides believed to bind to Flucloxacillin is not concretely known. With a typical proteomics shotgun/bottom up approach, intact proteins are digested under the protease enzyme trypsin, which is known to cut protein sequence after lysine (K) or arginine (R). Thus the subsequent matching of spectra to peptides will only consider peptide sequences that could possibly be generated with such an enzyme specificity. This peptide universe is much smaller than the universe of all potential peptides in the human proteome.

The peptide sequences that are generated as part of the immune system apparatus may not have such a recognised pattern, as tryptic peptide sequences do, so usually when analysing MHC presented peptides, search engines are asked to consider all possible human peptides of the appropriate mass. This is a very large population of peptide sequences, and suffers from an enormous loss of statistical power, particularly when using traditional database

search engines such as Mascot. “Hybrid” search engines, such as PeaksDB, which perform a *de novo* sequencing step before sequence database search can partially avoid the enormous loss in statistical power. However, the additional problem of the varied and complex neutral losses observed for Flucloxacillin makes for a second challenge, yet further increasing the search space. As noted below, while PeaksDB claims to offer support for customising neutral losses for a given modification, in practice this appears not to work well for reasons that cannot easily be further investigated. We thus finally decided on a method that controls the search space for database size by an alternative approach (selecting plausible MHC peptides based on known motifs for a given MHC molecule), and using an easily customizable free search engine, called MS Amanda.

6.5.2 Methods & Results

6.5.2.1 Flucloxacillin Neutral Losses

The Flucloxacillin drug molecule is thought to be frequently broken in two separate places in its structure under ionisation and fragmentation conditions.

The first break results in a loss of 159.0354, known as a thiazolidine ring loss. This is thought to be the most common neutral loss observed, and we have observed the drug and drug bound to peptide with this mass lost from them. We also frequently see the non-neutral (charged) version of this lost molecule, i.e. charged with a single proton, with a mass of roughly ~160 Da.

The second potential break results in a loss of 194.9887, known as a two ring loss. This seems less common than the thiazolidine ring loss – this may be anecdotal though as the charged ~160 Da ion is the prime piece of evidence for the thiazolidine loss. A loss that is always lost neutrally will never be visible in spectra.

6.5.2.2 Preliminary Analysis

Data was initially provided that contained peptides derived from a trypsin digest of Flucloxacillin bound to proteins.

This tryptic peptide data was searched in the Mascot search engine with the Flucloxacillin modification (complete with the two standard neutral losses as described above). Specifying ‘Trypsin’ as the enzyme constrains the universe of potential peptides. This initial test was a success as a handful of Flucloxacillin modified peptides were visible in the search results.

However, when attempting to identify spectra that were not a result of a tryptic digest, but were instead a result of immune system mechanisms and MHC interaction, we had difficulty. No Flucloxacillin modified peptides were identified with acceptable scores, and there were many false positive peptide identifications.

This meant that we needed to re-evaluate our search strategy and possibly our understanding of the Flucloxacillin modification.

6.5.2.3 Profiling Neutral Losses

To help understand and clarify the potential losses of/from Flucloxacillin, we used a python script to look at the gaps between peptide precursor mass and every peak in each spectra of

our data thought to contain real peptide-drug combinations. We placed the absolute difference between peaks into 'bins', each 0.1 Da wide, to account for slight differences in mass. We would then inspect the bins with the highest frequency in our datasets as candidates for being Flucloxacillin-derived losses.

The ~159 Da loss always seemed to be placed as the most frequent loss observed, and the ~453 Da loss was similarly highly placed, however there were a number of very frequent losses for which we have no explanation. Further work would be necessary to identify the composition of these, if they are relevant. A python script is included as a deliverable to help automate this step.

6.5.2.4 PEAKS DB

We explored an alternative approach, using the benefits of PeaksDB (hybrid search methodology) to identify Flucloxacillin-modified peptides. First, we demonstrated on tryptic digests for HSA-modified samples that by switching on and off the requirement for peptides to be tryptic, there is almost no loss of statistical power (unlike in Mascot, where there is a very large loss of sensitivity). We next tried searching datasets in which a small proportion of genuine MHC presented peptides had been modified with Flucloxacillin, by specifying the Flucloxacillin modification in Peaks with neutral loss settings as discussed above. Results from this process were poor, with no plausible high scoring Flucloxacillin peptides identified. Further investigation (with manually created spectra of various types) revealed that Peaks appears to have a wider problem with custom modifications and neutral losses, in that peptide scores seem to be strongly down-weighted versus peptides with no modifications. This unusual behaviour was reported back to the Peaks customer support team, but they could not offer a solution.

A further approach was attempted in Peaks. Given that we had observed Peaks performance for general MHC peptide identification (unmodified) is very good, we attempted to apply a transform to candidate modified spectra to make them appear to Peaks as if they were unmodified spectra. In this pipeline, we first identified candidate Flucloxacillin-modified spectra via presence of a characteristic ion at 454 Daltons. These spectra were then "mass shifted", so that measured precursor mass ion mass (prior to fragmentation) was altered by -454 Daltons (the mass of Flucloxacillin). Additional fragment ion peaks were also inserted into the spectrum, based on deducting likely neutral losses from potentially modified fragment ion peaks. On test data, this approach appeared promising, and produced high-scoring peptide-spectrum matches (PSMs) from Peaks DB. However, performance on real data was again poor for reasons not well understood, so this approach was also discarded.

6.5.2.5 CUSTOM PEPTIDE DATABASE

Since it was believed that the universe of all human peptides would be too large a search space for the search engines to handle, we decided to create our own peptide sequence databases, based on some of the rules related to MHC peptide presentation. It is thought that peptides that bind to the MHC molecule presenting peptides in this sample (B*57:01) must contain a lysine (K), terminate in phenylalanine (F), tryptophan (W) or leucine (L), and should have serine (S), threonine (T) or alanine (A) at position 2 in the peptide sequence. The length of matching peptides should also be constrained between 9 and 11 amino acid residues. An example of a peptide matching such criteria would be 'LTHGKDEPL'. We generated a sequence database of potential MHC and Flucloxacillin binding peptides, using Uniprot human sequences as a starting point.

We also created equivalent databases where potential peptides could additionally terminate with tyrosine (Y), methionine (M) or isoleucine (I), and also where the peptide sequence could be required to contain an arginine (R) instead of being required to contain a lysine (K) residue.

6.5.2.5.1 MASCOT

We initially used the Mascot search engine to search our spectra against our custom peptide databases. However, Mascot does not offer complete and flexible control on what results it will report, including low scoring peptides. As such, we followed some alternative approaches.

6.5.2.5.2 CUSTOM SEARCH ENGINE

To validate our custom peptide database approach and allow us to see all peptide-spectrum hits (PSMs), we constructed a very simple MS1-only search engine in Python. This was done to be able to find all peptide hits within a specific tolerance, and then be able to inspect these results and understand how we could improve our process.

6.5.2.5.3 MS AMANDA

The MS Amanda[1] approach builds on our custom search engine, however it also performs MS2 matching, which is crucial for Flucloxacillin localisation on the peptide. The MS2 matching also gives us a probability or score for each peptide spectrum match (PSM) which we can use for thresholding and false discovery rate estimation. Like our custom search engine, MS Amanda will give us a full set of matching results within the criteria that we set. It supersedes the custom search engine we made, and should be the preferred way of searching for Flucloxacillin modified peptides.

The Flucloxacillin modification needs to be specified in the unimod XML file that accompanies MS Amanda (and describes modifications). The specification also includes both of the common neutral losses for Flucloxacillin.

While MS Amanda supports MS2 ion matching, it does not give detailed information of MS2 ions that match for a PSM. This information is included when MS Amanda is run via Proteome Discoverer, however the standalone version of MS Amanda used here did not support detailed MS2 ion matching information at the time of analysis.

We think this MS Amanda approach is the best here. When we apply a 5% FDR threshold (as described below) we are able to see a handful of plausible Flucloxacillin-modified peptide spectrum hits. We see fewer hits with a 1% FDR threshold, as expected. Choosing a lower FDR threshold would be optimal in theory, however we know that finding flucloxicilin-modified peptides remain challenging, due to neutral losses, and they may score more weakly than regular peptides, thus a 5% FDR threshold seems acceptable here, so long as it is understood that some of the hits may be false positives.

MS Amanda is available as part of ProtomeDiscoverer and SearchGUI software packages, however the need to specify the custom Flucloxacillin modification in the unimod XML file may preclude the use of MS Amanda through these avenues.

6.5.2.6 False Discovery Rate (and Q-value) Calculation

The output from MS Amanda searches is a tabular format which can be read by Microsoft Excel and similar packages. For the purpose of quality control, MS Amanda does not give any information about false discovery rate. FDR can be calculated from the MS Amanda scores if decoy (known false peptides) are included in the search. We created a python script to generate FDR values (and Q values) for every peptide-spectrum match, add these extra columns to the tabular output, and generate a new file.

6.5.2.7 Shiny Web Application

For browsing of the MS Amanda result data we developed an R/Shiny web application[2]. This mainly allows the user to filter results to only show Flucloxacillin-modified peptides, and to apply an FDR threshold to exclude low quality hits.

6.5.2.8 Data Files

Included with this report are MS Amanda result files (including FDR and Q-values) with the following permutations:

- Control (not expected to be modified), manually curated (annotated as being modified) or (potentially) Flucloxacillin modified.
- 0.1 or 0.04 Da MS2 matching tolerance.
- Searches including neutral loss searching or not.

6.5.3 Deliverables

This report is one of the deliverables for this project.

6.5.3.1 Python Scripts

6.5.3.1.1 Generate MHC peptides.

This script allows the user to generate candidate peptides for drug binding, given a protein sequence database, and a pattern for peptide sequence to match.

6.5.3.1.2 Find neutral losses in spectra.

This script allows the user to profile any potential neutral losses occurring in spectra. It does this by calculating the size of the 'gaps' between peptide precursor and other peaks and reporting those gaps which are the most frequent.

6.5.3.1.3 FDR calculator for MS Amanda.

This script allows the user to calculate the false discovery rate and q-value for a particular MS Amanda hit. This required that a 'decoy' search has been performed, i.e. there are known false positive hits interleaved with the true positive hits.

6.5.4 GLOSSARY

6.5.4.1 NEUTRAL LOSS

A ‘loss’ is a drop in recorded m/z or Da of a peptide ion or peptide fragment ion, when compared to an expected mass, due to the partial or complete fragmentation of particular chemical group.

6.5.5 MS AMANDA USAGE

MS Amanda may be used either as part of ProteomeDiscoverer, as part of SearchGUI, or as a standalone console version. If custom modifications may be defined as part of PD or SG, then non-neutral loss searching will be possible through these softwares, with an appropriate protein/peptide database. The Flucloxacillin modification would be defined with its full mass.

If neutral loss searching is desired, it is anticipated that the standalone console version of MS Amanda will have to be used. This is because the neutral loss values have to be configured as part of the modifications in the Unimod XML file for MS Amanda. The Unimod XML file us included with this report. It can be used for Flucloxacillin searching with MS Amanda (or potentially other search engines that support the unimod format), could be modified to include further information about Flucloxacillin fragmentation, and can form the basis of a specification for another drug modification, different to Flucloxacillin. The specification of Flucloxacillin begins at line 49 and ends at line 99 of the XML file.

Additionally included is the template for specifying the search settings for MS Amanda – the ‘settings_prototype.xml’ file. This was used for automating searched with Python code. The parameters values with capital letters were substituted automatically with appropriate values programmatically. This is included for completeness – the MS Amanda user guide and support should be used for further information. MS Amanda standalone versions may be downloaded from their website[3].

6.5.6 REFERENCES

1. Dorfer, Viktoria et al. “MS Amanda, a universal identification algorithm optimized for high accuracy tandem mass spectra.” Journal of proteome research vol. 13,8 (2014): 3679-84. doi:10.1021/pr500202e
2. http://pgb.liv.ac.uk/shiny/simon/drug-peptide/
3. http://ms.imp.ac.at/?goto=msamanda

CHAPTER 7 - FINAL DISCUSSION

Early studies attempting to elucidate the mechanisms attributing to the development of idiosyncratic drug hypersensitivity reactions believed the 'combination of compounds with protein' leads to sensitization in animal models (Landsteiner and Jacobs, 1935). In a separate study, formalinized proteins were capable of sensitizing rabbits to formaldehyde (Horsfall Frank L., 1934), further confirming this hypothesis. Indeed, in more recent years accumulative evidence has shown that covalent drug binding plays an important role in ADRs (Ju and Uetrecht, 2002; Pichler, 2003; Singh *et al.*, 2011). In the case of immune mediated ADRs, T cell activation through covalent binding is observed (El-Ghaiesh *et al.*, 2012; Monshi *et al.*, 2013; Yaseen *et al.*, 2015; Meng *et al.*, 2017). Attempts to calculate the critical binding threshold for the activation of T lymphocytes has been made using piperacillin, a β-lactam antibiotic. The level of modification on Lys-541 in HSA was found to be between 2.6-4.8% in tolerant and hypersensitive patients. When HSA from the incubation media of activated T cell was characterized, 2.6% modification of Lys-541 was observed. Importantly, this indicated that the threshold level of drug antigen required for T cell activation is formed in all patients (Meng *et al.*, 2017). While this drug-protein adduct may provide signal 1 to T cells, current theory suggests that tolerance will develop without further co-stimulatory activity or disruption of immune regulation (Yun *et al.*, 2016). In immediate hypersensitivity reactions, IgE and IgG antibodies reactive to the side chain of β-lactam antibodies have been detected in allergic patients (Harle and Baldo, 1990; Torres *et al.*, 1997). However, despite decades of intensive research, the precise antigens required for T cell mediated immune activation have not yet been defined.

The association of covalent drug-protein binding and ADRs has a significant input into drug development, both with respect to pharmacology and toxicology. While covalent drug binding is involved in toxicological effects, the high potency and prolonged effects that can

be achieved by covalent binding can also be used to enhance drug efficacy (Wilson *et al.*, 2013). In turn, this can reduce the frequency of dosing and increases therapeutic margins; offering benefit to the patients (Bauer, 2015). In the discovery of some novel therapeutic compounds, reactive functional groups are designed to form a covalent bond with their target, chemically inhibiting protein activity. Currently, it is estimated that up to 30% of all marketed drugs targeting enzymes contain reactive war-heads, making them an important tool in the treatment of disease (Wen *et al.*, 2019). The requirement for pharmaceutical companies to screen millions of compounds for covalent activity comes at significant cost to perform high throughput experimental analysis. To improve drug efficacy a current approach in narrowing down the number of lead compounds comes from virtual screening. Molecular docking is particularly useful when the 3D crystal structure of the target protein is available (Wang and Zhu, 2016). While this is valuable in envisaging covalent binding with the protein target, off target toxicity presently cannot be predicted.

At the start of this thesis, a number of aims were set out to investigate the role of flucloxacillin protein binding, and the implications in the onset of ADRs. A high titre flucloxacillin specific antibody was successfully generated, with minimal cross reactivity with other β-lactam antibiotics (piperacillin, amoxicillin, benzyl penicillin and penicillin V). Cross reactivity to oxacillin, cloxacillin and dicloxacillin was identified, indicating that the epitope for recognition is the isoxazole ring. Importantly, antibody could be selectively inhibited through pre incubation with flucloxacillin modified N-acetyl lysine. A number of techniques were used to determine flucloxacillin binding in cellular systems. Relevant cell lines such as antigen presenting C1R-B*57:01 cells, and liver like HepG2 and HepaRG cells, were investigated in addition to primary human hepatocytes. In all instances, SDS-PAGE revealed no observable change to protein abundance when cells were incubated with flucloxacillin. However, through the use of the flucloxacillin-specific antibody, Western blot was utilised to reveal a number of proteins irreversibly modified.

The pharmacological target for flucloxacillin, along with other β-lactam antibiotics, is the penicillin binding protein (PBP). As analogues of the D-alanyl-D-alanine peptide subunits involved in the synthesis of the bacterial peptidoglycan layer, β-lactams bind to the Ser403 residue of the PBP active site. Through this covalent acyl-enzyme complex with the nucleophilic serine, β-lactam antibiotics inhibit PBP activity, halting bacterial cell wall synthesis (Kelly *et al.*, 1989; Lovering *et al.*, 2012). Mass spectrometric analysis of C1R-B*57:01 cells incubated with flucloxacillin identified the master regulator of MAPK signalling, 14-3-3, to be selectively modified across different isoforms. Furthermore, flucloxacillin was identified to preferentially bind to Lys-53 in P38α, another master regulator of MAPK pathways. While the function of this binding is not known, off target protein binding can undoubtedly impair physiological processes and signalling pathways.

Immunofluorescence microscopy was used to localise flucloxacillin binding in HepG2, HepaRG and C1R-B*57:01 cells. Extracellular modification was largely seen on HepG2 cells, and is believed to be due to reduced expression of influx membrane transporters. Secondly PXR expression, responsible for maintaining bile acid homeostasis, is high in HepG2 cells. As flucloxacillin has been identified as a potential PXR ligand, this may explain rapid clearance of flucloxacillin upon entering the cell (Andrews *et al.*, 2010). Localisation within biliary canaliculi of HepaRG cells was identified, and was weakly dependant on MRP2 and P-gp activity. Interestingly, MRP2/P-gp activity was found to increase with prolonged exposure to flucloxacillin treatment. However, CMFDA accumulation, used as a marker for MRP2/P-gp activity, may have increased due to the observed dilation of the bile canaliculi. Both intracellular and extracellular protein modification was observed in the antigen presenting C1R-B*57:01 cells, generating a potential source for the generation of drug modified antigenic MHC peptides.

Another major aim outlined at the start of this thesis was to identify the precise molecular signatures involved in flucloxacillin related hypersensitivity. Flucloxacillin-modified proteins

involved in subsequent proteolytic digestion and presentation through TAP processing were believed to result in the presentation of drug-modified MHC peptides. The pMHC complex presented to T lymphocytes is crucial for providing signal 1, resulting in activation. T cells are heavily regulated to prevent alloreactive populations causing damage to healthy tissue. Therefore, a major aim of this thesis was to identify the precise molecular signatures involved in flucloxacillin related hypersensitivity. Due to the strong association of flucloxacillin DILI with HLA-B*57:01 (Daly *et al.*, 2009), antigen presenting cells expressing HLA-B*57:01 (C1R-B*57:01 cells) were used for immunoaffinity capture of MHC peptides. The implementation of the required methods for this section of work were successful.

Flucloxacillin modification of MHC peptides presented by C1R-B*57:01 cells incubated with drug for 10 minutes were identified (chapter 5). Due to the complexity involved with antigen processing pathways, this points towards direct modification of MHC peptides already presented (Figure 7.1, i). When antigen presenting cells were incubated with flucloxacillin for 48 hours, the modification of MHC peptides was too observed. Interestingly, upon manual characterization of peptide sequences, it was found that flucloxacillin modified peptides were not found in untreated controls. This could indicate differential protesomal processing, perhaps due to the introduction of missed cleavages (Figure 7.1, ii). Furthermore, ligands uniquely presented in the presence of flucloxacillin showed an overall weaker binding affinity to HLA-B*57:01. Interestingly, both the parent drug and degradation product (penicilloic acid) were found to interact non-covalently with HLA-B*57:01. This could provide an additional anchor in the C-F pockets of the peptide binding groove, with a preference for aromatic ring terminating peptides. Indeed, a significant increase in phenylalanine and decrease in tryptophan at the C-terminal amino acid position of flucloxacillin unique HLA-B*57:01 peptides was identified, leading to the generation of novel self-peptides (Figure 7.1, iii). While amino acid preferences and HLA binding affinity was altered in the presence of flucloxacillin treatment, the overall proteome appears the same.

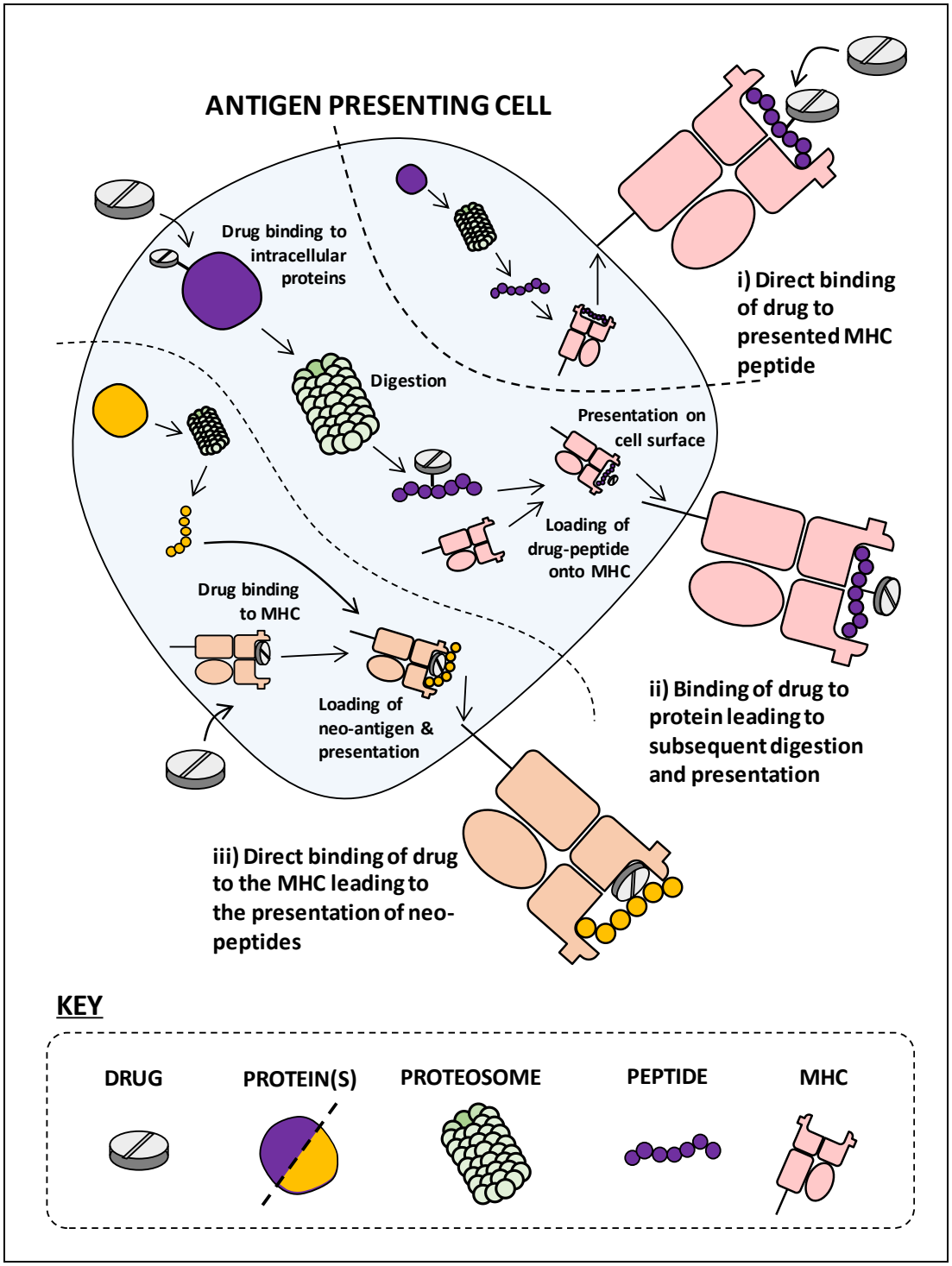


Figure 7.1. Schematic visualzation of the hypotheses for presentation of flucloxacillin-haptented peptides. Flucloxacillin covalently binds to peptides that are presented by MHC molecules on the cell surface (i, direct binding); or intracellular proteasomal processing of flucloxacillin-haptenated proteins generates haptenated peptides of an appropriate size, which are loaded on to MHC molecules, transported, and presented on the cell surface (ii); or flucloxacillin and its penicilloic acid binds to MHC molecules through either covalent or non-covalent interaction, leading to the presentation of neo-peptides (iii).

Analysis of MHC peptide data acquired using mass spectrometry is challenging due to the inability to hypothesise the protein source of the protease used for digestion. Further complications arise from biological PTMs and peptides of no single protein source, i.e. spliced peptides. Software is used for the automatic characterization of MHC peptides, as manual interpretation all spectra would not be feasible. Statistical validation can also be drawn from the peptide spectral matches made by search engines, making the data analysis more robust. Drug modification of peptides can, in theory, be searched as a potential 'PTM'. While different software packages allow for the addition of user specific search parameters, none are particularly good at identifying drug modified peptides, where the drug fragments in the mass spectrometer. While flucloxacillin fragmentation is helpful to determine the presence of drug on a peptide, it hinders automatic characterization of peptide sequences. A number of different approaches were used to automatically detect flucloxacillin modified peptides with limited success. While some annotations were made using a custom built peptide library, manual interpretation of flucloxacillin modified spectra were more convincing. Further investigation into bioinformatic workflows for the characterization of flucloxacillin modified MHC peptides is required. However, the information generated from this section of the thesis highlighted the complexity associated with drug-MHC peptide analysis.

Good progress was made in characterizing the peptide antigens that may have involvement in the activation of T cells. However, experimental limitations do still exist. HepaRG cells were primarily used to investigate the localisation of flucloxacillin in liver-like cells. Though the use of primary hepatocytes may indeed be preferred, the cost associated, and the availability of fresh/cryopreserved stocks make this impractical in most situations. Furthermore, the freeze thaw process is believed to disrupt their ability to perform metabolic processes of primary hepatocytes, and batch to batch variation hinders experimental replication. In addition to the above, immortalized cell lines are attractive for both the ability to perform genetic manipulation and culture expandable cell populations. That said, an advantage of using

primary hepatocytes is the ability to used genotype matched donors. While not primary tissue, the capability of HepaRG cells to differentiate into functional biliary epithelial cells makes them one of the most physiologically relevant liver like cell lines (Sison-Young *et al.*, 2015).

The challenge remains to develop predictive models for DILI. Organs, such as the liver, contain complex mixtures of different cell types each responding to a diverse range of environmental stresses. Even within single cell populations, there are very different physiological mechanisms taking place depending on external events. All of these processes are somewhat disregarded in immortalized cell culture systems. Work has begun to generate co-cultures of different cell types; however, these are still often limited to 2D monolayers of cells. More recently, spheroid-based models are being developed to improve physiological relevance, with some success. For example, monolayers of HepG2 cells lack the ability to form bile canaliculi with localised active transporters such as P-gp and MRP2. When grown in 3D cell culture, HepG2 cells can successfully form biliary structures, expressing both of these transporters (Gaskell *et al.*, 2016). While this is reassuring, the limited cell numbers often make this approach impractical for the study of immunopeptidomes. Another important factor to consider in the interpretation of immunopeptidomic results from cell lines is protein turnover. In cell culture conditions, cells are maintained in nutrient rich growth conditions while undergoing rapid expansion. This generates a bias towards proteins required for cell growth, and therefore may not reflect the physiological balance of protein degradation and synthesis (Claydon and Beynon, 2012). This is noteworthy when MHC peptide presentation is heavily dependent on the proteosomal digestion of proteins marked for degradation. Of course, while primary tissue isolated from patients undergoing flucloxacillin treatment would have been the ideal source for MHC peptide characterization, this is not possible.

A further challenge with immunoaffinity purification of HLA molecules is the vast number of polymorphisms in the human genome. The W6/32 antibody utilised in this thesis captured all class I MHC molecules. In a cell line, such as C1R-B*57:01 cells, this approach is possible due to single HLA allele expression. While in the C1R-B*57:01 cells a small amount of HLA-C*04:01 is present, the stark difference in anchor residues allows these to be removed from the analysis workflows. In the case of primary tissue, up to 6 different MHC class I alleles will be present, making it challenging to decipher the original source of MHC peptide presentation. Although the same approach of discriminating between anchor residues can be employed, it is often that the preference for amino acids at certain positions overlaps between unrelated alleles. Indeed, it may be possible to develop antibodies specific to HLA types, however this requires a level of skill and is potentially costly, making its feasibility limited.

One component that was not addressed in this thesis is the metabolism of flucloxacillin to its 5'-hydroxymethy metabolite (5'OH-flucloxacillin) which is mediated by CYP3A4 (Lakehal *et al.*, 2001). In a study examining flucloxacillin metabolism in patients with renal failure, it was determined that 5'OH-flucloxacillin has a plasma half-life double that of the parent compound. Plasma protein binding was determined, with 91.6% and 82.6% of flucloxacillin and 5'OH-flucloxacillin, respectively, covalently bound. This indicated that the unbound fraction of the 5'OH-flucloxacillin metabolite was double that of flucloxacillin. Furthermore, it was discovered that in patients with renal failure the plasma concentration of 5'OH-flucloxacillin was consistently higher than in healthy controls (Thijssen and Wolters, 1982). More recently the bioactivation of flucloxacillin by human CYPs has been investigated in detail. A panel of CYP enzymes were used to identify those responsible for flucloxacillin metabolism, with CYP3A4 being confirmed as a major component in the formation of 5'OH-flucloxacillin. In human liver, meta-analysis has revealed that CYP3A4 levels range between 0 and 601-pmol.mg^{-1}; indicating a huge range for interindividual variability of flucloxacillin

metabolism capability. CYP3A7 was also identified as a player in flucloxacillin metabolism, however, until recently it has only been believed to be expressed in foetal tissue. In fact, in approximately 10% of adult livers, CYP3A7 is expressed at up to 90-pmol.mg^{-1} protein; the mean level of CYP3A4. This increased expression of CYP3A4 is strongly associated with the CYP3A7*1C allele (Dekker *et al.*, 2019). Previous studies have identified 5'OH-flucloxacillin as a biliary epithelial cell toxin (Lakehal *et al.*, 2001). The tolerogenic environment promoted by the liver makes the threshold for immune activation high. Therefore, for T cell activation to occur immune stimulation must also be high. The low incidence of flucloxacillin induced DILI is exacerbated by the carriage of HLA-B*57:01, however this alone is not enough for the development of ADRs. Further patient susceptibility factors are likely involved in the breakdown of tolerance; for example CYP expression resulting in higher levels of 5'OH-flucloxacillin, infection providing PAMP signals in addition to DAMPs, and the disruption of immune regulation (Thijssen and Wolters, 1982; Holt and Ju, 2006; Dekker *et al.*, 2019).

Flucloxacillin can stimulate T cells without the requirement for drug metabolism (Monshi *et al.*, 2013). A number of flucloxacillin modified MHC peptides presented by HLA-B*57:01 were characterized in this study, providing a pool of antigens required for signal 1. A critical question still not answered is "why does flucloxacillin result in a high incidence of liver reactions when compared with other β-lactam antibiotics?". Flucloxacillin was found to localise within the bile canaliculi regions of HepaRG cells. This was an important discovery as flucloxacillin induced DILI often results in cholestasis (Lakehal *et al.*, 2001; Kaplowitz, 2004). Cell damage results in the release of DAMPs, recruiting circulatory T cells through signalling of the innate immune system. It is well characterized that HepaRG cells express high levels of CYP3A4 (Sison-Young *et al.*, 2015), therefore it is anticipated that flucloxacillin is metabolised by hepatocytes resulting in the localisation of 5'OH-flucloxacillin in the bile canaliculi in addition to the parent drug.

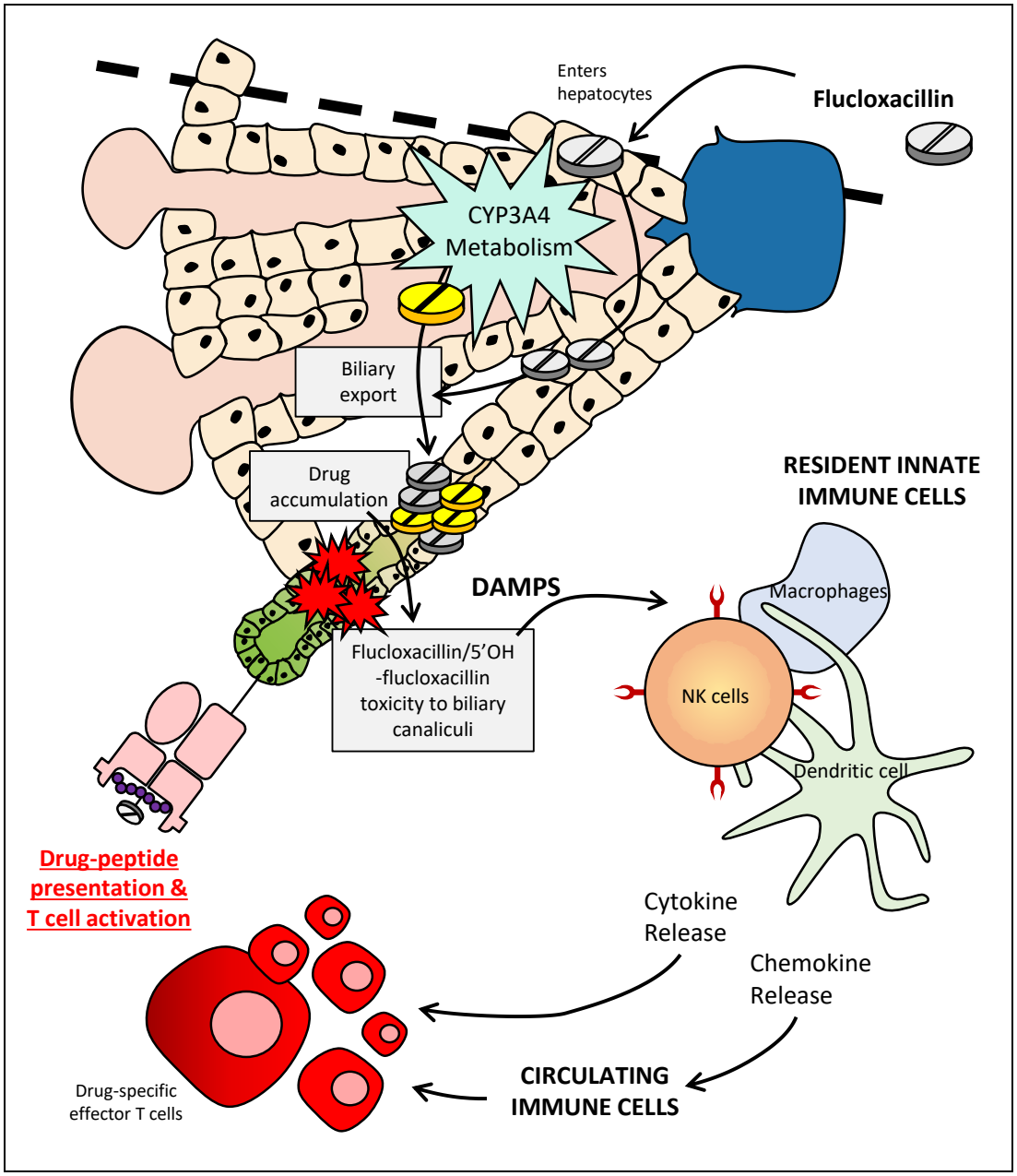


Figure 7.2. T lymphocyte responses to flucloxacillin-associated MHC peptides resulting in biliary canalicular cell death. Flucloxacillin enters hepatocytes through the circulatory system whereby it may be metabolised to 5-hydroxymethly flucloxacillin by CYP3A4. Flucloxacillin (and metabolite) localise within biliary canalicular regions through membrane transporter activity. Biliary toxicity results in the release of DAMPs, activiating the innate immune system providing signal 2 and 3 to circulatory T cells. Drug-specific effector T cells migrate to the bile canalicli, where they are exposed to signal 1 in the form of drug-modified/drug-associated MHC peptides.

Due to toxicity to biliary epithelial cells, this could in theory provide the required DAMPs to aid in the recruitment of immune cells to the site (Figure 7.2). These endogenous danger signals are believed to provide both signal 2 and signal 3 to the antigen presenting cells, upregulating the expression of co-stimulatory molecules on the cell surface and providing

polarizing cytokines needed for T cell priming. On arrival of circulating T cells, signal 1 is presented in the form of flucloxacillin-modified/flucloxacillin-associated MHC peptides, resulting in T cell mediated cytotoxicity (Park *et al.*, 2001; Pirmohamed *et al.*, 2002).

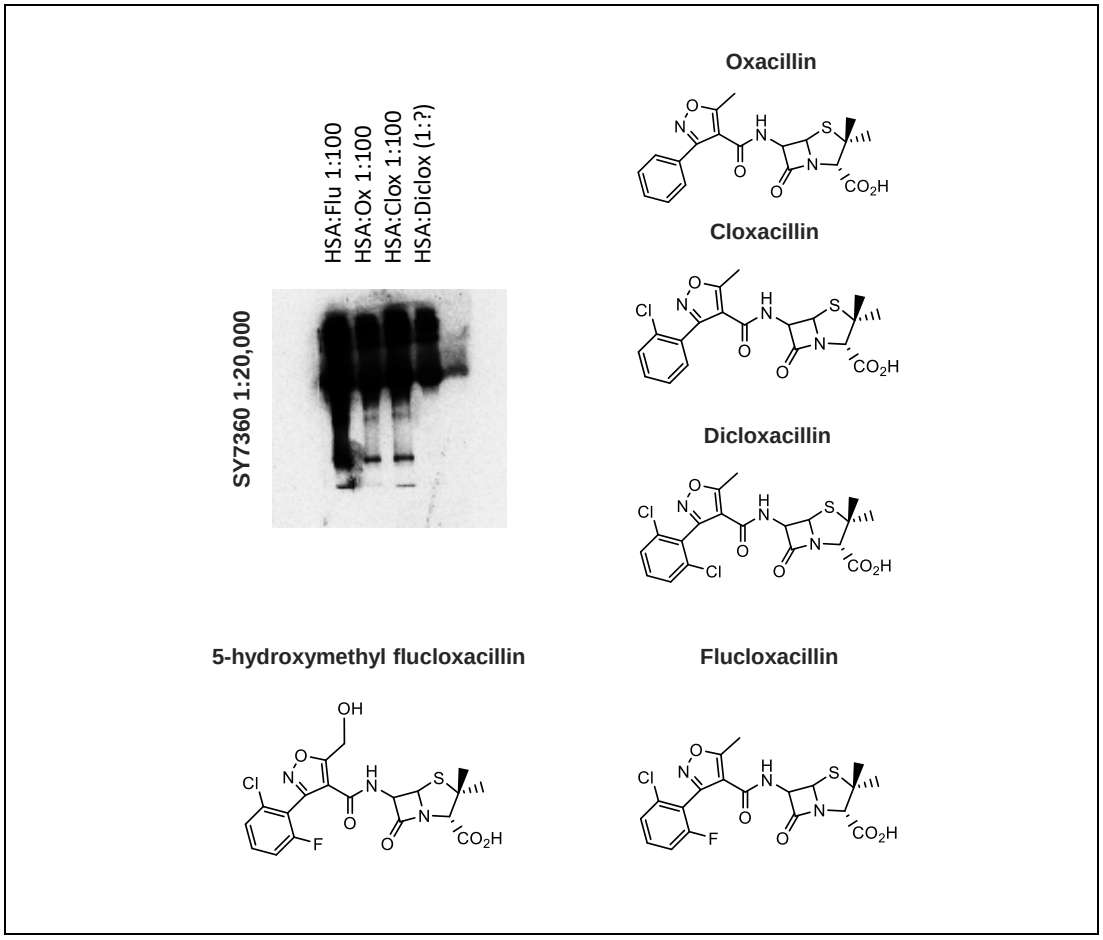


Figure 7.3. Anti-flucloxacillin antibody cross reactivity with isoxazole containing β-lactam antibiotics. As described in Chapter 2, the anti-flucloxacillin antibody (antisera) developed cross reacts with the isoxazole ring of oxacillin, cloxacillin and dicloxacillin. Due to the conservation of the isozaxole ring in the 5'OH-flucloxacillin metabolite, cross reactivity is too anticipated.

Due to the limited availability of the 5'OH-flucloxacillin metabolite, it was not possible to assess its role in the HLA-B*57:01 immunopeptidome. However, it is anticipated that 5'OH-flucloxacillin would too be capable of modifying MHC peptides. Binding of 5'OH-flucloxacillin occurs through the same mechanism as the parent drug; nucleophilic attack of the β-lactam ring by a free amine. In chapter 2, cross reactivity of the anti-flucloxacillin antisera to oxacillin, cloxacillin and dicloxacillin was identified (Figure 7.3). Due to the conservation of the isoxazole ring, it is therefore anticipated that the anti-flucloxacillin antisera would also

cross react with 5'OH-flucloxacillin. This makes it impossible to definitively discriminate between the two when localisation was observed in the bile canaliculi.

The primary aim of this research is to assist in the development of immunogenic antigens that can be used as a diagnostic marker for naïve patients prior to drug exposure. Therefore, the production of a synthetic, drug modified peptide was achieved for subsequent functional analysis with patient derived T cells. However, using HSATQK*EHGW T cell activation has so far been unsuccessful. This is attributed to several factors. Firstly, the availability of T cells derived from patients presenting with flucloxacillin DILI is limited due to the rarity of the disease. Priming to healthy volunteers can be optimised through the availability of genotyped donors, therefore HLA-B*57:01 positive cells can be used, however the relatively low yield of modified peptide makes it challenging to effectively prime T cells. Future work is focussing on improving the production of drug modified synthetic peptides to provide the tools to thoroughly investigate their ability to initiate immune responses. Indeed, colleagues within the group (unpublished data) have characterized T cell activation by amoxicillin bound to synthetic designer peptides, similar to that of other studies (Padovan *et al.*, 1997). Peptides were designed to fit within the HLA-DRB*15:01 binding cleft, a risk allele for amoxicillin induced DILI, with drug modification at various positions across the peptide backbone. T cells responded to amoxicillin modified designer peptides, while remaining inactivated by the unmodified counterpart. Encouragingly, T cell activation was specific to the location of the modification on the peptide. Although multiple T cell clones were generated to modifications in different positions, none were cross reactive. As amoxicillin was found to activate T cells when bound to designer peptides, the argument that flucloxacillin triggers T cell responses through similar mechanisms is strengthened.

Within the field there is a great need for the characterization of antigens involved in immune mediated ADRs. In future work, drug-modified peptides shown to be functional T cell antigens will be used for the generation of tetramers. MHC peptide tetramer constructs,

containing a core PE-conjugated streptavidin bound to 4 biotinylated recombinant class I heavy chain proteins, have been developed to enable direct visualisation, quantification and phenotypic characterization of antigen-specific T cells (Sims, Willberg and Klenerman, 2010). A similar technique would enable tetramers to be utilized for quantification and phenotypic assessment of T cells derived from hypersensitive patient PBMCs. The ability to counterstain T cells using antibodies for specific surface/intracellular protein markers would enable further characterization of responsive cells. This approach can be further applied to tolerant patients and healthy donors to more accurately determine precursor frequency and HLA binding epitopes on drug susceptibility. Of course, a major goal in this research is to successfully crystallise pMHC complexes containing drug-modified peptides with responsive T cells receptors, providing unequivocal evidence for the interaction between these molecules.

In this thesis binding and localisation of flucloxacillin in cell culture systems, including both primary and immortalized cell sources, has been defined. Indeed, this intracellular modification of proteins has always been believed to result in proteosomal digestion and presentation of drug- associated antigens on the cell surface. Due to the HLA predisposition of flucloxacillin-related DILI, the HLA-B*57:01 immunopeptidome was characterized in the presence of flucloxacillin. For the first time, we have demonstrated that drug-modified peptides are presented on the cell surface of antigen presenting cells, providing a potential antigen for the activation of T cells in idiosyncratic drug hypersensitivity reactions. More work is needed to ascertain the precise immunogenicity of these peptides, leading to the development of diagnostic tools with the aim to further improve patient outcomes.

BIBLIOGRAPHY

Abbas, A. K. *et al.* (2014) *Cellular and molecular immunology*. 8th edn, *J. Exp. Med.* 8th edn.

Abbas, A. K., Lichtman, A. H. and Shiv, P. (2018) *Cellular and molecular immunology*. 9th edn.

Abbas, W., Kumar, A. and Herbein, G. (2015) 'The eEF1A Proteins: At the Crossroads of Oncogenesis, Apoptosis, and Viral Infections', *Front. Oncol.*, 5.

Abe, R. *et al.* (2003) 'Toxic epidermal necrolysis and Stevens-Johnson syndrome are induced by soluble Fas ligand.', *Am. J. Pathol.*, 162(5), pp. 1515–20.

Abelin, J. G. *et al.* (2017) 'Mass Spectrometry Profiling of HLA-Associated Peptidomes in Mono-allelic Cells Enables More Accurate Epitope Prediction', *Immunity*, 46(2), pp. 315–326.

Abul K. Abbas, Andrew H. Lichtman, S. P. (2012) 'Cellular and Molecular Immunology', *J. Exp. Med.*, p. 513.

Acuto, O. and Michel, F. (2003) 'CD28-mediated co-stimulation: a quantitative support for TCR signalling.', *Nat. Rev. Immunol.*, 3(12), pp. 939–51.

Adkins, J. N. (2002) 'Toward a Human Blood Serum Proteome: Analysis By Multidimensional Separation Coupled With Mass Spectrometry', *Mol. Cell. Proteomics*, 1(12), pp. 947–955.

Agarwal, V. K. *et al.* (2010) 'Important elements for the diagnosis of drug-induced liver injury.', *Clin. Gastroenterol. Hepatol.*, 8(5), pp. 463–70.

Aiba, S. *et al.* (2003) 'p38 Mitogen-activated protein kinase and extracellular signal-regulated kinases play distinct roles in the activation of dendritic cells by two representative haptens, NiCl2 and 2, 4-dinitrochlorobenzene', *J. Invest. Dermatol.*, 120(3), pp. 390–399.

Akdis, C. A. and Akdis, M. (2009) 'Mechanisms and treatment of allergic disease in the big picture of regulatory T cells', *J. Allergy Clin. Immunol.*, 123(4), pp. 735–746.

Alberts, B., Johnson, A. and Lewis, J. (2002) *Molecular Biology of the Cell.* 4th Editio.

Alfirevic, A. and Pirmohamed, M. (2010) 'Drug-induced hypersensitivity reactions and pharmacogenomics: past, present and future.', *Pharmacogenomics*, 11(4), pp. 497–9.

Amiri, A. *et al.* (2007) 'eEF1A2 activates Akt and stimulates Akt-dependent actin remodeling, invasion and migration', *Oncogene*, 26(21), pp. 3027–3040.

Andersen, G. R., Nissen, P. and Nyborg, J. (2003) 'Elongation factors in protein biosynthesis', *Trends Biochem. Sci.*, 28(8), pp. 434–441.

Andrade, R. J. *et al.* (2005) 'Drug-induced liver injury: an analysis of 461 incidences submitted to the Spanish registry over a 10-year period.', *Gastroenterology*, 129(2), pp. 512–21.

Andrews, E. *et al.* (2010) 'A role for the pregnane X receptor in flucloxacillin-induced liver injury.', *Hepatology*, 51(5), pp. 1656–64.

Angel, T. E. *et al.* (2012) 'Mass spectrometry-based proteomics: existing capabilities and future directions', *Chem. Soc. Rev.*, 41(10), pp. 3912–3928.

Ariza, A. *et al.* (2012) 'Protein haptenation by amoxicillin: high resolution mass spectrometry analysis and identification of target proteins in serum.', *J. Proteomics*. 2012/10/09, 77, pp. 504–20.

Arrighi, J.-F. *et al.* (2001) 'A Critical Role for p38 Mitogen-Activated Protein Kinase in the Maturation of Human Blood-Derived Dendritic Cells Induced by Lipopolysaccharide, TNF- ,

and Contact Sensitizers', *J. Immunol.*, 166(6), pp. 3837–3845.

Atkinson, T. . and Kaliner, M. . (1992) 'Anaphylaxis', *Med. Clin. North Am.*, 76(4), pp. 841–855.

Autieri, M. V *et al.* (1996) 'Expression of 14-3-3 gamma in injured arteries and growth factor- and cytokine-stimulated human vascular smooth muscle cells.', *Cell Growth Differ.*, 7(11), pp. 1453–60.

Azoury, M. E. *et al.* (2018) 'Identification of T-cell epitopes from benzylpenicillin conjugated to human serum albumin and implication in penicillin allergy', *Allergy*, 73(8), pp. 1662–1672.

Bagshaw, R. D., Callahan, J. W. and Mahuran, D. J. (2000) 'Desalting of in-gel-digested protein sample with mini-C18 columns for matrix-assisted laser desorption ionization time of flight peptide mass fingerprinting.', *Anal. Biochem.*, 284(2), pp. 432–435.

Baldo, B. A., Pham, N. H. and Weiner, J. (1995) 'Detection and side-chain specificity of IgE antibodies to flucloxacillin in allergic subjects.', *J. Mol. Recognit.*, 8(3), pp. 171–7.

Banchereau, J. *et al.* (2000) 'Immunobiology of dendritic cells.', *Annu. Rev. Immunol.*, 18, pp. 767–811.

Batchelor, F. R., Dewdney, J. M. and Gazzard, D. (1965) 'Penicillin Allergy: The Formation of the Penicilloyl Determinant', *Nature*, 206(4982), pp. 362–364.

Bauer, R. A. (2015) 'Covalent inhibitors in drug discovery: from accidental discoveries to avoided liabilities and designed therapies', *Drug Discov. Today*, 20(9), pp. 1061–1073.

Beeler, A. *et al.* (2006) 'Long-lasting reactivity and high frequency of drug-specific T cells after severe systemic drug hypersensitivity reactions', *J. Allergy Clin. Immunol.*, 117(2), pp. 455–462.

Bertolino, P. *et al.* (1995) 'Peripheral deletion of autoreactive CD8+ T cells in transgenic mice expressing H-2Kb in the liver.', *Eur. J. Immunol.*, 25(7), pp. 1932–42.

Beylot, C., Doutre, M. S. and Beylot-Barry, M. (1996) 'Acute generalized exanthematous pustulosis.', *Semin. Cutan. Med. Surg.*, 15(4), pp. 244–9.

Bharadwaj, M. *et al.* (2012) 'Drug hypersensitivity and human leukocyte antigens of the major histocompatibility complex.', *Annu. Rev. Pharmacol. Toxicol.*, 52, pp. 401–31.

Bilzer, M., Roggel, F. and Gerbes, A. L. (2006) 'Role of Kupffer cells in host defense and liver disease', *Liver Int.*, 26(10), pp. 1175–1186.

Bleibel, W. *et al.* (2007) 'Drug-induced liver injury: review article.', *Dig. Dis. Sci.*, 52(10), pp. 2463–71.

Bogdanov, B. and Smith, R. D. (2005) 'Proteomics by fticr mass spectrometry: TOP down and bottom up', *Mass Spectrom. Rev.*, 24(2), pp. 168–200.

Bourdi, M. *et al.* (2002) 'Protection against acetaminophen-induced liver injury and lethality by interleukin 10: role of inducible nitric oxide synthase.', *Hepatology*, 35(2), pp. 289–98.

Bouvier, M. and Wiley, D. C. (1994) 'Importance of peptide amino and carboxyl termini to the stability of MHC class I molecules.', *Science*, 265(5170), pp. 398–402.

Breedveld, A. *et al.* (2017) 'Granulocytes as modulators of dendritic cell function.', *J. Leukoc. Biol.*, 102(4), pp. 1003–1016.

Buettiker, U. *et al.* (2006) 'Oral prednisolone induced acute generalized exanthematous pustulosis due to corticosteroids of group A confirmed by epicutaneous testing and lymphocyte transformation tests.', *Dermatology*, 213(1), pp. 40–3.

Burban, A. *et al.* (2017) 'Penicillinase-resistant antibiotics induce non-immune-mediated cholestasis through HSP27 activation associated with PKC/P38 and PI3K/AKT signaling pathways', *Sci. Rep.*, 7(1), pp. 1–17.

Burbank, M. *et al.* (2015) 'Selective bile canalicular changes induced by cholestatic drugs', *Toxicol. Lett.*, 238(2), p. S301.

Byerly, F. L. *et al.* (2005) 'Valdecoxib-associated acute generalized exanthematous pustulosis.', *Burns*, 31(3), pp. 383–7.

Cacoub, P. *et al.* (2011) 'The DRESS syndrome: a literature review.', *Am. J. Med.*, 124(7), pp. 588–97.

Caligiuri, M. A. (2008) 'Human natural killer cells', *Blood*, 112(3), pp. 461–469.

Callery, M. P., Kamei, T. and Flye, M. W. (1989) 'Kupffer cell blockade inhibits induction of tolerance by the portal venous route.', *Transplantation*, 47(6), pp. 1092–4.

Cameron, B. J. *et al.* (2013) 'Identification of a Titin-Derived HLA-A1-Presented Peptide as a Cross-Reactive Target for Engineered MAGE A3-Directed T Cells', *Sci. Transl. Med.*, 5(197), p. 197ra103-197ra103.

Carey, M. A. and van Pelt, F. (2005) 'Immunochemical detection of flucloxacillin adduct formation in livers of treated rats.', *Toxicology*, 216(1), pp. 41–8.

Caron, E. *et al.* (2015) 'Analysis of Major Histocompatibility Complex (MHC) Immunopeptidomes Using Mass Spectrometry', *Mol. Cell. Proteomics*, 14(12), pp. 3105–3117.

Carr, D. F. *et al.* (2017) 'Genome-wide association study of nevirapine hypersensitivity in a sub-Saharan African HIV-infected population', *J. Antimicrob. Chemother.*, p. dkw545.

Carter, L. L. and Dutton, R. W. (1995) 'Relative perforin- and Fas-mediated lysis in T1 and T2 CD8 effector populations.', *J. Immunol.*, 155(3), pp. 1028–31.

Castrejon, J. L. *et al.* (2010) 'Stimulation of human T cells with sulfonamides and sulfonamide metabolites.', *J. Allergy Clin. Immunol.*, 125(2), p. 411–418.e4.

Cerec, V. *et al.* (2007) 'Transdifferentiation of hepatocyte-like cells from the human hepatoma HepaRG cell line through bipotent progenitor', *Hepatology*, 45(4), pp. 957–967.

Chatterjee, S. and Annaert, P. (2018) 'Drug-induced Cholestasis: Mechanisms, Models, and Markers.', *Curr. Drug Metab.*, 19(10), pp. 808–818.

Chen, L. and Flies, D. B. (2013) 'Molecular mechanisms of T cell co-stimulation and co-inhibition.', *Nat. Rev. Immunol.*, 13(4), pp. 227–42.

Chen, P. *et al.* (2011) 'Carbamazepine-Induced Toxic Effects and HLA-B*1502 Screening in Taiwan', *N. Engl. J. Med.*, 364(12), pp. 1126–1133.

Chessman, D. *et al.* (2008) 'Human leukocyte antigen class I-restricted activation of CD8+ T cells provides the immunogenetic basis of a systemic drug hypersensitivity.', *Immunity*, 28(6), pp. 822–32.

Chester, C., Fritsch, K. and Kohrt, H. E. (2015) 'Natural Killer Cell Immunomodulation: Targeting Activating, Inhibitory, and Co-stimulatory Receptor Signaling for Cancer Immunotherapy', *Front. Immunol.*, 6.

Chipinda, I. *et al.* (2010) 'Rapid and simple kinetics screening assay for electrophilic dermal sensitizers using nitrobenzenethiol.', *Chem. Res. Toxicol.*, 23(5), pp. 918–25.

Chung, W. H. *et al.* (2008) 'Granulysin is a key mediator for disseminated keratinocyte death

in Stevens-Johnson syndrome and toxic epidermal necrolysis', *Nat. Med.*, 14(12), pp. 1343–1350.

Clark, S. R. *et al.* (2007) 'Platelet TLR4 activates neutrophil extracellular traps to ensnare bacteria in septic blood.', *Nat. Med.*, 13(4), pp. 463–9.

Claydon, A. J. and Beynon, R. (2012) 'Proteome Dynamics: Revisiting Turnover with a Global Perspective', *Mol. Cell. Proteomics*, 11(12), pp. 1551–1565.

Codrea, M. C. and Nahnsen, S. (2016) 'Platforms and Pipelines for Proteomics Data Analysis and Management.', *Adv. Exp. Med. Biol.*, 919, pp. 203–215.

Colaert, N. *et al.* (2009) 'Improved visualization of protein consensus sequences by iceLogo.', *Nat. Methods*, 6(11), pp. 786–7.

Cornejo Castro, E. M. *et al.* (2015) 'HLA-allelotype associations with nevirapine-induced hypersensitivity reactions and hepatotoxicity: a systematic review of the literature and meta-analysis.', *Pharmacogenet. Genomics*, 25(4), pp. 186–98.

Corr, M. *et al.* (1994) 'T cell receptor-MHC class I peptide interactions: affinity, kinetics, and specificity.', *Science*, 265(5174), pp. 946–9.

Craig, R., Cortens, J. P. and Beavis, R. C. (2004) 'Open source system for analyzing, validating, and storing protein identification data.', *J. Proteome Res.*, 3(6), pp. 1234–42.

Crispe, I. N. *et al.* (2000) 'The liver as a site of T-cell apoptosis: graveyard, or killing field?', *Immunol. Rev.*, 174, pp. 47–62.

Curtsinger, J. M. *et al.* (1999) 'Inflammatory cytokines provide a third signal for activation of naive CD4+ and CD8+ T cells.', *J. Immunol.*, 162(6), pp. 3256–62.

Curtsinger, J. M. and Mescher, M. F. (2010) 'Inflammatory cytokines as a third signal for T cell activation.', *Curr. Opin. Immunol.*, 22(3), pp. 333–40.

D'Ambrosio, D. *et al.* (1994) 'Involvement of p21ras activation in T cell CD69 expression.', *Eur. J. Immunol.*, 24(3), pp. 616–20.

Dalton, H. R. *et al.* (2007) 'The role of hepatitis E virus testing in drug-induced liver injury.', *Aliment. Pharmacol. Ther.*, 26(10), pp. 1429–35.

Daly, A. K. *et al.* (2009) 'HLA-B*5701 genotype is a major determinant of drug-induced liver injury due to flucloxacillin.', *Nat. Genet.*, 41(7), pp. 816–9.

Daly, A. K. (2010) 'Drug-induced liver injury: past, present and future', *Pharmacogenomics*, 11(5), pp. 607–611.

Davern, T. J. *et al.* (2011) 'Acute hepatitis E infection accounts for some cases of suspected drug-induced liver injury.', *Gastroenterology*, 141(5), p. 1665–1672.e9.

Decker, K. (1990) 'Biologically active products of stimulated liver macrophages (Kupffer cells).', *Eur. J. Biochem.*, 192(2), pp. 245–61.

DeJarnatt, A. C. and Grant, J. A. (1992) 'Basic mechanisms of anaphylaxis and anaphylactoid reactions', *Immunol. Allergy Clin. North Am.*, 12(3), pp. 505–515.

Dekker, S. J. *et al.* (2019) 'Characterization of kinetics of human cytochrome P450s involved in bioactivation of flucloxacillin: inhibition of CYP3A-catalysed hydroxylation by sulfaphenazole', *Br. J. Pharmacol.*, 176(3), pp. 466–477.

deLemos, A. S. *et al.* (2016) 'Amoxicillin–Clavulanate-Induced Liver Injury', *Dig. Dis. Sci.*, 61(8), pp. 2406–2416.

Descotes, J. and Choquet-Kastylevsky, G. (2001) 'Gell and Coombs's classification: is it still

valid?', *Toxicology*, 158(1–2), pp. 43–49.

DeShazo, R. D. (1997) 'Allergic Reactions to Drugs and Biologic Agents', *JAMA J. Am. Med. Assoc.*, 278(22), p. 1895.

Diken, M. *et al.* (2017) 'Discovery and Subtyping of Neo-Epitope Specific T-Cell Responses for Cancer Immunotherapy: Addressing the Mutanome', in, pp. 223–236.

Domon, B. and Aebersold, R. (2006) 'Mass spectrometry and protein analysis', *Science*, 312(5771), pp. 212–217.

Dorfer, V. *et al.* (2014) 'MS Amanda, a universal identification algorithm optimized for high accuracy tandem mass spectra.', *J. Proteome Res.*, 13(8), pp. 3679–84.

Dudek, N. L. *et al.* (2016) 'A Systems Approach to Understand Antigen Presentation and the Immune Response', in *Methods Mol. Biol.*, pp. 189–209.

Durbin, K. R. *et al.* (2016) 'Quantitation and Identification of Thousands of Human Proteoforms below 30 kDa.', *J. Proteome Res.*, 15(3), pp. 976–82.

Ebert, R. and Florey, H. (1939) 'The Extravascular Development of the Monocyte Observed In vivo', *Br J Exp Pathol*, 20(4), pp. 342–356.

Ebstein, F. *et al.* (2016) 'Proteasomes generate spliced epitopes by two different mechanisms and as efficiently as non-spliced epitopes.', *Sci. Rep.*, 6, p. 24032.

Edwards, I. R. and Aronson, J. K. (2000) 'Adverse drug reactions: definitions, diagnosis, and management.', *Lancet (London, England)*, 356(9237), pp. 1255–9.

El-Ghaiesh, S. *et al.* (2012) 'Characterization of the antigen specificity of T-cell clones from piperacillin-hypersensitive patients with cystic fibrosis.', *J. Pharmacol. Exp. Ther.*, 341(3), pp. 597–610.

Elsadek, B. and Kratz, F. (2012) 'Impact of albumin on drug delivery - New applications on the horizon', *J. Control. Release*, 157(1), pp. 4–28.

Eng, J. K., Jahan, T. A. and Hoopmann, M. R. (2013) 'Comet: an open-source MS/MS sequence database search tool.', *Proteomics*, 13(1), pp. 22–4.

Exley, M. A. and Koziel, M. J. (2004) 'To be or not to be NKT: natural killer T cells in the liver.', *Hepatology*, 40(5), pp. 1033–40.

Fang, X. and Zhang, W. (2008) 'Affinity separation and enrichment methods in proteomic analysis.', *J. Proteomics*, 71(3), pp. 284–303.

Faridi, P. *et al.* (2018) 'A subset of HLA-I peptides are not genomically templated: Evidence for cis- and trans-spliced peptide ligands', *Sci. Immunol.*, 3(28), p. eaar3947.

Fasciglione, G. F. *et al.* (1996) 'Hapten-carrier interactions and their role in the production of monoclonal antibodies against hydrophobic haptens.', *Hybridoma*, 15(1), pp. 1–9.

Fernández-Murga, M. L. *et al.* (2018) 'Advances in drug-induced cholestasis: Clinical perspectives, potential mechanisms and in vitro systems.', *Food Chem. Toxicol.*, 120, pp. 196–212.

Fernandez, T. D. *et al.* (2008) 'Cytokine and chemokine expression in the skin from patients with maculopapular exanthema to drugs', *Allergy*, 63(6), pp. 712–719.

Fortier, M. *et al.* (2008) 'The MHC class I peptide repertoire is molded by the transcriptome', *J. Exp. Med.*, 205(3), pp. 595–610.

Gao, B., Jeong, W. and Tian, Z. (2008) 'Liver: An organ with predominant innate immunity.', *Hepatology*, 47(2), pp. 729–36.

Garzon, D. *et al.* (2014) 'Mass spectrometric strategies for the identification and characterization of human serum albumin covalently adducted by amoxicillin: Ex Vivo Studies', *Chem. Res. Toxicol.*, 27(9), pp. 1566–1574.

Gaskell, H. *et al.* (2016) 'Characterization of a functional C3A liver spheroid model', *Toxicol. Res. (Camb).*, 5(4), pp. 1053–1065.

Gebreselassie, D., Spiegel, H. and Vukmanovic, S. (2006) 'Sampling of Major Histocompatibility Complex Class I-Associated Peptidome Suggests Relatively Looser Global Association of HLA-B*5101 With Peptides', *Hum. Immunol.*, 67(11), pp. 894–906.

Gefen, T. *et al.* (2015) 'The effect of haptens on protein-carrier immunogenicity.', *Immunology*, 144(1), pp. 116–26.

Geissmann, F. *et al.* (2010) 'Development of Monocytes, Macrophages, and Dendritic Cells', *Science*, 327(5966), pp. 656–661.

Gell, P. G. and Coombs, R. . (1963) *Clinical Aspects of Immunology*. Edited by P. Gell and R. Coombs.

Gillette, M. A. and Carr, S. A. (2013) 'Quantitative analysis of peptides and proteins in biomedicine by targeted mass spectrometry.', *Nat. Methods*, 10(1), pp. 28–34.

Gomes, E. R. and Demoly, P. (2005) 'Epidemiology of hypersensitivity drug reactions.', *Curr. Opin. Allergy Clin. Immunol.*, 5(4), pp. 309–16.

Gonçalo, M. *et al.* (2013) 'HLA-B*58:01 is a risk factor for allopurinol-induced DRESS and Stevens-Johnson syndrome/toxic epidermal necrolysis in a Portuguese population', *Br. J. Dermatol.*, 169(3), pp. 660–665.

Gordon, S. (2003) 'Alternative activation of macrophages', *Nat. Rev. Immunol.*, 3(1), pp. 23–35.

Gordon, S. and Plüddemann, A. (2017) 'Tissue macrophages: heterogeneity and functions', *BMC Biol.*, 15(1), p. 53.

Grakoui, A. *et al.* (1999) 'The immunological synapse: a molecular machine controlling T cell activation.', *Science*, 285(5425), pp. 221–7.

Granot, T. *et al.* (2017) 'Dendritic Cells Display Subset and Tissue-Specific Maturation Dynamics over Human Life', *Immunity*, 46(3), pp. 504–515.

von Greyerz, S. *et al.* (2001) 'Degeneracy and additional alloreactivity of drug-specific human alpha beta(+) T cell clones', *Int Immunol*. 2001/06/30, 13(7), pp. 877–885.

Gripon, P. *et al.* (2002) 'Infection of a human hepatoma cell line by hepatitis B virus', *Proc. Natl. Acad. Sci.*, 99(24), pp. 15655–15660.

Hamey, J. J. and Wilkins, M. R. (2018) 'Methylation of Elongation Factor 1A: Where, Who, and Why?', *Trends Biochem. Sci.*, 43(3), pp. 211–223.

Hammerich, L. and Tacke, F. (2014) 'Role of gamma-delta T cells in liver inflammation and fibrosis.', *World J. Gastrointest. Pathophysiol.*, 5(2), pp. 107–13.

Hammond, T. G. *et al.* (2014) 'Mass spectrometric characterization of circulating covalent protein adducts derived from a drug acyl glucuronide metabolite: multiple albumin adductions in diclofenac patients.', *J. Pharmacol. Exp. Ther.* 2014/06/07, 350(2), pp. 387–402.

Han, D. *et al.* (2013) 'Regulation of drug-induced liver injury by signal transduction pathways: critical role of mitochondria.', *Trends Pharmacol. Sci.*, 34(4), pp. 243–53.

Hanada, K.-I., Yewdell, J. W. and Yang, J. C. (2004) 'Immune recognition of a human renal cancer antigen through post-translational protein splicing.', *Nature*, 427(6971), pp. 252–6.

Harle, D. G. and Baldo, B. A. (1990) 'Identification of penicillin allergenic determinants that bind IgE antibodies in the sera of subjects with penicillin allergy', *Mol. Immunol.*, 27(11), pp. 1063–1071.

Hashimoto, W. *et al.* (1995) 'Cytotoxic NK1.1 Ag+ alpha beta T cells with intermediate TCR induced in the liver of mice by IL-12.', *J. Immunol.*, 154(9), pp. 4333–40.

Hasim, A. *et al.* (2012) 'Post-transcriptional and epigenetic regulation of antigen processing machinery (APM) components and HLA-I in cervical cancers from Uighur women.', *PLoS One*, 7(9), p. e44952.

Hautekeete, M. L. *et al.* (1999) 'HLA association of amoxicillin-clavulanate-induced hepatitis', *Gastroenterology*, 117(5), pp. 1181–1186.

Helm, D. *et al.* (2014) 'Ion mobility tandem mass spectrometry enhances performance of bottom-up proteomics.', *Mol. Cell. Proteomics*, 13(12), pp. 1–24.

Hickey, M. J. and Kubes, P. (2009) 'Intravascular immunity: the host-pathogen encounter in blood vessels.', *Nat. Rev. Immunol.*, 9(5), pp. 364–75.

de Hoffmann, E. (1996) 'Tandem mass spectrometry: A primer', *J. Mass Spectrom.*, 31(2), pp. 129–137.

Holt, M. P. and Ju, C. (2006) 'Mechanisms of drug-induced liver injury.', *AAPS J.*, 8(1), pp. E48-54.

Horsfall Frank L. (1934) 'Formaldehyde and Serum Proteins', *J. Immunol.*, 27(6), p. 553 LP-567.

Hsiao, Y. H. *et al.* (2014) 'Genotype-phenotype association between HLA and carbamazepine-induced hypersensitivity reactions: Strength and clinical correlations', *J. Dermatol. Sci.*, 73(2), pp. 101–109.

Huang, L. *et al.* (1994) 'The liver eliminates T cells undergoing antigen-triggered apoptosis in vivo.', *Immunity*, 1(9), pp. 741–9.

Hunt, D. F. *et al.* (1992) 'Peptides presented to the immune system by the murine class II major histocompatibility complex molecule I-Ad', *Science*, 256(5065), pp. 1817–1820.

Illing, P. T. *et al.* (2012) 'Immune self-reactivity triggered by drug-modified HLA-peptide repertoire', *Nature*, 486(7404), pp. 554–558.

Ishida, Y. *et al.* (2002) 'A pivotal involvement of IFN-gamma in the pathogenesis of acetaminophen-induced acute liver injury.', *FASEB J.*, 16(10), pp. 1227–36.

Jenkins, R. E. *et al.* (2009) 'Characterisation of flucloxacillin and 5-hydroxymethyl flucloxacillin haptenated HSA in vitro and in vivo.', *Proteomics. Clin. Appl.*, 3(6), pp. 720–9.

Jenkins, R. E. *et al.* (2013) 'β-lactam antibiotics form distinct haptenic structures on albumin and activate drug-specific T-lymphocyte responses in multiallergic patients with cystic fibrosis', *Chem. Res. Toxicol.* 2013/05/15, 26(6), pp. 963–975.

Jones, G. *et al.* (1997) 'Development and validation of a genetic algorithm for flexible docking 1 1Edited by F. E. Cohen', *J. Mol. Biol.*, 267(3), pp. 727–748.

Ju, C. *et al.* (2002) 'Protective role of Kupffer cells in acetaminophen-induced hepatic injury in mice.', *Chem. Res. Toxicol.*, 15(12), pp. 1504–13.

Ju, C. and Reilly, T. (2012) 'Role of immune reactions in drug-induced liver injury (DILI)', *Drug*

Metab. Rev., 44(1), pp. 107–115.

Ju, C. and Uetrecht, J. (2002) 'Mechanism of Idiosyncratic Drug Reactions: Reactive Metabolites Formation, Protein Binding and the Regulation of the Immune System', *Curr. Drug Metab.*, 3(4), pp. 367–377.

Kambayashi, T. and Laufer, T. M. (2014) 'Atypical MHC class II-expressing antigen-presenting cells: can anything replace a dendritic cell?', *Nat. Rev. Immunol.*, 14(11), pp. 719–30.

Kaniwa, N. *et al.* (2013) 'Specific HLA types are associated with antiepileptic drug-induced Stevens-Johnson syndrome and toxic epidermal necrolysis in Japanese subjects.', *Pharmacogenomics*, 14(15), pp. 1821–1831.

Kaplan, M. H. (2013) 'Th9 cells: differentiation and disease.', *Immunol. Rev.*, 252(1), pp. 104–15.

Kaplowitz, N. (2004) 'Drug-Induced Liver Injury', *Clin. Infect. Dis.*, 38(Supplement 2), pp. S44–S48.

Kapsenberg, M. L. (2003) 'Dendritic-cell control of pathogen-driven T-cell polarization.', *Nat. Rev. Immunol.*, 3(12), pp. 984–93.

Keir, M. E., Freeman, G. J. and Sharpe, A. H. (2007) 'PD-1 regulates self-reactive CD8+ T cell responses to antigen in lymph nodes and tissues.', *J. Immunol.*, 179(8), pp. 5064–70.

Kelleher, N. L. *et al.* (1999) 'Top Down versus Bottom Up Protein Characterization by Tandem High-Resolution Mass Spectrometry', *J. Am. Chem. Soc.*, 121(4), pp. 806–812.

Kelly, J. A. *et al.* (1989) 'Crystallographic mapping of β-lactams bound to a d-alanyl-d-alanine peptidase target enzyme', *J. Mol. Biol.*, 209(2), pp. 281–295.

Kenna, J. G. (1997) 'Immunoallergic drug-induced hepatitis: lessons from halothane.', *J. Hepatol.*, 26 Suppl 1, pp. 5–12.

Kim, M.-S. *et al.* (2014) 'A draft map of the human proteome.', *Nature*, 509(7502), pp. 575–81.

Kim, S.-H. H. *et al.* (2015) 'Characterization of amoxicillin- and clavulanic acid-specific T cells in patients with amoxicillin-clavulanate-induced liver injury', *Hepatology*, 62(3), pp. 887–899.

Kim, S. and Lee, K. W. (2010) 'HLA-B * 5901 is strongly associated with methazolamide-induced Steven-Johnson syndrome/toxic epidermal necrolysis', *Pharmacogenomics*, 11, pp. 879–884.

Kim, S. and Pevzner, P. A. (2014) 'MS-GF+ makes progress towards a universal database search tool for proteomics.', *Nat. Commun.*, 5, p. 5277.

Kirkwood, K. J. *et al.* (2013) 'Characterization of native protein complexes and protein isoform variation using size-fractionation-based quantitative proteomics.', *Mol. Cell. Proteomics*, 12(12), pp. 3851–73.

Kitteringham, N. R. *et al.* (2009) 'Multiple reaction monitoring for quantitative biomarker analysis in proteomics and metabolomics.', *J. Chromatogr. B. Analyt. Technol. Biomed. Life Sci.*, 877(13), pp. 1229–39.

Knolle, P. A. *et al.* (1999) 'Induction of cytokine production in naive CD4(+) T cells by antigen-presenting murine liver sinusoidal endothelial cells but failure to induce differentiation toward Th1 cells.', *Gastroenterology*, 116(6), pp. 1428–40.

Knowles, B., Howe, C. and Aden, D. (1980) 'Human hepatocellular carcinoma cell lines secrete the major plasma proteins and hepatitis B surface antigen', *Science*, 209(4455), pp. 497–499.

Ko, T. *et al.* (2011) 'Shared and restricted T-cell receptor use is crucial for carbamazepine-induced Stevens-Johnson syndrome', *J. Allergy Clin. Immunol.*, 128(6), p. 1266–1276.e11.

Kobayashi, T. *et al.* (2017) 'Follicular helper T cells mediate IgE antibody response to airborne allergens.', *J. Allergy Clin. Immunol.*, 139(1), p. 300–313.e7.

Kongpan, T. *et al.* (2015) 'Candidate HLA genes for prediction of co-trimoxazole-induced severe cutaneous reactions.', *Pharmacogenet. Genomics*, 25(8), pp. 402–411.

Kumar, H., Kawai, T. and Akira, S. (2011) 'Pathogen Recognition by the Innate Immune System', *Int. Rev. Immunol.*, 30(1), pp. 16–34.

Lakehal, F. *et al.* (2001) 'Indirect cytotoxicity of flucloxacillin toward human biliary epithelium via metabolite formation in hepatocytes.', *Chem. Res. Toxicol.*, 14(6), pp. 694–701.

Lam, H. *et al.* (2007) 'Development and validation of a spectral library searching method for peptide identification from MS/MS.', *Proteomics*, 7(5), pp. 655–67.

Landsteiner, K. and Jacobs, J. (1935) 'Studies on the sensitization of animals with simple chemical compounds.', *J. Exp. Med.*, 61(5), pp. 643–56.

Lang, C. *et al.* (2007) 'Mutations and polymorphisms in the bile salt export pump and the multidrug resistance protein 3 associated with drug-induced liver injury.', *Pharmacogenet. Genomics*, 17(1), pp. 47–60.

Laurence, D. R. and Carpenter, J. R. (1998) *A Dictionary of Pharmacology and Allied Topics*. 2nd Editio. Edited by D. R. Laurence and J. R. Carpenter.

Lavergne, S. N. *et al.* (2009) '"Danger" Conditions Increase Sulfamethoxazole-Protein Adduct Formation in Human Antigen-Presenting Cells', *J. Pharmacol. Exp. Ther.*, 331(2), pp. 372–381.

Leal, M. F. *et al.* (2012) 'Clinical implication of 14-3-3 epsilon expression in gastric cancer.', *World J. Gastroenterol.*, 18(13), pp. 1531–7.

Levine, B. B. and Ovary, Z. (1961) 'Studies on the mechanism of the formation of the penicillin antigen: III. The N-(D-alpha-benzylpenicilloyl) group as an antigenic determinant responsible for hypersensitivity to penicillin G', *J. Exp. Med.*, 114(6), pp. 875–940.

Li, M. O. and Flavell, R. A. (2008) 'TGF-beta: a master of all T cell trades.', *Cell*, 134(3), pp. 392–404.

Li, Z. and Diehl, A. M. (2003) 'Innate immunity in the liver.', *Curr. Opin. Gastroenterol.*, 19(6), pp. 565–71.

Liang, S. *et al.* (2000) 'Indentification of venom proteins of spider S. huwena on Two-dimensional electrophoresis gel by N-terminal microsequencing and mass spectrometric peptide mapping', *J. Protein Chem.*, 19(3), pp. 225–9.

Licona-Limón, P. *et al.* (2013) 'Th9 Cells Drive Host Immunity against Gastrointestinal Worm Infection.', *Immunity*, 39(4), pp. 744–57.

Liepe, J. *et al.* (2016) 'A large fraction of HLA class I ligands are proteasome-generated spliced peptides.', *Science*, 354(6310), pp. 354–358.

Limmer, A. *et al.* (2000) 'Efficient presentation of exogenous antigen by liver endothelial cells to CD8+ T cells results in antigen-specific T-cell tolerance.', *Nat. Med.*, 6(12), pp. 1348–54.

Linderkamp, O. *et al.* (1998) 'Passive deformability of mature, immature, and active neutrophils in healthy and septicemic neonates.', *Pediatr. Res.*, 44(6), pp. 946–50.

Liu, K.-Q. *et al.* (1998) 'T Cell Receptor–initiated Calcium Release Is Uncoupled from Capacitative Calcium Entry in Itk-deficient T Cells', *J. Exp. Med.*, 187(10), pp. 1721–1727.

Liu, Z.-X., Govindarajan, S. and Kaplowitz, N. (2004) 'Innate immune system plays a critical role in determining the progression and severity of acetaminophen hepatotoxicity.', *Gastroenterology*, 127(6), pp. 1760–74.

Lovering, A. L. *et al.* (2012) 'Structural Insights into the Anti-methicillin-resistant Staphylococcus aureus (MRSA) Activity of Ceftobiprole', *J. Biol. Chem.*, 287(38), pp. 32096–32102.

Lucas, A. *et al.* (2015) 'Abacavir-reactive memory T cells are present in drug naïve individuals', *PLoS One*, 10(2), pp. 1–16.

Mackaness, G. B. (1962) 'Cellular Resistance to Infection', *J. Exp. Med.*, 116(3), pp. 381–406.

Madden, D. R. (1995) 'The three-dimensional structure of peptide-MHC complexes.', *Annu. Rev. Immunol.*, 13, pp. 587–622.

Magalhaes, I. *et al.* (2019) 'Facing the future: challenges and opportunities in adoptive T cell therapy in cancer', *Expert Opin. Biol. Ther.*, p. 14712598.2019.1608179.

Maier-Salamon, A. *et al.* (2017) 'Pharmacokinetics of flucloxacillin and its metabolites in patients with renal failure: Impact on liver toxicity', *Int. J. Clin. Pharmacol. Ther.*

Mallal, S. *et al.* (2002) 'Association between presence of HLA-B*5701, HLA-DR7, and HLA-DQ3 and hypersensitivity to HIV-1 reverse-transcriptase inhibitor abacavir.', *Lancet (London, England)*, 359(9308), pp. 727–32.

Mallal, S. *et al.* (2008) 'HLA-B*5701 screening for hypersensitivity to abacavir.', *N. Engl. J. Med.*, 358(6), pp. 568–579.

Manchanda, T. *et al.* (2002) 'Haptenation of sulfonamide reactive metabolites to cellular proteins.', *Mol. Pharmacol.* 2002/10/23, 62(5), pp. 1011–26.

Mannering, S. I. *et al.* (2018) 'Shuffling peptides to create T-cell epitopes: does the immune system play cards?', *Immunol. Cell Biol.*, 96(1), pp. 34–40.

Manuyakorn, W. *et al.* (2013) 'Phenobarbital-induced severe cutaneous adverse drug reactions are associated with CYP2C19*2 in Thai children.', *Pediatr. Allergy Immunol.*, 24(3), pp. 299–303.

Martin, M. A. and Kroetz, D. L. (2013) 'Abacavir pharmacogenetics--from initial reports to standard of care.', *Pharmacotherapy*, 33(7), pp. 765–75.

Martin, S. *et al.* (1992) 'Role of hapten-anchoring peptides in defining hapten-epitopes for MHC-restricted cytotoxic T cells. Cross-reactive TNP-determinants on different peptides.', *J. Immunol.*, 149(8), pp. 2569–75.

Masubuchi, Y. *et al.* (2003) 'Role of interleukin-6 in hepatic heat shock protein expression and protection against acetaminophen-induced liver disease.', *Biochem. Biophys. Res. Commun.*, 304(1), pp. 207–12.

Matsumoto, Y. *et al.* (2008) 'Case of acute generalized exanthematous pustulosis caused by ampicillin/cloxacillin sodium in a pregnant woman.', *J. Dermatol.*, 35(6), pp. 362–4.

Mauri-Hellweg, D. *et al.* (1996) 'Cross-reactivity of T cell lines and clones to beta-lactam antibiotics.', *J. Immunol.*, 157(3), pp. 1071–9.

McGill, M. R. *et al.* (2012) 'Acetaminophen-induced liver injury in rats and mice: comparison of protein adducts, mitochondrial dysfunction, and oxidative stress in the mechanism of toxicity.', *Toxicol. Appl. Pharmacol.*, 264(3), pp. 387–94.

McVey, C. E. *et al.* (2001) 'Crystal structures of penicillin acylase enzyme-substrate

complexes: structural insights into the catalytic mechanism.', *J. Mol. Biol.*, 313(1), pp. 139–50.

Medzhitov, R., Preston-Hurlburt, P. and Janeway, C. A. (1997) 'A human homologue of the Drosophila Toll protein signals activation of adaptive immunity.', *Nature*, 388(6640), pp. 394–7.

Mehta, A. M. *et al.* (2007) 'Genetic variation of antigen processing machinery components and association with cervical carcinoma.', *Genes. Chromosomes Cancer*, 46(6), pp. 577–86.

Meng, X. *et al.* (2011) 'Direct evidence for the formation of diastereoisomeric benzylpenicilloyl haptens from benzylpenicillin and benzylpenicillenic acid in patients.', *J. Pharmacol. Exp. Ther.*, 338(3), pp. 841–9.

Meng, X. *et al.* (2014) 'Abacavir forms novel cross-linking abacavir protein adducts in patients.', *Chem. Res. Toxicol.* 2014/02/28, 27(4), pp. 524–35.

Meng, X. *et al.* (2015) 'Auto-oxidation of Isoniazid Leads to Isonicotinic-Lysine Adducts on Human Serum Albumin.', *Chem. Res. Toxicol.* 2014/12/10, 28(1), pp. 51–8.

Meng, X. *et al.* (2016) 'Amoxicillin and Clavulanate Form Chemically and Immunologically Distinct Multiple Haptenic Structures in Patients.', *Chem. Res. Toxicol.*, 29(10), pp. 1762–1772.

Meng, X. *et al.* (2017) 'Definition of the Nature and Hapten Threshold of the β-Lactam Antigen Required for T Cell Activation In Vitro and in Patients', *J. Immunol.*, 198(11), pp. 4217–4227.

Mescher, M. F. *et al.* (2006) 'Signals required for programming effector and memory development by CD8+ T cells.', *Immunol. Rev.*, 211, pp. 81–92.

Mi, H., Muruganujan, A., Ebert, D., *et al.* (2019) 'PANTHER version 14: more genomes, a new PANTHER GO-slim and improvements in enrichment analysis tools', *Nucleic Acids Res.*, 47(D1), pp. D419–D426.

Mi, H., Muruganujan, A., Huang, X., *et al.* (2019) 'Protocol Update for large-scale genome and gene function analysis with the PANTHER classification system (v.14.0)', *Nat. Protoc.*, 14(3), pp. 703–721.

Michalski, A. *et al.* (2012) 'Ultra High Resolution Linear Ion Trap Orbitrap Mass Spectrometer (Orbitrap Elite) Facilitates Top Down LC MS/MS and Versatile Peptide Fragmentation Modes', *Mol. Cell. Proteomics*, 11(3), p. O111.013698.

Mishto, M. and Liepe, J. (2017) 'Post-Translational Peptide Splicing and T Cell Responses.', *Trends Immunol.*, 38(12), pp. 904–915.

Mizui, M. *et al.* (2008) 'Bimodal regulation of T cell-mediated immune responses by TIM-4.', *Int. Immunol.*, 20(5), pp. 695–708.

Monshi, M. M. *et al.* (2013) 'Human leukocyte antigen (HLA)-B*57:01-restricted activation of drug-specific T cells provides the immunological basis for flucloxacillin-induced liver injury.', *Hepatology*, 57(2), pp. 727–39.

Mosmann, T. R. and Coffman, R. L. (1989) 'TH1 and TH2 cells: different patterns of lymphokine secretion lead to different functional properties.', *Annu. Rev. Immunol.*, 7, pp. 145–73.

Nacaroglu, H. T. *et al.* (2014) 'Acute generalized exanthematous pustulosis induced by ceftriaxone use', *Adv. Dermatology Allergol.*, 4(12), pp. 269–271.

Nahrendorf, M., Pittet, M. J. and Swirski, F. K. (2010) 'Monocytes: Protagonists of Infarct

Inflammation and Repair After Myocardial Infarction', *Circulation*, 121(22), pp. 2437–2445.

Naisbitt, D. J. *et al.* (1999) 'Cellular disposition of sulphamethoxazole and its metabolites: implications for hypersensitivity.', *Br. J. Pharmacol.*, 126(6), pp. 1393–407.

Naisbitt, D. J. *et al.* (2001) 'Reactive metabolites and their role in drug reactions.', *Curr. Opin. Allergy Clin. Immunol.*, 1(4), pp. 317–25.

Naisbitt, D. J. *et al.* (2007) 'Investigation of the immunogenicity of diclofenac and diclofenac metabolites.', *Toxicol. Lett.*, 168(1), pp. 45–50.

Nathan, C. F. (1987) 'Secretory products of macrophages.', *J. Clin. Invest.*, 79(2), pp. 319–26.

Navarro, V. J. and Seeff, L. B. (2013) 'Liver injury induced by herbal complementary and alternative medicine.', *Clin. Liver Dis.*, 17(4), p. 715–35, x.

Nicoletti, P. *et al.* (2019) 'Drug-induced injury due to flucloxacillin: relevance of multiple HLA alleles', *Clin. Pharmacol. Ther.*

Ninkovic, T. and Hanisch, F.-G. (2007) 'O-glycosylated human MUC1 repeats are processed in vitro by immunoproteasomes.', *J. Immunol.*, 179(4), pp. 2380–8.

Nishimura, H. *et al.* (1999) 'Development of lupus-like autoimmune diseases by disruption of the PD-1 gene encoding an ITIM motif-carrying immunoreceptor.', *Immunity*, 11(2), pp. 141–51.

Niu, J. *et al.* (2015) 'Association of CD8+ T lymphocyte repertoire spreading with the severity of DRESS syndrome', *Sci. Rep.*, 5(1), p. 9913.

Norcross, M. A. *et al.* (2012) 'Abacavir induces loading of novel self-peptides into HLA-B*57: 01: an autoimmune model for HLA-associated drug hypersensitivity', *AIDS*, 26(11), pp. F21-9.

Ogese, M. *et al.* (2014) 'HLA-DQ restricted activation of nitroso-sulfamethoxazole-specific CD4+ T-lymphocytes', *Clin. Transl. Allergy*, 4(Suppl 3), p. P110.

Ohradanova-Repic, A. *et al.* (2016) 'Differentiation of human monocytes and derived subsets of macrophages and dendritic cells by the HLDA10 monoclonal antibody panel', *Clin. Transl. Immunol.*, 5(1), p. e55.

Olsen, J. V, Ong, S.-E. and Mann, M. (2004) 'Trypsin cleaves exclusively C-terminal to arginine and lysine residues.', *Mol. Cell. Proteomics*, 3(6), pp. 608–14.

Oosting, M. *et al.* (2014) 'Human TLR10 is an anti-inflammatory pattern-recognition receptor', *Proc. Natl. Acad. Sci.*, 111(42), pp. E4478–E4484.

Ortmann, B. *et al.* (1992) 'Synthetic peptides anchor T cell-specific TNP epitopes to MHC antigens.', *J. Immunol.*, 148(5), pp. 1445–50.

Ostapowicz, G. *et al.* (2002) 'Results of a prospective study of acute liver failure at 17 tertiary care centers in the United States.', *Ann. Intern. Med.*, 137(12), pp. 947–54.

Ostrov, D. *et al.* (2012) 'Drug hypersensitivity caused by alteration of the MHC-presented self-peptide repertoire.', *Proc. Natl. Acad. Sci. U. S. A.*, 109(25), pp. 9959–64.

Padovan, E. *et al.* (1997) 'Penicilloyl peptides are recognized as T cell antigenic determinants in penicillin allergy.', *Eur. J. Immunol.*, 27(6), pp. 1303–7.

Park, B. K. *et al.* (2001) 'Metabolic activation in drug allergies', *Toxicology*, 158(1–2), pp. 11–23.

Park, B. K. *et al.* (2011) 'Managing the challenge of chemically reactive metabolites in drug development', *Nat. Rev. Drug Discov.*, 10(4), pp. 292–306.

Pavlos, R. *et al.* (2014) 'Fever, Rash, and Systemic Symptoms: Understanding the Role of Virus and HLA in Severe Cutaneous Drug Allergy', *J. Allergy Clin. Immunol. Pract.*, 2(1), pp. 21–33.

Pennino, D. *et al.* (2010) 'IL-17 Amplifies Human Contact Hypersensitivity by Licensing Hapten Nonspecific Th1 Cells to Kill Autologous Keratinocytes', *J. Immunol.*, 184(9), pp. 4880–4888.

Pessayre, D. and Larrey, D. (2008) 'Drug-Induced Liver Injury', in *Textb. Hepatol.*, pp. 1209–1268.

Pétrilli, V. *et al.* (2007) 'The inflammasome: a danger sensing complex triggering innate immunity.', *Curr. Opin. Immunol.*, 19(6), pp. 615–22.

Petrov, P. D. *et al.* (2018) 'Predicting drug-induced cholestasis: preclinical models.', *Expert Opin. Drug Metab. Toxicol.*, 14(7), pp. 721–738.

Peyrière, H. *et al.* (2006) 'Variability in the clinical pattern of cutaneous side-effects of drugs with systemic symptoms: does a DRESS syndrome really exist?', *Br. J. Dermatol.*, 155(2), pp. 422–8.

Phillips, E. J. *et al.* (2011) 'Drug hypersensitivity: Pharmacogenetics and clinical syndromes', *J. Allergy Clin. Immunol.*, 127(3), pp. S60–S66.

Picard, D. *et al.* (2010) 'Drug Reaction with Eosinophilia and Systemic Symptoms (DRESS): A Multiorgan Antiviral T Cell Response', *Sci. Transl. Med.*, 2(46), p. 46ra62-46ra62.

Pichler, W. J. (2003) 'Delayed drug hypersensitivity reactions', *Ann. Intern. Med.*, 139(8), pp. 683–693.

Pichler, W. J. (2005) 'Direct T-cell stimulations by drugs--bypassing the innate immune system.', *Toxicology*, 209(2), pp. 95–100.

Piddock, L. J. and Wise, R. (1986) 'Properties of the penicillin-binding proteins of four species of the genus Bacteroides.', *Antimicrob. Agents Chemother.*, 29(5), pp. 825–32.

Pillai, V. B. *et al.* (2011) 'Acetylation of a Conserved Lysine Residue in the ATP Binding Pocket of p38 Augments Its Kinase Activity during Hypertrophy of Cardiomyocytes', *Mol. Cell. Biol.*, 31(11), pp. 2349–2363.

Pirmohamed, M. *et al.* (2002) 'The danger hypothesis—potential role in idiosyncratic drug reactions', *Toxicology*, 181–182, pp. 55–63.

Pirmohamed, M. *et al.* (2004) 'Adverse drug reactions as cause of admission to hospital', *BMJ Br. Med. J.*, 329(7463), p. 460.

Platteel, A. C. M. *et al.* (2017) 'Multi-level Strategy for Identifying Proteasome-Catalyzed Spliced Epitopes Targeted by CD8+ T Cells during Bacterial Infection.', *Cell Rep.*, 20(5), pp. 1242–1253.

Polak, M. E. *et al.* (2014) 'Vancomycin-specific T Cell responses in allergic and non-allergic individuals', *Clin. Transl. Allergy*, 4(Suppl 3), p. 2.

Posadas, S. J. and Pichler, W. J. (2007) 'Delayed drug hypersensitivity reactions - new concepts.', *Clin. Exp. Allergy*, 37(7), pp. 989–99.

Pritchard, A. L. *et al.* (2015) 'Exploration of peptides bound to MHC class I molecules in melanoma', *Pigment Cell Melanoma Res.*, 28(3), pp. 281–294.

Purcell, A. W., van Driel, I. R. and Gleeson, P. A. (2008) 'Impact of glycans on T-cell tolerance to glycosylated self-antigens', *Immunol. Cell Biol.*, 86(7), pp. 574–579.

Purcell, A. W., Ramarathinam, S. H. and Ternette, N. (2019) 'Mass spectrometry-based

identification of MHC-bound peptides for immunopeptidomics.', *Nat. Protoc.*

Pymm, P. *et al.* (2017) 'MHC-I peptides get out of the groove and enable a novel mechanism of HIV-1 escape', *Nat. Struct. Mol. Biol.*, 24(4), pp. 387–394.

Qi, Y. *et al.* (2010) 'Human basophils express amphiregulin in response to T cell-derived IL-3.', *J. Allergy Clin. Immunol.*, 126(6), p. 1260–6.e4.

Rammensee, H. G. (1995) 'Chemistry of peptides associated with MHC class I and class II molecules.', *Curr. Opin. Immunol.*, 7(1), pp. 85–96.

Rapoport, T. A. (2007) 'Protein translocation across the eukaryotic endoplasmic reticulum and bacterial plasma membranes', *Nature*, 450(7170), pp. 663–669.

Rashid, M., Goldin, R. and Wright, M. (2004) 'Drugs and the liver', *Hosp. Med.*, 65(8), pp. 456–461.

Reeves, E. and James, E. (2017) 'Antigen processing and immune regulation in the response to tumours.', *Immunology*, 150(1), pp. 16–24.

Reilly, T. P. *et al.* (2001) 'A protective role for cyclooxygenase-2 in drug-induced liver injury in mice.', *Chem. Res. Toxicol.*, 14(12), pp. 1620–8.

Renn, C. N. *et al.* (2002) 'Amoxicillin-induced exanthema in young adults with infectious mononucleosis: demonstration of drug-specific lymphocyte reactivity', *Br. J. Dermatol.*, 147(6), pp. 1166–1170.

Robinson, J. *et al.* (2015) 'The IPD and IMGT/HLA database: allele variant databases.', *Nucleic Acids Res.*, 43(Database issue), pp. D423-31.

Robinson, M. W., Harmon, C. and O'Farrelly, C. (2016) 'Liver immunology and its role in inflammation and homeostasis.', *Cell. Mol. Immunol.*, 13(3), pp. 267–76.

Romano, A. *et al.* (1995) 'Evaluation of adverse cutaneous reactions to aminopenicillins with emphasis on those manifested by maculopapular rashes.', *Allergy*, 50(2), pp. 113–8.

Rosenberg, J. and Huang, J. (2018) 'CD8 + T cells and NK cells: parallel and complementary soldiers of immunotherapy', *Curr. Opin. Chem. Eng.*, 19, pp. 9–20.

Rothenberg, M. E. and Hogan, S. P. (2006) 'The Eosinophil', *Annu. Rev. Immunol.*, 24(1), pp. 147–174.

Rozanov, D. V. *et al.* (2018) 'MHC class I loaded ligands from breast cancer cell lines: A potential HLA-I-typed antigen collection', *J. Proteomics*, 176(September 2017), pp. 13–23.

Rubegni, P. *et al.* (2008) 'Terbinafine-induced acute generalized exanthematous pustulosis.', *G. Ital. di dermatologia e Venereol. organo Uff. Soc. Ital. di dermatologia e Sifilogr.*, 143(2), pp. 151–5.

Rubinstein, D., Roska, A. K. and Lipsky, P. E. (1986) 'Liver sinusoidal lining cells express class II major histocompatibility antigens but are poor stimulators of fresh allogeneic T lymphocytes.', *J. Immunol.*, 137(6), pp. 1803–10.

Saissi, E.-H. *et al.* (no date) 'Drugs associated with acute generalized exanthematic pustulosis.', *Ann. dermatologie vénéréologie*, 130(6–7), pp. 612–8.

Saito, N. *et al.* (2013) 'Stevens-Johnson syndrome/toxic epidermal necrolysis mouse model generated by using PBMCs and the skin of patients', *J. Allergy Clin. Immunol.*, 131(2), p. 434–441.e9.

Schaid, D. J. *et al.* (2014) 'Prospective validation of HLA-DRB1*07:01 allele carriage as a predictive risk factor for lapatinib-induced liver injury.', *J. Clin. Oncol.*, 32(22), pp. 2296–303.

Schlapbach, C. *et al.* (2014) 'Human TH9 Cells Are Skin-Tropic and Have Autocrine and Paracrine Proinflammatory Capacity', *Sci. Transl. Med.*, 6(219), p. 219ra8-219ra8.

Schneck, J. *et al.* (2008) 'Effects of treatments on the mortality of Stevens-Johnson syndrome and toxic epidermal necrolysis: A retrospective study on patients included in the prospective EuroSCAR Study.', *J. Am. Acad. Dermatol.*, 58(1), pp. 33–40.

Schnyder, B. *et al.* (1997) 'Direct, MHC-dependent presentation of the drug sulfamethoxazole to human alphabeta T cell clones.', *J. Clin. Invest.*, 100(1), pp. 136–41.

Schnyder, B. *et al.* (2000) 'Recognition of sulfamethoxazole and its reactive metabolites by drug-specific CD4+ T cells from allergic individuals.', *J. Immunol.*, 164(12), pp. 6647–54.

Schuster, H. *et al.* (2018) 'A tissue-based draft map of the murine MHC class I immunopeptidome.', *Sci. data*, 5, p. 180157.

Schwartz, R. A., McDonough, P. H. and Lee, B. W. (2013) 'Toxic epidermal necrolysis: Part II. Prognosis, sequelae, diagnosis, differential diagnosis, prevention, and treatment.', *J. Am. Acad. Dermatol.*, 69(2), p. 187.e1-16.

Sgro, C. *et al.* (2002) 'Incidence of drug-induced hepatic injuries: a French population-based study.', *Hepatology*, 36(2), pp. 451–5.

Sharanek, A. *et al.* (2016) 'Rho-kinase/myosin light chain kinase pathway plays a key role in the impairment of bile canaliculi dynamics induced by cholestatic drugs', *Sci. Rep.*, 6(1), p. 24709.

Sharpe, A. H. and Abbas, A. K. (2006) 'T-cell costimulation--biology, therapeutic potential, and challenges.', *N. Engl. J. Med.*, 355(10), pp. 973–5.

Sheppard, K.-A. *et al.* (2004) 'PD-1 inhibits T-cell receptor induced phosphorylation of the ZAP70/CD3zeta signalosome and downstream signaling to PKCtheta.', *FEBS Lett.*, 574(1–3), pp. 37–41.

Shi, C. and Pamer, E. G. (2011) 'Monocyte recruitment during infection and inflammation', *Nat. Rev. Immunol.*, 11(11), pp. 762–774.

Shteynberg, D. *et al.* (2011) 'iProphet: multi-level integrative analysis of shotgun proteomic data improves peptide and protein identification rates and error estimates.', *Mol. Cell. Proteomics*, 10(12), p. M111.007690.

Shteynberg, D. *et al.* (2013) 'Combining results of multiple search engines in proteomics.', *Mol. Cell. Proteomics*, 12(9), pp. 2383–93.

Sims, S., Willberg, C. and Klenerman, P. (2010) 'MHC–peptide tetramers for the analysis of antigen-specific T cells', *Expert Rev. Vaccines*, 9(7), pp. 765–774.

Singh, J. *et al.* (2011) 'The resurgence of covalent drugs.', *Nat. Rev. Drug Discov.*, 10(4), pp. 307–17.

Sison-Young, R. L. C. *et al.* (2015) 'Comparative Proteomic Characterization of 4 Human Liver-Derived Single Cell Culture Models Reveals Significant Variation in the Capacity for Drug Disposition, Bioactivation, and Detoxication', *Toxicol. Sci.*, 147(2), pp. 412–424.

Siuti, N. and Kelleher, N. L. (2007) 'Decoding protein modifications using top-down mass spectrometry.', *Nat. Methods*, 4(10), pp. 817–21.

Slone, S. R., Lesort, M. and Yacoubian, T. A. (2011) '14-3-3theta Protects against Neurotoxicity in a Cellular Parkinson's Disease Model through Inhibition of the Apoptotic Factor Bax', *PLoS One*. Edited by A. C. LeBlanc, 6(7), p. e21720.

Smith, G. S. *et al.* (1998) 'Role of neutrophils in hepatotoxicity induced by oral acetaminophen administration in rats.', *J. Surg. Res.*, 80(2), pp. 252–8.

Smith, N. F., Figg, W. D. and Sparreboom, A. (2005) 'Role of the liver-specific transporters OATP1B1 and OATP1B3 in governing drug elimination', *Expert Opin. Drug Metab. Toxicol.*, 1(3), pp. 429–445.

Soen, Y. *et al.* (2003) 'Detection and characterization of cellular immune responses using peptide-MHC microarrays.', *PLoS Biol.*, 1(3), p. E65.

Spraggs, C. F. *et al.* (2011) 'HLA-DQA1*02:01 is a major risk factor for lapatinib-induced hepatotoxicity in women with advanced breast cancer.', *J. Clin. Oncol.*, 29(6), pp. 667–73.

Stave, J. W. and Lindpaintner, K. (2013) 'Antibody and Antigen Contact Residues Define Epitope and Paratope Size and Structure', *J. Immunol.*, 191(3), pp. 1428–1435.

Steel, L. F. *et al.* (2003) 'Efficient and specific removal of albumin from human serum samples.', *Mol. Cell. Proteomics*, 2(4), pp. 262–270.

Steinman, R. M. (1991) 'The dendritic cell system and its role in immunogenicity.', *Annu. Rev. Immunol.*, 9, pp. 271–96.

Steinman, R. M., Hawiger, D. and Nussenzweig, M. C. (2003) 'Tolerogenic Dendritic Cells', *Annu. Rev. Immunol.*, 21(1), pp. 685–711.

Stephens, C., Andrade, R. J. and Lucena, M. I. (2014) 'Mechanisms of drug-induced liver injury.', *Curr. Opin. Allergy Clin. Immunol.*, 14(4), pp. 286–92.

Storkus, W. J. *et al.* (1987) 'NK susceptibility varies inversely with target cell class I HLA antigen expression.', *J. Immunol.*, 138(6), pp. 1657–9.

Sundaram, V. and Björnsson, E. S. (2017) 'Drug-induced cholestasis.', *Hepatol. Commun.*, 1(8), pp. 726–735.

Tamai, I. *et al.* (2000) 'Molecular Identification and Characterization of Novel Members of the Human Organic Anion Transporter (OATP) Family', *Biochem. Biophys. Res. Commun.*, 273(1), pp. 251–260.

Tassaneeyakul, W. *et al.* (2009) 'Strong association between HLA-B*5801 and allopurinol-induced Stevens-Johnson syndrome and toxic epidermal necrolysis in a Thai population.', *Pharmacogenet. Genomics*, 19(9), pp. 704–9.

Tatlidil, D., Ucuncu, M. and Akdogan, Y. (2015) 'Physiological concentrations of albumin favor drug binding', *Phys. Chem. Chem. Phys.*, 17(35), pp. 22678–22685.

Teixeira, M., Silva, E. and Selores, M. (2006) 'Acute generalized exanthematous pustulosis induced by nimesulide.', *Dermatol. Online J.*, 12(6), p. 20.

Tenzer, S. *et al.* (2005) 'Modeling the MHC class I pathway by combining predictions of proteasomal cleavage,TAP transport and MHC class I binding', *C. Cell. Mol. Life Sci.*, 62(9), pp. 1025–1037.

Thiede, B. *et al.* (2000) 'Analysis of missed cleavage sites, tryptophan oxidation and N-terminal pyroglutamylation after in-gel tryptic digestion.', *Rapid Commun. mass Spectrom.*, 14(6), pp. 496–502.

Thijssen, H. H. and Wolters, J. (1982) 'The metabolic disposition of flucloxacillin in patients with impaired kidney function.', *Eur. J. Clin. Pharmacol.*, 22(5), pp. 429–34.

Torres, M. J. *et al.* (1997) 'IgG and IgE Antibodies in Subjects Allergic to Penicillins Recognize Different Parts of the Penicillin Molecule', *Int. Arch. Allergy Immunol.*, 113(1–3), pp. 342–

344.

Tripathi, C. *et al.* (2011) 'Drug-induced Stevens-Johnson syndrome (SJS), toxic epidermal necrolysis (TEN), and SJS-TEN overlap: A multicentric retrospective study', *J. Postgrad. Med.*, 57(2), p. 115.

Tsutsui, H. *et al.* (1997) 'IL-18 accounts for both TNF-alpha- and Fas ligand-mediated hepatotoxic pathways in endotoxin-induced liver injury in mice.', *J. Immunol.*, 159(8), pp. 3961–7.

Uetrecht, J. and Naisbitt, D. J. (2013) 'Idiosyncratic adverse drug reactions: current concepts.', *Pharmacol. Rev.*, 65(2), pp. 779–808.

Veldhoen, M. *et al.* (2008) 'Transforming growth factor-beta "reprograms" the differentiation of T helper 2 cells and promotes an interleukin 9-producing subset.', *Nat. Immunol.*, 9(12), pp. 1341–6.

Verma, S. and Kaplowitz, N. (2009) 'Diagnosis, management and prevention of drug-induced liver injury.', *Gut*, 58(11), pp. 1555–64.

Vigneron, N. *et al.* (2004) 'An antigenic peptide produced by peptide splicing in the proteasome.', *Science*, 304(5670), pp. 587–90.

Vinuesa, C. G. *et al.* (2016) 'Follicular Helper T Cells.', *Annu. Rev. Immunol.*, 34(1), pp. 335–68.

Vivier, E. *et al.* (2011) 'Innate or adaptive immunity? The example of natural killer cells.', *Science*, 331(6013), pp. 44–9.

Wang, G. and Zhu, W. (2016) 'Molecular docking for drug discovery and development: a widely used approach but far from perfect', *Future Med. Chem.*, 8(14), pp. 1707–1710.

Wang, J. huai and Reinherz, E. L. (2012) 'The structural basis of αβ T-lineage immune recognition: TCR docking topologies, mechanotransduction, and co-receptor function', *Immunol. Rev.*, 250(1), pp. 102–119.

Watkins W., S. and P. (2013) 'Activating interactions of sulfanilamides with T cell receptors', *Open J. Immunol.*, 3(3), pp. 139–157.

Wehenkel, M. *et al.* (2012) 'A selective inhibitor of the immunoproteasome subunit LMP2 induces apoptosis in PC-3 cells and suppresses tumour growth in nude mice.', *Br. J. Cancer*, 107(1), pp. 53–62.

Wei, C. Y. *et al.* (2012) 'Direct interaction between HLA-B and carbamazepine activates T cells in patients with Stevens-Johnson syndrome', *J. Allergy Clin. Immunol.*, 129(6), pp. 1562–1569.

Wen, C. *et al.* (2019) 'Systematic Studies on the Protocol and Criteria for Selecting a Covalent Docking Tool', *Molecules*, 24(11), p. 2183.

Whitaker, P. *et al.* (2011) 'Mass spectrometric characterization of circulating and functional antigens derived from piperacillin in patients with cystic fibrosis.', *J. Immunol.*, 187(1), pp. 200–11.

WHO (1972) *International drug monitoring : the role of national centres , report of a WHO meeting [held in Geneva from 20 to 25 September 1971]*.

Wieczorek, M. *et al.* (2017) 'Major Histocompatibility Complex (MHC) Class I and MHC Class II Proteins: Conformational Plasticity in Antigen Presentation.', *Front. Immunol.*, 8, p. 292.

Wilson, A. J. *et al.* (2013) 'Keap Calm, and Carry on Covalently', *J. Med. Chem.*, 56(19), pp.

7463–7476.

Wuillemin, N. *et al.* (2013) 'HLA haplotype determines hapten or p-i T cell reactivity to flucloxacillin.', *J. Immunol.*, 190(10), pp. 4956–64.

Yamada, M. *et al.* (2002) 'Identification of low-abundance proteins of bovine colostral and mature milk using two-dimensional electrophoresis followed by microsequencing and mass spectrometry', *Electrophoresis*, 23(7–8), pp. 1153–60.

Yang, J. *et al.* (2014) 'Monocyte and macrophage differentiation: circulation inflammatory monocyte as biomarker for inflammatory diseases.', *Biomark. Res.*, 2(1), p. 1.

Yaseen, F. S. *et al.* (2015) 'Promiscuous T-cell responses to drugs and drug-haptens', *J. Allergy Clin. Immunol.*, 136(2), pp. 474–476.

Yip, V. L. *et al.* (2017) 'Mass Spectrometric Characterization of Circulating Covalent Protein Adducts Derived from Epoxide Metabolites of Carbamazepine in Patients.', *Chem. Res. Toxicol.*, 30(7), pp. 1419–1435.

Yuan, L. and Kaplowitz, N. (2013) 'Mechanisms of drug-induced liver injury.', *Clin. Liver Dis.*, 17(4), pp. 507–18.

Yun, J. *et al.* (2016) 'T-cell-mediated drug hypersensitivity: immune mechanisms and their clinical relevance.', *Asia Pac. Allergy*, 6(2), pp. 77–89.

Zanni, M. P. *et al.* (1998) 'HLA-restricted, processing- and metabolism-independent pathway of drug recognition by human alpha beta T lymphocytes.', *J. Clin. Invest.*, 102(8), pp. 1591–8.

Zemmour, J. *et al.* (1992) 'The HLA-A,B "negative" mutant cell line C1R expresses a novel HLA-B35 allele, which also has a point mutation in the translation initiation codon.', *J. Immunol.*, 148(6), pp. 1941–8.

Zeng, T. *et al.* (2015) 'Association of HLA-B*1502 allele with lamotrigine-induced Stevens-Johnson syndrome and toxic epidermal necrolysis in Han Chinese subjects: a meta-analysis.', *Int. J. Dermatol.*, 54(4), pp. 488–93.

Zerr, I. *et al.* (1998) 'Detection of 14-3-3 protein in the cerebrospinal fluid supports the diagnosis of Creutzfeldt-Jakob disease.', *Ann. Neurol.*, 43(1), pp. 32–40.

Zhang, F.-R. *et al.* (2013) 'HLA-B*13:01 and the Dapsone Hypersensitivity Syndrome', *N. Engl. J. Med.*, 369(17), pp. 1620–1628.

Zhang, H. and Ge, Y. (2011) 'Comprehensive analysis of protein modifications by top-down mass spectrometry.', *Circ. Cardiovasc. Genet.*, 4(6), p. 711.

Zhang, N. and Bevan, M. J. (2011) 'CD8+ T Cells: Foot Soldiers of the Immune System', *Immunity*, 35(2), pp. 161–168.

Zheng, Y.-R. *et al.* (2014) 'Pt(IV) prodrugs designed to bind non-covalently to human serum albumin for drug delivery', *J. Am. Chem. Soc.*, 136(24), pp. 8790–8798.

Zimmerman, H. J. (2000) 'Drug-induced liver disease.', *Clin. Liver Dis.*, 4(1), p. 73–96, vi.

Zubarev, R. A. *et al.* (2000) 'Electron capture dissociation for structural characterization of multiply charged protein cations.', *Anal. Chem.*, 72(3), pp. 563–73.