University of Lille – Graduate School of Biology and Health Sciences

Faculty of Pharmacy

THESIS

Submitted to attain the degree of

Doctor of Philosophy (Ph.D.) in Biomolecules, Pharmacology, Therapeutics

Presented by

PONZONI Adele Asia

Development and Application of Mass Spectrometry Methods to Qualify Models and Compounds to Support the Discovery of Small Molecules Modulators of Antigen Presentation

Publicly defended on December 17, 2024, before the jury composed of

DEPREZ Benoit, PhD, Prof, Dir, University of Lille – CoDirector

RAMOS Corinne, PhD, Aliri Bioanalysis - CoDirector

SINCE Marc, PhD, MCU-HDR, University of Caen - Reviewer

ANDREN Per, PhD, Prof, Uppsala University - Reviewer

FRUCI Doriana, PhD, PI, Paediatric Hospital Bambino Gesù - Examiner

Supported by the European Union's Horizon 2020 research and innovation program under the Marie Skłodowska-Curie grant agreement n° 954992 – Project Capstone-ETN

Abstract and key words

The challenges of drug discovery can be addressed through a multidisciplinary approach that integrates various scientific disciplines. Analytical chemistry has contributed significantly to advances in research and development of new drugs, with mass spectrometry-based techniques being widely applied in this context. The primary objective of this thesis is to support a drug discovery program targeting the endoplasmic reticulum aminopeptidases, ERAP1 and ERAP2, by developing and applying advanced analytical methods, specifically spectrometry imaging (MSI) and liquid chromatography-tandem mass spectrometry (LC-MS/MS). ERAPs are involved in processing and presenting immunogenic antigens to T cells, and their inhibition is considered an innovative therapeutic strategy for autoimmune and oncological diseases. Three research projects targeted the validation of preclinical models and the characterization of candidate compounds to evaluate their pharmacodynamic (PD) and pharmacokinetic (PK) profiles.

In the first project, matrix-assisted laser desorption/ionization (MALDI)-MSI was used to study biochemical changes in a preclinical model of ankylosing spondylitis, the HLA-B27 transgenic rat. Spatially resolved, untargeted metabolomic and lipidomic analyses revealed candidate biomarkers of two deregulated mechanisms linked to gut inflammation, intestinal permeability and immune infiltration. These findings validate the relevance of this model for testing the efficacy of ERAPs inhibitors.

In the second project, a quantitative LC-MS/MS method was developed to measure the ERAP1-dependent tumour antigen GSW11, serving as a biomarker of efficacy for an ERAP1 inhibitor, and further validating the compound's mechanism of action *in vitro*. Despite method optimization, the endogenous peptide could not be detected, highlighting the technical challenges associated with peptide extraction and analysis from complex biological samples. This led to the identification of factors affecting peptide recovery and detection, offering insights for future method development in peptidomics for pharmacodynamic studies.

The third project focused on the biodistribution and metabolism of a candidate ERAP2 inhibitor *in vivo*. Quantitative MSI was successfully employed to measure the compound's concentration in various organs, revealing its potential therapeutic applicability. Biotransformation products of the compound were detected and investigated both *in vitro* and in *vivo*. These findings pave the way for further characterization of the compound, supporting decision-making processes in future stages of drug development.

This research highlights the importance of continuously evaluating preclinical models and compounds to support informed decision-making in drug discovery, ultimately reducing the risk of failures in advanced stages. Future research will focus on integrating these approaches to provide a comprehensive characterization of ERAP inhibitors, combining quantitative tissue analysis with

biomarker modulation in preclinical models. The ultimate goal is to identify ERAPs inhibitors with a favourable therapeutic profile for clinical development in autoimmune and oncological indications.

Mass spectrometry - Biomarkers - Pharmacokinetics/Pharmacodynamics - Drug discovery - ERAP - Antigen

Résumé et mots-clés

Les défis de la découverte de médicaments peuvent être abordés par une approche multidisciplinaire qui intègre diverses disciplines scientifiques. La chimie analytique a contribué de manière significative aux avancées dans la recherche et le développement de nouveaux médicaments, les techniques basées sur la spectrométrie de masse étant largement appliquées dans ce contexte. L'objectif principal de cette thèse est de soutenir un programme de découverte de médicaments ciblant les aminopeptidases du réticulum endoplasmique, ERAP1 et ERAP2, en développant et en appliquant des méthodes analytiques avancées, spécifiquement l'imagerie par spectrométrie (MSI) et la spectrométrie de masse en tandem avec chromatographie liquide (LC-MS/MS). Les ERAPs sont impliquées dans le traitement et la présentation d'antigènes immunogènes aux lymphocytes T, et leur inhibition est considérée comme une stratégie thérapeutique innovante pour les maladies auto-immunes et oncologiques. Trois projets de recherche ont ciblé la validation des modèles précliniques et la caractérisation des composés candidats pour évaluer leurs profils pharmacodynamiques (PD) et pharmacocinétiques (PK).

Dans le premier projet, l'imagerie par spectrométrie de masse à désorption-ionisation laser assistée par matrice a été utilisée pour étudier les changements biochimiques dans un modèle préclinique de spondylarthrite ankylosante, le rat transgénique HLA-B27. Des analyses métabolomiques et lipidomiques untargeted résolues spatialement ont révélé des biomarqueurs candidats de deux mécanismes dérégulés liés à l'inflammation intestinale, la perméabilité intestinale et l'infiltration immunitaire. Ces résultats valident la pertinence de ce modèle pour tester l'efficacité des inhibiteurs des ERAP.

Dans le deuxième projet, une méthode LC-MS/MS quantitative a été développée pour mesurer l'antigène tumoral dépendant de l'ERAP1, GSW11, servant de biomarqueur d'efficacité pour un inhibiteur de l'ERAP1, et validant davantage le mécanisme d'action du composé *in vitro*. Malgré l'optimisation de la méthode, le peptide endogène n'a pas pu être détecté, soulignant les défis techniques associés à l'extraction et à l'analyse de peptides à partir d'échantillons biologiques complexes. Cela a conduit à l'identification de facteurs affectant la récupération et la détection des peptides, offrant des perspectives pour le développement futur de méthodes en peptidomique pour des études pharmacodynamiques.

Le troisième projet a porté sur la biodistribution et le métabolisme d'un inhibiteur candidat de l'ERAP2 *in vivo*. La MSI quantitative a été employée avec succès pour mesurer la concentration du composé dans divers organes, révélant son applicabilité thérapeutique potentielle. Des produits de biotransformation du composé ont été détectés et étudiés à la fois *in vitro* et *in vivo*. Ces résultats ouvrent la voie à une caractérisation plus approfondie du composé, soutenant les processus de prise de décision dans les étapes futures du développement de médicaments.

Cette recherche souligne l'importance d'évaluer en permanence les modèles précliniques et les composés pour soutenir une prise de décision éclairée dans la découverte de médicaments, réduisant ainsi le risque d'échecs dans les phases avancées. Les recherches futures se concentreront sur l'intégration de ces approches pour fournir une caractérisation complète des inhibiteurs de l'ERAPs, en combinant l'analyse tissulaire quantitative avec la modulation des biomarqueurs dans des modèles précliniques. L'objectif ultime est d'identifier des inhibiteurs des ERAP ayant un profil thérapeutique favorable pour le développement clinique dans les indications auto-immunes et oncologiques.

Spectrométrie de masse - Biomarqueurs - Pharmacocinétique/Pharmacodynamiques - Découverte de médicaments - ERAPs - Antigène

Acknowledgements

First and foremost, I would like to express my gratitude to the members of the evaluation committee, Dr Since Marc, Prof Andrén Per, and Dr Fruci Doriana, for accepting the responsibility of reviewing my thesis. Your time, expertise, and dedication are invaluable, and I am honoured to have my work evaluated by such esteemed professionals.

I am deeply grateful to my thesis co-directors for their guidance, insight, and support throughout this journey. Dr Corinne Ramos, thank you for selecting me for this position, allowing me to conduct this doctoral project. Thank you for welcoming me into your team and for protecting me when I first arrived at the company, introverted and lost. Under your guidance, I have grown in many different directions. Prof Bent Deprez, thank you for being such an amazing mentor and for sharing your expertise and passion with me. You found the time to offer your support, even when it took me a bit of courage to ask for it. Your encouragement and feedback have been fundamental to my scientific and personal growth.

Je tiens à remercier l'équipe d'Aliri pour m'avoir accueillie durant ces trois années de doctorat. Un grand merci au Dr Mathieu Gaudin pour avoir supervisé mon travail de thèse : tes conseils et orientations scientifiques ont été essentiels pour surmonter les défis quotidiens du laboratoire et pour affiner mes capacités analytiques. Merci au Dr David Bonnel pour avoir accepté de faire partie de mon comité de suivi individuel et pour les échanges scientifiques enrichissants que nous avons eus lors de ces rencontres.

Je souhaite également exprimer toute ma reconnaissance aux membres des équipes Histo (Stan, Clémence, Raphaël), Biomol (Mélodie, Sandra, Richard) et MSI (Aurore, Paul, Guillaume, Axelle) pour votre soutien constant en laboratoire, vos astuces de "pro" qui m'ont tant aidée et pour l'ambiance conviviale qui rendait chaque journée plus agréable. Un merci infini à mes collègues de bureau, l'incroyable équipe IT/data science. Fabien, merci pour les cafés décents, les muffins au chocolat, et tes dépannages informatiques chaque fois que j'avais un souci technique. Ta prévenance envers chacun de nous a fait de toi un véritable pilier du bureau. Victor, même si tu ne me le montres pas, je sais que tu tiens à moi : merci pour nos blagues, nos cafés après le déjeuner, nos discussions de fin de journée et les mille liens secureshare. Amandine, merci de m'avoir permis de répondre "strongly agree" à la question "as-tu un/e meilleure ami/e au travail ?" : tu as été plus qu'une collègue, une véritable amie, toujours présente dans et en dehors du bureau. Merci pour les discussions profondes quand le stress me submergeait, et pour m'avoir initiée aux "boundaries" à la française. Lisa, ton arrivée au bureau a profondément changé mon parcours ; tu as été l'origine et le lien du magnifique trio "ImaFamily", et pour cela, je te serai toujours reconnaissante. On sait bien que tu as de meilleures raisons qui t'amènent à Lille, mais merci d'avoir toujours trouvé le temps de nous voir.

Merci aux formidables stagiaires Ima, JD, Keo, Laure, Simon, Marie, Marine : votre présence a apporté de la joie et de la couleur aux journées de travail, et je suis heureuse d'avoir fait votre connaissance. Un grand merci également à Alba pour être la personne merveilleuse que tu es. Merci à tous les autres membres de l'équipe, votre amitié et votre soutien ont enrichi mon parcours.
Je tiens à remercier toute l'unité de recherche U1177, pour votre accueil chaleureux lors de mes participations aux réunions de laboratoire, ainsi que pour vos questions et conseils précieux. Merci spécialement à Catherine et à l'équipe ADME (Alex, Dominique, Maxime) pour m'avoir accueillie et formée dans votre laboratoire durant mon séjour. Mes remerciements vont à la Prof. Rebecca Deprez-Poulain, pour son soutien indéfectible, pour sa compréhension et son aide lors des moments cruciaux, et pour avoir façonné ce projet en me permettant de travailler sur ses composes. Un remerciement particulier pour avoir créé et dirigé le consortium CAPSTONE.

To my fellow ESRs, this experience would not have been nearly as memorable without you. Vasilis, for sharing this journey closely with me, your companionship has been invaluable in navigating French bureaucracy and in sharing joyful moments outside of work, whether testing new venues or revisiting familiar spots (Chez Raoul). Thank you for our conversations on life, future, love, and science. Ben, meeting you has filled me with joy, and this part of the world will miss you dearly. Alice, Sandra, Filipa, Paula, you have been true friends and confidantes making me feel a little less alone in this perfect imperfection. Thank you for being there, even from afar, during the tough times, the joyful ones, and for opening your hearts to me. To the wonderful duo in the UK, Shario, thank you for consistently making a memorable entrance at each meetup, for the laughs and your friendship. Niki, Galateia, Kamila, Martha, Nadia, Aroosha, I feel enriched by your presence and life experiences, and I cherish the moments we shared together.
To all the professors -Dr E. Stratikos, Dr D. Georgiadis, Prof P. van Endert, Dr J. Kuiper, Dr A. Afantadis- thank you for hosting and organizing the summits, and for the invaluable insights, questions, and guidance that have allowed me to develop interdisciplinary skills. A special thanks to Dr. Laetitia Lesire for being a pillar in the consortium. My deep gratitude also goes to Prof E. James for generously providing the cell lines for my experiments, and again to Dr Doriana Fruci, for accepting to be part of my individual evaluation committee, for welcoming me warmly in her lab during my secondment in Rome, and for her incredible kindness and support on all occasions.
My gratitude goes to the EU's H2020 research and innovation program under the MSC actions for funding this project, allowing me not only to conduct my research but also to engage in interdisciplinary and personal exchanges, meeting all the wonderful people mentioned above.
Thank you to the collaborators of the research unit U1286 for providing the animal model studied in this thesis project.

Grazie ai professionisti che mi hanno accompagnata durante questo percorso. Alla Dott.ssa Viola Zulian, per avermi insegnato l'importanza di non esagerare e di prendersi cura di sé stessi per portare a termine un buon lavoro; grazie per avermi aiutata a trovare il mio ritmo e la giusta misura in ogni cosa. Al Dott. Paolo Simeone, per avermi insegnato tanto sul mio modo di essere (anche con i miei difetti), per la tua pazienza quando serviva e la fermezza quando era necessaria; grazie per aver mantenuto la promessa di accompagnarmi fino in fondo a questo percorso, facendo molto di più di quanto mi sarei mai aspettata. Ero persa nella nebbia che riempiva la mia strada, e mi hai insegnato a guidarci attraverso.

Grazie a "Les Meufs de Luca", gli italiani a Lille, Barbara, Alex, Andrea, Alessandro, Chiara, Nicola, Valentina, Carlotta, Monica, Riccardo, Marco, Fabio, Sara. Non sapete quanto sia stato fondamentale per me avere un gruppo di amici con cui condividere momenti piacevoli in questa città, dove inizialmente mi sono sentita molto sola, e con cui poter parlare la mia lingua madre almeno una volta alla settimana. Sono immensamente grata al destino per averci fatto incontrare e anche se ora le nostre strade si divideranno, spero che la nostra amicizia duri nel tempo.

Sono inoltre grata agli amici che mi accompagnano da anni e che nonostante tutto restano vicini. Mirko, perché incontrarci sia a Bellinzona che poi, per caso, nel nord della Francia è sicuramente un segno che devo custodire la nostra amicizia. Grazie per tutti i tuoi consigli a cui do molto valore. Matteo, per il tuo supporto, i consigli e la presenza dall'altra parte della cornetta ogni volta che ne avevo bisogno; sono infinitamente grata di aver condiviso questo percorso con te. Atea, grazie perchè, pur di girare un po', sei venuta fino alla ridente cittadina di Lille (e Lione) e per la tua amicizia che dura da anni. Daniele, collega in questo mondo di spettrometria di massa, grazie per condividere i dolori della ricerca e le gioie della ASMS. Francesco, fin dai tempi delle sbobine punto di riferimento per impaginazione e stile, grazie per il tuo aiuto con le immagini di questa tesi e per le serate di Skribble. Grazie agli amici patavini, per continuare a mantenere gli imperdibili (anche se qualcuno l'ho mancato) incontri estivi e natalizi. Grazie, Thaui, per esserti aperto e per aver condiviso la tua bella anima con me.

Un grazie di cuore alle mie amiche e ai miei amici di casa, Alessandra, Michela, Elena, Federica, Manuel, Davide, Alessandra, Mauri. Anche se i miei rientri diventano sempre più rari, sento sempre la vostra vicinanza e vi ringrazio per ogni momento condiviso e per ogni traguardo festeggiato insieme. Vi voglio bene, e vi penso molto più spesso di quanto vi scriva.

Infine, un doveroso e imprescindibile grazie alla mia incredibile famiglia. Ai miei genitori, i primi e più importanti insegnanti della mia vita. Grazie per non aver mai limitato i miei orizzonti, per avermi insegnato a puntare in alto, anche se questo mi porta lontano da casa e vi fa soffrire. Grazie per aver sempre messo il mio benessere e il mio futuro sopra ad ogni cosa. Alle mie sorelle, le mie più grandi amiche e confidenti: siete il pezzo che completa il mio cuore e la mia anima. Mi mancate ogni giorno,

e vi sono infinitamente grata per essere al mio fianco, rendendo ogni successo e ogni sconfitta degni di essere condivisi, ascoltati e accettati.

Ai miei zii di sangue, sì, anche a te zio Bruno: anche se ci ripudi, non puoi impedirci di definirci tue nipoti e di esserne fiere. A te, zia Vaniera, grazie per aver sempre rispettato i miei bisogni, per avermi incoraggiata e per trovare ogni volta le parole giuste che mi permettono di guardare i problemi da una prospettiva diversa.

Ai miei zii adottivi, Franca e Marco, il nostro viaggio a New York è stata per me la prima esperienza di un'aria che forse, per un pezzetto della mia vita, chiamerò "aria di casa". Grazie per questa incredibile esperienza ed il vostro affetto incondizionato.

Alla famiglia Frigerio, che considero un po' mia: Donatella, Cesare, Sara, Claudio, Valentina e la piccola Bianca, grazie per avermi sempre accolta come in una seconda casa, per la pazienza verso la mia lontananza, i miei piani dell'ultimo minuto, e le mie apparizioni sporadiche.

E a Daniele, grazie per sopportarmi da così tanto tempo, anche quando sono una vera spina nel fianco; per aver accolto tutte le mie insicurezze e per avermi abbracciata in silenzio quando ero sopraffatta. Grazie per esserti impegnato profondamente per costruire un futuro migliore per noi, e per essere corso in mio aiuto in questi ultimi mesi difficili come un cavaliere, trovandoti costretto a fare Cenerentola. *"Io non capisco dove sia casa mia, mi chiedevo se salto nel vuoto vieni con me?"*

"It is imperfection - not perfection - that is the end result of the program written into that formidably complex engine that is the human brain, and of the influences exerted upon us by the environment and whoever takes care of us during the long years of our physical, psychological and intellectual development."

-Rita Levi Montalcini

To each and every one of you who has taken care of me

Table of Contents

Publications

Accepted Publications

- Brina D, **Ponzoni A**, Troiani M, et al. The Akt/mTOR and MNK/eIF4E pathways rewire the prostate cancer translatome to secrete HGF, SPP1 and BGN and recruit suppressive myeloid cells. *Nat Cancer*. 2023;4(8):1102-1121. doi:10.1038/s43018-023-00594-z
- **Ponzoni A**, Speca S, Hartle M, et al. An untargeted metabolomic study using MALDI-mass spectrometry imaging reveals region-specific biomarkers associated with bowel inflammation. *Metabolomics*. 2024;21(1):5. doi:10.1007/s11306-024-02200-4

Oral communications

- Doctoral Colloquium - André Verbert Day: "Biomarker Discovery, Compound Distribution, Efficacy, and Toxicity Analysis by Mass Spectrometry in ERAP-Related Models" November 2023, Lille, France

Posters

- 71st ASMS conference on mass spectrometry and allied topics: "Untargeted spatial metabolomic analysis on Ankylosing Spondylitis preclinical model using MALDI-FTICR-Mass Spectrometry Imaging – Novel analysis approach for biomarker discovery" June 2023, Houston, TX, USA
- 1st IMSIS Annual Conference on Mass Spectrometry Imaging and Integrated Topics: "Mapping Spatial Biomarkers for Inflammatory Bowel Disease associated to Ankylosing Spondylitis: Exploring the Metabolic Landscape with MALDI-FTICR-MSI" October 2023, Montreal, Canada
- 72nd ASMS conference on mass spectrometry and allied topics: "Combining in Vivo Metabolism and Whole-Body Biodistribution for Comprehensive Insight into ERAP2 Inhibitor Pharmacokinetic" June 2024, Anaheim, CA, USA

List of figures

List of tables

List of Abbreviations

Δppm Delta Part Per Million

AA Acetic Acid

ACN Acetonitrile

ADME Absorption, Distribution, Metabolism and Excretion

ADME-T Absorption, Distribution, Metabolism, Excretion, and Toxicity

AS Ankylosing Spondylitis

BBKNN Batch Balanced k Nearest Neighbours

BEH Ethylene Bridged Hybrid

BSA Bovine Serum Albumin

Capstone Controlling Antigen Processing in Autoimmune diseases and Cancer

CASI Continuous Accumulations of Selected Ions

Cer Ceramide

CID Collision Induced Dissociation

CMC Carboxymethyl Cellulose

CS Cholesterol Sulphate

CSH Charged Surface Hybrid

CTL Cytotoxic T Lymphocyte

CYP450 Cytochrome P450

DAN Diaminonaphthalene

DESI Desorption-Electrospray Ionization

DHB Dihydroxybenzoic Acid

DMPK Drug Metabolism and Pharmacokinetic

DMSO Dimethyl Sulfoxide

ER Endoplasmic Reticulum

ERAP Endoplasmic Reticulum AminoPeptidase

ESI Electrospray Ionization

FA Formic Acid

FAs Fatty Acyls

FS Full Scan

FTICR Fourier-Transform Ion Cyclotron Resonance

GPLs Glycerophospholipids

H&E Haematoxylin and Eosin

H_2O Water

HILIC Hydrophilic Interaction Liquid Chromatography

HLA Human Leucocyte Antigen

HLA-B27^{+} Human Leucocyte Antigen B27 Transgenic

HLB Hydrophilic Lipophilic Balanced

HRMS High-Resolution Mass Spectrometry

IBD Inflammatory Bowel Disease

IC_{50} Half Maximal Inhibitory Concentration

ICR Ion Cyclotron Resonance

IF Immunofluorescence

IP Intraperitoneal

IS Internal Standard

ITO Indium Tin Oxide

KO Knocked Out

LBAs Ligand Binding Assays

LC-MS/MS Liquid Chromatography coupled to tandem Mass Spectrometry

LC-PUFAs Long-Chain PolyUnsaturated Fatty Acids

LLOQ Lower Limit of Quantification

LOD Limit of Detection

LPC LysoPhosphatidylCholine

LPE LysoPhoshatidylEthanolamine

MALDI Matrix-Assisted Laser Desorption/Ionization

MeOH Methanol

MHC I Major Histocompatibility Complex Class I

MoA Mechanism of Action

MRM Multiple Reaction Monitoring

MSI Mass Spectrometry Imaging

NADP Nicotinamide Adenine Dinucleotide Phosphate

NK Natural Killer

OC_{50} Half Maximal Occupancy Concentration

OCT Optimal Cutting Temperature

OTCD On-Tissue Chemical Derivatization

PBS Phosphate Buffered Saline

PC PhosphatidylCholine

PD Pharmacodynamic

PK Pharmacokinetic

QC Quality Control

qWBA Quantitative Whole-Body Autoradiography

R^2 Coefficient of Determination

RMS Root Mean Square

ROIs Regions of Interest

RP Peversed Phase

Rp Resolving power

RSD% Relative Standard Deviation Percentile

RT Ritention Time

SLs Sphingolipids

SMs Sphingomyelins

SpA Spondylarthritis

SPE Solid Phase Extraction

ST Sterol Lipid

TECs Tissue Extraction Coefficients

TFA Frifluoroacetic Acid

TIC Total Ion Count

Tox Toxicity

TPP Target Product Profile

UHPLC Ultra High Performance Liquid Chromatograhy

ULOQ Upper Limit of Quantification

UMAP Uniform Manifold Approximation and Projection

VP Vaporizer Temperature

WS Working Solution

WT Wild Type

1 Chapter 1 - Introduction

1.1 Drug discovery as a matter of valid target, valid models and valid compounds

The drug discovery pipeline requires experts from different research fields joining forces to understand the mechanism underlying a disease status and identify a target, to discover chemical entities that can be used to act on this target, and to test them against a series of attributes that must be met to be successful in advancing toward clinical trials. The aim of a drug discovery program is to deliver to clinical practice one or more candidate molecules with substantial benefits over currently available therapies.

A successful drug discovery program depends on three fundamental components: valid targets, valid models, and valid compounds.

1.1.1 Basic research – valid targets

The unmet need of effective medical treatments for a human disease drives the efforts of basic researchers to study and characterize the pathological condition, aiming to identify deregulated mechanisms and pathways associate with its onset and development. Following the discovery of a deregulated process, further investigations are carried out to identify a player, a biological entity, that could serve as a target of that process. The next step involves biochemical studies to validate that the selected target is indeed involved in the pathological process and that acting on it can lead to a therapeutic activity without causing undesired side effects. At this stage, several complementary approaches provide evidences of the target's essential function, its druggability and selectivity (Wyatt et al., 2011). By understanding how potential drugs interact with the target, the mechanism of action (MoA) for the therapeutic intervention is studied and proposed. Target validation is crucial in drug discovery as the development of effective and safe drugs depends on it.

1.1.2 Biomarker discovery – valid models

Experimental models range from simple systems such as isolated targets, to cell lines grown in monolayers or in 3D organoids to mimic *in vivo* condition and extend to animal models. These models offer varying levels of biological complexity, each suited for different application. Throughout a drug discovery program, it is essential to identify and develop appropriate experimental *in vitro* and *in vivo* models to thoroughly study the pathological conditions, validate the pharmacological target and test candidate compounds. Models that mirror the human disease state along with physiological controls are warranted to assess efficacy and address safety issues.

All models are characterized by limitations and potential. To be considered valid, a model has to mimic as closely as possible the condition that aim to model, the condition needs to arise through the appropriate mechanism and the outcomes measured in the model need to be predictive of the ones in humans (Tadenev & Burgess, 2019).

Identifying, in such models, endogenous molecules that can serve as indicators of pathological and physiological processes is beneficial. The process of biomarker discovery supports decision-making by improving the understanding of the disease, validating mechanism of action and efficacy and anticipating compound's behaviour in humans. Biomarkers are defined as measurable characteristics that are indicators of physiological processes, pathogenic processes, or biological responses to interventions, including therapeutic ones (FDA-NIH Biomarker Working Group, 2016). They allow earlier and more robust measurements of drug safety and efficacy and can be used as readouts for pharmacodynamic studies.

1.1.3 DMPK and PD properties for a successful program – valid compounds

With identified and validated pharmaceutical target, and available disease models, medicinal chemists can proceed to screen chemical libraries or rationally designed probes that can interact with the target and modulate its activity *in vitro* and *in vivo*. *In silico* models are also beneficial at early stages to prioritize compounds with a higher probability of succeeding based on drug-likeness properties and predicted interactions with the target structure.

Several features of the compounds, including but not limited to toxicity, bioavailability, selectivity, stability, potency, are tested and optimized during the discovery phase. The compound's physicochemical properties and behaviour in physiological systems depend on their chemical structure. The objective is to obtain drug candidates that fit the target product profile (TPP) required to enter the development phase. If the multiparametric characterization shows that a compound fails in meeting the TPP criteria, it may either be optimized or removed from the list of candidates.

Absorption, Distribution, Metabolism, Excretion, and Toxicity (ADME-T) are evaluated to understand the fate of a compound once administered, as well as its potential to reach therapeutic concentrations at the target site without accumulating off-target (Hughes et al., 2011). The most promising candidates undergo *in vitro* and *in vivo* testing in validated models to assess their pharmacokinetic (PK) and pharmacodynamic (PD) profiles. These two branches of pharmacology focus on studying the compound's journey from administration to target interaction, and, ultimately, to elimination from the body. Combining PK and PD studies allows the correlation of the candidate compounds concentration in a body compartment with the effects (Negus & Banks, 2018).

Demonstrating appropriate target exposure, target engagement, and pharmacological activity at early stages is essential to predict a candidate compound's likelihood of success in advanced clinical

phases. Target exposure refers to the compound's concentration at site of action, which must be high enough to ensure an effect. Target engagement is the binding of a compound to its pharmacological target, which is essential to modulate target activity. Lastly, pharmacological activity is the functional modification of a target to obtain the desired therapeutic effect, measurable through *in vitro* and *in vivo* studies (Morgan et al., 2012).

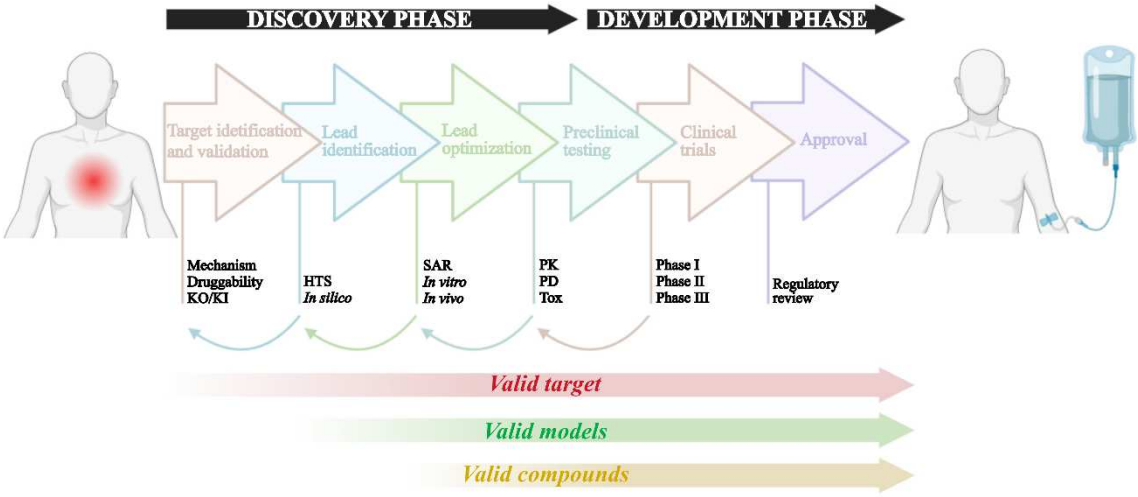


Figure 1. 1 Graphical representation of the drug discovery and development pipeline and phases
Adapted from (Ain et al., 2020; Bano et al., 2023), created with BioRender.

1.1.4 An interdisciplinary and in-parallel process

The success of a drug discovery program is dependent on the integration and validation of targets, models, and compounds, ultimately aiming to deliver one or more candidate molecules to clinical practice, with substantial benefits over currently available therapies.

Although preclinical drug discovery programs have well-defined milestones, the process is neither linear nor straightforward. The duration, costs and attrition rates of drug discovery programs are evidence of how challenging it is to successfully transition compounds from discovery phases to clinical trials.

The 100 years from 1900 to 2000 have marked a significant advancement in drug discovery, largely due to the development of collaborations across various disciplines (Drews, 2000). Originally led by medicinal chemistry, the field has expanded to integrate expertise from pharmaceutical technology, toxicology, pharmacology and analytical chemistry. Interdisciplinary collaboration has enhanced the efficiency in drug discovery, by promoting an iterative approach of learning from mistakes and revisiting previous stages.

1.2 Analytical methods to support drug discovery: Mass spectrometry

Cutting edge technologies and new methodologies lead to breakthroughs in the development of new drugs, with mass spectrometry emerging as a powerful tool in analytical chemistry. Its advantages over alternative analytical methods in throughput, sensitivity, and specificity, make MS invaluable across various stages of drug discovery, from studying drug targets and characterizing compounds to quantification and preclinical testing.

1.2.1 Liquid chromatography coupled to mass spectrometry: Enhancing sensitivity and specificity in bioanalysis

A common cause of candidate compounds failure during clinical phases is poor ADME properties. The complete assessment of drug metabolism and pharmacokinetic (DMPK) profile at the preclinical stage is beneficial for a successful drug discovery campaign. These studies rely on quantification of the compound and its metabolites in biofluids. MS/MS fragmentation in liquid chromatography coupled to tandem mass spectrometry LC-MS/MS) is the preferred method for drug quantification in biological matrices (Ackermann et al., 2002; Korfmacher, 2005). This technique isolates and fragments the ion of interest, using the intensity of the resulting fragment ions for quantification.

Beyond its role in PK and drug quantification, LC-MS/MS's quantitative capabilities extent to biomarker study. Efficacy and toxicity biomarkers can be measured alongside compounds during PD evaluation. Compared to ligand-binding assays (LBA), LC-MS/MS offers flexibility for diverse analytes without the need for specific antibodies, greater specificity and selectivity, faster method development and the ability to multiplex (Zheng et al., 2014).

LC-MS/MS holds qualitative potential in biomarker discovery studies as it is the gold standard technique for molecular annotation. The online fragmentation facilitates the elucidation of molecular structures, while the high molecular coverage enables a broad application to metabolites, lipids, peptides, and proteins.

1.2.2 Achievement of the spatial dimension with mass spectrometry imaging

Although measuring drugs and metabolites concentrations in biofluids is standard practice in bioanalytical PK assays, these matrices can be poor surrogates for determining drug’s concentration within the target tissue, where therapeutic effects occur. LC-MS/MS is employed for drug quantification in tissue homogenates; however, this method lacks spatial information, which is crucial to understand a drug's efficacy and safety within specific regions.

To address the need for studying the spatial distribution of drugs within tissues, complementary techniques have been employed in drug discovery. Quantitative whole-body autoradiography (qWBA) has long been considered the gold standard, accepted by regulatory agencies, for

characterizing drug biodistribution in animal models (McEwen & Henson, 2015). This technique still holds major drawbacks, such as the requirement for radiolabelling and the limited specificity to distinguish drugs from metabolites.

The advent of soft ionization techniques, such as desorption-electrospray ionization (DESI) and matrix-assisted laser desorption/ionization (MALDI), enabled the development of mass spectrometry-based methods to generate molecular images of biomolecules directly from biological tissue samples (Caprioli et al., 1997). Since its first application in 1999 to visualize pharmaceuticals within tissues, mass spectrometry imaging (MSI) was employed for whole-body drugs visualization, with comparisons to LC-MS/MS demonstrating its quantitative potential (Reyzer et al., 2003; Troendle et al., 1999).

The development of novel methodologies, such as the deposition of compound's serial dilution on-tissue and the introduction of the isotope-labelled (or "internal standard") approach, have made quantification via MSI possible and helped addressed limitations like tissue-specific ion suppression and the irreproducibility of ion signals (Nilsson et al., 2010; Pirman, Reich, et al., 2013). The major advantages of qMSI over LC-MS/MS and qWBA are the ability to visualize and quantify both the drug and its metabolites directly on tissue in a single, untargeted, and unlabelled analysis. qMSI is applied in drug discovery to assess tissue exposure, study the blood-brain barrier penetration, and identify excretion routes. Additionally, qMSI can anticipate potential safety issues by revealing the distribution of drugs and metabolites in specific tissues (Hochart et al., 2014; Nilsson et al., 2015).

The untargeted capability of MSI is also valuable for biomarker discovery, supporting the study and validation of animal models of the disease as well as the PD evaluation of compounds. Disease progression can induce modifications in the molecular fingerprint of samples, which can aid in diagnosis, progression assessment, and determination of compounds therapeutic effect. MALDI MSI offers a powerful means to detect these modifications while preserving spatial integrity, enabling the discovery of spatially resolved biomarkers. Moreover, MSI can be combined with other imaging techniques, such as microscopy and immunofluorescence, allowing the correlation of biomarkers with histological and molecular changes to provide a more comprehensive interpretation of biochemical alterations in disease context (Neumann et al., 2020). Given its capabilities, MSI is applied at multiple stages of drug discovery, including pharmacodynamics (PD), pharmacokinetics (PK), and toxicity (Tox) assessments (Figure 1. 2).

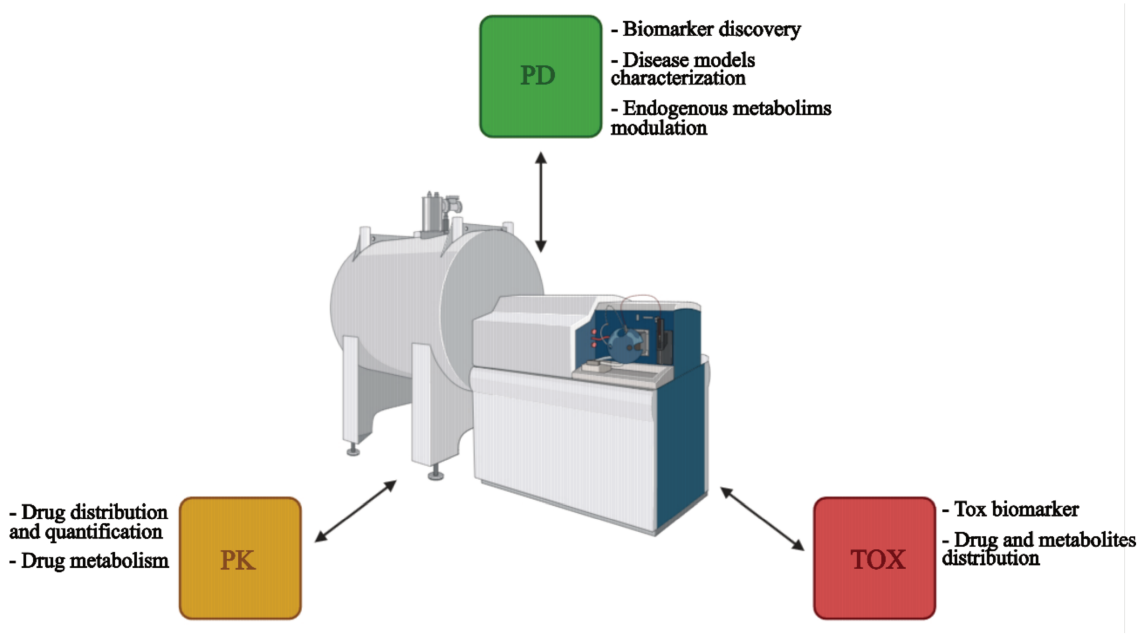


Figure 1. 2 The application of MSI in drug discovery for PK, PD, and Tox studies
Adapted from (Hochart et al., 2014), created with BioRender.

1.2.2.1 Challenges in MSI

The sample preparation for MSI analyses requires many careful measures to ensure stability, integrity, and compatibility with the technique. To prevent degradation, samples are usually snap-frozen and stored below -70°C and freezing-thaw cycles are kept to minimum to preserve the quality. Sample preparation steps, such as washing, derivatization and matrix application, are performed to remove contaminants that can interfere with the analysis and cause ion suppression of the analytes of interest, to improve analyte detection, and to support the ionization process, respectively. All these procedures must be performed with attention and, when possible, automation to ensure reproducibility. Any modifications in protocols, timelines, operators, instruments, or days of analysis can contribute to overall variability (Goodwin, 2012).

Variability introduces artifacts and batch effect to the MSI analysis, that can occur at multiple levels, including inter-pixel, inter-sample, and inter-slide, potentially misleading observations and conclusions that are not truly related to biological variations. Unavoidable variabilities can be normalized and compensated with techniques such as Total Ion Count (TIC), root mean square (RMS), tissue-extraction coefficients (TECs), or internal standards (IS). For large datasets and exploratory analysis, bioinformatics tools can help correct unavoidable artifacts. However, caution is needed to avoid introducing additional artifacts with data processing (Balluff et al., 2021).

Data analysis is also integral for interpreting the results. It is used to detect differences between samples and investigate biochemical changes related to disease conditions or treatment regimens. Statistical methods used in this context can be divided into unsupervised and supervised.

Unsupervised methods do not require prior knowledge and aim to reveal underlying biochemical changes. Supervised methods require predefined groups and are used to identify discriminating features between them (Alexandrov, 2012).

1.3 Small molecules-based immunotherapy

1.3.1 Unmet needs in cancer and autoimmune diseases

The unmet need in the management of complex pathological conditions such as cancer and autoimmune diseases drives the research for identifying new therapeutic strategies. For a long time, these conditions have been studied and treated independently, but the shared role of the immune system opens opportunities for common research pathways.

The advent of biotherapeutics in the early 20th century, and their advancement as immunotherapeutic agents in the past few decades, have transformed the treatment landscape for both cancer and autoimmune diseases. Nevertheless, these advanced therapeutic strategies tend to be effective for only a small percentage of patients (A. R. Kumar et al., 2021).

Small molecules still hold several benefits over biotherapeutics such as patient compliance, manufacturing, preservation and distribution costs (Makurvet, 2021; Poduval et al., 2023). In addition to addressing disease treatment, the primary target is the patients, making it essential to ensure that new drug research and development prioritize accessibility for end users.

Small molecules can be employed to target intracellular targets as a means of obtaining immune regulation (Zhong et al., 2021).

1.3.2 Antigen processing and presentation in immunosurveillance - ERAPs

Immunosurveillance activity is finely regulated by mechanisms of processing and presentation of antigens, which are recognized by cytotoxic lymphocytes, T or natural killer (NK) cells. The peptide repertoire is presented on cell's surface by major histocompatibility complex class I (MHC I) molecules. Research efforts by various groups in the early 2000s led to the identification of aminopeptidases located in the endoplasmic reticulum (G. E. Hammer et al., 2006; Saric et al., 2002; Saveanu et al., 2005; Serwold et al., 2002; York et al., 2002). Endoplasmic reticulum aminopeptidases 1 and 2 (ERAP1 and ERAP2) are members of the oxytocinase subfamily of the M1-family of zinc metalloproteases. These enzymes trim peptides to achieve the optimal length for stable binding to MHC I molecules. Genes encoding for ERAP1 and 2 are highly polymorphic, with several single nucleotide polymorphisms (SNPs) that have been identified as risk genes of MHC-I associated inflammatory conditions (or "MHC-I-opathies"), amongst which Ankylosing Spondylitis, Bechet's disease, Birdshot chorioretinopathy, Crohn's disease, and Psoriasis (Evans et al., 2011; Franke et al., 2010; Guasp et al., 2019; Kirino et al., 2013; Kuiper et al., 2018; Lorente et al., 2020; Strange et al.,

2010; Wiśniewski et al., 2018). Recent researches have associated ERAPs SNPs with cancer progression and predisposition to infectious disease (Cifaldi et al., 2012; Compagnone et al., 2019; Hamilton et al., 2023; Schott et al., 2022).

ERAPs have emerged as interesting therapeutic targets due to their role in immunosurveillance and their implication in numerous pathological conditions. The intervention mechanism suggests that by modulating the ERAPs trimming activity, it is possible to alter the repertoire of antigens presented by cells and consequently modify the response of cytotoxic immune cells (Figure 1. 1). This could work by dampening immune responses in autoimmune diseases, where reducing the activity of autoreactive cytotoxic cells is desired, and by enhancing immune detection and response in cancers, where immune evasion plays a central role. ERAP1 and ERAP2 are paralogues with independent enzymatic activity and different substrate specificity. Their investigation as therapeutic targets can be concurrent, with pan inhibitor, or independent, with selective inhibitors, as their inhibition can alter the presentation of distinct antigen repertoires, and thus may serve different therapeutic purposes (Fougiaxis et al., 2024).

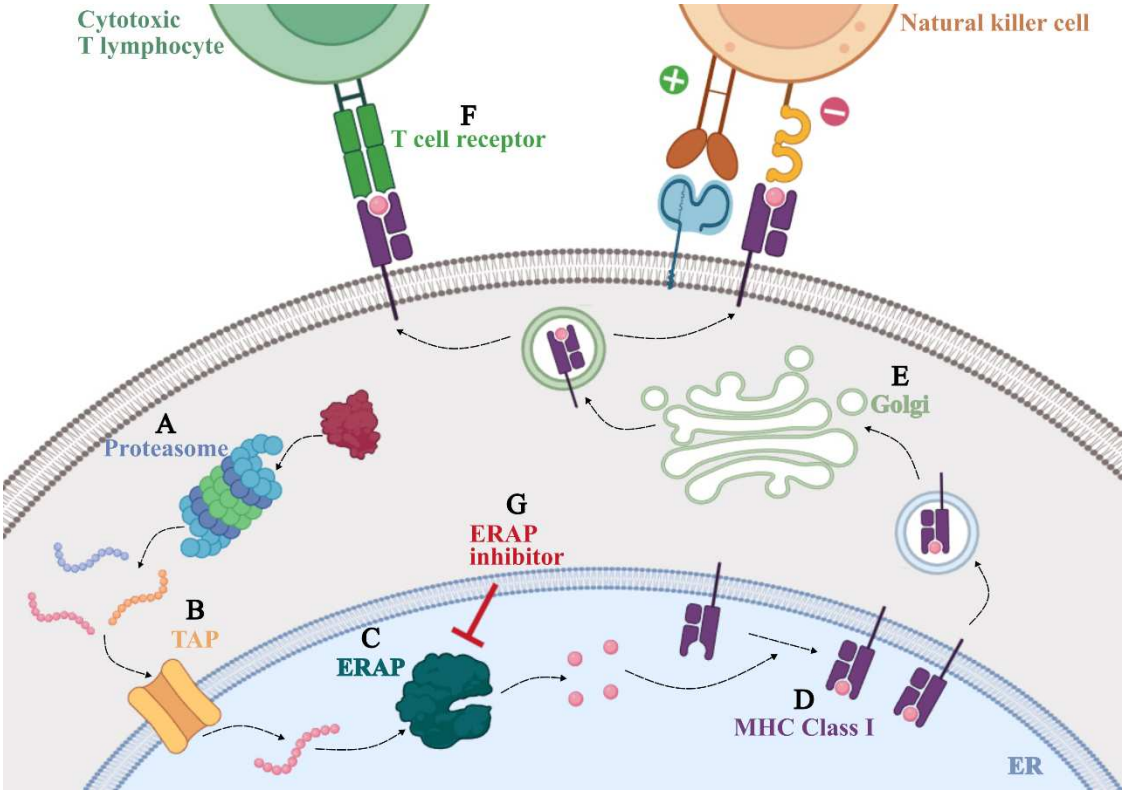


Figure 1. 3 Graphical representation of ERAP role in MHC class I antigen processing and presentation

(A) Proteins are first broken down into peptides by the proteasome, and (B) are then transported into the endoplasmic reticulum (ER) by transporter associated with antigen processing (TAP) proteins. (C) In the ER, the peptides undergo N-terminal trimming by ERAPs to the proper length to be (D) loaded onto major histocompatibility complex (MHC) class I protein. The MHC class I-peptide is then (E) transported to the cell surface through the Golgi apparatus, where it (F) triggers an antigen-specific immune response in cytotoxic T

lymphocytes and NK cells. (G) Small molecules targeting ERAPs could modify their trimming activity and alter the repertoire of presented antigens. Adapted from (Rosenbaum et al., 2021), created with BioRender.com

1.4 Research objectives

There is still an unmet need for potent and selective ERAP1 and ERAP2 inhibitors -with only one ERAP1 inhibitor in clinical phase- that could improve the therapeutic options for patients affected by cancer and autoimmune diseases. Mass spectrometry is among the tools that can support small molecules drug discovery programs to identify and test candidate ERAPs inhibitors and advance the research in this context.

The scope of this thesis is to support the drug discovery program of ERAPs inhibitors through the development and application of advanced mass spectrometry based analytical methods. With the projects that will be presented in the following chapters, we aimed to tackle the aspects of *valid models* and *valid compounds* of the drug discovery pipeline. These researchers were conducted for current and future application in PD, PK, and Tox studies (Figure 1. 4).

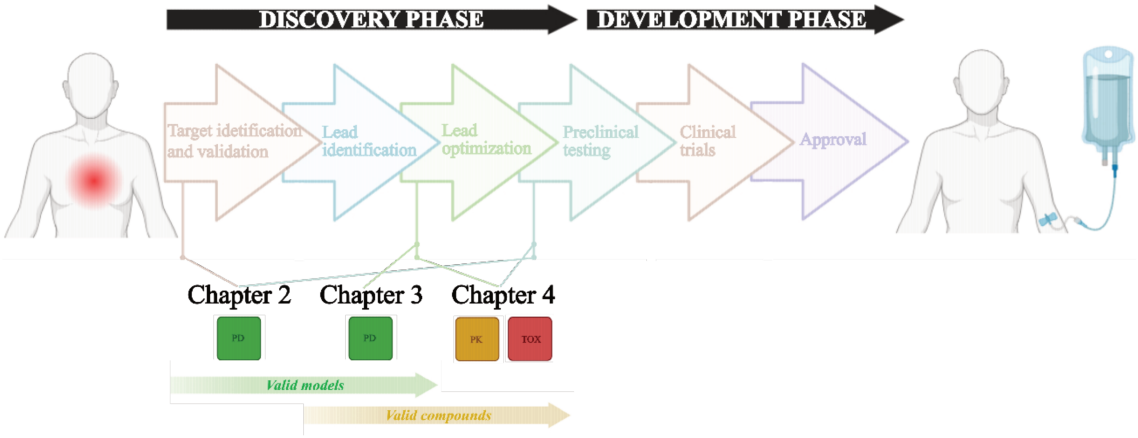


Figure 1. 4 Research projects presented in the chapters of this thesis and their role and application in the phases of drug discovery

An MSI-based untargeted metabolomic study was conducted to discover biomarkers associated to an ERAP-related disease preclinical model, the HLA-B27 transgenic rat, which develops ankylosing spondylitis. Chapter 2 presents and discusses challenges and opportunities of developing and applying this methodology to different tissue samples. A batch integration algorithm was developed and implemented as a method to enhance comparability among different runs. Significant differences were detected in intestinal compartment samples, suggesting a metabolic shift associated with the inflammatory bowel disease condition. Discriminative features between the disease condition and wild-type controls were molecularly annotated as metabolites and lipids, using LC-MS/MS. The spatial dimension was leveraged to interpret the role of these metabolites and lipids into disease

mechanisms by spatially correlating them with inflammatory markers. These could serve as candidate biomarkers for PD studies of ERAP inhibitors in this pathological condition.

Later, we aimed to develop a quantitative method to analyse the *in vitro* modulation of an ERAP1-dependent tumour antigen, GSW11, as a biomarker of efficacy during pharmacodynamic evaluation. We employed a murine colon carcinoma cell line (CT26) in which the pharmacological target, ERAAP, was knocked out, to validate the mechanism of action of a candidate ERAP1 inhibitor. In Chapter 3 we present the optimization and discuss the challenges of using LC-MS/MS technology to detect and quantify GSW11 from cells extracts. This project is critical to evaluate the target engagement of compounds, the validity of preclinical models, and the applicability of the technology.

In Chapter 4, we conducted a comprehensive investigation into the DMPK profile of a candidate ERAP2 inhibitor by investigating its whole-body biodistribution and combining *in vitro/in vivo* metabolism studies. Quantitative MSI was used to measure concentration in various organs and tissues after intraperitoneal (IP) administration, and the results were evaluated considering the concentration (exposure) required for target engagement and inhibition. Biotransformation and degradation products were identified, and their biodistribution within the organism was investigated. Overall, these results provide valuable insights into the candidate compound's therapeutic potential and safety profile, guiding future development efforts.

This thesis project was carried on within the EU founded Capstone ETN consortium, a multidisciplinary platform that involve experts from various disciplines to increase the research impact into ERAP aminopeptidases in various aspects of human and animal disease.

2 Chapter 2 - Development and application of an untargeted metabolomic study using MALDI mass spectrometry imaging for the discovery of region-specific biomarkers in the HLA-B27/hβ2m transgenic rat

The results of this chapter were included in Ponzoni A. et al. 2024 "An untargeted metabolomic study using MALDI-mass spectrometry imaging reveals region-specific biomarkers associated with bowel inflammation" under review before publication in Metabolomics

2.1 Rationale

Preclinical models are designed to closely replicate human conditions and predict the outcomes of administering a candidate therapeutic compound before moving to clinical trials. As the statistician George Box said "All models are wrong, but some are useful" (Box & Draper, 1987). To find application and utility in a drug discovery program, models need to be validated.

Beyond establishing the validity in mirroring human conditions, the identification of biomarkers it is crucial for the practical applicability of these preclinical models, particularly in pharmacodynamic studies. Biomarkers are measurable characteristics that can serve as indicators of pathological and physiological states, they can be followed upon treatment to study candidate compounds efficacy and safety.

In this chapter, we investigated via spatial analytical techniques whether the genetic modifications introduced to induce the disease state in a transgenic rat model resulted in meaningful molecular changes. Our research aimed to understand the disease mechanism and identify candidate biomarkers that are differentially expressed between diseased condition and healthy controls. These biomarkers could support the evaluation of the disease status and the development of new treatments.

We employed MALDI-MSI to study the disease model directly from affected tissues, providing spatially resolved information. The ability to simultaneously detect thousands of molecules in a single experiment and to directly mapping their distribution in tissues provides deeper insights into pathological processes and has great potential for biomarker discovery (Gessel et al., 2014). This technique has already proved an invaluable role for in situ metabolomic research and biomarker discovery in cancer research (Ma & Fernández, 2024).

Combining MSI with other imaging techniques, like microscopy or immunofluorescence, creates new opportunities to link metabolic data with histological and molecular changes. This multimodal approach allows for a more complete understanding of biochemical changes within the context of disease (Neumann et al., 2020).

Molecular annotation of some spatially resolved MSI signals was performed by LC-MS/MS. LC-MS/MS remains the gold standard for molecular annotation, since performing data-dependent MS/MS online is not feasible in MSI, and computational tools for the automatic annotation of MSI data are still lacking (Alexandrov et al., 2019; Palmer, 2017; Palmer et al., 2016).

2.2 Context

The HLA-B27/hβ2m transgenic rat develops a spontaneous multisystem inflammatory disease with clinical and histologic resemblance to the human spondylarthritis (SpA) (R. E. Hammer et al., 1990). This model, and specifically the F344 inbred background, presents inflammatory bowel disease (IBD), peripheral arthritis, spinal and skin lesions among its clinical features, making it a powerful preclinical model for ankylosing spondylitis (AS) and associated conditions (Milia et al., 2009; Taugor et al., 1999).

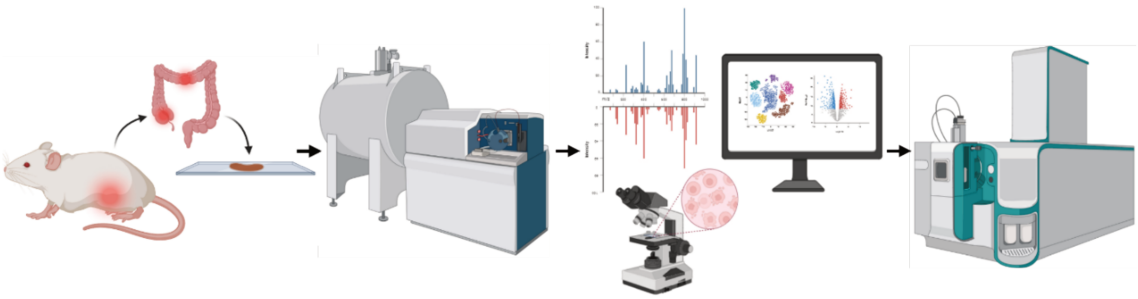
AS is a chronic immune-mediated inflammatory arthritis and the major subtype of the group of rheumatic diseases known as SpAs (Dougados & Baeten, 2011). The first clinical symptoms are inflammatory back pain, peripheral arthritis and enthesitis (Boel et al., 2022). Extra-articular manifestations, such as uveitis, or co-morbidities, like psoriasis and IBD, often affect AS patients and influence the prognosis and the treatment decision (Stolwijk et al., 2015). The incidence of IBD among AS patients is estimated between 5% and 10%, and up to 50% present subclinical gut inflammation. Therapeutic options for AS management include symptomatic relief, with line of treatment including physiotherapy, painkillers, anti-inflammatory and disease-modifying antirheumatic drugs (Braun & Sieper, 2002). Two decades ago, biological therapies targeting the tumour necrosis factor (TNF) were developed, improving AS control in patients unresponsive to conventional treatment (Braun et al., 2005). More recent progress led to the development of interleukin-17 (IL-17) inhibitors to manage AS in patients with inadequate response to TNF inhibitors (Baeten et al., 2015). However, a subset of patients remains unresponsive to biological treatments (Webers et al., 2023). The lack of effective treatments, and the resulting pain and disability experienced by some patients, contribute to the socio-economic burden of AS.

Genetic factors are strongly linked to AS susceptibility: the major histocompatibility complex gene encoding for the human leucocyte antigen B27 (HLA-B27) has long been associated with AS and to contribute to about 20% of its heritability (Brewerton et al., 1973; B. Chen et al., 2017; Schlosstein Lee et al., 1973). However, other genetic components have been uncovered, with genome-wide association studies (GWAS) revealing the correlation between AS and the genes encoding for the endoplasmic reticulum aminopeptidases 1 and 2 (ERAP1 and ERAP2) and the cytokine receptor IL23R (Robinson et al., 2015; Wellcome Trust Case Control Consortium et al., 2007). ERAPs could

act through a pathogenic mechanism involving aberrant peptide trimming and presentation, and its inhibition could be explored as a new therapeutic option for HLA-B27$^+$ AS patients.

Despite progresses made in understanding the genetic and clinical aspects of AS, diagnosis of the disease remains complex, with the current methodology relying in medical imaging and clinical criteria. Few available biomarkers to assess AS activity and progression are prevalently markers of inflammation, such as C-reactive protein (CRP), erythrocyte sedimentation rate, cytokines of the T helper 17 (Th17) pathway, vascular endothelial growth factor (VEGF) (Danve & O'Dell, 2015; Reveille, 2015). Moreover, to predict the manifestation of IBD in AS patients, the calprotectin protein, which is released by activated immune cells that migrate to inflamed joints or intestinal mucosa, is also measured in serum and stool (Klingberg et al., 2017). From a metabolomic perspective, several studies converged in finding that carbohydrate, lipid, and amino acid metabolisms are deregulated in this pathological condition (Bauset et al., 2021; Huang et al., 2022). The lack of studies performed directly on tissue and the need to further understand the pathogenic aspect underlying the interplay between AS and IBD provides opportunities for additional biomarker research.

2.3 Experimental design



Graphical abstract 1 MALDI-MSI untargeted metabolomic analysis on selected organs from HLA-B27 transgenic preclinical model for biomarker discovery

Compartments of the large intestine were extracted from HLA-B27 transgenic and wildtype control rats. Samples were cryosectioned and metabolomic and lipidomic profiles were acquired using a 7T-MALDI-FTICR mass spectrometer. Data analysis was performed to identify discriminating features between HLA-B27$^+$ and WT samples, which were annotated via LC-MS/MS. Interpretation of the biomarker's role in the disease context was hypothesized based on their spatial distribution within histological regions and colocalization with cellular subtypes. Created with BioRender.com

2.3.1 Experimental cohort and sample preparation

2.3.1.1 Study animals and collected organs

The experimental cohort consisted of HLA-B27 transgenic (HLA-B27$^+$) and wild type control (WT) rats (n=9 per group) provided by Prof Dubuquoy (University of Lille). Each collected organ, namely spleen, eyes, ankles, colon and cecum, was extracted at the 19th week of life, when a worsening of the

inflammatory phase is expected in this model. Organs were selected based on the known model features and the clinical symptoms in the human pathological condition. The spleen was collected to study the lymphoid system and therefore the immunological function. The eyes were selected to study the anterior uveitis, which is a human feature associated to the pathology. Ankle joints were collected to study peripheral arthritis, one of the first and most important clinical symptoms of AS. Colon and cecum compartments of the large intestine were extracted to characterize the IBD condition. All the samples were collected, snap frozen in dry ice and stored at -80°C until use.

2.3.1.2 Sample preparation for MSI analysis

Sections from spleen, colon, and cecum were obtained by cryosectioning to 10 µm thickness in a cryo-microtome at -20°C. Eyes were embedded in 1.5% carboxymethyl cellulose (CMC) in water and incubated at -80°C until solidification prior sectioning at 10 µm in a cryo-microtome at -20°C. Consecutive tissue sections from the 18 samples per compartment were randomly thaw-mounted on indium tin oxide (ITO) microscope slides, ensuring an even distribution of samples per group for each analysis. On each ITO slides analyses, a section of "quality control" tissue, liver homogenate spiked with a small molecule ionizable in positive (olanzapine) or negative (rutin) mode, was placed and acquired. Ankle joints were embedded in 5% CMC in water and incubated at -80°C until solidification. Sectioning was performed collecting 10 µm thick sections on tape using a Cryostat at -25°C. On-tape tissue sections were adhered to the ground steel target MALDI plate with the support of double-sided tape.

2.3.1.3 Sample preparation for histology and immunofluorescence

Colon and cecum sectioning was performed by cryosectioning to 5 µm thickness in a cryo-microtome at -20°C. Tissue sections were thaw-mounted onto SuperFrost Plus™ Adhesion slides.

2.3.1.4 Sample preparation for LC-MS/MS analysis

Four to six 10 µm thick cryosections per sample were transferred in one 2 mL tube pre-filled with beads (Thermo Fisher Scientific, Waltham, MA, USA) per group: WT colon, HLA-B27$^+$ colon, WT cecum and HLA-B27$^+$ cecum samples. Aqueous and organic extracts were obtained following the protocol from (Want et al., 2013). Briefly 0.5 mL of prechilled 1:1 MeOH:H2O was added to the tube. Samples were homogenized and centrifuged. The supernatant was dried using SpeedVac SPD120. 0.5 mL of prechilled 3:1 CH_2Cl_2:MeOH was added to the solid precipitate. Samples were homogenized again. After centrifugation, the supernatant was dried in fume hood (6 hours, RT). 200 µL of prechilled 1:1 MeOH:H_2O was added to the remaining aqueous solvent and dried.

2.3.2 MALDI MSI method development and QCs

MALDI MSI analysis was performed in both negative and positive ionization modes. Commonly used MALDI matrices were selected to acquire mass profiles ranging from 75 to 1200 Da. 2,5-dihydroxybenzoic acid (DHB) was employed for analysis in positive ionization mode, while 1,5-diaminonaphtalene (DAN) for the analysis in negative. HTX TM-sprayer was used to evenly deposit the MALDI matrices over the samples.

Ion transfer and analyser parameters were modified during method optimization for each tissue type and polarity, by directly assessing the resolution and population of the spectra. The sample analysis was designed to randomly acquire images across the slides, considering both position and grouping. Lateral spatial resolution was selected based on tissue dimension and histological layers thickness. Spleen, eyes and ankle joints were analysed at 80 μm resolution, while colon and cecum at 40 μm. During the images acquisition, online calibration was active, and a list of matrix peaks and prevalent endogenous lipids was used to recalibrate the spectra. Data reduction factor was set at 97% because of the large data generated by the MALDI when imaging regions. Save reduced profile spectrum is an option that allows for the compression of acquired data, including both the spectrum and the image, to prevent the size from becoming excessively large during imaging acquisitions.

Quality checks were conducted after each analysis to ensure the validity and usability of the collected data. Mean signal intensity of matrix peaks, *m/z* 273.04 for DHB and *m/z* 315.16 for DAN, and QCs, *m/z* 313.14 for olanzapine and *m/z* 609.15 for rutin, were compared inter-analysis to calculate global relative standard deviation percentile. An RSD% below 35% was considered acceptable. Calibration error (Δ part per million), defined as the difference between theoretical and experimental *m/z* of matrix peaks and endogenous lipids, was calculated, with acceptable threshold set at Δppm lower than 3.

2.3.3 Unsupervised data analysis

Proprietary software, Multimaging™ v1.2.8.4, was used for images visualization and data analysis. Data pretreatment included peak picking, creation of matrix of intensities for each peak and data normalization. Built-in modules were used for clustering and statistical comparison. We applied one of the best nonlinear dimensionality reduction techniques for MSI called uniform manifold approximation and projection (UMAP) to the pixels of an MSI dataset and their intensities for each *m/z* and project the results into a 3D space (McInnes et al., 2018; Smets et al., 2019). Each pixel was coloured based on its position in the space, with similar colours indicating biological proximity and aiding cluster identification. K-means clustering using Elbow method and cosine distance was then applied to the reduced data to identify biochemical clusters (Winderbaum et al., 2015).

2.3.4 Statistical comparison

To identify mass features that distinguish the pathological state in the transgenic model from the WT control, we conducted a statistical analysis directly comparing samples from both groups. We performed statistical comparison via Student's *t*-test and adjusted the p-values using the Benjamini-Hochberg (BH) correction.

2.3.5 LC-MS/MS method development

We prepared aqueous and organic tissue extracts from pooled tissue sections obtained from the two compartments, cecum and colon, and groups, HLA-B27$^+$ and WT, separately. We employed a two-step extraction protocol to enhance the extraction of polar and apolar metabolites (Want et al., 2013). We developed and applied two LC methods: one for the detection and separation of small polar metabolites in the aqueous extracts and one for the detection of apolar lipids in the organic extract, to obtain good coverage of the molecular classes that composed the final list of 275 features. For the analysis of metabolites, we developed a method which employed an ethylene bridged hybrid (BEH) hydrophilic interaction liquid chromatography (HILIC) stationary phase and a gradient from strongly organic, 95% ACN, to aqueous, 25 mM ammonium carbonate in water. Additionally, we performed, on the same extracts, the analysis using a BEH C_{18} reversed-phase column with a gradient from highly aqueous, 0.1% FA in water, to highly organic, 0.1% FA in ACN. This strategy was selected to take advantage of the complementarity of these two analyses which better retain and separate highly polar metabolites, in the case of the HILIC column, and moderately polar metabolites, in the case of the C_{18} column.

For the lipidomic LC protocol, we adapted the gradient and mobile phases from a previous research (Sarafian et al., 2014). We used isopropanol in mobile phase B to improve the elution of apolar lipids, and added ammonium formate and formic acid to help maintain a stable pH and support ionization in both positive and negative modes. We tested multiple columns, BEH C_{18}, BEH C_8, charged surface hybrid (CSH) C_{18}, and monitor the performance in eluting and separating a mixture of standard lipids (EquiSPLASH, Avanti Polar Lipids, Alabaster, AL, USA), which we selected for including different classes that could be present in our samples.

Mixture Component	Formula	Molecular Weight	$[M-H]^-$	$[M+H]^+$	$[M+NH_4]^+$
15:0-18:1(d7) PC	C41H73D7NO8P	752.6061		753.6134	
18:1(d7) Lyso PC	C26H45D7NO7P	528.3921		529.3994	
15:0-18:1(d7) PE	C38H67D7NO8P	710.5591	709.5519	711.5664	
18:1(d7) Lyso PE	C23H39D7NO8P	486.3451	485.3379	487.3524	
15:0-18:1(d7) PG	C39H68D7O10P	741.5537	740.5464		759.5875
15:0-18:1(d7) PI	C42H72D7O13P	829.5698	829.5625		847.6036
15:0-18:1(d7) PS	C39H67D7NO10P	754.5490	753.5417	755.5562	
15:0-18:1(d7)-15:0 TAG	C51H89D7O6	811.7646			829.7985
15:0-18:1(d7) DAG	C36H61D7O5	587.5506			605.5844
18:1(d7) MAG	C21H33D7O4	363.3366		364.3429	381.3704
18:1(d7) Chol Ester	C45H71D7O2	657.6441			657.6779
d18:1-18:1(d9) SM	C41H72D9N2O6P	737.6397		738.6470	
C15 Ceramide-d7	C33H58D7NO3	530.5404	520.5331	531.5477	

Table 2. 1 List of lipids of the standard mix used for validating LC-MS lipidomic method and selecting the column.

Each discriminative feature present in our final list was investigated in the sample's extracts in which it was detected by MALDI-MSI analysis. We employed a multiple reaction monitoring (MRM) mass spectrometry method and set the list of precursor masses to be fragmented. The investigation, as for MSI analysis, was carried out in both polarities.

2.4 Results

2.4.1 Unsupervised data analysis showed differences in the biochemical content of HLA-B27$^+$ samples compared to WT

2.4.1.1 Biochemical profiles changes between HLA-B27$^+$ and WT in features intensity levels

To investigate the biochemical changes, we acquired and compared mass profiles acquired from disease-bearing rat's samples (HLA-B27$^+$) with that of healthy controls (WT) for each compartment under analysis. By comparing the spectra from the two groups, we started to observe some differences in features intensity (Figure 2. 1). Resolving power (Rp) of the analyses was over 120k at *m/z* 200 and over 50k at *m/z* 400.

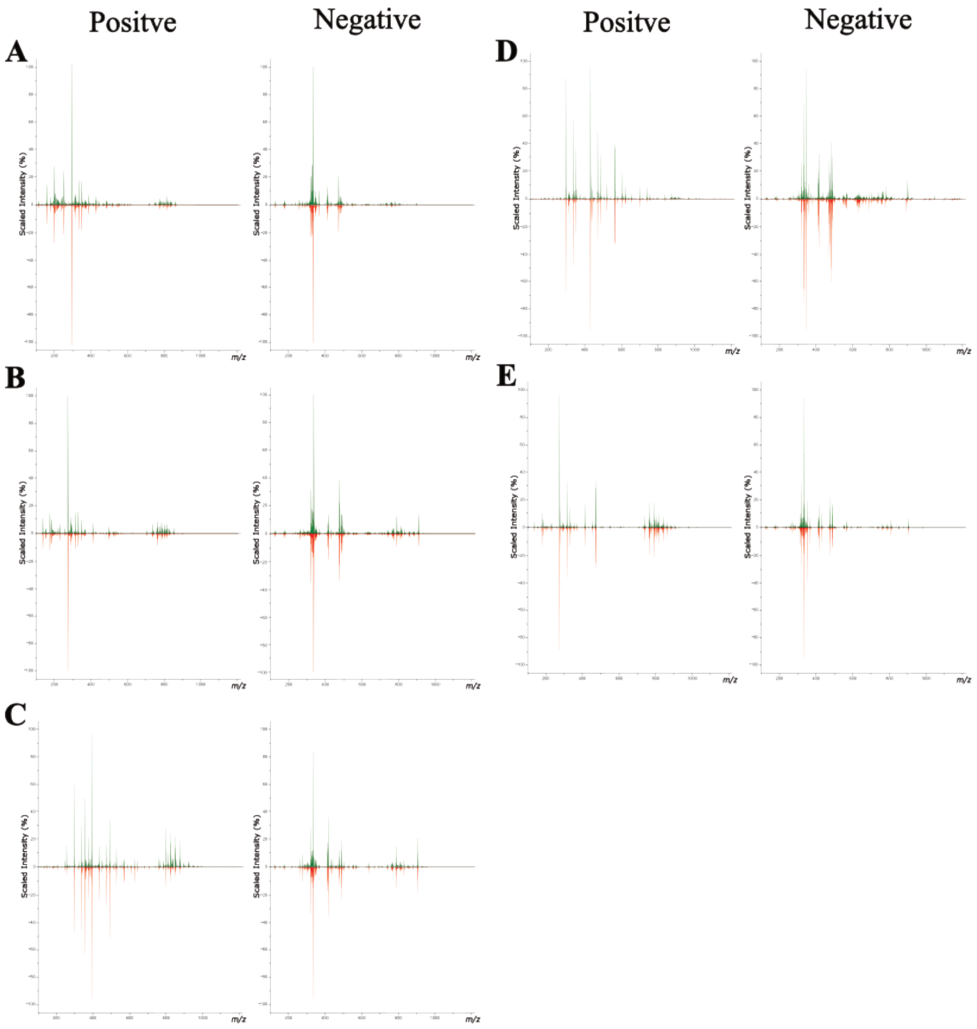


Figure 2. 1 Metabolomic and lipidomic mass profiles from HLAB27^{+} and WT samples

Spectra ranging from 75 to 1200 Da, of colon (A), cecum (B), spleen (C), ankle (D) and eyes (E) samples acquired by MALDI-MSI in positive and negative ionization mode. Each spectrum compares the profile acquired from HLA-B27^{+} (green) and WT (red) showing the different features intensity between the groups.

We performed an exploratory data analysis on each run to investigate whether differences could be identified in an unbiased fashion. *K*-means clustering applied on MSI data projected in a UMAP 3D plot was used to identify biochemical clusters. The first attempt to perform unsupervised data analysis revealed that slides acquired on different days, even if they passed the QC and validation criteria, and despite study design to attenuate batch effect, proved to be not comparable in the analysis, since different clusters were identified inter-batches (Figure 2. 2 A-F). Only the datasets acquired from colon and cecum compartments already showed the presence of clusters discriminating between the two groups intra-batch (Figure 2. 2 A-D).

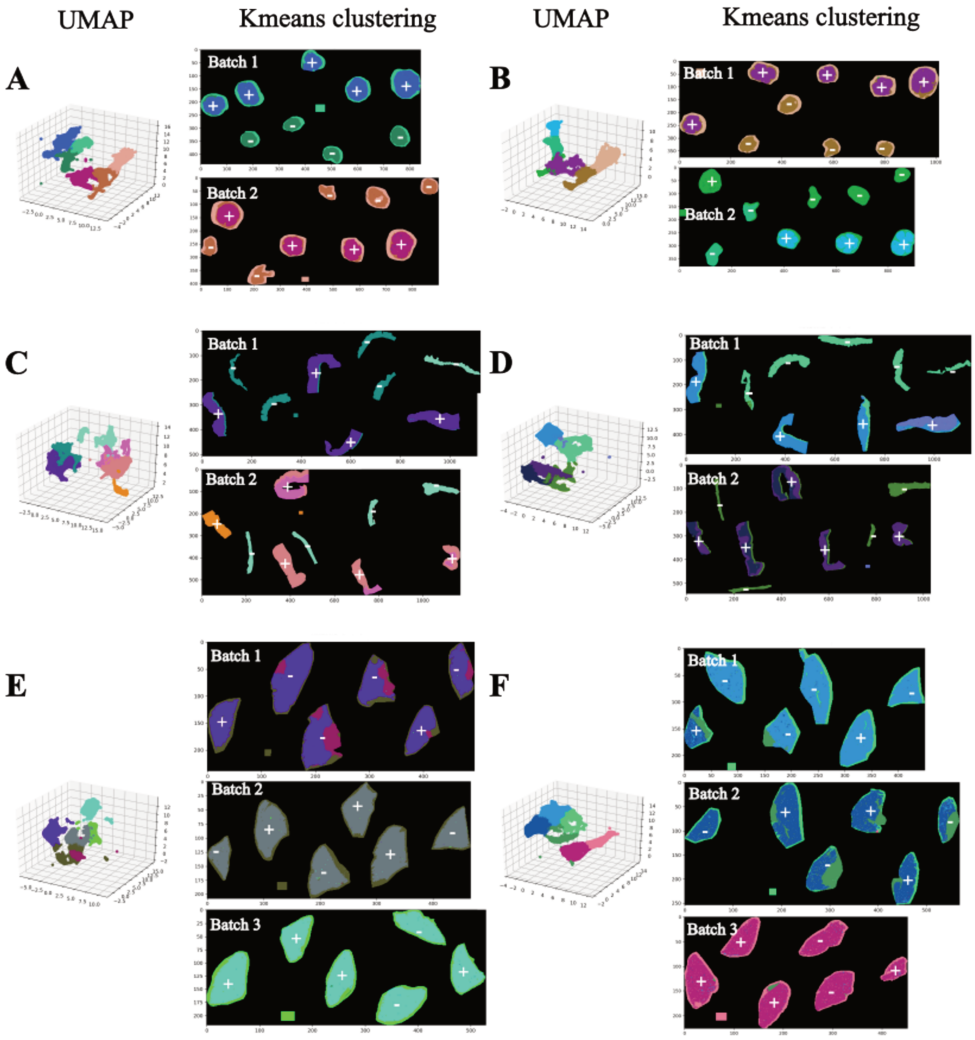


Figure 2. 2 MSI data projected in UMAP tri-dimensional plot and clustered by K-means

Data analysis was performed on dataset from colon analysed in positive (A) and negative (B) ion mode, cecum analysed in positive (C) and negative (D) ion mode and spleen analysed in positive (E) and negative (F) ion mode. Pixels on the image are coloured based on the membership clusters. HLA-B27^{+} samples are identified with the symbol "+", while WT samples with the symbol "-".

2.4.1.2 Implementation and validation of an integration algorithm to compensate for batch effect

Technical artefacts can be introduced by the operator or the instrument and could induce inter-day variability. Data processing can attenuate these effects, *i.e.* with the implementation of batch integration algorithm. Drawing upon data analysis tools used for large-scale single-cell RNA-sequencing, we found an extremely fast and graph-based data integration algorithm, batch balanced k nearest neighbours (BBKNN), that we adapted for MSI data (Figure 2. 3) (Polański et al., 2020). BBKNN finds a smaller set of *k* nearest neighbours for each cell, or pixel in the case of MSI data, within each batch individually, rather than across the entire dataset.

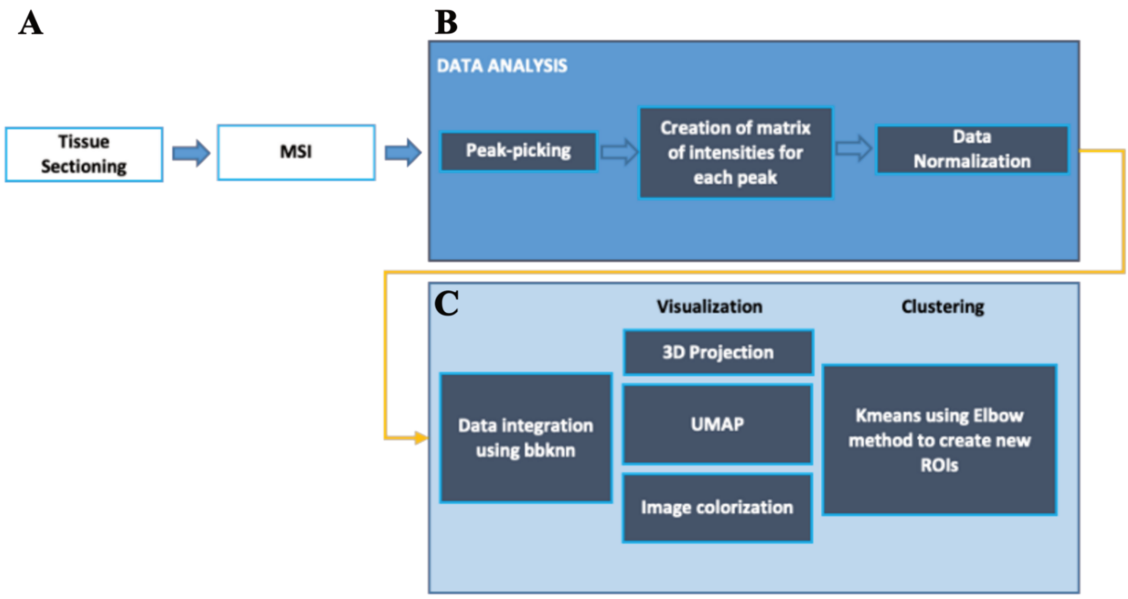


Figure 2. 3 Schematic workflow for data analysis performed on MSI data to identify biochemical clusters (A) The flow starts with sample preparation and acquisition. (B) Pretreatment of the data is performed prior data reduction and clustering. (C) Batch integration is integrated at this stage.

We validated the algorithm using a tumour cohort from a previous metabolomic project. The implementation of the batch integration algorithm before data reduction and clustering proved to be capable of integrating data from samples analysed on different days (Figure 2. 4 A-D).

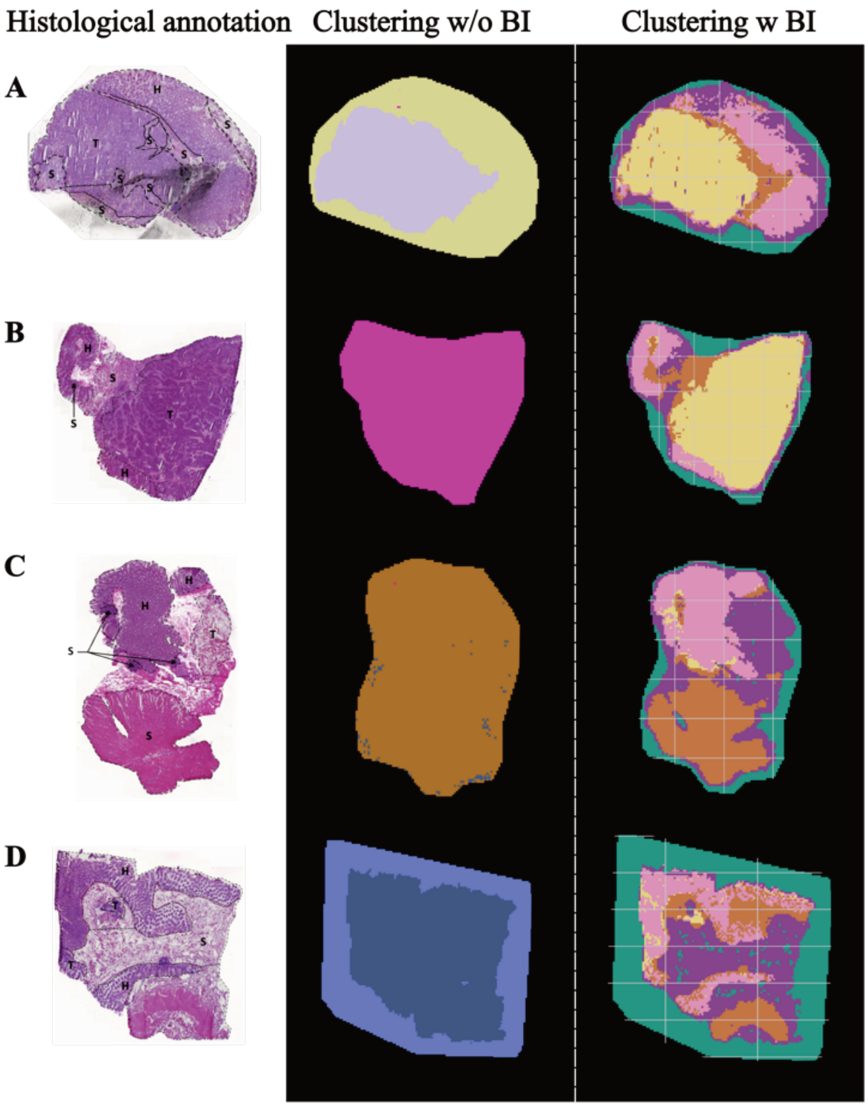


Figure 2. 4 Cancer samples used for batch integration algorithm validation
(A-D) For each representative sample H&E coloration and segmentation in "H" healthy tissue, "S" stroma, and "T" tumour regions, clusters on sections without batch integration and with batch integration are showed

Moreover, the clusters detected on tissue samples correlated with histological regions annotated by a trained pathologist. To assess how accurately the clustering method matched these annotated regions, we used multi-class receiver operating characteristic (ROC) curves. A ROC curve is a graph that shows a model's ability to distinguish between different categories, tumour and stroma regions in this case, by balancing sensitivity (how well the method correctly identifies true regions) and specificity (how well it avoids incorrect identifications). We compared the clustering performance with and without batch integration, to see how closely each approach matched the pathologist's annotation. For both tumour (Figure 2. 5 A) and stroma (Figure 2. 5 B) regions, the area under the curve (AUC) score was higher when batch integration was used, indicating better alignment with the annotated regions.

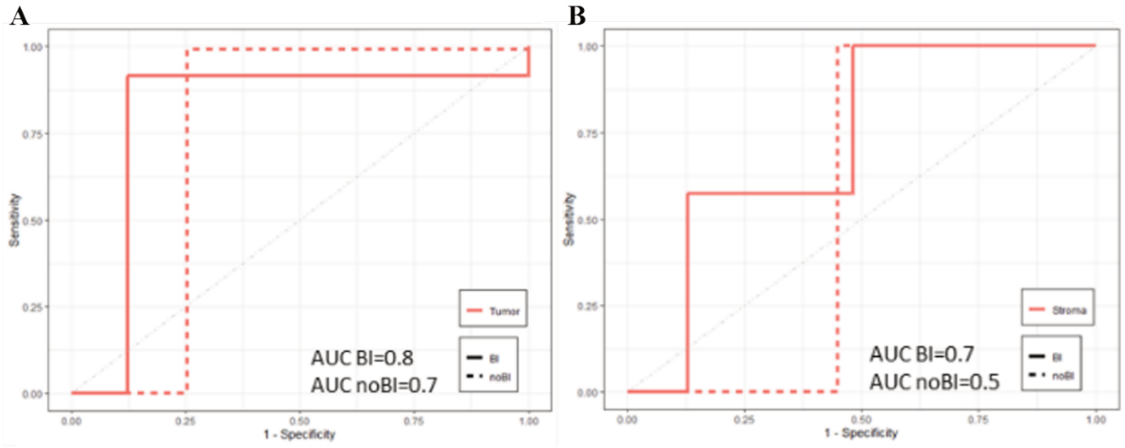


Figure 2. 5 Multi-class ROC curves showing clusters correlation performance

Tumour (A) and Stroma (B) regions correlation with clusters identified using batch integration (solid line) or without batch integration (dashed line)

2.4.1.3 Implementing batch integration algorithm facilitated the identification of biochemically relevant cluster inter batches

In the experimental cohort from HLA-B27$^+$ rats, the batch integration algorithm enabled the detection of clusters inter-batch on each compartment analysed in both ion modalities. Clusters evenly distributed inter-batches for spleen (Figure 2. 6 E&F), eyes (Figure 2. 6 G&H) and ankles (Figure 2. 6 I&J), though no clusters differentiating HLA-B27$^+$ from WT samples were detected in these samples. In colon (Figure 2. 6 A&B) and cecum (Figure 2. 6 C&D) compartments, where clusters discriminative for the two groups were already detected prior batch correction, we observed multiple clusters with homogeneous distribution inter-batches.

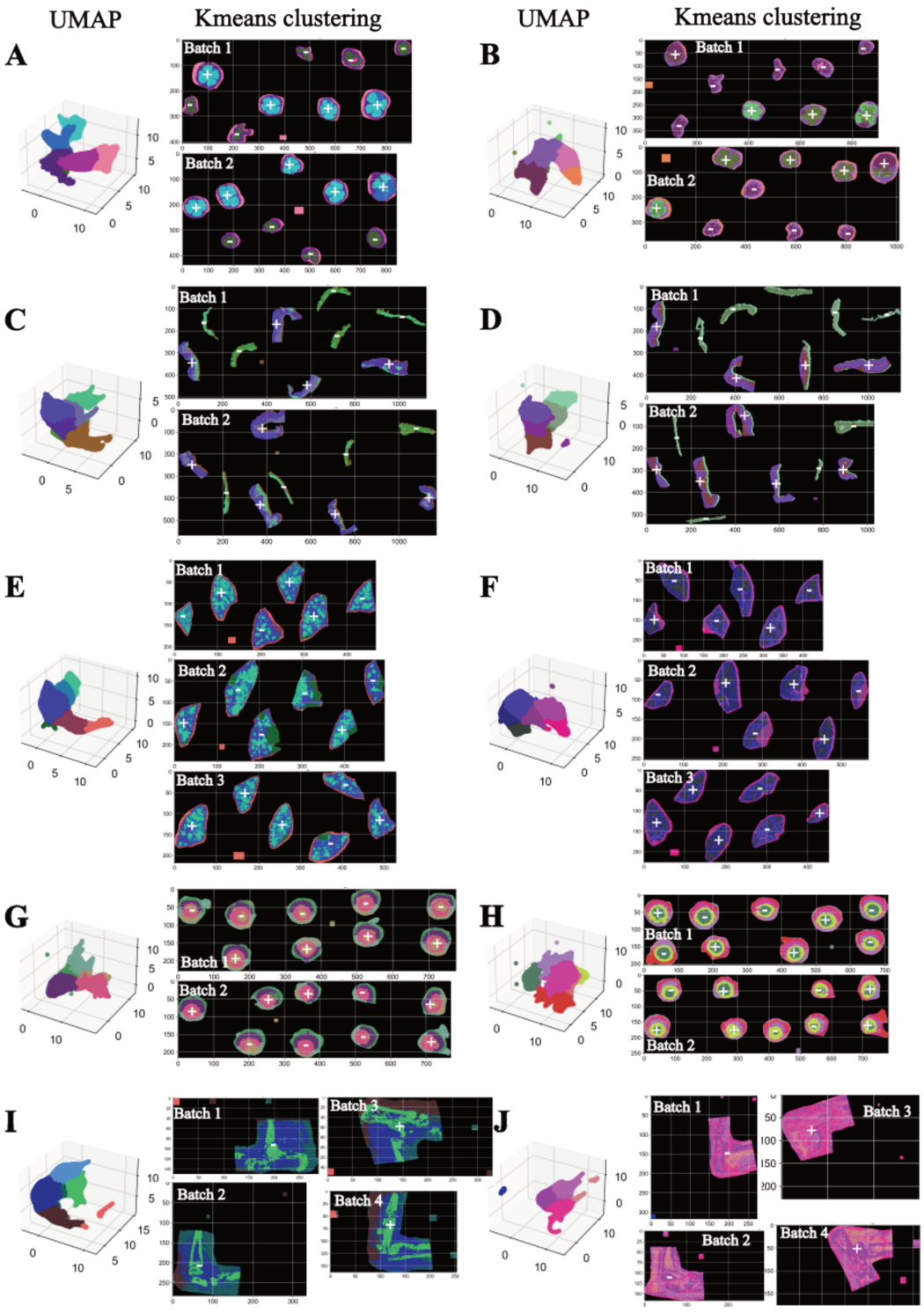


Figure 2. 6 Batch corrected MSI data projected in UMAP 3D plot

Colon samples analysed in positive (A) and negative (B) ion mode, cecum samples analysed in positive (C), and negative (D) ion mode, spleen samples analysed in positive (E), and negative (F) ion mode, eyes samples analysed in positive (G), and negative (H) ion modes and ankle samples (4 representative samples displayed) analysed in positive (I), and negative (J) ion modes. Pixels on the image are coloured based on the membership clusters. HLA-B27^{+} samples are identified with the "+" symbol, while WT samples with the "-" symbol

We overlaid the H&E-stained images of the analysed tissue sections with MSI molecular images and clusters projection on tissue. By closely observing the spatial distribution of clusters in relation to histology, we noted a spatial correlation between the clusters and the segmented histological layers in each compartment (Figure 2. 7).

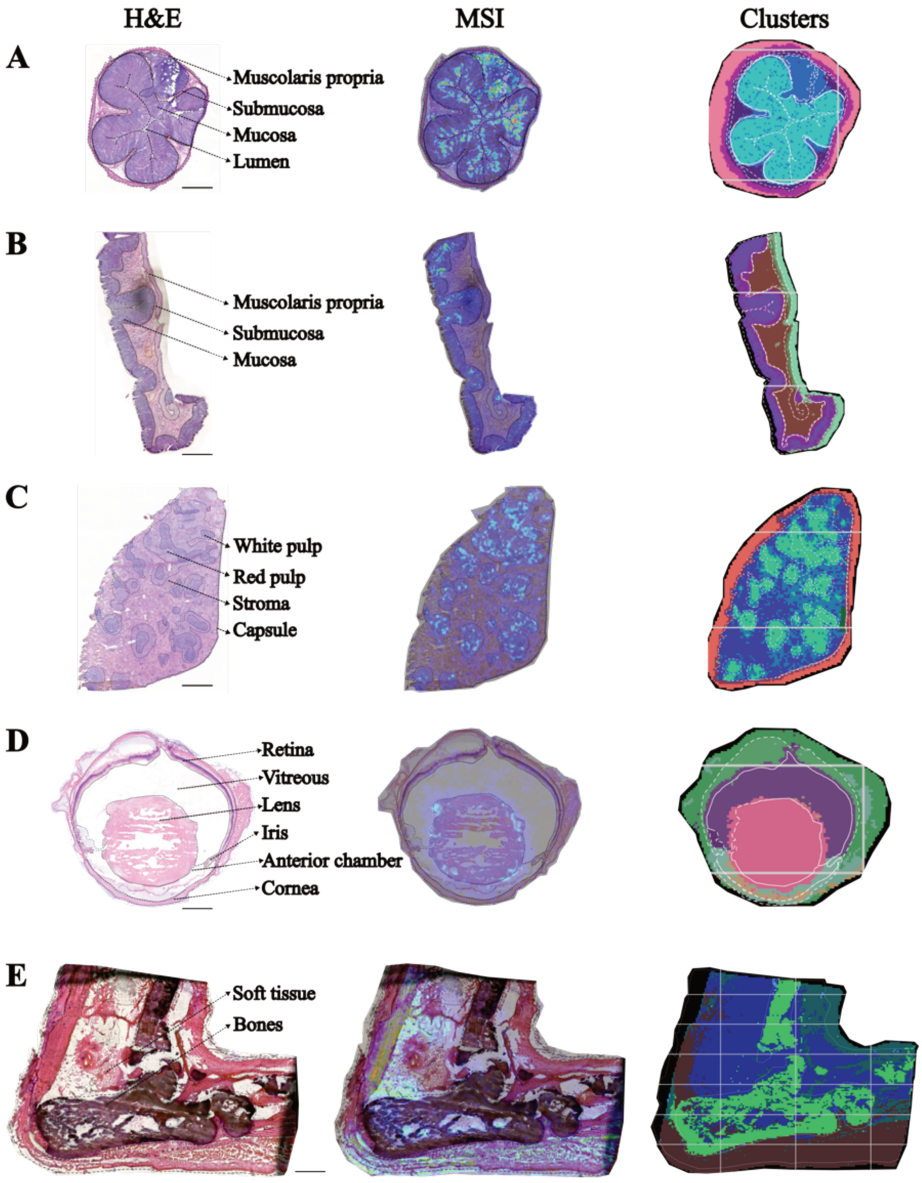


Figure 2. 7 H&E scans overlaid with MSI molecular images and clusters projection

One representative sample per compartment, colon (A), cecum (B), spleen (C), eyes (D), and ankle (E) stained with H&E on which histological regions were segmented. Scale bars 1 mm.

2.4.1.4 Clusters discriminative of the two groups were identified in samples from bowel compartment

Clusters differentiating HLA-B27^{+} samples from WT ones were only detected in the bowel compartments, colon and cecum. Clusters 1 and 2 in colon (Figure 2. 8 A&B) and cluster 1 in cecum

(Figure 2. 8 C&D) were present across all samples of each run, regardless of the groups. In addition, cluster 2 on cecum samples and cluster 3 on colon samples were specific to the group WT, and clusters 3, 4 and 5 on cecum samples and clusters 4 and 5 on colon samples were specific to the HLA-B27^{+} group. These discriminative clusters, distributed in a region-dependent manner, were detected in the mucosa and submucosa histological layers of the large intestine.

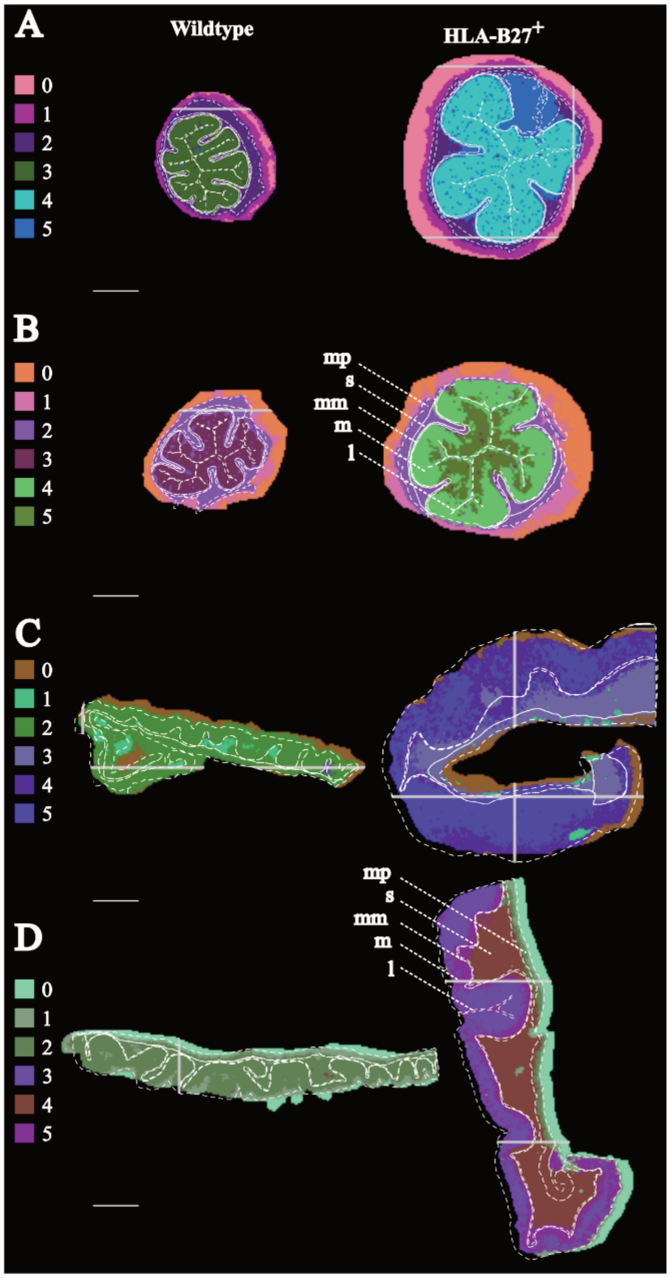


Figure 2. 8 K-means clustering by spectral or biochemical pattern similarities on colon and cecum samples in both ionization modes

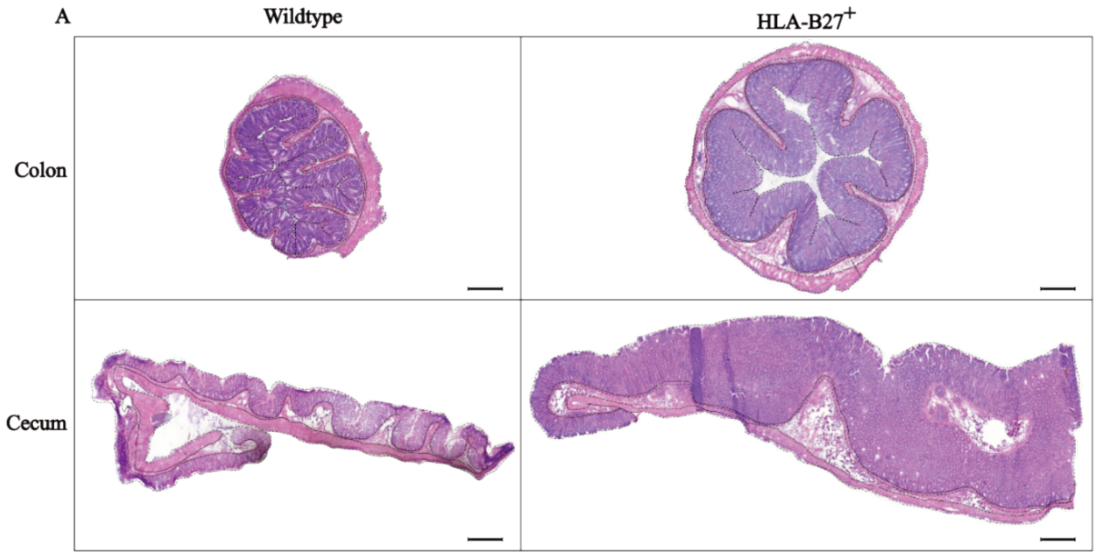
Representative images of clusters distribution on (A) colon analysed in positive ionization mode, (B) colon analysed in negative ionization mode, (C) cecum analysed in positive ionization mode and (D) cecum analysed

in negative ionization mode. The large intestine layers are marked on each section ad follow "mp" for muscularis propria, "s" for submucosa, "mm" for muscularis mucosae, "m" for mucosa and "l" for lumen. Scale bars 1 mm.

2.4.2 Histopathology and immunofluorescence analysis validated spontaneous colitis in HLA-B27$^+$ rats

2.4.2.1 Qualitative histopathological evaluation of the inflammation in HLA-B27$^+$ samples

Further analysis was focused on colon and cecum compartments, since they proved to be the ones expressing the most evident biochemical differences between HLA-B27$^+$ and WT groups. Tissues were extracted from the animal model at the 19th week of life. At this age, the spontaneous intestinal inflammation worsens and can be assessed histologically (Asquith et al., 2016). To ensure the suitability of the experimental cohort for linking the biochemical and metabolic shift to the inflammatory condition, adjacent transversal tissue sections from cecum and colon samples were stained with haematoxylin and eosin (H&E) and underwent histopathological examination (Figure 2. 9 A). Qualitative examination of macroscopic features revealed layers thickening, swelling, oedema and ulceration. Microscopically, we observed epithelial distortion and immune cells infiltration into the lamina propria of HLA-B27$^+$ samples compared to WT (Figure 2. 9 B).



B

Macroscopic findings	Microscopic findings
•Thick colonic wall with visible sign of inflammation such as swelling and oedema •Presence of ulcer on the mucosa surface •Narrowing of the colon lumen	•Distortion of the glandular structure •Infiltration inflammatory cell in the lamina propria •Fibrous tissue in the submucosa and muscularis propria

Figure 2. 9 Histopathological evaluation of colon and cecum HLA-B27$^+$ and WT samples

(A) H&E coloration on colon and cecum from HLA-B27$^+$ rats and wildtype controls. Scale bars 500 µm. (B) Table summarizing histopathological findings on HLA-B27$^+$ samples.

2.4.2.2 Staining of immune cells from innate and adaptive immune response

Inflammatory cells were infiltrating the connective tissue of the mucosa, known as the lamina propria, of HLA-B27$^+$ samples, therefore we conducted an immunofluorescence (IF) staining for markers of cells from innate and adaptive immune responses. We developed a two-colour staining protocol targeting CD68, marker for macrophages, and CD3, marker for T lymphocytes, with DAPI to counterstain the nuclei. To verify antibody specificity, we compared the signals from tissue sections stained with the primary mouse anti-rat CD68 [KP1] antibody coupled with the fluorescent dye Alexa Fluor 647 (Figure 2. 10 A) to signals from adjacent tissue sections stained with a mouse isotype control antibody coupled with the same fluorophore (Figure 2. 10 B). To test the rabbit anti-rat CD3 [SP162] antibody, which is not directly coupled to a fluorophore, we compare the signals from sections stained with the primary antibody followed by a secondary donkey anti-rabbit secondary antibody coupled with the fluorescent dye Alexa Fluor 594 (Figure 2. 10 C), to signals from adjacent sections stained with a primary rabbit isotype control followed by the same secondary antibody (Figure 2. 10 D) or from sections stained with the secondary antibody only (Figure 2. 10 E). An isotype control is an antibody that matches the primary antibody in terms of isotype (e.g., IgG) but lacks specificity for the target antigen; it serves as a negative control to determine if the observed signal results from specific binding of the primary antibody or non-specific binding. Specific signals at the nuclear level were detected in stained tissue sections and absent in sections stained with control antibodies.

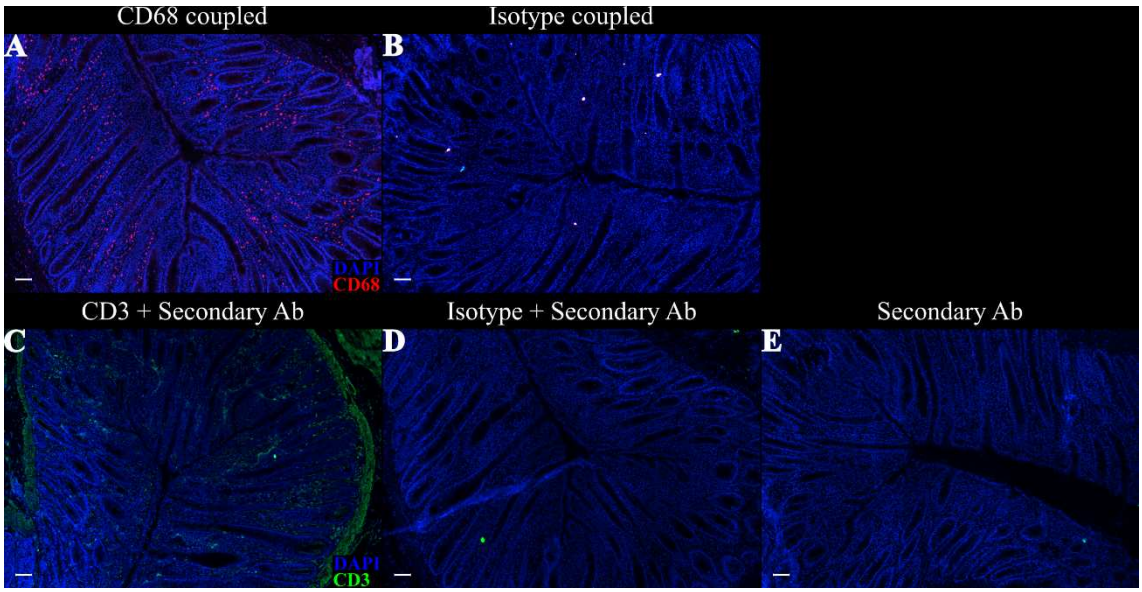


Figure 2. 10 Immunofluorescence staining for CD68$^+$ macrophages (A) and CD3$^+$ T lymphocytes (C) on large intestine transversal tissue sections counterstained with nuclear marker DAPI

(B) Mouse isotype control coupled with AF647 was used as a control of specific $CD68^+$ staining. (D) Rabbit isotype control followed by incubation with a secondary donkey anti-rabbit secondary antibody coupled with AF594 or incubation with secondary antibody only were used as control of specific $CD3^+$ staining.

We optimized antibodies dilution and incubation times to obtain the final protocol, detailed in the experimental section, for optimal signal detection. We assayed large intestine samples with the IF staining protocol and detected $CD3^+$ and $CD68^+$ cells in colon sections from WT (Figure 2. 11 A-C) and HLA-$B27^+$ (Figure 2. 11 D-F) groups and cecum samples from WT (Figure 2. 11 G-I) and HLA-$B27^+$ (Figure 2. 11 J-L) groups. The higher density of positive cells was detected in the lamina propria of HLA-$B27^+$ samples (Figure 2. 11 F, L black arrows).

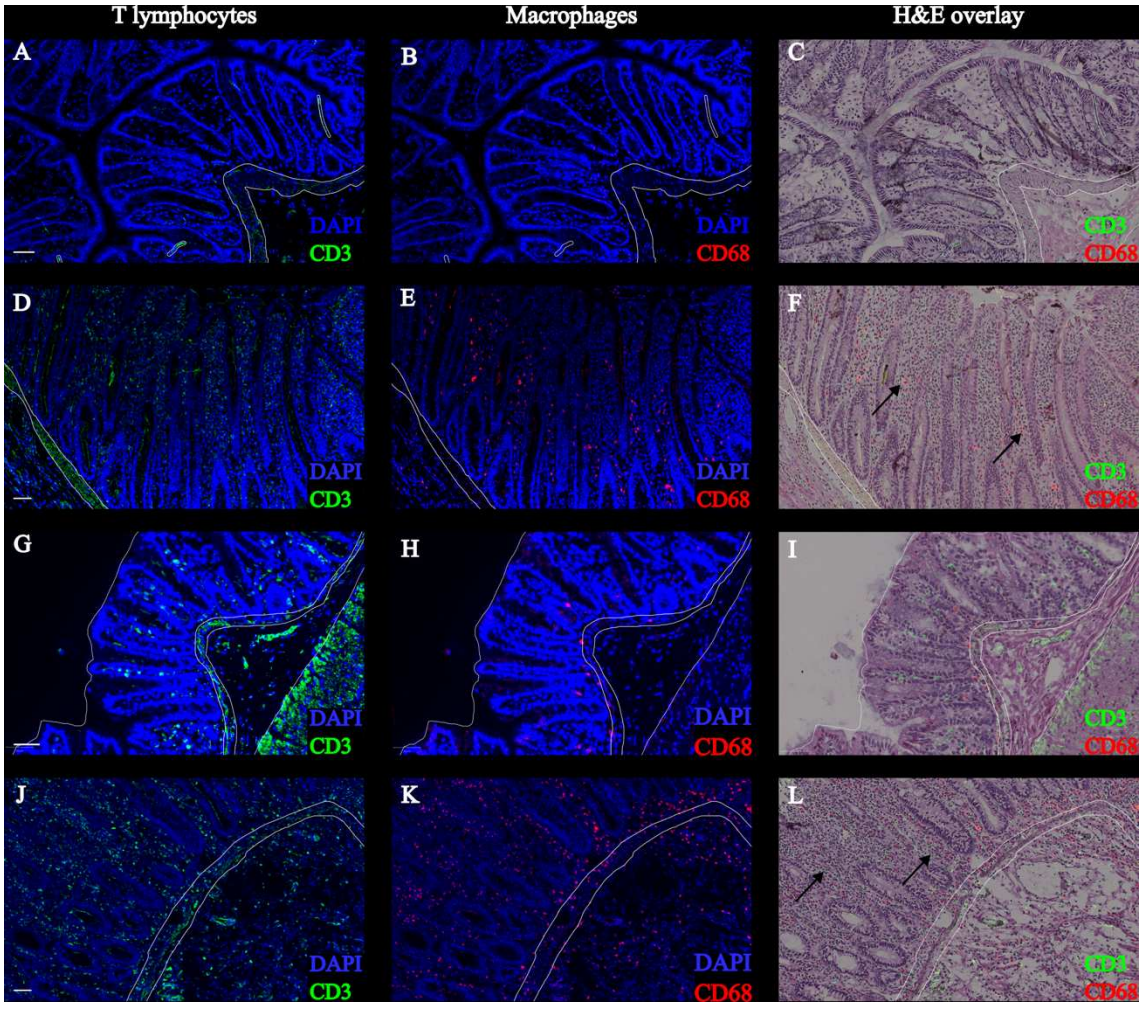


Figure 2. 11 Double immunofluorescence staining for $CD3^+$ T lymphocytes and $CD68^+$ macrophages on large intestine transversal sections counterstained with nuclear marker DAPI

Colon and cecum samples from the HLA-$B27^+$ group (D-F, J-L) present more $CD3^+$ and $CD68^+$ cells compared to colon and cecum samples from the WT group (A-C, G-I). The overlay of immunostaining with H&E coloration on the same sections (C, F, I, L) enable the localize most of the immune cells within the lamina propria (black arrows). Scale bars 50 µm.

2.4.3 Statistical analysis identified discriminating features between HLA-B27$^+$ and WT colon and cecum samples, annotated as metabolites and lipids

2.4.3.1 Supervised statistical comparison between ROIs

We selected regions of interest (ROIs) on sample tissue sections, framing the histological layers mucosa and submucosa, where the discriminative clusters distributed, and performed statistical comparison (Figure 2. 12 A). The statistical analysis identified a list of 275 features discriminating the HLA-B27$^+$ group from WT, with an adjusted p-value <0.001 and a fold change lower than 0.5 or greater than 2 (Figure 2. 12 B-E). A total of 207 features were more prevalent in the HLA-B27$^+$ group, and 68 were more prevalent in the WT group. Out of the 207 features associated with the HLA-B27$^+$ samples, 91 were identified in colon and 116 in cecum (Figure 2. 12 F). Out of the 68 features associated with WT samples, 40 were identified in colon and 28 in cecum compartments (Figure 2. 12 G). Notably, 26 of the identified discriminative features were shared between colon and cecum, 19 being more abundant and 7 being less abundant in HLA-B27$^+$ compared to WT samples, showing a partially similar disease impact on the biochemical patterns of the two compartments.

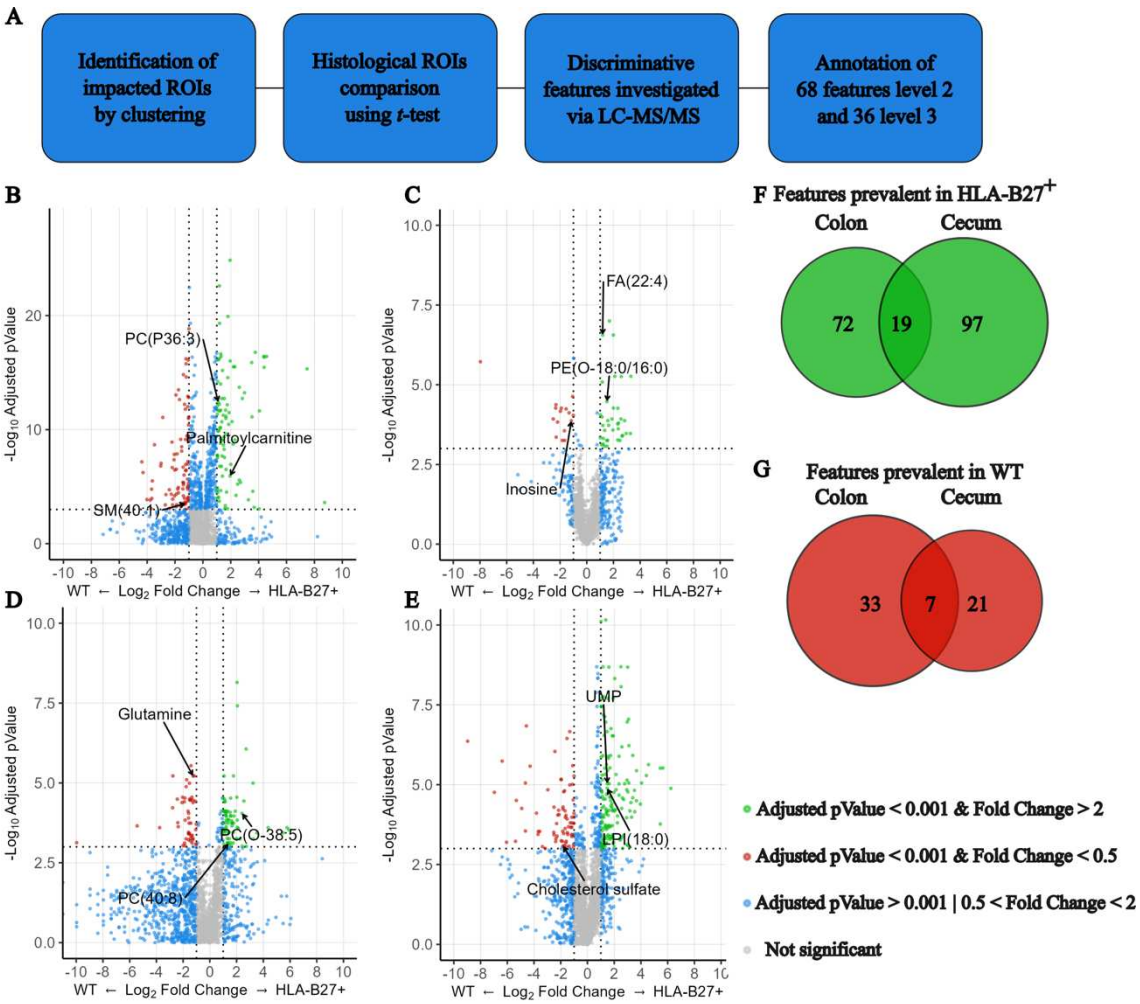


Figure 2. 12 Volcano plots displaying -Log(Adjusted p-Value) and log2(Fold change) from Student's t-test results

(A) Schematic workflow of the statistical analysis performed to identify discriminative features. Volcano plots obtained from the comparative analysis of ROIs on colon samples analysed in positive (B) and negative (B) ion mode and cecum samples analysed in positive (D) and negative (E) ion mode from HLA-B27$^+$ versus WT. By setting adjusted p-value and Fold change thresholds, we identified the most significant features prevalent in HLA-B27$^+$ (F) or WT (G) samples, to be investigated and annotated via LC-MS/MS.

2.4.3.2 LC-MS/MS method validation

Feature's molecular annotation was achieved using LC-MS/MS due to constrains already presented in the introduction and lower sensitivity of MALDI-FTICR performed on tissues in obtaining MS/MS data on least abundant ions.

Lipidomic LC method was validated analysing a lipids standard mixture, and analysing chromatograms obtained to detect the peaks corresponding to each standard composing the mixture. The maximum coverage was obtained with the analysis employing the CSH C_{18} column (Figure 2. 13), therefore we used this column for further analysis on the experimental cohort.

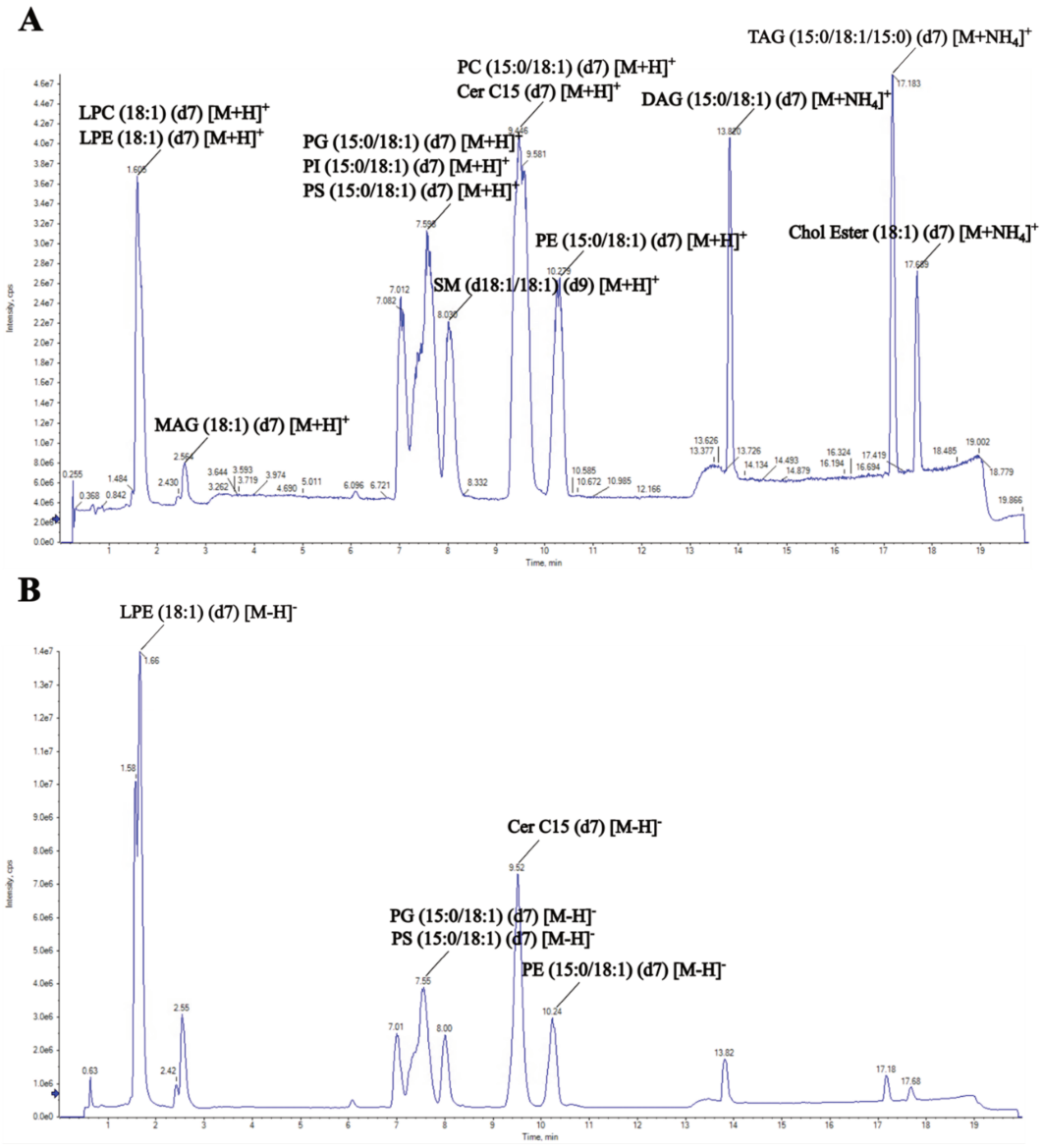


Figure 2. 13 Chromatographic separation of the standard lipid mixture using the CSH C_{18} column in positive (A) and negative (B) ionization mode (total ion count – TIC)

2.4.3.3 Annotated molecular classes

The combined metabolomic and lipidomic analysis enabled the annotation of various molecular classes, supporting the identification of biomarkers. The metabolomic standard initiative guidelines were followed to define the level of identification (Sumner et al., 2007). Analysis of the fragmentation patterns led to the putative annotation of 68 out of 275 discriminative features at identification level 2, based on spectral similarity with public spectral libraries. An additional 36 discriminative features were putatively annotated at identification level 3, based on physicochemical properties such as exact mass and ionization modality. Multiple adducts were consolidated into a final list of 83 unique annotations. We acknowledge the limitation of level 2 and level 3 annotations, which lack orthogonal validation for confirmation.

Discriminative features and annotations are listed in Annex 2 and discussed in the following paragraph.

The most represented classes were amino acids, nucleotides, nucleosides, purines and pyrimidines, and lipids, among which molecular species possessing quaternary nitrogen and phospholipids were highly represented.

A shift in the amino acid signature was observed between the two groups: glutamine was less abundant in both HLA-B27$^+$ colon and cecum samples, whereas aspartic acid, phosphoserine, glutamyl glutamic acid, and argininosuccinic acid were more abundant in colon, cecum, or both samples of the same group compared to WT rats. Nucleosides that were less abundant in HLA-B27$^+$ samples included imidazoleacetic acid ribotide in both compartments and inosine in the cecum, whereas guanosine and cytidine were more abundant in cecum and colon HLA-B27$^+$ samples, respectively. As for nucleotides, we detected uridine monophosphate (UMP) more abundant in cecum samples from the HLA-B27$^+$ group compared to WT controls.

The categories (Figure 2. 15) comprising the over expressed lipids in HLA-B27$^+$ samples were glycerophospholipids (GPLs), representing 72% of the lipids; fatty acyls (FAs), accounting for 21%; and sphingolipids (SLs), which represented the remaining 7%. In WT samples, the lipid profile comprised 55% GPLs, 31% SLs, one FA 7%, and one sterol lipid (ST) 7%. Glycerophospholipids (GPLs) that were more abundant in HLA-B27$^+$ samples were primarily glycerophosphocholines (47%), followed by glycerophosphoethanolamines (24%), glycerophosphoserines (10%), glycerophosphoglycerols (8%), glycerophosphoinositols (8%), and one glycerophosphates (3%). The GPLs main classes less abundant in HLA-B27$^+$ samples included glycerophosphocholines, LPCs and PCs, and glycerophosphoethanolamines LPEs and PEs.

Among fatty acyls (FAs), several long-chain acyl carnitines accumulated in both HLA-B27$^+$ colon and cecum samples, while the short-chain fatty acid ester butyrylcarnitine was underrepresented in HLA-B27$^+$ colon samples. Another main class of FAs, long-chain polyunsaturated fatty acids (LC-PUFAs), specifically FA(22:6), FA(22:5), and FA(22:4), were more abundant in HLA-B27$^+$ colon samples. Ceramides (Cers), a class of sphingolipids (SLs), and specifically Cer(34:0), Cer(34:1), and Cer(42:2), were prevalent in HLA-B27+ samples, while sphingomyelins (SMs), including SM(36:1), SM(38:1), SM(40:1), and SM(42:1), were more abundant in both cecum and colon samples of the WT group. Cholesterol sulphate was underrepresented in the cecum of HLA-B27$^+$ samples.

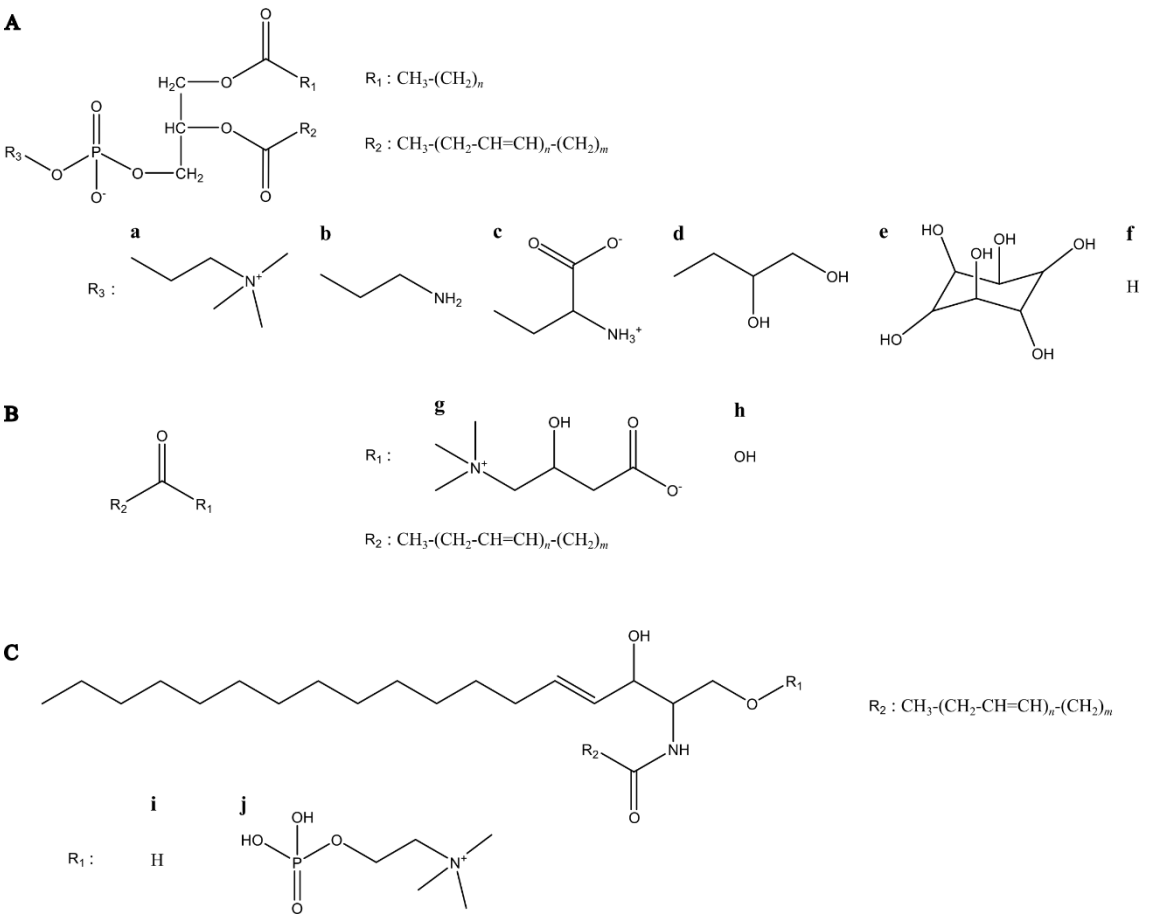


Figure 2. 14 General structure of lipidic categories and classes annotated

(A) Structure of glycerophospholipids (GLPs), (a) glycerophosphocholines (PCs), (b) glycerophosphoethanolamines (PEs), (c) glycerophosphoserines (PSs), (d) glycerophosphoglycerols (PGs), (e) glycerophosphoinositols (PIs) and (f) glycerophosphates (Pas). (B) Structure of fatty acyls (FAs), (g) acylcarnitines and (b) fatty acids. (C) Structure of sphingolipids SLs, (i) ceramides (Cers) and (j) sphingomyelins (SMs).

2.4.4 Spatial distribution of features discriminating HLA-B27$^+$ from WT changes across histological regions

We investigated the spatial distribution of discriminative metabolites and lipids across tissue sections and observed a distinct region-specific distribution pattern. Glutamine was predominantly localized in the mucosa and submucosa of WT cecum samples (Figure 2. 16 B). Inosine, a nucleoside, was more abundant in the submucosa of WT colon samples (Figure 2. 15 F), while the nucleotide uridine monophosphate was more concentrated in the mucosa of HLA-B27$^+$ cecum samples, specifically colocalizing with the epithelium (Figure 2. 16 F). The glycerophospholipids that accumulated in HLA-B27$^+$ samples were primarily associated with the lamina propria, including phosphatidylcholine PC(P-36:3) (Figure 2. 15 D), PC(O-38:5) (Figure 2. 16 C), and lysophosphatidylinositol LPI(18:0)

(Figure 2. 16 H), or with the epithelium, as seen with phosphatidylethanolamine PE(O-18:0/16:1) (Figure 2. 15 H). The acylcarnitines, particularly palmitoyl carnitine, were distributed in the lumen of HLA-B27$^+$ colon and cecum samples (Figure 2. 15 B). Long-chain fatty acids, such as fatty acid FA(22:4), co-localized with lamina propria cells in HLA-B27$^+$ colon samples (Figure 2. 15 G). The sphingolipid profile revealed ceramides in the epithelium of HLA-B27$^+$ samples, while sphingomyelins, including SM(40:1), were more prevalent in the submucosa and muscularis propria of WT colon and cecum samples (Figure 2. 15 C). Cholesterol sulphate, a sterol lipid ester, was primarily located in the epithelium and lumen of WT cecum samples (Figure 2. 16 G). The observed spatial correlation between these metabolites and specific histological regions, particularly those associated with inflammation and immune infiltration, provides valuable insights into their potential roles in the pathophysiological mechanisms of the disease.

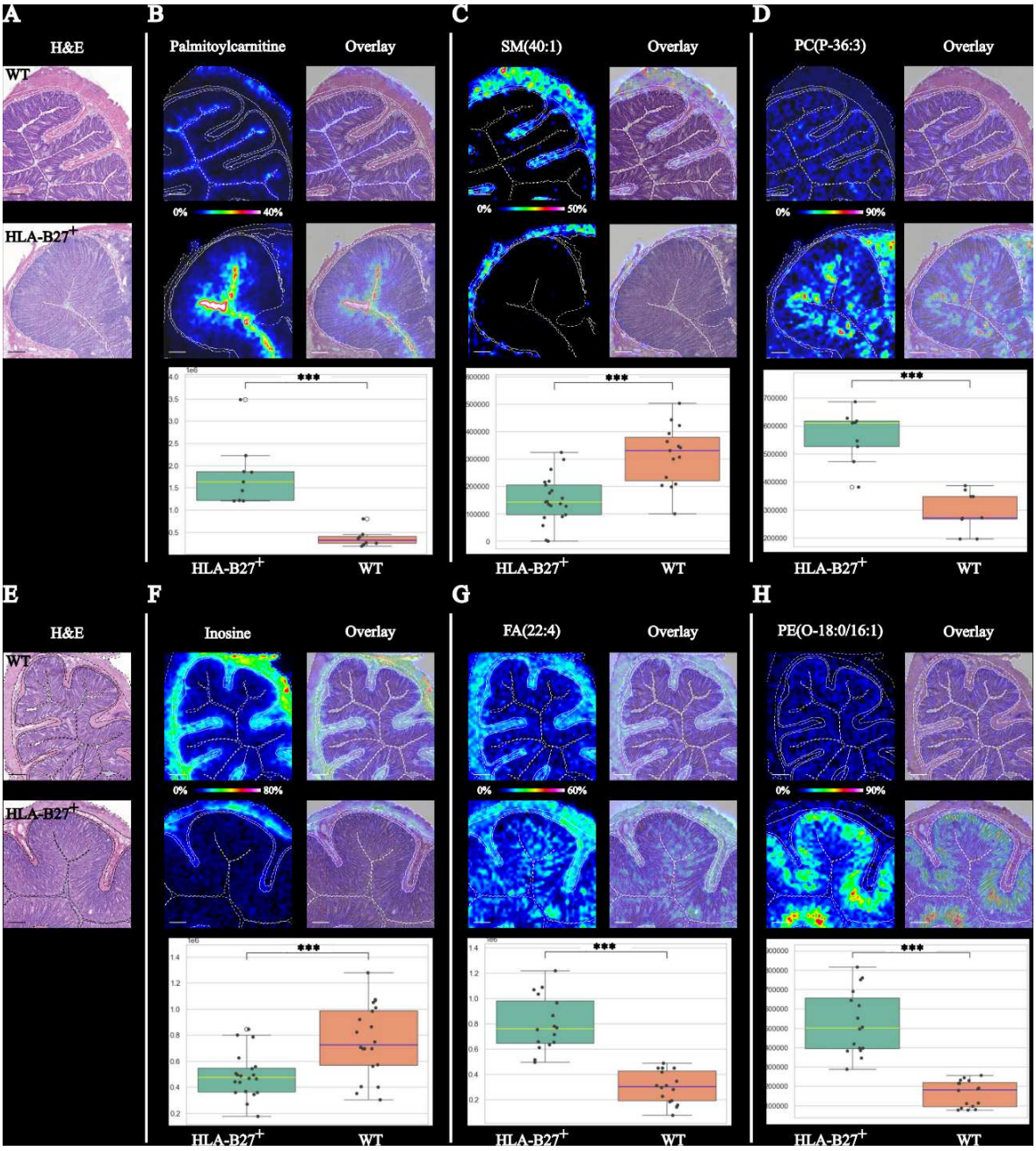


Figure 2. 15 Distribution and modulation of region-specific features between HLA-B27$^+$ and WT colon samples Spatial distribution was investigated by overlaying mass spectrometry images (B-D, F-H) with H&E coloration (A&E) on the same tissue section. Plots shows the modulated intensity of features discriminating between HLA-B27$^+$ and WT in the specific histological region where they distribute *** Adjusted p-value<0.001 (B) *m/z* 400.3436, signal of palmitoylcarnitine $[M+H]^+$, was positively regulated in the lumen of HLA-B27$^+$ samples. (C) *m/z* 787.6646 and (F) *m/z* 267.0704, signals of SM(40:1) $[M+H]^+$ and Inosine $[M-H]^-$ respectively, were less abundant in the submucosa of HLA-B27$^+$ relative to WT samples. Within the mucosa, different features colocalized with lamina propria, (D) *m/z* 768.5953 PC(P-36:3) $[M+H]^+$ and (G) *m/z* 331.2645 FA(22:4) $[M-H]^-$, or epithelium, (H) *m/z* 702.5424 PE(O-18:0/16:1) $[M+H]^+$ of HLA-B27$^+$ samples.

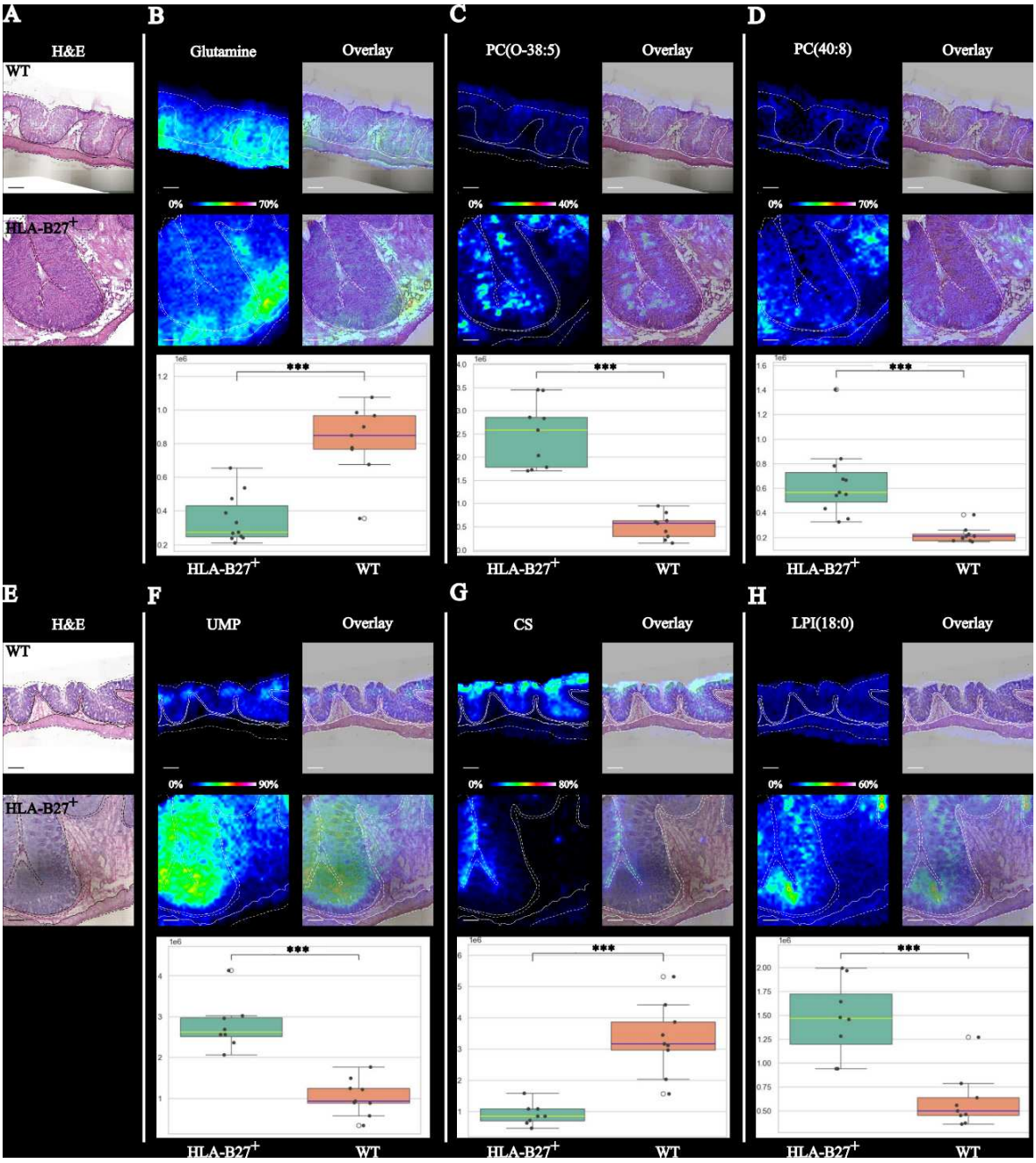


Figure 2. 16 Distribution and modulation of region-specific features between HLA-B27$^+$ and WT cecum samples

Spatial distribution of region-specific features was investigated by overlaying mass spectrometry images (B-D, F-H) with the H&E coloration (A&E) on the same tissue section. Box plots shows the modulated intensity of features discriminating between HLA-B27$^+$ and WT in the specific histological region where they are distributed *** Adjusted p-value<0.001 (B) *m/z* 147.0758 $[M+H]^+$ glutamine was less prevalent in the mucosa of HLA-B27$^+$ samples. (D) *m/z* 830.5696 $[M+H]^+$ was prevalent in the submucosa of HLA-B27$^+$ samples. Within the mucosa, different features colocalized with lamina propria, (C) *m/z* 794.6085 PC(O-38:5) $[M+H]^+$ and (H) *m/z* 599.3215 LPI(18:0) $[M-H]^-$ or epithelium, (F) *m/z* 323.0292 uridine monophosphate $[M-H]^-$ of HA-

$B27^+$ samples. (G) *m/z* 465.3056, signal of cholesterol sulphate $[M-H]^-$, was negatively regulated in the lumen of HLA-$B27^+$ relative to WT samples.

2.5 Biochemical interpretation and discussion

In this study, using MALDI-MSI, we detected biochemical changes in the inflamed gut of HLA-$B27^+$ rats, compared to the healthy gut of WT rats. With a spatial resolution of 40 µm, we were able to distinguish signal distribution within histological regions (lumen, mucosa, submucosa, muscularis propria) and specific tissues (epithelium, lamina propria). Unsupervised data analysis identified clusters discriminating the two groups. These clusters distribute in areas of the bowel sections with similar shape and size as the histological regions delineated with haematoxylin and eosin staining. The mucosa and submucosa layers were the most affected by the pathological conditions, showing non-uniform disease impact on the biochemical and metabolomic content across the tissue. The statistical study, focusing on the selected histological ROIs, identified molecular features differentially expressed in the HLA-$B27^+$ compared to WT samples, and characterized by region-specific distribution. The existing link between these metabolites and disease mechanisms, enriched with additional information provided by the spatial dimension, helped us formulate hypothesis on the role of these metabolites and lipids in the pathogenesis of IBD. Hereafter we discuss their involvement and potential as biomarkers of gut barrier integrity and immune infiltration, two key aspects of IBD.

Although the aetiology of IBD is not fully understood, gut barrier integrity plays a critical role in disease onset and progression (Michielan & D'Incà, 2015). Epithelial cells protect underlying tissue by forming tight junctions (TJs) and secreting mucins, crucial components of the mucus layer (Wells et al., 2017). Experimental evidence showed that the increase of the intestinal permeability is associated with histopathological inflammation, characterized by altered morphology and immune cell infiltration in the lamina propria, in the HLA-$B27^+$ rat model (Kerr et al., 1999). The increased permeability allows environmental and potentially pathogenic antigens to penetrate the gut barrier and trigger immune-mediated inflammatory response.

We observed significantly lower levels of glutamine in the mucosa of both colon and cecum HLA-$B27^+$ samples compared to WT controls. Glutamine deprivation or inhibition of its synthesis has been shown to decrease transepithelial resistance, increase permeability, and reduce the expression of tight junction components in a cell line model of gut barrier function (DeMarco et al., 2003; Li et al., 2004). The deficiency of this amino acid could indicate a loss of gut barrier integrity. We also observed the accumulation of uridine monophosphate and cytidine in epithelial cells of HLA-$B27^+$ samples in both colon and cecum compartments. Both molecules are actively converted into uridine through salvage pathways in epithelial cells. Uridine has shown potential activity in improving the gut barrier by ameliorating intestinal morphology when supplemented in weaned piglets (Xie et al., 2019).

Therefore, the accumulation of these mandatory precursors in HLA-B27$^+$ gut samples may suggest an impaired salvage pathway, negatively affecting protective uridine levels.

We detected several differences in the lipid signature of HLA-B27$^+$ samples compared to WT controls. The sphingolipidomic profile showed a prevalence of ceramides in HLA-B27$^+$ samples versus sphingomyelins in WT controls. Inflammatory signals in IBD can activate sphingomyelinase, which catalyses the hydrolysis of sphingomyelins into ceramides in tight junctions (Schütze et al., 1992). This altered TJ composition could contribute to increased transepithelial permeability (Bock et al., 2007). Additionally, we identified several long-chain acylcarnitines accumulating in the lumen of colon HLA-B27$^+$ samples. Increased levels of long-chain acylcarnitines were measured in the stools of IBD patients and mouse model, as a consequence of intestinal epithelial mitochondrial dysfunction (Smith et al., 2021). The mechanism and causality of this process with the pathology are still unclear, but deregulation in energy metabolism significantly impacts intestinal barrier integrity (Novak & Mollen, 2015). Moreover, the absence of cholesterol sulphate (CS) in the diseased model, prevalent in epithelial cells and the lumen of cecum WT control samples, may be another candidate marker of enhanced intestinal permeability. CS depletion was observed in a mouse model of induced colitis, and dietary supplementation of CS showed beneficial effects in this condition (Xu et al., 2022). Altogether, these metabolites could serve as candidate biomarkers signature for gut barrier impairment and increased permeability, key factors in IBD development.

We found a higher density of CD68$^+$ and CD3$^+$ positive cells infiltrating the lamina propria of HLA-B27$^+$ bowel compartments. This region of the intestine contains mature immune cells with effector roles as part of the gut-associated lymphoid tissue (Mörbe et al., 2021). In the same region, we co-localized discriminative lipids, which may serve as potential markers of a conditioned phenotype of residing macrophages and T lymphocytes. Pro-inflammatory CD68$^+$ cells in the lamina propria of inflamed intestines can exhibit an excessive inflammatory response to antigens (Mahida, 2000). This can exacerbate inflammatory processes in IBD, as seen with CD3$^+$ cells exhibiting a cytotoxic phenotype (Kappeler & Mueller, 2020). We detected an accumulation of long-chain polyunsaturated fatty acids (LC-PUFAs) in the lamina propria of HLA-B27$^+$ colon samples. The role of fatty acids (FAs) in lymphocyte and macrophage functions remains to be clarified, but FAs metabolism is involved in immune cell differentiation and function, influencing the pro versus anti-inflammatory balance (Neto et al., 2021). The lipidomic profile of the lamina propria in HLA-B27$^+$ samples was also marked by several lysophospholipids (LPLs), produced by the hydrolysis of phospholipids by phospholipases. This metabolic process affects various pathological processes. LPLs aggravate inflammation by inducing chemotaxis, altering lymphocyte immune functions, and enhancing macrophage phagocytic and pro-inflammatory responses (Liu et al., 2020; Sevastou et al., 2013). We also observed lower inosine levels in HLA-B27$^+$ colon samples compared to WT. Nucleoside

metabolism can impact immune modulation, and inosine inhibits the release of pro-inflammatory chemokines and cytokines by macrophages (Haskó et al., 2000). This activity reduces inflammation and provides protective effects in a murine model of induced colitis (Mabley et al., 2003). The lack of inosine in our disease model may indicate a pro-inflammatory state of macrophages. The identified metabolites can serve as candidate biomarkers of deregulated biological functions in the immune cells residing in the bowel compartments of this experimental model.

Through this study we were able to identify mechanistical properties of the animal model, that can be used to investigate deregulated processes of gut permeability and immune response in IBD. Given ERAPs involvement in immune response through aberrant peptide trimming and presentation, its role can be linked to at least one of the identified deregulated mechanisms. Its inhibition using small molecules could be beneficial, and the outcome could be investigated through the modulation of the candidate biomarkers identified and listed in this chapter. Overall, this study validated the HLA-B27 transgenic rat as a preclinical model to evaluate the efficacy of ERAPs inhibitors and identified detectable and measurable features that can be applied for pharmacodynamic studies.

2.6 Technical interpretation: MS potential and limitations

When conducting an untargeted metabolomic study, it must be kept in mind that the diversity of the metabolome doesn't allow to cover all the molecular classes in a single experiment. Sample preparation methods and MS platforms employed strongly impact the number and class of metabolites that can be detected. LC-MS/MS remains the most widely used analytical platform for untargeted metabolomics, because of its high sensitivity and broad range of detectable molecular classes. In this study however, we decided to employ MSI, since we consider that the spatial information attainable by this type of analysis are of greater importance and utility to study and interpret deregulated mechanism associated with disease conditions. We could still be missing relevant information that would further support and improve the validation and applicability of the preclinical model studied in this chapter. For example, among the most detected class of lipids listed in this research, several choline or carnitine containing phospholipids were detected. Choline and carnitine are quaternary ammonium compounds that are permanently, positively ionized can easily be detect in positive mode using soft ionization techniques such as electrospray ionization (ESI) and MALDI. To account for this type of issue, we used two matrices in two polarities to collect as much information as possible, while maintaining a workflow with a manageable number of variables.

A characteristic feature of mass spectrometry imaging is technical variability, due to sample preparation being performed as multiple batches of individual slides, and in this chapter we presented a successful application of a bioinformatic tool to compensate for it (Balluff et al., 2021). The batch correction algorithm has enabled the comparison among datasets acquired on different days and the

detection of clusters discriminative of the pathological conditions and independent of the analytical variability. Further efforts are required in this direction to continue developing and standardizing MALDI-MSI-based assays, advancing their application in clinical studies.

Overall, this research project showed a successful application of MALDI-MSI for an unbiased metabolomic study to discover candidate biomarkers for IBD on colon and cecum tissue samples. Taken together, our findings bridge the gap between existing knowledge about metabolic signature changes between IBD condition and healthy controls and the potential role of *in situ* metabolites in the pathogenesis of the disease, based on their spatial distribution. This greater informative level is attainable only by working on intact tissue, instead of bulk, and using imaging techniques.

2.7 Future works

Not all the organs selected for the study, despite being potentially impacted by the genetic modification, showed the expected difference between HLA-B27$^+$ and WT groups, likely due to intrinsic biological or technical unsuitability. The analysis on the eyes confirmed previous findings stating that, despite anterior uveitis being a clinical feature associated with AS and SpA, the HLA-B27 transgenic rat doesn't mirror this condition (Baggia et al., 1997). Even though we observed a macroscopic enlargement of the spleen in transgenic rats, molecular content did not show significant differences from WT. We selected the spleen as a secondary lymphoid organ since it harbours the antigen-driven proliferation and maturation of lymphocytes. It could be interesting to include lymph nodes or thymus, where lymphocytes differentiate, in further analysis to complete the study of the adaptive immune systems and identify potential differences.

The analysis on fresh frozen ankle joints, which are composed by soft and hard tissues, posed significant technical challenges. We decided not to opt for a decalcification protocol, because it could wash out, delocalize and degrade the molecular content. This procedure, however, removes unwanted signal caused by minerals in the bones and facilitates the sectioning. To overcome the difficulty in sectioning this unprepared and complex tissue, we embedded the samples in CMC and used tape to assist the collection of tissue sections. Additionally, the tissue section on tape was adhered to the plate for analysis using double-sided tape. All the layers added to support sectioning and mounting may have induced ion suppression. As an alternative protocol, the collection of tissue sections directly on double-sided tape, as well as different types of tape, including conductive copper one, should be tested to improve the mass sensitivity while maintaining tissue integrity and histological information (Bender et al., 2023). Alternative steps of sample preparation, such as freeze-drying and thaw mounting, can be employed to remove residual water content (Good et al., 2022). Considering the current level of optimization, we are unable to conclude whether a difference in metabolic footprint on the rat ankle compartment is absent or it was not detectable with the protocols and techniques

employed in this study. This opens the way for further development in sample preparation to acquire spatially resolved metabolic and lipidomic information from hard undecalcified tissue.

Further validation studies are warranted to utilize the list of discriminative features identified and listed in this study as biomarkers for diagnosis and assessment of disease progression in inflammatory bowel disease associated with ankylosing spondylitis. These biomarkers could then be used to evaluate the *in vivo* effects of ERAPs inhibitor administered to HLA-B27 transgenic rat. This study would help to address the benefits of ERAPs inhibition in restoring the physiological gut barrier integrity and regulating the immune response in ankylosing spondylitis.

3 Chapter 3 - Development of a mass method for the quantification of the ERAAP-dependent tumour antigen GSW11 in a cancer cell line and modulation upon ERAP1 inhibitory treatment

3.1 Rationale

In vitro studies are critical for the validation of therapeutic compounds, allowing for the confirmation of their mechanism of action and efficacy, before moving toward the development phase.

In this chapter, we focus on building an assay to assess the ability of a test compound to restore the level of the ERAAP-dependent tumor antigen W11 through an inhibitory mechanism targeting the ERAP1 enzyme. To validate the mechanism of action of the ERAP1 inhibitor, we chose to test the efficacy of the compound *in vitro* in both naïve cell line and the same cell line in which the pharmacological target, ERAAP (murine homologue of human ERAP1), was knocked out (Graphical abstract 2).

We employed liquid chromatography-tandem mass spectrometry (LC-MS/MS) to develop a method for the detection and absolute quantification of GSW11 in cell extracts. While ligand binding assays (LBAs) are commonly used to quantify pharmacologically active peptides and biomarkers in biological matrices, LC-MS/MS holds several advantages including faster and more cost-effective development, enhanced specificity through direct unlabeled analysis, ability to analyze peptides regardless of pharmacological activity, and the use of internal standards (IS) to monitor and compensate for variability of the process (Zhang & Jian, 2014). In this study, we employed a stable isotope-labelled peptide introduced at early stage of sample preparation. We used multiple reaction monitoring (MRM) approach for top-down peptide quantification. In this approach the ion of the intact peptide is selected as the parent for collision-induced dissociation (CID) fragmentation, after which a product ion is chosen for quantification.

The direct and quantitative analysis of this biomarker aimed to enhance our understanding of the compound's pharmacodynamic properties, including its role in managing antigen levels as a therapeutic option in cancer treatment.

3.2 Context

ERAPs play a key role in processing and presenting antigens on MHC class I molecules by trimming peptides at the N-terminal. They generate the antigens repertoire required for the proper stimulation of cytotoxic immune cells, thereby regulating immunosurveillance. The activity of ERAPs can

modulate the expression of antigenic peptides, which may increase, decrease, or remain unchanged depending on the enzyme's expression and activation levels (G. E. Hammer et al., 2006).

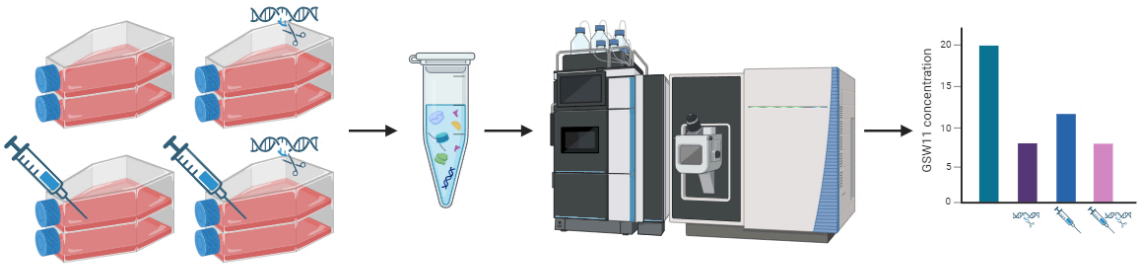
GSW11 is a tumor-specific peptide discovered in CT26 murine colon cancer cells, which has been identified as a target of cytotoxic T lymphocyte (CTL) responses. The immunostimulatory ability was identified by CCD2Z T-cell hybridoma. Moreover, GSW11 was found to have a role mediating cross-protective immunity in various tumor models (James et al., 2010).

Pharmacological inhibition of various proteases demonstrated that ERAAP overactivity leads to the excessive trimming of GSW11, resulting in its destruction and a reduced capacity to stimulate immune responses. The role of ERAAP was confirmed by knocking it out in CT26 cells showing that CCD2Z T-cell hybridoma stimulation, dependent on GSW11 levels, was proportional to the ERAAP expression *in vitro*. Attenuation of tumor growth *in vivo* and prolonged survival after inhibition of ERAAP provided proof that inhibitors of ERAAP might have a therapeutic value in oncology (James et al., 2013).

Zervoudi et al. investigated the impact of two pseudopeptidic ERAP1 inhibitors on the presentation of endogenous antigens. They evaluated the generation and cell surface presentation of GSW11 by assessing the activation of the T-cell hybridoma CCD2Z. Inhibiting ERAAP with DG013A enhanced GSW11 tumor antigen presentation in CT26, stimulating cytotoxic T lymphocyte responses in a dose-dependent manner (Zervoudi et al., 2013).

In this chapter, we aimed to assess the efficacy of a small molecule in increasing GSW11 quantity in CT26 cells through the mechanism of ERAP1 inhibition.

3.3 Experimental design



Graphical abstract 2 LC-MS/MS analysis for ERAP1-dependent GSW11 peptide quantification

CT26 ERAAP^{naïve} and ERAAP^{ko} cells were treated with ERAP1 inhibitor or vehicle. Extraction protocols were applied to extract the endogenous ERAP1-dependent tumour antigen GSW11. A quantitative method, employing surrogate matrix and isotope-labelled internal standard approaches, was developed to measure GSW11 expression in the different conditions. Created with BioRender.com

3.3.1 Experimental cohort and sample preparation

3.3.1.1 Cell lines

CT26 ERAAP^{naive} and CT26 ERAAP^{ko} cells, provided by Prof James (University of Southampton), were maintained in RPMI 1640 medium supplemented with 10% foetal bovine serum, 2 mM L-glutamine, 100 I.U./mL penicillin, 100 µg/ml streptomycin, 50 µM 2-Mercaptoethanol, 1 mM sodium pyruvate, and 1 mM HEPES.

3.3.1.2 Cells expansion and treatment

100 millions of CT26 ERAAP^{naive} and ERAAP^{ko} cells per replicate (n=3) at 60-70% confluency were treated with an ERAP1 inhibitor provided by Prof Deprez-Poulain (University of Lille) or vehicle (DMSO) to a final concentration of 10 µM for 48 hours. Cells were lifted from culture flasks with EDTA 0.2 mM, resuspended in PBS and centrifuged at 2000 rpm for 3 minutes to wash the pellet. After PBS aspiration, pellets were stored at -80°C until use.

3.3.1.3 Peptide extraction from cell pellet

The protocol for the extraction of GSW11 peptide (Gly-Gly-Pro-Glu-Ser-Phe-Tyr-Cys-Ala-Ser-Trp) (Figure 3. 3 A) from CT26 cells was developed by James et al., and adapted from Hammer et al. (G. E. Hammer et al., 2006; James et al., 2013). Briefly, 100 million CT26 ERAAP^{naive} and ERAAP^{ko} cell pellets were resuspended in 800 µl of 10% formic acid in H_2O with 5 µM of "martyr" irrelevant peptide and 100 nM of IS. The solution was heated at 95°C for 10 minutes and centrifuged at max speed for 15 minutes. The supernatant was filtered using a 10 kDa cutoff filter and the filtrate was dried in a SpeedVac SPD120 (Figure 3. 1). We tested either the complete or the partial drying of the filtrate prior to LC-MS/MS injection.

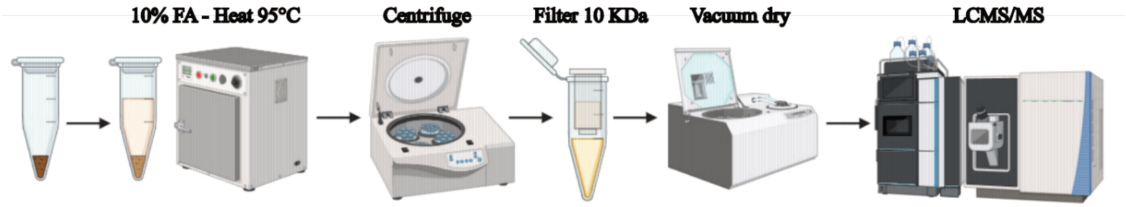


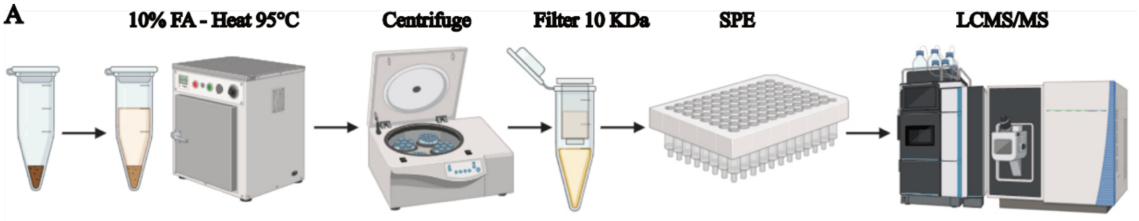
Figure 3. 1 Protocol for GSW11 peptide extraction from CT26 cell pellets
Created with BioRender.com

3.3.1.4 Optimization of solid phase extraction step

To clean the extract from potential contaminants and optimize the recovery of endogenous GSW11 peptide from cell pellets, we added a step of solid phase extraction (SPE) to the original protocol using an Oasis HLB 96-well plate (Figure 3. 2 A). We tested four methods using different preparation steps to optimize the elution of the peptide (Figure 3. 2 B). Three methods were proposed by the vendor for generic application ("Standard") or optimized based on the analyte dissociation constant

("Basic" and "Acid"). Moreover, we applied a method ("Peptide") developed from previous researches employing the Oasis HLB plate for peptide purification from biological matrices such as plasma and organoids (George et al., 2023; Kay et al., 2018; Miedzybrodzka et al., 2020). For all the steps, the volume of solvent used was 200 µL, except for the elution step that was repeated twice with 25 µl.

A



B

Step	Method			
	Standard	**Acid**	**Basic**	**Peptide**
Condition	MeOH	MeOH	MeOH	MeOH
Equilibration	H_2O	H_2O	H_2O	H_2O
Wash n1	5% MeOH	5% MeOH	5% MeOH	0.1% FA
Wash n2	5% MeOH	5% NH_4OH in MeOH/H_2O	2% FA in MeOH/H_2O	5% MeOH 1% AA
Elution	MeOH	2% FA in MeOH/H_2O	5% NH_4OH in MeOH/H_2O	60% MeOH 10% AA

Figure 3. 2 Protocol for GSW11 peptide extraction from CT26 cell pellets employing SPE

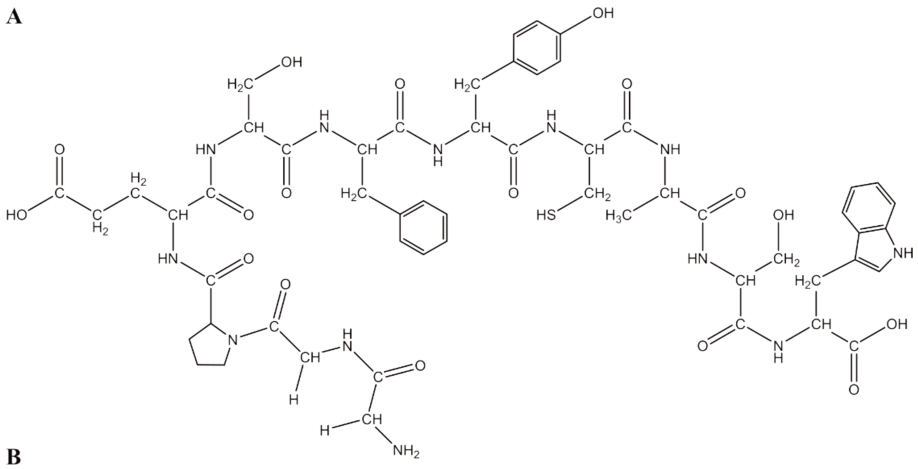
(A) Schematic workflow of the protocol, including solid phase extraction step. Created with BioRender.com

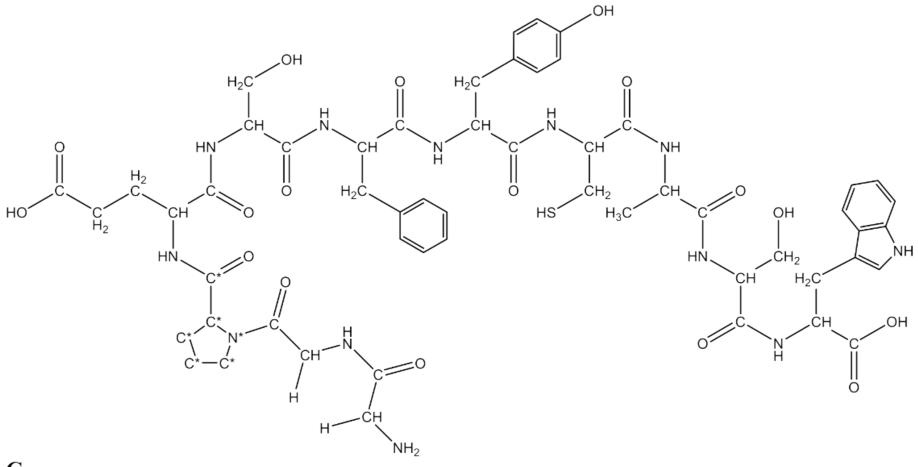
(B) Table reporting the steps tested in the different methods for SPE plate preparation, extract washing and elution

3.3.2 LC-MS/MS method development

3.3.2.1 Synthetic peptides

Synthetic standards of the native GSW11 peptide (Figure 3. 3 A) and its isotope labelled IS (Figure 3. 3 B) were synthesized by Thermo Scientific. In the IS, the five ^{12}C and the ^{14}N of the proline were replaces with ^{13}C and ^{15}N isotopes, with a mass shift of 6 Da from the native peptide (Figure 3. 3 C). Stock solutions of the synthetic peptides were prepared at a concentration of 1 mM by in DMSO.





C

Name	Sequence	Modification	Formula	MW (g/mol)
Native GSW11	[NH2]-GGPESFYCASW-[COOH]		$C_{55}H_{70}N_{12}O_{17}S$	1203.3
Internal standard	[NH2]-GG(P)ESFYCASW-[COOH]	Proline-$^{13}C_5$,^{15}N	$C_{50}{}^{13}C_5H_{70}N_{11}{}^{15}NO_{17}S$	1209.3

Figure 3. 3 Structures of native GSW11 (A) and its isotope labelled IS (B) peptides
Asterisks indicate the ^{12}C and ^{14}N atoms on proline substituted with the isotopes ^{13}C and ^{15}N (C) Table reporting the chemical properties of the synthetic peptides

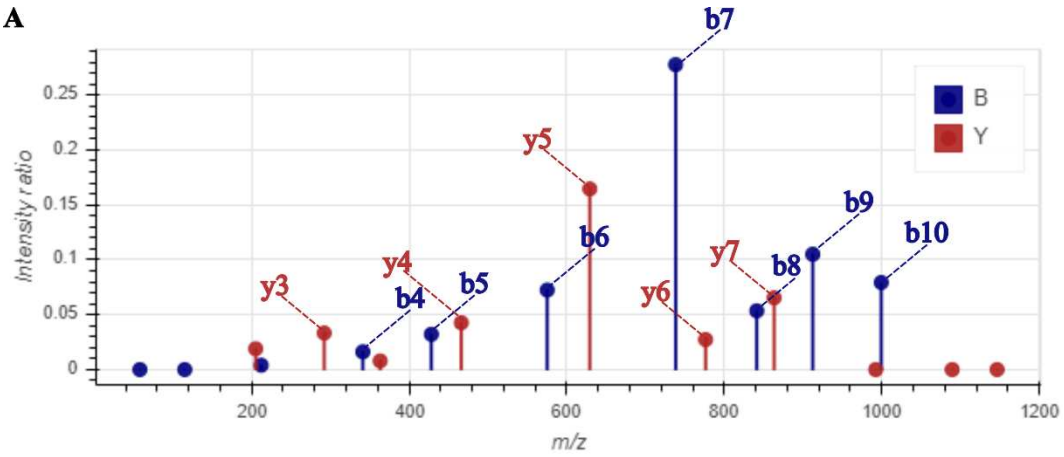
3.3.2.2 MS/MS method development

To develop and optimize the MS/MS method, a solution at 10 µM, prepared in ACN:H_2O 1:1 (v:v) + 0.1% FA, of the native GSW11 peptide and its IS were infused directly in the triple quadrupole mass spectrometer using a syringe at a flow rate of 20 µl/min. The analytes ionization was performed in negative and positive electrospray ionization mode. Four vaporizer temperatures, 300, 325, 350, 375°C, were assayed to identify the optimal one for analyte's detection based on the native GSW11 peptide normalized peak area.

In negative ion mode, the detection and optimization of the product ions, originated by the fragmentation of the native GSW11 or IS peptides, were performed using the vendor's tuning application. The software identified the three most intense production ions and optimized the collision energy for fragmentation.

In positive ion mode, we used the bioinformatic tool MS2PIP Server - CompOmics to predict the fragmentation pattern by collision induced dissociation (CID), including peak intensities, from the native GSW11 peptide sequence (Declercq et al., 2023). The software provided a peak list and a spectrum (Figure 3. 4), which was cross-checked with the experimental data to select the product ions to be used for detection and quantification. The vendor's tuning application was used to optimize the collision energy for fragmentation.

A



B

b4 : H-Gly-Gly-Pro-Glu-CO$^+$
b5 : H-Gly-Gly-Pro-Glu-Ser-CO$^+$
b6 : H-Gly-Gly-Pro-Glu-Ser-Phe-CO$^+$
b7 : H-Gly-Gly-Pro-Glu-Ser-Phe-Tyr-CO$^+$
b8 : H-Gly-Gly-Pro-Glu-Ser-Phe-Tyr-Cys-CO$^+$
b9 : H-Gly-Gly-Pro-Glu-Ser-Phe-Tyr-Cys-Ala-CO$^+$
b10 : H-Gly-Gly-Pro-Glu-Ser-Phe-Tyr-Cys-Ala-Ser-CO$^+$

y3 : NH3$^+$-Ala-Ser-Trp-COOH
y4 : NH3$^+$-Cys-Ala-Ser-Trp-COOH
y5 : NH3$^+$-Tyr-Cys-Ala-Ser-Trp-COOH
y6 : NH3$^+$-Phe-Tyr-Cys-Ala-Ser-Trp-COOH
y7 : NH3$^+$-Ser-Phe-Tyr-Cys-Ala-Ser-Trp-COOH

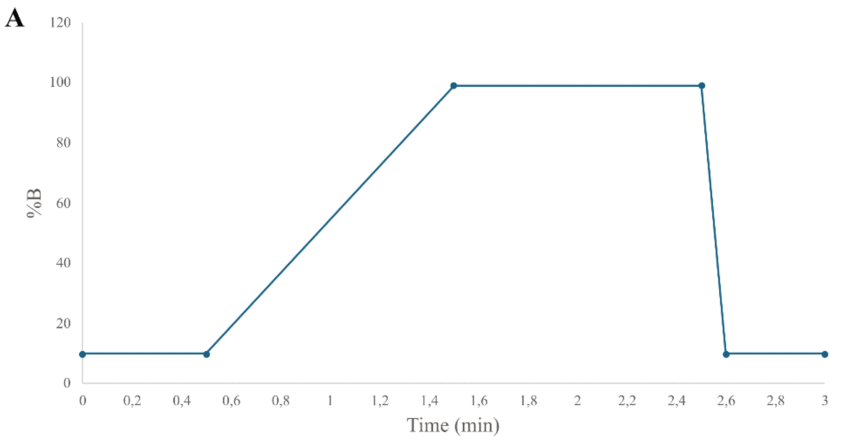
Figure 3. 4 Theoretical fragmentation pattern of the native GSW11 peptide

(A) Spectrum of the peptide fragmentation pattern. (B) Sequence of the b, acylium, and y, ammonium, fragment ions generated. Created with MS2PIP Server - CompOmics using the collision induced dissociation (CID) model.

3.3.2.3 Chromatography

Elution was conducted in gradient mode, using 0.1% FA in H_2O as mobile phase A and 0.1% FA in ACN as mobile phase B on an Ultimate 3000 UHPLC system. Gradient program is detailed in Figure 3. 5. Column and autosampler temperature were maintained at 20°C and 4°C respectively. The flow rate was set at 0.4 ml/min.

A



B

Time (min)	%B	Flow rate (ml/min)
0	10	0,4
0,5	10	0,4
1,5	99	0,4
2,5	99	0,4
2,6	10	0,4
3	10	0,4

Figure 3. 5 Graphic (A) and table (B) reporting the gradient program for chromatographic separation

3.3.3 Parallelism test

Parallelism test was conducted to choose the suitable surrogate matrix for the quantification of GSW11 in CT26 cells extracts. Mouse plasma, mouse colon homogenate and BSA were assayed. Calibration curves were prepared by spiking serial dilution of the native GSW11 peptide in matrix extracts prepared according to the original protocol for endogenous peptide extraction from cells (Fig.3.1). Thermo Excalibur Quan Browser software was used to build the titration curve by plotting the native GSW11 peaks area, normalized by the IS peaks area, against the concentrations.

3.3.4 Extraction recovery and matrix effect

Extraction recovery was measured for the two extraction protocols tested. Native GSW11 peptide was spiked in surrogate matrix before, at each step and after the extraction protocol. Samples were analysed by LC-MS/MS and recovery was calculated with the formula:

$$\%\ Recovery = \frac{Normalized\ peak\ area\ (pre\ spike)}{Normalized\ peak\ area\ (post\ spike)} * 100$$

We measured the recovery of the extraction protocol from James et al., in the two cases of vacuum drying performed to completion, or partially, with about 50 µl of solvent remaining. The same amount of native GSW11 peptide was spiked in colon homogenate either before the extraction protocol or in the final extract obtained under both conditions (Figure 3. 6).

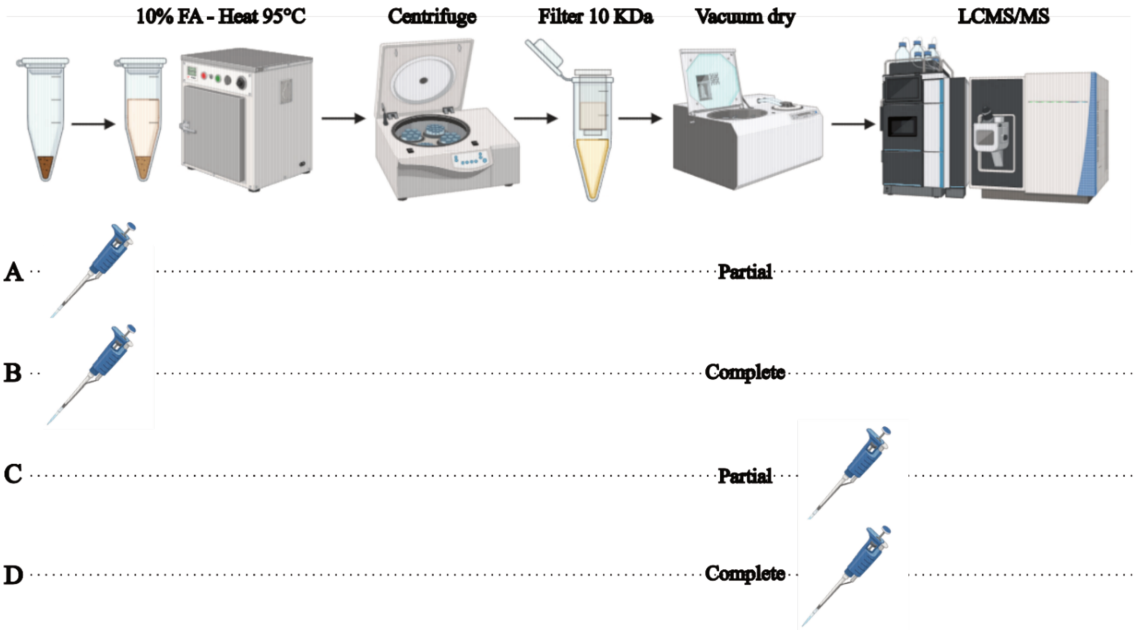


Figure 3. 6 Schematic workflow of the recovery assay for native GSW11 peptide extraction using classic protocol

The pipette icon indicates the moment of spiking in surrogate matrix. The same amount of synthetic peptide was spiked before the extraction in two matrix tubes, that underwent either (A) partial or (B) complete drying. The same amount of synthetic peptide was spiked in the two final extracts, which were (C) partially or (D) completely dried. Created with BioRender.com

We measured the recovery of the extraction protocol using the SPE with the "Peptide" method. The same amount of native GSW11 peptide was spiked in murine cell pellets before each step of the protocol (Figure 3. 7). The same analysis was performed with or without the addition of 5 µM of "martyr", irrelevant peptide in the extraction solvent.

The same amount of native GSW11 peptide was prepared in solution (ACN:H_2O 1:1 v/v + 0.1% FA). We measured the matrix effect of the surrogate matrix, and the analytical method employed. The variation between the synthetic peptide signal measured in surrogate matrix or in solution was calculated using the formula:

$$\%\, Variation = \frac{Peak\ area\ (pre\ spike) - Peak\ area\ (in\ sol)}{Peak\ area\ (\text{in sol})} * 100$$

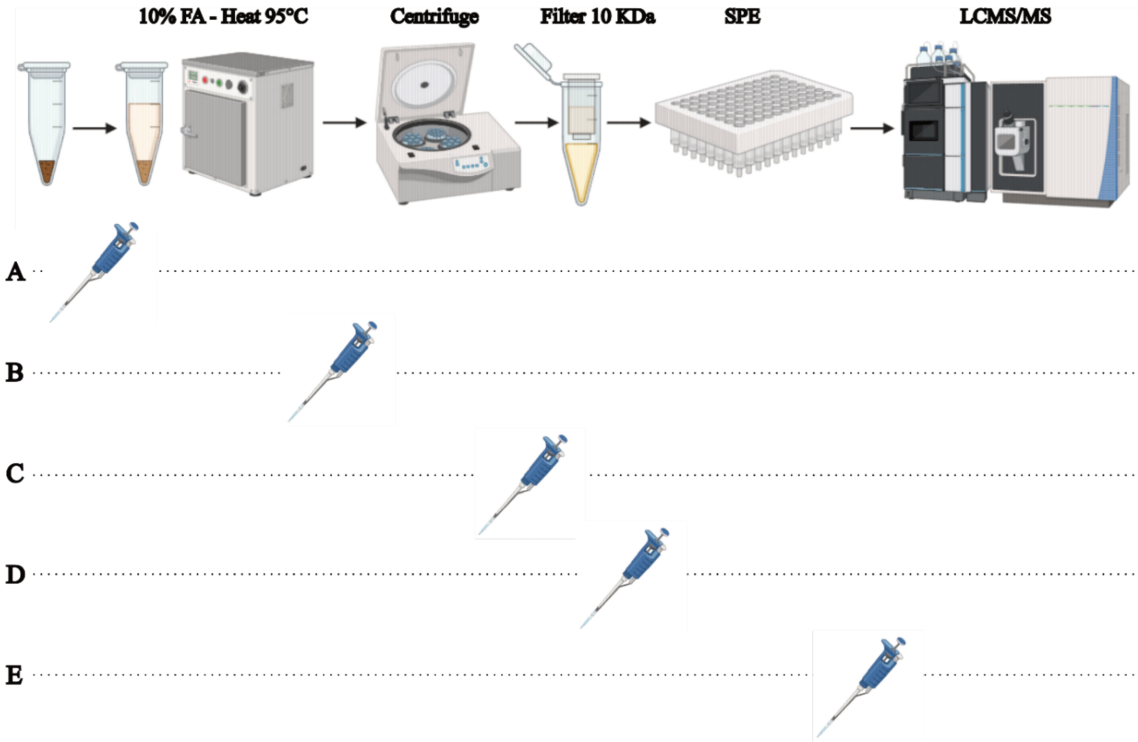


Figure 3. 7 Schematic workflow of the recovery assay for native GSW11 peptide extraction using protocol employing SPE

The pipette icons indicate the moment of spiking in surrogate matrix. The same amount of synthetic peptide was spiked in the matrix (A) before the extraction, (B) before the centrifugation, (C) before the filtration, (D) before the SPE, or (E) in the extracts after the extraction. Created with BioRender.com

3.4 Results

3.4.1 LC-MS/MS method development for optimal detection of native GSW11 peptide

3.4.1.1 MRM transitions development in negative ion mode

The most intense peaks for native GSW11 and IS peptide were detected at *m/z* 1201.5 and *m/z* 1207.5 respectively (Figure 3. 8 C, F), corresponding to the singly negatively charged ions. Ionization was performed in electrospray negative ionization mode, with a spray voltage of 2500 V. The tuning application identified the three most intense product ions (transitions) at *m/z* 1171.417, *m/z* 1183.417 and *m/z* 1167.417 for the native GSW11 peptide and *m/z* 1177.333, *m/z* 1189.333 and *m/z* 1173.333 for the IS peptide (Figure 3. 8 D, G). Collision voltages were optimized between 40 and 50 V for both native GSW11and IS peptides (Figure 3. 8 E, H). A minor cross contamination (1:1000) was detected between the native GSW11 and IS peptides when analysed simultaneously, justified by the presence of a peak at *m/z* 1207.5, accounting for a 0.111% of the native GSW11 peptide isotopic pattern (Figure 3. 8 A), and the minimal percentage of impurity always present in isotope labelled amino acids. To

correct for this issue, the MRM transitions for the IS peptide were selected by shifting the precursor and product ions by +1 Da.

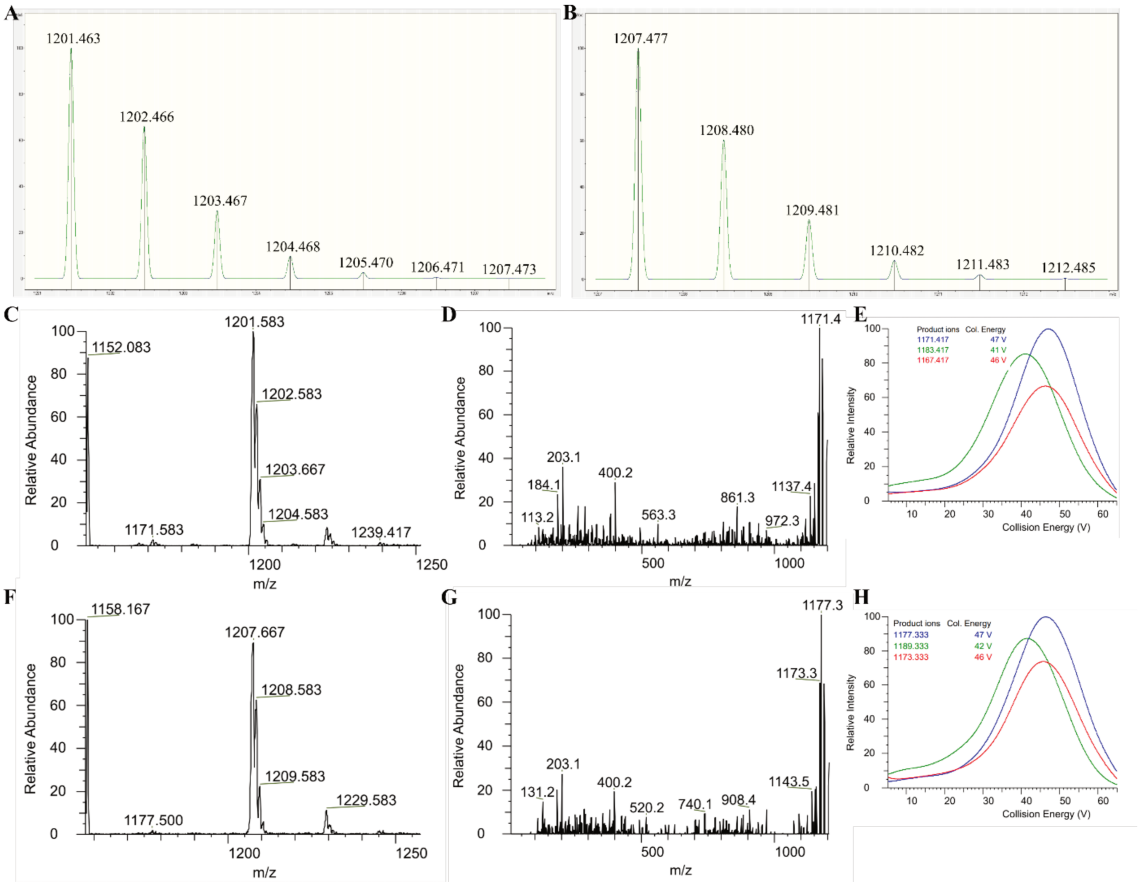


Figure 3. 8 Transitions development in negative ion mode

Theoretical spectra predicted from Compass Isotope Pattern application based on the chemical formula of (A) native GSW11 and (B) IS peptides. Precursor ion spectrum, product scan and breakdown curve at 1.5 mTorr are shown for the native GSW11 ion *m/z* 1201.5 (C, D, E) and the IS ion *m/z* 1207.5 (F, G, H).

3.4.1.2 Vaporizer temperature optimization

The performance of the four vaporizer temperatures (VP) tested was evaluated by comparing native GSW11 normalized peak area for each condition. Rising the VP from 300 to 325°C led to a 15% increase of the normalized peak area (Figure 3. 9). The further rise of the VP didn't yield to significant additional increase, therefore VP was set to 325°C.

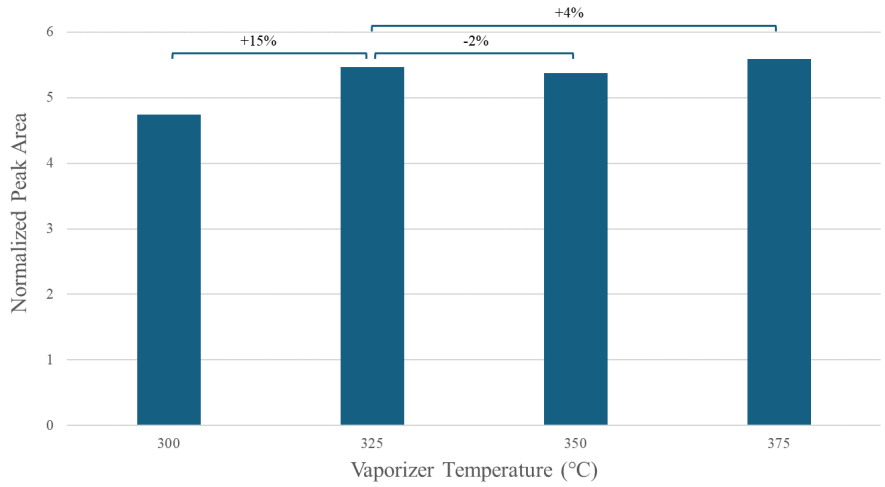
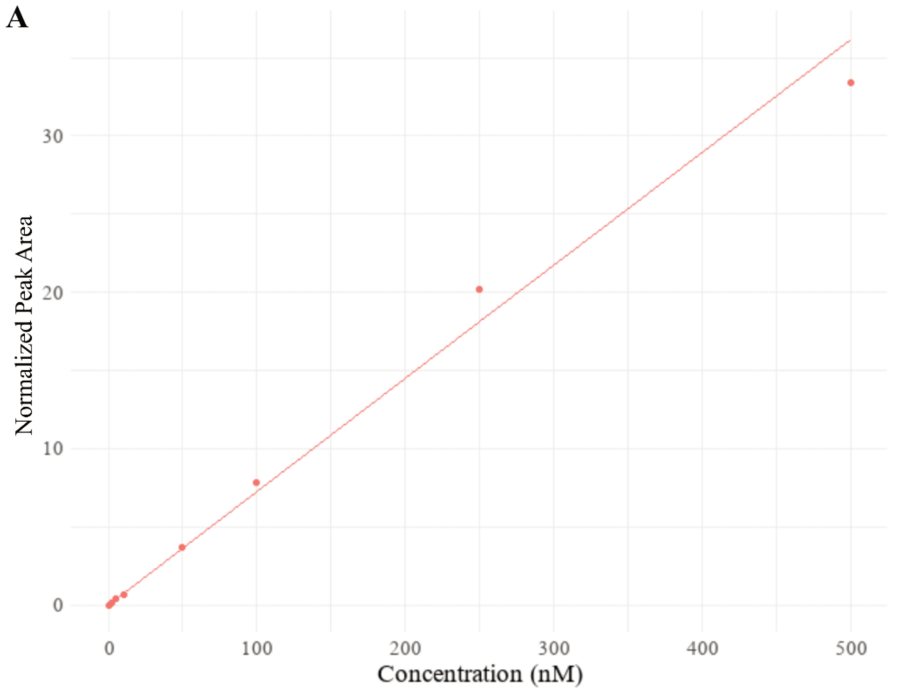


Figure 3. 9 Vaporizer temperature optimization results

Bar graph showing the modulation of the GSW11 normalized peak area as a result of the rising of the vaporiser temperature (VP). The percentages of increase or decrease are reported above the respective temperatures.

3.4.1.3 Evaluation of method sensitivity in solution

The estimation of limit of detection (LOD) and lower limit of quantification (LLOQ) was based on a titration curve of serial dilution of native GSW11 peptide to generate twelve (12) points, with concentrations from 0.075 to 500 nM prepared in ACN:H_2O 1:1 (v/v) + 0.1% FA in presence of IS at 25 nM and "martyr" irrelevant peptide at 5 µM. The chromatographic peak area of the native GSW11 peptide was normalized by the chromatographic peak area of the IS. The titration curve was generated using the Quan Browser software. A weighted linear 1/X curve with coefficient of determination R^2 0.9922 was obtained. LOD and LLOQ of the native GSW11 peptide were found at 0.1 and 0.25 nM respectively (Figure 3. 10).



B

Curve	Equation	R^2	LOD (nM)	LLOQ (nM)
Titration curve in solution	y = -0.0110 + 0.0724x	0.9922	0.1	0.25

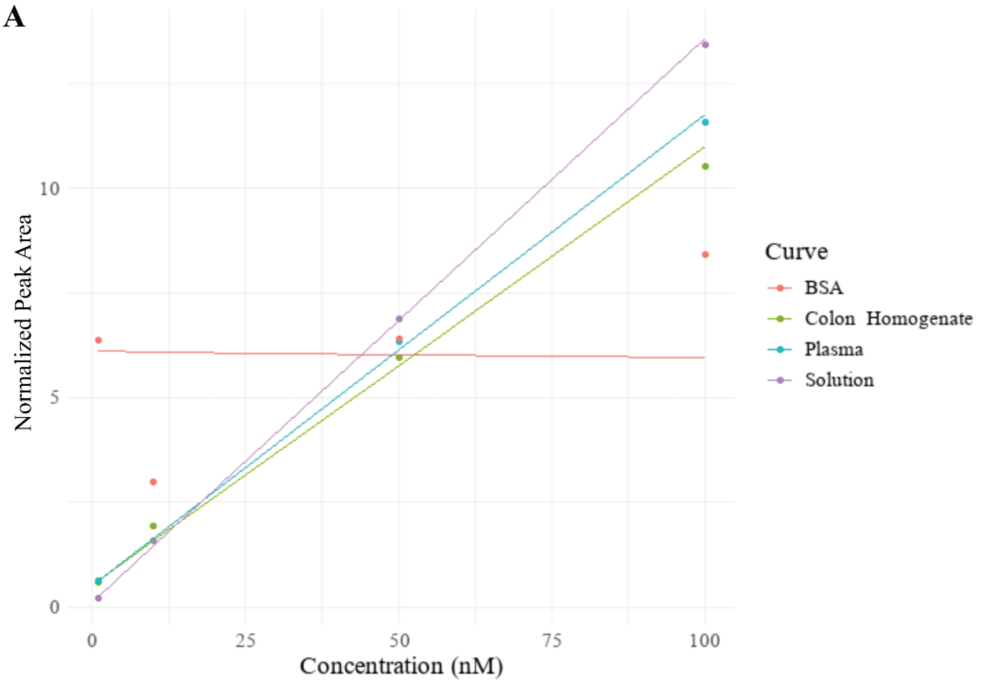
Figure 3. 10 Native GSW11 peptide titration curve in solution

(A) Linear weighted 1/X titration curve of the native GSW11 peptide in ACN:H_2O 1:1 (v/v) + 0.1% FA plotted in a graph normalized peak area over concentration (nM). (B) Table reporting equation, R^2 and limits of the titration curve.

3.4.1.4 Parallelism test for selecting the surrogate matrix

Quantifying endogenous molecules is challenging due to the difficulty in obtaining matrices that are both comparable to the samples being analysed and free of the analyte. Several approaches have been developed to address this issue, including the one selected for this study, which uses a surrogate matrix spiked with analyte's standards. To demonstrate the appropriate choice of a surrogate matrix for accurate quantification, parallelism needs to be demonstrated (Jones et al., 2012). Titration curves were prepared in mouse colon homogenate, mouse plasma and BSA as surrogate matrices and compared to a titration curve prepared in solution (ACN:H_2O 1:1 (v/v) + 0.1% FA). Good parallelism was observed for curves of the native GSW11 peptide prepared in colon homogenate and plasma, with a relative standard deviation percentile of 13% from the one in solution. The BSA, however, didn't prove to be a suitable matrix to build titration curve, as no correlation was observed between native GSW11 peptide normalized peaks areas and concentrations (Figure 3. 11). This was possibly due to the peptide's hydrophobicity and ammino acid composition, including tryptophan, tyrosine, and phenylalanine, which may confer high affinity for serum albumin. Given that CT26 are cells from

murine colon carcinoma, the mouse colon homogenate was expected to have more biological similarity and was therefore selected as the surrogate matrix.



B

Curve	Equation	R^2
Solution	y = 0.1120 + 0.1346x	0.9992
Colon Homogenate	y = 0.5431 + 0.1045x	0.9904
Plasma	y = 0.5254 + 0.1124x	0.9994
BSA	y = 6.1042 + -0.0015x	0.0003
RSD%	13%	

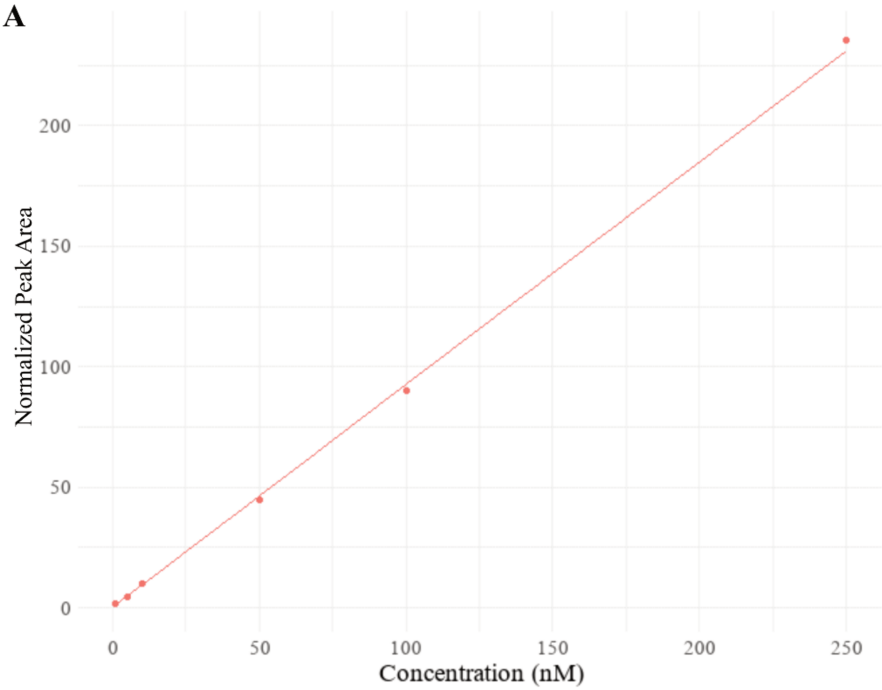
Figure 3. 11 Results of the parallelism test for surrogate matrix selection

(A) Linear weighted 1/x titration curves proving parallelism of the native GSW11 peptide in ACN:H_2O 1:1 (v/v) + 0.1% FA, colon homogenate and plasma, with an RSD% of 13%. The titration curve of the native GSW11peptide in BSA didn't show correlation between normalized peaks area and concentrations. (B) Table reporting equation, R^2 and variability (RSD%) of each curve.

3.4.1.5 Evaluation of method sensitivity in surrogate matrix

The estimation of LOD and LLOQ was based on a titration curve of serial dilution of native GSW11 peptide to generate twelve (12) points, with concentrations ranging from 0.075 to 250 nM, spiked in mouse colon homogenate extracts prepared according to James et al. protocol (Figure 3. 1). The chromatographic peak area of the native GSW11 peptide was normalized by the chromatographic peak area of the IS. The titration curve was generated using Quan Browser software. A weighted linear 1/X curve with R^2 0.9989 was obtained. LOD and LLOQ of the native GSW11 peptide were found at 0.25 and 1 nM respectively (Figure 3. 12).

A



B

Curve	Equation	R^2	LOD (nM)	LLOQ (nM)
Titration curve in surrogate matrix	y = 0.5086 + 0.9217x	0.9989	0.25	1

Figure 3. 12 Native GSW11 peptide titration curve in surrogate matrix

(A) Linear weighted 1/X titration curve of the native GSW11 peptide in colon homogenate plotted in a graph normalized peak area over concentration (nM). (B) Table reporting equation, R^2 and limits of the titration curve.

3.4.2 Endogenous GSW11 peptide extraction from CT26 cell pellets

3.4.2.1 Evaluation of the "Martyr" peptide's effect on native GSW11 peptide chromatographic elution

The presence of the "Martyr" peptide, added in excess to the extraction solution, was justified by the author as it binds the HPLC column, allowing higher concentration of the peptide of interest to be eluted rather than to bind the column. We evaluated the benefit of the "Martyr" peptide on the method developed to quantify the native GSW11 peptide via LC-MS/MS. Four concentrations of native GSW11 peptide were prepared in ACN:H_2O 1:1 (v/v) + 0.1% FA with IS 25 nM and with or without 5 µM of "Martyr" irrelevant peptide. We compared the chromatographic peaks areas of the native GSW11 peptide to determine if the presence of the "Martyr" peptide improved the elution. The normalized peak areas were comparable between the two conditions, with or without "Martyr" peptide (Figure 3. 13 A), however a considerable difference was observed in the raw areas, with the presence of the "Martyr" peptide negatively impacting the elution of the native GSW11 peptide (Figure 3. 13 B). We decided to perform further extractions without the "Martyr" peptide addition.

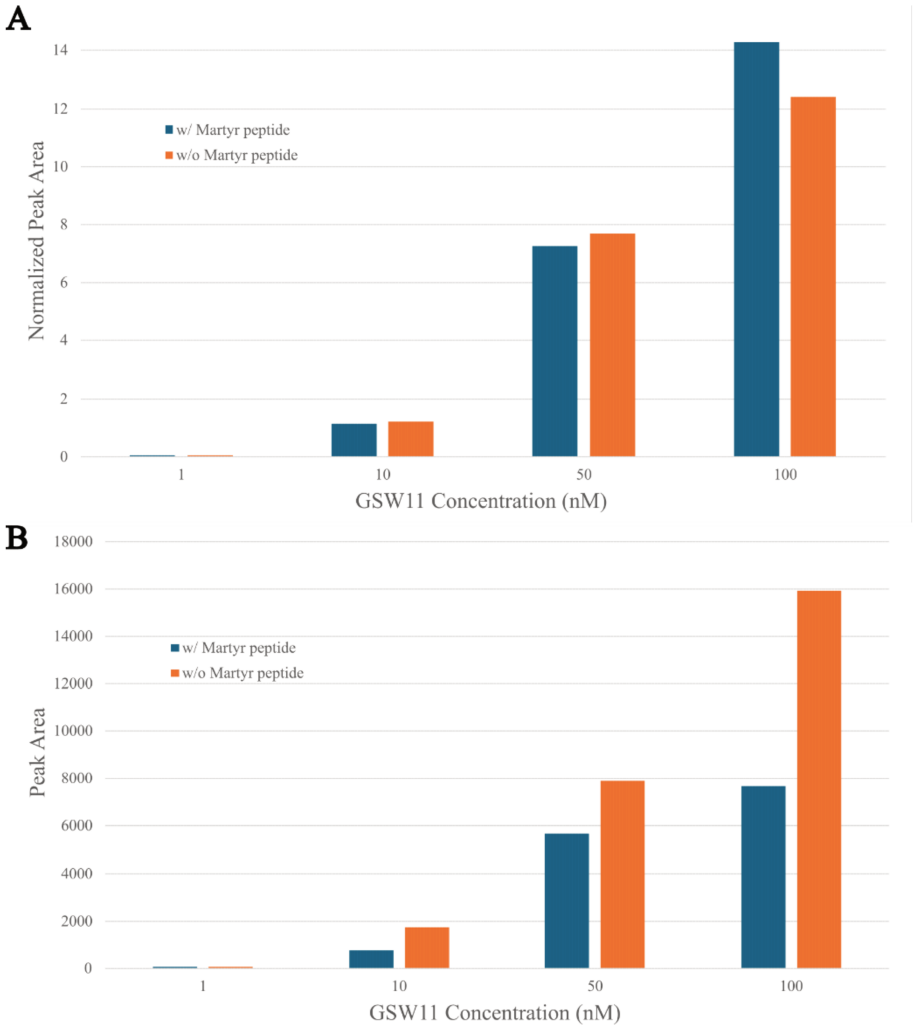


Figure 3. 13 Investigation of "Martyr" peptide role results

Bar graphs showing the modulation of the (A) normalized and (B) non-normalized chromatographic peak area of native GSW11 peptide at different concentration in presence (w/ blue) or absence (w/o orange) of "martyr" irrelevant peptide.

3.4.2.2 Assessment of the extraction protocol recovery

The percentage of an analyte recovered after the sample preparation represents the efficiency of an extraction protocol. The extraction protocol from James et al. was assayed to measure the recovery from surrogate matrix spiked with native GSW11 peptide pre or post extraction. To investigate the impact of the vacuum drying step, we performed the experiment in duplicate: two samples were fully dried, while two were only partially dried. We first observed that complete drying negatively impacted the recovery of the native GSW11 peptide compared to partial drying, in both samples spiked before (pre-spike) or after (post-spike) extraction. The recovery was 75%, calculated as the ratio between the normalized peak area of pre-spike completely dried samples and pre-spike partially dried samples (the same ratio applied for post-spike). In both drying conditions, the overall recovery of the

extraction protocol, determined by comparing the pre- and post-spike areas, was 47% (Figure 3. 14). We proceeded to test both protocols on control, untreated CT26 ERAAP^{naive} and ERAAP^{ko} cell pellets, to determine whether this recovery was sufficient to obtain a signal for the endogenous GSW11 peptide.

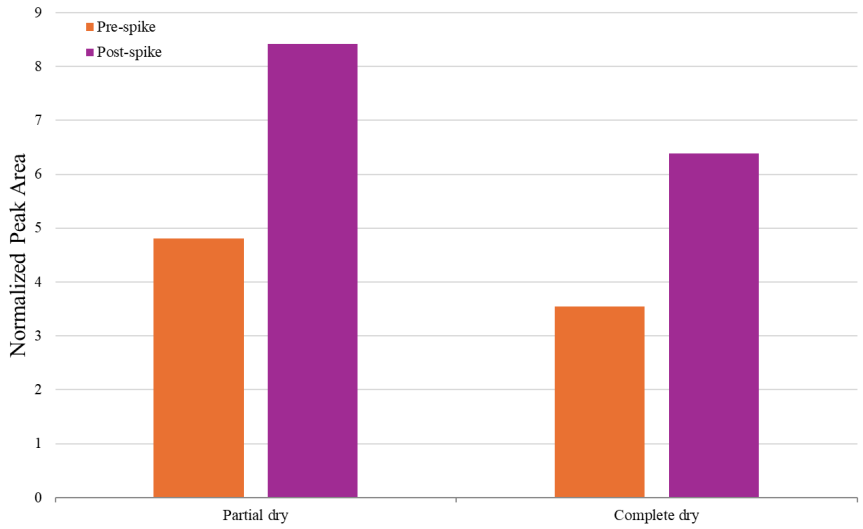


Figure 3. 14 Recovery assay for native GSW11 peptide extraction using classic protocol results
Bar graph showing the modulation of the normalized chromatographic peak area of native GSW11 peptide spike in surrogate matrix pre (orange) or post (purple) extraction, which underwent either complete or partial drying.

3.4.2.3 Application of the extraction protocol on cell pellets

We applied the developed extraction protocol and LC-MS/MS analytical method on untreated CT26 ERAAP^{naive} and ERAAP^{ko}. We observed the chromatographic peak corresponding to the native GSW11 peptide in the surrogate matrix, eluting at 1.41 minutes (Figure 3. 15 A). However, in the extracts from both ERAAP naïve and knocked out CT26 cell pellets, subjected to complete vacuum drying, the chromatographic peak at the expected retention time was absent, and we observed a prominent peak at 1.36 minutes (Figure 3. 15 B, C). Similarly, no specific peak corresponding to the endogenous GSW11 peptide was present in the pellets extracted using the method with only partial drying, from which we anticipated a 25% higher recovery (Figure 3. 15 D, E). Two factors may contribute to this issue: the poor recovery of the extraction process and the sensitivity and specificity of the analytical method, which might not be adequate to detect the peak corresponding to the endogenous peptide, expected to be present in low abundance in ERAAP naïve CT26 cells. In the knocked-out cells, where the absence of ERAAP should result in reduced degradation of the GSW11 peptide, we were expecting to observe a peak, but no differences were observed between the two conditions.

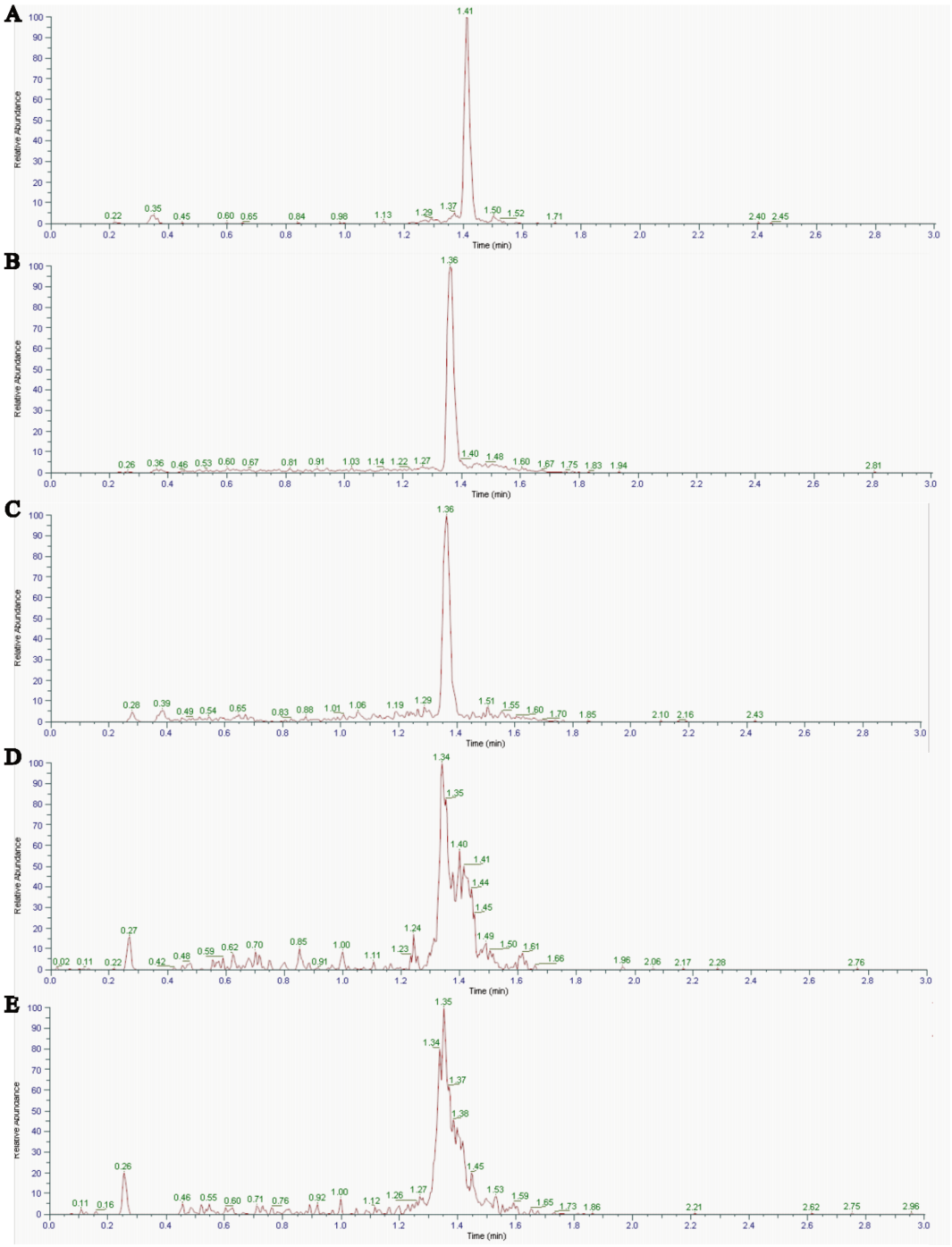


Figure 3. 15 Chromatograms of the *m/z* 1201.5>1183.46 transition after GSW11 peptide extraction

(A) Signal of the synthetic native peptide was detected in surrogate matrix at RT 1.41 min. No peaks were observed at the expected RT (1.41 min) in CT26 ERAAP^{naïve} cell extracts prepared with extraction protocol using complete (B) or partial (D) drying. Same results were observed for CT26 ERAAP^{ko} undergoing complete (C) or partial (E) drying during the extraction protocol.

3.4.2.4 Optimization of the extraction protocol with SPE

To address the low recovery of the extraction protocol (47%), one of the reasons for the missing endogenous GSW11 peptide detection, we incorporated a solid-phase extraction (SPE) step, aiming to obtain cleaner and more concentrated cell pellet extracts. Different washing and elution solutions were tested to optimize the method on surrogate matrix pre-spiked with native GSW11 peptide. The standard method, which used an aqueous solution (5% MeOH in H_2O) for washing and an organic solution (100% MeOH) for elution, led to the detection of a peak at RT 1.42 min, although it was not well resolved (Figure 3. 16 A). The "acid" method, which employed a basic solution (5% NH_4OH in ACN:H_2O 1:1, v/v) for washing and an acidic solution (2% FA in ACN:H_2O 1:1, v/v) for elution, resulted in a major peak at 1.40 min, but several other peaks were present at close RT (Figure 3. 16 B). The "basic" method, which used the acidic solution for washing and the basic solution for elution, failed to show any specific peak attributable to the peptide in the chromatogram (Figure 3. 16 C). Finally, the so-called "peptide" method, which used an aqueous, slightly acidic solution (5% MeOH, 1% AA in H_2O) for washing and an organic, more acidic solution (60% MeOH, 10% AA in H_2O), combining the advantages of the "standard" and "acid" methods, resulted in the better elution of the native GSW11 peptide and the detection of a single, well-resolved peak at 1.42 min (Figure 3. 16 D).

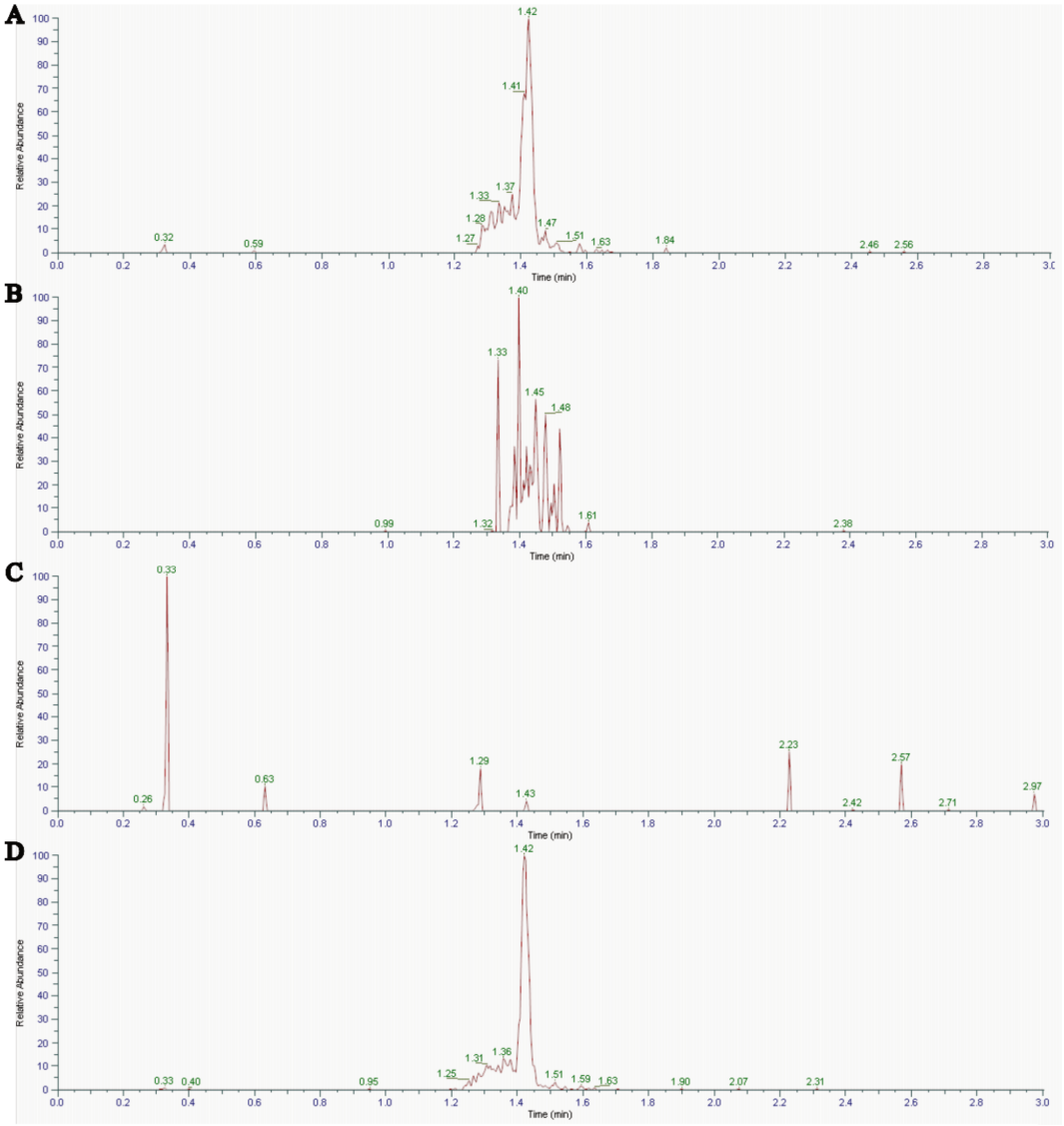


Figure 3. 16 Chromatograms of the *m/z* 1201.5>1183.46 transition for SPE optimization

Extracts were prepared with the extraction protocol applying SPE and using the (A) “standard”, (B) “acid”, (C) “basic” or (D) “peptide” methods.

3.4.2.5 Assessment of the recovery of the protocol, including an SPE step

The optimized extraction protocol, including the SPE step with the “peptide” method, was assayed to measure the recovery. The surrogate matrix was spiked with the native GSW11 peptide before each step of the protocol, to identify limiting steps. We reconsidered the role of the "Martyr" peptide, which was originally explained to influence the chromatographic elution of the peptide. We hypothesized that it might also play a role in the extraction process, potentially affecting phenomena such as adsorption onto the plastic materials used during the sample preparation steps. Therefore, we conducted the experiment in duplicate, in presence of absence of the “Martyr” peptide in the extraction solution.

We observed that the presence of the "Martyr" peptide didn't significantly impact the recovery of the native GSW11 peptide from the surrogate matrix, with an average deviation lower than 10%.

The overall recovery of the extraction protocol without the "Martyr" peptide, determined by comparing the pre- and post-spike areas, was 39%. The low recovery appears to be due to cumulative losses during the entire extraction protocol, with no specific step identified as the primary cause (Figure 3. 17).

For this experiment, we used murine cell pellets as surrogate matrix, to closely mimic the effects on the endogenous GSW11 peptide's extraction and detection from the experimental samples. The effect of the matrix, calculated as variation percentile between the native GSW11 peptide peak area detected in surrogate matrix and in solution, caused an ion enhancement of 16%.

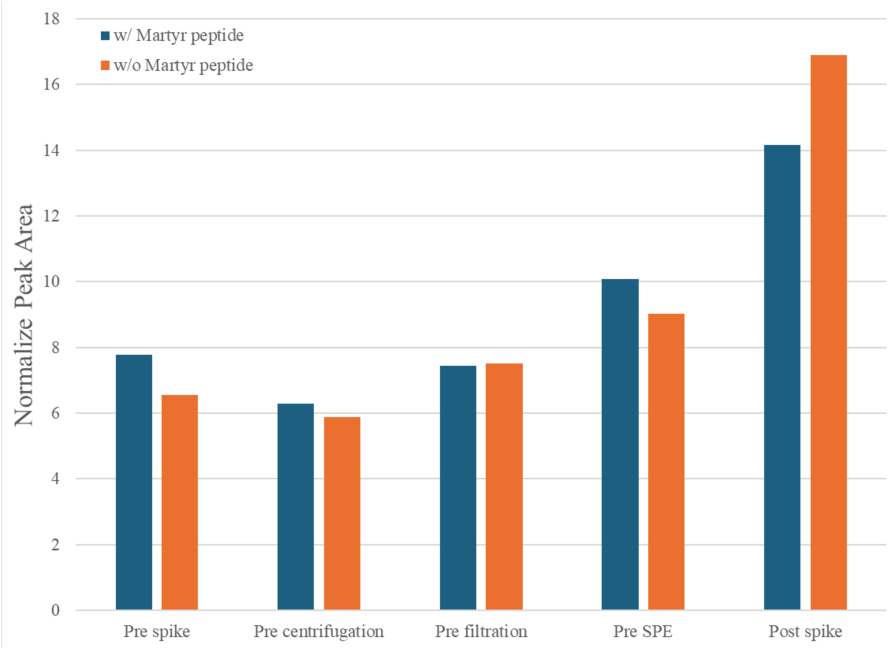


Figure 3. 17 Recovery assay for native GSW11 peptide extraction using protocol employing SPE results

Bar graph showing the modulation of the normalized chromatographic peak area of native GSW11 peptide spiked in surrogate matrix at different steps of the extraction protocol conducted in presence (w/ blue) or absence (w/o orange) of "martyr" irrelevant peptide.

Further evidence of the insufficient recovery of this method, which was comparable to, if not worse than, the previously one tested, was the absence of a peak at the expected retention time (1.41 min) in chromatograms from extracted control, untreated CT26 ERAAP^{naive} and ERAAP^{ko} cell pellets (Figure 3. 18 A). The interfering peak at 1.37 min, also observed in previous analyses, was detected in both ERAAP^{naive} and ERAAP^{ko} conditions (Figure 3. 18 B, C).

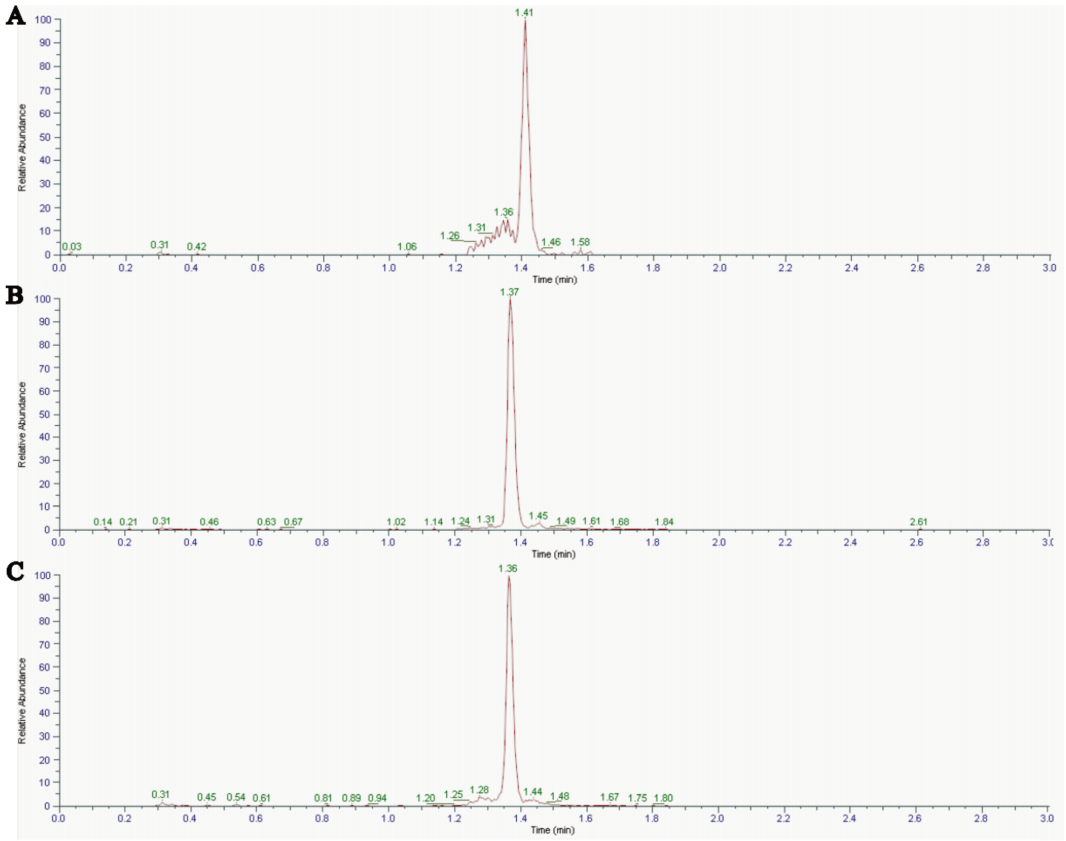


Figure 3. 18 Chromatograms of the *m/z* 1201.5>1183.46 transition after GSW11 peptide extraction employing SPE

(A) Signal of the synthetic native peptide was detected in surrogate matrix at RT 1.41 min. No peaks were observed at the expected RT (1.41 min) in (B) CT26 ERAAP^{naïve} or (C) ERAAP^{ko} cell extracts prepared with extraction protocol applying SPE and using the "peptide" method

3.4.3 Re-optimization of the MS method development for specific detection of native GSW11 peptide

3.4.3.1 GSW11 fragmentation by collision induced dissociation in positive and negative ion mode

We re-optimized the analytical method, aiming to improve its specificity and achieving the detection of the endogenous GSW11 peptide. The product ions obtained by the initial method, developed in negative ion mode, corresponded to the loss of CH_2O from a serine, the loss of H_2S from a cysteine, and the loss of H_2O from a glutamic acid (Figure 3. 19 A) (Bowie et al., 2002). These fragmentation products may not have been highly specific. We optimized a new method by performing mass analysis in positive ion mode. A bioinformatic tool (MS2PIP Server - CompOmics) was employed to predict the precursor ion fragmentation in CID, and three product ions were found experimentally (Figure 3. 19 B).

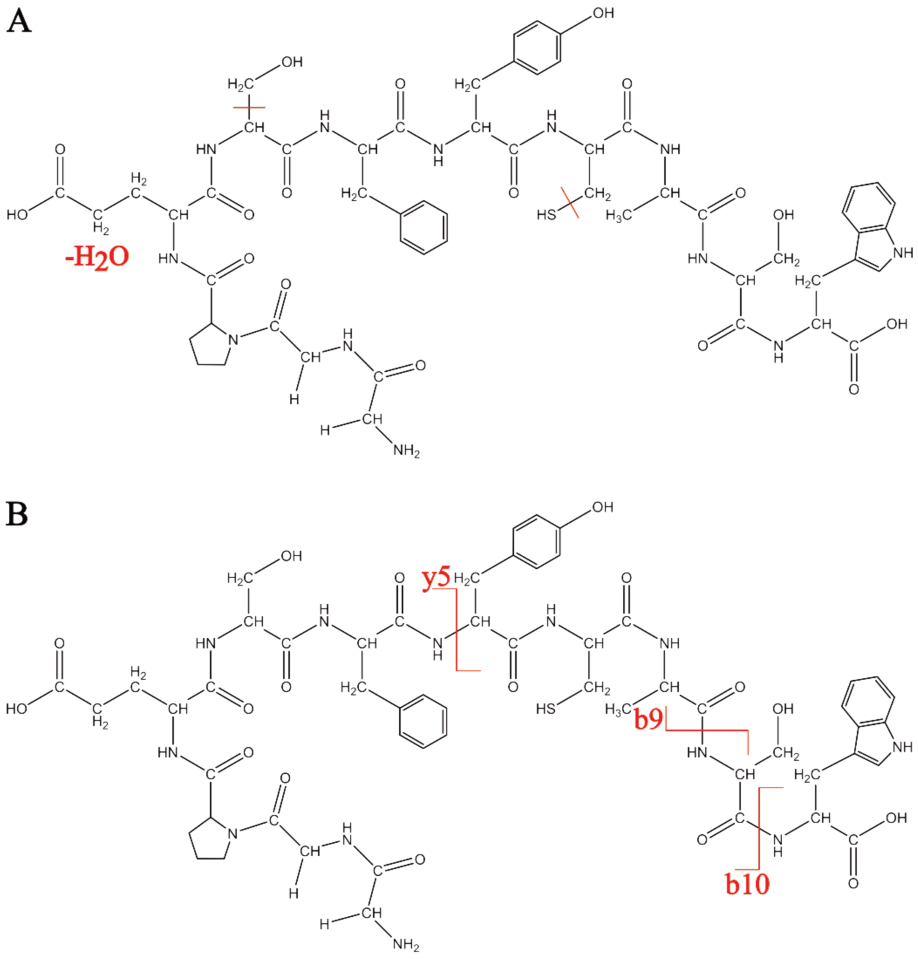


Figure 3. 19 Sites of fragmentation on the native GSW11 peptide structure generated by experimental CID in (A) negative and (B) positive ionization mode.

3.4.3.2 MRM transitions development in positive ion mode

Ionization was performed in electrospray in positive ionization mode, with a spray voltage of 3500 V. The native GSW11 and IS peaks were detected at *m/z* 1203.5 and *m/z* 1209.5 respectively (Figure 3. 20 A, D). Three product ions (transitions), cross detected between experimental and predicted native GSW11 peptide fragmentation, at *m/z* 999.35, *m/z* 912.35 and *m/z* 629.24 (Figure 3. 20 B), corresponded to the b10, b9 and y5 fragments (Figure 3. 20 G, H, I). Two corresponding product ions at *m/z* 918.35 and 1005.35 were detected for the IS peptide (Figure 3. 20 E). Collision voltages were optimized between 20 and 35 V for both native GSW11and IS peptides (Figure 3. 20 C, F).

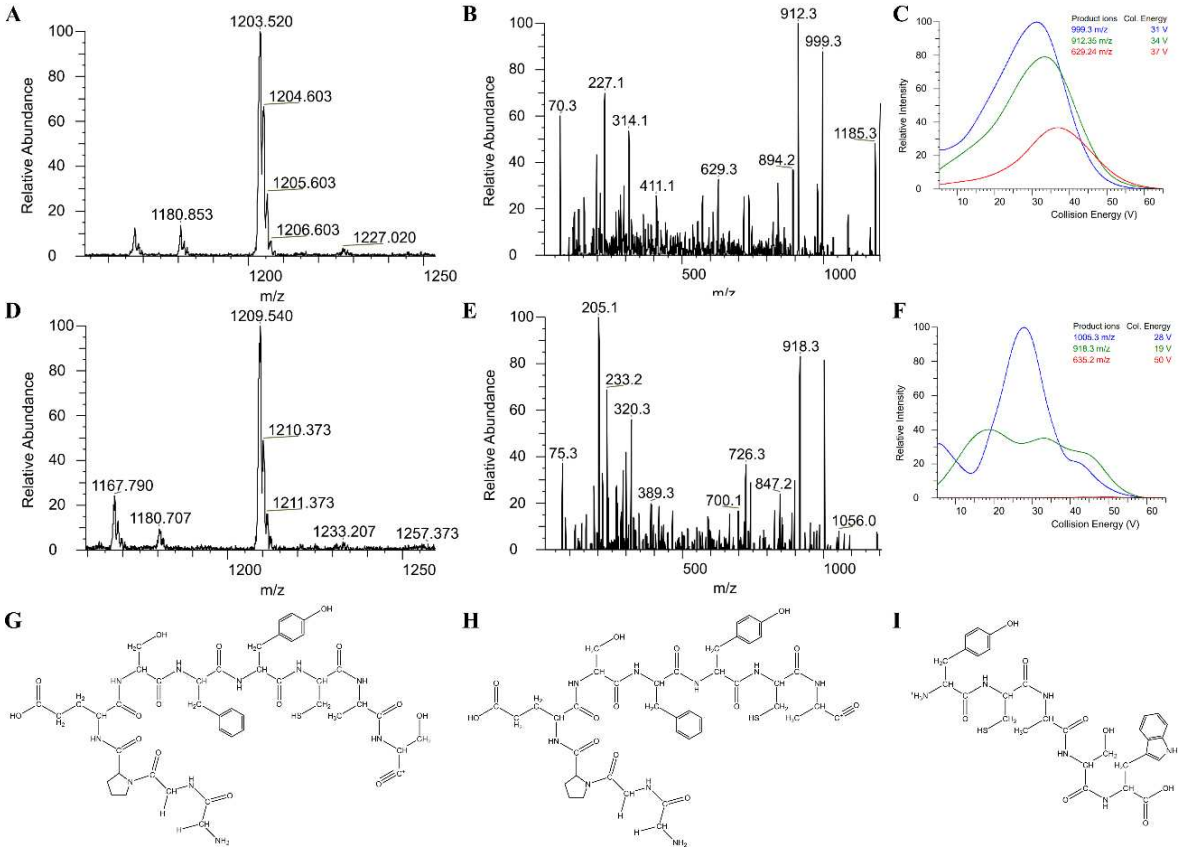


Figure 3. 20 Transitions development in positive ion mode

Precursor ion spectrum, product scan and breakdown curve at 1.5 mTorr of the native GSW11 ion *m/z* 1203.5 (A, B, C) and the IS ion *m/z* 1209.5 (D, E, F). Structures of the acylium fragment ions (G) b10, (H) b9, and the ammonium fragment ion (I) y5.

3.4.3.3 Analysis of cells extract using the MS method in positive ion mode

Despite the optimization of the analytical method to improve the specificity, our efforts to detect the endogenous GSW11 peptide remained unsuccessful. We reanalysed the cell extracts prepared using both the original extraction protocol and the one employing SPE, applying the new mass spectrometry conditions in positive ionization mode. However, we were still unable to observe any signal corresponding to the GSW11 peptide in the samples.

3.5 Technical interpretation: MS potentials and limitations

The project aimed to develop an LC-MS/MS based method for the quantification of the endogenous ERAP1-dependent antigen, GSW11, which ultimately failed in detecting the target peptide, despite optimization of both analytical method and extraction protocol. Hereafter, we discuss the technical challenges encountered during the project and identify potential missteps.

The LC-MS/MS method for the quantification of GSW11 peptide, employing surrogate matrix and isotope-labelled internal standard approaches, proved a sensitivity (LLOQ) in the nanomolar range.

James et al. reported that approximately 500 pM of GSW11 were generated from 1 million CT26 cells, based on the stimulation of CCD2Z T-cell hybridoma using cells extract compared to synthetic GSW11 (James et al., 2010). However, formulating assumption about the adequate sensitivity of the method developed is complicated by the lack of knowledge about specific expected quantity. Moreover, the titration curve was prepared by spiking the synthetic native GSW11 peptide in surrogate matrix after extraction, method that doesn't consider the loss of the peptide during the extraction process.

Low and variable recovery from biological matrix is among the challenges associated with bioanalytical methods. Analytes can be poorly extracted or lost at various steps of the sample preparation (D. Kumar et al., 2022). The evaluation of the recovery from each step of the extraction protocol tested, proved that each one contributed to the cumulative losses of the analyte. The extraction protocols applied in our study yielded suboptimal overall recovery rate (47% for the original method and 39% for the method employing SPE). Generally, a recovery rate is considered optimal if between 80% and 110%. However, lower recoveries, from 50% to 70%, may still be acceptable if they are consistent. In our case, the measured values were relatively constant, but fell below the minimum acceptance threshold, resulting in insufficient peptide extraction for detection by the analytical method.

The components, other than the analyte, present in the matrix, can contaminate the analysis and cause ion suppression, which further impact the recovery rate (Matuszewski et al., 2003). When evaluating the method's sensitivity in surrogate matrices, compared to the one in solution, a decline in performance was noted, indicating that ion suppression can be caused by the matrix effect. On the contrary, when we assessed the matrix effect on cell pellets as surrogate matrix, we measured an ion enhancement. These conflicting results proved a non-reproducible matrix effect that can impact the accuracy of the analysis and should be investigated further.

With the current level of optimization, which resulted in the inability to detect the endogenous GSW11 peptide in ERAAP naïve and knocked-out CT26 cells, and therefore investigate the modulation of its expression upon ERAP1 inhibition, we can't draw conclusions on the pharmacodynamic profile and mechanism of action of the compound under investigation.

3.6 Future works

To achieve the detection of endogenous GSW11 peptide, future research should prioritize optimizing the extraction protocol. Several reviews and research articles describing extraction protocols for peptides from various biofluids can be found in literature, although, very few report extraction protocols from cell pellets or tissues (Aristoteli et al., 2007; Chambers, 2013; W. Chen et al., 2019; Kawashima et al., 2010; Miedzybrodzka et al., 2020; Svensson et al., 2003; Vitorino et al., 2012).

Various organic solvents, for protein precipitation, in presence of acids, for peptide-protein binding disruption, could be tested. Guanidine hydrochloride and buffer containing urea, which have stronger denaturation capacity, may be required. Detergents can be helpful to extract and solubilize the peptide in solution, however they can cause severe ion suppression and therefore impact the MS analysis, requiring additional steps of removal from the extracts. Other reagents and contamination species, i.e. salts and lipids, can be removed using SPE. Alternative to the Oasis® HLB copolymer, which uses hydrophilic-lipophilic balance, are the reversed-phase C8 SPE, which would lead to a reduced retention of the peptide, anion exchange Oasis® MAX SPE, suitable for acidic compounds, or mixed mode (MM) SPE, using both reversed-phase and ion-exchange retention mechanisms.

The role of the "Martyr" peptide in the extraction protocol should also be re-evaluated. Since it did not enhance the elution as expected, we could either remove it from the protocol entirely or conduct further experiments to determine under which conditions, if any, it provides a benefit.

The matrix effect should be thoroughly studied. The surrogated matrix chosen may not be the right substitute to mimic the context created by a 100 million cell pellets. The quantity of material required to recreate this context for each point of the titration curve was not readily available, but additional investment can be beneficial to test the matrix effect in conditions similar to the experimental ones. Multiple biological matrices should be tested as surrogate matrix to reveal potential inconsistencies that affect peptide detection and quantification. Experimental samples spiked with native GSW11 peptide could be used as a control of the detection ability of the bioanalytical assay developed, and to investigate recovery and matrix effect.

In addition to the biological matrix, sample preparation steps and ionization modality and polarity can influence the matrix effects. Electrospray ionization is more subject to ion suppression, while negative ionization mode is less affected by this phenomenon, due to the fewer molecules that ionize in this polarity (D. Kumar et al., 2022; Panuwet et al., 2016). Experiment with alternative ionization techniques, such as MALDI, or different ionization conditions could be tested to enhance peptide ionization and detectability. In these new settings, doubly or triply charged ions should be reinvestigated, since highest charge typically led to better fragmentation. And in this context, more specific, and not necessarily more intense, product ion could be identified.

If the species, detected in the experimental samples, eluting at retention time 1.36 min remains in future analysis, extracts should be post-spiked with native synthetic peptide to confirm whether it originates from a contamination and its origin shall be investigated further by high-resolution mass spectrometry (HRMS).

All the further optimization steps will be conducted employing a clear workflow timeline, systematically changing one variable at the time, and using controls at each stage, like known concentrations of the native GSW11 peptide spiked into the appropriate surrogate matrix. This

approach will help assessing the variability and reliability of the method and making more informed decisions about the best paths to follow.

4 Chapter 4 - Comprehensive insights into the DMPK profile of an ERAP2 inhibitor: integration of whole-body biodistribution and *in vivo/in vitro* metabolism analysis

4.1 Rationale

In early drug development, the primary objective is to select promising compounds that will enable safe and effective doses and regimens. Integrating pharmacokinetics (PK) and pharmacodynamics (PD) plays a key role in guiding compound selection. The support from preclinical models and bioanalytical assays, ensure that compounds meet critical attributes, such as safety, efficacy, and pharmacokinetics, defined by the target product profile (TPP). Meeting the TPP is essential to advance through the drug development program and minimize the risk of failure in clinical stages (Miller et al., 2005; Wyatt et al., 2011).

Understanding compound's exposure and bioavailability at the target site is essential to determine its potential target engagement and therapeutic value. The determination of the level of exposure at target site or in plasma that triggers the observed pharmacological effect will be used as a target concentration when dosing a new animal species with different PK or bioavailability in the R&D process. Effects of the drug candidates must be evaluated on- and off-target, to address issues with secondary, undesirable engagement and toxicity. Measuring the concentration within relevant tissues is essential to evaluate whether the compound reaches therapeutic levels necessary for target engagement or if it, and its metabolites, distributes in secondary organs (Reichel & Lienau, 2016). The study of compound's metabolism is equally important to understand the stability and anticipate potential toxic effects or discover active metabolites.

Spectral information about compounds can be obtained by analysing biofluids and tissue extracts, and LC-MS/MS remains the gold standard for quantification. Depending on the analyzer used, this technique can achieve high mass accuracy and measure the fragmentation patterns, as multiple analytes might contribute to the same signal. However, LC-MS/MS cannot provide a better spatial resolution than the dissection procedure.

MALDI mass spectrometry imaging (MSI) serves as a complementary technique, providing spatially resolved visualization of compound's biodistribution. The lateral resolution of a MALDI-MSI experiment typically ranges from a few tens of micrometers for isolated organs, to a few hundred micrometers for whole-body analysis. This resolution allows to investigate the distribution within histological regions when examining isolated organs, and within tissues and organs when analysing whole-bodies. Using an ICR mass analyzer in this context allows to obtain high mass accuracy in the low mass range. The unlabeled and multiplexed capabilities of MSI make it a valuable complement

to classic methods to investigate in a single assay compound and metabolites localization and quantification. The detailed distribution of drugs and metabolites provides insights for evaluating both efficacy and safety (Granborg et al., 2022).

In this chapter, we aim to characterize the behaviour and DMPK profile of a potent and selective ERAP2 inhibitor by investigating its biodistribution and metabolism, combining whole-body qMSI and *in vivo/in vitro* metabolism studies. By mapping the spatial distribution of the compound and its metabolites in various tissues, and measure its exposure in target organs, we can gain insights about the compound's target engagement, accumulating organs, and potential effects off-target.

4.2 Context

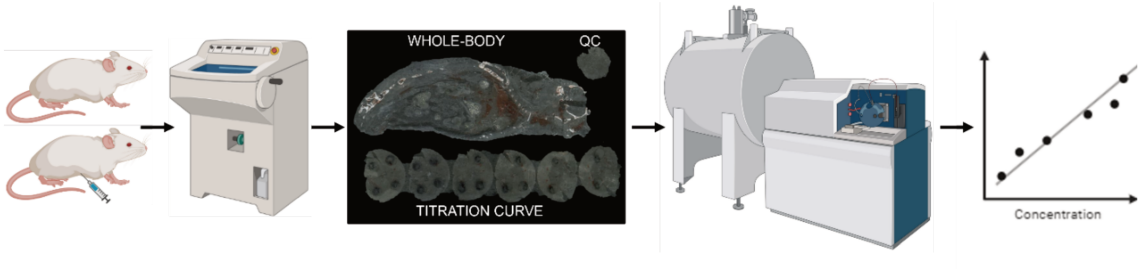
ERAP2 is implicated in autoimmune diseases and cancer, due to its role in modulating MHC-I-mediated antigen presentation. It has been identified as a risk factor in ankylosing spondylitis (HLA-B27), Behcet's Disease (HLA-B51), psoriasis (HLA-C06:02), and Birdshot chorioretinopathy (HLA-A29) (Guasp et al., 2019; Kuiper et al., 2018; Lorente et al., 2020; Wiśniewski et al., 2018). In oncology, altered ERAP2 expression correlates with prognosis, survival and therapy response in oral cavity squamous cell carcinoma and bladder cancer (Kuo et al., 2017; Lim et al., 2018).

ERAP1 is the most studied paralogue between the two endoplasmic reticulum aminopeptidases, with far fewer inhibitors developed for ERAP2 (Fougiaxis et al., 2024). ERAP2 has a distinct substrate specificity and targeting it can reshape a distinct repertoire of presented antigens and rewire immunosurveillance in conditions affected by immune-escape and dysregulation.

The first selective nanomolar ERAP2 inhibitor was reported in 2022, demonstrating promising *in vitro* effects on antigen presentation, with a favourable ADME profile and *in vivo* exposure levels consistent with concentrations needed for target engagement (Camberlein et al., 2022). Confirming these results through *in vivo* biodistribution studies will allows to better understand the compound's behaviour in the body and assess its therapeutic potential.

This chapter of the thesis is a critical component of the *compound validation* phase within of the ERAPs inhibitor drug discovery program, providing an analysis of its distribution, metabolism, and potential therapeutic and toxic implications.

4.3 Experimental design



Graphical abstract 3 Quantitative mass spectrometry imaging for biodistribution study
Whole-body sections were obtained from treated animals and vehicle controls. Spatially resolved mass spectra were acquired to enable a quantitative assessment of the compound's biodistribution.

4.3.1 Experimental cohort and sample preparation

A 10 ml/kg dose of ERAP2 inhibitor, provided by Prof Deprez-Poulain (University of Lille), formulated in 30% kleptose aqueous solution, was administered intraperitoneally (IP) to female CD1 mice, resulting in final dose of 50 mg/kg, tolerable *in vivo*. Two mice received the compound, called Treated 1 and Treated 2, while one mouse, referred to as Vehicle, received the vehicle only. The time point for sacrifice, 0.33h (20 min), was selected based on previous experimental assessment of the t_{max} from a PK study performed in plasma under the same conditions (Figure 4. 1).

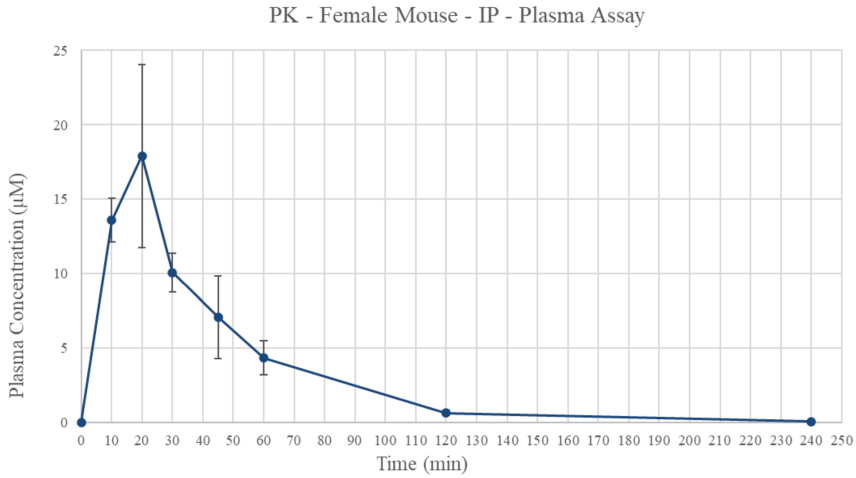
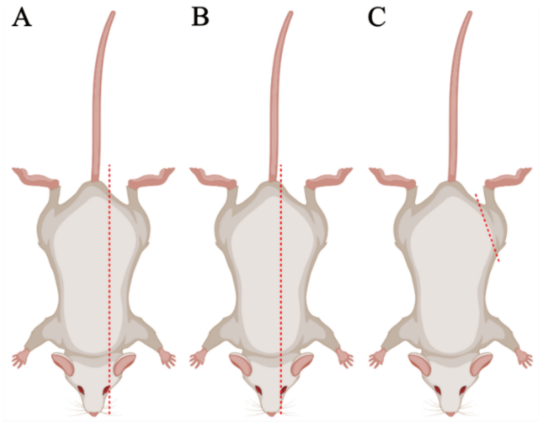


Figure 4. 1 *In vivo* PK profile of the ERAP2 inhibitor
Graph of compound's plasmatic concentration (µM) over time (min) in plasma from female CD1 mouse which received 50 mg/kg of ERAP2 inhibitor administered IP

Carcasses were snap frozen in an isopentane bath cooled down with dry ice for at least 5 minutes and stored at -80°C until use. 20 µm thick tissue sections were collected for each animal using a Microm HM 560 cryostat set at -25°C. Tissue sections were thaw-mounted on indium tin oxide (ITO) slides

for MSI analysis. Two sagittal section planes, intersecting at the left eye plane, were selected to expose most of the organs to the analysis. Additionally, the left hind leg of the animals was collected, and sections of the ankle joints were added onto the ITO slides (Figure 4. 2).



D

	Eye	Brain	Brown adipose tissue	Lung	Liver	Stomach	Kidney	Spleen	Intestine	White adipose tissue
Section plane 1	++	-	+	+	++	++	+	++	++	+
Section plane 2	+	+	++	++	+	+	++	-	++	++

Figure 4. 2 Study animals sectioning plan

(A) Section plane 1, (B) section plane 2, (C) plane of extraction of the ankle joint. (D) Non exhaustive list of main organs observed per section plane. "-" absent, "+" present, "++" present and recommended.

4.3.2 MALDI MSI method development

MALDI matrices were tested to select the polarity for optimal desorption and ionization of the candidate compound under analysis, hereafter referred to as "analyte" ($C_{23}H_{23}N_5O_5S_2$) (Figure 4. 3 A), and its 4-deuterated analogue, hereafter referred to as "internal standard" (IS) ($C_{23}H_{19}N_5O_5S_2D_4$) (Figure 4. 3 B). For this purpose, 2,5-dihydroxybenzoic acid (DHB) and 1,5-diaminonaphthalene (DAN) were chosen based on the molecular weight of the analyte, a small molecule weighing 513.60 g/mol. The analyte and IS were dissolved in pure DMSO at 10 mM creating standard solutions used for development and quantification purposes. 1 µl of analyte or IS solution at 10 or 100 µM were spotted onto an ITO slide and mixed with a saturated solution of either DHB or DAN (Figure 4. 3 C). Two polarities were tested: positive ionization mode for the analysis using DHB matrix, and negative ionization mode for the analysis using DAN.

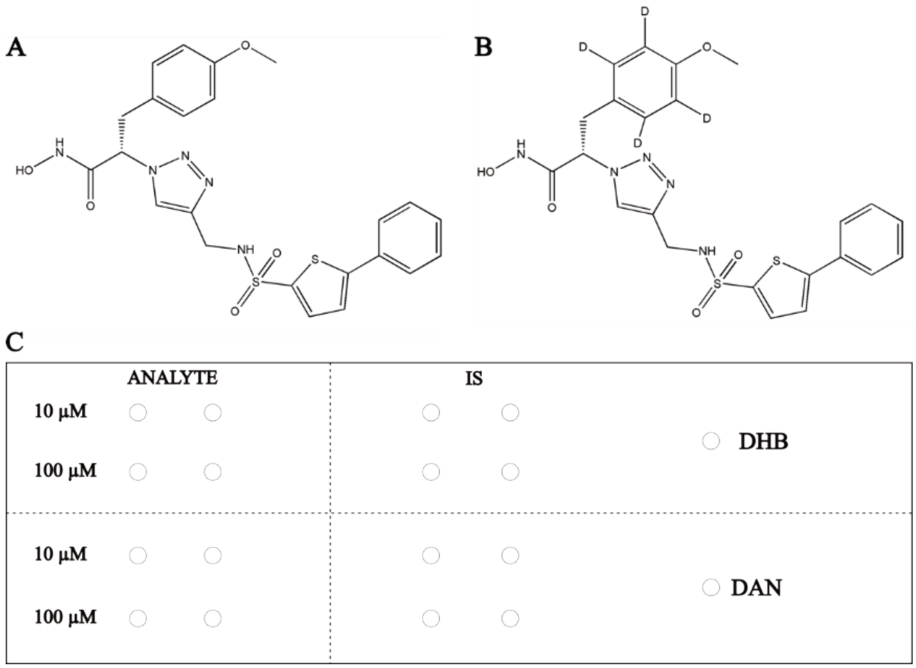


Figure 4. 3 Schematic representation of the experiment to test the optimal matrix for analyte's detection Chemical structure of the candidate ERAP2 inhibitor under analysis (A), referred to as "analyte" and its 4-deuterated analogue (B) referred to as "IS". (C) Schematic representation of the ITO slide prepared to test optimal matrix and polarities to ionize the analyte and the IS.

Once the matrix was selected, different ratios of solvents with the addition of 0.1% TFA were tested to optimize ionization conditions. This and further optimizations were performed by spotting 1 µl of analyte and IS solutions at 10 and 100 µM on surrogate matrix, mouse liver tissue sections. For this experiment, one slide was prepared and sprayed with either a solution of DHB at a concentration of 40 mg/mL in MeOH/H_2O at rations of 1:1 or 7:3 (v/v), with each covering half of the slide (Figure 4. 4). MSI acquisition was then performed with the same method in a single run.

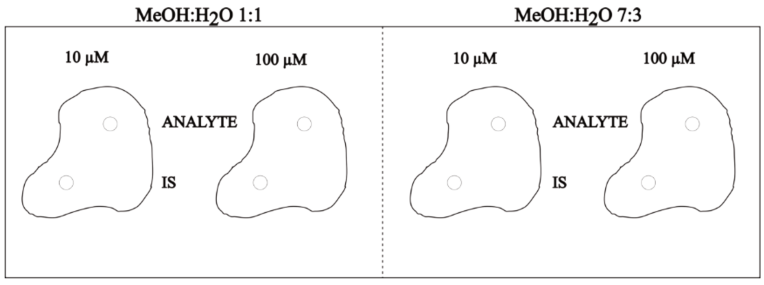


Figure 4. 4 Schematic representation of the experiment to test the optimal solvents ratio

4.3.2.1 Washing step

1 µl of analyte and IS solutions at a concentration of 10 and 100 µM, or an 8-points titration curve, were spotted onto mouse liver tissue sections (Figure 4. 5). Tissue sections were subjected to a

washing procedure, conducted by dipping portion of the slides for 15 seconds in prechilled solutions. The solutions were prepared by diluting 10 mM ammonium hydroxide in distilled water to create a basic solution. Additionally, K^+-doped water was made by adding 10 mM KCl to distilled water, while Zn^{2+}-doped water was prepared by dissolving 100 mM zinc sulphate in distilled water. The selection of washing solutions was based on the solubility and properties of the analyte to ensure effective contaminant removal while preventing loss of the analyte. After washing, slides were air-dried at room temperature before DHB matrix application.

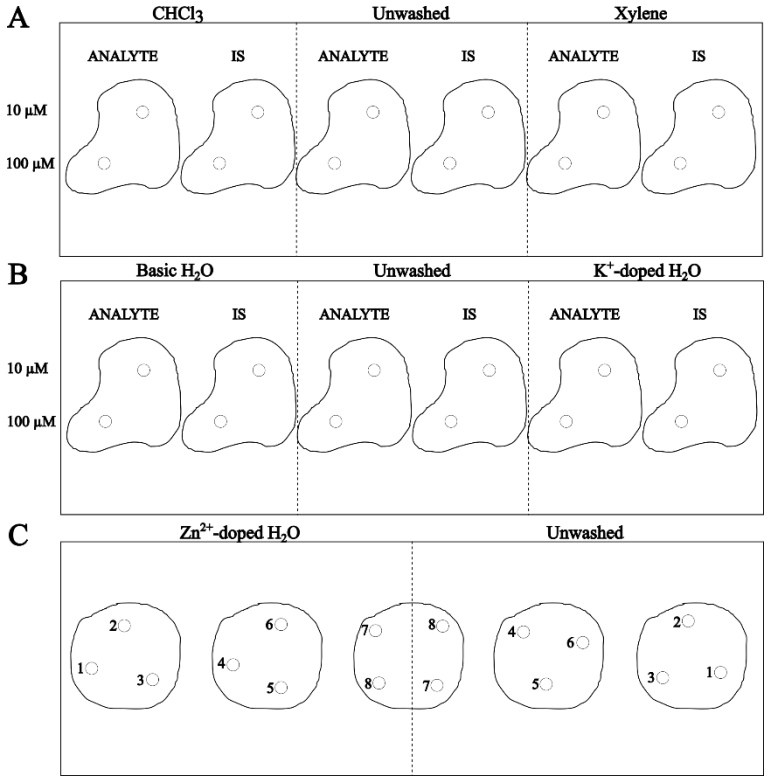
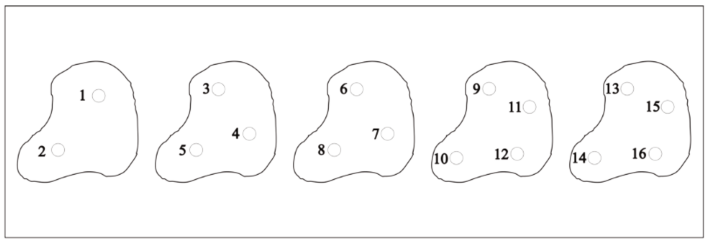


Figure 4. 5 Schematic representation of the experiments to test the washing solutions

4.3.2.2 Dilution series preparation

For quantification purposes, including determining LOD, LLOQ, upper limit of quantification (ULOQ), fifteen (15) dilutions of the analyte ranging from 0.15 µM to 250 µM were prepared in MeOH/H_2O 1:1 (v/v) (Figure 4. 6 B). One (1) microliter of each working solution (WS) was spotted on mouse liver tissue sections and placed in a desiccator for 15 min before MALDI matrix deposition. A spot of pure solvent was added as a negative control (Figure 4. 6 A).

A



B

ID	WS1	WS2	WS3	WS4	WS5	WS6	WS7	WS8	WS9	WS10	WS11	WS12	WS13	WS14	WS15	WS16
Concentration (µM)	250	125	100	75	50	30	25	12.5	6	3	1.25	0.8	0.5	0.25	0.15	0

Figure 4. 6 Schematic representation of experiments to test method's limits

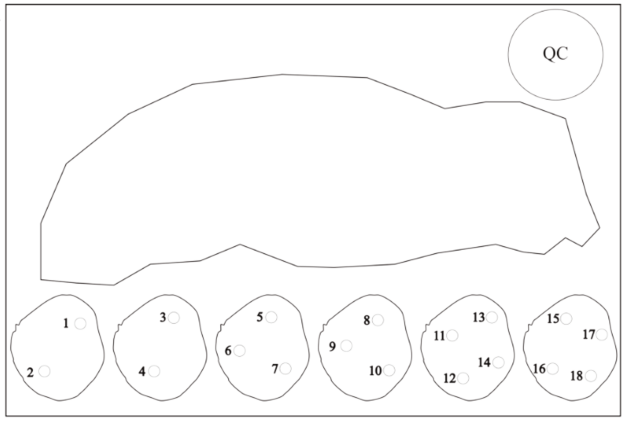
(A) ITO slide prepared to test LOD, LLOQ and ULOQ. (B) Table reporting concentrations of the standard solutions prepared to build the titration curve.

4.3.3 Biodistribution study

Whole-body tissue sections on ITO slides were placed in a desiccator for 15 min prior to matrix deposition. This step allows a perfect drying of the tissue and optimized matrix effects. A QC, consisting of liver homogenate spiked with olanzapine, which efficiently ionizes in positive mode using DHB as MALDI matrix, at 40 µg/g of tissue, was added to each ITO slide to monitor signal variability across acquisitions (Figure 4. 7 A). An optical image of each slide was acquired with a scanner to synchronize positions of the tissue sections with the target of the laser. The DHB MALDI matrix was spiked with the IS at a final concentration of 1 µM and sprayed with an automatic sprayer (TM-Sprayer, HTX-Imaging). The same spraying method was applied to all subsequent images for optimal data consistency.

One or two whole-body images were acquired daily over a week. An eighteen-points titration curve was acquired each day of analysis, for a total of five curves (Figure 4. 7 B).

A



B

ID	WS1	WS2	WS3	WS4	WS5	WS6	WS7	WS8	WS9	WS10	WS11	WS12	WS13	WS14	WS15	WS16	WS17	WS18
Concentration (µM)	250	125	100	75	50	30	25	12.5	6	4	3	2	1.25	0.8	0.5	0.25	0.15	0

Figure 4. 7 Schematic representation of the experiment to study biodistribution and quantification of the ERAP2 inhibitor *in vivo*

(A) ITO slide holding one whole-body section, one QC section, and 6 sections of mouse liver tissue sections as surrogate matrix for spotting the titration curve. (B) Table reporting concentrations of the standard solutions prepared to build the titration curve.

4.3.3.1 MALDI MSI analysis

All MSI acquisitions were performed with the same mass method: positive ionization, continuous accumulation of selected ions (CASI) mode, mass range between 464 and 564 Da, and laser frequency at 2000 Hz. CASI enhances sensitivity by selecting specific ions with a mass selection quadrupole and trapping ions generated from multiple laser shots for stronger signal detection. Online calibration was performed using the IS signal at *m/z* 518.1464. The lateral resolution was set at 350 µm for whole-body imaging and 200 µm for the titration curve points and the olanzapine QC. Data reduction factor was set at 98%. Save reduced profile spectrum is an option that allows for the compression of acquired data, including both the spectrum and the image, to prevent the size from becoming excessively large during imaging acquisitions.

4.3.3.2 Data treatment

Normalization of the data was performed by the IS signal introduced in the MALDI matrix solution before the spraying step. The normalization was based on a ratio between analyte and IS intensities per pixel. Convolution step was performed on the images using a normalized uniform kernel which simply averages the values around a position. The signal of the analyte was evaluated in each whole-body section imaged to check if a real peak or only background noise was observed through the analysis prior to quantification. Intensity scales representing the molecular signal were adjusted for each image to discriminate the noise from the molecular signal and give the best visualization of the signal across the sections.

4.3.3.3 Absolute quantification

Optical scans of the whole-body sections were directly used to annotate the histological regions of interest (ROIs). The following organs and tissues were selected for ROIs demarcation on Multimaging™ software: brain, eye, spleen, liver, kidney, stomach, heart, lung, spleen, ankle, pancreas, and white and brown adipose tissue. These organs were exposed in the section planes that were selected to investigate the pharmacological relevance and safety of the compound. The quantitation of the analyte was performed using the IS approach. Each image was quantified using the titration curve generated by acquiring a series of dilutions on the same day. At least 75% of the calibration points on the calibration curve had back-calculated concentrations with an accuracy within ±30% of the nominal concentration. A correlation between the calibration curve and the signals obtained from the tissue samples allowed to determine the analyte concentration in each histological ROI, reported in micrograms per gram (µg/g) of tissue and micromolar (µM) units.

4.3.4 Metabolism study

Phase I metabolism was predicted using structural analysis and the bioinformatic tool BioTransformer 3.0 (Wishart et al., 2022). Phase I (CYP450) transformation in combined mode, using both rule-base method and machine learning model and combining the results, were selected and three reaction iteration were calculated.

BioTransformer 3.0 was also used to predict Phase II metabolism, selecting Phase II transformation and calculating one reaction iteration.

4.3.4.1 Microsomal incubation for in vitro metabolism study

The analyte was incubated at 50 µM with 1 mg/ml of mouse female microsomes at 37°C for 2 hours. The reaction mix was prepared as follow: 50 mM phosphate buffer, pH 7.4, in the presence of $MgCl_2$ (5 mM), nicotinamide adenine dinucleotide phosphate (NADP) (1 mM), glucose-6-phosphate dehydrogenase (0.4 U/mL), and glucose-6-phosphate (5 mM). Two control reactions were prepared: one lacking NADP (enzymatic control) and one using PBS instead of the reaction mixture (matrix control). To study *in vitro* metabolism a 150 µL aliquot was taken at intervals of 0, 30, 60, 90, and 180 minutes. The reaction was stopped by adding 4 volumes of acetonitrile. After centrifugation at 12000 g for 10 minutes at 4°C, the supernatants were evaporated to dryness (35°C, 5h) and reconstituted in 150 µl of ACN/H_2O 1:1 (v/v). 1 µl of each experimental and control sample was spotted on a MTP 384 target plate and mixed with 1 µl of DHB (40 mg/mL in MeOH/H_2O 1:1 v/v) for MALDI MS analysis.

For each experimental and control sample, spectra were obtained by averaging signals from 32 scans. The mass method used positive ionization, CASI mode, mass range from 414 to 614 Da, and laser frequency at 2000 Hz.

4.3.4.2 Whole-body images acquisition for in vivo *metabolism study*

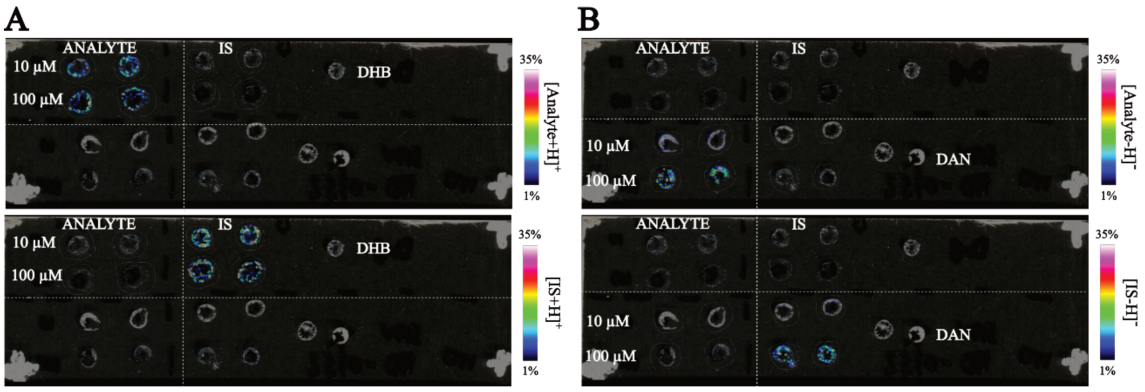
Treated whole-body tissue sections (section plane 2) were prepared and sprayed as described previously. To investigate the predicted metabolites and performed the exploratory analysis of Phase I metabolism, MSI acquisitions was performed with the mass method: positive ionization, CASI mode, mass range between 414 and 614 Da, and laser frequency at 2000 Hz. To investigate the predicted metabolites from Phase II metabolism, MSI acquisitions was performed with the mass method: positive ionization, full scan (FS) mode, and laser frequency at 2000 Hz. Online calibration was performed using the DHB signal at *m/z* 273.0393. The lateral resolution was set at 350 µm. Data reduction factor was set at 98%, because of the large data generated by the imaging analysis.

4.4 Results

4.4.1 MALDI MSI method development for optimal analyte's detection

4.4.1.1 Selection of MALDI matrix and polarity

The selection of the optimal MALDI matrix for the ionization of the analyte and IS was based on signal intensity from pure standards spotted on ITO glass slides at concentration of 10 and 100 µM and mixed with DHB or DAN. In MALDI, a pulsed laser irradiates the sample and excites the matrix, which redistributes the energy for desorption and ionization of analytes, facilitated by proton transfer, thereby generating stable ionized species. The analyte and IS were detected by FTICR-MSI in both positive and negative ionization modalities (Figure 4. 8 A, B). Higher average intensity was measured for standards mixed with DHB and analysed in positive ionization mode, showing 0.5 to 1.5 log difference compared to those mixed with DAN and analysed in negative ionization mode (Figure 4. 8 C).



C

	Concentration	Matrix	*m/z*	Adduct	Average Intensity
Analyte	10 µM	DHB	514,1213	$[M+H]^+$	1,40E+07
	100 µM				1,67E+07
	10 µM	DAN	512,1067	$[M-H]^-$	7,89E+05
	100 µM				5,65E+06
IS	10 µM	DHB	518,1464	$[M+H]^+$	2,42E+07
	100 µM				2,49E+07
	10 µM	DAN	516,1318	$[M-H]^-$	6,64E+05
	100 µM				5,87E+06

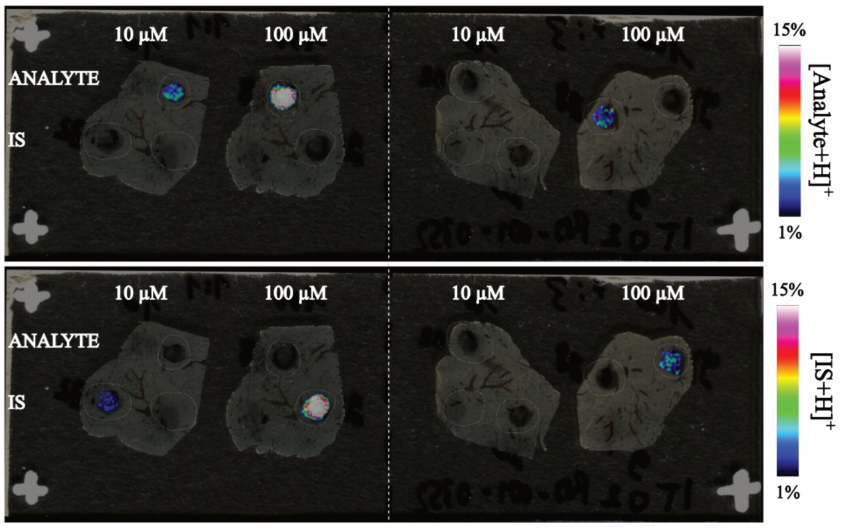
Figure 4. 8 Optimal matrix section results

Overlay of ITO slide scans with MSI images of analyte and IS pure standards spotted, mixed with DHB or DAN, and acquired in positive (A) or negative (B) ion mode. (C) Table reporting the average intensity detected in the acquired regions around the spotted drop of pure standard and matrix.

The optimization of the MS method and MALDI matrix solution for the detection of analyte and IS was based on the signal intensity measured from pure standards spotted on mouse liver tissue sections at 10 and 100 µM followed by two different matrix spraying (Figure 4. 9 A). The MALDI matrix 2,5-DHB prepared at 40 mg/ml in MeOH/H_2O 1:1 (v/v) with 0.1% TFA (v/v), yielded the highest average

intensities for both analyte and IS under $[M+H]^+$ adduct (Figure 4. 9 B). This preparation was selected for further evaluation of the LOD and LLOQ.

A



B

	Concentration	Matrix	Solvent	*m/z*	Adduct	Average Intensity
Analyte	10 µM	DHB	MeOH:H_2O 1:1 w/ 0.1% TFA	514,1213	$[M+H]^+$	8,61E+04
	100 µM					5,85E+05
	10 µM		MeOH:H_2O 7:3 w/ 0.1% TFA			1,48E+04
	100 µM					7,13E+04
IS	10 µM	DHB	MeOH:H_2O 1:1 w/ 0.1% TFA	518,1464	$[M+H]^+$	5,96E+04
	100 µM					7,92E+05
	10 µM		MeOH:H_2O 7:3 w/ 0.1% TFA			2,07E+04
	100 µM					1,16E+05

Figure 4. 9 Solvent selection results

(A) Overlay of ITO slide scans with MSI images of analyte and IS pure standards spotted on mouse liver tissue section, sprayed with DHB matrices, and acquired in positive ion mode. (C) Table reporting the average intensity detected in the acquired regions around the spotted drop of pure standard.

4.4.1.2 Evaluation of method sensitivity

The estimation of the LOD and LLOQ was based on a titration curve of diluted analyte standard spotted on mouse liver tissue sections at sixteen (16) concentrations, ranging from 0.15 µM to 250 µM, and followed by matrix spraying with the IS at 1 µM. Data normalization was based on the per-pixel intensities ratio between the analyte and IS. The LOD is defined as the first calibration point showing a peak at the expected *m/z* and with a signal intensity that is greater than three times the noise level. The LLOQ was defined as the first point with signal and within the linear calibration range. Two weighted linear 1/X curves for high and low intensities ranges with R^2 0.975 and 0.991 were obtained. The LOD and LLOQ of the analyte were found at 1.54 µg/g of tissue (Figure 4. 10).

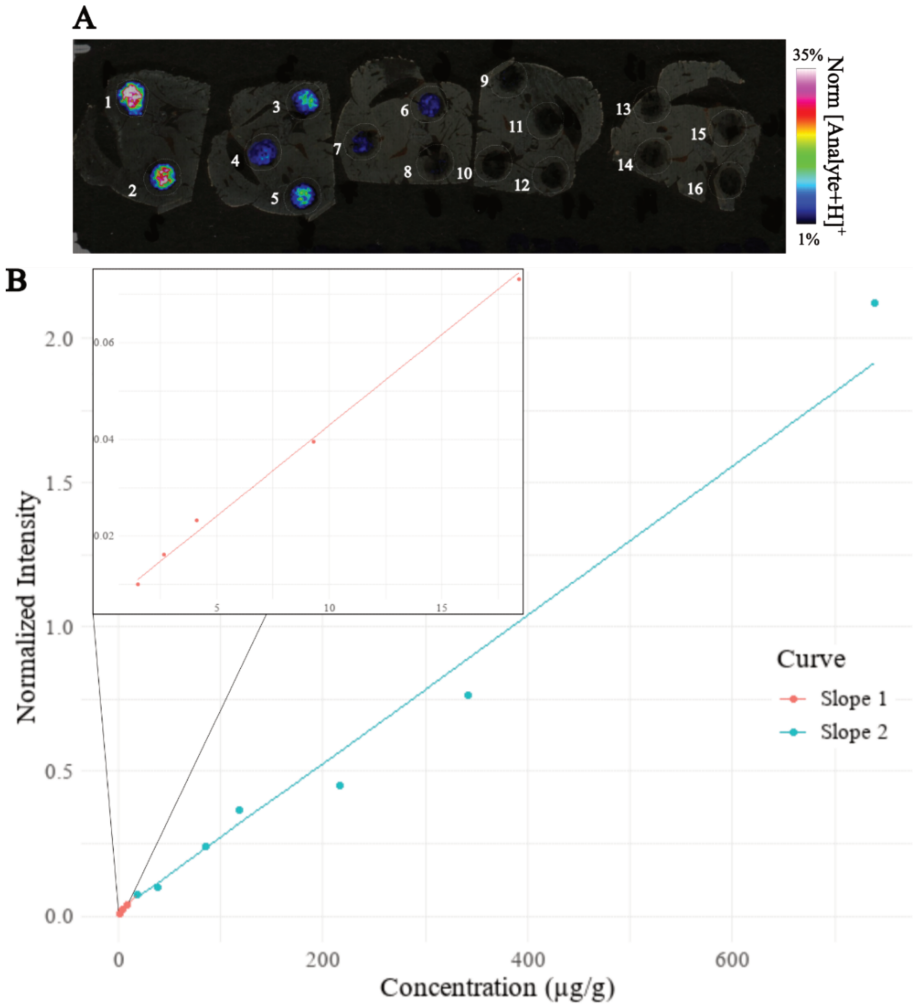


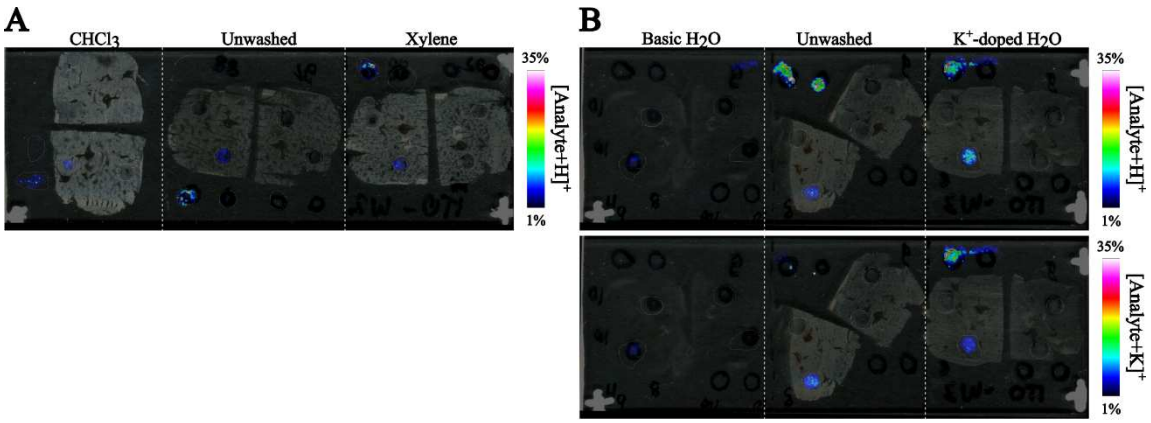
Figure 4. 10 Analyte's titration curve to evaluate method sensitivity

(A) Overlay of ITO slide scan with MSI images of the analyte's titration curve spotter on mouse liver tissue section, sprayed with DHB matrices, and acquired in positive ion mode. (B) Linear weighted 1/X curves used for LOD and LLOQ estimation, plotted in a graph intensity over concentration (µg/g).

4.4.1.3 Washing steps to improve the sensitivity

We incorporated washing steps to improve the MSI analysis sensitivity and the LOD. Given the solubility of the analyte in polar solvents such as ACN and MeOH, we selected organic and highly apolar solvents, xylene and chloroform, as washing solutions, to remove potential apolar contaminants without washing away the compound (Figure 4. 11 A). Additionally, the ionization in positive mode indicated a potential basic character of the analyte. This led to the hypothesis that washing with a basic solution could enhance the detection. However, it appears that this approach may have increased the solubility of the compound, affecting its retention during analysis (Figure 4. 11 B). We observed that the compound could also ionize under the $[M+K]^+$ form. We hypothesized that switching the adduct balance from $[M+H]^+$ to $[M+K]^+$ could lead to improved sensitivity. To

address this, we used potassium-doped water as a washing solution (Figure 4. 11 B). To evaluate the performance of the washing solutions, we measured the intensities ratio between signals from standard solutions spotted on mouse liver tissue sections with or without washing steps prior to matrix spraying. None of the tested washing solutions consistently or significantly improved the sensitivity in detecting the investigated analyte adducts (Figure 4. 11 C).



C

	Concentration	*m/z*	Adduct	Washing solution	Intensity Ratio
Analyte	10 µM	514,1213	$[M+H]^+$	$CHCl_3$	0.59
	100 µM				0.67
	10 µM			Xylene	0.5
	100 µM				0.79
	10 µM			Basic H_2O	1.51
	100 µM				0.36
	10 µM	552,0772	$[M+K]^+$	K^+-doped H_2O	0.83
	100 µM				0.54

Figure 4. 11 Washing steps to improve method sensitivity results

(A-B) Overlay of ITO slide scans with MSI images of analyte pure standard spotted on mouse liver tissue section, washed with one of the solutions, sprayed with DHB matrices, and acquired in positive ion mode. (C) Table reporting intensity ratios between signal from washed regions and unwashed.

To exploit the high affinity of the analyte for Zn (II) ion, which is linked to the presence of a hydroxamate function, we tested an additional washing solution: zinc-doped water. We aimed to detect one or more of the adducts that could originate from the coordination complex between the analyte and Zn (II): $[M+Zn]^{2+}$, $[2M+Zn]^{2+}$, $[3M+Zn]^{2+}$, $[4M+Zn]^{2+}$. The presence of the adducts was investigated by accumulating spectra using MALDI-FTICR-MS from washed and unwashed spots of the analyte's pure standard spotted on mouse liver tissue sections at different concentrations. Spectra were acquired in full scan and CASI mode, selecting the *m/z* range of interest for each adduct. A signal at *m/z* 545.07, potentially representing the $[2M+Zn]^{2+}$ adduct, was detected (Figure 4. 12 B). However, it was also present in the unwashed control and the acquisition of MS image from pure standards spotted on mouse liver tissue sections, suggested that the signal likely originated from the matrix, based on its spatial distribution (Figure 4. 12 E). No other complex between Zn (II) ion and the analyte was detected after the washing step with zinc-doped water (Figure 4. 12 A, C, D).

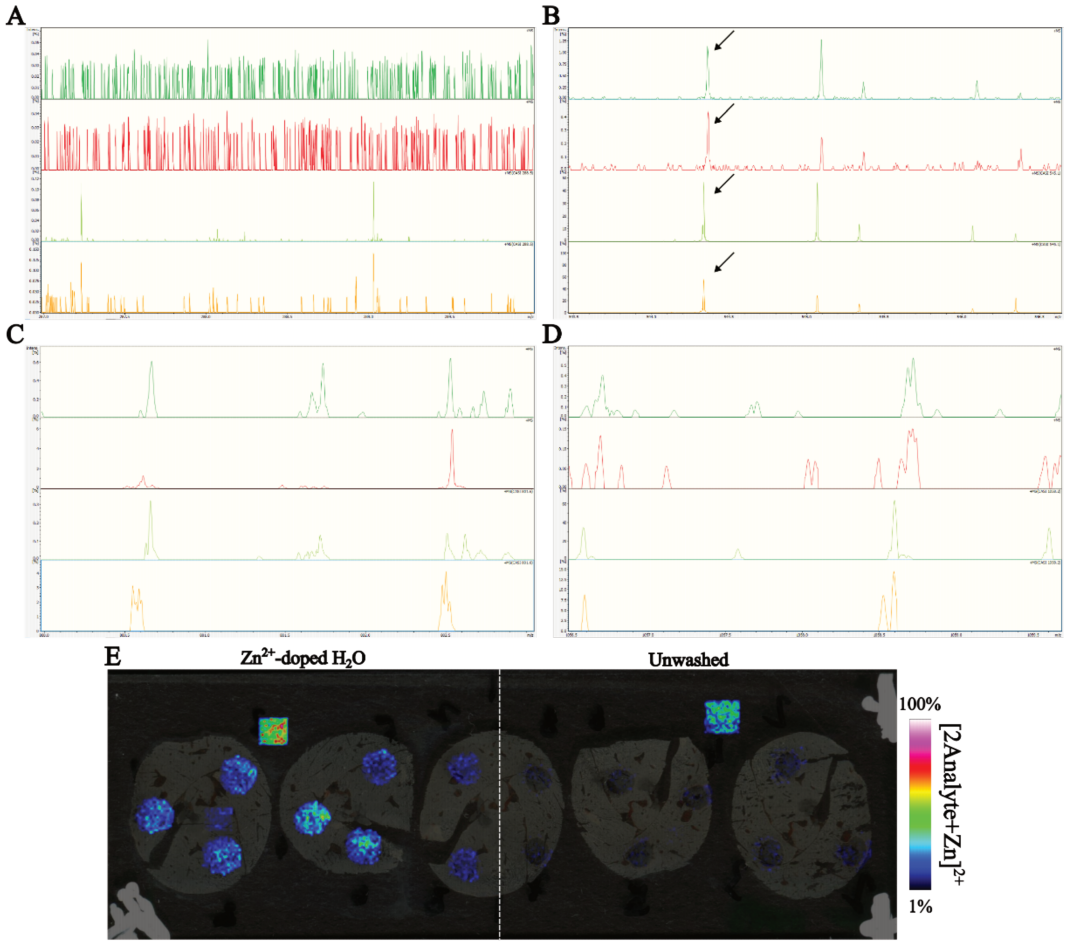


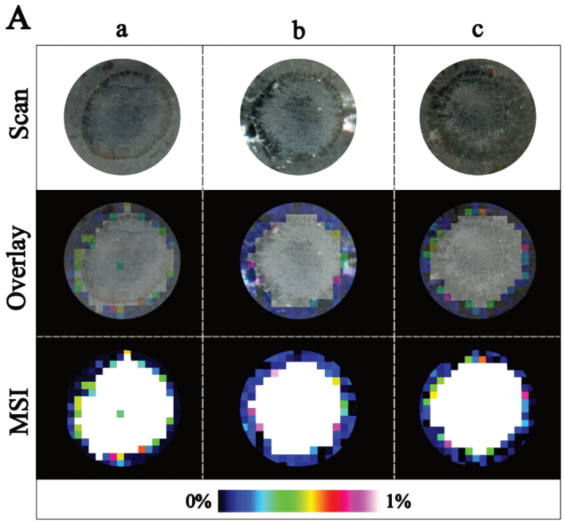
Figure 4. 12 Washing step with Zn^{2+}-doped water results

Spectra focusing on the ions $[M+Zn]^{2+}$ *m/z* 288.52 (A) $[2M+Zn]^{2+}$ *m/z* 545.07 (B) $[3M+Zn]^{2+}$ *m/z* 801.63 (C) $[4M+Zn]^{2+}$ *m/z* 1058.19 (D) obtained by MALDI-MS in full scan mode from washed (red) or unwashed (dark green) pure standard's drop or in CASI mode from washed (orange) or unwashed (light green) pure standard's drop. (E) Overlay of ITO slide scans with MSI images of analyte pure standard spotted on mouse liver tissue section, washed with the solution, sprayed with DHB matrix, and acquired in positive ion mode.

4.4.1.4 Optimization of spraying conditions to minimize delocalization

When acquiring images from treated mice whole-body sections, we observed an intense analyte signal surrounding the visceral tissue, which may have delocalized into nearby organs. Delocalization can occur at different stages of sample preparation, including mounting, washing, and matrix spraying. Thaw mounting minimizes delocalization due to minimal cytosolic liquid present in tissue sections. Since no washing steps were performed, the spraying step was identified as potential source of delocalization. We then optimized the spraying parameter to minimize delocalization and assessed the analyte's signal spread at high saturation around the mark of the pure standard drop spotted on control tissue (Figure 4. 13 A). Initially, we decreased the spray flow rate (ml/min) and improved the number of passes, to reduce the volume of solvent applied, which can penetrate the tissue and cause

the delocalization (Figure 4. 13 B-b). This modification led to a reduction in signal dispersion around the drop (Figure 4. 13 A-b), but also resulted in a higher LOD (7.83 µg/g), due to a reduced matrix deposition. Secondly, we increased the margin length (mm) to allow additional drying time to the solvent between passes (Figure 4. 13 B-c). With this method, we observed a reduction in signal dispersion around the drop (Figure 4. 13 A-c), while maintaining an LOD comparable to that achieved during previous optimization steps (1.44 µg/g). These optimized spraying parameters were maintained for the acquisition of the entire dataset.



B

	Temp (°C)	Passes (count)	Flow rate (ml/min)	Velocity (mm/min)	Pressure (psi)	Gas flow rate (l/min)	Overspray margins (mm)
a	75	12	0,09	1200	10	3	2 (r, t) 4 (l, b)
b	75	16	0,05	1200	10	3	2 (r, t) 4 (l, b)
c	75	12	0,09	1200	10	3	15 (r, l, t, b)

Figure 4. 13 Study of the analyte's delocalization following spraying step

(A) Overlay of scans of the drop marks on mouse liver tissue sections with MSI images of analyte pure standard signal, sprayed with the matrix using different methods (a, b, c), and acquired in positive ion mode. (B) Table reporting the parameters applied for each spraying method.

4.4.2 Whole-body quantification of the analyte in selected organs

4.4.2.1 Analytical method performance

To evaluate the performance and consistency of the analysis, we compared the calibration curves generated during the study. For each calibration point, we calculated the relative standard deviation (RSD%) for both the signal intensity and calculated concentration (µg/g) at each dilution of pure standard. The average RSD% for intensity and concentration were 18% and 8% respectively (Table 4. 1 A&B).

A

	Curve	WS15	WS14	WS13	WS12	WS11	WS10	WS9	WS8	WS7	WS6	WS5	WS4	WS3	WS2
Intensity	a	-	0,05	0,05	0,07	0,09	0,09	0,13	0,26	0,42	0,51	0,72	1,18	1,47	1,89
	b	-	0,04	0,05	0,06	-	0,10	0,15	0,31	0,46	0,54	-	1,28	-	2,19
	c	-	-	0,05	0,05	0,07	0,08	0,11	-	0,35	0,39	0,78	0,93	1,54	1,71
	d	-	-	0,05	0,06	0,08	0,08	0,10	-	0,36	0,39	0,69	1,02	1,49	1,93
	e	0,02	-	0,03	0,03	0,04	-	0,08	0,15	-	0,32	0,58	0,92	1,34	1,54
	Mean	0,018	0,047	0,046	0,055	0,071	0,088	0,114	0,239	0,397	0,431	0,690	1,064	1,461	1,850
	SD	-	0,006	0,009	0,013	0,022	0,008	0,027	0,083	0,052	0,094	0,081	0,160	0,085	0,244
	RSD%	-	13%	19%	24%	30%	9%	24%	35%	13%	22%	12%	15%	6%	13%

B

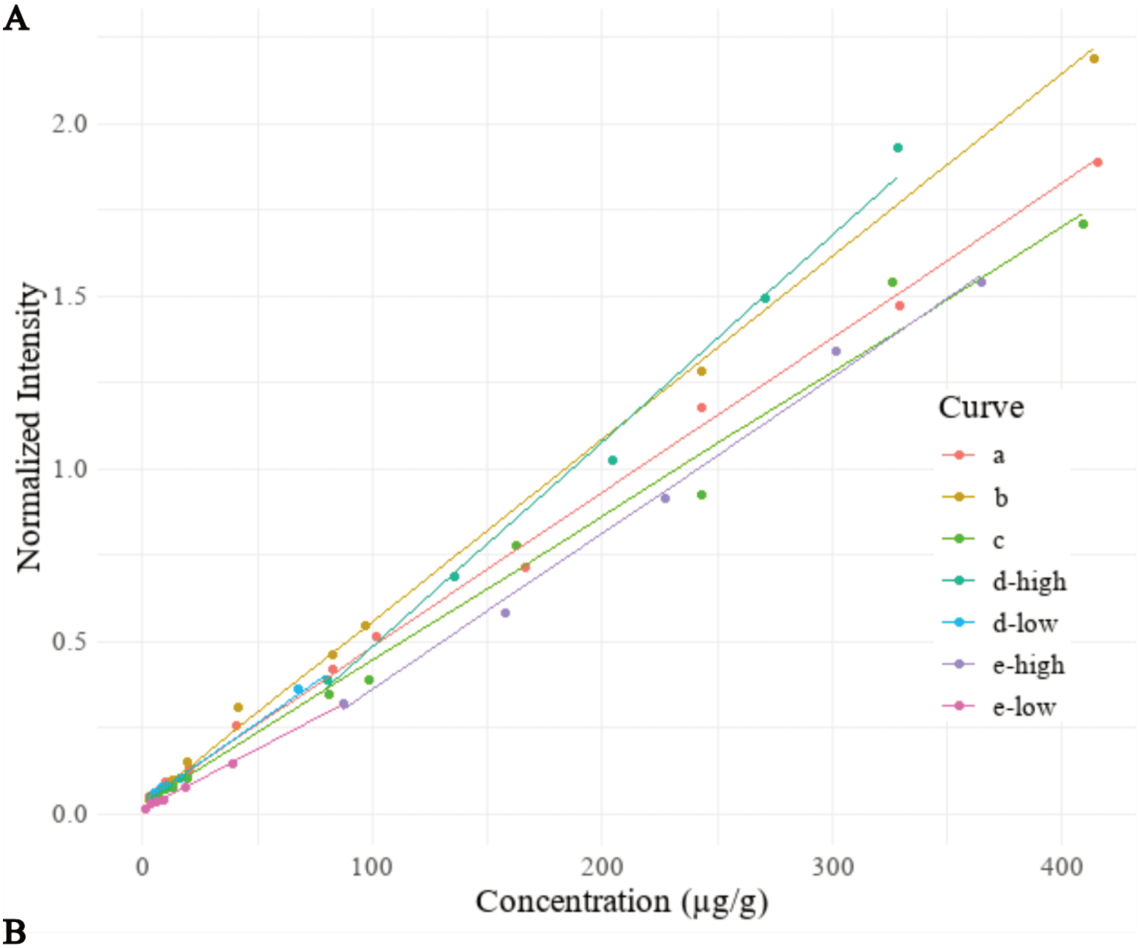
	Curve	WS15	WS14	WS13	WS12	WS11	WS10	WS9	WS8	WS7	WS6	WS5	WS4	WS3	WS2
Conc. (µg/g)	a	-	2,68	4,05	6,52	9,93	13,10	19,86	41,15	82,31	101,37	166,32	243,18	329,23	415,80
	b	-	2,61	4,53	6,58	-	12,97	19,36	41,58	82,31	96,78	-	243,20	-	413,70
	c	-	-	3,97	6,45	9,88	12,97	19,75	-	81,06	98,77	162,12	243,18	325,89	409,44
	d	-	-	3,34	5,39	8,03	10,70	16,25	-	67,44	80,25	135,44	204,03	270,89	328,89
	e	1,49	-	3,73	5,86	9,04	-	18,43	39,15	-	87,55	158,13	227,12	301,41	364,77
	Mean	1,490	2,645	3,924	6,160	9,220	12,435	18,730	40,627	78,280	92,944	155,503	232,142	306,855	386,520
	SD	-	0,049	0,437	0,518	0,892	1,158	1,496	1,297	7,251	8,802	13,787	17,186	26,994	38,457
	RSD%	-	2%	11%	8%	10%	9%	8%	3%	9%	9%	9%	7%	9%	10%

Table 4. 1 Titration curve performance

Reporting values of intensity (A) and calculated concentration (B) for each point of the titration curve for which analyte's signal was detected. The variability among the five titration curves was evaluated by calculating the RSD as a ratio of the standard deviation to the mean.

We examined the distribution of slopes (Figure 4. 14 A), as well as the LODs, LLOQs, and ULOQs to assess the reproducibility of the method and the consistency of the analytical performance throughout the dataset. All the limits were expressed in µg/g of tissue and µM units. The five calibration curves were all weighted linear 1/X. For two curves, we generated separate calibrations for low and high intensity signals. The average LOD and LLOQ were 1.75 µg/g and 2.82 µg/g of tissue, corresponding to 3.4 µM and 5.59 µM in tissue, respectively. The coefficient of determination, R^2, was greater than 0.99 for each curve (Figure 4. 14 B).

A



B

Curve	Equation	R^2	LOD (µg/g)	LOD (µM)	LLOQ (µg/g)	LLOQ (µM)	ULOQ (µg/g)	ULOQ (µM)
a	y = 0,0386 + 0,0045x	0,9961	1,65	3,22	2,68	5,21	415,80	809,59
b	y = 0,0309 + 0,0053x	0,9949	1,64	3,19	2,61	5,08	413,66	805,41
c	y = 0,0284 + 0,0042x	0,9904	2,65	5,15	3,97	7,74	409,44	797,19
d-low	y = 0,0351 + 0,0046x	0,9942	1,31	2,55	3,34	6,51	80,25	156,25
d-high	y = -0,1082 + 0,0059x	0,9927	1,31	2,55	80,25	156,25	328,89	640,37
e-low	y = 0,0138 + 0,0035x	0,9949	1,49	2,91	1,49	2,91	87,55	170,45
e-high	y = -0,0948 + 0,0045x	0,9938	1,49	2,91	87,55	170,45	364,77	710,23

Figure 4. 14 Titration curves for analyte's quantification *in vivo* result and performance

(A) Linear weighted 1/X curves created for each day of sample analysis plotted in a graph intensity over concentration (µg/g). (B) Table reporting equation, R^2, and limits expressed in µg/g and µM for each of the five acquired titration curves.

4.4.2.2 Analyte's biodistribution and quantification

Visceral adipose tissue and fat intestines, which make the sectioning of the whole-body more difficult, were observed in the vehicle animal carcass. Sections were carefully collected by gathering the pieces of tissue when separation occurred. The following organs and tissues were delineated, when present on the whole-body section: intestine, kidney, liver, lung, spleen, pancreas, brain, eye, ankle, white and brown adipose tissue. We observed retention and high analyte signal in the white adipose tissue. Careful ROI drawing was performed to exclude potential contamination from surrounding tissue. If

contamination was observed around a tissue, the non-contaminated area was delineated for qMSI analysis.

The analyte was detected at *m/z* 514.1211 in all treated whole-body sections. The resolving power of the method developed was 65k at *m/z* 514.1 and over 100k at *m/z* 400. The molecular distribution of the analyte in treated animals is presented in figures 4.15 and 4.16. The molecular images of the analyte obtained from vehicle animal is presented in figures 4.17. Normalization of raw molecular signals by the IS allowed correcting the data per position (pixel) revealing the true distributions of the analyte.

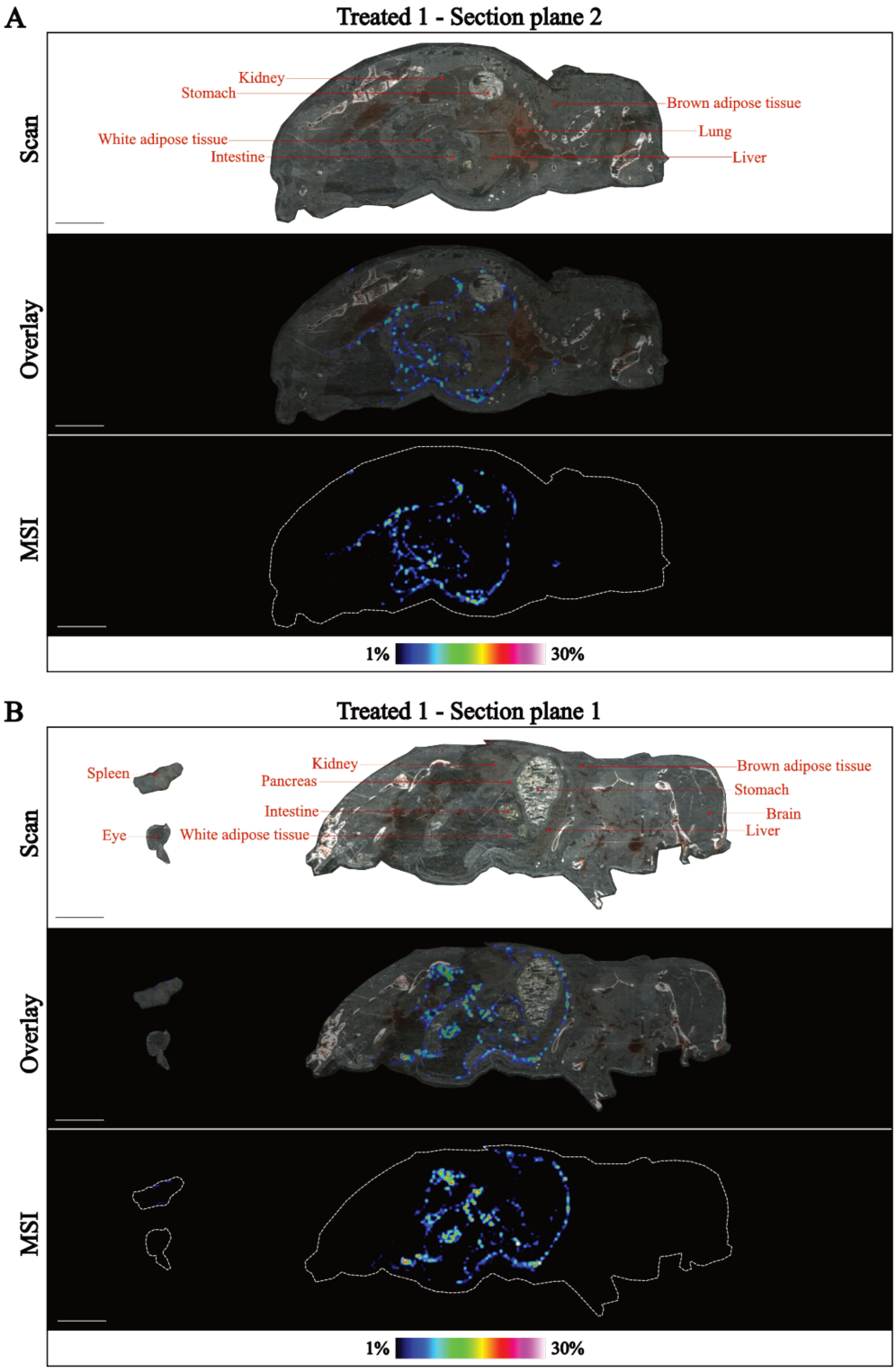


Figure 4. 15 Normalized signal of the analyte in the first treated animal

Overlay of the whole-body section scans exposing organs of interest with normalized MSI molecular images acquired from section plane 1 (A) and 2 (B).

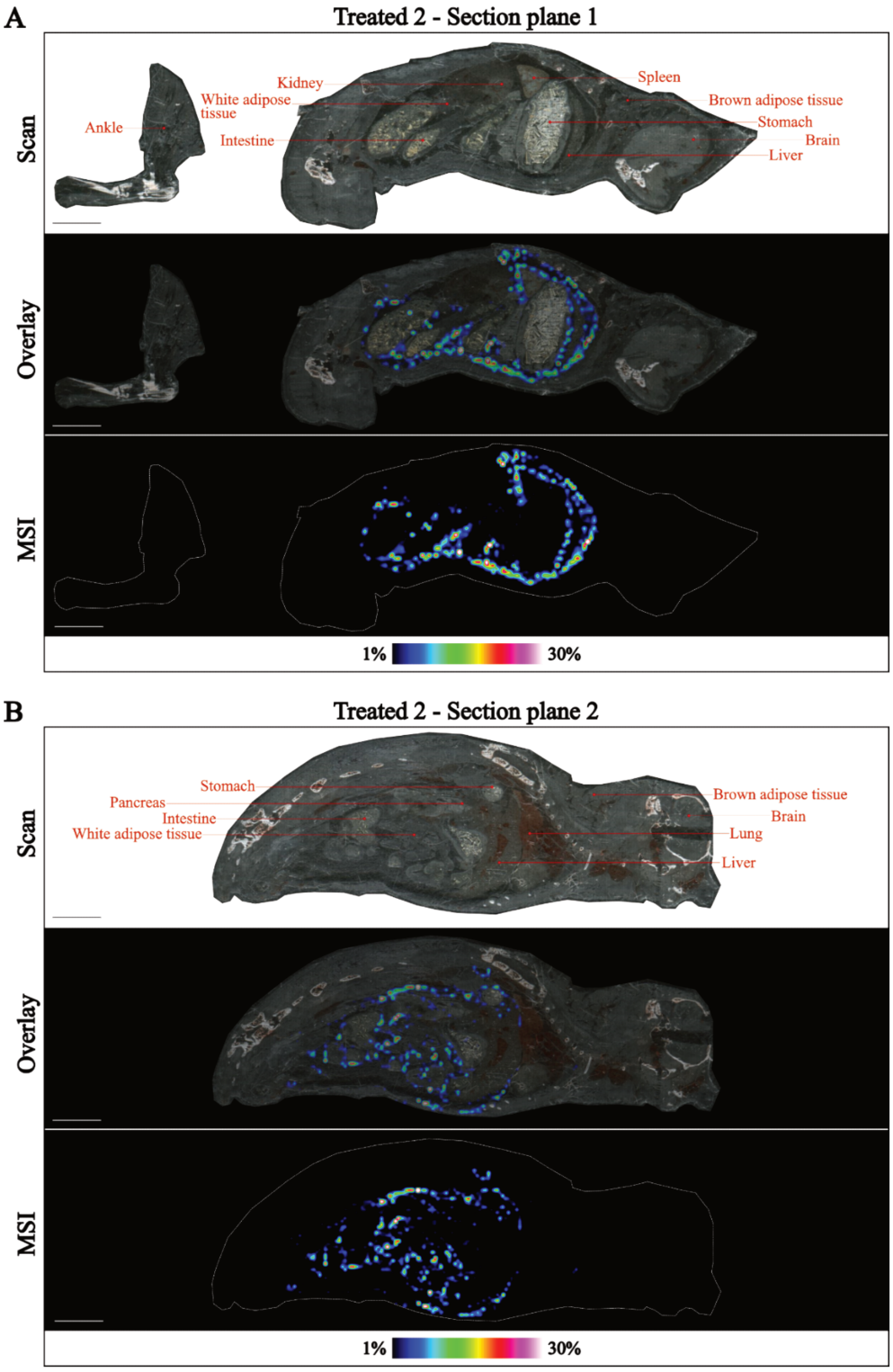


Figure 4. 16 Normalized signal of the analyte in the second treated animal

Overlay of the whole-body section scans exposing organs of interest with normalized MSI molecular images acquired from section plane 1 (A) and 2 (B).

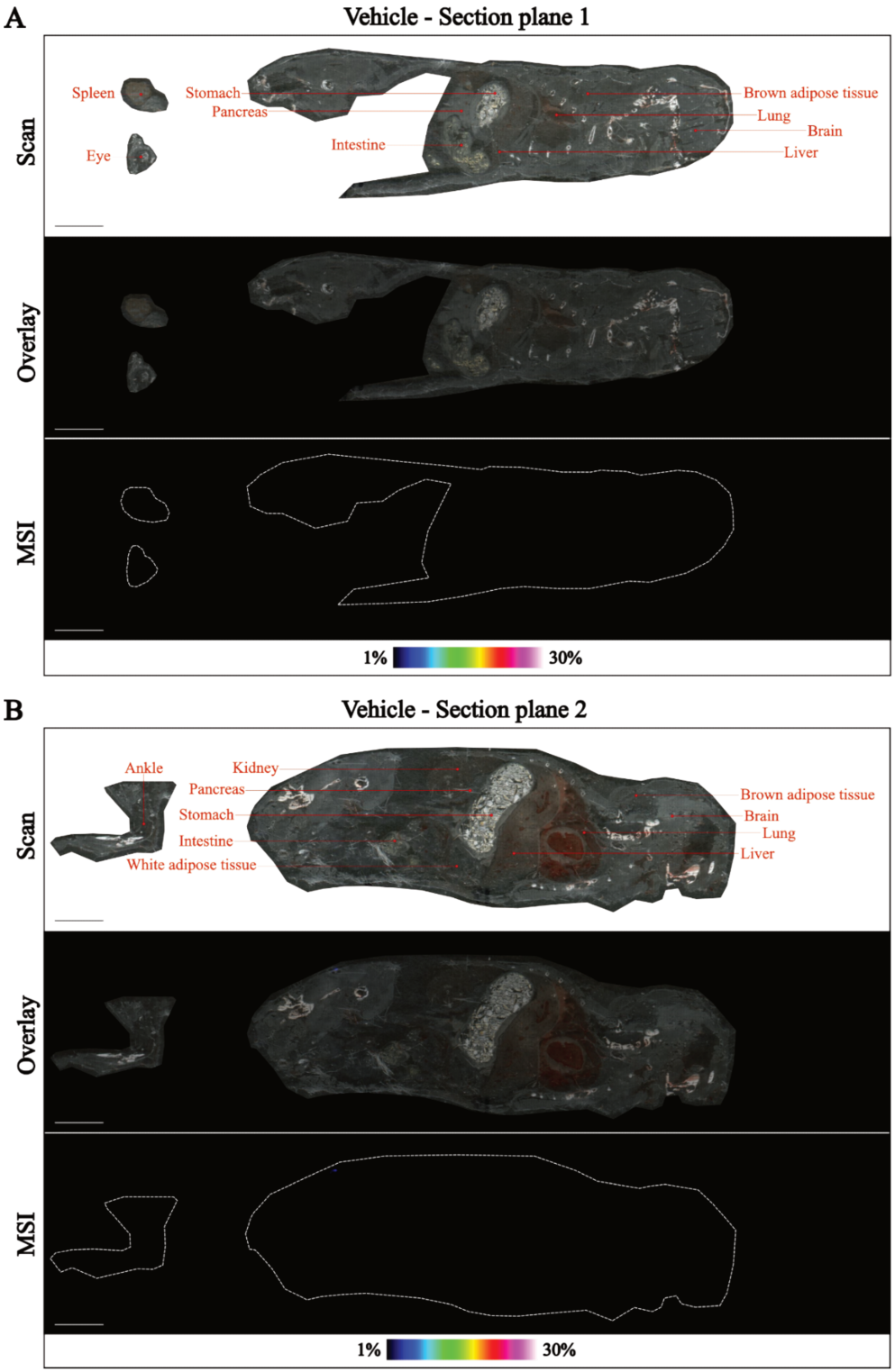


Figure 4. 17 Normalized signal of the analyte in the Vehicle animal

Overlay of the whole-body section scans exposing organs of interest with normalized MSI molecular images acquired from section plane 1 (A) and 2 (B).

The highest concentration of the analyte was measured in white adipose tissue, with an average concentration of 332.69 µg/g of tissue or 647.77 µM. Organs presenting higher analyte signal were,

in decreasing order, the kidney (87.63 µg/g, 170.61 µM), the pancreas (59.59 µg/g, 115.96 µM), the liver (48.51 µg/g, 94.46 µM), the intestine (44.04 µg/g, 85.74 µM), and the spleen (38.21 µg/g, 74.40 µM). The analyte also distributed in the lung (17.61 µg/g, 34.28 µM), stomach (13.52 µg/g, 23.66 µM), and brown adipose tissue (4.46 µg/g, 8.67 µM). In investigated compartments such as the brain, the eye and the ankle joint, the analyte was not detected (Table 4. 2). This either means that the organ was not exposed to the compound, or it was exposed at levels below the limit of detection of the analytical method. Two (2) biological replicates were included in the study and inter sample variability was higher in white (49%) and brown (63%) adipose tissue and for organs such as the kidney (42%) and stomach (52%) than the other tissues. This variability covers both biological and analytical variabilities, the latter estimated at 33.5% and 38.1% on DHB matrix and on QC signals, respectively.

A

Animal ID	Section plane	LOD	LLOQ	ULOQ	Concentration in organs (µg/g)										
					Intestine	Kidney	Liver	Lung	Spleen	Pancreas	Stomach	White adipose tissue	Brown adipose tissue	Brain	Eye	Ankle
Treated 1	**1**	2,65	3,97	409,44	41,60	82,43	54,18	NP	28,07	63,23	7,63	502,5 (*)	ND	ND	ND	NP
	2	2,65	3,97	409,44	36,18	126,32	38,82	15,15	NP	NP	7,38	159,43	6,44	NP	NP	NP
Treated 2	**1**	1,31	3,34	328,89	42,35	54,13	68,94	NP	48,36	64,05	20,17	436,15	ND	ND	NP	ND
	2	1,49	1,49	634,77	56,01	NP	32,11	20,07	NP	51,39	18,90	232,69	2,47	ND	NP	NP

B

Animal ID	Section plane	LOD	LLOQ	ULOQ	Concentration in organs (µM)										
					Intestine	Kidney	Liver	Lung	Spleen	Pancreas	Stomach	White adipose tissue	Brown adipose tissue	Brain	Eye	Ankle
Treated 1	**1**	5,15	7,74	797,19	81,00	160,50	105,48	NP	54,65	123,12	14,86	978,39	ND	ND	ND	NP
	2	5,15	7,74	797,19	70,45	245,95	75,58	29,49	NP	NP	14,37	310,41	12,53	NP	NP	NP
Treated 2	**1**	2,55	6,51	640,37	82,46	105,39	134,24	NP	94,15	124,70	39,26	849,20	ND	ND	NP	ND
	2	2,91	2,91	1235,93	109,05	NP	62,53	39,07	NP	100,05	36,81	453,06	4,82	ND	NP	NP

Table 4. 2 Analyte's concentrations in each histological region of interest of the whole-body

Concentration, per sections planes analysed and per animal, obtained by comparing the analyte intensity in each ROI to the titration curve, reported in μg/g of tissue (A) and μM (B). (*) Value extrapolated because above the ULOQ.

4.4.3 Identification and biodistribution of metabolites

4.4.3.1 Phase I metabolism prediction and investigation in vitro and in vivo

Structural analysis and predictive bioinformatic tools (BioTransformer 3.0) were used to predict biotransformation sites on the analyte and identify potential Phase I metabolites (Table 4. 3). Hydroxylation could occur at multiple positions, including the two phenyl groups or the methine and methylene bridges (Figure 4. 18 A b-d). The generated secondary alcohols could be further oxidized into the corresponding carbonyls (Figure 4. 18 A-e). Other transformations include O-dealkylation of the methoxyphenyl group (Figure 4. 18 A-h) and S-oxidation of the thiophene ring (Figure 4. 18 A-j). The inherent instability of the hydroxamic acid moiety can lead to the corresponding carboxamide and carboxylic acid. These structures were also listed and investigated (Figure 4. 18 A-m, n). We extended our analysis by adding biotransformation reactions up to three iterations (Figure 4. 18 A-f, g, i, k, l) and investigated biotransformation occurring on the degradation products (Figure 4. 18 A-o, r). To experimentally validate these predictions, the list of candidate metabolites was investigated *in vitro* and *in vivo*.

Modification		**Formula**	***m/z***	***in vitro***	***in vivo***
Parent compound	**a**	C23H23N5O5S2	514.1213	yes	yes
Hydroxylation	**b**	C23H23N5O6S2	530.1162	yes (*)	no
Hydroxylation/Hydroxilation	**c**	C23H23N5O7S2	546.1111	no	no
Hydroxylation/Hydroxilation/Hydroxilation	**d**	C23H23N5O8S2	562.1061	no	no
Hydroxylation/Oxidation of secondary alcohol	**e**	C23H21N5O6S2	528.1006	no	no
Hydroxylation/Oxidation of secondary alcohol/Hydroxylation	**f**	C23H21N5O7S2	544.0955	no	no
Hydroxylation/Oxidation of secondary alcohol/Hydroxylation/O-Dealkylation	**g**	C22H19N5O7S2	530.0798	no	no
O-Dealkylation	**h**	C22H21N5O5S2	500.1056	yes (*)	no
O-Dealkylation/Hydroxylation	**i**	C22H21N5O6S2	516.1006	no	no
S-Oxidation	**j**	C23H23N5O6S2	530.1162	yes (*)	no
S-Oxidation/Hydroxylation	**k**	C23H23N5O7S2	546.1111	no	no
S-Oxidation/Hydroxylation/O-Dealkylation	**l**	C22H21N5O7S2	532.0955	no	no
Hydroxamic acid > Carboxamide	**m**	C23H23N5O4S2	498.1264	yes (^)	yes
Hydroxamic acid > Carboxamide > Carboxylic acid	**n**	C23H22N4O5S2	499.1104	yes (^)	yes
Hydroxamic acid > Carboxamide /Hydroxylation	**o**	C23H23N5O5S2	514.1213	-	-
Hydroxamic acid > Carboxamide /O-Dealkylation		C22H21N5O4S2	484.1107	yes	no
Hydroxamic acid > Carboxamide /Hydroxylation/O-Dealkylation	**p**	C22H21N5O5S2	500.1056	yes (*)	no
Hydroxamic acid > Carboxamide > Carboxylic acid /Hydroxylation	**q**	C23H22N4O6S2	515.1053	no	no
Hydroxamic acid > Carboxamide > Carboxylic acid /O-Dealkylation		C22H20N4O5S2	485.0947	yes	no
Hydroxamic acid > Carboxamide > Carboxylic acid /Hydroxylation/O-Dealkylation	**r**	C22H20N4O6S2	501.0897	no	no

Table 4. 3 List of Phase I predicted metabolites

For the investigated molecules we listed the letter relative to the structure in Figure 4. 18, the formula, the theoretical *m/z*, and whether the signal was detected *in vitro* or *in vivo*. Asterisks (*) indicate different modification leading to the same nominal *m/z*. Carets (^) indicate ions detected in the *in vivo* experimental samples, that where however detected in control samples too.

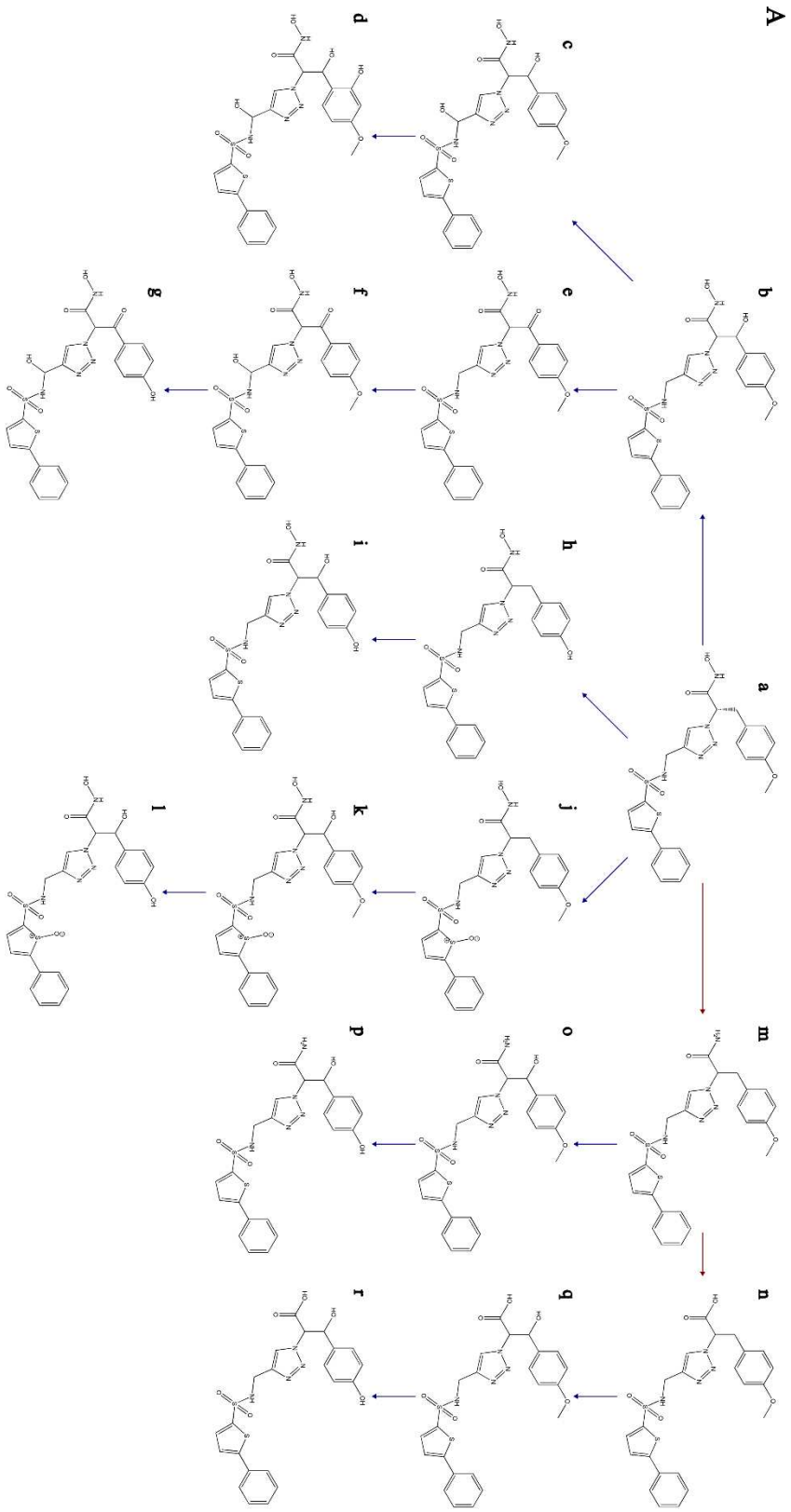
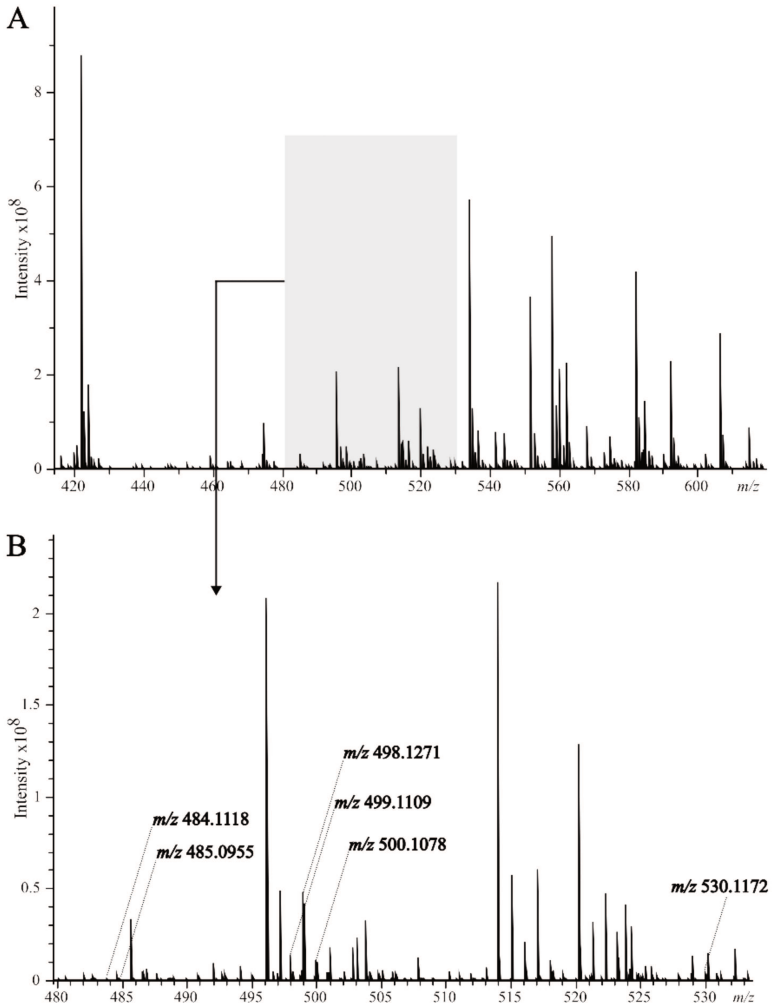


Figure 4. 18 Predicted metabolic pathways

(A) Structure of the metabolites predicted to result by the biotransformation (blue arrows) or degradation (red arrows) of the analyte (a).

In vitro, the analyte was incubated with female mouse microsomes to investigate the phase I metabolism. Experimental samples were compared to time point 0 and to enzymatic and matrix controls to identify specific peaks. Among the biotransformation products, we detected the presence of peaks at *m/z* 500.1078, corresponding to the O-demethylation, and *m/z* 530.1172, corresponding to either hydroxylation or thiophene S-oxidation. Both the carboxamide (*m/z* 498.1271) and carboxylic acid (*m/z* 499.1109) signals were detected, though they were also present in the controls, as this modification is independent of the CYP activity. Additionally, we detected the O-demethylation product of both the carboxamide (*m/z* 482.1118) and carboxylic acid (*m/z* 485.0955), and potentially the O-demethylation followed by hydroxylation of the carboxamide (*m/z* 500.1172), though this mass corresponds to the one of O-demethylation product of the parent compound (Figure 4. 19 B). The shift between theoretical and experimental *m/z*, expressed in parts per million (Δppm), was 2.10 ppm on average and below 5 ppm globally (Figure 4. 19 C).



C

Modification	Formula	Theoretical *m/z*	Experimental *m/z*	Δppm
Hydroxylation *or* S-Oxidation	C23H23N5O6S2	530.1162	530.1172	1.89
O-Dealkylation *or* Hydroxamic acid > Carboxamide /Hydroxylation/O-Dealkylation	C22H21N5O5S2	500.1056	500.1078	4.40
Hydroxamic acid > Carboxamide	C23H23N5O4S2	498.1264	498.1271	1.41
Hydroxamic acid > Carboxamide > Carboxylic acid	C23H22N4O5S2	499.1104	499.1109	1.00
Hydroxamic acid > Carboxamide /O-Dealkylation	C22H21N5O4S2	484.1107	484.1118	2.27
Hydroxamic acid > Carboxamide > Carboxylic acid /O-Dealkylation	C22H20N4O5S2	485.0947	485.0955	1.65

Figure 4. 19 Phase I metabolites detected *in vitro* on microsomal stability study samples
(A) Representative spectrum acquired from experimental sample collected at t=60 min during the microsomal stability study acquired using MALDI-MS with a 200 Da wide CASI window. (B) Enlarged view of the region of the spectrum where the metabolites were detected, each labelled with the experimental mass detected. (C) Table reporting metabolites modifications, formula, theoretical *m/z*, *m/z* detected experimentally and the shift (Δppm) between the two values.

For the *in vivo* analysis, we acquired an MSI image from a treated whole-body section in CASI mode selecting a wide window (200 Da) centred on the analyte mass. Signals at *m/z* 498.1264 and *m/z* 499.1104 were detected, allowing to investigate the biodistribution of the carboxamide and carboxylic acid derivate from the hydroxamic acid moiety, which were compared to that of the analyte. Notably, they showed similar biodistribution to that of the parent compound, but in addition they were highly present in the liver and in the content of a region that was annotated as the small intestine of the treated animals (Figure 4. 20, Annex 3). No other metabolites that were present *in vitro* were detected *in vivo*.

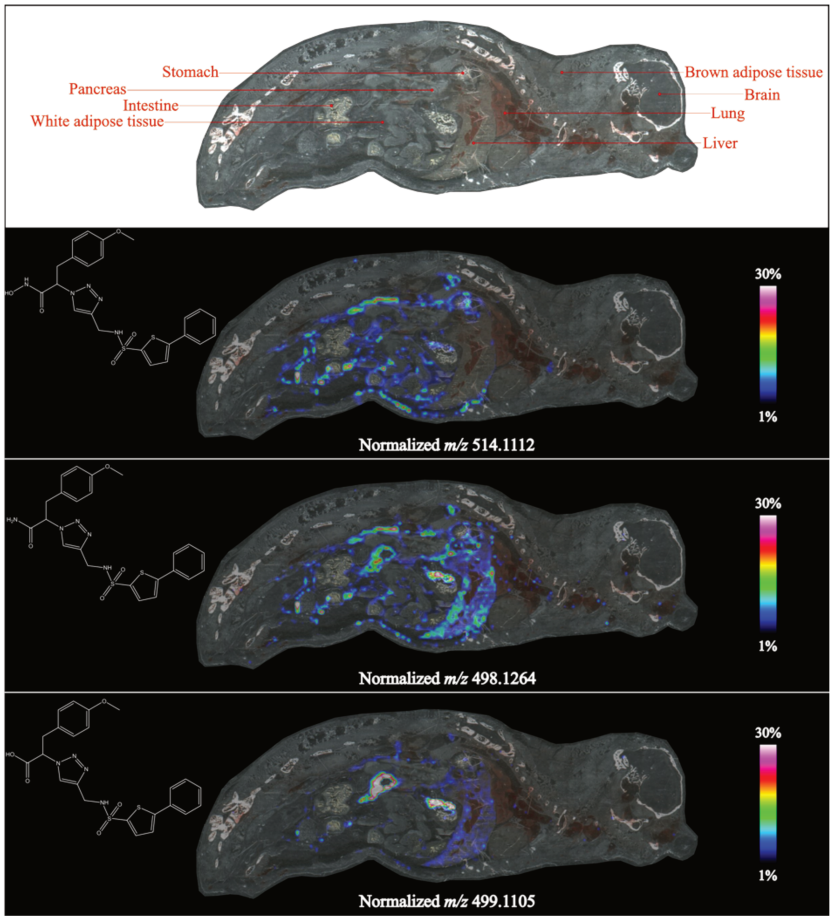


Figure 4. 20 Metabolism investigation *in vivo*
Normalized signals of the analyte and its degradation products detected in a treated animal. Overlay of the whole-body section scans with MSI molecular images acquired.

4.4.3.2 *Phase II metabolism prediction and investigation in vivo*

The bioinformatic tool BioTransformer 3.0 was used to predict potential Phase II metabolites. The software predicted the N-glucuronidation of the sulphonamide (Figure 4. 21 A). We acquired an MSI image in full scan mode and detected *in vivo* a peak with the characteristic mass shift of the glucuronic acid conjugate (+176 Da). Upon biodistribution investigation, the conjugate was found in the content of a region that was annotated as the small intestine of the treated animals (Figure 4. 21 B, Annex 3). Additionally, we investigated the presence of sulphate and glutathione conjugates, which however were not detected.

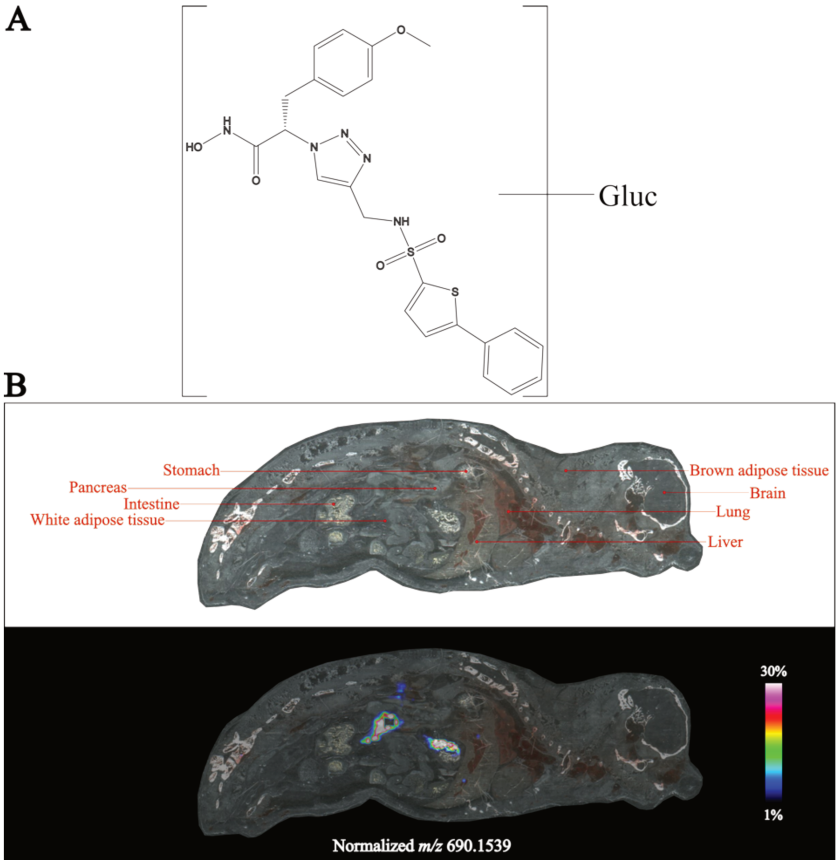


Figure 4. 21 Phase II metabolism prediction and investigation *in vivo*

(A) Analyte's glucuronidation reaction. (B) Normalized signal of the analyte's glucuronic acid conjugate detected in a treated animal. Overlay of the whole-body section scans with MSI molecular images acquired.

4.4.3.3 *Metabolism exploratory analysis*

To further investigate the metabolism *in vivo*, we compared spectra from treated animals to those from the vehicle control, in the mass range 464 and 564 Da, focusing on Phase I biotransformation. We detected three additional peaks at *m/z* 536.1028, *m/z* 538.0988, and *m/z* 552.0772, specific to the treated condition (Figure 4. 22 B-E). Their spatial biodistribution colocalized with that of the parent compound (Figure 4. 22 A). By cross-checking the presence of these peaks *in vitro*, only the signal at *m/z* 552.0772 was detected, but it was also present at t=0 min and in control samples.

A similar exploratory analysis was performed *in vitro* by comparing the spectra of the experimental samples at each time point with those from the t=0 min and control samples. Peaks unique to the experimental samples were identified as likely products of cytochrome-mediated biotransformation (Figure 4. 22 F). None of these peaks, present *in vitro*, were detected *in vivo*.

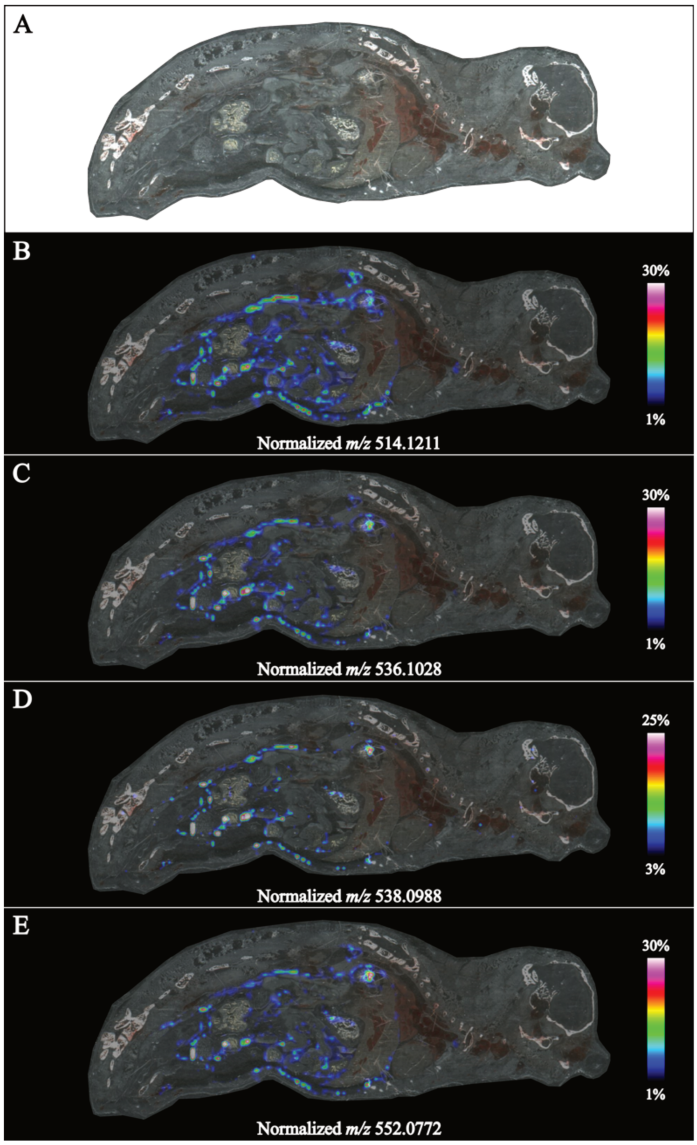


Figure 4. 22 *In vivo* and *in vitro* metabolism exploratory analysis results

(A) Normalized signal detected for the analyte and its potential biotransformation or degradation products *in vivo*. Overlay of the whole-body section scans with MSI molecular images acquired. (B) List of *m/z* detected *in vitro* potentially corresponding to biotransformation products of the analyte.

4.5 Biochemical interpretation and discussion

The candidate compound under analysis is one of the first described potent (half maximal inhibitory concentration IC_{50}=0.019 µM) and selective (IC_{50}(ERAP1)>100 µM) ERAP2 inhibitors. Its characterization already demonstrated favourable *in vitro* activity and *in vivo* systemic

pharmacokinetic profile: following intraperitoneal administration at the tolerated dose of 50 mg/kg, the C_{max} was in the range of the concentration needed for target engagement (half maximal occupancy concentration OC_{50}=23 μM) (Camberlein et al., 2022).

In this chapter, we presented the results of an *in vivo* biodistribution and quantification analysis employing quantitative mass spectrometry imaging. We aimed to further investigate the metabolism and pharmacokinetics (DMPK) properties of the compound, essential for assessing its therapeutic potential. As discussed in the thesis introduction, valid compounds must demonstrate appropriate target exposure, target engagement, and pharmacological activity. The combination of pharmacokinetic (PK) and pharmacodynamic (PD) investigations allows to correlate the concentration of candidate compounds in a specific body compartment and their potential effects. In this study, we measured the concentration (exposure) of the compound in various organs at the time point identified as t_{max} for this therapeutic regimen and will hereafter discuss these findings in the context of the concentrations required for target engagement and inhibition.

Organs for which the ERAP2 inhibitory activity is desirable for therapeutic purposes, based on its involvement in various diseases, include the spine and sacroiliac joints (ankylosing spondylitis), the eyes (Behçet's disease, Birdshot chorioretinopathy, ankylosing spondylitis), the skin (psoriasis, Behçet's disease, ankylosing spondylitis), the lungs (Behçet's disease), the digestive system (ankylosing spondylitis, Behçet's disease, oral cavity squamous cell carcinoma), and the urogenital tract (Behçet's disease, bladder cancer) (El Maghraoui, 2011; Feller & Lemmer, 2012; Kirkali et al., 2005; Minos et al., 2016; Zeidan et al., 2016). The histological regions of interest on the whole-body sections, exposed to the compound at concentrations in the range of target engagement were the kidney (170.61 μM), the pancreas (115.96 μM), the liver (94.46 μM), the intestine (85.74 μM), the spleen (74.40 μM), the lung (34.28 μM), and the stomach (23.66 μM). This indicates promising therapeutic potential for the ERAP2 inhibitor in addressing pathologies affecting the digestive system and lungs. However, the compound's limited ability to reach peripheral organs (skin, eyes, joints), which are involved in various pathological conditions that would benefit from ERAP2 inhibition, limits its therapeutic potential, under the current therapeutic regimen. Alternative delivery methods such as intra-articular injections, topical creams, or ophthalmic drops could be explored to enhance the compound's bioavailability in these regions.

The high concentration reached by the compound in organs such as the kidneys, pancreas, and liver, which are involved in metabolization and elimination of unnecessary or toxic substances and in maintaining hormonal homeostasis, should be monitored for potential undesirable and toxic effects. Moreover, the compound showed a high affinity for adipose tissue, with white adipose tissue exhibiting the highest exposure post IP administration, with an average measured concentration of 647.77 μM. This finding suggests the potential need for further pharmacokinetic studies to determine

clearance from this tissue and possibly adjust doses accordingly to achieve the desired therapeutic effect while minimizing unwanted accumulations.

The absence of detectable signal in the brain of the animals suggests that the compound doesn't cross the blood-brain barrier (BBB) under this therapeutic regimen.

Subsequently, we conducted a combined *in vivo*/*in vitro* metabolism study of the ERAP2 inhibitor. The simultaneous evaluation of the candidate compound's pharmacokinetics, efficacy, and safety is crucial in a drug discovery program, as these properties must meet the TPP requirements to advance through subsequent phases. Assessing the presence of potential metabolites provides valuable insights into the drug's fate post administration and its DMPK properties. Elevate distribution in specific regions may anticipate potential secondary target engagement and toxicity, which would need further investigation and monitoring.

We used structural analysis and predictive bioinformatic tools to identify potential biotransformation products of the candidate ERAP2 inhibitor compound, specifically focusing on phase I metabolism mediated by cytochrome P450. Moreover, given the known instability of the hydroxamic acid moiety, we specifically investigated the corresponding carboxamide and carboxylic acid derivates.

In vitro, we identified various biotransformation products, including hydroxylation or S-oxidation, and O-demethylation. However, these products were not detected *in vivo*, possibly due to the sacrifice time point (20 min) selected for the biodistribution study. While this time point was appropriate to study the compound's maximal concentration and evaluate the target exposure, it may have been too early to detect phase I metabolism products. Alternatively, we hypothesized a rapid glucuronic acid conjugation of the hydroxylated and O-demethylated metabolites *in vivo*, but the corresponding ions were not detected in the FS spectra.

The degradation products of the parent compound, carboxamide and carboxylic acid, were detected both *in vitro* and *in vivo*, since their generation is independent of enzymatic activity. *In vivo*, their signals were particularly intense in the liver and the small intestine of treated animals, suggesting elimination via bile. Biliary excretion involves the active secretion of drug and metabolites from hepatocytes to the bile and could contribute to the observed distribution of the compound's byproducts in these tissues. Further in vivo studies could help to clarify the extent of biliary clearance of these metabolites.

Following these initial findings on metabolic activity *in vivo*, we conducted a broader exploratory investigation. Specific peaks were detected in treated animals, potentially corresponding to additional metabolites. The presence of one peak only, at *m/z* 552.0772, was confirmed *in vitro*, however it was also detected in control samples. This suggests that the others may result from reactions independent of cytochrome P450 enzymes, potentially implicating alternative metabolic pathways.

Overall, these analyses were crucial to determine the compound's metabolism and pharmacokinetic (DMPK) profile, supporting its suitability for therapeutic application and further development.

4.6 Technical interpretation, MS potentials and limitations

Performing assays employing MALDI-mass spectrometry imaging is challenging due to the low ion signal reproducibility inter and intra analysis. Whole-body images are intrinsically more susceptible to analytical variability since they are analysed on different days. The technical variability measured in this study for matrix (DHB) and QCs (olanzapine) signals, that was 33.5% and 38.1% respectively, was above LC-MS/MS standards, relying on FDA/EMA guidelines for bioanalytical validation, but acceptable by MALDI qMSI standards.

Additional source of variability are tissue heterogeneity and matrix application, that lead to variation in analyte extraction and ionization. The different compartments and organs that compose a whole-body cause tissue-specific effects on the analyte. In whole-body MSI, for reasons linked with development and acquisition time management, they are analysed with the same analytical method, while they should be addressed and developed separately. The MALDI matrix application systems have evolved to combine good reproducibility and image resolution, but matrix crystal size and evenly coating can vary significantly among runs. To address these numerous sources of variability, in this study we employed the IS approach, using a chemical analogue of the analyte under investigation, its isotope-labelled form, homogeneously applied on the tissue section, and acquired simultaneously with the analyte (Pirman, Kiss, et al., 2013; Pirman & Yost, 2011). The signal of the analyte was then normalized by the signal of the IS, to compensate for variability in ionization efficiency caused by differences in tissue composition and preparation methods.

A potential limitation of the analytical method developed is the use of one organ, mouse liver, as a surrogate matrix for the whole-body, to build on-tissue the titration curve used for compound's quantification (Tobias & Hummon, 2020). As already discussed, there is tissue-specific variability in the ionization efficiency of analytes, based on the organ in which they are located. However, the impracticality of building titration curves using a surrogate matrix for each quantified organ, together with the cost associated with obtaining authentic blank tissue for each quantified organ, necessitates the selection of a surrogate matrix that is suitable, available, and affordable (Ho & Gao, 2015).

Taken together, the internal standard approach combined with the use of mouse liver as single blank matrix is an appropriate choice for a fit-for-purpose MALDI qMSI method application to investigate drug distribution at an early stage of drug discovery.

The study of compound metabolism *in vivo*, conducted via MALDI-FTICR-MSI, allowed the simultaneous localization of the compound and its metabolites, supporting the DMPK profile investigation and anticipating potential safety concerns that shall require attention when designing

toxicological studies at later stages of drug development. The added value of whole-body MSI from a Tox perspective is to inform decision making for future studies.

Finally, one of the major benefits of working with an unlabeled and spatial technique such as MSI, is the ability to investigate the biodistribution of an administered compound, while simultaneously collecting signals from hundreds of endogenous molecules. This technique allows to gain information about the compound's journey in the organism and evaluate the response of the organism to the presence of the compound, enabling PK/PD evaluation in a single assay while maintaining the spatial information.

4.7 Future works

Even if the high mass resolution power of Fourier transform ion cyclotron resonance (FTICR) MS allows to resolve analyte's signal from interfering species, the sensitivity of the mass method can be further improved through on-tissue chemical derivatization (OTCD). Several agents have been successfully applied for OTCD of drugs, aiming to enhance their detection in MSI studies. Some non-exhaustive example are the Girard's reagent T (GirT), which reacts with a carbonyl group, N,N,N-trimethylamonioanilyl N-hydroxysuccinimidyl carbamate iodide (TAHS), which reacts with an amine group, *N*-(2-(bromomethyl)benzyl)-*N*,*N*-diethylethanaminium bromide or cationic xylyl-bromide (CAX-B), which reacts with an active hydrogen, and 2-fluoro-1-methylpyridinium (FMP)-based reagents, which reacts with primary and secondary ammine and phenolic hydroxyl groups. Based on the analyte's chemical structure, GirT and FMP-based reagents could be assayed for OTCD. The introduction of an additional step during sample preparation must be carefully monitored to avoid problems associated with OTCD such as sample delocalization or tissue degradation (Harkin et al., 2022; Merdas et al., 2021).

Based on the results obtained from this study, it may be worthwhile to compare the current findings with analogues of the investigated drug candidate or alternative administration routes and evaluate the benefits in terms of distribution in peripheral organs, affinity for adipose tissue, and metabolism. While studying the *in vivo/in vitro* metabolism, we realized that the time point selected to study the biodistribution may have been too early to identify products of phase I metabolism. A dedicated experiment with longer time points and multiple sampling intervals can be conducted to further investigate *in vivo* metabolism and facilitate a more comprehensive characterization of the DMPK profile and the ADME processes of the compound. In this context, additional Phase 2 metabolites characterization shall be performed first *in vitro*, using microsomal supplementation with Phase 2 cofactors or direct hepatocyte incubation, and then *in vivo*.

The peaks detected *in vitro* and *vivo* during the exploratory analysis will be investigated to identify the specific structural modifications associated with biotransformation. For this purpose, HRMS will

be used to generate fragmentation patterns of the identified peaks and structurally characterize the detected metabolites.

5 General conclusion and future perspectives

The primary objective of this thesis was to support a drug discovery program by developing and applying powerful analytical methods, specifically focusing on mass spectrometry techniques, namely MSI and LC-MS/MS. The drug targets of this program are the endoplasmic reticulum aminopeptidases, ERAP1 and ERAP2. Their inhibition is considered an innovative therapeutic strategy in the context of autoimmune and oncological diseases. This research aims to demonstrate the importance of a multidisciplinary approach in tackling the complex challenges of drug discovery. The integration of analytical chemistry, along with various other scientific disciplines, has led to significant advancements in the research and development of new drugs, with mass spectrometry-based techniques being widely applied in this context. LC-MS/MS has been extensively used in analytical biochemistry due to its ability to separate, identify, and quantify molecules in complex matrices. Its high sensitivity and specificity have made it applicable in research, diagnostic, and clinical studies. It has a broad range of applications, allowing the investigation of metabolites, peptides, proteins, and lipids, as well as drugs. MSI is a technique that has already found considerable application in investigating endogenous and exogenous molecules at the tissue level in biomedical research. Numerous examples report its application in the context of biomarker discovery and the study of drug biodistribution and pharmacokinetics, as well as the simultaneous investigation of these two entities to further study the pharmacodynamic profile and efficacy of drugs. The three research projects presented in this thesis have allowed to explore of various aspects of the drug discovery pipeline: from biomarker discovery for validating preclinical models and characterizing drug candidates to their pharmacodynamic and pharmacokinetic evaluation.

In Chapter 2, we explored the potential use of the HLA-B27 transgenic rat preclinical model, commonly utilized to study ankylosing spondylitis (AS) and the associated chronic intestinal inflammation (IBD), to investigate the efficacy of drugs targeting ERAPs through the modulation of specific biomarkers. AS is an MHC-I-associated autoimmune disorder. Both ERAP1 and ERAP2 enzymes are risk factors for the disease, and their inhibition could represent an innovative therapeutic approach. Through untargeted metabolomic analysis performed using MALDI-MSI, we identified numerous features discriminating samples from the preclinical model from those derived from wild-type animals. The analysis of the fragmentation pattern, by LC-MS/MS, enabled us to identify some of these features as metabolites and lipids. We interpreted the role of the identified candidate biomarkers, already known for their involvement in the disease, based on spatial information and correlation of biochemical and morphological data. This revealed that two main mechanisms were deregulated in the investigated disease model, intestinal permeability and immune infiltration. This research opens new avenues for studying the pathogenesis of the disease and validating the efficacy

of therapeutic compounds that may alleviate the burden of this condition, which still requires alternative therapeutic strategies for those patients not responding to current lines of treatment. Evaluating the levels of the candidate biomarkers following treatment with ERAP1 and ERAP2 inhibitors will allow us to assess whether these drugs can target the pathological mechanisms, resolving the acute inflammatory state affecting the intestinal compartment of patients with AS. In this context, our study validated the HLA-B27 transgenic rat as a preclinical model for evaluating the efficacy of the inhibitors. The main contribution lies in MSI's ability to provide spatial dimensions to biochemical analysis, offering a more comprehensive view of pathogenic processes compared to traditional "bulk" analysis approaches. A significant advancement of this project is the implementation of a batch integration algorithm. Data manipulation can mitigate the technical variability associated with MSI analysis and represents a step forward in improving the reliability and reproducibility of this technique, making it more applicable in preclinical contexts and, potentially, clinical settings.

In Chapter 3, we focused on developing an LC-MS/MS-based method for quantifying an antigen whose expression depends on ERAP1 activity, the peptide GSW11. This peptide plays a significant role in stimulating T-cell-mediated immune responses in a preclinical model of colon cancer and in conferring cross-protective immunity in various tumours. The goal was to use it as a biomarker to evaluate the efficacy of an ERAP1 inhibitor through a direct bioassay, rather than relying on the stimulation of the CCD2Z T-cell hybridoma, as done previously. Furthermore, employing a genetic approach to study the modulation of GSW11 induced by the compound in a preclinical model where the pharmacological target had been knocked out, would have enabled us to validate the mechanism of action of the candidate drug.

Although the method's sensitivity was within the nanomolar range, and optimization efforts were made to enhance the specificity of the analysis, it was not possible to detect the endogenous peptide from cellular extracts. This project highlighted several technical challenges associated with bioanalytical methods, such as the low recovery rates from extraction processes and ion suppression related to complex biological matrices. In this context, the study also evaluated the validity of the preclinical cellular model for application in peptidomics studies based on the LC-MS/MS technique. Despite these challenges, the work conducted has provided valuable insights for the future, suggesting that further optimizations in sample preparation and the use of different ionization techniques could improve the method's outcomes.

In Chapter 4, we delved into the analysis of the pharmacokinetic and metabolic profile (DMPK) of a candidate ERAP2 inhibitor using quantitative MSI to study the compound's distribution across various tissues. ERAP2 and ERAP1 are paralogues with independent activity and different substrate

preferences. Numerous inhibitors of ERAP1 are available, but only recently research efforts have been directed toward discovering selective inhibitors for ERAP2. It is important to further expand research activities on these inhibitors to enhance the therapeutic applications of these drugs, as the two enzymes can generate distinct antigen repertoires that may result in complementary therapeutic effects.

The results of the quantitative biodistribution analysis showed that the compartments most exposed to the compound were white adipose tissue, kidney, pancreas, liver, intestine, and spleen. These data were interpreted in light of previous evidence regarding the compounds target engagement (OC_{50}) and inhibition (IC_{50}) abilities and based on the desired therapeutic applications. Sufficient concentrations for target engagement were achieved in organs associated with conditions such as Behçet's disease and ankylosing spondylitis. The compound's limited ability to reach peripheral organs (eyes, skin, joints), suggesting low bioavailability, highlights the room for further optimizations of the formulation and administration strategies, to make it an effective therapeutic option for conditions like Birdshot chorioretinopathy, psoriasis, and bladder cancer. In addition to these findings, the distribution of the compound in tissues and organs which are not targets for therapeutic purposes shall be taken into consideration for further optimization of dosage and administration routes.

This project demonstrated a successful application of the MALDI-MSI technique to investigate compound validity and opened avenues for further studies to complete its characterization, as well as additional optimization strategies.

One of the main contributions of this thesis is demonstrating how the constant evaluation of the validity of preclinical models and compounds can support the drug discovery process. The preclinical data gathered in this research will aid the decision-making process necessary to advance candidate drugs through development, and subsequently into clinical phases, thereby reducing the risk of failures.

Despite the positive outcomes of this research, several limitations warrant thorough consideration. One significant challenge is the inherent variability associated with data collection in mass spectrometry imaging, which can mask true biological signals and significantly affect the reliability of the results. To address this, approaches to monitor signals during analysis were introduced (QCs), with thresholds to ensure only datasets successfully meeting the acceptance criteria are considered for biological interpretation. Moreover, maintaining consistent conditions throughout the experiments avoided introducing technical variability. Nevertheless, we observed inter-batch variability that could have impaired the comprehensive analysis of the experimental cohort. To address this, the support of

data analysis tools was essential, allowing us to combine and interpret diverse datasets effectively, enhancing the robustness of our findings.

Challenges that we were unable to address arose during sample preparation, especially in the analysis of ankle joints for the spatial metabolomic study, and in the extraction protocol from cells for quantitative peptidomic analysis. The complexity of biological matrices and the requirements to make samples compatible with mass spectrometry techniques, limits the applicability in one end, while leave room for unlimited optimization in the other. Therefore, optimizing sample preparation techniques remains a critical focus area to improve.

The versatility of mass spectrometry-based techniques allows their application in both untargeted studies (Chapter 2) and targeted investigations (Chapters 3 and 4), facilitating biomarker discovery (Chapter 2) as well as the study of pharmacodynamic (Chapter 3) and pharmacokinetic profiles (Chapter 4).

Future perspectives for this research certainly involve the integration of the methods developed in the various projects, aiming for a comprehensive pharmacokinetic and pharmacodynamic characterization of the compounds. Chapters 2 and 4, in particular, represent preliminary phases of the advanced potential application of mass spectrometry imaging. The unlabelled and multiplexing properties of this technique, combined with the spatial dimension, make it a unique tool for directly and simultaneously visualizing both the candidate drug and biomarkers of efficacy and toxicity within a specific region of the body, correlating the presence with observable effects.

Once an ERAP inhibitor with a safe therapeutic regimen and an adequate DMPK profile will be identified, its efficacy will be tested *in vivo* using an ERAP-associated preclinical model, measuring the modulation of specific biomarkers related to the pathology. The studies presented in Chapters 2 and 4 were conducted using the same instrument and compatible sample preparation steps, enabling the future combination of the two methods to quantify the presence of the ERAP inhibitor in various organs while simultaneously visualizing its effects on the signature of candidate biomarkers in the HLA-B27 transgenic rat model.

6 Materials and methods

6.1 Chemicals

Bovine serum albumin (BSA), phosphate buffered saline (PBS), 2,5-Dihydroxybenzoic acid (2,5-DHB), 1,5-Diaminonaphthalene (1,5-DAN), trifluoroacetic acid (TFA), dichloromethane (CH_2Cl_2), formalin solution neutral buffered 10%, isopropanol, formic acid, ammonium formate, ammonium bicarbonate, ammonium hydroxide, and haematoxylin and eosin used for staining were purchased from Sigma Aldrich (St. Louis, MO, USA). LC-MS-grade methanol (MeOH), LC-MS-grade water (H_2O) and LC-MS-grade acetonitrile (ACN) were purchased from Thermo Fisher Scientific (Waltham, MA, USA). Ethanol, xylene and Pertex® mounting medium were purchased from VWR (Fontenay-sous-Bois, France). The optimal cutting temperature (OCT) embedding matrix was purchased from CellPath (Newtown, UK).

6.1.1 Compounds and standards stock and working solutions preparation

5.21 mg of ERAP1 inhibitor (449.499 g/mol) were resuspended in 382.5 µl of DMSO to prepare a stock solution at 30 mM. A working treatment solution at 60 µM was prepare by diluting 185 µl of stock solution in 92.5 ml of complete RPMI. A working vehicle solution was prepared by diluting 185 µl of DMSO in 92.5 ml of complete RPMI. 5 ml of the working solutions were added to each flask containing 25 ml of complete RPMI, for a final treatment concentration of 10 µM, and DMSO concentration <10%.

9.88 mg ERAP2 inhibitor (513.5983 g/mol) and 6.02 mg of is 4-deuterated analogue (517.6 g/mol) were resuspended in 1924 and 1163 µl of DMSO respectively to prepare a stock solutions at 10 mM. To use it as internal standard, the 4-deutereted analogue was diluted in MeOH:H_2O 1:1 (v/v) to a final concentration of 1 µM. To build the titration curve, serial dilution of the ERAP2 inhibitor were prepared in MeOH:H_2O 1:1 (v/v) as per Table 6. 1.

1 mg of Native GSW11 (1203.3 g/mol) and 1 mg of isotope labelled IS (1209.3 g/mol) peptides were resuspended in 831 and 826.9 µl of DMSO respectively to prepare a stock solutions at 0.95 mM. To use it as internal standard, the isotope labelled peptide was diluted in H_2O + 10% FA to a final concentration of 100 nM. To build the titration curve, serial dilution of the Native GSW11 were prepared in ACN:H_2O 1:1 (v/v) + 0,1% FA from an intermediate dilution at 1 µM as per Table 6. 2.

Volume required (µL)	ID	C (µM)	Dilution solvent	V solvent (µL)	Added solution	
					C (µM)	V (µL)
ERAP2 inhibitor	SS	10000	DMSO			
200	WS1	250	MeOH:H_2O 1:1 v/v	195	10000	5,0
200	WS2	125	MeOH:H_2O 1:1 v/v	100	250	100,0
200	WS3	100	MeOH:H_2O 1:1 v/v	40	125	160,0
200	WS4	75	MeOH:H_2O 1:1 v/v	50	100	150,0
200	WS5	50	MeOH:H_2O 1:1 v/v	67	75	133,3
200	WS6	30	MeOH:H_2O 1:1 v/v	80	50	120,0
200	WS7	25	MeOH:H_2O 1:1 v/v	33	30	166,7
200	WS8	12,5	MeOH:H_2O 1:1 v/v	100	25	100,0
200	WS9	6	MeOH:H_2O 1:1 v/v	104	12,5	96,0
200	WS10	4	MeOH:H_2O 1:1 v/v	67	6	133,3
200	WS11	3	MeOH:H_2O 1:1 v/v	50	4	150,0
200	WS12	2	MeOH:H_2O 1:1 v/v	67	3	133,3
200	WS13	1,25	MeOH:H_2O 1:1 v/v	75	2	125,0
200	WS14	0,8	MeOH:H_2O 1:1 v/v	72	1,25	128,0
200	WS15	0,5	MeOH:H_2O 1:1 v/v	75	0,8	125,0
200	WS16	0,25	MeOH:H_2O 1:1 v/v	100	0,5	100,0
200	WS17	0,15	MeOH:H_2O 1:1 v/v	80	0,25	120,0
200	WS18	0	MeOH:H_2O 1:1 v/v	200	0,15	0,0

Table 6. 1 Serial dilution of ERAP2 inhibitor to prepare the working solutions to build the titration curve for compound quantification.

Volume required (µL)	ID	C (nM)	Dilution solvent	V solvent (µL)	Added solution	
					C (nM)	V (µL)
Native GSW11	SS	1000	ACN:Water 1:1 v/v + 0,1% FA			
500	WS1	250	ACN:Water 1:1 v/v + 0,1% FA	375	1000	125,0
500	WS2	100	ACN:Water 1:1 v/v + 0,1% FA	300	250	200,0
500	WS3	50	ACN:Water 1:1 v/v + 0,1% FA	250	100	250,0
500	WS4	10	ACN:Water 1:1 v/v + 0,1% FA	400	50	100,0
500	WS5	5	ACN:Water 1:1 v/v + 0,1% FA	250	10	250,0
500	WS6	2	ACN:Water 1:1 v/v + 0,1% FA	300	5	200,0
500	WS7	1	ACN:Water 1:1 v/v + 0,1% FA	250	2	250,0
500	WS8	0,5	ACN:Water 1:1 v/v + 0,1% FA	250	1	250,0
500	WS9	0,25	ACN:Water 1:1 v/v + 0,1% FA	250	0,5	250,0
500	WS10	0,1	ACN:Water 1:1 v/v + 0,1% FA	300	0,25	200,0
500	WS11	0,075	ACN:Water 1:1 v/v + 0,1% FA	125	0,1	375,0
500	WS12	0	ACN:Water 1:1 v/v + 0,1% FA	500	0,075	0,0

Table 6. 2 Serial dilution of Native GSW11 to prepare the working solutions to build the titration curve for peptide quantification.

6.2 Preclinical models

6.2.1 Animals

HLA-B27/human β2-microglobulin (33-3 line) transgenic rats are backcrossed with the Fischer 344 strain. Rats were maintained in a specific pathogen-free facility at the Institute Pasteur (Lille, France)

and were fed a standard diet with free access to water. The animal rooms were maintained at 21–22 °C, ~ 55% relative humidity, and 12 h light/dark cycle. Experiments were carried out according to the European directive 2016/63/UE enforced by the decree n°2013-118 and authorized by the departmental ethics committee (n° reference APAFIS #34626-2022011018321312 v3).
Female CD1 mice (7 weeks old) were purchased by Charles River Laboratories. Animals were housed in a temperature-controlled room (20-24°C) and maintained in a 12h light/ 12h dark cycle. Animals had access ad libitum to drinking water and food before the study. All experimental procedures were approved by and conducted in accordance with the regulations of the local Animal Welfare authorities (n° reference APAFIS #27003-2020082815505003).

6.2.2 Cells

Cells were cultured in a cell culture incubator under constant humidity at 37°C and with 5% CO2, using RPMI 1640 (EuroClone, Pero, MI, Italy) supplemented with 10% foetal bovine serum (Glibco, Grand Island, NY, USA), 2 mM L-glutamine (EuroClone, Pero, MI, Italy), 100 I.U./mL penicillin (EuroClone, Pero, MI, Italy), 100 µg/ml streptomycin (EuroClone, Pero, MI, Italy), 50 µM 2-Mercaptoethanol (Glibco, Grand Island, NY, USA), 1 mM sodium pyruvate (EuroClone, Pero, MI, Italy), and 1 mM HEPES (EuroClone, Pero, MI, Italy). Cells were seeded sub-confluent and splitted when reaching confluency in sterile condition. Bacterial contamination was assessed using a cell culture microscope and tests for mycoplasma contamination were performed. CT26 is a murine colorectal carcinoma cell line which is from a BALB/c mouse. The cell is a clone of the N-nitroso-N-methylrethane-induced undifferentiated CT26 colon carcinoma cell line. These cells are adherent and have a fibroblast morphology.

6.3 Sample collection

6.3.1 Whole-body mice flash freezing procedure

Carcasses of sacrificed animals were places in a cryo-mould prepared by making a squared aluminium mould to maintain the animal. The mouse has to be laid down smoothly on the dorsal region without compression of the organs to avoid any displacement. The cryo-mould holding the carcass was places in previously prepared isopentane bath cooled down with dry ice. Freezing time was adapted to the size of the animal. In general mice around 30g take 5 min in isopentane/dry ice to freeze. Excess of isopentane was removed by gentle tapping onto paper towel and frozen carcasses were stored at -80°C until use.

6.3.2 Organ's extraction and flash freezing procedure

Clean scalpels and forceps were used to isolate the organ of interest from rat carcasses after sacrifice. Samples were laid down on prechilled aluminium foils covering dry ice. The freezing time for medium sized tissues on dry ice is about 3 min. Frozen samples were stored at -80°C until use.

6.3.3 Cells collection

CT26 cells were lifted from culture flasks with EDTA (Sigma Aldrich, St. Louis, MO, USA) 0.2 mM, resuspended in PBS (EuroClone, Pero, MI, Italy) and centrifuged at 2000 rpm for 3 minutes to wash the pellet. After PBS aspiration, pellets were stored at -80°C until use.

6.4 Sample preparation

6.4.1 Tissue sectioning

Spleen, colon, and cecum sectioning was performed by mounting samples onto cryostat chunk with minimal OCT embedding matrix and cryosectioning in a 6250 cryo-microtome (Dakewe Medical, Shenzhen, China) at -20°C. Eyes were embedded in 1.5% carboxymethylcellulose (CMC) in water and incubated at -80°C until solidification, before mounting onto cryostat chunk with minimal OCT and sectioning in a 6250 cryo-microtome (Dakewe Medical, Shenzhen, China) at -20°C. 10 µm thick sections were thaw-mounted on indium tin oxide (ITO)-coated microscope slides (Delta Technologies, Loveland, CO, USA) for MSI analysis. 5 µm thick sections were thaw-mounted onto SuperFrost Plus™ Adhesion slides (Epredia, Kalamazoo, MI, USA) for histological and immunofluorescence analysis. Four to six 10 µm thick sections were transferred in one 2 mL tube pre-filled with beads (Thermo Fisher Scientific, Waltham, MA, USA) for LC-MS/MS analysis.
Ankle joints were embedded in 5% CMC in water and incubated at -80°C until solidification, before mounting onto cryostat chunk with minimal OCT and collecting 10 µm thick sections on tape using a Microm HM 560 Cryostat (Thermo Fisher Scientific, Waltham, MA, USA) at -25°C.
Whole-body sectioning was performed by shaving the left side of the carcasses, removing ears, tail and front paws stabilizing the carcasses onto cryostat chunk with minimal OCT and collecting 20 µm thick sections using a Microm HM 560 Cryostat (Thermo Fisher Scientific, Waltham, MA, USA) at -25°C. Sections were thaw-mounted on indium tin oxide (ITO)-coated microscope slides (Delta Technologies, Loveland, CO, USA) for MSI analysis.

6.4.2 Microsomal stability

The analyte was incubated at 50 µM with 1 mg/ml of female CD1 mouse microsomes at 37°C for 2 hours. The reaction mix was prepared as follow: 50 mM phosphate buffer, pH 7.4, in the presence of

$MgCl_2$ (5 mM), NADP (1 mM), glucose-6-phosphate dehydrogenase (0.4 U/mL), and glucose-6-phosphate (5 mM). Two control reactions were prepared: one lacking NADP and one using PBS instead of the reaction mixture. To study *in vitro* metabolism an aliquot of 150 µL was taken at 0, 30, 60, 90, and 180 minutes. The reaction was stopped by adding 4 volumes of acetonitrile. After centrifugation at 12000 g for 10 minutes at 4°C, the supernatants were dry evaporated (35°C, 5h) and stored at -20°C until use.

6.4.3 Extraction protocol from tissue sections for LC-MS/MS metabolomic analysis

Four to six 10 µm thick cryosections per sample were transferred in one 2 mL tube pre-filled with beads (Thermo Fisher Scientific, Waltham, MA, USA) per group: WT colon, HLA-$B27^+$ colon, WT cecum and HLA-$B27^+$ cecum samples. Aqueous and organic extracts were obtained following the protocol from (Want et al., 2013). Briefly 0.5 mL of prechilled 1:1 MeOH:H_2O was added in the tube. Samples were homogenized using Bead Mill 24 homogenizer (5 m/s, 10 sec cycle) (Thermo Fisher Scientific, Waltham, MA, USA). After centrifugation (16000g, 10 min, 4°C) (Thermo Fisher Scientific, Waltham, MA, USA) the supernatant was dried using SpeedVac SPD120 (Program 1, 40°C, 2 hours cycle) (Thermo Fisher Scientific, Waltham, MA, USA). 0.5 mL of prechilled 3:1 CH_2Cl_2:MeOH was added to the solid precipitate. Samples were homogenized again. After centrifugation, the supernatant was dried in fume hood (6 hours, RT). 200 µL of prechilled 1:1 MeOH:H_2O was added to the remaining aqueous solvent and dried.

6.4.4 Extraction protocol from cells pellets for LC-MS/MS peptide analysis

100 million CT26 cell pellets were resuspended in 800 µl of 10% formic acid in H_2O with 100 nM of IS and with or without 5 µM of "martyr" irrelevant peptide. The solution was heated at 95°C for 10 minutes in a Heratherm oven (Thermo Fisher Scientific, Waltham, MA, USA). After centrifugation (16000g, 10 min) (Thermo Fisher Scientific, Waltham, MA, USA), the supernatant was filtered using a Microcon-10kDa centrifugal filter unit with ultracel-10 membrane (MRCPRT010 Merck, Rahway, NJ, USA) through centrifugation (16000g, 30 min). The filtrate was dried using a SpeedVac SPD120 (Program 3, 2 hours cycle) (Thermo Fisher Scientific, Waltham, MA, USA).

6.5 MALDI-FTICR-MSI analysis

6.5.1 Spray

Nitrogen flow pressure, flow rate, slide selection (single or big), overspray margins, and temperature are set as per method. Parameters of the spraying method for the matrices used for spatial metabolomics and biodistribution studies are listed in Table 6. 3. TM sprayer control software v 4.1 is used to set up the parameter and start the run. ITO slide is secured inside the spraying chamber

using paper scotch. With the valve on load and the flow rate directed to the needle, the required volume of matrix solution is loaded using a syringe. Once all the parameters have reached the required nominal value and matrix solution is loaded, on the instrument valve is moved to spray and on the software spraying method is started. The ITO slides are weighted before and after spraying, to monitor the quantity of matrix deposed over time, used as an additional QC.

	Temp (°C)	Passes (count)	Flow rate (ml/min)	Velocity (mm/min)	Pressure (psi)	Gas flow rate (l/min)	Overspray margins (mm)
DHB organs	75	12	0,09	1200	10	3	2 (r, t) 4 (l, b)
DAN organs	60	10	0,1	1200	10	2	2 (r, t) 4 (l, b)
DHB whole-body	75	12	0,09	1200	10	3	15 (r, l, t, b)

Table 6. 3 Parameters of the spraying methods for the matrices used for metabolomics and biodistribution studies.

6.5.2 Acquisition methods

6.5.2.1 MALDI-MSI for spatial metabolomics analysis

MALDI-MSI for spatial metabolomics analysis was performed on colon and cecum tissue sections at a lateral resolution of 40 µm using a SolariX 7T-MALDI-FTICR (Bruker Daltonics, Bremen, Germany) in positive and negative ionization mode. The instrument was equipped with Smartbeam II laser operated in small mode at a repetition rate of 2000 Hz; laser power was optimized prior to acquisition. Spectra were collected by summing signals from 300 laser shots per pixel. Mass spectra were acquired in full scan mode over the mass range *m/z* 75-1200 Da. The quadrupole isolation *m*/*z* (Q1 mass) was set at *m*/*z* 100. A list of peaks, listed in Table 6. 4, was used as a lock masses for internal *m*/*z* calibration. The TOF value was adjusted to 0.600 ms and 0.500 ms, in positive in negative ion mode respectively, and the transfer optics frequency was 4 MHz. Samples were acquired in a random order to prevent bias arising from matrix vacuum instability or changes in mass spectrometer sensitivity.

Positive ionization mode		**Negative ionization mode**	
Name	***m/z***	**Name**	***m/z***
DHB	273.0394	G6P	259.02244
DHB	313.0318	GSH	306.07653
DHB	335.01372	DAN	312.138045
DHB	409.0554	DAN	315.16152
PC32:0	734.5654	AMP	346.05581
PC34:1	760.5856	PA34:1	673.4814
PC36:4	782.5654	PA36:2	699.497
PC34:1	798.5415	PE34:1	716.5236
PC34:0	800.5566	PE36:2	742.5386
PC36:4	820.5259	PE36:1	744.5548
PC38:4	848.5572	PE38:5	764.5229
		PE38:4	766.5392
		PE38:2	770.5699
		PE-p40:6	776.5589
		PS36:2	786.5286
		PS36:1	788.5447
		PE40:6	790.5392
		PE40:4	794.5699
		PS38:4	810.5284
		PI34:0	837.5494
		PI38:4	885.5493

Table 6. 4 List of peaks used as a lock masses for internal *m*/*z* calibration.

6.5.2.2 MALDI-MSI for biodistribution analysis

MALDI-MSI for biodistribution analysis was performed on whole-body sections at a lateral resolution of 350 µm using a SolariX 7T-MALDI-FTICR (Bruker Daltonics, Bremen, Germany) in positive ionization mode. The instrument was equipped with Smartbeam II laser operated in medium mode at a repetition rate of 2000 Hz; laser power was optimized prior to acquisition. Spectra were collected by summing signals from 300 laser shots per pixel. Mass spectra were acquired in CASI mode. The quadrupole isolation *m*/*z* (Q1 mass) was set at *m*/*z* 514.12, with an isolation window of 100 Da. The TOF and transfer optics frequency values were adjusted to 0.800 ms and 6 MHz, respectively. The IS peak at *m*/*z* 518.1464 was used as a lock mass for internal *m*/*z* calibration.

6.5.2.3 MALDI-MSI for in vivo metabolism analysis

MALDI-MSI for in vivo metabolism analysis was performed on whole-body sections at a lateral resolution of 350 µm using a SolariX 7T-MALDI-FTICR (Bruker Daltonics, Bremen, Germany) in positive ionization mode. The instrument was equipped with Smartbeam II laser operated in medium mode at a repetition rate of 2000 Hz; laser power was optimized prior to acquisition. Spectra were collected by summing signals from 300 laser shots per pixel. Mass spectra were acquired in CASI mode. To investigate Phase I metabolism, the quadrupole isolation *m*/*z* (Q1 mass) was set at *m*/*z* 514, with an isolation window of 200 Da. The TOF and transfer optics frequency values were adjusted to 0.700 ms and 4 MHz, respectively. To investigate Phase II metabolism, mass spectrum was acquired in full scan mode over the mass range *m/z* 100-1000 Da. The quadrupole isolation *m*/*z* (Q1 mass) was

set at *m*/*z* 100. The TOF and transfer optics frequency values were adjusted to 0.700 ms and 4 MHz, respectively. The DHB peak at *m*/*z* 273.0393 was used as a lock mass for internal *m*/*z* calibration.

6.6 MALDI-FTICR-MS for microsomal stability sample analysis

6.6.1 Acquisition method

MALDI-MS for microsomal stability sample analysis was performed on experimental and control sample spotted on a MTP 384 target plate and mixed with DHB (40 mg/mL in MeOH/H_2O 1:1 v/v) using a SolariX 7T-MALDI-FTICR (Bruker Daltonics, Bremen, Germany) in positive ionization mode. The instrument was equipped with Smartbeam II laser operated in medium mode at a repetition rate of 2000 Hz; laser power was optimized prior to acquisition. Spectra were obtained by averaging 32 scans with 300 laser shots per scan. Mass spectra were acquired in CASI mode. The quadrupole isolation *m*/*z* (Q1 mass) was set at *m*/*z* 514.12, with an isolation window of 200 Da. The TOF and transfer optics frequency values were adjusted to 0.700 ms and 4 MHz, respectively.

6.7 LC-MS/MS analysis

6.7.1 Chromatographic method, gradient and mobile phases

6.7.1.1 LC metabolomic

Tissue aqueous dry extracts were resuspended in 1:1 ACN:H_2O (Figure 6. 1). One half of the extract was analysed in reversed-phase mode. Water was added to obtain a final injection solution 1:5 ACN:H_2O and chromatographic separation was performed on a LC30AD UPLC (Shimadzu, Kyoto, Japan) using a ACQUITY UPLC Peptide BEH C18 reversed-phase column (100x2.1mm, particle size 1.7µm, pore size 300Å, Waters 186003686, Milford, MA, USA) at 65°C and 0.4 mL/min. Mobile phase A consists of 0.1% formic acid in water, and mobile phase B consists of 0.1% formic acid in acetonitrile. The injection volume was 5.0 µL. The gradient was as follows: 0 min-1% B, 1 min-1% B, 14 min-80% B, 14.1 min-98% B, 15 min-98% B, 15.1 min-1% B and stop at 20 min. An integrated valve diverted the LC flow to waste before 0.25 minutes and after 17 minutes. The second half of the extract was analysed using a HILIC stationary phase. Acetonitrile was added to obtain a final injection solution 5:1 ACN:H_2O and chromatographic separation was performed on a LC30AD UPLC using an Atlantis Premier BEH Z-HILIC Column (100x2.1mm, particle size 1.7µm, pore size 95Å, Waters 186009979, Milford, MA, USA) at 65°C and 0.4 mL/min. Mobile phase A consists of 25 mM ammonium carbonate in water, and mobile phase B consists of 95:5 ACN:H_2O. The injection volume was 2.0 µL. The gradient was as follows: 0 min–99%B, 2 min–85%B, 5 min–75%B, 8 min–60%B, 9 min–40%B, 10 min–40%B, 11 min–1%B, 12 min–1%B, 12.1 min–99%B and stop at 20 min. An integrated valve diverted the LC flow to waste before 0.25 minutes and after 15.1 minutes.

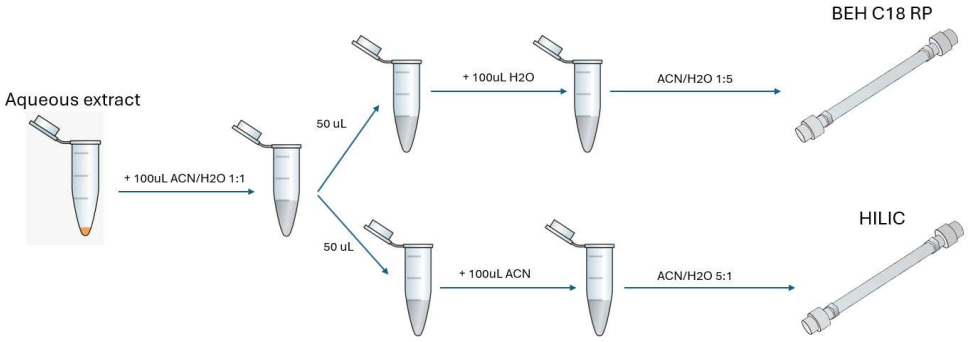


Figure 6. 1 Schematic workflow of the aqueous extract's reconstitution for metabolomic LC-MS/MS analysis

6.7.1.2 LC lipidomic

Tissue organic dry extracts were resuspended in 2:1:1 Isopropanol:ACN:H_2O (Figure 6. 2) and chromatographic separation was performed on a LC30AD UPLC using an ACQUITY Premier CSH C18 column (100x2.1mmm particle size 1.7µm, pore size 130Å Waters 186009464, Milford, MA, USA) at 55°C and 0.4 mL/min. Mobile phase A consists of 10 mM ammonium formate and 0.1% formic acid in 3:2 ACN:H_2O, and mobile phase B consists of 10 mM ammonium formate and 0.1% formic acid in 9:1 Isopropanol:ACN. The injection volume was 5.0 µL. The gradient was as follows: 0 min–40%B 0 curve, 2 min–43%B 6 curve, 2.1 min–50%B 1 curve, 12 min–54%B 6 curve, 12.1 min–70%B 1 curve, 18 min-99%B 6 curve, 18.1 min–40%B 6 curve, and stop at 20 min. An integrated valve diverted the LC flow to waste before 0.25 minutes and after 19.9 minutes.

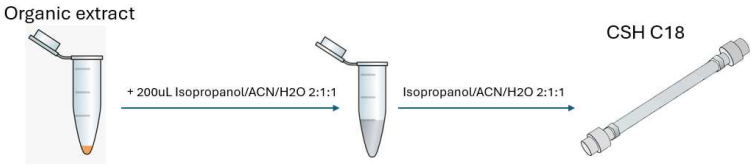


Figure 6. 2 Schematic workflow of the organic extract's reconstitution for lipidomic LC-MS/MS analysis

6.7.1.3 LC peptidomic

Cells extracts were resuspended in 1:1 ACN:H_2O + 0.1% FA and chromatographic separation was performed on a UltiMate 3000 UHPLC (Thermo Fisher Scientific, Waltham, MA, USA) using an ACQUITY UPLC BEH C18 column (50x2.1mmm particle size 1.7µm, pore size 130Å Waters 186002350, Milford, MA, USA) at 20°C and 0.4 mL/min. Mobile phase A consists of 0.1% formic acid in H_2O, and mobile phase B consists of 0.1% formic acid in ACN. The injection volume was 5.0 µL. The gradient was as follows: 0 min–10%B, 0.5 min–10%B, 1.5 min–99%B, 2.5 min–99%B, 2.6 min–10%B, and stop at 3 min.

6.7.2 MS Acquisition method

6.7.2.1 MS metabolomic and lipidomic

Analysis was performed on a ZenoTOF 7600 system (SCIEX, Framingham, MA, USA) operated in positive and negative electrospray ionization modes. The following ion source parameters were applied: capillary temperature 500 °C, nebulizer gas (GS1) 20 psi, heating gas (GS2) 30 psi, curtain gas 35 psi, and CAD gas 12 psi. The ion spray voltage (ISV) was 5000 V for positive mode and -4000 V for negative. The MRMHR acquisition method consisted of TOF MS and TOF MSMS scans collected concurrently. The TOF MS scan uses an accumulation time of 0.05 s from 40 to 1200 Da using a declustering potential (DP) of 30 eV and collision energy (CE) of 5.0 eV for positive mode. The negative mode uses an accumulation time of 1.0 s with a DP of -20 eV and CE of -5.0 eV. The TOF MSMS scans used a list of precursor masses with an accumulation time 0.005 s. The positive mode used a DP at 30 eV and CE at 30 eV. The negative ion mode used a DP of -20 eV and CE of -30 eV.

6.7.2.2 MS peptidomic

Analysis was performed on a TSQ Quantiva system (Thermo Fisher Scientific, Waltham, MA, USA) equipped with a heated electrospray (HESI) and operated in positive and negative ionization modes. The spray voltage on the HESI probe was 3500 V in positive and 2500 V in negative. Nitrogen was used as sheath gas, auxiliary gas and sweep gas at 30, 15 and 1 (arbitrary units), respectively. The ion transfer tube was kept at 350 °C and the vaporizer temperature was 325 °C. A cycle time of 0.2 s was used and resolution (FWHM) at both Q1 and Q3 was 0.7. Argon was used as collision gas at a pressure of 1.5 mTorr. The MRM acquisition is listed on Table 6. 5.

		Negative ion mode		
Analyte	**Precurson ion**	**Product ion**	**Collision energy (V)**	**Min dwell time (ms)**
Native GSW11	1201.5	1171.417	47	31
		1183.417	41	31
		1167.417	46	31
IS	1208.5	1174.467	46	31
		1178.383	47	31
		1190.647	42	31
		Positive ion mode		
Analyte	**Precurson ion**	**Product ion**	**Collision energy (V)**	**Min dwell time (ms)**
Native GSW11	1203.5	629.2	37	38
		912.4	34	38
		999.35	31	38
IS	1209.5	918.35	19	38
		1005.35	28	38

Table 6. 5 Precursor and product ions acquired for MRM peptidomic analysis

6.8 Microscopy

6.8.1 H&E

Haematoxylin and eosin (H&E) staining was performed with the Automatic sample preparation system ST4020 (Leica Biosystems, Heidelberger, Germany) that dip sequentially the slides for 60 seconds in holders containing: haematoxylin 3 times, H_2O 3 times, eosin 1:3 in ethanol, ethanol 100%, ethanol 90%, ethanol 70%, xylene.

6.8.2 IF

Slides were fixed with formalin solution neutral buffered 10% for 30 min at 4°C and incubated with 3% bovine serum albumin to block non-specific binding for 1h at room temperature. Slides were incubated overnight at 4°C with primary antibodies diluted in 0.5% bovine serum albumin: CD3 1:500 (AB245731, Abcam, Cambridge, UK); CD68 1:100 (sc-20060 AF647, Santa Cruz Biotechnology, Dallas, TA, USA). Fluorochrome-conjugated secondary antibodies were diluted in PBS and incubated for 1h at room temperature: Donkey anti-Rabbit AF594 1:2000 (A21207, Invitrogen, Waltham, MA, USA).

6.9 Data Analysis

6.9.1 MALDI-MSI data processing and analysis

6.9.1.1 Untargeted spatial metabolomic study

Bruker software suite including FTMS control v2.1, FlexImaging v5.0, and DataAnalysis v5.0 were used for recording and previsualization of the data. Data treatment, optical and molecular images were processed with Multimaging™ 1.2.8.4 software (Aliri France SAS, Loos, France). Multimaging™ software was used to perform statistical analysis. MALDI-MSI data collected by separate analysis runs were integrated and batch corrected by applying BBKNN (batch balanced k nearest neighbours), a batch effect removal tool (Polański et al., 2020). Data were projected in a UMAP tri-dimensional plot, the best nonlinear dimensionality reduction technique for MSI, based on the spectra's peaks intensities (McInnes et al., 2018; Smets et al., 2019). K-means clustering, using the cosine distance, was applied to identify biologically similar clusters (Winderbaum et al., 2015). Student's *t*-test was performed on Multimaging™ to compare regions of interest from the two groups (HLA-B27$^+$ vs WT) based on feature's signal intensity. p-Values were adjusted using the Benjamini-Hochberg (BH) correction (Benjamini & Hochberg, 1995).

6.9.1.2 Biodistribution study

Bruker software suite including FTMS control v2.1, FlexImaging v5.0, and DataAnalysis v5.0 were used for recording and previsualization of the data. Data treatment, optical and molecular images were processed with Multimaging™ 1.2.8.4 software (Aliri France SAS, Loos, France). Based on the MSI data set on the tissue sections and the dilution series, the quantitation of the analyte was performed with the internal standard approach. Multimaging™ software was used for quantitative analysis. The "normalization" function of the software was used to normalize the signal of the analyte in the spectra by the signal of the IS. The "quantification" function of the software A correlation between the calibration curve and the signal obtained on the tissues was performed to determine the concentration of the test items per histological structure in µg/g of tissue and µM.

6.9.2 LC-MS/MS data processing and analysis

6.9.2.1 Metabolomic study

Sciex software suite including Analyst v1.4.2 and SciexOS v3.1.6.44 was used for acquisition and visualization of the data. The metabolomic standard initiative guidelines were used to define the level of identification (Sumner et al., 2007). Level 2 identification was achieved by manual comparison of the MS/MS fragmentation pattern for each discriminative feature's precursor mass with open-source databases (HMDB, Lipid Maps, METASPACE) and relevant literature. Level 3 identification was obtained by manually entering each discriminative feature's precursor mass on open-source databases (HMDB, Lipid Maps, METASPACE) to seek verification (O'Donnell et al., 2019; Palmer et al., 2017; Wishart et al., 2007).

6.9.2.2 Peptidomic study

Thermo Scientific suite including TSQ Quantiva Tune Application v3.4.3293, Xcalibur Sequence Setup, Xcalibur Instrument Setup and Freestyle was used for acquisition and visualization of the data. Excalibur Quan Browser was used for quantitative analysis. The chromatographic peak areas of the native GSW11 peptide were normalized by the signal of the internal standard and the titration curve was built.

6.10 Microscopy data analysis

6.10.1.1 Histopathological evaluation

The assessment of intestinal inflammation was conducted by a pathologist through direct macroscopic and microscopic examination of H&E-stained tissue sections.

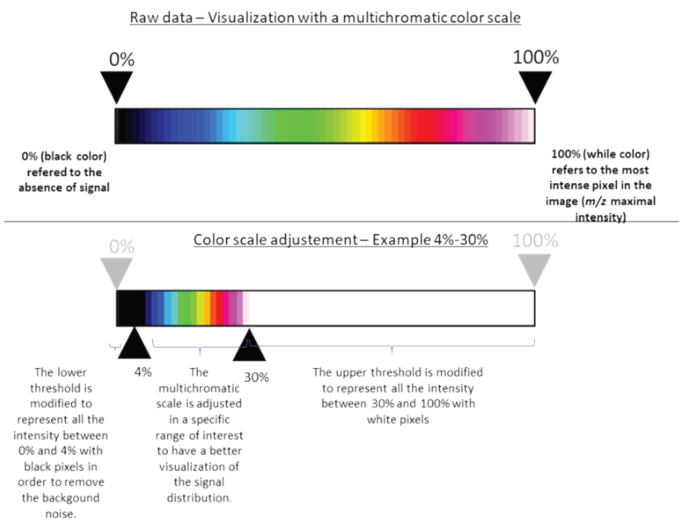
6.10.1.2 Immunofluorescence data processing and analysis

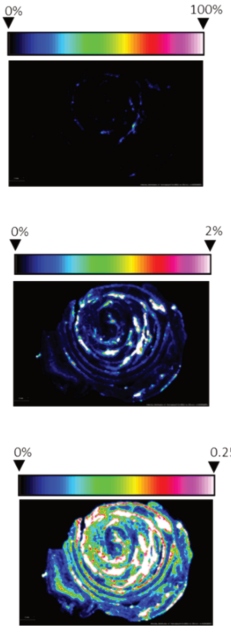
Single-cell segmentation was performed on QuPath v0.5 using StarDist, a nuclei-based deep learning algorithm (Schmidt et al., 2018; Weigert et al., 2020; Weigert & Schmidt, 2022). Cells were classified as CD3 or CD68 positive based on the mean intensity of these markers. Overlay between immunofluorescence and H&E images and density maps were created using the ImageCombinerWarpy extension in QuPath v3.2 (Chiaruttini et al., 2022).

7 Annexes

7.1 Annex 1

Colour Scale Adjustment for Optimized Visualization of a Molecular Distribution





7.2 Annex 2

List of features discriminating between HLA-B27$^+$ and WT samples in colon and cecum compartments, including their *m/z*, ion mode and histological region in which they were detected and details on annotation

Compartment	Group	Feature	Ion Mode	Region	Annotation	Adduct	Category	Class	Id level
Cecum	HLA-B27+	mz 124.0074	-	Mucosa&Submucosa					4
Cecum	HLA-B27+	mz 132.0304	-	Submucosa	Aspartic acid	M-H		Aminoacid	2
Cecum	HLA-B27+	mz 140.0118	-	Mucosa&Submucosa	Phosphoethanolamine	M-H		Phosphate esters	2
Colon	HLA-B27+	mz 141.332	+	Mucosa					4
Cecum	HLA-B27+	mz 146.0459	-	Mucosa&Submucosa					4

Cecum	HLA-B27+	mz 146.1647	+	Mucosa	Spermidine	M+H			3
Colon	WT	mz 147.0758	+	Lumen&Mucosa&Submucosa	Glutamine	M+H		Aminoacid	2
Cecum	WT	mz 147.076	+	Mucosa&Submucosa	Glutamine	M+H		Aminoacid	2
Colon	HLA-B27+	mz 158.0398	+	Lumen&Epithelium					4
Cecum	HLA-B27+	mz 160.1111	-	Epithelium					4
Colon	HLA-B27+	mz 162.1121	+	Submucosa&Muscularis Propria	Carnitine (L or S)	M+H		Quaternary ammonium compound	2
Cecum	WT	mz 169.058	+	Mucosa&Submucosa	Glutamine	M+Na		Aminoacid	3
Colon	WT	mz 169.0581	+	Lumen&Mucosa&Submucosa	Glutamine	M+Na		Aminoacid	3
Cecum	HLA-B27+	mz 174.0409	-	Submucosa					4
Colon	HLA-B27+	mz 179.982	+	Patches	Phosphoethanolamine	M+K		Phosphate esters	3
Cecum	WT	mz 185.0319	+	Mucosa&Submucosa					4
Colon	HLA-B27+	mz 200.0682	+	Muscularis Propria					4
Colon	HLA-B27+	mz 201.9638	+	Patches					4
Colon	WT	mz 203.0528	+	Mucosa					4
Colon	WT	mz 207.0138	+	Mucosa	3-Oxalomalic acid	M+H		Carboxylic acids and derivatives	3
Colon	HLA-B27+	mz 211.9963	+	Patches&Lamina Propria					4
Cecum	WT	mz 213.9372	-	Submucosa					4
Colon	HLA-B27+	mz 214.0488	-	Lumen&Epithelium	Glycerylphosphorylethanolamine	M-H	Glycerophospholipids	Glycerophosphoethanolamines	2
Cecum	HLA-B27+	mz 214.0488	-	Epithelium	Glycerylphosphorylethanolamine	M-H	Glycerophospholipids	Glycerophosphoethanolamines	2
Cecum	HLA-B27+	mz 215.052	-	Mucosa&Submucosa					4

Colon	HLA-B27+	mz 217.0832	-	Lumen&Mucosa	Glutamylalanine	M-H		Dipeptide	3
Cecum	HLA-B27+	mz 217.0832	-	Lumen&Mucosa&Submucosa	Glutamylalanine	M-H		Dipeptide	3
Colon	HLA-B27+	mz 217.9378	+	Patches&Lamina Propria					4
Cecum	HLA-B27+	mz 217.9382	+	Patches					4
Colon	WT	mz 219.0266	+	Mucosa					4
Cecum	HLA-B27+	mz 219.9788	-	Submucosa	Phosphoserine	M+Cl		Phosphate esters	3
Colon	HLA-B27+	mz 221.186	+	Lumen					4
Cecum	HLA-B27+	mz 222.029	+	Mucosa					4
Colon	HLA-B27+	mz 226.049	-	Mucosa					4
Colon	WT	mz 232.1542	+	Mucosa	Butyrylcarnitine	M+H	Fatty Acyls	Acyl carnitines	2
Colon	HLA-B27+	mz 242.0776	-	Epithelium	Cytidine	M-H		Nucleoside	2
Cecum	HLA-B27+	mz 244.0828	-	Submucosa					4
Colon	HLA-B27+	mz 244.8704	+	Patches&Lamina Propria					4
Cecum	HLA-B27+	mz 244.8732	+	Patche&Mucosa					4
Cecum	HLA-B27+	mz 245.0434	-	Epithelium	Glycerophosphoglycerol	M-H	Glycerophospholipids	Glycerophosphoglycerols	2
Colon	HLA-B27+	mz 254.019	+	Epithelium					4
Colon	HLA-B27+	mz 256.8542	+	Epithelium					4
Colon	HLA-B27+	mz 257.5224	+	Patches&Mucosa					4

Cecum	HLA-B27+	mz 257.5262	+	Patches&Mucosa				4
Colon	HLA-B27+	mz 259.985	+	Patches&Lamina Propria				4
Cecum	HLA-B27+	mz 260.0892	-	Mucosa				4
Colon	HLA-B27+	mz 264.04	+	Lamina propria				4
Cecum	WT	mz 265.0818	+	Submucosa				4
Colon	WT	mz 267.0704	-	Submucosa	Inosine	M-H	Nucleoside	2
Colon	WT	mz 269.1076	+	Muscularis Propria				4
Cecum	HLA-B27+	mz 269.5464	+	Mucosa&Submucosa				4
Colon	WT	mz 270.1092	+	Mucosa				4
Cecum	HLA-B27+	mz 270.199	-	Muscularis Propria				4
Cecum	HLA-B27+	mz 270.2218	+	Mucosa&Submucosa				4
Cecum	HLA-B27+	mz 270.557	+	Mucosa&Submucosa				4
Cecum	HLA-B27+	mz 273.5284	+	Mucosa&Submucosa				4
Cecum	HLA-B27+	mz 274.1414	-	Submucosa				4
Cecum	HLA-B27+	mz 275.0888	-	Patches&Lamina Propria	Glutamylglutamic acid	M-H	Dipeptide	3
Colon	HLA-B27+	mz 275.089	-	Lamina Propria	Glutamylglutamic acid	M-H	Dipeptide	3
Cecum	HLA-B27+	mz 275.108	-	Mucosa				4
Colon	HLA-B27+	mz 276.0014	+	Lumen&Mucosa				4
Cecum	HLA-B27+	mz 276.091	-	Mucosa&Submucosa				4

Cecum	WT	mz 282.0112	-	Submucosa					4
Cecum	HLA-B27+	mz 282.0853	-	Submucosa	Guanosine	M-H		Nucleoside	2
Cecum	HLA-B27+	mz 282.8736	+	Mucosa&Submucosa					4
Colon	WT	mz 283.0928	+	Mucosa					4
Cecum	WT	mz 283.0928	+	Mucosa&Submucosa					4
Cecum	HLA-B27+	mz 283.9794	-	Mucosa&Submucosa					4
Cecum	HLA-B27+	mz 286.0934	-	Submucosa					4
Cecum	HLA-B27+	mz 289.1156	-	Submucosa	Argininosuccinic acid	M-H		Aminoacid	3
Cecum	HLA-B27+	mz 290.0884	-	Submucosa					4
Colon	HLA-B27+	mz 290.0886	-	Mucosa&Submucosa					4
Cecum	HLA-B27+	mz 291.9612	-	Mucosa&Submucosa					4
Colon	HLA-B27+	mz 291.9754	+	Lumen&Mucosa					4
Cecum	HLA-B27+	mz 292.0726	-	Mucosa&Muscularis Mucosae	Glycerophosphocholine	M+Cl	Glycerophospholipids	Glycerophosphocholines	3
Cecum	WT	mz 297.011	-	Mucosa					4
Colon	HLA-B27+	mz 300.0396	-	Mucosa	N-Acetylgalactosamine 6/4-sulphate	M-H		Carbohydrates and carbohydrate conjugates	2
Cecum	HLA-B27+	mz 300.0402	-	Mucosa&Submucosa	N-Acetylgalactosamine 6/4-sulphate	M-H		Carbohydrates and carbohydrate conjugates	2
Cecum	HLA-B27+	mz 300.0495	-	Submucosa					4
Cecum	HLA-B27+	mz 302.9829	-	Mucosa&Submucosa					4
Colon	HLA-B27+	mz 307.2651	-	Epithelium					4

Cecum	WT	mz 308.0328	-	Submucosa					4
Colon	HLA-B27+	mz 309.2804	-	Mucosa	FA(20:1)	M-H	Fatty Acyls	Fatty acids	3
Cecum	HLA-B27+	mz 316.0209	-	Submucosa					4
Cecum	HLA-B27+	mz 323.0292	-	Epithelium	UMP	M-H		Nucleotide	2
Cecum	HLA-B27+	mz 324.0135	-	Patches					4
Cecum	HLA-B27+	mz 325.0449	-	Mucosa					4
Colon	HLA-B27+	mz 327.2337	-	Lamina propria	FA(22:6)	M-H	Fatty Acyls	Fatty acids	3
Cecum	HLA-B27+	mz 329.0868	+	Mucosa					4
Colon	HLA-B27+	mz 329.2491	-	Lamina propria	FA(22:5)	M-H	Fatty Acyls	Fatty acids	3
Colon	HLA-B27+	mz 331.2645	-	Lamina propria	FA(22:4)	M-H	Fatty Acyls	Fatty acids	2
Cecum	HLA-B27+	mz 331.9947	-	Patches					4
Cecum	HLA-B27+	mz 333.0592	-	Epithelium	Glycerophosphoinositol	M-H	Glycerophospholipids	Glycerophosphoinositols	2
Cecum	HLA-B27+	mz 334.1151	-	Submucosa					4
Cecum	WT	mz 337.0446	-	Submucosa	Imidazoleacetic acid ribotide	M-H		Nucleoside	3
Colon	WT	mz 337.0449	-	Submucosa	Imidazoleacetic acid ribotide	M-H		Nucleoside	3
Colon	WT	mz 338.2776	+	Lumen					4
Cecum	HLA-B27+	mz 348.072	-	Patches					4
Cecum	WT	mz 349.0596	-	Submucosa					4
Cecum	HLA-B27+	mz 350.1097	-	Mucosa&Submucosa					4

Cecum	WT	mz 357.2802	+	Mucosa					4
Colon	HLA-B27+	mz 361.2108	+	Epithelium					4
Colon	HLA-B27+	mz 362.0163	+	Mucosa					4
Cecum	HLA-B27+	mz 363.0009	+	Mucosa	UMP	M+K		Nucleotide	3
Cecum	HLA-B27+	mz 365.1354	-	Submucosa					4
Colon	WT	mz 366.105	-	Mucosa					4
Colon	HLA-B27+	mz 366.2784	+	Epithelium					4
Colon	HLA-B27+	mz 366.9033	+	Lamina Propria					4
Colon	HLA-B27+	mz 367.3005	+	Patches&Mucosa					4
Cecum	HLA-B27+	mz 367.3032	+	Patches&Mucosa					4
Colon	HLA-B27+	mz 369.0264	+	Epithelium					4
Colon	HLA-B27+	mz 372.3123	+	Lumen	Tetradecanoylcarnitine	M+H	Fatty Acyls	Acyl carnitines	2
Cecum	HLA-B27+	mz 377.2467	-	Mucosa					4
Colon	HLA-B27+	mz 382.8948	+	Patches&Mucosa					4
Colon	HLA-B27+	mz 384.0008	+	Patches&Mucosa					4
Cecum	HLA-B27+	mz 384.9825	+	Mucosa&Submucosa					4
Cecum	HLA-B27+	mz 385.2627	+	Mucosa					4
Colon	HLA-B27+	mz 386.2638	+	Patches&Mucosa					4

Cecum	WT	mz 388.9836	-	Submucosa					4
Colon	HLA-B27+	mz 389.1222	+	Patches&Mucosa					4
Cecum	WT	mz 389.2707	-	Lumen					4
Cecum	HLA-B27+	mz 398.0397	+	Mucosa					4
Colon	HLA-B27+	mz 398.328	+	Lumen	Hexadecenoylcarnitine	M+H	Fatty Acyls	Acyl carnitines	2
Colon	HLA-B27+	mz 399.9726	+	Patches&Mucosa					4
Colon	HLA-B27+	mz 400.3436	+	Lumen	Palmitoylcarnitine	M+H	Fatty Acyls	Acyl carnitines	2
Colon	HLA-B27+	mz 405.0968	+	Patches&Mucosa					4
Cecum	HLA-B27+	mz 405.3276	+	Submucosa					4
Cecum	HLA-B27+	mz 410.2844	+	Mucosa&Submucosa					4
Colon	HLA-B27+	mz 414.0136	+	Patches&Mucosa					4
Colon	HLA-B27+	mz 416.3388	+	Lumen	2-Hydroxyhexadecanoylcarnitine	M+H	Fatty Acyls	Acyl carnitines	2
Cecum	WT	mz 417.1204	-	Mucosa					4
Cecum	HLA-B27+	mz 424.3016	+	Mucosa&Submucosa					4
Colon	HLA-B27+	mz 424.3436	+	Lumen	Linoleyl carnitine	M+H	Fatty Acyls	Acyl carnitines	2
Colon	HLA-B27+	mz 426.36	+	Lumen	CAR(18:1)	M+H	Fatty Acyls	Acyl carnitines	3
Colon	HLA-B27+	mz 428.3751	+	Lumen	Stearoylcarnitine	M+H	Fatty Acyls	Acyl carnitines	2
Colon	HLA-B27+	mz 432.0852	+	Epithelium					4
Cecum	HLA-	mz 435.12	-	Patches&Mucosa					4

	B27+								
Colon	HLA-B27+	mz 435.9952	+	Patches					4
Colon	HLA-B27+	mz 436.2836	-	Lamina Propria	LPE(16:0)	M-H	Glycerophospholipids	Glycerophosphoethanolamines	2
Cecum	HLA-B27+	mz 436.2836	-	Lamina Propria	LPE(16:0)	M-H	Glycerophospholipids	Glycerophosphoethanolamines	2
Cecum	HLA-B27+	mz 439.2261	-	Patches					4
Colon	HLA-B27+	mz 439.9784	+	Mucosa					4
Cecum	HLA-B27+	mz 442.7732	-	Patches					4
Cecum	HLA-B27+	mz 458.1656	-	Mucosa					4
Cecum	HLA-B27+	mz 460.284	-	Mucosa					4
Cecum	HLA-B27+	mz 463.3032	-	Mucosa					4
Colon	HLA-B27+	mz 464.0832	+	Mucosa					4
Cecum	WT	mz 465.3056	-	Epithelium&Lumen	Cholesterol sulphate	M-H	Sterol Lipids	Cholesterol esters	2
Cecum	HLA-B27+	mz 472.1672	-	Mucosa					4
Cecum	WT	mz 473.1464	-	Mucosa	Enterolactone 3-glucuronide	M-H		Lignan glycosides	2
Cecum	WT	mz 478.2972	+	Mucosa&Lumen	LPE(18:2)	M+H	Glycerophospholipids	Glycerophosphoethanolamines	2
Colon	HLA-B27+	mz 480.924	+	Mucosa					4
Colon	HLA-B27+	mz 481.99	+	Patches					4
Cecum	HLA-B27+	mz 482.362	+	Lamina Propria	LPC(O-16:0)	M+H	Glycerophospholipids	Glycerophosphocholines	2
Colon	HLA-B27+	mz 482.3632	+	Lamina Propria	LPC(O-16:0)	M+H	Glycerophospholipids	Glycerophosphocholines	2

Colon	HLA-B27+	mz 489.118	+	Mucosa					4
Colon	HLA-B27+	mz 502.9055	+	Lamina Propria					4
Colon	WT	mz 504.2735	-	Lumen					4
Colon	WT	mz 508.341	-	Epithelium	LPE(20:0)	M-H	Glycerophospholipids	Glycerophosphoethanolamines	3
Colon	HLA-B27+	mz 508.3785	+	Lamina Propria	LPC(P-18:0) or LPC(O-18:1)	M+H	Glycerophospholipids	Glycerophosphocholines	2
Cecum	HLA-B27+	mz 508.379	+	Lamina Propria	LPC(P-18:0) or LPC(O-18:1)	M+H	Glycerophospholipids	Glycerophosphocholines	2
Colon	HLA-B27+	mz 511.0995	+	Lamina Propria					4
Cecum	HLA-B27+	mz 512.1665	-	Submucosa					4
Cecum	WT	mz 516.2535	+	Lumen	LPE(18:2)	M+K	Glycerophospholipids	Glycerophosphoethanolamines	3
Colon	HLA-B27+	mz 518.8792	+	Lamina Propria					4
Colon	HLA-B27+	mz 519.946	+	Patches&Lamina Propria					4
Colon	HLA-B27+	mz 520.3192	+	Lamina Propria	LPC(18:2)	M+H	Glycerophospholipids	Glycerophosphocholines	2
Cecum	HLA-B27+	mz 520.3205	+	Lamina Propria	LPC(18:2)	M+H	Glycerophospholipids	Glycerophosphocholines	2
Colon	WT	mz 521.3465	+	Lumen					4
Cecum	HLA-B27+	mz 524.2995	-	Lamina propria	LPS(18:0)	M-H	Glycerophospholipids	Glycerophosphoserines	2
Colon	HLA-B27+	mz 527.073	+	Lamina Propria					4
Cecum	HLA-B27+	mz 529.3055	-	Mucosa&Submucosa					4
Cecum	HLA-B27+	mz 536.5055	-	Epithelium	Cer(34:1)	M-H	Sphingolipids	Ceramides	3
Colon	WT	mz 542.071	+	Muscularis Propria&Muscularis Mucosae					4

Colon	WT	mz 542.3255	+	Lumen&Mucosa	LPC(20:5)	M+H	Glycerophospholipids	Glycerophosphocholines	3
Colon	HLA-B27+	mz 546.3345	+	Lamina Propria	LPC(20:3)	M+H	Glycerophospholipids	Glycerophosphocholines	2
Colon	WT	mz 547.3605	+	Mucosa					4
Colon	HLA-B27+	mz 549.056	+	Mucosa					4
Colon	HLA-B27+	mz 554.5155	-	Epithelium	Cer(t18:0/16:0)	M-H	Sphingolipids	Ceramides	3
Colon	WT	mz 559.304	+	Lumen					4
Colon	HLA-B27+	mz 565.0315	+	Mucosa					4
Colon	HLA-B27+	mz 599.321	-	Lamina propria	LPI(18:0)	M-H	Glycerophospholipids	Glycerophosphoinositols	3
Cecum	HLA-B27+	mz 599.3215	-	Lamina propria	LPI(18:0)	M-H	Glycerophospholipids	Glycerophosphoinositols	2
Colon	WT	mz 602.0094	+	Muscularis Propria&Muscularis Mucosa3					4
Cecum	WT	mz 613.5618	+	Mucosa					4
Colon	HLA-B27+	mz 615.9702	+	Mucosa					4
Cecum	HLA-B27+	mz 632.6364	+	Mucosa&Submucosa	Cer(42:2)	M+H	Sphingolipids	Ceramides	2
Cecum	HLA-B27+	mz 635.4528	-	Submucosa					4
Colon	HLA-B27+	mz 647.4654	-	Patches	PA(32:0)	M-H	Glycerophospholipids	Glycerophosphates	3
Cecum	HLA-B27+	mz 647.466	-	Patches	PA(32:0)	M-H	Glycerophospholipids	Glycerophosphates	3
Colon	WT	mz 664.1232	+	Muscularis Propria&Muscularis Mucosae					4
Cecum	WT	mz 664.1238	+	Muscularis Propria&Submucosa					4
Colon	HLA-B27+	mz 688.4916	-	Epithelium	PE(16:0/16:1) or PE(16:1/16:0)	M-H	Glycerophospholipids	Glycerophosphoethanolamines	2
Colon	HLA-	mz 690.5064	-	Epithelium	PE(16:0/16:0)	M-H	Glycerophospholipids	Glycerophosphoethanolamines	2

	B27+								
Cecum	WT	mz 698.5158	-	Lumen					4
Colon	HLA-B27+	mz 702.5424	-	Epithelium	PE(O-18:0/16:1)	M-H	Glycerophospholipids	Glycerophosphoethanolamines	2
Colon	HLA-B27+	mz 718.5367	-	Patches&Submucosa	PE(16:0/18:0) or PE(18:0/16:0)	M-H	Glycerophospholipids	Glycerophosphoethanolamines	2
Cecum	HLA-B27+	mz 720.5905	+	Lamina propria	PC(O-32:0)	M+H	Glycerophospholipids	Glycerophosphocholines	2
Colon	HLA-B27+	mz 720.5926	+	Lamina Propria	PC(O-32:0)	M+H	Glycerophospholipids	Glycerophosphocholines	3
Colon	HLA-B27+	mz 721.5943	+	Patches&Lamina Propria					4
Cecum	WT	mz 726.5475	-	Lumen	PE(P-18:0/18:2)	M-H	Glycerophospholipids	Glycerophosphoethanolamines	2
Colon	HLA-B27+	mz 730.5732	-	Epithelium	PE(O-18:0/18:1)	M-H	Glycerophospholipids	Glycerophosphoethanolamines	2
Cecum	HLA-B27+	mz 734.5261	+	Patches&Mucosa					4
Cecum	HLA-B27+	mz 734.5686	+	Mucosa&Submucosa	PC(32:0)	M+H	Glycerophospholipids	Glycerophosphocholines	2
Colon	WT	mz 736.5274	-	Mucosa	PE(20:4/16:1) or PE(16:1/20:4)	M-H	Glycerophospholipids	Glycerophosphocholines	2
Colon	HLA-B27+	mz 744.5035	+	Epithelium	PE(O-38:9)	M+H	Glycerophospholipids	Glycerophosphoethanolamines	3
Colon	HLA-B27+	mz 746.6056	+	Patches&Lamina Propria	PC(O-34:1)	M+H	Glycerophospholipids	Glycerophosphocholines	2
Colon	WT	mz 753.5934	+	Submucosa	SM(36:1)	M+Na	Sphingolipids	Sphingomyelins	3
Colon	HLA-B27+	mz 760.5135	-	Mucosa	PS(34:1)	M-H	Glycerophospholipids	Glycerophosphoserines	2
Cecum	HLA-B27+	mz 763.5096	-	Muscularis propria					4
Cecum	HLA-B27+	mz 763.5964	+	Mucosa	SM(d18:0/18:1(12Z)-2OH(9,10))	M+H	Sphingolipids	Sphingomyelins	2
Cecum	HLA-B27+	mz 766.6001	-	Muscularis propria	PE(18:0/20:4) or PE(20:4/18:0)	M-H	Glycerophospholipids	Glycerophosphoethanolamines	2

Colon	HLA-B27+	mz 768.5853	+	Lamina Propria	PC(P-36:3) or PC(O-36:4)	M+H	Glycerophospholipids	Glycerophosphocholines	2
Cecum	HLA-B27+	mz 768.5894	+	Lamina Propria	PC(P-36:3) or PC(O-36:4)	M+H	Glycerophospholipids	Glycerophosphocholines	2
Cecum	HLA-B27+	mz 771.5169	-	Mucosa	PG(36:3)	M-H	Glycerophospholipids	Glycerophosphoglycerols	3
Cecum	HLA-B27+	mz 772.5004	+	Patches&Mucosa					4
Cecum	HLA-B27+	mz 772.5221	+	Patches&Mucosa					4
Colon	HLA-B27+	mz 772.5229	+	Patches&Mucosa					4
Colon	WT	mz 774.6061	+	Mucosa	PC(35:1)	M+H	Glycerophospholipids	Glycerophosphocholines	2
Colon	WT	mz 781.6169	+	Submucosa	SM(38:1)	M+Na	Sphingolipids	Sphingomyelins	3
Cecum	HLA-B27+	mz 782.4936	-	Patches	PS(36:4)	M-H	Glycerophospholipids	Glycerophosphoserines	2
Colon	WT	mz 787.6646	+	Submucosa	SM(40:1)	M+H	Sphingolipids	Sphingomyelins	2
Colon	WT	mz 788.6263	-	Lumen&Epithelium					4
Cecum	WT	mz 788.6263	-	Lumen&Epithelium					4
Colon	HLA-B27+	mz 789.4888	+	Lamina Propria					4
Cecum	HLA-B27+	mz 790.5716	+	Submucosa&Lamina Proproa	PC(O-38:7) or PC(P-38:6)	M+H	Glycerophospholipids	Glycerophosphocholines	2
Cecum	HLA-B27+	mz 792.582	+	Submucosa&Lamina Proproa	PC(O-38:6) or PC(P-38:5)	M+H	Glycerophospholipids	Glycerophosphocholines	2
Cecum	HLA-B27+	mz 794.5854	+	Submucosa&Lamina Proproa	PC(37:5)	M+H	Glycerophospholipids	Glycerophosphocholines	2
Cecum	HLA-B27+	mz 794.6085	+	Submucosa&Lamina Proproa	PC(O-38:5)	M+H	Glycerophospholipids	Glycerophosphocholines	2
Colon	HLA-B27+	mz 794.6114	+	Patches&Lamina Propria					4
Cecum	HLA-B27+	mz 795.6123	+	Epithelium					4

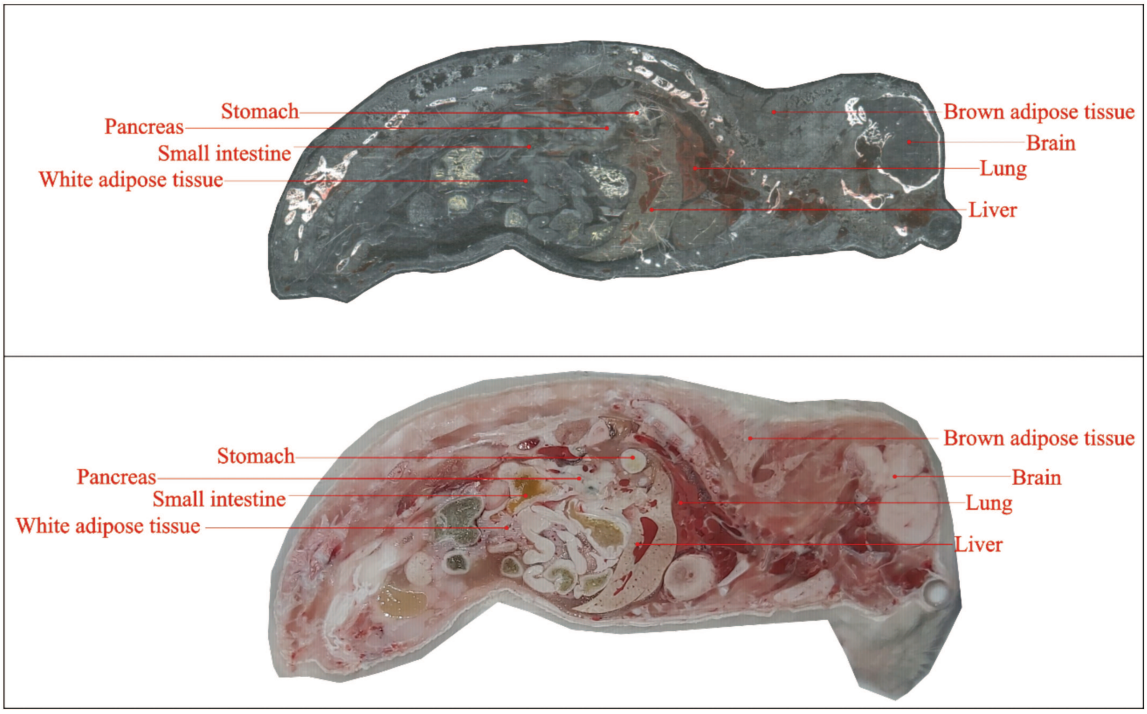
Cecum	HLA-B27+	mz 795.6382	-	Submucosa					4
Colon	WT	mz 796.5762	+	Lumen					4
Colon	HLA-B27+	mz 796.6133	+	Patches&Lamina Propria					4
Cecum	HLA-B27+	mz 806.5136	+	Patches&Mucosa	PS(38:7)	M+H	Glycerophospholipids	Glycerophosphoserines	2
Cecum	HLA-B27+	mz 808.5792	+	Mucosa&Submucosa	PC(38:5)	M+H	Glycerophospholipids	Glycerophosphocholines	2
Colon	WT	mz 809.6496	+	Submucosa	SM(40:1)	M+Na	Sphingolipids	Sphingomyelins	3
Cecum	HLA-B27+	mz 810.6056	+	Mucosa&Submucosa	PC(38:4)	M+H	Glycerophospholipids	Glycerophosphocholines	2
Cecum	HLA-B27+	mz 811.5352	-	Patches&Submucosa					4
Cecum	HLA-B27+	mz 811.6128	+	Mucosa&Submucosa					4
Colon	WT	mz 812.5289	+	Lumen	PE(42:10)	M+H	Glycerophospholipids	Glycerophosphoethanolamines	3
Colon	WT	mz 812.5624	+	Lumen					4
Colon	WT	mz 815.6869	+	Submucosa	SM(42:1)	M+H	Sphingolipids	Sphingomyelins	3
Cecum	HLA-B27+	mz 816.5912	+	Patches&Lamina Propria	PC(P-40:7)	M+H	Glycerophospholipids	Glycerophosphocholines	2
Colon	HLA-B27+	mz 816.5928	+	Patches&Lamina Propria	PC(P-40:7)	M+H	Glycerophospholipids	Glycerophosphocholines	2
Cecum	WT	mz 816.6537	-	Mucosa&Lumen					4
Cecum	HLA-B27+	mz 817.5041	-	Lamina propria	PG(18:2/22:6) or PG(22:6/18:2)	M-H	Glycerophospholipids	Glycerophosphoglycerols	2
Cecum	HLA-B27+	mz 820.5224	+	Mucosa&Submucosa					4
Cecum	HLA-B27+	mz 830.5696	+	Submucosa	PC(40:8)	M+H	Glycerophospholipids	Glycerophosphocholines	2
Cecum	HLA-B27+	mz 831.5664	-	Patches&Submucosa					4

Cecum	HLA-B27+	mz 832.568	+	Submucosa					4
Cecum	HLA-B27+	mz 836.6208	+	Patches&Mucosa	PC(40:5)	M+H	Glycerophospholipids	Glycerophosphocholines	2
Colon	WT	mz 837.6888	+	Muscularis Propria&Submucosa					4
Cecum	WT	mz 844.6905	-	Mucosa&Lumen					4
Colon	WT	mz 844.6937	+	Lumen					4
Cecum	HLA-B27+	mz 846.5367	+	Patches&Mucosa					4
Cecum	HLA-B27+	mz 848.5264	+	Mucosa&Submucosa					4
Cecum	HLA-B27+	mz 848.5512	+	Mucosa&Submucosa					4
Colon	WT	mz 854.7281	+	Lumen					4
Colon	WT	mz 856.6236	+	Lumen					4
Colon	WT	mz 866.68	+	Mucosa					4
Colon	HLA-B27+	mz 870.5848	+	Patches&Mucosa					4
Colon	HLA-B27+	mz 872.5564	+	Patches					4
Colon	HLA-B27+	mz 875.5801	+	Patches&Mucosa					4
Colon	WT	mz 882.6329	+	Lumen&Mucosa					4
Cecum	HLA-B27+	mz 884.5344	-	Patches					4
Colon	WT	mz 884.668	+	Lumen					4
Cecum	HLA-B27+	mz 885.3328	-	Patches					4
Cecum	HLA-B27+	mz 885.4264	-	Patches					4
Cecum	WT	mz 902.4867	-	Submucosa&Lumen					4

Cecum	HLA-B27+	mz 909.5481	+	Submucosa	PI(40:7)	M+H	Glycerophospholipids	Glycerophosphoinositols	3
Cecum	HLA-B27+	mz 912.58	-	Patches&Submucosa					4
Cecum	WT	mz 924.6349	-	Lumen					4
Cecum	HLA-B27+	mz 925.5229	+	Patches&Mucosa					4
Cecum	HLA-B27+	mz 933.5927	-	Submucosa					4

7.3 Annex 3

Comparison between whole-body scan and photograph of the carcass in the same sectional plane was performed to annotate the regions where metabolites are distributed as the small intestine.



8 Bibliography

Ackermann, B. L., Berna, M. J., & Murphy, A. T. (2002). Recent Advances in use of LC/MS/MS for Quantitative High-Throughput Bioanalytical Support of Drug Discovery. *Current Topics in Medicinal Chemistry*, *2*(1), 53–66. https://doi.org/10.2174/1568026023394605

Ain, Q. ul, Batool, M., & Choi, S. (2020). TLR4-Targeting Therapeutics: Structural Basis and Computer-Aided Drug Discovery Approaches. *Molecules*, *25*(3), 3. https://doi.org/10.3390/molecules25030627

Alexandrov, T. (2012). MALDI imaging mass spectrometry: Statistical data analysis and current computational challenges. *BMC Bioinformatics*, *13*(Suppl 16), S11. https://doi.org/10.1186/1471-2105-13-S16-S11

Alexandrov, T., Ovchinnikova, K., Palmer, A., Kovalev, V., Tarasov, A., Stuart, L., Nigmetzianov, R., & Fay, D. (2019). METASPACE: A community-populated knowledge base of spatial metabolomes in health and disease. *BioRxiv*. https://doi.org/10.1101/539478

Aristoteli, L. P., Molloy, M. P., & Baker, M. S. (2007). Evaluation of Endogenous Plasma Peptide Extraction Methods for Mass Spectrometric Biomarker Discovery. *Journal of Proteome Research*, *6*(2), 571–581. https://doi.org/10.1021/pr0602996

Asquith, M., Stauffer, P., Davin, S., Mitchell, C., Lin, P., & Rosenbaum, J. T. (2016). Perturbed Mucosal Immunity and Dysbiosis Accompany Clinical Disease in a Rat Model of Spondyloarthritis. *Arthritis & Rheumatology (Hoboken, N.J.)*, *68*(9), 2151–2162. https://doi.org/10.1002/art.39681

Baeten, D., Sieper, J., Braun, J., Baraliakos, X., Dougados, M., Emery, P., Deodhar, A., Porter, B., Martin, R., Andersson, M., Mpofu, S., & Richards, H. B. (2015). Secukinumab, an Interleukin-17A Inhibitor, in Ankylosing Spondylitis. *New England Journal of Medicine*, *373*(26), 2534–2548. https://doi.org/10.1056/NEJMoa1505066

Baggia, S., Lyons, J. L., Angell, E., Barkhuizen, A., Han, Y. B., Planck, S. R., Taurog, J. D., & Rosenbaum, J. T. (1997). A novel model of bacterially-induced acute anterior uveitis in rats

and the lack of effect from HLA-B27 expression. *Journal of Investigative Medicine: The Official Publication of the American Federation for Clinical Research*, *45*(5), 295–301.

Balluff, B., Hopf, C., Porta Siegel, T., Grabsch, H. I., & Heeren, R. M. A. (2021). Batch Effects in MALDI Mass Spectrometry Imaging. *Journal of the American Society for Mass Spectrometry*, *32*(3), 628–635. https://doi.org/10.1021/jasms.0c00393

Bano, I., Butt, U. D., & Mohsan, S. A. H. (2023). Chapter 25—New challenges in drug discovery. In S. Das, S. Thomas, & P. P. Das (Eds.), *Novel Platforms for Drug Delivery Applications* (pp. 619–643). Woodhead Publishing. https://doi.org/10.1016/B978-0-323-91376-8.00021-5

Bauset, C., Gisbert-Ferrándiz, L., & Cosín-Roger, J. (2021). Metabolomics as a Promising Resource Identifying Potential Biomarkers for Inflammatory Bowel Disease. *Journal of Clinical Medicine*, *10*(4), 4. https://doi.org/10.3390/jcm10040622

Bender, K. J., Wang, Y., Zhai, C. Y., Saenz, Z., Wang, A., & Neumann, E. K. (2023). Sample Preparation Method for MALDI Mass Spectrometry Imaging of Fresh-Frozen Spines. *Analytical Chemistry*, *95*(47), 17337–17346. https://doi.org/10.1021/acs.analchem.3c03672

Benjamini, Y., & Hochberg, Y. (1995). Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society: Series B (Methodological)*, *57*(1), 289–300. https://doi.org/10.1111/j.2517-6161.1995.tb02031.x

Bock, J., Liebisch, G., Schweimer, J., Schmitz, G., & Rogler, G. (2007). Exogenous sphingomyelinase causes impaired intestinal epithelial barrier function. *World Journal of Gastroenterology : WJG*, *13*(39), 5217–5225. https://doi.org/10.3748/wjg.v13.i39.5217

Boel, A., López-Medina, C., van der Heijde, D. M. F. M., & van Gaalen, F. A. (2022). Age at onset in axial spondyloarthritis around the world: Data from the Assessment in SpondyloArthritis international Society Peripheral Involvement in Spondyloarthritis study. *Rheumatology*, *61*(4), 1468–1475. https://doi.org/10.1093/rheumatology/keab544

Bowie, J. H., Brinkworth, C. S., & Dua, S. (2002). Collision-induced fragmentations of the (M-H)− parent anions of underivatized peptides: An aid to structure determination and some unusual

negative ion cleavages. *Mass Spectrometry Reviews*, *21*(2), 87–107. https://doi.org/10.1002/mas.10022

Box, G. E. P., & Draper, N. R. (1987). *Empirical model-building and response surfaces* (pp. xiv, 669). John Wiley & Sons.

Braun, J., Baraliakos, X., Brandt, J., & Sieper, J. (2005). Therapy of ankylosing spondylitis. Part II: Biological therapies in the spondyloarthritides: Scandinavian Journal of Rheumatology. *Scandinavian Journal of Rheumatology*, *34*(3), 178–190. https://doi.org/10.1080/03009740510026599

Braun, J., & Sieper, J. (2002). Therapy of ankylosing spondylitis and other spondyloarthritides: Established medical treatment, anti-TNF-α therapy and other novel approaches. *Arthritis Research & Therapy*, *4*(5), 307. https://doi.org/10.1186/ar592

Brewerton, D. A., Hart, F. D., Nicholls, A., Caffrey, M., James, D. C. O., & Sturrock, R. D. (1973). ANKYLOSING SPONDYLITIS AND HL-A 27. *The Lancet*, *301*(7809), 904–907. https://doi.org/10.1016/S0140-6736(73)91360-3

Camberlein, V., Fléau, C., Sierocki, P., Li, L., Gealageas, R., Bosc, D., Guillaume, V., Warenghem, S., Leroux, F., Rosell, M., Cheng, K., Medve, L., Prigent, M., Decanter, M., Piveteau, C., Biela, A., Eveque, M., Dumont, J., Mpakali, A., … Deprez-Poulain, R. (2022). Discovery of the First Selective Nanomolar Inhibitors of ERAP2 by Kinetic Target-Guided Synthesis. *Angewandte Chemie (International Ed. in English)*, *61*(39), e202203560. https://doi.org/10.1002/anie.202203560

Caprioli, R. M., Farmer, T. B., & Gile, J. (1997). Molecular Imaging of Biological Samples: Localization of Peptides and Proteins Using MALDI-TOF MS. *Analytical Chemistry*, *69*(23), 4751–4760. https://doi.org/10.1021/ac970888i

Chambers, E. E. (2013). Quantitative Analysis of Therapeutic and Endogenous Peptides using LC/MS/MS Methods. In G. Chen (Ed.), *Characterization of Protein Therapeutics using Mass Spectrometry* (pp. 59–94). Springer US. https://doi.org/10.1007/978-1-4419-7862-2_2

Chen, B., Li, J., He, C., Li, D., Tong, W., Zou, Y., & Xu, W. (2017). Role of HLA-B27 in the pathogenesis of ankylosing spondylitis. *Molecular Medicine Reports*, *15*(4), 1943–1951. https://doi.org/10.3892/mmr.2017.6248

Chen, W., Hwang, Y. Y., Gleaton, J. W., Titus, J. K., & Hamlin, N. J. (2019). Optimization of a Peptide Extraction and LC–MS Protocol for Quantitative Analysis of Antimicrobial Peptides. *Future Science OA*, *5*(1), FSO348. https://doi.org/10.4155/fsoa-2018-0073

Chiaruttini, N., Burri, O., Haub, P., Guiet, R., Sordet-Dessimoz, J., & Seitz, A. (2022). An Open-Source Whole Slide Image Registration Workflow at Cellular Precision Using Fiji, QuPath and Elastix. *Frontiers in Computer Science*, *3*. https://doi.org/10.3389/fcomp.2021.780026

Cifaldi, L., Romania, P., Lorenzi, S., Locatelli, F., & Fruci, D. (2012). Role of Endoplasmic Reticulum Aminopeptidases in Health and Disease: From Infection to Cancer. *International Journal of Molecular Sciences*, *13*(7), 8338–8352. https://doi.org/10.3390/ijms13078338

Compagnone, M., Cifaldi, L., & Fruci, D. (2019). Regulation of ERAP1 and ERAP2 genes and their disfunction in human cancer. *Human Immunology*, *80*(5), 318–324. https://doi.org/10.1016/j.humimm.2019.02.014

Danve, A., & O'Dell, J. (2015). The ongoing quest for biomarkers in Ankylosing Spondylitis. *International Journal of Rheumatic Diseases*, *18*(8), 826–834. https://doi.org/10.1111/1756-185X.12779

Declercq, A., Bouwmeester, R., Chiva, C., Sabidó, E., Hirschler, A., Carapito, C., Martens, L., Degroeve, S., & Gabriels, R. (2023). Updated MS^2PIP web server supports cutting-edge proteomics applications. *Nucleic Acids Research*, *51*(W1), W338–W342. https://doi.org/10.1093/nar/gkad335

DeMarco, V. G., Li, N., Thomas, J., Neu, J., & West, C. M. (2003). Glutamine and Barrier Function in Cultured Caco-2 Epithelial Cell Monolayers. *The Journal of Nutrition*, *133*(7), 2176–2179. https://doi.org/10.1093/jn/133.7.2176

Dougados, M., & Baeten, D. (2011). Spondyloarthritis. *The Lancet*, *377*, 2127–2137. https://doi.org/10.1016/S0140-6736(11)60071-8

Drews, J. (2000). Drug Discovery: A Historical Perspective. *Science*, *287*(5460), 1960–1964. https://doi.org/10.1126/science.287.5460.1960

El Maghraoui, A. (2011). Extra-articular manifestations of ankylosing spondylitis: Prevalence, characteristics and therapeutic implications. *European Journal of Internal Medicine*, *22*(6), 554–560. https://doi.org/10.1016/j.ejim.2011.06.006

Evans, D. M., Spencer, C. C. A., Pointon, J. J., Su, Z., Harvey, D., Kochan, G., Oppermann, U., Dilthey, A., Pirinen, M., Stone, M. A., Appleton, L., Moutsianas, L., Leslie, S., Wordsworth, T., Kenna, T. J., Karaderi, T., Thomas, G. P., Ward, M. M., Weisman, M. H., … Donnelly, P. (2011). Interaction between ERAP1 and HLA-B27 in ankylosing spondylitis implicates peptide handling in the mechanism for HLA-B27 in disease susceptibility. *Nature Genetics*, *43*(8), 761–767. https://doi.org/10.1038/ng.873

FDA-NIH Biomarker Working Group. (2016). *BEST (Biomarkers, EndpointS, and other Tools) Resource*. Food and Drug Administration (US). http://www.ncbi.nlm.nih.gov/books/NBK326791/

Feller, L., & Lemmer, J. (2012). *Oral Squamous Cell Carcinoma: Epidemiology, Clinical Presentation and Treatment*. *2012*. https://doi.org/10.4236/jct.2012.34037

Fougiaxis, V., He, B., Khan, T., Vatinel, R., Koutroumpa, N. M., Afantitis, A., Lesire, L., Sierocki, P., Deprez, B., & Deprez-Poulain, R. (2024). ERAP Inhibitors in Autoimmunity and Immuno-Oncology: Medicinal Chemistry Insights. *Journal of Medicinal Chemistry*, *67*(14), 11597–11621. https://doi.org/10.1021/acs.jmedchem.4c00840

Franke, A., McGovern, D. P. B., Barrett, J. C., Wang, K., Radford-Smith, G. L., Ahmad, T., Lees, C. W., Balschun, T., Lee, J., Roberts, R., Anderson, C. A., Bis, J. C., Bumpstead, S., Ellinghaus, D., Festen, E. M., Georges, M., Green, T., Haritunians, T., Jostins, L., … Parkes, M. (2010). Genome-wide meta-analysis increases to 71 the number of confirmed Crohn's disease susceptibility loci. *Nature Genetics*, *42*(12), 1118–1125. https://doi.org/10.1038/ng.717

George, A. L., Foreman, R. E., Sayda, M. H., Reimann, F., Gribble, F. M., & Kay, R. G. (2023). Rapid and Quantitative Enrichment of Peptides from Plasma for Mass Spectrometric Analysis.

Methods in Molecular Biology (Clifton, N.J.), *2628*, 477–488. https://doi.org/10.1007/978-1-0716-2978-9_28

Gessel, M. M., Norris, J. L., & Caprioli, R. M. (2014). MALDI imaging mass spectrometry: Spatial molecular analysis to enable a new age of discovery. *JOURNAL OF PROTEOMICS*. http://dx.doi.org/10.1016/j.jprot.2014.03.021

Good, C. J., Neumann, E. K., Butrico, C. E., Cassat, J. E., Caprioli, R. M., & Spraggins, J. M. (2022). High Spatial Resolution MALDI Imaging Mass Spectrometry of Fresh-Frozen Bone. *Analytical Chemistry*, *94*(7), 3165–3172. https://doi.org/10.1021/acs.analchem.1c04604

Goodwin, R. J. A. (2012). Sample preparation for mass spectrometry imaging: Small mistakes can lead to big consequences. *Journal of Proteomics*, *75*(16), 4893–4911. https://doi.org/10.1016/j.jprot.2012.04.012

Granborg, J. R., Handler, A. M., & Janfelt, C. (2022). Mass spectrometry imaging in drug distribution and drug metabolism studies – Principles, applications and perspectives. *TrAC Trends in Analytical Chemistry*, *146*, 116482. https://doi.org/10.1016/j.trac.2021.116482

Guasp, P., Lorente, E., Martín-Esteban, A., Barnea, E., Romania, P., Fruci, D., Kuiper, J. W., Admon, A., & Castro, J. A. L. de. (2019). Redundancy and Complementarity between ERAP1 and ERAP2 Revealed by their Effects on the Behcet's Disease-associated HLA-B*51 Peptidome *[S]. *Molecular & Cellular Proteomics*, *18*(8), 1491–1510. https://doi.org/10.1074/mcp.RA119.001515

Hamilton, F., Mentzer, A. J., Parks, T., Baillie, J. K., Smith, G. D., Ghazal, P., & Timpson, N. J. (2023). Variation in ERAP2 has opposing effects on severe respiratory infection and autoimmune disease. *American Journal of Human Genetics*, *110*(4), 691–702. https://doi.org/10.1016/j.ajhg.2023.02.008

Hammer, G. E., Gonzalez, F., Champsaur, M., Cado, D., & Shastri, N. (2006). The aminopeptidase ERAAP shapes the peptide repertoire displayed by major histocompatibility complex class I molecules. *Nature Immunology*, *7*(1), 103–112. https://doi.org/10.1038/ni1286

Hammer, R. E., Maika, S. D., Richardson, J. A., Tang, J.-P., & Taurog, J. D. (1990). Spontaneous inflammatory disease in transgenic rats expressing HLA-B27 and human *β*2m: An animal model of HLA-B27-associated human disorders. *Cell*, *63*(5), 1099–1112. https://doi.org/10.1016/0092-8674(90)90512-D

Harkin, C., Smith, K. W., Cruickshank, F. L., Logan Mackay, C., Flinders, B., Heeren, R. M. A., Moore, T., Brockbank, S., & Cobice, D. F. (2022). On-tissue chemical derivatization in mass spectrometry imaging. *Mass Spectrometry Reviews*, *41*(5), 662–694. https://doi.org/10.1002/mas.21680

Haskó, G., Kuhel, D. G., Németh, Z. H., Mabley, J. G., Stachlewitz, R. F., Virág, L., Lohinai, Z., Southan, G. J., Salzman, A. L., & Szabó, C. (2000). Inosine Inhibits Inflammatory Cytokine Production by a Posttranscriptional Mechanism and Protects Against Endotoxin-Induced Shock1. *The Journal of Immunology*, *164*(2), 1013–1019. https://doi.org/10.4049/jimmunol.164.2.1013

Ho, S., & Gao, H. (2015). Surrogate Matrix: Opportunities and Challenges for Tissue Sample Analysis. *Bioanalysis*, *7*(18), 2419–2433. https://doi.org/10.4155/bio.15.161

Hochart, G., Hamm, G., & Stauber, J. (2014). Label-free MS imaging from drug discovery to preclinical development. *Bioanalysis*, *6*(20), 2775–2788. https://doi.org/10.4155/bio.14.202

Huang, T., Pu, Y., Wang, X., Li, Y., Yang, H., Luo, Y., & Liu, Y. (2022). Metabolomic analysis in spondyloarthritis: A systematic review. *Frontiers in Microbiology*, *13*. https://www.frontiersin.org/articles/10.3389/fmicb.2022.965709

Hughes, J., Rees, S., Kalindjian, S., & Philpott, K. (2011). Principles of early drug discovery. *British Journal of Pharmacology*, *162*(6), 1239–1249. https://doi.org/10.1111/j.1476-5381.2010.01127.x

James, E., Bailey, I., Sugiyarto, G., & Elliott, T. (2013). Induction of Protective Antitumor Immunity through Attenuation of ERAAP Function. *The Journal of Immunology*, *190*(11), 5839–5846. https://doi.org/10.4049/jimmunol.1300220

James, E., Yeh, A., King, C., Korangy, F., Bailey, I., Boulanger, D. S., Van den Eynde, B. J., Murray, N., & Elliott, T. J. (2010). Differential Suppression of Tumor-Specific CD8+ T Cells by Regulatory T Cells. *The Journal of Immunology*, *185*(9), 5048–5055. https://doi.org/10.4049/jimmunol.1000134

Jones, B. R., Schultz, G. A., Eckstein, J. A., & Ackermann, B. L. (2012). Surrogate Matrix And Surrogate Analyte Approaches For Definitive Quantitation of Endogenous Biomolecules. *Bioanalysis*, *4*(19), 2343–2356. https://doi.org/10.4155/bio.12.200

Kappeler, A., & Mueller, C. (2020). The role of activated cytotoxic T cells in inflammatory bowel disease. *Histology and Histopathology*, *15*, 167–172. https://doi.org/001: 10.14670/HH-15.167

Kawashima, Y., Fukutomi, T., Tomonaga, T., Takahashi, H., Nomura, F., Maeda, T., & Kodera, Y. (2010). High-Yield Peptide-Extraction Method for the Discovery of Subnanomolar Biomarkers from Small Serum Samples. *Journal of Proteome Research*, *9*(4), 1694–1705. https://doi.org/10.1021/pr9008018

Kay, R. G., Challis, B. G., Casey, R. T., Roberts, G. P., Meek, C. L., Reimann, F., & Gribble, F. M. (2018). Peptidomic analysis of endogenous plasma peptides from patients with pancreatic neuroendocrine tumours. *Rapid Communications in Mass Spectrometry*, *32*(16), 1414–1424. https://doi.org/10.1002/rcm.8183

Kerr, S. W., Wolyniec, W. W., Filipovic, Z., Nodop, S. G., Braza, F., Winquist, R. J., & Noonan, T. C. (1999). Repeated Measurement of Intestinal Permeability as an Assessment of Colitis Severity in HLA-B27 Transgenic Rats. *Journal of Pharmacology and Experimental Therapeutics*, *291*(2), 903–910.

Kirino, Y., Bertsias, G., Ishigatsubo, Y., Mizuki, N., Tugal-Tutkun, I., Seyahi, E., Ozyazgan, Y., Sacli, F. S., Erer, B., Inoko, H., Emrence, Z., Cakar, A., Abaci, N., Ustek, D., Satorius, C., Ueda, A., Takeno, M., Kim, Y., Wood, G. M., … Kastner, D. L. (2013). Genome-wide association analysis identifies new susceptibility loci for Behçet's disease and epistasis between HLA-B*51 and ERAP1. *Nature Genetics*, *45*(2), 202–207. https://doi.org/10.1038/ng.2520

Kirkali, Z., Chan, T., Manoharan, M., Algaba, F., Busch, C., Cheng, L., Kiemeney, L., Kriegmair, M., Montironi, R., Murphy, W. M., Sesterhenn, I. A., Tachibana, M., & Weider, J. (2005). Bladder cancer: Epidemiology, staging and grading, and diagnosis. *Urology*, *66*(6, Supplement 1), 4–34. https://doi.org/10.1016/j.urology.2005.07.062

Klingberg, E., Strid, H., Ståhl, A., Deminger, A., Carlsten, H., Öhman, L., & Forsblad-d'Elia, H. (2017). A longitudinal study of fecal calprotectin and the development of inflammatory bowel disease in ankylosing spondylitis. *Arthritis Research & Therapy*, *19*, 21. https://doi.org/10.1186/s13075-017-1223-2

Korfmacher, W. A. (2005). Foundation review: Principles and applications of LC-MS in new drug discovery. *Drug Discovery Today*, *10*(20), 1357–1367. https://doi.org/10.1016/S1359-6446(05)03620-2

Kuiper, J. J. W., van Setten, J., Devall, M., Cretu-Stancu, M., Hiddingh, S., Ophoff, R. A., Missotten, T. O. A. R., van Velthoven, M., Den Hollander, A. I., Hoyng, C. B., James, E., Reeves, E., Cordero-Coma, M., Fonollosa, A., Adán, A., Martín, J., Koeleman, B. P. C., de Boer, J. H., Pulit, S. L., … Radstake, T. R. D. J. (2018). Functionally distinct ERAP1 and ERAP2 are a hallmark of HLA-A29-(Birdshot) Uveitis. *Human Molecular Genetics*, *27*(24), 4333–4343. https://doi.org/10.1093/hmg/ddy319

Kumar, A. R., Devan, A. R., Nair, B., Vinod, B. S., & Nath, L. R. (2021). Harnessing the immune system against cancer: Current immunotherapy approaches and therapeutic targets. *Molecular Biology Reports*, *48*(12), 8075. https://doi.org/10.1007/s11033-021-06752-9

Kumar, D., Gautam, N., & Alnouti, Y. (2022). Analyte recovery in LC-MS/MS bioanalysis: An old issue revisited. *Analytica Chimica Acta*, *1198*, 339512. https://doi.org/10.1016/j.aca.2022.339512

Kuo, I.-C., Kao, H.-K., Huang, Y., Wang, C.-I., Yi, J.-S., Liang, Y., Liao, C.-T., Yen, T.-C., Wu, C.-C., & Chang, K.-P. (2017). Endoplasmic reticulum aminopeptidase 2 involvement in metastasis of oral cavity squamous cell carcinoma discovered by proteome profiling of

primary cancer cells. *Oncotarget*, *8*(37), 61698–61708. https://doi.org/10.18632/oncotarget.18680

Li, N., Lewis, P., Samuelson, D., Liboni, K., & Neu, J. (2004). Glutamine regulates Caco-2 cell tight junction proteins. *American Journal of Physiology-Gastrointestinal and Liver Physiology*, *287*(3), G726–G733. https://doi.org/10.1152/ajpgi.00012.2004

Lim, Y. W., Chen-Harris, H., Mayba, O., Lianoglou, S., Wuster, A., Bhangale, T., Khan, Z., Mariathasan, S., Daemen, A., Reeder, J., Haverty, P. M., Forrest, W. F., Brauer, M., Mellman, I., & Albert, M. L. (2018). Germline genetic polymorphisms influence tumor gene expression and immune cell infiltration. *Proceedings of the National Academy of Sciences*, *115*(50), E11701–E11710. https://doi.org/10.1073/pnas.1804506115

Liu, P., Zhu, W., Chen, C., Yan, B., Zhu, L., Chen, X., & Peng, C. (2020). The mechanisms of lysophosphatidylcholine in the development of diseases. *Life Sciences*, *247*, 117443. https://doi.org/10.1016/j.lfs.2020.117443

Lorente, E., Fontela, M. G., Barnea, E., Martín-Galiano, A. J., Mir, C., Galocha, B., Admon, A., Lauzurica, P., & López, D. (2020). Modulation of Natural HLA-B*27:05 Ligandome by Ankylosing Spondylitis-associated Endoplasmic Reticulum Aminopeptidase 2 (ERAP2) *. *Molecular & Cellular Proteomics*, *19*(6), 994–1004. https://doi.org/10.1074/mcp.RA120.002014

Ma, X., & Fernández, F. M. (2024). Advances in mass spectrometry imaging for spatial cancer metabolomics. *Mass Spectrometry Reviews*, *43*(2), 235–268. https://doi.org/10.1002/mas.21804

Mabley, J. G., Pacher, P., Liaudet, L., Soriano, F. G., Haskó, G., Marton, A., Szabó, C., & Salzman, A. L. (2003). Inosine reduces inflammation and improves survival in a murine model of colitis. *American Journal of Physiology-Gastrointestinal and Liver Physiology*, *284*(1), G138–G144. https://doi.org/10.1152/ajpgi.00060.2002

Mahida, Y. R. (2000). The Key Role of Macrophages in the Immunopathogenesis of Inflammatory Bowel Disease. *Inflammatory Bowel Diseases*, *6*(1), 21–33. https://doi.org/10.1097/00054725-200002000-00004

Makurvet, F. D. (2021). Biologics vs. small molecules: Drug costs and patient access. *Medicine in Drug Discovery*, *9*, 100075. https://doi.org/10.1016/j.medidd.2020.100075

Matuszewski, B. K., Constanzer, M. L., & Chavez-Eng, C. M. (2003). Strategies for the Assessment of Matrix Effect in Quantitative Bioanalytical Methods Based on HPLC−MS/MS. *Analytical Chemistry*, *75*(13), 3019–3030. https://doi.org/10.1021/ac020361s

McEwen, A., & Henson, C. (2015). Quantitative Whole-Body Autoradiography: Past, Present and Future. *Bioanalysis*, *7*(5), 557–568. https://doi.org/10.4155/bio.15.9

McInnes, L., Healy, J., Saul, N., & Großberger, L. (2018). UMAP: Uniform Manifold Approximation and Projection. *Journal of Open Source Software*, *3*(29), 861. https://doi.org/10.21105/joss.00861

Merdas, M., Lagarrigue, M., Vanbellingen, Q., Umbdenstock, T., Da Violante, G., & Pineau, C. (2021). On-tissue chemical derivatization reagents for matrix-assisted laser desorption/ionization mass spectrometry imaging. *Journal of Mass Spectrometry: JMS*, *56*(10), e4731. https://doi.org/10.1002/jms.4731

Michielan, A., & D'Incà, R. (2015). Intestinal Permeability in Inflammatory Bowel Disease: Pathogenesis, Clinical Evaluation, and Therapy of Leaky Gut. *Mediators of Inflammation*, *2015*, 628157. https://doi.org/10.1155/2015/628157

Miedzybrodzka, E. L., Foreman, R. E., Galvin, S. G., Larraufie, P., George, A. L., Goldspink, D. A., Reimann, F., Gribble, F. M., & Kay, R. G. (2020). Organoid Sample Preparation and Extraction for LC-MS Peptidomics. *STAR Protocols*, *1*(3), 100164. https://doi.org/10.1016/j.xpro.2020.100164

Milia, A. F., Ibba-Manneschi, L., Manetti, M., Benelli, G., Messerini, L., & Matucci-Cerinic, M. (2009). HLA-B27 Transgenic Rat. *Annals of the New York Academy of Sciences*, *1173*(1), 570–574. https://doi.org/10.1111/j.1749-6632.2009.04757.x

Miller, R., Ewy, W., Corrigan, B. W., Ouellet, D., Hermann, D., Kowalski, K. G., Lockwood, P., Koup, J. R., Donevan, S., El-Kattan, A., Li, C. S. W., Werth, J. L., Feltner, D. E., & Lalonde, R. L. (2005). How modeling and simulation have enhanced decision making in new drug development. *Journal of Pharmacokinetics and Pharmacodynamics*, *32*(2), 185–197. https://doi.org/10.1007/s10928-005-0074-7

Minos, E., Barry, R. J., Southworth, S., Folkard, A., Murray, P. I., Duker, J. S., Keane, P. A., & Denniston, A. K. (2016). Birdshot chorioretinopathy: Current knowledge and new concepts in pathophysiology, diagnosis, monitoring and treatment. *Orphanet Journal of Rare Diseases*, *11*(1), 61. https://doi.org/10.1186/s13023-016-0429-8

Mörbe, U. M., Jørgensen, P. B., Fenton, T. M., von Burg, N., Riis, L. B., Spencer, J., & Agace, W. W. (2021). Human gut-associated lymphoid tissues (GALT); diversity, structure, and function. *Mucosal Immunology*, *14*(4), 793–802. https://doi.org/10.1038/s41385-021-00389-4

Negus, S. S., & Banks, M. L. (2018). Pharmacokinetic—Pharmacodynamic (PKPD) Analysis with Drug Discrimination. *Current Topics in Behavioral Neurosciences*, *39*, 245–259. https://doi.org/10.1007/7854_2016_36

Neto, J. C. R., Calder, P. C., Curi, R., Newsholme, P., Sethi, J. K., & Silveira, L. S. (2021). The Immunometabolic Roles of Various Fatty Acids in Macrophages and Lymphocytes. *Int. J. Mol. Sci.*, *22*, 8460. https://doi.org/10.3390/ijms22168460

Neumann, E. K., Djambazova, K. V., Caprioli, R. M., & Spraggins, J. M. (2020). Multimodal Imaging Mass Spectrometry: Next Generation Molecular Mapping in Biology and Medicine. *Journal of the American Society for Mass Spectrometry*, *31*(12), 2401–2415. https://doi.org/10.1021/jasms.0c00232

Nilsson, A., Fehniger, T. E., Gustavsson, L., Andersson, M., Kenne, K., Marko-Varga, G., & Andrén, P. E. (2010). Fine Mapping the Spatial Distribution and Concentration of Unlabeled Drugs within Tissue Micro-Compartments Using Imaging Mass Spectrometry. *PLOS ONE*, *5*(7), e11411. https://doi.org/10.1371/journal.pone.0011411

Nilsson, A., Goodwin, R. J. A., Shariatgorji, M., Vallianatou, T., Webborn, P. J. H., & Andrén, P. E. (2015). Mass Spectrometry Imaging in Drug Development. *Analytical Chemistry*, *87*(3), 1437–1455. https://doi.org/10.1021/ac504734s

Novak, E. A., & Mollen, K. P. (2015). Mitochondrial dysfunction in inflammatory bowel disease. *Frontiers in Cell and Developmental Biology*, *3*. https://doi.org/10.3389/fcell.2015.00062

O'Donnell, V. B., Dennis, E. A., Wakelam, M. J. O., & Subramaniam, S. (2019). LIPID MAPS: Serving the next generation of lipid researchers with tools, resources, data, and training. *Science Signaling*, *12*(563), eaaw2964. https://doi.org/10.1126/scisignal.aaw2964

Palmer, A. (2017). *FDR-controlled metabolite annotation for high-resolution imaging mass spectrometry*.

Palmer, A., Phapale, P., Chernyavsky, I., Lavigne, R., Fay, D., Tarasov, A., Kovalev, V., Fuchser, J., Nikolenko, S., Pineau, C., Becker, M., & Alexandrov, T. (2017). FDR-controlled metabolite annotation for high-resolution imaging mass spectrometry. *Nature Methods*, *14*(1), 57–60. https://doi.org/10.1038/nmeth.4072

Palmer, A., Trede, D., & Alexandrov, T. (2016). Where imaging mass spectrometry stands: Here are the numbers. *Metabolomics*, *12*(6). https://doi.org/10.1007/s11306-016-1047-0

Panuwet, P., Hunter Jr., R. E., D'Souza, P. E., Chen, X., Radford, S. A., Cohen, J. R., Marder, M. E., Kartavenka, K., Ryan, P. B., & Barr, D. B. (2016). Biological Matrix Effects in Quantitative Tandem Mass Spectrometry-Based Analytical Methods: Advancing Biomonitoring. *Critical Reviews in Analytical Chemistry*, *46*(2), 93–105. https://doi.org/10.1080/10408347.2014.980775

Pirman, D. A., Kiss, A., Heeren, R. M. A., & Yost, R. A. (2013). Identifying Tissue-Specific Signal Variation in MALDI Mass Spectrometric Imaging by Use of an Internal Standard. *Analytical Chemistry*, *85*(2), 1090–1096. https://doi.org/10.1021/ac3029618

Pirman, D. A., Reich, R. F., Kiss, A., Heeren, R. M. A., & Yost, R. A. (2013). Quantitative MALDI Tandem Mass Spectrometric Imaging of Cocaine from Brain Tissue with a Deuterated

Internal Standard. *Analytical Chemistry*, *85*(2), 1081–1089. https://doi.org/10.1021/ac302960j

Pirman, D. A., & Yost, R. A. (2011). Quantitative Tandem Mass Spectrometric Imaging of Endogenous Acetyl-l-carnitine from Piglet Brain Tissue Using an Internal Standard. *Analytical Chemistry*, *83*(22), 8575–8581. https://doi.org/10.1021/ac201949b

Poduval, P., Parsekar, S., & Meena, S. N. (2023). Chapter 10—Small molecules vs biologics. In S. N. Meena, V. Nandre, K. Kodam, & R. S. Meena (Eds.), *New Horizons in Natural Compound Research* (pp. 179–199). Academic Press. https://doi.org/10.1016/B978-0-443-15232-0.00001-1

Polański, K., Young, M. D., Miao, Z., Meyer, K. B., Teichmann, S. A., & Park, J.-E. (2020). BBKNN: Fast batch alignment of single cell transcriptomes. *Bioinformatics (Oxford, England)*, *36*(3), 964–965. https://doi.org/10.1093/bioinformatics/btz625

Reichel, A., & Lienau, P. (2016). Pharmacokinetics in Drug Discovery: An Exposure-Centred Approach to Optimising and Predicting Drug Efficacy and Safety. *Handbook of Experimental Pharmacology*, *232*, 235–260. https://doi.org/10.1007/164_2015_26

Reveille, J. D. (2015). Biomarkers for diagnosis, monitoring of progression, and treatment responses in ankylosing spondylitis and axial spondyloarthritis. *Clinical Rheumatology*, *34*(6), 1009–1018. https://doi.org/10.1007/s10067-015-2949-3

Reyzer, M. L., Hsieh, Y., Ng, K., Korfmacher, W. A., & Caprioli, R. M. (2003). Direct analysis of drug candidates in tissue by matrix-assisted laser desorption/ionization mass spectrometry. *Journal of Mass Spectrometry: JMS*, *38*(10), 1081–1092. https://doi.org/10.1002/jms.525

Robinson, P. C., Costello, M.-E., Leo, P., Bradbury, L. A., Hollis, K., Cortes, A., Lee, S., Joo, K. B., Shim, S.-C., Weisman, M., Ward, M., Zhou, X., Garchon, H.-J., Chiocchia, G., Nossent, J., Lie, B. A., Førre, Ø., Tuomilehto, J., Laiho, K., … Brown, M. A. (2015). ERAP2 is associated with ankylosing spondylitis in HLA-B27-positive and HLA-B27-negative patients. *Annals of the Rheumatic Diseases*, *74*(8), 1627–1629. https://doi.org/10.1136/annrheumdis-2015-207416

Rosenbaum, J. T., Gill, T., & Martin, T. M. (2021). EnRAPtured: Is Endoplasmic Reticulum Aminopeptidase a New Clue to the Pathogenesis and ThERAPy of Uveitis? *Ophthalmology Science*, *1*(3), 100056. https://doi.org/10.1016/j.xops.2021.100056

Sarafian, M. H., Gaudin, M., Lewis, M. R., Martin, F.-P., Holmes, E., Nicholson, J. K., & Dumas, M.-E. (2014). Objective Set of Criteria for Optimization of Sample Preparation Procedures for Ultra-High Throughput Untargeted Blood Plasma Lipid Profiling by Ultra Performance Liquid Chromatography–Mass Spectrometry. *Analytical Chemistry*, *86*(12), 5766–5774. https://doi.org/10.1021/ac500317c

Saric, T., Chang, S.-C., Hattori, A., York, I. A., Markant, S., Rock, K. L., Tsujimoto, M., & Goldberg, A. L. (2002). An IFN-γ–induced aminopeptidase in the ER, ERAP1, trims precursors to MHC class I–presented peptides. *Nature Immunology*, *3*(12), 1169–1176. https://doi.org/10.1038/ni859

Saveanu, L., Carroll, O., Lindo, V., Del Val, M., Lopez, D., Lepelletier, Y., Greer, F., Schomburg, L., Fruci, D., Niedermann, G., & van Endert, P. M. (2005). Concerted peptide trimming by human ERAP1 and ERAP2 aminopeptidase complexes in the endoplasmic reticulum. *Nature Immunology*, *6*(7), 689–697. https://doi.org/10.1038/ni1208

Schlosstein Lee, Terasaki Paul I., Bluestone Rodney, & Pearson Carl M. (1973). High Association of an HL-A Antigen, W27, with Ankylosing Spondylitis. *New England Journal of Medicine*, *288*(14), 704–706. https://doi.org/10.1056/NEJM197304052881403

Schmidt, U., Weigert, M., Broaddus, C., & Myers, G. (2018). Cell Detection with Star-Convex Polygons. *Medical Image Computing and Computer Assisted Intervention – MICCAI 2018*, 265–273. https://doi.org/10.1007/978-3-030-00934-2_30

Schott, B. H., Wang, L., Zhu, X., Harding, A. T., Ko, E. R., Bourgeois, J. S., Washington, E. J., Burke, T. W., Anderson, J., Bergstrom, E., Gardener, Z., Paterson, S., Brennan, R. G., Chiu, C., McClain, M. T., Woods, C. W., Gregory, S. G., Heaton, N. S., & Ko, D. C. (2022). Single-cell genome-wide association reveals that a nonsynonymous variant in ERAP1 confers

increased susceptibility to influenza virus. *Cell Genomics*, *2*(11), 100207. https://doi.org/10.1016/j.xgen.2022.100207

Schütze, S., Potthoff, K., Machleidt, T., Berkovic, D., Wiegmann, K., & Krönke, M. (1992). TNF activates NF-κB by phosphatidylcholine-specific phospholipase C-induced "Acidic" sphingomyelin breakdown. *Cell*, *71*(5), 765–776. https://doi.org/10.1016/0092-8674(92)90553-O

Serwold, T., Gonzalez, F., Kim, J., Jacob, R., & Shastri, N. (2002). ERAAP customizes peptides for MHC class I molecules in the endoplasmic reticulum. *Nature*, *419*(6906), 480–483. https://doi.org/10.1038/nature01074

Sevastou, I., Kaffe, E., Mouratis, M.-A., & Aidinis, V. (2013). Lysoglycerophospholipids in chronic inflammatory disorders: The PLA2/LPC and ATX/LPA axes. *Biochimica et Biophysica Acta (BBA) - Molecular and Cell Biology of Lipids*, *1831*(1), 42–60. https://doi.org/10.1016/j.bbalip.2012.07.019

Smets, T., Verbeeck, N., Claesen, M., Asperger, A., Griffioen, G., Tousseyn, T., Waelput, W., Waelkens, E., & De Moor, B. (2019). Evaluation of Distance Metrics and Spatial Autocorrelation in Uniform Manifold Approximation and Projection Applied to Mass Spectrometry Imaging Data. *Analytical Chemistry*, *91*(9), 5706–5714. https://doi.org/10.1021/acs.analchem.8b05827

Smith, S. A., Ogawa, S. A., Chau, L., Whelan, K. A., Hamilton, K. E., Chen, J., Tan, L., Chen, E. Z., Keilbaugh, S., Fogt, F., Bewtra, M., Braun, J., Xavier, R. J., Clish, C. B., Slaff, B., Weljie, A. M., Bushman, F. D., Lewis, J. D., Li, H., … Wu, G. D. (2021). Mitochondrial dysfunction in inflammatory bowel disease alters intestinal epithelial metabolism of hepatic acylcarnitines. *The Journal of Clinical Investigation*, *131*(1), e133371. https://doi.org/10.1172/JCI133371

Stolwijk, C., Tubergen, A. van, Castillo-Ortiz, J. D., & Boonen, A. (2015). Prevalence of extra-articular manifestations in patients with ankylosing spondylitis: A systematic review and meta-analysis. *Annals of the Rheumatic Diseases*, *74*(1), 65–73. https://doi.org/10.1136/annrheumdis-2013-203582

Strange, A., Capon, F., Spencer, C. C., Knight, J., Weale, M. E., Allen, M. H., Barton, A., Band, G., Bellenguez, C., Bergboer, J. G., Blackwell, J. M., Bramon, E., Bumpstead, S. J., Casas, J. P., Cork, M. J., Corvin, A., Deloukas, P., Dilthey, A., Duncanson, A., … Trembath, R. C. (2010). Genome-wide association study identifies new psoriasis susceptibility loci and an interaction between HLA-C and ERAP1. *Nature Genetics*, *42*(11), 985–990. https://doi.org/10.1038/ng.694

Sumner, L. W., Amberg, A., Barrett, D., Beale, M. H., Beger, R., Daykin, C. A., Fan, T. W.-M., Fiehn, O., Goodacre, R., Griffin, J. L., Hankemeier, T., Hardy, N., Harnly, J., Higashi, R., Kopka, J., Lane, A. N., Lindon, J. C., Marriott, P., Nicholls, A. W., … Viant, M. R. (2007). Proposed minimum reporting standards for chemical analysis Chemical Analysis Working Group (CAWG) Metabolomics Standards Initiative (MSI). *Metabolomics : Official Journal of the Metabolomic Society*, *3*(3), 211–221. https://doi.org/10.1007/s11306-007-0082-2

Svensson, M., Sköld, K., Svenningsson, P., & Andren, P. E. (2003). Peptidomics-Based Discovery of Novel Neuropeptides. *Journal of Proteome Research*, *2*(2), 213–219. https://doi.org/10.1021/pr020010u

Tadenev, A. L. D., & Burgess, R. W. (2019). Model validity for preclinical studies in precision medicine: Precisely how precise do we need to be? *Mammalian Genome*, *30*(5), 111–122. https://doi.org/10.1007/s00335-019-09798-0

Taugor, J. D., Maika, S. D., Satumtira, N., Dorris, M. L., McLean, I. L., Yanagisawa, H., Sayad, A., Stagg, A. J., Fox, G. M., Le O'Brein, A., Rehman, M., Zhou, M., Weiner, A. L., Splawski, J. B., Richardson, J. A., & Hammer, R. E. (1999). Inflammatory disease in HLA-B27 transgenic rats. *Immunological Reviews*, *169*(1), 209–223. https://doi.org/10.1111/j.1600-065X.1999.tb01317.x

Tobias, F., & Hummon, A. B. (2020). Considerations for MALDI-Based Quantitative Mass Spectrometry Imaging Studies. *Journal of Proteome Research*, *19*(9), 3620–3630. https://doi.org/10.1021/acs.jproteome.0c00443

Troendle, F. J., Reddick, C. D., & Yost, R. A. (1999). Detection of pharmaceutical compounds in tissue by matrix-assisted laser desorption/ionization and laser desorption/chemical ionization tandem mass spectrometry with a quadrupole ion trap. *Journal of the American Society for Mass Spectrometry*, *10*(12), 1315–1321. https://doi.org/10.1016/S1044-0305(99)00103-8

Vitorino, R., Barros, A. S., Caseiro, A., Ferreira, R., & Amado, F. (2012). Evaluation of different extraction procedures for salivary peptide analysis. *Talanta*, *94*, 209–215. https://doi.org/10.1016/j.talanta.2012.03.023

Want, E. J., Masson, P., Michopoulos, F., Wilson, I. D., Theodoridis, G., Plumb, R. S., Shockcor, J., Loftus, N., Holmes, E., & Nicholson, J. K. (2013). Global metabolic profiling of animal and human tissues via UPLC-MS. *Nature Protocols*, *8*(1), 17–32. https://doi.org/10.1038/nprot.2012.135

Webers, C., Ortolan, A., Sepriano, A., Falzon, L., Baraliakos, X., Landewé, R. B. M., Ramiro, S., van der Heijde, D., & Nikiphorou, E. (2023). Efficacy and safety of biological DMARDs: A systematic literature review informing the 2022 update of the ASAS-EULAR recommendations for the management of axial spondyloarthritis. *Annals of the Rheumatic Diseases*, *82*(1), 130–141. https://doi.org/10.1136/ard-2022-223298

Weigert, M., & Schmidt, U. (2022). Nuclei Instance Segmentation and Classification in Histopathology Images with Stardist. *The IEEE International Symposium on Biomedical Imaging Challenges (ISBIC)*, 1–4. https://doi.org/10.1109/ISBIC56247.2022.9854534

Weigert, M., Schmidt, U., Haase, R., Sugawara, K., & Myers, G. (2020). Star-convex Polyhedra for 3D Object Detection and Segmentation in Microscopy. *The IEEE Winter Conference on Applications of Computer Vision (WACV)*. https://doi.org/10.1109/WACV45572.2020.9093435

Wellcome Trust Case Control Consortium, Australo-Anglo-American Spondylitis Consortium (TASC), Burton, P. R., Clayton, D. G., Cardon, L. R., Craddock, N., Deloukas, P., Duncanson, A., Kwiatkowski, D. P., McCarthy, M. I., Ouwehand, W. H., Samani, N. J., Todd, J. A., Donnelly, P., Barrett, J. C., Davison, D., Easton, D., Evans, D. M., Leung, H.-T., … Brown,

M. (2007). Association scan of 14,500 nonsynonymous SNPs in four diseases identifies autoimmunity variants. *Nature Genetics*, *39*(11), 1329–1337. https://doi.org/10.1038/ng.2007.17

Wells, J. M., Brummer, R. J., Derrien, M., MacDonald, T. T., Troost, F., Cani, P. D., Theodorou, V., Dekker, J., Méheust, A., de Vos, W. M., Mercenier, A., Nauta, A., & Garcia-Rodenas, C. L. (2017). Homeostasis of the gut barrier and potential biomarkers. *American Journal of Physiology-Gastrointestinal and Liver Physiology*, *312*(3), G171–G193. https://doi.org/10.1152/ajpgi.00048.2015

Winderbaum, L. J., Koch, I., Gustafsson, O. J. R., Meding, S., & Hoffmann, P. (2015). Feature extraction for proteomics imaging mass spectrometry data. *The Annals of Applied Statistics*, *9*(4), 1973–1996. https://doi.org/10.1214/15-AOAS870

Wishart, D. S., Tian, S., Allen, D., Oler, E., Peters, H., Lui, V. W., Gautam, V., Djoumbou-Feunang, Y., Greiner, R., & Metz, T. O. (2022). BioTransformer 3.0—A web server for accurately predicting metabolic transformation products. *Nucleic Acids Research*, *50*(W1), W115–W123. https://doi.org/10.1093/nar/gkac313

Wishart, D. S., Tzur, D., Knox, C., Eisner, R., Guo, A. C., Young, N., Cheng, D., Jewell, K., Arndt, D., Sawhney, S., Fung, C., Nikolai, L., Lewis, M., Coutouly, M.-A., Forsythe, I., Tang, P., Shrivastava, S., Jeroncic, K., Stothard, P., … Querengesser, L. (2007). HMDB: The Human Metabolome Database. *Nucleic Acids Research*, *35*(suppl_1), D521–D526. https://doi.org/10.1093/nar/gkl923

Wiśniewski, A., Matusiak, Ł., Szczerkowska-Dobosz, A., Nowak, I., Łuszczek, W., & Kuśnierczyk, P. (2018). The association of ERAP1 and ERAP2 single nucleotide polymorphisms and their haplotypes with psoriasis vulgaris is dependent on the presence or absence of the HLA-C*06:02 allele and age at disease onset. *Human Immunology*, *79*(2), 109–116. https://doi.org/10.1016/j.humimm.2017.11.010

Wyatt, P. G., Gilbert, I. H., Read, K. D., & Fairlamb, A. H. (2011). Target Validation: Linking Target and Chemical Properties to Desired Product Profile. *Current Topics in Medicinal Chemistry*, *11*(10), 1275–1283. https://doi.org/10.2174/156802611795429185

Xie, C., Wang, Q., Li, G., Fan, Z., Wang, H., & Wu, X. (2019). Dietary supplement with nucleotides in the form of uridine monophosphate or uridine stimulate intestinal development and promote nucleotide transport in weaned piglets. *Journal of the Science of Food and Agriculture*, *99*(13), 6108–6113. https://doi.org/10.1002/jsfa.9850

Xu, D., Ma, R., Ju, Y., Song, X., Niu, B., Hong, W., Wang, R., Yang, Q., Zhao, Z., Zhang, Y., Zheng, Y., Bai, Q., Lv, M., Sun, N., & Li, X. (2022). Cholesterol sulfate alleviates ulcerative colitis by promoting cholesterol biosynthesis in colonic epithelial cells. *Nature Communications*, *13*, 4428. https://doi.org/10.1038/s41467-022-32158-7

York, I. A., Chang, S.-C., Saric, T., Keys, J. A., Favreau, J. M., Goldberg, A. L., & Rock, K. L. (2002). The ER aminopeptidase ERAP1 enhances or limits antigen presentation by trimming epitopes to 8–9 residues. *Nature Immunology*, *3*(12), 1177–1184. https://doi.org/10.1038/ni860

Zeidan, M. J., Saadoun, D., Garrido, M., Klatzmann, D., Six, A., & Cacoub, P. (2016). Behçet's disease physiopathology: A contemporary review. *Autoimmunity Highlights*, *7*(1), 4. https://doi.org/10.1007/s13317-016-0074-1

Zervoudi, E., Saridakis, E., Birtley, J. R., Seregin, S. S., Reeves, E., Kokkala, P., Aldhamen, Y. A., Amalfitano, A., Mavridis, I. M., James, E., Georgiadis, D., & Stratikos, E. (2013). Rationally designed inhibitor targeting antigen-trimming aminopeptidases enhances antigen presentation and cytotoxic T-cell responses. *Proceedings of the National Academy of Sciences*, *110*(49), 19890–19895. https://doi.org/10.1073/pnas.1309781110

Zhang, S. (Weihua), & Jian, W. (2014). Recent advances in absolute quantification of peptides and proteins using LC-MS. *Reviews in Analytical Chemistry*, *33*(1), 31–47. https://doi.org/10.1515/revac-2013-0019

Zheng, J., Mehl, J., Zhu, Y., Xin, B., & Olah, T. (2014). Application and Challenges in Using LC–MS Assays for Absolute Quantitative Analysis of Therapeutic Proteins in Drug Discovery. *Bioanalysis*, *6*(6), 859–879. https://doi.org/10.4155/bio.14.36

Zhong, L., Li, Y., Xiong, L., Wang, W., Wu, M., Yuan, T., Yang, W., Tian, C., Miao, Z., Wang, T., & Yang, S. (2021). Small molecules in targeted cancer therapy: Advances, challenges, and future perspectives. *Signal Transduction and Targeted Therapy*, *6*(1), 1–48. https://doi.org/10.1038/s41392-021-00572-w