Characterisation of a bacterial P450 system using surface mass spectrometry

A thesis submitted to The University of Manchester for the degree of Master of Philosophy in the Faculty of Engineering and Physical Sciences

2016

Andres Galindo Garcia

University of Manchester
Faculty of Engineering and Physical Sciences
School of Chemistry
## Contents

List of Figures .......................................................................................................................... 5  
List of Tables .......................................................................................................................... 8  
List of Equations ..................................................................................................................... 9  
Abbreviations .......................................................................................................................... 10  
Abstract .................................................................................................................................... 12  
Declaration ............................................................................................................................... 13  
Copyright Statement .............................................................................................................. 14  
Acknowledgements .................................................................................................................. 15  

Chapter 1 Literature review ................................................................................................. 16  
1.1 Introduction ....................................................................................................................... 16  
1.2.1 Mass spectrometry ........................................................................................................ 16  
1.2.2 Mass spectrometry in microbiology .............................................................................. 17  
1.2.3 Mass spectrometry imaging ........................................................................................... 18  
1.3 Matrix-assisted laser desorption ionisation mass spectrometry .................................... 18  
1.4.1 Secondary ion mass spectrometry ................................................................................ 19  
1.4.2 The basic equation ........................................................................................................ 20  
1.4.3 The static limit ................................................................................................................ 20  
1.5 Desorption electrospray ionisation .................................................................................. 21  
1.6 Rapid evaporative ionisation mass spectrometry ............................................................ 22  
1.7.1 Mass analysers .............................................................................................................. 25  
1.7.2 Mass resolution ............................................................................................................. 25  
1.7.3 Mass accuracy ............................................................................................................... 27  
1.7.4 Time-of-flight mass analysers ...................................................................................... 28  
1.7.5 Quadrupole analysers ................................................................................................. 28  
1.7.6.1 Hybrid instruments ............................................................................................. 30
Chapter 2 Instrumentation ................................................................. 37

2.1.1 J105 3D chemical imager ......................................................... 37
2.1.2 C_{60} ion source ........................................................................ 39
2.1.3 Sample handling system ............................................................. 40
2.2 DESI 2D ion source ..................................................................... 41
2.3 iKnife .......................................................................................... 42
2.4 Xevo G2-XS QToF ....................................................................... 43

Chapter 3 Materials and methodology .............................................. 45

3.1.1 Standard samples ...................................................................... 45
3.1.2 Diclofenac ................................................................................ 45
3.1.3 5-Hydroxydiclofenac ................................................................. 46
3.2.1 Biological system ..................................................................... 47
3.2.2 Transformation protocol ........................................................... 47
3.1.3 Biological samples ................................................................... 49
3.3.1 SIMS sample preparation .......................................................... 49
3.2.2 Substrate cleaning procedure ................................................... 49
3.2.3 Spin coating ............................................................................ 50
3.1.2 Desalting ............................................................................... 51
3.4 SIMS experiment conditions ....................................................... 51
3.5 DESI sample preparation .............................................................. 51
3.6 DESI experiment conditions ....................................................... 51
List of Figures

Figure 1.1: Different ionisation mechanisms used in mass spectrometry techniques. Reprinted with permission from (Cooks et al., 2006) ..................................................... 17
Figure 1.2: Labeled schematic depicting the ionisation process in a DESI experiment. Reprinted with permission from (Cooks et al., 2006) .......................... 22
Figure 1.3: Cross-validation results accounting for the identification of 28 bacterial species by REIMS, as a function of the bin size on species, genus and Gram-stain levels. Reprinted with permission from (Strittmatter et al., 2014) ......................... 23
Figure 1.4: Δm measurement for each given definition. Reprinted with permission from (Muenster, 2009) ................................................................................. 26
Figure 1.5: Comparison of the spectral output from a particular analyte at different mass resolution values. Reprinted with permission from (Muenster, 2009) ...... 27
Figure 1.6: Quadrupole mass analyser schematic (a) and photograph (b). Reprinted with permission from Jurgen. (Gross, 2004) ......................................................... 30
Figure 1.7: Schematic illustrating the orthogonal acceleration process in a ReToF instrument. Modified from (Waters, 2013a) .............................................................. 31
Figure 1.8: (A) Typical absorption spectrum for a CYP enzyme showing the 450nm band shift generated by the complex with CO alongside a main (Soret) band at ∼420nm. (Munro et al., 2007) (B) Structure of heme prosthetic group iron protoporphyrin IX, heme b. (Sono et al., 1996) ........................................... 32
Figure 1.9: Catalytic cycle of CYP monooxygenation. Substrate (S) is oxidised through steps 1-6 to form SO, with intermediates shown (A-F); uncoupled reactions lead to H$_2$O$_2$ formation pathways (steps 2a and 3a) while Compound I generation via the shunt pathway is also depicted (Steps 7 and 8) For simplicity, the chemical bond (Fe$^{III}$–O) is omitted, alongside the charge state of the iron atom and the porphyrin moiety, which is present in all intermediates but shown only for Compound I. Modified from (Hrycay and Bandiera, 2015) ......................................................... 33
Figure 1.10: Schematic representations of class VII P450 RhF. (A) Topographical arrangement of the fused protein, with N and C terminals shown and a NADPH binding motif implied. (B) Domain organisation based on sequence analysis with the heme-ligand pocket displayed. Reprinted with permission from (O’Reilly et al., 2013) and modified with permission from (Hunter et al., 2005) respectively ...... 35
Figure 2.1: Schematic indicating the main components of the J105-3D Chemical Imager. Reprinted with permission from (Fletcher and Vickerman, 2010) ............ 37
Figure 2.2: Schematic of the shaped field buncher and reflectron ToF mass analyser design. Reproduced with permission from (Hill et al., 2011) .......................................................... 38
Figure 2.3: Schematic of the C_{60}^{+} ion gun optics. Reproduced with permission from (Weibel et al., 2003) ........................................................................................................ 40
Figure 2.4: Copper stub with samples mounted using silicon wafers as a substrate ................................................................................................................................. 41
Figure 2.5: Prosolia’s DESI 2D ion source ............................................................... 42
Figure 2.6: Schematic of the components and workflow for REIMS. Reprinted from (Waters, 2015) ........................................................................................................ 43
Figure 2.7: Schematic of the ion optics in the Xevo G2-XS QTof high definition mass spectrometry system. Reprinted with permission from (Waters, 2014) ............... 44
Figure 3.1: Diclofenac sodium salt. ChemDraw Professional (version 15.1.0.144) software .......................................................................................................... 46
Figure 3.2: P450 RhF monooxidation reaction over diclofenac to produce 5-hydroxydiclofenac. Relevant properties from both molecules displayed. Figures and information generated with ChemDraw Professional (version 15.1.0.144) software ........................................................................................................ 47
Figure 3.3: (A) E. coli transformed with an empty plasmid, thus lacking the RhF gene. Colonies are small with a bright slightly yellow colour. (B) RhF expressing E. coli colonies display a darker colour, a translucent yellow with a mild red tonality, due to the heme group in the enzyme ........................................................................ 48
Figure 4.1: Spectrum of indium standard (MW: 114.90) analysed in positive ion mode with 20 kV C_{60}^{+} using an ion dose of 8.25E11 ions/cm². The signal intensity obtained is optimal, however, a ramp can be seen within the of 116-118 Da range in the 110-120 Da spectrum, due to a technical issue with the buncher plates .......... 55
Figure 4.2: Expected molecular signal for (A) diclofenac and (B) 5-hydroxydiclofenac, according to Surface Identity Software .................................................. 56
Figure 4.3: Diclofenac spectrum acquired in positive ion mode using the 20 kV C_{60}^{+} ion source with an ion dose of 1.32E12 ions/cm². (A) The intact molecule is not detected in the spectrum. (B) Two main fragments peaks were assigned in the spectrum ........................................................................................................ 57
Figure 4.4: 5-Hydroxydiclofenac spectrum acquired using the 20 kV C_{60}^{+} ion source with an ion dose of 1.32E12 ions/cm². The isotopic pattern displayed matches with the expected one ........................................................................................................ 58
Figure 4.5: Mass spectrum generated from RhF (A) and control E. coli (B) bacterial
samples. Defined peaks only appear in the lower mass region. Imprinting technique

Figure 4.6: Mass spectra acquired from rat brain tissue during parameter optimisation. Modifying the geometry of the sample with respect to the inlet, and instrumental settings such as voltage or temperature all contribute to obtaining good quality spectra with a low S/N ratio ................................. 59

Figure 4.7: DESI spectra from bacterial samples in negative ion mode where the peaks from phospholipids and fatty acids clearly dominate ................................. 61

Figure 4.8: Diclofenac and 5-Hydroxydiclofenac spectra acquired from a glass surface. MS/MS capability was used to properly characterise the molecule and its fragments ................................. 62

Figure 4.9: Imaging experiment revealing bacterial colonies (green), diclofenac (red) and 5-hydroxydiclofenac (blue) over the background signals (white), based on the abundance of distinctive signals selected for each of the species. Image generated and processed in HDI v1.4 software using TIC normalisation ................................. 63

Figure 4.10: Bacterial colonies of control (red) and RhF E. coli samples (blue). Images processed using different colour gradients to show the analytes with highest maximum intensity for each sample. Normalised by TIC ................................. 64

Figure 4.11: (A) Combined spectrum from selected ROIs from control and RhF bacterial samples. (B) Combined spectrum after maximum normalisation ................................. 66

Figure 4.12: Loading plot for PC1 ................................. 67

Figure 4.13: Scatter plot visually separating bacterial colonies through PC1, PC2 and PC3. Control colonies are depicted in red, RhF samples are coloured in blue ................................. 68

Figure 4.14: Zoom in of the negative ion DESI spectra from E. coli using shows differences of 0.333 m/z between peaks denote a charged state (+3 in this case) which implies the detection of an analyte of 3300 Da approximately ................................. 70

Figure 4.15: Spectral output for DESI (A) and REIMS (B) for the analysis of colonies belonging to the same cultures. Green spectra belong to control samples, red spectra belongs to RhF samples ................................. 72
List of tables

Table 1.1: Specifications, advantages and disadvantages of the MS techniques.24
Table 1.2: Mass analysers in mass spectrometry. Reprinted with permission from
(De Hoffmann and Stroobant, 2007) ...........................................................25
Table 4.1: Variance captured by PCA model for the ROIs after maximum
normalisation ...........................................................................................67
Table 4.2: Identity of lipids accounting for major variations registered in PC1.
FA: fatty acids, PE: phosphatidylethanolamine, PG: phosphatidylglicerol. Assigned in
accordance with previous lipid profiling studies in E. coli. (Garrett et al., 2011,
Oursel et al., 2007) ...................................................................................69
List of equations

Equation 1.1 .............................................................................................................. 20
Equation 1.2 .............................................................................................................. 26
Equation 1.3 .............................................................................................................. 29
Equation 2.1 .............................................................................................................. 39
**Abbreviations**

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Full Form</th>
</tr>
</thead>
<tbody>
<tr>
<td>MS</td>
<td>Mass spectrometry</td>
</tr>
<tr>
<td>m/z</td>
<td>Mass-to-charge ratio</td>
</tr>
<tr>
<td>FDA</td>
<td>Food and drug administration</td>
</tr>
<tr>
<td>MSI</td>
<td>Mass spectrometry imaging</td>
</tr>
<tr>
<td>SIMS</td>
<td>Secondary ion mass spectrometry</td>
</tr>
<tr>
<td>2D</td>
<td>Two-dimensional</td>
</tr>
<tr>
<td>3D</td>
<td>Three-dimensional</td>
</tr>
<tr>
<td>MALDI</td>
<td>Matrix-assisted laser desorption ionisation</td>
</tr>
<tr>
<td>UV</td>
<td>Ultra-violet</td>
</tr>
<tr>
<td>ToF</td>
<td>Time-of-flight</td>
</tr>
<tr>
<td>DESI</td>
<td>Desorption electrospray ionisation</td>
</tr>
<tr>
<td>REIMS</td>
<td>Rapid evaporative ionisation mass spectrometry</td>
</tr>
<tr>
<td>RF</td>
<td>Radio-frequency</td>
</tr>
<tr>
<td>MS/MS</td>
<td>Tandem mass spectrometry</td>
</tr>
<tr>
<td>FT-ICR</td>
<td>Fourier transform ion cyclotron resonance</td>
</tr>
<tr>
<td>FWHM</td>
<td>Full width at half maximum</td>
</tr>
<tr>
<td>DC</td>
<td>Direct current</td>
</tr>
<tr>
<td>AC</td>
<td>Alternating current</td>
</tr>
<tr>
<td>oa-ToF</td>
<td>Orthogonal acceleration time-of-flight</td>
</tr>
<tr>
<td>Re-ToF</td>
<td>Reflectron time-of-flight</td>
</tr>
<tr>
<td>P450 or CYP</td>
<td>Cytochrome P450</td>
</tr>
<tr>
<td>NAD(P)H</td>
<td>Nicotidamide adenine dinucleotide phosphate</td>
</tr>
<tr>
<td>Cpd I</td>
<td>Compound I</td>
</tr>
<tr>
<td>FMN</td>
<td>Flavin mono-nucleotide</td>
</tr>
<tr>
<td>SARC</td>
<td>Surface analysis research centre</td>
</tr>
<tr>
<td>SAC</td>
<td>Sample analysis chamber</td>
</tr>
<tr>
<td>PREP</td>
<td>Preparation chamber</td>
</tr>
<tr>
<td>ADC</td>
<td>Analogue-to-digital converter</td>
</tr>
<tr>
<td>RMS</td>
<td>Root mean square</td>
</tr>
<tr>
<td>NSAID</td>
<td>Non-steroidal anti-inflammatory drug</td>
</tr>
<tr>
<td>PTGS</td>
<td>Prostaglandin-endoperoxide synthase</td>
</tr>
<tr>
<td>COX-1</td>
<td>Cyclooxygenase-2</td>
</tr>
<tr>
<td>COX-2</td>
<td>Cyclooxygenase-2</td>
</tr>
</tbody>
</table>
HPLC  High-performance liquid chromatography
NMR  Nuclear magnetic resonance
E. coli  Escherichia coli
SOC  Super optimal broth with catabolic repressor
IPTG  Isopropyl-β-D-thiogalactopyranoside
LB  Lysogeny broth
BSL-1  Biosafety level 1
S/N  Signal-to-noise ratio
TIC  Total ion current
MSn  Multiple mass spectrometry
PCA  Principal components analysis
PC  Principal components
ROI  Region of interest
PG  Phosphatidylglycerol
PE  Phosphatidylethanolamine
FA  Fatty acids
WT  Wild type
LC-MS  Liquid chromatography-mass spectrometry
ESI-MS  Electrospray ionisation-mass spectrometry
Abstract

The University of Manchester

Candidate’s name: Andres Galindo Garcia

Degree title: Master of Philosophy

Thesis title: Characterisation of a bacterial P450 system using surface mass spectrometry

In this study, the components of a biotechnologically relevant *E. coli* expression system were characterised through the use of mass spectrometry techniques. These include the bacterial host itself, the enzymatic product of the P450 RhF gene expression which catalyses a selective hydroxylation reaction that produces 5-hydroxydiclofenac from a diclofenac substrate; and both the substrate and product of the reaction. The identification of molecular signals as phosphatidylglycerol (PG), phosphatidylethanolamine (PE) and fatty acids (FA) led to the generation of a lipid profile. Instrumentation relying on different ionisation mechanisms was explored and bioinformatic tools were used to extract useful information from large imaging data sets through normalisation and PCA analysis.

September 2016
Declaration

No portion of the work referred to in the thesis has been submitted in support of an application for another degree or qualification of this or any other university or other institute of learning.
Copyright Statement

i. The author of this thesis (including any appendices and/or schedules to this thesis) owns certain copyright or related rights in it (the "Copyright") and s/he has given The University of Manchester certain rights to use such Copyright, including for administrative purposes.

ii. Copies of this thesis, either in full or in extracts and whether in hard or electronic copy, may be made only in accordance with the Copyright, Designs and Patents Act 1998 (as amended) and regulations issued under it or, where appropriate, in accordance with licensing agreements which the University has from time to time. This page must form part of any such copies made.

iii. The ownership of certain Copyright, patents, designs, trade marks and other intellectual property (the "Intellectual Property") and any reproductions of copyright works in the thesis, for example graphs and tables ("Reproductions"), which may be described in this thesis, may not be owed by the author and may be owned by third parties. Such Intellectual Property and Reproductions cannot and must not be made available for use without the prior written permission of the owner(s) of the relevant Intellectual Property and/or Reproductions.

iv. Further information on the conditions under which disclosure, publication and commercialisation of this thesis, the Copyright and any Intellectual Property and/or Reproductions described in it may take place is available in the University IP Policy (see http://documents.manchester.ac.uk/DocuInfo.aspx?DocID=487), in any relevant Thesis restriction declarations deposited in the University Library, The University Library’s regulations (see http://manchester.ac.uk/library/aboutus/regulations) and in The University’s policy on Presentation of Theses.
Acknowledgement

First, I would like to thank my dear mum for her never-ending love and understanding, for always encouraging me to follow my heart, and for her endless support through every challenge I have ever faced in my life.

I would like to acknowledge the Mexican National Council for Science and Technology (CONACyT), for funding my MPhil studies.

I would like to express my gratitude towards my supervisor, Dr. Nick Lockyer, for his guidance, patience, and support throughout this project. I would also like to thank Prof. Sabine Flitsch for co-supervising my project and providing me with access to the biological facilities; as well as Dr. Joanne Porter, for her assistance in samples preparation.

I am thoroughly grateful towards members of the surface analysis research centre group who shared their knowledge and experience with me, always being supportive and friendly. Special thanks to Emrys Jones, for his invaluable help with several steps of this project, that made its completion a reality.

Last but not least, I would like to thank old and new friends for being there for me, for they’ve always been able to cast a smile upon my face in spite of circumstance.
Chapter 1 Literature review

1.1 Introduction

The constant development of new technologies and instrumental advances has seen mass spectrometry benefit as previous analytical limitations, such as the confinement to small molecule analysis, are being overcome with new ionisation methods that offer softer, more efficient desorption processes, which larger molecules can now survive either intact or with much less fragmentation. Likewise, the development of the appropriate bioinformatic tools has made it possible to draw reliable conclusions by extracting relevant information from the large data sets generated that would be otherwise deceiving and too difficult to interpret. This has made biomolecular analysis feasible, and the study of complete biosystems has followed with the introduction of mass spectrometry as a tool in the fields of metabolomics, genomics, proteomics and lipidomics.

In biotechnology, the study of enzymatic conversions holds utmost importance since they not only they regulate the chemical processes required for life, but have also been implicated in industrial processes, pharmacology, and bioremediation. This has led to the use of a wide range of analytical techniques and workflows in the struggle to efficiently elucidate and characterise enzymatic activities. In multidisciplinary studies such as this, however, the methodology is normally dictated by the availability of resources where the data processing is critical, given the large data sets produced when biological samples are implicated. (Liesener and Karst, 2005)

1.2.1 Mass spectrometry

Mass spectrometry (MS) is based upon ion generation from samples that can be either organic or inorganic compounds, for which multiple methods exist (Figure 1.1), followed by the separation of those ions according to their mass-to-charge ratio \((m/z)\), and qualitative and quantitative detection in proportion to their abundance. This information is displayed in a mass spectrum, a two-dimensional representation that plots relative abundance (y-axis) against \(m/z\) ratio (x-axis).
In order to display it in a more comprehensive and comparable manner, the most intense peak in the mass spectrum, also called base peak, is normalised to a 100% relative abundance, with every other peak holding a proportional value, as percentage of it. (Gross, 2004)

With unequalled sensitivity, detection limits, speed and a broad range of applications, mass spectrometry has become an important analytical tool for scientists aiming to unravel biochemical problems found in fields such as microbiology, clinical diagnostics and environmental research. The development of new instrumentation with novel technology and workflows has seen mass spectrometry step up as a sensitive technique to be used for microbial sample analysis. (De Hoffmann and Stroobant, 2007, Sauer and Kliem, 2010, Fang and Dorrestein, 2014)

Figure 1.1: Different ionisation mechanisms used in mass spectrometry techniques. Reprinted with permission from (Watrous et al., 2011)

1.2.2 Mass spectrometry in microbiology

Catherine Clark Fenselau’s pioneering research in the 1970s demonstrated that MS analysis of microbes produces fingerprint type signatures, chemotypes for particular samples or microbial taxa, which can be correlated to a database of known species by identifying cultured bacteria. (Anhalt and Fenselau, 1975) The idea gave way to current platforms such as the Biotyper from Bruker and the VITEX MS from bioMérieux, both based on the acquisition of a molecular fingerprint such as the most abundant proteins present, for clinical microbiology identification. Cleared in 2013 by the Food and Drug Administration (FDA), these systems are likely to be widely spread in the clinical setting along other MS techniques in the near future due to the advantages they hold over biochemical tests and other traditional methods, being faster, cheaper and less labour intensive. (Fang and Dorrestein,
1.2.3 Mass spectrometry imaging

Mass spectrometry imaging (MSI) was demonstrated with secondary ion mass spectrometry (SIMS) in the 1960s for the first time; however, it has been implemented in microbiology only during the last decade. Even so, its utility in microbiology has already gone beyond microbial identification, enabling scientists to detect metabolites, proteins and lipids directly from microbial colonies. The introduction of revolutionary techniques and workflows in mass spectrometry meant a two-dimensional (2D) and three-dimensional (3D) visualisation of metabolites and lipids is now possible in a simultaneous and spatially resolved manner. This allows chemical hypotheses to be drawn based on the spatial mapping of the atomic and molecular information, which can then be correlated with the phenotypical output. (Fang and Dorrestein, 2014)

In an imaging experiment, the desorption and ionisation occurs at predefined positions controlled by a computer controlled xy-stage when a desorption probe hits the surface of the sample, dislodging secondary ions that are then guided into the mass spectrometer and towards the detector, generating a mass spectrum for each position, that once collected, will be shown as a single image with each specific mass displayed as a false-colour gradient where the signal intensity at each sampling location is indicated by the intensity of the colour, revealing the specific molecular distribution and relative abundance of the metabolites on the sample. Thousands of unique signals can be generated from a sample, and the molecular distribution zones in the images can help reveal the function of unknown compounds or cell structures. At present, no other current imaging method is able to reveal surface chemistry in a similar fashion. (Watrous and Dorrestein, 2011)

1.3 Matrix-assisted laser desorption ionisation mass spectrometry

Although this method was not employed in this study and therefore is not reviewed in depth, it is worth mentioning, given the success it has had in the field of mass spectrometry in recent years, especially in the analysis of biological samples, currently being the most extensively implemented MSI technique. As its name implies, it is based upon laser ablation of a matrix coated sample,
where the ultra-violet (UV) absorbing properties of the matrix, normally an organic acid, facilitate the desorption and ionisation of analytes.

At present, the microbial identification accuracy of matrix-assisted laser desorption ionisation mass spectrometry (MALDI) based instruments equals or even surpasses that from commercial biochemical identification tests, matching accordingly with the results from 16S rRNA sequencing, the current basis for microbiological taxonomy. With an analysis time of 30-60 s, microorganism isolation and culture remains the main limitation. (Strittmatter et al., 2013) It has also been applied for metabolite characterisation in cultured microbes, revealing aspects of biochemical processes that were previously not possible to track such as metabolite transfer within polymicrobial colonies and biofilms. (Fang and Dorrestein, 2014)

Nowadays, MALDI is regarded as a complementary technique to SIMS (see below) when it comes to the analysis of biological systems since it is capable of yielding a great amount of information on high-mass constituents, such as proteins and oligosaccharides while SIMS is particularly good revealing low mass molecules such as lipids and drugs. (Cassiday, 2008)

1.4.1 Secondary ion mass spectrometry

SIMS is an ultra-high vacuum technique in which highly energised primary particles impact the surface of the sample, ejecting secondary particles during the collisional process, called sputtering, which are guided into the instrument for analysis. (Armitage et al., 2013)

There are two main variants of the technique, dynamic SIMS and static SIMS, differing in the amount of primary ions directed at the sample, static SIMS uses a low flux so that each primary ion hits an undisturbed area of the sample, whereas dynamic SIMS floods the surface with primary ions to provide in-depth information. Static SIMS is the one employed in this study. (Fletcher et al., 2006)

Even though the highly energetic nature of the process limits the scope of the technique to lower mass ranges in the analysis of biological systems, it has been successfully employed to study membrane lipids at the single cell level, and to detect the lateral distribution of metabolites in microbial samples. (Watrous and Dorrestein, 2011, Fang and Dorrestein, 2014, Denbigh and Lockyer, 2015)
1.4.2 The basic equation

Two types of ions are present in a SIMS experiment: primary ions, which are used to sputter the surface of the sample, and secondary ions, that result from the sputtering process. While being ejected, the sputtered particles become ionised as a result of the collision process and the electron exchange in the surface. The electron exchange will depend on the electronic affinity and ionisation potentials of both the ejected particles and the surrounding sample material. This relation between the secondary ion yield and the electronic state of the analyte translates into quantitative analysis limitations.

In the basic SIMS equation, the secondary ion intensity (Coulombs/sec) of a particular analyte in the charged state $q$ is given by the equation:

Equation 1.1

$$I_m^q = I_p y_m \propto^q \theta_m \eta$$

Where:

- $I_m$ = Secondary ion current of species $m$
- $I_p$ = Primary ion current
- $y_m$ = Sputter yield of the chemistry $m$
- $\propto^q$ = Ionisation probability to ions of charge $q$
- $\theta_m$ = Fractional concentration of the chemistry $m$ in the surface layer
- $\eta$ = Transmission efficiency of the instrumental setup.

(Vickerman and Briggs, 2013).

1.4.3 The static limit

When a damaging atomic or small cluster ion beam (Ga+, Au+, etc.) is used in SIMS, if the number of primary ions impacting on the sample’s surface is too high, it can be damaged and rendered useless for analysis. The loss of molecular signal makes that structural and chemical damage to the sample evident, and it’s due to
the fragmentation of the molecules present in the sample’s surface. Therefore the static limit, a figure of $10^{12}$ ions per square centimetre, refers to the maximum ion dose to be used in SIMS analysis. Approaching the static limit during analysis will maximise the superficial data output, without accepting data from damaged parts of the sample’s surface. It does not apply, however, to fullerene or argon gas clusters, since damage produced by these beams is significantly less due to the cluster’s disintegration on impact and the subsequent energy dissipation into the few superficial monolayers, which prevents deep penetration. This allows for the use of an ion dose several orders higher during analysis with a $C_{60}$ beam, which is of particular importance due to the fact that the transmission of modern time-of-flight (ToF) instruments is close to optimum, meaning that signal intensity can only increase by increasing the ion dose and the ion yield. (Weibel et al., 2003) (Vickerman and Briggs, 2013)

1.5 Desorption electrospray ionisation

First reported in 2004 by Prof. Cook’s group, desorption electrospray ionisation (DESI) is a pneumatically assisted electrospray technique in which ions are generated externally to the mass spectrometer under ambient conditions. In this method, a charged solvent mixture is directed to the sample, causing desorption of the analyte species present at or near the surface. The hypothesised mechanism, described as droplet pick-up, involves the initial wetting of the sample surface, which generates an aqueous solvent layer that aids to analyte extraction and is followed by the emission of secondary micron sized droplets due to the splashing caused by collisions between droplets and the thin liquid film. (Ifa et al., 2010, Costa and Graham Cooks, 2008)

The secondary microdroplets containing the dissolved analyte are propelled towards a standard heated inlet capillary interface, where free gas-phase ions are produced by evaporation (Figure 1.2) (Wiseman et al., 2008).

Given that the desorption process of the analytes thoroughly depends on their dissolution efficiency, an appropriate solvent choice is paramount in order to achieve an optimal mass spectrum. The success of a particular solvent in DESI is therefore correlated with the solubility of a particular analyte, with methanol/water being employed as a standard solvent for polar molecules both in positive and negative ion modes. (Ifa et al., 2010)
The signal intensity observed will rely upon interdependent geometrical factors, such as the angle and distance between the sprayer and the surface, and between the surface and the mass spectrometer inlet. These factors are influenced as well by operating conditions such as the nebulising gas pressure and flow rate; therefore, optimisation of all these parameters is crucial in order to obtain good quality mass spectra for a particular sample. (Wiseman et al., 2008)

1.6 Rapid evaporative ionisation mass spectrometry

Recently developed by Takats’s group as a real-time, in vivo, intraoperative analytical approach for human tissue analysis, rapid evaporative ionisation mass spectrometry (REIMS) is based on a Joule-heating ionisation process generated by a surgical diathermy device, that causes the sample to undergo thermal disintegration by applying an alternating current at radio frequency (RF), producing an aerosol which is then analysed by a mass spectrometer. The device, named the iKnife, has achieved an accuracy of 92-100% for intraoperative tissue identification. (Golf et al., 2015)

REIMS is a novel technique when it comes to identification in the field of microbiology, but has already been successful in generating species-specific mass spectral fingerprints of intact bacterial cells analysed directly from solid culture.
medium, proving to be specific enough to be regarded as a potential tool for microorganism characterisation and identification in the clinical setting.

A characterisation and identification study by Strittmatter, the only one reported for REIMS involving clinically relevant microorganisms so far, yielded in percentages, 95.5%, 97.8% and 100% correct cross-validation results at species, genus and Gram level, respectively, as well as 87.3% at a strain-level, and a 98.8% correct classification of pathogenic yeasts species belonging to the same genus using leave-one-out cross validation (Figure 1.3). (Strittmatter et al., 2014)

![Figure 1.3: Cross-validation results accounting for the identification of 28 bacterial species by REIMS, as a function of the bin size on species, genus and Gram-stain levels. Reprinted with permission from (Strittmatter et al., 2014).](image)

The main advantages of the *modus operandi* include a short analysis time (2-3s), no solvent flow requirement; as an ambient technique, it requires no previous sample preparation steps and unlike pyrolysis systems, which also rely on a thermal process, it induces a soft ionisation which translates in a high molecular ion yield. The only disadvantage is the degree of uncertainty in the validation step, since the analysis is highly destructive, which means that the biomass producing a spectral output cannot be tested any further and identification is hence based on
interpolation of the environment of the evaporated biomass, although this does not represent a significant setback in the case of microbial identification. (Golf et al., 2015, Strittmatter et al., 2014)

Table 1.1: Specifications, advantages and disadvantages of the MS techniques.

<table>
<thead>
<tr>
<th>Method</th>
<th>Specifications</th>
<th>Advantages</th>
<th>Disadvantages</th>
</tr>
</thead>
<tbody>
<tr>
<td>Static SIMS</td>
<td>A primary beam hits the sample surface sputtering secondary ions, which are then analysed by the mass spectrometer.</td>
<td>High spatial resolution. Different primary ion sources available. Good for low mass molecules such as lipids and drugs.</td>
<td>Molecular fragmentation due to the energetic process limits the technique to low mass ranges. Sample must be evenly coated on a conductive surface.</td>
</tr>
<tr>
<td>MALDI</td>
<td>Laser ablation of a matrix coated sample, which facilitates desorption and ionisation of analytes.</td>
<td>Widest mass range available. Yields a great amount of information on high mass ranges.</td>
<td>Samples must be covered with organic matrix. Quantitative limitations. Lower mass ranges.</td>
</tr>
<tr>
<td>DESI</td>
<td>A charged solvent mixture is directed to the sample, causing the desorption of analytes present at the surface.</td>
<td>Soft ionisation. Solvent mixture can be optimised for different analytes and reactive imaging is possible. Atmospheric pressure analysis. Analysis at colony level possible.</td>
<td>Low spatial resolution. Quantitative limitations. The high-pressure gas can destroy samples. Many geometrical factors influence acquisition.</td>
</tr>
<tr>
<td>REIMS</td>
<td>A heating ionisation process caused by a diathermy device generates an aerosol that is directed to the spectrometer for analysis.</td>
<td>Has proved to be useful for tissue analysis. Produces bacterial fingerprints with from samples with no preparation steps at ambient conditions.</td>
<td>Uncertainty in validation steps due to the highly destructive nature of the analysis. Imaging is not possible.</td>
</tr>
</tbody>
</table>
1.7.1 Mass analysers

Once ions are generated from the sample, a mass analyser separates them according to their mass-to-charge ratio ($m/z$) by making use of electric and magnetic fields that can be static or dynamic, alone or in combination, with the way in which these are exerted varying between different types of mass analysers. A mass analyser suitable for all purposes has not yet been developed, with each type of analyser holding its own advantages and limitations and being based on a particular separation principle (Table 1.2).

<table>
<thead>
<tr>
<th>Type of analyser</th>
<th>Symbol</th>
<th>Principle of separation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Electric sector</td>
<td>E or ESA</td>
<td>Kinetic energy</td>
</tr>
<tr>
<td>Magnetic sector</td>
<td>B</td>
<td>Momentum</td>
</tr>
<tr>
<td>Quadrupole</td>
<td>Q</td>
<td>$m/z$ (trajectory stability)</td>
</tr>
<tr>
<td>Ion trap</td>
<td>IT</td>
<td>$m/z$ (resonance frequency)</td>
</tr>
<tr>
<td>Time-of-flight</td>
<td>TOF</td>
<td>Velocity (flight time)</td>
</tr>
<tr>
<td>Fourier transform ion cyclotron resonance</td>
<td>FTICR</td>
<td>$m/z$ (resonance frequency)</td>
</tr>
<tr>
<td>Fourier transform orbitrap</td>
<td>FT-OT</td>
<td>$m/z$ (resonance frequency)</td>
</tr>
</tbody>
</table>

Table 1.2: Mass analysers in mass spectrometry. Reprinted with permission from (De Hoffmann and Stroobant, 2007)

In general terms, their performance can be measured by 5 interrelated parameters, which must be considered when selecting a mass analyser for a specific application; these are the mass range, scan speed (time taken to scan a decade in mass), resolving power, mass accuracy, and transmission (ratio between the number of ions reaching the detector and the number of ions entering the mass analyser). Additionally, the ease of use with ancillary equipment or the prospect of tandem mass spectrometry (MS/MS) analysis should be taken into account if required. (De Hoffmann and Stroobant, 2007, Jennings and Dolnikowski, 1990)

1.7.2 Mass resolution

Mass resolution or resolving power, is a measurable attribute of mass analysers referring to their capability to generate discernible signals for ions with a minimal $m/z$ difference. It can be calculated using the following equation:
Equation 1.2

\[ R = \frac{m/z}{\Delta(m/z)} \]

Where:

- \( R \) = mass resolution
- \( m \) = mass of measured signal (peak)
- \( z \) = absolute charge number
- \( \Delta m \) = mass difference between two peaks at masses \( m \) and \( m+\Delta m \)

There are 3 different ways to define \( \Delta m \) (Figure 1.4), two of which are equivalent, the 10% valley definition, which involves two different peaks of the same height, and the 5% peak height; these are used for magnetic sensor or ion cyclotron resonance instruments. The width at half maximum (FWHM) definition is commonly employed in the case of Quadrupole MS, FT-ICR MS, Orbitrap MS and ToF MS.

![Figure 1.4: \( \Delta m \) measurement for each given definition. Reprinted with permission from (Muenster, 2009).](image)

The masking effect caused by isobaric matrix interferences encountered in samples of high chemical complexity, such as environmental or biological samples, hinders the detection of analytes present at low concentrations, making it necessary for a mass analyser to perform at a high resolution in order to yield high quality, information rich mass spectra (Figure 1.5). (Prohaska et al., 2015, Muenster, 2009)
With no exact reference value to describe the boundary between high and low resolution, a mass analyser is usually considered to perform at a high resolution when the R value is greater than 10 000 (FWHM). (De Hoffmann and Stroobant, 2007)

1.7.3 Mass accuracy

Mass accuracy is a measure of the instrument’s analytical deviation from a known exact mass. It can be expressed as a relative value, in parts per million (ppm), as a percentage, or as absolute units of u (mmu), the latter being more practical given that the calculated absolute mass accuracy in mass spectrometers tends to remain similar over a wide range. Sharp resolved peaks are a requirement if a reasonable mass accuracy is to be achieved; hence, mass resolution and mass accuracy are key considerations when considering an instrument for a particular experiment. (Gross, 2004)
1.7.4 Time-of-flight mass analysers

In a ToF mass analyser ions of different $m/z$ leaving the source, accelerated by an electric field are dispersed in a field-free drift path of known length, according to their velocities. Provided that all ions have the same kinetic energy, lighter ions hit the detector before heavier ones.

Advantages of ToF analysers include a theoretically unlimited upper mass range, high sensitivity, fast acquisition speed (<1ms for most $m/z$ values) and a simple design. An integrated reflectron improves the resolving power by compensating the influence of factors such as the spatial distribution and kinetic energy spread of ions. It is basically an electrostatic ion mirror located behind the field-free region opposite to the ion source that consists of a series of ring-shaped electrodes that generate a homogeneous electric field in the centre that deflects ions, sending them back through the flight tube in a slightly different angle. When ions of the same mass but different kinetic energy enter the reflectron, those with a higher kinetic energy will penetrate deeper into the reflectron than the ones a lower kinetic energy, and although their energy does not change, the length of their flight path increases in proportion to this difference, which means the ToF is increased so that ions of a similar mass reach the detector simultaneously. (Vickerman and Briggs, 2013, De Hoffmann and Stroobant, 2007, Jennings and Dolnikowski, 1990, Gross, 2004)

1.7.5 Quadrupole mass analysers

A quadrupole mass analyser exploits the stability of ions’ trajectories under oscillating electric fields to separate and filter them according to their $m/z$ ratios while traveling through it in the z-direction. It is made up of four parallel, circular or ideally hyperbolic shaped rods mounted in the xy-plane, with each pair of opposite rods held to the same potential, with one being a direct current (DC) time-independent component and the other an alternating (AC) time-dependent component, opposite in polarities (Figure 1.6). The potential to which ions are subjected is described by the equation:
Equation 1.3

\[
\Phi_0 = (U + V \cos \omega t)
\]

Where:
- \( \Phi_0 \) = applied potential
- \( U \) = DC voltage
- \( V \) = RF drive voltage
- \( \omega \) = frequency

Ions entering the assembly are attracted towards a rod with a polarity opposite to their own; a change in potential before they collide with the rod, generates attraction and repulsion forces that accelerate them in both the x- and y-planes, modifying their direction, since the sign of the electric force changes periodically in time. The nature of the applied potential of the electrode that operates in the x-z plane forms a high pass mass filter, that eliminates ions below a certain \( m/z \) value from the ion beam, whereas the conditions in the y-z plane are set to form a low pass mass filter that eliminates ions over a certain \( m/z \). Thus, in order to reach the detector, ions must comply with a \( m/z \) range requirement, defined by a given set of \( U \), \( V \), and \( \omega \), in order remain stable in both planes and avoid collision with the rods that would see them discharged and eliminated as neutrals. (De Hoffmann and Stroobant, 2007, Gross, 2004, Miller and Denton, 1986)

When a quadrupole is operated with a voltage \( U \) set to zero, it becomes a wide band pass for ions, focusing all ions within the transmission mass range in the centre of the rods. This has led to their use as ion guides and collision cells in what is known as RF-only mode and has been especially useful when an atmospheric pressure source is used and an efficient ion transfer from a functional part of a mass spectrometer to another with a minimal transmission loss is required. (De Hoffmann and Stroobant, 2007)
1.7.6.1 Hybrid instruments

In modern mass spectrometry, instruments that combine different mass analysers in a single instrument in order to exploit the virtues of each individual component have come to dominate the market. This study makes full use of such systems in order to take advantage of their versatility, along with the exploration of different ionisation methods, in an assorted range of instruments with different configurations.

1.7.6.2 Orthogonal acceleration time-of-flight mass analysers

Orthogonal acceleration time-of-flight (oa-ToF) instruments efficiently couple a continuous ionisation source to a ToF mass analyser in a setup that holds many advantages, such as a duty cycle within 5-50%, high transmission, high mass resolution, high acquisition rates, and high mass accuracy (up to 1ppm).

Once generated in the source in a continuous way, ions are focused into a parallel ion beam and guided towards an orthogonal accelerator that once filled with an ion package, exerts a voltage of 5-10kV, pushing ions orthogonally with respect to their original direction (x-axis) and into the ToF or ReToF analyser (y-axis). During the travel time, the ion beam refills the accelerator and once the heaviest m/z ion
reached the detector, the flight cycle ends and the process is repeated (Figure 1.7). (Gross, 2004)

Figure 1.7: Schematic illustrating the orthogonal acceleration process in a ReToF instrument. Modified from (Waters, 2013)

Quadrupole-oa-ReToF hybrids are currently the most common systems in the market, with oa-ToF instruments normally including an RF multipole ion guide in order to produce a stable beam independently of the quality of the beam generated by the source, enhancing the attributes of the system by decoupling the ion source and the mass analyser. A successful QqToF arrangement that performs optimally in both MS and MS/MS modes integrates a quadrupole analyser and a quadrupolar collision cell before the oa-ToF analyser. (De Hoffmann and Stroobant, 2007)

1.8.1 Cytochromes P450

The P450 terminology is based on a spectral feature rather than function as, unlike other cytochromes, an absorption band is displayed at 450 nm when bound to carbon monoxide in its reduced form, for which a phylogenetically conserved
cysteine thiolate ligand is responsible (Figure 1.8 A).

Cytochromes P450 (P450s or CYPs) are an ever-growing superfamily of proteins widely distributed in nature, being found in all three phylogenetic domains of life. They are thought of as monooxygenases, heme b-containing enzymes (Figure 1.8 B) responsible for a wide range of oxidative reactions of an enormous number of organic substrates that can be either endogenous compounds (e.g., steroids, fatty acids, pheromones, leukotrienes and prostaglandins) or xenobiotics (e.g., drugs, carcinogens). With many of them being able to catalyse multiple reactions, a nomenclature based on structural homology has been implemented for which 267 families with more than 5000 genes have been reported. (Hannemann et al., 2007, O’Reilly et al., 2013)

These enzymes catalyse the two-electron activation and scission of molecular oxygen that binds to a ferrous heme iron at its core, leading to the regio- and stereospecific insertion of one atom of oxygen into the substrate and reduction of the other to water, according to the equation: \( RH + 2e^- + 2H^+ \rightarrow ROH + H_2O \);

where \( R \) is the substrate, electrons are provided from reduced pyridine coenzymes NAD(P)H and protons are donated by active-site residues from amino acids conforming the proton relay mechanism. (McLean et al., 2005)

![Figure 1.8](image_url)

Figure 1.8: (A) Typical absorption spectrum for a CYP enzyme showing the 450nm band shift generated by the complex with CO alongside a main (Soret) band at ~420nm. (Munro et al., 2007) (B) Structure of heme prosthetic group iron protoporphyrin IX, heme b. (Sono et al., 1996)
Figure 1.9: Catalytic cycle of CYP monooxygenation. Substrate (S) is oxidised through steps 1-6 to form SO, with intermediates shown (A-F); uncoupled reactions lead to \( \text{H}_2\text{O}_2 \) formation pathways (steps 2a and 3a) while Compound I generation via the shunt pathway is also depicted (Steps 7 and 8).

For simplicity, the chemical bond (\( \text{Fe}^{\text{III}}\text{O} \)) is omitted, alongside the charge state of the iron atom and the porphyrin moiety, which is present in all intermediates but shown only for Compound I. Modified from (Hrycay and Bandiera, 2015)

The monooxygenation cycle begins when an organic substrate binds to a heme iron, with the dissociation of a weakly bound water ligand molecule trans to the cysteinate, shifting the spin-state equilibrium of the heme ferric iron which results in the increase of its reduction potential, favouring electron transfer from NAD(P)H through the reductase domain and further reduction the heme iron from \( \text{Fe}^{3+} \) to \( \text{Fe}^{2+} \). Ferrous heme iron then binds molecular oxygen, which undergoes activation and scission through a second electron transfer from NAD(P)H followed by two
consecutive protonation steps via a P450 proton relay system, resulting in the production of a porphyrin π radical ferril (Por•^{+}Fe^{IV}=O) intermediate known as Compound I species (Cpd I), the active oxidant to which the catalytical prowess of the system is attributed. CpdI effects the oxygenation of the substrate, with the release of the product and re-binding of a water molecule to the ferric iron, completing the cycle (Figure 1.9).

The importance of a phylogenetically conserved cysteine thiolate as a fifth ligand bound to the heme b iron must be emphasised since its reducing nature has a major impact in the enzyme’s catalytic activity. Appropriate binding of the substrate is necessary to start the catalytic cycle, as a switch on mechanism that allows for high substrate specificity of bacterial CYPs with uncoupled reactions normally resulting in H₂O₂ production. (Munro et al., 2007, Hrycay and Bandiera, 2015)

1.8.2 P450 RhF

Microbial CYPs have been classified based on the arrangement of the protein components involved in the electron transfer to the P450 enzyme, which have been discovered to be quite diverse, with ten classes currently assigned. CYP116B2 (P450 RhF) from Rhodococcus sp. Strain NCIMB 9784 is the first class VII P450 to be reported, with a unique primary structural organisation consisting of a N-terminal heme monooxygenase domain fused to a FMN-containing reductase with a [2Fe-2S] ferredoxin like centre C-terminal domain; it is synthesised as a single polypeptide with both domains linked via a short 21 amino acid region. (Fig. 1.10)

The reductase partner displays sequence conservation for dioxygenase reductase proteins (34%) and shows three functional parts: a FMN-binding motif, a NADPH-binding motif and a [2Fe-2S] ferredoxin domain. (Hannemann et al., 2007, Kelly and Kelly, 2013) It has become a valuable tool when the native electron partners for proteins are unknown, allowing catalytically active proteins to be obtained. The enzyme’s self-sufficient arrangement enables electron transfer to the P450 domain without the need for auxiliary reductase partners, as it is generally the case for prokaryotic CYPs.
Figure 1.10: Schematic representations of class VII P450 RhF. (A) Topographical arrangement of the fused protein, with N and C terminals shown and a NADPH binding motif implied. (B) Domain organisation based on sequence analysis with the heme-ligand pocket displayed. Reprinted with permission from (O’Reilly et al., 2013) and modified with permission from (Hunter et al., 2005) respectively.

Even though the physiological role remains unclear, it has proven to be a promising template for directed evolution due to its reported substrate promiscuity, being capable of catalysing the O-dealkylation, aromatic hydroxylation, olefin epoxidation and asymmetric sulfoxidation of a range of substituted aromatics, a rare attribute for wild type CYPs that might suggest a detoxification role. The oxidation’s regioselectivity is affected by subtle changes in the substrate’s structure.(O’Reilly et al., 2013)

1.8.3 P450 RhF origin

The P450 RhF gene was isolated from *Rhodococcus*, a bacterial genus of Gram positive, aerobic chemoorganotrophic bacteria found in a broad range of environments such as diverse soils, seawater and eukaryotic cells. Rhodococci oxidise chemical bonds present in organic compounds they use as a lone source of carbon and energy required for cell functions. Their powerful oxidative metabolism enables them to degrade complex molecules such as aromatic compounds and steroids through catabolic pathways that rely on mono- and dioxygenases. The abundant presence of these enzymes bestows an important catabolic versatility, which makes them highly valuable for biotransformations and bioremediation. This makes evident their contribution to the sustainment of the biosphere and explains
their presence in sites such as sludge from bioreactors for chemical waste treatment, from which they have been isolated. (Larkin et al., 2005, Alvarez, 2010)

1.9 Aim of the study

The main goal of this study is to effectively characterise the components of an enzymatic system of relevance in biotechnology through a methodology implicating different ionisation methods and instrumental configurations, highlighting their usefulness to accomplish the task of biological sample analysis. These components are the bacterial host, the expressed enzyme, as well as the substrate and product of its activity.

Bearing great significance is the development of an efficient workflow for the generation of chemical profiles from bacterial samples through the implementation of appropriate sampling techniques and experimental approaches, along with the use of bioinformatic resources designed to meet such purpose. A successful characterisation implicates the identification of distinctive signals for each of the species, as well as the generation of a lipid profile.

Given the nature of the selected system the experimental outcome is expected to hold meaning in other areas inherently related to it, with these being pharmacology and bioremediation, given the implication of diclofenac, which has been a massively produced drug since its discovery known to pollute water bodies due to it being difficult to degrade and an enzyme which hydroxylates the molecule, with this process being the basis of pollutant degradation. (Pérez-Estrada et al., 2005)
Chapter 2 Instrumentation

2.1.1 Ionoptika J105 – 3D Chemical Imager

A result of combining expertise in SIMS instrumentation design, novel mass spectrometry technology and experience in cluster beam SIMS analysis of organic samples, this mass spectrometer provides rapid data acquisition with simultaneous high spatial resolution and high mass resolution (up to 7 000 FWHM) by sampling the secondary ion beam into a two-stage time of flight system. The instrument is equipped with a fullerene source ($C_{60}^+$ projectile) and a gas cluster ion beam source, which can run in direct current mode or slow pulsed mode (if charge neutralisation is required). (Hill et al., 2011)

![Schematic of the J105-3D Chemical Imager](image)

Figure 2.1: Schematic indicating the main components of the J105-3D Chemical Imager. Reprinted with permission from (Fletcher and Vickerman, 2010).

The J105-3D Chemical Imager at The University of Manchester, illustrated in figure 2.1, was built with the collaboration from the Surface analysis research centre (SARC) group at The University of Manchester, Ionoptika Ltd. (Southampton, UK) and SAI Ltd. (Manchester, UK), which aimed to overcome the limitations in existing ToF-SIMS spectrometers owing to the requirement of short primary ion beam pulses that hindered the high spatial resolution capabilities inherent to the primary ion beam, along with the related low duty cycle that resulted in a long acquisition time.
A unique linear buncher consisting of 30 coupled apertures pulsed every 100 µs in conjunction with an harmonic reflectron ToF analyser (Figure 2.2) was developed as a solution to effectively decouple mass and spatial resolution using DC primary beams without an outstanding loss in transmission and mass resolution.

Figure 2.2: Schematic of the shaped field buncher and reflectron ToF mass analyser design. Reproduced with permission from (Hill et al., 2011).

The secondary ions produced by the impact of a direct current (DC) 40 keV C$_{60}^+$ primary ion beam into a sampling surface undergo a collision gas cooling process in a radio frequency-only, gas filled (N$_2$) quadrupole, before being filtered to a 1 eV energy spread by an electrostatic analyser and then injected into the shaped field buncher, which creates a time focus at the entrance of the ToF analyser through an accelerating field varying from 7kV to 1kV. The secondary ions will impact the detector after going through a nonlinear reflectron that corrects their kinetic energy dispersion. Thus, the full advantages of continuous primary ion beams are achieved alongside the inherent mass detection capabilities of a ToF analyser. (Fletcher et al., 2008, Hill et al., 2011)

The transmission of the buncher obeys the following equation:
Equation 2.1

\[
\text{transmission} = \frac{l}{t} \sqrt{\frac{m}{2Vq}} \times 100
\]

Where:
\( l \) = length of the buncher
\( V \) = transport voltage
\( t \) = buncher filling time

2.1.2 C\(_{60}\) ion source

A high performance tool for static SIMS analysis in the fields of biochemistry and polymer science, amongst others; the IOG-C60 coupled in the J105 instrument produces high mass primary ions named buckminsterfullerenes, which allow for a substantial yield increase of intact molecules and large fragments during the analysis of organic surfaces (Figure 2.3).

For this process, a charge of C\(_{60}\) powder is heated in small oven located in the rear of the source, where C\(_{60}\) vapour is produced by sublimation. The vapour passes into the centre of a chamber through a nozzle, where electron bombardment occurs as a result of the application of current to a thoriated tungsten filament, which causes the ionisation of the gaseous clusters. The C\(_{60}\) ions are then extracted along with other positively charged fragments and accelerated to full voltage (40kV) as they enter the optical column. The filtering of small carbon clusters is achieved by a Wein filter, a simple device capable of an effective \( m/z \) ratio separation of ions with a molecular weight up to 1000 amu. During operation, ions of similar mass with the same kinetic energy (since velocity is dependent on mass) pass through the filter undeflected, whereas ions of different masses are deflected from the optical axis.

Although the main species obtained tends to be C\(_{60}\)\(^+\), with a lesser amount of C\(_{60}\)\(^{++}\) and C\(_{60}\)\(^{+++}\), the energy of the ionising electron applied influences their formation ratio. (Vickerman and Briggs, 2013, Ionoptika, 2015)

The 40 kV C\(_{60}\) ion gun fitted in the J105-3D Chemical Imager is capable of emitting a beam with energy up to 120 KeV (for C\(_{60}\)\(^{+++}\)) and delivering a minimum spot size of 200nm.
2.1.3 Sample handling

Silicon wafers are the preferred substrate for SIMS samples, which are normally coated or imprinted onto them, as they are advantageous considering the characteristics of the ionisation process. Although other substrates can also be used for analysis, an optimal sputtering process is usually achieved with silicon wafers, with efficient energy coupling into the surface atomic layers and hence into surface-bound analytes, which translates into high signal intensity. Regardless of the chosen substrate, samples are mounted into a stub, tightened with bolts before being introduced to the mass spectrometer. The stub is 40mm long x 30mm wide; it is made of copper and has an array of holes for sample fastening. It also has a Faraday cup for ion beam current measurement beneath one of the holes (Figure 2.4). (Ionoptika, 2012)
The J105 instrument has a fully automated sample handling system that comprises 3 motorised modules controlled by a single software to allow the user to introduce the samples into sample analysis chamber (SAC), where the analysis takes place. The first one, named Z-lift, moves vertically inside the preparation chamber (PREP) and allows the user to put the stub holding the samples in and out of the high-vacuum, with a lock position that can be vented and serves for sample insertion and a lower tier position always kept under high-vacuum for sample storage. The second is a transfer arm that moves horizontally and takes the stub from PREP to SAC for analysis; it goes through a gate valve that protects SAC from venting. The third one is the XYZ-stage inside the SAC, through which the sample positioning is controlled along, as its name implies, the x, y and z-axis.

### 2.2 DESI 2D ion source

The DESI 2D ion source (Prosolia, USA) was coupled to a compatible mass spectrometer for DESI experiments. It consists of the sample stage surface, which can hold two 76 x 26 mm glass microscope slides for sample analysis, a spray head, a camera and a safety enclosure to avoid the spreading of potentially dangerous aerosols (Figure 2.5). The system is controlled by Prosolia’s Omni spray 2D motion software, which operates alongside the mass spectrometer’s software for a synchronised acquisition process. Geometrical settings related to the spray head can be manually optimised to allow for acquisition of good quality spectra. (Takáts et al., 2005)
2.3 iKnife

The iKnife diathermy device from Waters (Manchester, UK) was used with blade and needle tips for REIMS analysis of biological samples. It was coupled to a Xevo® G2-XS QTof mass spectrometer through plastic tubing and a pumped Venturi device for the analysis of the vapour produced from the samples. In REIMS, vapour is produced when an alternating current generated from an external power supply is applied via the tip of the iKnife onto the sample, and ions are produced from the vapor flow through interaction with a heated surface before they go through the mass spectrometer's ion optics. The workflow and components for the technique are illustrated in figure 2.6.
2.4 Xevo® G2-XS QTof

The Xevo® G2-XS QTof (Waters Corporation, Manchester, UK) is a hybrid instrument that combines a high stability quadrupole mass analyser with a high performance orthogonal acceleration time of flight (oaTof) mass analyser, in a setup that has been named QuanToF™, and yields fast, accurate, high-resolution results, with a mass range of $m/z$ 20 to 100 000 and a mass resolution up to 40 000 FWHM.
StepWave™ ion optics technology in the instrument enhances transmission while reducing contamination by eliminating neutral species, gas and non-desolvated material; whereas the XS Collision Cell technology, based on the application of a DC gradient by a segmented quadrupole, is responsible for MS/MS performance, an attribute explored in this study. (Figure 2.7)

Figure 2.7: Schematic of the ion optics in the Xevo G2-XS QTof high definition mass spectrometry system. Reprinted with permission from (Waters, 2014).

Different ion sources can be easily coupled to the instrument, in this study, the DESI 2D system (Prosolia, Inc. Indianapolis, IN, USA) was used for DESI experiments; while the iKnife™ and Venturi Device setup was used for REIMS analysis. Once ions arrive to the QuanTof detector, the signals produced are recorded over time by an analogue-to-digital converter (ADC), which enables accurate data to be obtained with a high dynamic range at acquisition rates up to 30 composite mass spectra per second. A further advantage of the analyser is that it is not affected by charge-capacity limitations, which normally hinder the performance of detectors in other kinds of mass analysers, especially when analysing complex samples.

MassLynx 4.1 and HDI v1.4 imaging software were used to control the instrumental parameters and setup the experiments, whereas the Omni Spray 2D control software was employed to control the DESI 2D ion source.
Chapter 3 Materials and methodology

3.1.1 Standard samples

Standard samples were analysed in order to optimise the experimental conditions for each and every technique employed in this study. The standards samples analysed, however, were not the same for all methods, and were chosen according to the workflow established for the particular instrument at the time.

For SIMS analysis, an indium coating in one of the corners of the stub was analysed as a standard whenever an assessment of the instrument’s status was deemed necessary. A biological sample consisting of fixed rat brain tissue over a glass slide was used as the standard for DESI analysis, which allowed for a comprehensive parameter optimisation given the nature of the samples to be analysed.

In the case of REIMS, an in-house workflow had been previously established involving the analysis of pork meat samples, yielding good quality spectra; therefore, these were used to test the instrument conditions.

Additionally to the standards already mentioned, as the purpose of the study involves the whole enzymatic system, the reaction’s substrate and the purified product were analysed in SIMS and DESI, in order to prove whether the techniques could resolve them, which would give not only a clue of the nature of these molecules, but also provide an insight into the potential of the techniques for analysis of biosystems with a biotechnological approach. This was not the case for REIMS given the nature of the process requiring them to be immersed in the biological matrix (bacterial colonies) in order to be readily detected, which was not possible due to the technical difficulties involved in the analysis of a running reaction.

3.1.2 Diclofenac

Diclofenac is a non-steroidal anti-inflammatory drug (NSAID) used to relieve pain and inflammation associated with multiple disorders. Synthesised with an acidity constant of 4.0, a partition coefficient of 13.4 and a structural arrangement including a phenylacetic acid group, a secondary amino group and a phenyl ring
containing chlorine atoms in its ortho positions that generate a maximum twisting of the ring, it meets the ideal attributes for an effective antirheumatic agent. (Sallmann, 1986)

It's mechanism of action involves the inhibition of prostaglandin synthesis through the nonselective binding of the drug molecule to both isoforms of the prostaglandin-endoperoxide synthase (PTGS) enzyme, commonly referred to as cyclooxygenase-1 (COX-1) and cyclooxygenase-2 (COX-2), which catalyses the conversion of arachidonic acid to prostaglandins, peripheral pain receptors that bind histamine and bradykinin to produce immediate pain and inflammation. (Williams and DuBois, 1996, Kantor, 1986)

Diclofenac sodium salt (2-[(2,6-Dichlorophenyl)amino]benzeneacetic acid sodium salt) (Figure 3.1) with a molecular weight of 318.13 and a solubility in water of 50mg/mL was purchased from Sigma-Aldrich for MS analysis. (Sigma)

![Figure 3.1: Diclofenac sodium salt. ChemDraw Professional (version 15.1.0.144) software.](image)

### 3.1.3 5-Hydroxydiclofenac

5-Hydroxydiclofenac, was kindly provided by Prof. Sabine Flitsch’s group at the Manchester Institute of Biotechnology (Manchester, UK), obtained through the scale up of the enzymatic reaction process in which the enzyme is expressed in a baffle flask at 20ºC and whole cells are used for a 24 hour-long biotransformation with diclofenac, which is followed by pelleting, acidification and supernatant extraction with ethyl acetate. The solvent is then evaporated and separated by high-performance liquid chromatography (HPLC) and the purity of the product is confirmed through nuclear magnetic resonance (NMR).
3.2.1 Biological system

A bacterial system was used for this study, which consisted of *E. coli* cells, genetically modified for CYP116B2 expression (described in section 1.8.2). The enzymatic reaction was set up using diclofenac sodium (Sigma-Aldrich, UK) as a substrate.

According to previous studies (Yan et al., 2016), 5-Hydroxydiclofenac was obtained as a lone catalytical product of the selective aromatic hydroxylation reaction of the P450 RhF system with diclofenac as a substrate, according to the following chemical equation:

![Diagram of P450 RhF monooxidation reaction over diclofenac to produce 5-hydroxydiclofenac. Relevant properties from both molecules displayed. The weights of the most abundant isotopes are shown (m/z values). Figures and information generated with ChemDraw Professional (version 15.1.0.144) software.]

The system is particularly good given that the expression of CYPs has not been reported to happen in any known wild type *E. coli* strains. This offers an exciting approach to the analysis and highlights the potential of the study to be applied to biotechnologically important biosystems.

3.2.2 Transformation protocol

The transformation of chemically competent BL21(DE3) *E. coli* cells was performed
as stated in the high efficiency transformation protocol C2527 by New England Biolabs (UK) Ltd. with modifications to steps 8 and 10 including the use of nitrocellulose membranes and different incubation conditions, based on previous experience on the system, according to which; a cryotube of competent BL21(DE3) *E.coli* cells stored at -80°C is thawed on ice for 10 min, followed by the addition of 1-5 µl containing 1 pg-100 ng of plasmid DNA to the cell mixture, which is flicked gently 4-5 times to mix cells and DNA. With no further mixing, the mixture is placed on ice for 30 min and then heat shocked in a water bath at 42°C for 10 s and placed on ice for 5 min. 950 µl of room temperature super optimal broth with catabolic repressor (SOC) is pipetted into the mixture, which is then placed in a shaker (250 rpm) at 37°C for 60 min. Selection plates are warmed up to 37°C and nitrocellulose membranes are placed on top of the agar. The cells are mixed thoroughly and 10-fold serial dilutions in SOC are prepared. 100 µl aliquots of the desired dilutions are pipetted and spread over the nitrocellulose membranes, to be incubated for 24h at 37°C. The nitrocellulose membranes are then aseptically transferred to IPTG plates, which are incubated for 24h at 20°C. Afterwards, the bacterial colonies are ready for analysis, they can as well be stored at 4°C for several weeks.

**Figure 3.3:** (A) *E. coli* transformed with an empty plasmid, thus lacking the RhF gene. Colonies are small with a bright slightly yellow colour. (B) RhF expressing *E. coli* colonies display a darker colour, a translucent yellow with a mild red tonality, due to the heme group in the enzyme.

A pET28a plasmid vector was used for the bacterial transformation, which carries a T7lac promoter, N-terminal His tag, a thrombin cleavage site, an internal T7 epitope
tag, a C-terminal His tag, and a kanamycin resistance gene in its construction. The RhF gene was inserted through restriction enzyme cloning for the transformation of RhF expressing cultures, while an empty vector was used for transformation in the case of non-expressing cultures for control samples. Isopropyl-β-D-thiogalactopyranoside (IPTG) is a molecular mimic of allolactose that activates the transcription of lac operon that was added to the lysogeny broth (LB) media with agar to induce protein expression. The concentrations for kanamycin and IPTG for the plates were of 50 µg/mL and 1 mM, respectively. (Hannig and Makrides, 1998)

3.2.3 Biological samples

All of the biological samples used in this study were of a bacterial nature and were handled according to the safety guidelines and protocols that apply for the system in particular. In this case a biosafety level 1 (BSL-1) suffices according to the pertinent risk assessment. The samples consisted of colonies of the transformed bacteria, both positive and negative for RhF expression.

3.3.1 SIMS sample preparation

Diclofenac and 5-hydroxydiclofenac were dissolved in water and methanol, respectively, up to a 8 mM concentration before being spin coated onto the silicon wafer substrates. Bacterial colonies were imprinted in silicon wafers for a set of experiments, and spread on them using a sterile aluminium spatula for another set of experiments in efforts of improving the signal intensity.

3.3.2 Substrate cleaning procedure

Pre-cut 5mm x 5mm low-resistivity silicon wafers with a 460-530 µm thickness, polished on one side were purchased from Agar Scientific (Essex, UK) to be used as substrates for SIMS analysis. The conductive nature of these substrates reduces up to some extent the excess charging effects caused by the primary ion beam on the sample surface, a beneficial attribute given recent studies reporting damage being caused by electron irradiation from flood guns used to neutralise those effects, over certain metabolites; with a lower secondary emission being the ultimate
consequence of their use. The cleanliness and purity offered by silicon wafers when compared to other substrates such as stainless steel or coated glass, as well as the flatness better than 1 µm of the whole surface, all contribute to obtaining an acceptable signal-to-noise (S/N) ratio as well as high intensity signals for metabolites. A low cost and ease of disposal are additional advantages. (Debois et al.)

In order to avoid sample contamination, a substrate cleaning protocol was used consisting of the sonication of the silicon wafers while immersed in the following solvents: Chloroform (Fisher Scientific UK Limited, analytical reagent grade), H₂O (HPLC grade Sigma-Aldrich, UK) and methanol (HPLC grade Sigma-Aldrich, UK). A Fischer Scientific FB15052 ultrasonic bath cleaner was used for this purpose with the order of the solvents being as follows: Chloroform (10 min), HPLC H₂O (10 min), Methanol (10 min), HPLC H₂O (4 hours).

### 3.3.3 Spin coating

The spin-coating technique is commonly employed in sample preparation routines for SIMS experiments due to the technique’s sensitivity towards surface irregularities, especially for depth profiling studies. Through its use, it is possible to obtain a thin and uniform film made of the sample of interest on top of the substrate. (Hall et al., 1998)

A WS-400BX-6NPP-Lite vacuum spin coater (Laurell Technologies Corporation, North Wales, USA) was used to prepare SIMS standard samples by drop casting 30 µL of the sample solution onto the polished side of clean silicon wafers with a micropipette in three 10 µL deposition steps, allowing some time to dry between droplets.

Static and dynamic dispensing modes, which only differ in whether the rotation is applied after or during the application of sample droplets in the surface of the substrate, respectively, were both used to prepare the samples. For those with water as a non-volatile solvent, the static mode was preferred, with a rotation speed of over 6000 rpm whereas a dynamic mode was used for volatile solvents, with lower rotations speeds of no more than 3000 rpm.
3.3.4 Desalting

A desalting step was required in order to obtain a decent signal for the diclofenac sample. It consisted in the washing of the wafer coated with the sample in a 1 mL ammonium formate (Sigma-Aldrich, UK) solution at a 100 mM concentration for 1-5 s. Alternatively, an aliquot of ammonium formate can be added to the sample solution before the spin coating, with similar results in efficiency. This protocol was established based on previous reports on the effectivity of ammonium formate for desalting in MS (Wang et al., 2011), and was not required for any other sample as a decent signal could be readily obtained.

3.4 SIMS experiment conditions

Samples were analysed with a $C_{60}^+$ beam in positive ion mode, exposed to an ion dose not exceeding a figure of $10^{12}$ ions/cm$^2$. A beam current between 25-40 pA, a shot rate of 10 kHz, 16 x 16 pixels and a raster area of 400 µm$^2$ raster size as a constant, modifying the number of shots per pixel and the duty cycle as necessary, within 200-500 and 50-100%, respectively. Spectra were acquired with a mass range of 100-1000 and a mass accuracy of 5 ppm.

3.5 DESI sample preparation

The rat brain tissue standard sample had been previously fixed onto the slide and was ready for its role in parameter optimization. For diclofenac and 5-hydroxydiclofenac samples, a spotting technique was used consisting of droplet deposition onto a glass slide of 1mg dissolved with 1 mL of a methanol-water mixture (75:25) (V/V). A piece of the nitrocellulose membrane over which the transformed bacteria had grown was cut and fixed with double-sided tape onto a glass slide for imaging analysis. Additionally, diclofenac and 5-hydroxydiclofenac were spotted onto a clean spot of the nitrocellulose membrane for analysis.

3.6 DESI experiment conditions

A methanol/water (75:25) (V/V) solvent mixture was sprayed onto the samples using a flow rate of 2 µL/min, a 6 bar nitrogen pressure and a capillary voltage of
2.5 – 3.5 kV. The sampling cone voltage of 79 V and temperature of 150ºC were kept constant. The incident angle was adjusted to around 45º with a collection angle around 10º. Data was acquired in both positive and negative ion modes with a resolution of 20000 FWHM (sensitivity mode) and a mass range of 100-1700. For MS/MS analysis, the collision energy was set between at 6.0 V.

3.7 REIMS analysis

The standard sample for REIMS assessment did not require any kind of preparation. For the bacterial analysis, the whole Petri Dish was used and individual colonies were selected for the analysis. Thus, this technique does not require a sample preparation step as such and is very time-efficient.

3.8 REIMS experiment conditions

Data was acquired with a mass range of 50-1700 in negative ion mode with 20000 FWHM resolution (sensitivity mode). The capillary voltage (2.6 kV), source temperature (120ºC), and desolvation temperature (600ºC) were kept constant. The iKnife was held to a power of around 20 W in bipolar mode.

3.9.1 Data processing

A major challenge when dealing with biological samples is to make the most of the large data sets generated, picking up important aspects of it and not being deceived by those not relevant into drawing inaccurate conclusions. Data pretreatment is necessary to prevent factors other than metabolite concentrations from having a major influence over the data analysis and thus hindering data interpretation. (van den Berg et al., 2006)

In this study, Matlab v7.8.0.347 (R2009a) (The MathWorks Inc., USA) was employed for the processing of imported data using in-house written scripts.

3.9.2 Data conversion

As a result of instrument developers constantly aiming to excel in a competitive
market and to overcome existing analytical limitations, a broad range of mass spectrometers has been produced with unique attributes that include not only design and performance, but data systems as well. This means the data output is different between instruments; encoded in a closed, vendor-specific format that complicates data analysis and comparison, especially when different systems are involved. In order to address this issue, the native binary data produced by mass spectrometers is converted to mzXML format, which is basically an open generic XML (extensible markup language) representation for MS, MS/MS or multiple mass spectrometric (MSn) data that serves as an interface that works between different mass spectrometers and data analysis pipelines. (Pedrioli et al., 2004)

Experimental data files were converted from their native RAW format through the MSConvert tool from the ProteoWizard version 3 software, an open access converter available for Water’s MassLynx v4.1 software, into an mzXML format. (Chambers et al., 2012, Kessner et al., 2008)

### 3.9.3 Data normalisation

In imaging mass spectrometry, normalisation is applied in order to correct for pixels with lower intensity due to the influence of factors other than analyte concentration; in DESI it might be caused by variations in the geometrical factors for example. In this study, total ion current (TIC) normalisation was applied on mzXML data in order to compensate for the effect of physical and chemical phenomena that might lead to signal intensity variation. In this method, the obtained spectra are multiplied by a single arbitrary variable or normalisation factor in order to equal the total amount of detected ions between them. As with other global methods, the normalisation factor is calculated by taking all of the spectral features in a simultaneous manner. (Haglund, 2008)

For PCA analysis, maximum normalisation was used. It ensures all spectra are on the same scale by summing all the intensities in a spectrum to express each peak as a proportion of the total with the highest peak set to 100%.

### 3.9.4 Principal components analysis

The use of advanced statistical methods is key to successfully achieve a systematical interpretation of SIMS data from biological surfaces, and a crucial step
in the analytical process due to the complexity of the spectral output. Among these methods, with an already established history and wide availability, principal components analysis (PCA) is the most applied multivariate statistical technique used for exploratory analysis in ToF-SIMS. It reduces the amount of data generated from the analysis whilst maintaining a minimal loss of information, making it possible to distinguish spectral features that account for the chemical diversity of the sample. The identification of peaks arising from similar chemical species, which aids in resolving matrix and topographic effects from surface chemistry variations, and the improving of the image contrast are additional advantages of this method. Hundreds of individual ion images are reduced to a few identified principal components (PCs), combinations of variables representing the highest source of variation in the whole data set.

The PCA algorithm takes the variance from each column of data accounting for an individual mass channel or range and determines the PCs, orthogonal lines of best fit through the variance measures. The highest degree of variance is described by the first principal component (PC1), the most significant variation. The second principal component (PC2), orthogonal to the direction of PC1, describes the highest degree of remaining variation, it is the second most significant variation in the data. This determination carries on, with the following PCs produced always decreasing in variance. PCA highlights differences between the samples in the first few PCs, with the rest of the PCs containing noise. The output from a PCA analysis will include loadings, scores and the percentage of the total variance described by each PC. As the factors are compulsively orthogonal, with each of them capturing the maximum amount of remaining variance, they will often lack a chemical or physical meaning apart from the statistical value they hold.

Data pre-treatment will influence PCA results; an appropriate scaling and alignment of the data will be required to avoid the loss of information from low-intensity secondary ions of analytical significance, which often arise from intact or nearly intact molecules, in the case of ToF-SIMS.

When the user has information available beforehand regarding the outcome of the analysis, e.g. if group classification is involved and therefore a match is expected, a canonical variates analysis should be employed to make the most of the information from the data set. (Vickerman and Briggs, 2013)
Chapter 4 Results and discussion

4.1.1 SIMS

4.1.2 Standard analysis

Indium was analysed in order to ensure the instrumental settings were optimal for analysis (Figure 4.1). An unusual signal ramp was observed in the spectrum, which does not hinder the analytical process and has to do with the buncher. As it is present with low intensity always near to the molecular ion, it can be disregarded.

Figure 4.1: Spectrum of indium standard (MW: 114.90) analysed in positive ion mode with 20 kV C$_{60}^+$ using an ion dose of 8.25E11 ions/cm$^2$. The signal intensity obtained is optimal, however, a ramp can be seen within the of 116-118 Da range in the 110-120 Da spectrum, due to a technical issue with the buncher plates.
4.1.3 Diclofenac and 5-hydroxydiclofenac

The analysis of both molecules was carried out following the analytical protocol established in the methods section. Surface Identity Software v1.0 (SurfaceSpectra Ltd.), which takes the molecular formula of the analytes of interest as an input, was used for the prediction of the isotopic pattern, as shown in figure 4.2 (Henderson et al., 2013). The expected signal could not be detected in the case of diclofenac due to a high fragmentation of the molecule that was not possible to correct with the reduction of the total ion dose applied nor with the application of charge compensation voltage. In the case of 5-hydroxydiclofenac, however, the molecular signal could be readily seen (Figure 4.4), albeit at a much lower signal intensity when compared to the molecular peak corresponding to a fragment.

A \[ \text{C}_{14}\text{H}_{11}\text{Cl}_2\text{NO}_2 \]

B \[ \text{C}_{14}\text{H}_{11}\text{Cl}_2\text{NO}_3 \]

Figure 4.2: Expected molecular signal for (A) diclofenac and (B) 5-hydroxydiclofenac, according to Surface Identity Software.
Figure 4.3: Diclofenac spectrum acquired in positive ion mode using the 20 kV C$_{60}^{+}$ ion source with an ion dose of 1.32E12 ions/cm$^2$. (A) The intact molecule is not detected in the spectrum. (B) Two main fragments peaks were assigned in the spectrum.
Two intense signals corresponding to main fragments were detected in the spectra at \( m/z \) 215 and 180, which result from a molecular loss of 81 and 116 u respectively, in accordance with the suggested fragmentation pattern reported in the literature (Galmier et al., 2005), corresponding to \([\text{M+H–H}_2\text{O–CO–Cl}]^+\) and \([\text{MH–H}_2\text{O–CO–2Cl}]^+\) (Figure 4.3). The two isotopes of the chlorinated compound were clearly resolved.

The fact that the signals of two other fragments known to have been detected with other MS methods at \( m/z \) 278 and 250 were not present in the spectrum suggests that the method is far too energetic for the integer molecule appear as such in the spectrum.

Figure 4.4: 5-Hydroxydiclofenac spectrum acquired using the 20 kV \( \text{C}_6\text{O}^+ \) ion source with an ion dose of 1.32E12 ions/cm\(^2\). The isotopic pattern displayed matches with the expected one.
4.1.4 Bacterial colonies

The analysis of bacterial samples was carried out using 20 kV $C_{60}^+$ in positive ion mode at different ion doses aiming to obtain a lipid profile. Only spectra with signals at low mass regions ($<400 \, m/z$) could be acquired (Figure 4.5).

Figure 4.5: Mass spectrum generated from RhF (A) and control E. coli (B) bacterial samples. Defined peaks only appear in the lower mass region. Imprinting technique.
4.2.1 DESI

4.2.2 Standard analysis

The optimisation of experimental parameters was carried out through the analysis of a previously fixed rat brain tissue, instead of relying on the gold standard for the DESI 2D-ion source, rhodamine, in order to optimise the experimental parameters in a more appropriate fashion, which takes into account the kind of samples used in this study. As a complex biological sample, the chemically fixed tissue was deemed more useful than rhodamine for the purpose, given the spectral similarities shared with the bacterial samples; these being the mass ranges in which high intensity signals are expected to show up (200-400 and 600-800 m/z), and most importantly, the fact that the complexity of both of them require the geometric parameters and other instrumental settings such as cone voltage to be actively adjusted.

Figure 4.6: Mass spectra acquired from rat brain tissue during parameter optimisation. Modifying the geometry of the sample with respect to the inlet, and instrumental settings such as voltage or temperature all contribute to obtaining good quality spectra with a low S/N ratio.
4.2.3 Bacterial Samples

The experiment was carried out using the conditions established as optimal according to the standard analysis. Good quality spectra were obtained with strong signals from the 600-800 \(m/z\) regions, where mostly lipid signals are expected to show up, as shown in Figure 4.7.

![Figure 4.7: DESI spectra from bacterial samples in negative ion mode where the peaks from phospholipids and fatty acids clearly dominate.](image)

4.2.4 Diclofenac and 5-hydroxydiclofenac

Both compounds were analysed over a glass slide substrate in negative ion mode. The use of the instrument’s MS/MS capability allowed the confirmation of the compound’s presence through fragment identification. The proposed fragmentation signature for diclofenac was confirmed with the further identification of a \(m/z\) 250 fragment (Figure 4.8 A).

For 5-hydroxydiclofenac, a fragment ion was identified at \(m/z\) 266.01 as a result of the loss of the carboxyl moiety \([M–H–COOH]^+\) (Figure 4.8 B). A further loss of a hydrogen atom results in a peak of 265.15 with the gain of a positive charge (Figure 4.8 B).
Figure 4.8: Diclofenac and 5-Hydroxydiclofenac spectra acquired from a glass surface. MS/MS capability was used to properly characterise the molecule and its fragments.
4.2.5 Imaging mass spectrometry

Once spectra were effectively acquired for each of the components previously mentioned, an imaging experiment was setup with all of the components of the biosystem of interest present on the surface of the substrate, in this case the nitrocellulose membrane over which transformed colonies grow. Characteristic analytes from each one of them were selected to generate a composite image reflecting the advantage of such a discrimination based on a false-colour image (Figure 4.9); RhF transformed colonies were imaged and compared against each other on the basis of the most abundant ion signal present on their respective spectra (Figure 4.10).

Figure 4.9: Imaging experiment revealing bacterial colonies (green), diclofenac (red) and 5-hydroxydiclofenac (blue) over the background signals (white), based on the abundance of distinctive signals selected for each of the species. Image generated and processed in HDI v1.4 software using TIC normalisation.
Figure 4.10: Bacterial colonies of control (red) and RhF *E. coli* samples (blue). Images processed using different colour gradients to show the analytes with highest maximum intensity for each sample. Normalised by TIC.
4.2.6 Data analysis

The next step was to generate regions of interest (ROIs) from the large imaging data set acquired in order to be able to process it through Matlab with the appropriate scripts generated in-house for data processing through normalisation and PCA. Control and RhF bacteria were the samples chosen for this, aiming to discern from one another through the PCA method, which would highlight the utility of the technique for high-resolution profiling due to the difference between them residing only in the expression of the CYP RhF.

The ROIs were exported from HDI software and converted into the mzXML generic format with a mass range of 100-1700 and a bin size of 0.01. It was then imported into Matlab using the appropriate scripts for the purpose. A composite spectrum is shown before normalisation and after maximum normalisation, taking into account the spectral features of both data sets (Figure 4.11). Once data was normalised, PCA analysis was run, Table 4.1 shows the captured variance calculated for 5 PCs, showing PC1 responsible for 93.73% of the variation. The loading plot for PC1 (Figure 4.12) permitted the identification of the signals to which the major differences between the groups can be attributed.

TIC normalisation was also tested for this purpose but the loading plot generated significantly minimised intense signals and allowed for the introduction of more noise in the spectrum, for this reason, maximum normalisation was carried out for the analysis as the output was more reliable, even when both of them showed clear separation between the groups when a scatter plot was generated.

The imported ROIs were separated in two groups consisting of nine ROIs from each bacterial kind. They were separated and labeled for the analysis, separated by colour; blue for RhF and red for control. A 3D scatter plot was generated through Matlab that reflects the separation of the components within the groups between PC1, PC2, and PC3 (Figure 4.13).

Once the groups had been visually separated through the scatter plot and the main sources of variation identified through the loading plot, the main peaks responsible for the differences were identified using resources such as metabolic databases and existing reports on the lipid signature of E. coli.
Figure 4.11: (A) Combined spectrum from selected ROIs from control and RhF bacterial samples. (B) Combined spectrum after maximum normalisation.
A

Table 4.1: Variance captured by PCA model for the ROIs after maximum normalisation.

<table>
<thead>
<tr>
<th>Principal Component Number</th>
<th>Eigenvalue of Cov(X)</th>
<th>% Variance Captured</th>
<th>% Variance Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>4.38e+000</td>
<td>93.73</td>
<td>93.73</td>
</tr>
<tr>
<td>2</td>
<td>2.62e-001</td>
<td>5.61</td>
<td>99.34</td>
</tr>
<tr>
<td>3</td>
<td>1.33e-002</td>
<td>0.29</td>
<td>99.63</td>
</tr>
<tr>
<td>4</td>
<td>8.72e-003</td>
<td>0.19</td>
<td>99.82</td>
</tr>
<tr>
<td>5</td>
<td>4.67e-003</td>
<td>0.10</td>
<td>99.92</td>
</tr>
</tbody>
</table>

B

Figure 4.12: Loading plot for PC1.
Ions corresponding to \( m/z \) values of 773.49, 747.48, 719.45, 688.45, 281.22, 255.21 and 253.20 were identified as major contributors to variation between the groups. Single charged species \([M–H]^-\) of phosphatidyglycerol (PG) with 36:2, 34:1 and 32:1 acyl chains were assigned to peaks at \( m/z \) 773.48, 747.48 and 719.45, respectively; the peak at 688.45 \( m/z \) corresponds to a phosphatidylethanolamine (PE) derivative with 32:1 acyl chains; whereas ions with lower \( m/z \) correspond to fatty acids (FA), with these being the carboxylic anions of octadecenoate (18:1, 281.22 \( m/z \)), palmitate (16:0, 255.21 \( m/z \)) and palmitoleate (16:1, 253.19 \( m/z \)) as shown in Table 4.2. This identification was made by comparison with results of previous studies in which wild type (WT) \( E. coli \) lipid extracts were analysed through liquid chromatography-mass spectrometry (LC-MS), using electrospray ionisation mass spectrometry (ESI-MS) operating in negative ion mode to analyse fractioned lipid extracts. Lipid species were identified by exact mass measurements and MS/MS. (Garrett et al., 2011, Oursel et al., 2007)
Table 4.2: Identity of lipids accounting for major variations registered in PC1. FA: fatty acids, PE: phosphatidylethanolamine, PG: phosphatidylglycerol. Assigned in accordance with previous lipid profiling studies of *E. coli*. (Garrett et al., 2011, Oursel et al., 2007)

<table>
<thead>
<tr>
<th>m/z</th>
<th>Identity</th>
</tr>
</thead>
<tbody>
<tr>
<td>255.21</td>
<td>Palmitic acid (FA C16:0)</td>
</tr>
<tr>
<td>281.22</td>
<td>Octadecenoic acid (FA C18:1)</td>
</tr>
</tbody>
</table>
| 688.45 | PE C15:0/cyC17:0  
PE C16:0/C16:1 |
| 719.45 | PG C14:0/C18:1  
PG C15:0/cyC17:0  
PG C16:0/C16:1 |
| 747.48 | PG C34:1 |
| 773.49 | PG C36:2 |

**Multiply charged species in DESI**

With closer examination to the mass spectral output, a peculiarity of the DESI technique could be observed as peaks separated by 0.333 m/z could be seen, which denotes the existence of a 3300 m/z analyte responsible for the signal. This holds advantages and disadvantages for analysis, as it means a more sensitive analysis can be done and high mass molecules can be effectively detected. However this also decreases the S/N ratio and makes the data harder to interpret. (Figure 4.14)
Figure 4.14: Zoom in of the negative ion DESI spectra from *E. coli* using shows differences of 0.333 m/z between peaks denote a charged state (+3 in this case) which implies the detection of an analyte of 3300 Da approximately.

The presence of multiply charged species has been reported as a spectral characteristic shared between electrospray methods such as DESI and ESI, even when the principles of both techniques hold differences such as the phase of the
sample (condensed phase for DESI and solution phase for ESI), with the production of multiply charged molecular species \([M+nH]^{n+}\) and \([M−nH]^{−}\) has been deemed a complicating factor given its dependence on experimental conditions. (Ifa et al., 2007, Keller et al., 2008)

### 4.3 REIMS

Two variants of the REIMS technique were explored for analysis as both needle and blade tips were tested with the iKnife on bacterial samples. The quality of the spectra obtained from the needle was lower than the one from the blade, which was similar to those obtained from DESI (Figure 4.15). While it might seem that REIMS introduces more noise into the spectra than DESI, in which the peaks are clearly resolved with high intensity signals, peaks related well to the lipid findings through DESI when compared, suggesting a small trade-off between speed of analysis and S/N ratio. However, some of the signals that could be regarded as noise carry important lipid information, such as peaks at \(m/z\) 391 and 417 are fragmentation products corresponding to \([M−H−74−RCH_{2}COOH_{\text{sn-2}}]^{−}\) and \([M−H−74−RCH_{2}COOH_{\text{sn-1}}]^{−}\) (Oursel et al., 2007).

As in previous studies regarding the lipid composition of biosystems such as bacteria, one main concern with the profile generated is whether the lipid composition for the samples analysed is related to the environmental conditions at the time of analysis, such as temperature, pH or availability of nutrients in the medium, and if so, which is very likely, up to what extent. While efforts are being made in workflows to enhance the reliability of lipid profiling techniques, it is possible that the relative changes in concentration could hold more meaning than the determination of their absolute quantities not only of lipids but other important metabolites related with cell function (Ivanova et al., 2009). It is also important to consider the oxidation and hydrolisis reactions over small molecules that take place during the ionisation process and are caused by the breakup of water molecules.
Figure 4.15: Spectral output for DESI (A) and REIMS (B) for the analysis of colonies belonging to the same cultures. Green spectra belong to control samples, red spectra belongs to RhF samples. Similarities can be clearly appreciated between the outputs of both techniques, as it would be expected.
Chapter 5 Conclusions and future work

SIMS DESI and REIMS mass spectrometry techniques were explored for an analysis of a biotechnologically relevant system. The main compounds in the enzymatic reaction were readily identified using SIMS and DESI, while a well defined lipid profile, in agreement with previous studies, was acquired for the system based on the expression of a desired molecule on an *E. coli* system through DESI and REIMS. For such purpose, both methods shared spectral similarities, as expected with both being ambient techniques; with REIMS proving to be even an equally reliable, faster technique albeit at the expense of the loss of the analysed material due to the nature of its ionisation principle. DESI holds the advantage of its high resolution imaging capabilities, which allow the user to view and manipulate output data rather easily.

The use of the negative ion mode for biological analysis with SIMS has previously shown promise for lipid profiling in the past, a reason for this to be explored in the future given the high fragmentation resulting from positive ion mode in the lipid regions of the spectrum. A combination of both ion modes could be an ideal alternative given the usefulness shown for drug analysis in positive ion mode. The use of the appropriate bioinformatic tools was critical for this analysis in order to discern between the bacterial samples used for this study. The clear separation between transformed bacteria with positive and negative expression of the RhF gene using PCA analysis might suggest a small variation in the concentrations of the analytes reflecting major statistic variance and thus is somehow related to the process. Loading plot for PC1 exposed peaks with *m/z* values of 773.49, 747.48, 719.45, 688.45, 281.22, 255.21 and 253.20 as the main source of variation between groups, peaks identified according to previous lipid profiling experiments reported and lipidomics databases and identified as PG, PE and FA derivatives.

A further step in the characterisation process would be to use MS/MS not only to identify the reaction’s substrate and product, but to confirm the nature of the lipids assigned; the development of a protocol using internal standards to assess quantitation with the technique would also be a study of interest as well as real time reaction monitoring.
References


Haglund, O. 2008. *Qualitative comparison of normalization approaches in MALDI-MS*. Master in Biomedical Engineering, Royal Institute of Technology.


