Imperial College London

Department of Metabolism, Nutrition, and Reproduction

Development of novel untargeted SPE-NMR protocols for metabolite annotation

Daniel George McGill

Thesis submitted for the degree of

Doctor of Philosophy at Imperial College London, December 2019

I hereby declare that this thesis and the work reported herein was composed by and originated entirely from me. Information derived from the published and unpublished work of others has been acknowledged in the text and references are given in the list of sources.

Dan McGill (2019)

Abstract

Metabolite identification and annotation procedures are necessary for the discovery of biomarkers indicative of diet, lifestyle, and disease. NMR spectroscopy remains a powerful tool in metabolic profiling — however, despite its reproducibility, ease of use, and quantitative data generation, its relatively limited sensitivity remains a bottleneck; low-concentration compounds become indistinguishable from noise, and the signals of more concentrated compounds regularly obscure those of less concentrated compounds. The aim of the project was to develop SPE-NMR protocols utilising different cartridge chemistries, using both natural and artificial urine mixtures, to produce unique retention profiles useful for metabolic profiling. These methods can then be incorporated as part of an analyst's 'toolbox' for metabolite identification and structural elucidation.

Different cartridge sorbents and conditions were utilised in order to produce unique retention profiles for different compound classes. Each elution demonstrated differing retention profiles for each method — replicates of the same method, however, had little difference between spectra, guaranteeing the reproducibility of the spectra. We found that application of the developed SPE methods to biofluids (such as urine) can be used to selectively retain metabolites based on compound taxonomy or other key functional groups, reducing peak overlap through concentration and fractionation of unknowns, and hence promising greater control over the metabolite annotation/identification process. Several NMR peaks demonstrating the presence of 3-hydroxyhippurate, unreported in literature or databases, were revealed through the use of the developed methods on human urine — a range of analytical methods were subsequently utilised in order to annotate the revealed peaks. Finally, standard operating procedures were written and validated for an automated SPE system, such that the SPE methods can be more generally used by researchers; this automated procedure was used to attempt to selectively retain specific glucuronide metabolites found in human urine, with retention being demonstrated where the moiety contains aromatic or otherwise significantly hydrophobic functional groups.

Acknowledgements

Some parts of this project required the direct contribution of others to come to fruition.

In this field, the terms 'known knowns', 'known unknowns', and 'unknown unknowns' (infamously popularised, but not coined, by former US Secretary of Defense Donald Rumsfeld in 2002 to describe the Johari window) are ubiquitous when speaking informally about certainty in metabolite identification. Rumsfeld is better known for promoting 'enhanced interrogation techniques' — torture — among broader violations of international law, as in the invasions of Iraq and Afghanistan, based on fraudulent information. The work detailed here is dedicated to his victims and survivors.

Contents

List of Tables

List of Figures

Nomenclature

B	Magnetic field strength
E	Energy
I	Nuclear spin
M	Magnetisation
MW	Molecular Weight
pw	Pulse width
sw	Spectral width
v	FID frequency

Greek Symbols

Θ	Pulse angle
δ	Chemical shift
γ	Gyromagnetic ratio
μ	Magnetic moment
σ	Shielding factor
τ	Pulse time

List of Abbreviations

2-PY	N-methyl-2-PYridone-5-carboxamide
CID	Collision-Induced Dissociation
CLASSY	CLuster Analysis Statistical SpectroscopY
COSY	COrrelation SpectroscopY
CV	Cruciferous Vegetable

DFI	Direct Flow Injection
DMA	DiMethylAmine
ESI	ElectroSpray Ionisation
FID	Free Induction Decay
FT	Fourier Transform
FT-ICR	Fourier Transform Ion Cyclotron Resonance
GC-MS	Gas Chromatography-Mass Spectrometry
HILIC	HydrophILc Interaction Chromatography
HLB	Hydrophilic-Lipophilic Balance
HMBC	Heteronuclear Multiple Bond Correlation spectroscopy
HMDB	Human Metabolome DataBase
HPHPA	3-(3-HydroxyPhenyl)-3-HydroxyPropionic Acid
HSQC	Heteronuclear Single Quantum Coherence spectroscopy
ICP-MS	Inductively Coupled Plasma-Mass Spectrometry
JRES	J-RESolved spectroscopy
LC-MS	Liquid Chromatography-Mass Spectrometry
MICE	Metabolite Identification Carbon Efficiency
MS	Mass Spectrometry
NMR	Nuclear Magnetic Resonance
NOE	Nuclear Overhauser Effect
NOESY	Nuclear Overhauser Effect SpectroscopY
NSAID	Non-Steroidal Anti-Inflammatory Drug
NUS	Non-Uniform Sampling
PAG	PhenylAcetylGlutamine
PBA	PhenylBoronic Acid
PCA	Principal Component Analysis
PEG	PolyEthylene Glycol
qToF	quadrupole-Time of Flight
RF	RadioFrequency

SAX	Strong Anionic eXchange
SCX	Strong Cationic eXchange
SHY	Statistical HeterospectroscopY
SMCSO	S-Methyl Cysteine SulfOxide
SPE	Solid Phase Extraction
STOCSY	Statistical TOtal Correlational SpectroscopY
STORM	SubseT Optimisation by Reference Matching
TML	N,N,N-TriMethylLysine
TMS	TriMethylSilane
TSP	TrimethylSilylPropanoic acid
TOCSY	TOtal Correlation Spectroscopy
ToF	Time of Flight
WAX	Weak Anionic eXchange
WCX	Weak Cationic eXchange
WEFT	Water Eliminated Fourier Transform

It ain't what you don't know that gets you into trouble. It's what you know for sure that just ain't so.

- The Big Short, 2015 (Misattributed to Mark Twain)

Chapter 1

Background

1.1 Metabolic profiling

Metabonomics was originally described as 'the quantitative measurement of the dynamic multiparametric metabolic response of living systems to pathophysiological stimuli or genetic modification'[1]. The primary target for analysis is the metabolome, a concept used to denote the sum total of small molecules which exist within a given biological system; the metabolome can be probed using a myriad of bacteria-, plant-, animal-, or human-derived samples, including but not limited to cell extracts[2,3], tissue[4,5], breath condensate[6], and biofluids such as urine[7], plasma[8], synovial fluid[9], etc.

Metabolic profiling is a wide-ranging field which aims to analyse a very diverse range of small molecules, often exploiting the differences between those compounds in order to aid identification — for example, the exploitation of compound polarity by reversed-phase chromatography allows for separation of those compounds, and hence generates identifying data about those molecules in the form of retention times, which can be directly compared between samples. Broadly, metabolic profiling can be split into 'targeted' and 'untargeted' strategies; whereas 'targeted' strategies start from a pre-defined hypothesis where the molecular target is already known and identified, 'untargeted' strategies aim to take complex mixtures and use analytical methods in order to generate hypotheses[10] — the most basic structure of these hypotheses taking the form 'under these (biological) conditions, the concentration of metabolite X in this sample will

increase/decrease'.

The broad aim of metabolic profiling is to discover biomarkers of disease, diet, or lifestyle, which are ultimately expressed chemically in the metabolome; these biomarkers can then be used to broaden understanding of metabolic pathways, propose drug targets, or make policy recommendations. Untargeted approaches will often utilise powerful multivariate statistical methods in order to extract signal from noise; a whole branch of statistical Nuclear Magnetic Resonance (NMR) spectroscopy methods — including but not limited to Statistical Total Correlation Spectroscopy (STOCSY)[11], Subset Optimisaton by Reference Matching (STORM)[12], Cluster Analysis Statistical Spectroscopy (CLASSY)[13], and Statistical Heterospectroscopy (SHY)[14] — have hence been developed to aid the observation of correlated spectral peaks in data, and consequently to aid biomarker discovery. However, statistical methods are not the sole area of development, and physical methods also constitute an active area of research; for example, the use of nanoparticles by Zhang *et al.* allowed for the suppression of signals from cationic and anionic small molecule metabolites within a complex mixture by specifically binding small molecules with charged silica nanoparticles, reducing the information density in a given sample[15]. Studying the metabolome requires the use of powerful analytical instruments, the most common being NMR[16] and Mass Spectrometry (MS)[17], including Liquid Chromatography-Mass Spectrometry (LC-MS)[18]. In the framework of LC-MS analysis, different column chemistries have been explored to ensure wide metabolome coverage of biofluids, and to take account of the physicochemical diversity of their components. Integration of NMR spectroscopy allows acquisition of complementary information for definitive structure elucidation and confirmation. These instruments are used for metabolite annotation (the putative identification of metabolites based on spectral similarity to literature or external data) and identification (confirmation of molecular identity based on 'a minimum of two independent and orthogonal datasets relative to an authentic reference standard')[19], as they provide orthogonal qualitative data, as well as absolute and relative quantification, in a very precise and high-throughput manner.

When identifying metabolites within a sample, there is often ample room for misinterpretation; within a complex natural mixture, this problem becomes far more apparent. Due to overlap in NMR spectra, confidence in analysis must be established and communicated in order to reliably

fit metabolites ostensibly present within a sample into a biomedical narrative. In 2007, as part of the Metabolomics Standards Initiative, the Chemical Analysis Working Group established a classification system and proposed minimum reporting standards[19] — the four tiers of this system are:

- Level 1: Identified compounds — a minimum of two independent and orthogonal datasets compared to data from an authentic reference standard under identical experimental conditions;

- Level 2: Putatively annotated compounds — data containing spectral similarity with a known compound found in a spectral library, or with characteristic physicochemical properties of a chemical class;

- Level 3: Putatively characterised compound classes — data containing spectral similatity with a compound class found in a spectral library;

- Level 4: Unknown compounds.

This classification system has found use in day-to-day descriptions of spectral peaks of interest (along with the more vague terms 'known knowns' (roughly analogous to level 1 identification), 'known unknowns' (level 2-3), and 'unknown unknowns'), but has seen limited uptake; the four-tier system does not contain the granularity in order to adequately describe identification in special cases (such as chemical isomers, which may have comparable spectral characteristics)[20]. Alternative systems of classification expanding into quantitative metrics have been proposed[21]. One alternative index is Metabolite Identification Carbon Efficiency (MICE)[22]; targeted towards NMR metabolite identification and relying on the reproducibility and stability of NMR data, the MICE index for a metabolite can be simply calculated by adding the total number of 'bits' of metabolite information and dividing by the number of heavy atoms in that metabolite[22] — the relevant 'bits' of information here includes chemical shifts, multiplicities, coupling constants, 1st/2nd order signal nature, signal half-bandwidth, signal integral, 2D NMR crosspeaks, and signal rate of change. Naturally, annotated metabolites which have more spectra associated with them (e.g 2D NMR spectra, on top of 1D spectra) have more associated bits, hence improving

confidence[22]. For 1D NMR alone, theoretical MICE values average around 1.8 to 2.2 bits per carbon atom, whereas the inclusion of 2D NMR spectra can raise this to 4.6 bits per carbon atom. Where the MICE index is ≥ 1, the compound can generally be described as confidently identified (level 1); where the MICE index is under 1, it is only putatively annotated (level 2 and below). This system allows greater clarity in communicating confidence of annotations compared to the simple four-tier system.

In practice, the use of NMR as a tool for metabolic profiling is ubiquitous. In one example, biomarkers differentiating diets high in cruciferous vegetables (CVs) from low CV diets were identified using various NMR techniques coupled with multivariate statistical modelling on human urine samples[23]. Four singlet peaks, representing four similar but chemically distinct metabolites related to the compound S-methyl-cysteine sulfoxide (SMCSO), were found to be 'diagnostic of CV consumption'. The use of ^{1}H NMR coupled with the multivariate statistical methods O-PLS-DA and PCA has also been used to identify biomarkers in urinary profiles of humans adhering to vegetarian diets, and those adhering to low- and high-meat diets[24]; one such compound characterised by a singlet peak at 3.11 ppm, N,N,N-trimethyllysine (TML), was found to have a negative loading coefficient in the urine of individuals adhering to meat-free diet. TML is formed from the hydrolysis of proteins with trimethylated lysines as constituents — this process occurs in most eukaryotes. It can be used to synthesise the compound L-carnitine; if dietary intake of L-carnitine is low, the body uses endogenous supplies to TML to compensate. As a result, in meat-free (and hence low-TML) diets, TML levels in urine are lowered as the body consumes it for production of L-carnitine. Other biomarkers of vegetarian diets include a ^{1}H singlet peak at 3.45 ppm (with additional doublets at 6.87 and 7.17 ppm), indicative of 4-hydroxyphenylacetate — this is formed by transamination of tyrosine by gut bacteria. The presence of 4-hydroxyphenylacetate in urine hence suggests a difference in gut microbiota between dietary groups, as the enzyme for this process is not produced by humans[24].

Urine has found widespread use in metabolic profiling due to its non-invasive ease of collection; however, the human urinary metabolome remains only partially mapped. This is in part because of the sheer number and dynamic range of compounds within a given urine sample – and, as a corollary of this, because of peak overlap that frustrates annotation efforts. In a high-

throughput NMR study of human urine, it was found by Bingol *et al.*[25] that, in a sample ^{13}C-^{1}H Heteronuclear Single Quantum Coherence (HSQC) spectrum, of the 1012 peaks detected, only 437 peaks (belonging to 98 individual metabolites) could be assigned. In 2013, Bouatra *et al.*[7] utilised a variety of analytical platforms (NMR, GC-MS, DFI/LC-MS/MS, ICP-MS, and HPLC) in order to identify 445 and quantify 378 unique urine metabolites, and a literature review led to the identification of an additional 2206 compounds. The Urine Metabolome Database (http://www.urinemetabolome.ca) at the time of writing counts 4267 total metabolites, of which only 1613 are 'detected and quantified', 452 are 'detected but not quantified', and 2202 are 'expected but not detected'. There is hence a need to expand the understanding of the human urinary metabolome in order to improve our capacity for characterization of population phenotypes, and for discovery of biomarkers related to disease and diet.

1.2 Nuclear Magnetic Resonance

The bulk of the NMR theory described in this section is described in the textbook 'High-resolution NMR techniques in organic chemistry' by Claridge[26].

Nuclear Magnetic Resonance (NMR) spectroscopy is a powerful technique in the characterisation of metabolites — the quantitative, non-destructive, and rapid processing of samples, as well as the inherent reproducibility of the data produced, can give researchers a huge amount of structural information about the constituent compounds of complex liquid (or solid, in 'magic angle' NMR) samples.

1.2.1 Theory

All identical nuclei have the same integral properties. Some nuclei exhibit angular momentum; the total angular momentum of a nucleus is determined by the number of constituent protons and neutrons, and is represented by the concept of nuclear spin. Nuclear spin I is determined by the spin interactions between the quarks within protons and neutrons, but the following rules are generally applicable:

- In nuclei with even numbers of both protons and neutrons (e.g. ^{12}C, ^{16}O), $I = 0$;

- In nuclei with odd numbers of both protons and neutrons (e.g. ^{14}N, ^{2}H), I is a positive integer;

- In nuclei where the number of protons and neutrons are odd/even or even/odd respectively (e.g. ^{1}H, ^{13}C), I is a half integral.

The nuclear spin I is related to the magnetic moment μ of the nucleus by the equation:

$$\mu = \gamma I$$

Where γ represents the gyromagnetic ratio, a constant which is dependent on the nucleus of a given atom. It is hence apparent that nuclei with zero spin do not produce a magnetic moment, whereas nuclei with spin do. The magnetic moment produced is parallel to the axis of rotation, as determined by the right-hand rule.

In the absence of a magnetic field, there is no net magnetisation of any given substance due to the random arrangement of magnetic dipoles contained within the nuclei of that substance. When placed within an external magnetic field, splitting of the ground state occurs; this is known as the Zeeman Effect, and the degeneracy of the energy levels can be described as $2I + 1$. The energy E of a specific energy level is dictated by the equation:

$$E = -\mu B_0$$

Where B_0 is the strength of the external magnetic field. As a result, splitting of the energy levels occurs; spins aligned with the external magnetic field are lower in energy, whereas spins aligned against the field are higher in energy. The distribution of a sample of nuclei across energy levels is described by a standard Boltzmann distribution, as the distance between energy levels is small, and hence can be affected by energy transfer due to thermal collisions.

In a typical NMR experiment — assuming that we are analysing a nucleus with I $= \frac{1}{2}$ — radio-frequency electromagnetic radiation near the Larmor frequency of the nuclei being probed is used to excite those nuclei from the ground state to the excited state; this is known as saturating the spin system. Relaxation from the first energy level back to the ground state

causes a signal to be emitted, named the Larmor Precession Frequency — or, more commonly, the Free Induction Decay (FID) — which induces a voltage within the NMR detection coil.

The frequency of the FID is determined by the resonance frequency of the nuclei being probed, which in turn is determined by the strength of the external magnetic field B_0 being applied — however, electrons exposed to an external magnetic field will generate their own local magnetic fields against the external field, partially 'shielding' nuclei. As a result, the effective field B_{eff} experienced by the nuclei is lower than the external field B_0 — and hence the resonance frequency is lower. It can be expressed using the equation:

$$B_{eff} = B_0(1 - \sigma)$$

where σ is the shielding factor. The shielding factor can be affected by factors such as electron density, magnetic anisotropy, and hydrogen bonding. Electron density is a measure of the number of electrons in the chemical environment and can be influenced by inductive effects, hyperconjugation, etc. Magnetic anisotropy (literally 'non-uniform magnetism') refers to opposing magnetic fields created by electron π-systems (such as benzene) in response to an external magnetic field. Hydrogen bonding refers to the predominantly electrostatic interaction felt between a hydrogen covalently bonded to a more electronegative atom, and another highly electronegative atom.

After detection, the FID is Fourier transformed. The multi-peak spectrum is defined in reference to a standard peak — such as that produced by tetramethylsilane (TMS) — using the equation:

$$\frac{(v_{sample} - v_{reference}) * 10^6}{v_{reference}} ppm$$

where v refers to the resonance frequency:

$$v = \frac{\gamma}{2\pi} B_{eff}$$

Hence a larger shielding effect σ from the electronic environment leads to a lower chemical shift δ (this is known as being more 'upfield'), whereas a smaller σ from the electronic environment

leads to a higher δ (more 'downfield'). As a result, due to the unique magnetic environments of each (non-chemically identical) nucleus within a compound, structural information can be ascertained.

B_{eff} can also be affected by spin-spin coupling, where the magnetic field experienced by nuclei is affected by neighbouring atoms of a differing chemical environment. For example, consider a hydrogen atom H_A neighboured by one other hydrogen atom H_B in a uniform magnetic field B_0 — the B_{eff} experienced by that H_A can either be greater (where H_B is aligned with B_0), hence causing a lower σ (deshielding H_A) and, consequently, a higher B_{eff} — hence shifting the H_A signal slightly downfield to a higher ppm. Alternatively, if H_B is aligned against the magnetic field B_0, the slight shielding effect (and lower B_{eff}) from H_B shifts the H_A signal slightly upfield to a lower ppm. As the number of nuclei populating each energy state is roughly equal, this is represented in the spectra by a doublet separated by a distance J.

Considering instead a hydrogen atom H_C neighbouring two hydrogen atoms H_{D1} and H_{D2} within a uniform magnetic field B_0, there are four possible outcomes with respect to spin-spin coupling:

1. H_{D1} and H_{D2} are both aligned with the magnetic field, shifting the H_C signal downfield;

2. H_{D1} is aligned with the magnetic field, but H_{D2} is aligned against the magnetic field;

3. H_{D1} is aligned against the magnetic field, but H_{D2} is aligned with the magnetic field;

4. H_{D1} and H_{D2} are both aligned against the magnetic field, shifting the H_C signal upfield;

Since outcome 2 and 3 produce identical effects while H_{D1} and H_{D2} exert the same effect on H_C (i.e, while they are chemically identical), this resolves as a triplet with peaks of relative intensity 1:2:1. This is an example of first-order splitting, which can be generally predicted by the '(n+1) rule', where n is the number of chemically equivalent neighbouring protons (table 1.1). Where H_{D1} and H_{D2} are not chemically identical (and hence produce different J couplings with H_C), second order splitting can be observed, producing multiplicities such as doublets of doublets or triplets of doublets, and hence not adhering to the (n+1) rule. The total integral of the multiplet itself is always proportional to the number of resonant nuclei — hence, chemically

identical hydrogen atoms on a methyl -CH$_3$ group will produce a signal of three times the intensity of a single -CHR$_2$ resonance.

Table 1.1: Tabular representation of the n+1 rule

$n = 0$							1						
$n = 1$						1		1					
$n = 2$					1		2		1				
$n = 3$				1		3		3		1			
$n = 4$			1		4		6		4		1		
$n = 5$		1		5		10		10		5		1	
$n = 6$	1		6		15		20		15		6		1

The firing of RadioFrequency (RF) waves at samples was originally done using Continuous Wave (CW) spectroscopy, where the frequency of the radiation was varied within the constant external field. This method is known as the 'frequency sweep'. Similarly, one could vary the external field while applying waves of a constant frequency — a 'field sweep'. Continuous wave spectroscopy was made antiquated by the introduction of Fourier Transform NMR (FT-NMR), which uses pulses of RF radiation in order to excite nuclei. These pulses can be hard (where the bandwidth v of the pulse is greater than the spectral width sw — hence the pulse excites the whole range of chemical shifts of one nucleus, e.g ^{1}H), or they can be soft ($v < sw$; soft pulses selectively excite certain chemical shifts).

The bandwidth of a pulse is determined by the pulse width pw — the duration of time which the pulse is active for. The bandwidth v of the pulse is determined by the equation:

$$v = \frac{1}{4\tau}$$

Where τ is the pulse time. Hence, shorter pulses are 'harder', and lead to a wider frequency range of radiation.

We can now imagine a sample being probed by NMR macroscopically. When the sample is placed in the external magnetic field B_0, the net magnetisation of the sample can be described by a vector parallel to that of the external magnetic field — this vector is the equilibrium magnetisation, M_0. Conventionally, the external magnetic field — and net magnetisation at rest — are parallel to the Z axis, hence the net magnetisation is described by the longitudinal

magnetisation M_z. This means that at equilibrium, $M_z = M_0$. If a sample were to be irradiated with RF energy to induce a 90° change (a '90° pulse'), the net magnetisation becomes 0 — the magnetisation vectors of the individual spin packets will also undergo precession, which (in a basic experiment) causes them to be be aligned in the XY plane. This alignment induces the spin packets to rotate around the Z axis at a frequency equal to the photon which caused the precession (this frequency is the Larmor frequency).

Pulses also have an angle Θ attached to them — this angle determines the direction in which the net magnetisation vector rotates. It is determined by the equation:

$$\Theta = 2\pi\gamma\tau B_1$$

Where τ is the pulse time and B_1 is the magnitude of the applied field. Pulse angles are commonly either 90° (such that the net magnetisation vector is in the Y axis, causing rotation and detection by the receiver coil) or 180° (the net magnetisation vector is in the $-Z$ axis).

Relaxation from this state occurs via two separate processes. The first process is Spin-Lattice relaxation; characterised by the Spin-Lattice relaxation time (T_1), it describes the loss of the energy gained by nuclear absorption of the RF pulse into the chemical lattice. Spin-Lattice relaxation can be thought of as affecting the length of the net magnetisation vector in the Z direction to its equilibrium state over time. The equation describing the relationship between T_1 and magnetisation is:

$$M_Z = M_0(1 - e^{-\frac{t}{T_1}})$$

The second process, Spin-Spin relaxation (characterised by Spin-Spin relaxation time, T_2) describes the increasing loss of coherence of the individual spin packets as they rotate around the Z axis, due to the return to equilibrium of the transverse magnetisation. As every spin packet exists within a unique environment, their individual Larmor frequencies are different — this causes a phase difference which is exacerbated by time. Hence while the net magnetisation is initially in the Y axis, it eventually becomes 0 in the M_{XY} plane. T_2 is always less than or equal to T_1. The equation describing the relationship between T_2 and magnetisation is:

$$M_{XY} = M_{XY_0} e^{-\frac{t}{T_2}}$$

With this understanding of the principles of NMR spectroscopy, we are now in a position to examine different experiments used to extract chemical information from a sample.

1.2.2 NMR experiments

One of the most basic ^{1}H NMR experiments imaginable involves a 90° hard pulse, followed by acquisition of the FID (fig 1.1).
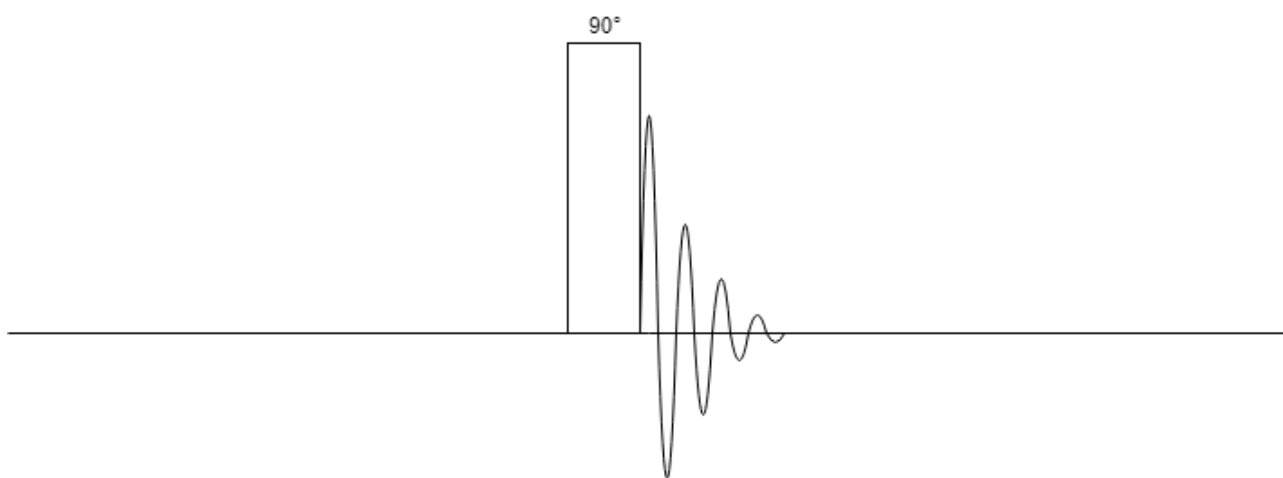
Figure 1.1: Pulse sequence diagram describing a single-pulse NMR experiment

In practice, this is rarely — if ever — used in metabolic profiling, as the amount of solvent (usually water) present usually results in suppression of small peaks within the spectrum by the solvent peak. Hence, solvent suppression is utilised in order to decrease the area of the spectrum taken up by the broader water signal[27]. This can be done by irradiating the sample with a low-power pulse for a long period, in a process called presaturation — however, this can interfere with the resonances of rapidly exchanging protons as well as the wider spectrum, leading to information loss.

One of the first solvent suppression approaches, Water Eliminated Fourier Transform (WEFT) NMR, incorporated one 180° hard pulse, mixing time, and a further 90° pulse, relying on the tendency of the solute to relax about 10-15 times faster than the solvent[28]. While this was appropriate when it was developed, advancements in the construction of more powerful

spectrometers has lead to this approach producing artifacts through radiation damping, a process whereby the solvent induces an electromagnetic field in the receiver coil[27]. The artifacts can be removed through the use of what has become the most common 1D ^{1}H NMR experiment for metabolic profiling[29]: the 1D-Nuclear Overhauser Effect SpectroscopY (NOESY) experiment, adapted from the 2D-NOESY experiment, utilises two 90° pulses instead of one 180° pulse (fig 1.2). Substitution of two 90° pulses allows phase cycling to remove artifacts from the final spectrum while minimising solvent signals; water suppression is also maintained by irradiation of the water frequency during the mixing time, leading in total to a much narrowed water peak. Later versions of this pulse sequence can also utilise a field gradient over a shorter mixing time.
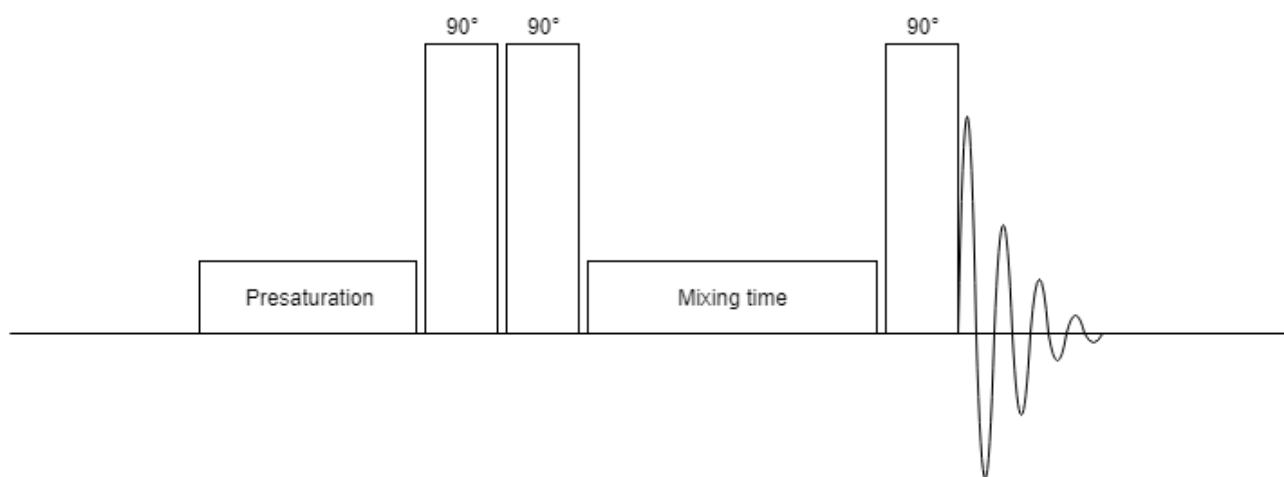
Figure 1.2: Pulse sequence diagram describing a 1D-NOESY experiment

1D ^{13}C NMR experiments, however, require other approaches. J-coupling between ^{1}H and ^{13}C causes unwanted splitting; since the ^{13}C experiment already produces much weaker signals (^{13}C has a prevalence approximately 99 times less than ^{1}H), the resulting multiplicity causes signals which are very difficult to interpret, both because of the complexity of the J-coupling, and of the weakness of the signal itself. To combat this, the ^{13}C atoms are decoupled from the protons by constantly irradiating the protons (nowadays this usually involves using a composite pulse, such as WALTZ-16, rather than a continuous wave) at a frequency higher than that of their Larmor frequency. Not only does this suppress proton J-coupling, it also causes the intensity of the carbon signals to be increased by up to four times via the Nuclear Overhauser Effect (NOE), whereby there is cross-relaxation of spatially close nuclei — however, this does mean that ^{13}C NMR peaks cannot be integrated like ^{1}H peaks, as the strength of the NOE will

vary amongst carbon nuclei.

2D NMR experiments are comprised of the standard 1D frequency scale as before, but with the addition of another dimension — 'true' 2D NMR incorporates a second frequency dimension, however some pseudo-2D methods can include other scalar couplings for the second dimension. The splitting caused by spin-spin coupling contains useful information about the chemical environment of a given hydrogen nucleus — however, in complex natural mixtures such as urine, peak overlap can lead to loss of information. The use of 2D spectroscopy — here, including the pseudo-2D method J-resolved spectroscopy (fig 1.3) — can hence be beneficial in retrieving this information:
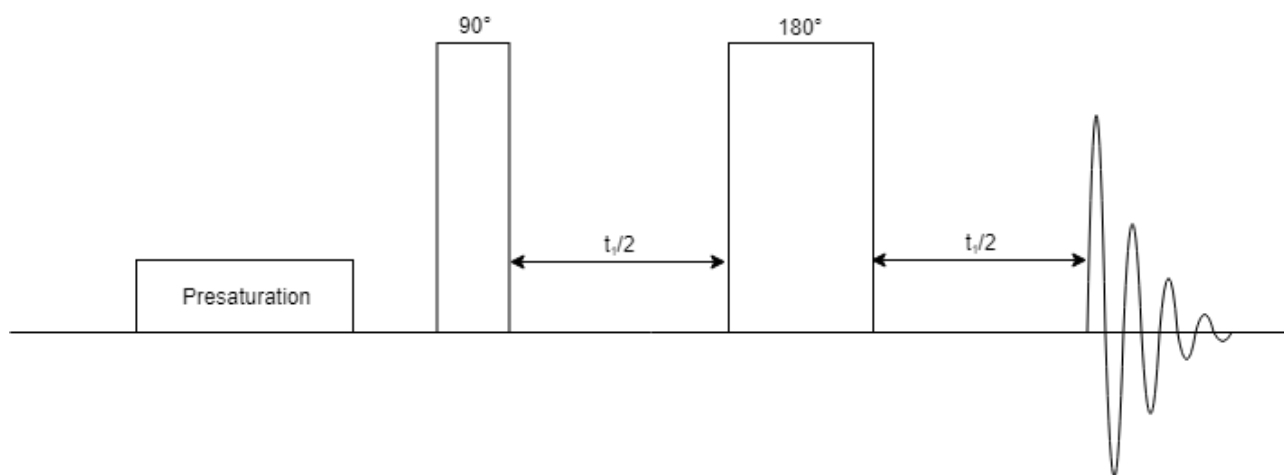
Figure 1.3: Pulse sequence diagram describing a J-resolved experiment.

Where t_1 describes the incremented evolution time.

The initial 90° pulse aligns the magnetisation alone the Y axis, where the nuclei begin to precess and lose coherency. After a specific time $t_1/2$, a second pulse causes a 180° flip of the spin packets in the Y axis. As a result, after a further $t_1/2$ seconds has passed, the signal regains coherency; this is known as a spin echo, which is detected by the spectrometer. This causes chemical shifts to refocus but couplings to continue evolving, hence causing shifts and couplings to emerge in F2, but only couplings in F1. After processing using a 45° tilt, the multiplet peaks are aligned in columns parallel to the f_1 dimension, hence allowing for facile analysis of the coupling patterns[30].

The 1D-NOESY and J-resolved spectroscopy experiments are the most common NMR experiments run, but the potential complexity of the biofluids being analysed means that information

extraction can be limited by peak overlap. In these cases, more broad ranging 2D through-bond correlation techniques can be used; the homonuclear COrrelation SpectroscopY (COSY, fig 1.4)[31] and TOtal Correlation SpectroscopY (TOCSY)[32], as well as the Heteronuclear Single-Quantum Correlation (HSQC)[33] and Heteronuclear Multiple-Bond Correlation spectroscopy[34] experiments. The different programs allow for unique information to be extracted from the sample; COSY reveals information about protons coupled to each other, while TOCSY reveals the links within an entire spin system. For the heteronuclear experiments, HSQC can be used to identify nuclei separated by one bond, whereas HMBC identifies nuclei separated by 2-4 bonds — this is usually applied (as it is here) to ^{13}C-^{1}H couplings, but can be used with other heteronuclei, such as ^{15}N and ^{31}P. Generally speaking, one single pulse program is not enough to give sufficient information for structural elucidation, even in simple samples — rather, a battery of experiments will be run in order to maximise the information available. These methods can be coupled with Non Uniform Sampling (NUS) approaches — usually based on linear prediction — in order to reduce experiment time and increase resolution[35].

The final approach used to deconvolute spectra — 1D selective TOCSY — involves using a shaped pulse to selectively excite a chosen resonance within the spectrum; tuning of the mixing time hence allows for greater or lesser magnetisation transfer between nuclei within the spin system[36]. There are a myriad of other pulse sequences devised in order to fit specific circumstances or requirements, or more generally to extract specific information from samples — these include pure shift, or HSQC-TOCSY — which are also described in the literature, but were not utilised in the course of the project.
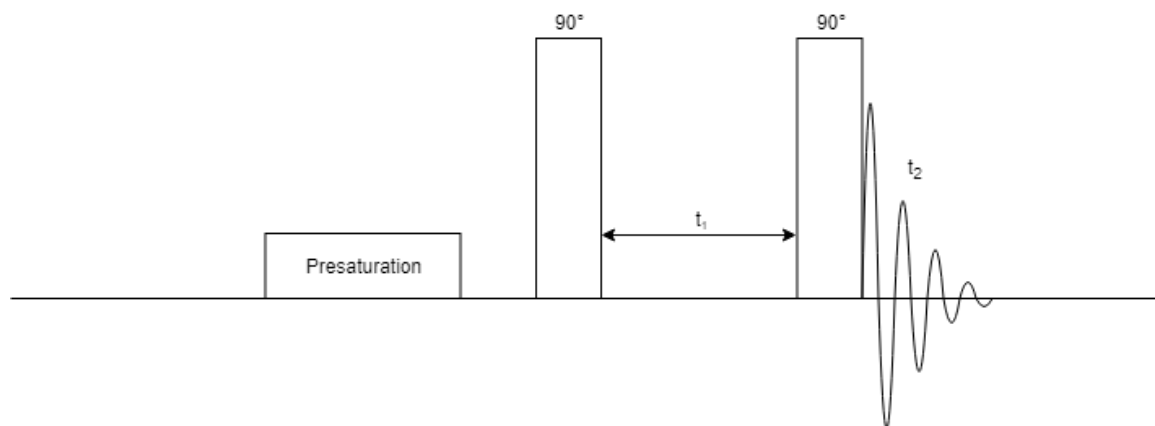
Figure 1.4: Pulse sequence diagram describing a COSY experiment

1.3   Chemometrics

The acquisition of multivariate data (such as that acquired through pursuing metabolic profiling of a cohort of samples) presents opportunities to extract a massive amount of information; however, the corollary of this wealth of information is the complexity of the data being produced. This is a classic problem resulting in an inevitable trade-off between sample authenticity and ease of analysis. In samples with reduced data, analysis becomes facile but limits the potential information available to the researcher; by contrast, samples with complex datasets assigned to them may be significantly harder to analyse, but may reap much greater insight. This philosophy has guided the entire project and indeed the wider -omics, but is particularly important in chemometrics.

Two primary statistical methods were used to great effect during the project: Principal Component Analysis (PCA), and STatistical COrrelational SpectroscopY (STOCSY). PCA is an 'unsupervised' technique — a technique which doesn't import knowledge about the samples being analysed (in contrast to supervised analysis, where samples are given pre-defined classes)[37] — which aims to explain variance within data using the smallest number of variables possible. First described in 1901[38], PCAs can be constructed to reduce the dimensionality of data and hence reduce visual representations of correlations between data points down to a given number of orthogonal 'principal components', which individually aim to represent as much of the data as possible. In this way, the first principal component PC1 accounts for the most variation in the data, the second PC2 accounts for the next most, and so on. A simple analogy for reducing dimensionality in this way describes the use of a flashlight to create 2D representations from the shadows cast by a 3D model (fig 1.5).

New graphs can hence be created with different principal components taking the place of the axes, in order to create a 'scores plot'. The related 'loadings' plot describes the importance of each variable to the principal components. PCA can be limited in some specific circumstances (while it is well placed to discovering orthogonal linear patterns, it may struggle with non-linear or non-orthogonal patterns), but overall its strength in being able to identify correlations in complex multivariate data makes it very worthwhile using.
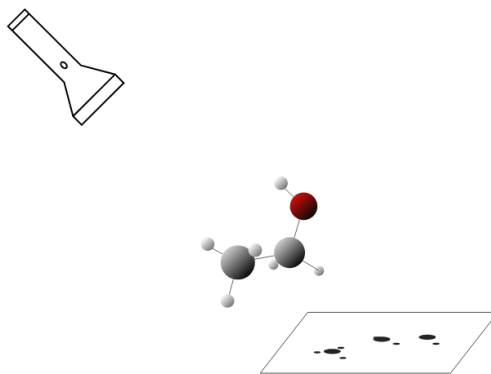
Figure 1.5: Diagram demonstrating the use of a torch to reduce the dimensionality of a 3D object.

STatistical COrrelational SpectroscopY (STOCSY) generates correlation matrices which suggest key metabolite peaks based on their covariance with a driver signal in a dataset[39,11]; the function relies on variations in intensity within a dataset due to natural changes in concentration between samples. Naturally, peaks belonging to the same molecule as the driver peak will correlate with each other; however, correlations may also be revealed with other metabolites which participate in other up- or downstream processes. Hence, statistical spectroscopic tools are particularly helpful in the analysis of complex mixtures, although must once again be selected based on the context of the experiment; for example, the difficulties STOCSY can face with overlapping peaks has encouraged the creation of alternative methods such as Subset OpTimisation by Reference Matching (STORM)[40], which also found some limited use during the project.

The use of different NMR experiments, as well as the statistical tools (here, primarily PCA and STOCSY) used to extract further information from them, allow for a broader understanding of 'unknown' metabolites. The structural information from 1D and 2D spectra can be coupled with inferences drawn from these methods, which can suggest the compound class of the unknown or its place within larger metabolic processes in order to give the analyst more control over the structural elucidation process. Other techniques like MS and LC-MS can additionally be used to gain relevant information about the compound being examined; the use of orthogonal techniques also increases certainty in identification, which can help in raising the confidence in an annotation from level 2 to level 1.

By examining the metabolites retained during solid phase extraction experiments (taking into account the cartridges chemistries through which these compounds are retained), further information can be deduced about the characteristics of the unknown metabolites. The use of solid phase extraction also has wider benefits for annotation of NMR spectra.

1.4   Solid Phase Extraction

Solid Phase Extraction (SPE) is a popular separation technique in analytical chemistry. In comparison to liquid-liquid extraction (a practice common to synthetic organic chemistry, used to separate compounds of differing polarity), SPE can be thought of as liquid-solid chromatography, in which a packed stationary column is used to separate out compounds within an aqueous complex mixture. The inherent chemical properties of the compounds in solution must be carefully considered in retention, but can also be taken advantage of for research purposes; for example, where the pH of a solution is higher than the pKa of a given metabolite within that solution, the metabolite will be ionised — this will consequently decrease retention on sorbents relying on hydrophobic forces (e.g reversed phase SPE cartridges), but will increase it for sorbents relying on ion-ion interactions (e.g ion exchange SPE cartridges). Similarly, the lipophilicity of a given metabolite — its solubility in non-polar solvents — will determine its retention through the hydrophobic force; the lipophilicity of a compound can be described by log P (where P, the partition coefficient, is the ratio of a solute dissolved in two immiscible solvents):

$$logP_{octanol/water} = log(\frac{[solute]_{octanol}}{[solute]_{water}})$$

Strictly speaking, the use of log k_w — the retention factor for pure water — provides the best prediction of retention for a compound on a reversed-phase cartridge; however, as log P is much more commonly recorded and is in a linear relationship with log k_w, it can be acceptable to use log P as an estimate for retention of compounds on reversed-phase cartridges[41]. In brief — within an analytical context, chemical manipulation of the constitutive metabolites allows for extraction, enrichment, and purification for further study.

1.4.1 Common sorbents

Most SPE sorbents are silica-based. Silica by itself is occasionally used as a sorbent in chemistry, such as in the use of a silica plug to isolate a compound with a high Retention Factor (R_f) from impurities with a low R_f. However, silica-based SPE cartridges are often made up of chemical modifications (such as C_{18} octadecyl chains, fig 1.6[42]) attached to a silica backbone. Varying these modifications allows for a different retention window for compounds in a solution, due to differences in intermolecular attractions stemming from the chemistries of the side-groups.
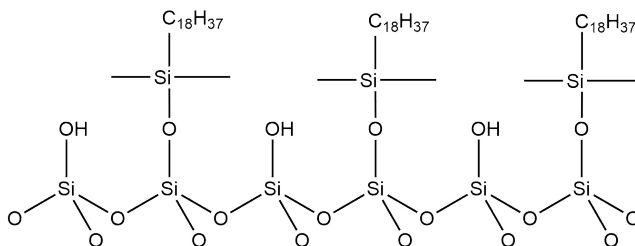
Figure 1.6: Silica-based C_{18} modification structure

Extractions are typically run in one of three phases:

- Reversed phase: a polar or moderately polar mobile phase is used with a non-polar stationary phase. The analytes to be retained will usually be non-polar to moderately polar. The relevant intermolecular forces are usually hydrophobic interactions: non-polar––non-polar interactions, and Van der Waals/Dispersion forces.

- Normal phase: a non-polar to moderately polar mobile phase is used with a polar stationary phase — hence the retained analytes are also polar. The relevant intermolecular forces are primarily hydrophilic: polar-polar interactions, hydrogen bonding, π-π interactions, δ-δ interactions, and dipole-induced dipole interactions.

- Ion exchange: Cationic and anionic compounds are retained — this primarily occurs through electrostatic interactions.

Silica-based SPE generally follows a general five step process, which can be exemplified with a reversed-phase method using a common C_{18} cartridge[43].

- Conditioning: When first encountered, the octadecyl chains of the cartridge are collapsed, resulting in a very low surface area for retention — there will also often be impurities present from the manufacturing process. To counter this, the sorbent is first conditioned using a suitable solvent such as methanol or acetonitrile; this causes the chains to 'float' freely, increasing the surface area for retention available, and eliminating some of the present impurities.

- Equilibration: Water or aqueous buffer is then eluted through the cartridge in an equilibration step — interstitial conditioning solvent often remains near the base of the octadecyl chains, but is otherwise replaced by the equilibrant. This allows solvation of the compounds in the sample such that retention can occur.

- Load: The sample is applied to the cartridge — the compound or compounds of interest are retained onto the sorbent, while impurities and analytes of the same polarity as the mobile phase pass through.

- Wash: The cartridge, with compounds of interest now bound to the sorbent, is washed with aqueous solvent to remove the remaining unretained impurities.

- Elution: The analytes of interest are eluted from the cartridge using an appropriate solvent or solvents.

Different silica modifications lend themselves to different applications. For example, two of the most common modifications are silica-based C_{18} and C_8 cartridges. These saturated chains provide points for hydrophobic interaction, while the remaining silanol SiOH groups keep some polar and ion exchange character. Endcapping — the replacement of the silanol groups with trimethylsilyl groups — can eliminate this polar/ion exchange character entirely, if very non-polar analytes are of interest (fig 1.7). A list of intermolecular forces and their relative strengths are detailed in table 1.2[44,45].
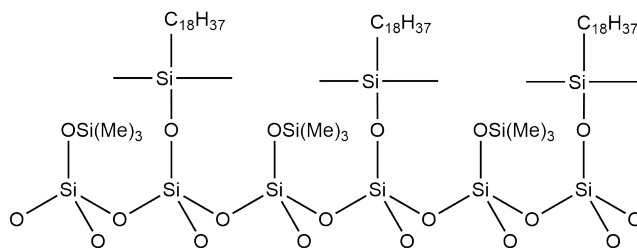
Figure 1.7: C_{18} stationary phase with endcapping

In addition to the simple hydrophobic effects common to the saturated chain modifications like C_{18} and C_8, π-π interactions in phenyl modifications (fig 1.8) allow for the phenyl ring to selectively attract aromatic and other conjugated analytes. These π interactions can vary in strength — for example, the strength of the π stacking bond has been shown to be bound to 8-12 kJ/mol[46].

The C_{18}, C_8, and phenyl modifications are all examples of reversed phase sorbents, with a polar mobile phase and a non-polar stationary phase. By comparison, the normal phase diol modification (which utilises a polar stationary phase with a non-polar mobile phase) primarily relies on strong hydrogen bonding between the hydroxyl groups of the sorbent and relevant functional groups on the analyte for retention — as well as π-π interactions, δ-δ interactions, δ-induced δ interactions, etc (table 1.2)[47]. The hydrocarbon spacer between the dimethylsilica group and the hydroxyl groups adds non-polar character, useful for the retention of analytes with hydrophobic functional groups (fig 1.9).
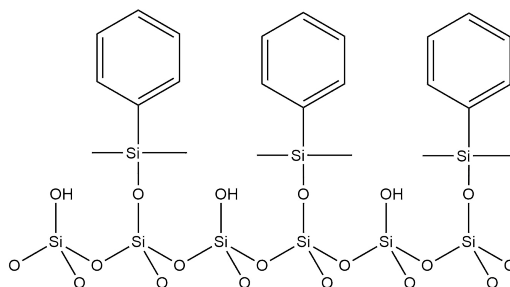
Figure 1.8: Silica-based phenyl modification structure

Table 1.2: Summary of intermolecular forces

Interaction	Cause	Strength Range (kJ/mol)
Ion-ion bond	"Attractive force between atoms with sharply different electronegativities" (IUPAC)	400-4000
Covalent bond	Attractive force between nuclei caused by sharing of electrons	200-1000
Metallic interaction	"Attractive force between conduction electrons and positively charged metal ions" (IUPAC)	100-1000
Ion-δ interaction	"Attractive force between charged molecule and permanent dipole (IUPAC)"	40-600
Hydrogen bond	"Attractive interaction between a hydrogen atom from a molecule or molecule fragment X-H in which X is more electronegative than H, and an atom or a group of atoms in the same or a different molecule, in which there is evidence of bond formation" (IUPAC)	10-40[48]
δ-δ Interactions (Keesom force)	Van der Waals force	5-25
Ion-induced δ interactions	Van der Waals force	3-15
δ-induced δ interactions (Debye force)	Attractive force between charged molecule and induced dipole	2-10
Hydrophobic force	Entropic effect from disruption of hydrogen bonds	<40
Induced $\delta-$ induced δ interactions (London Dispersion)	"Attractive force between apolar molecules" (IUPAC) — Van der Waals force	0.05-40
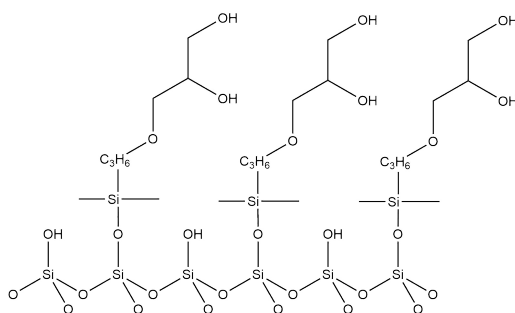
Figure 1.9: Silica-based diol modification structure

Some columns can be run in either normal or reversed phase, affording different results depending on the experimental conditions used. One such example is the CN modification (fig 1.10) — while typicaly used in normal phase applications, it can also be used in reversed phase to capture non-polar analytes which would otherwise irreversibly bind to the C18 cartridge. Like the diol modification, the CN modification allows access to the silanol groups — this adds

additional polar character to the column when in the reversed phase. Another sorbent which can be utilised in different phases is the aminopropyl 'NH$_2$' modification (fig 1.11) — this can be run either using normal phase conditions, or as a Weak Anion eXchanger (WAX). Similar in structure to the CN modification, the proximity of the NH$_2$ group to the silica base allows for additional polar character from the silanol groups. This allows for the retention of moderately polar compounds in the normal phase.
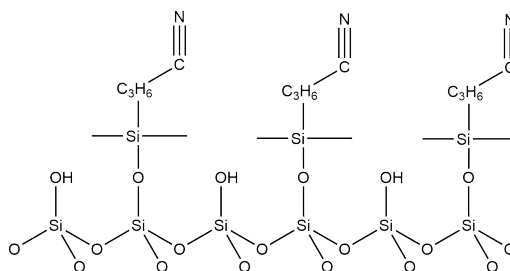
Figure 1.10: Silica-based CN modification structure
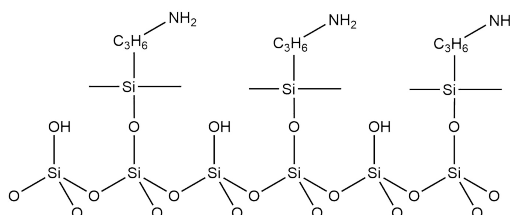
Figure 1.11: Silica-based NH$_2$ modification (WAX) structure

For use in ion exchange, the NH$_2$ sorbent must be kept at least two pH units below 9.8 — the pKa of the primary amine (fig 1.12). The pH used must also be at least two pH units above the pKa of an anionic analyte, such that both are charged. The electrostatic interaction between the charged molecules provides the retention capability.
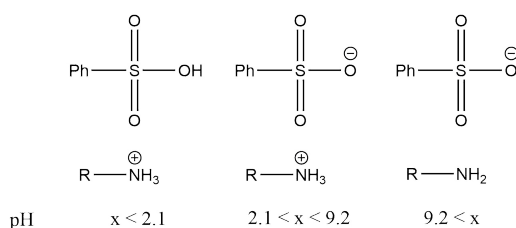
Figure 1.12: Effects of pH on charge of analyte (benzenesulfonic acid) and WAX sorbent

The primary amine group provides a weaker attraction to acids — which are positively

charged at physiological pH — than a quaternary amine group (which has a permanent positive charge). Due to the ability to capture all acids over a wider pH range, the quaternary amine modification is usually designated as a Strong Anionic eXchanger (SAX), by comparison with the aminopropyl modification as a Weak Anionic eXchanger (WAX).
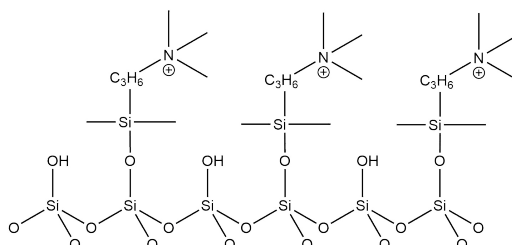
Figure 1.13: Silica-based quaternary amine (SAX) modification structure

A similar comparison can be made for the cationic exchange columns. The Weak Cationic eXchanger (WCX) sorbent depends on a carboxylic acid 'COOH' modification — hence requiring a certain pH window such that the carboxylic acid is deprotonated, while the desired (basic) analyte is protonated. By contrast, the Strong Cationic Exchange (SCX) column utilises a phenylsulfonic acid derivative with a very low pKa ($\approx$2.1) — it can hence be considered pseudo-permanently charged, as the operational pH range is unlikely to be below its pKa. This allows it to selectively retain most bases.
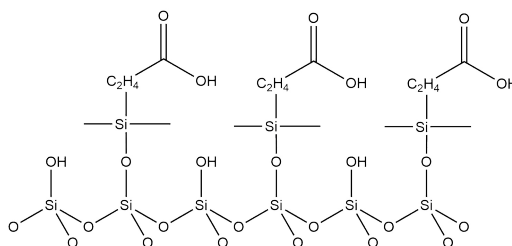
Figure 1.14: Silica-based carboxylic acid (WCX) modification structure
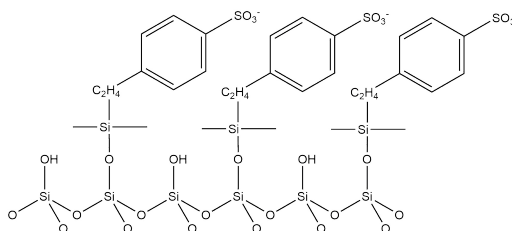
Figure 1.15: Silica-based benzenesulfonic acid (SCX) modification structure

There are two additional cartridges with unique, non-silica based chemistries worth examining. The first utilises a Hydrophilic-Lipophilic Balance 'HLB' copolymer sorbent (fig 1.16): formed from the conjugation of hydrophobic divinylbenzene and hydrophilic pyrrolidinone monomers, the juxtaposition of these two molecules utilising different interactions allows for simultaneous extraction of both polar and non-polar analytes[49]. One advantage of using a HLB cartridge over a standard C_{18} phase for sample cleanup is its tolerance to drying out. Drying on silica-based cartridges can negatively affect the recovery of analytes — however, the alternating hydrophilic-lipophilic units (as opposed to the lipophilic hydrocarbon chains of a C_{18} cartridge) allows the HLB cartridge to remain adequately wetted. The absence of the silica backbone — and corresponding silanols — also removes ion exchange character from the sorbent.
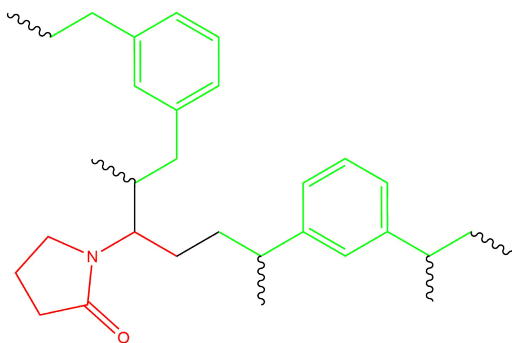
Figure 1.16: HLB copolymer (red: pyrrolidinone, green: divinylbenzene)

The final cartridge uses a unique reversible covalent approach to retain compounds with two heteroatoms located vicinially to each other (such as a diol or α-hydroxy ketone). The PhenylBoronic Acid (PBA) sorbent (fig 1.17)[50], when subjected to basic conditions, forms a reactive boronate. Retention of a diol analyte proceeds via an organic reaction, resulting in the loss of water — the analytes can then be eluted through acidification of the boronate column. Additional secondary interactions — hydrophobic interactions from the benzene ring, ionic interactions from the charged boron atom, and hydrogen bonding from the attached hydroxyl groups — are also present. This can cause additional unwanted analytes to be retained, such as through ionic bonding of the boronate and a cation. In addition, some compounds (such as α-hydroxycarboxylic acids or β-hydroxyamines) can react covalently — sometimes irreversibly — with the boronate. It is hence important to manage the sample pH carefully and to use appropriate buffers, such as phosphate buffer.
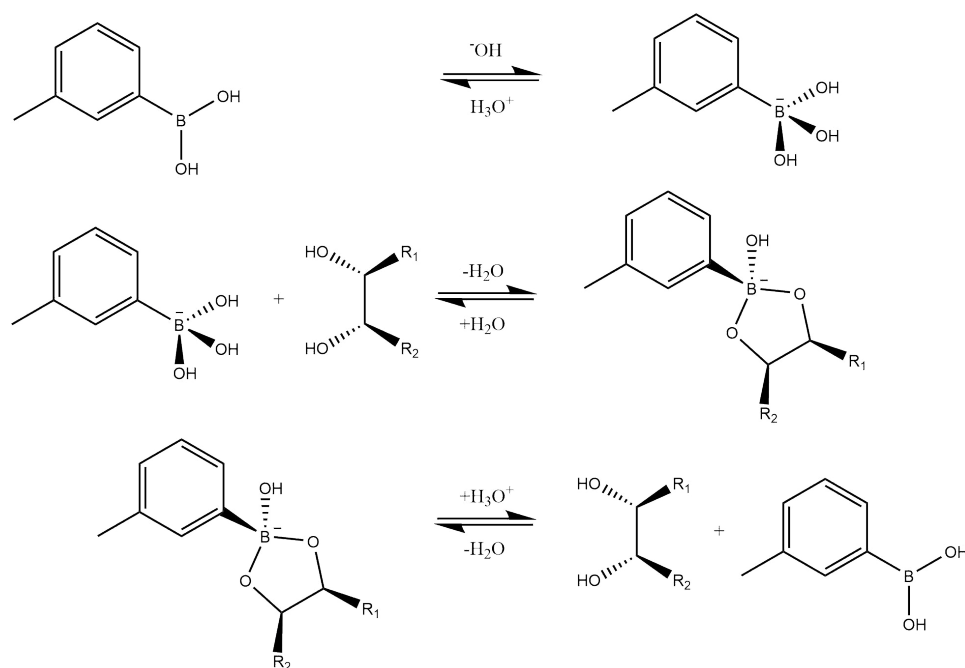
Figure 1.17: Mechanism of boronate retention. From top: Equilibration, Retention, Elution

Using these cartridges, the spectra produced by SPE-NMR can hence be simplified (by splitting a complex mixture into multiple fractions based on physical differences, such as polarity or pH) for metabolite annotation. Flexibility with solvent systems allows for a step-gradient approach, in which compounds are eluted separately through the use of different solvent mixtures of increasing/decreasing polarity. Additionally, SPE can be used to remove matrix effects from a sample, to increase concentration of analytes, and to minimise ion suppression in SPE-MS. Different cartridges can be used to this effect — for example, reversed phase cartridges can be used for the desalting of samples.

1.4.2 Targeted and untargeted SPE

Analytical methods — especially in metabolic profiling — can be generally distinguished into two approaches; targeted methods aim to measure specific known and characterised compounds or compound groups, whereas untargeted methods aim to analyse all measurable analytes in a sample, including chemical unknowns[51].

Previous use of SPE has generally aimed to utilise targeted methods to selectively extract one specific, known compound from a complex mixture. For example, an early use of SPE to

selectively extract target compounds included the spiking of radiolabelled β-blocker compounds into dog plasma and subsequent subjection to SPE, with the resulting eluants analysed using liquid scintillation counting[52]. In the same year, Wilson and Nicholson first described "SPEC-NMR", dosing healthy participants with therapeutic doses of common Non-Steroidal Anti-Inflammatory Drugs (NSAIDs) — aspirin, ibuprofen, paracetamol, naproxen, and oxpentifylline — and utilising SPE methods in order to selectively retain these individual drug metabolites, which could then be annotated via NMR spectroscopy[53,54]. SPE has also found significant use in hyphenated techniques, such as SPE-UPLC-MS[55].

Drug metabolites, both licit and illicit, are very common targets for solid phase extraction methods applied to human samples. In addition to the previously mentioned NSAIDs, targeted SPE methods have been devised for amphetamines[56,57], benzodiazepines[58,59,60,61], caffeine[62], cannabinoids[63,64,65], cocaine and heroin[66,67,68], irbesartan[69], quinolones[70], statins[71,72], and phenolphthalein glucuronate[73], as well as other common drugs[74], in biofluids such as urine and plasma.

Relatively few cases exist where solid phase extraction is used to attempt to selectively retain multiple metabolites in one experiment — examples of literature describing this include procedures to retain phthalates[75,76], steroids[77,78,79], and methylxanthines (including caffeine)[80], as well as the use of PBA sorbents to retain metabolites associated with apple consumption[81]. Similarly, there are very few cases of SPE being used as part of an untargeted method, where the breadth of the sample is analysed and a number of retained metabolites are unknown. Utilisation of SPE to achieve fractionation — and a corresponding reduction in peak overlap in human urine samples — was demonstrated by Yang *et al.* using C_{18} cartridges[82]; Jacobs *et al.* achieved similar results using HLB cartridges, including the quantitative measurement of specific polyphenols from dietary sources[83].

1.5 Mass spectrometry

The mass spectrometry theory below is described in the textbook 'Mass spectrometry: principles and applications' by de Hoffman and Stroobant[84].

While NMR represented the primary form of analysis during the course of this project, there were situations where analysis benefited from the use of mass spectrometry (MS) — an orthogonal technique which, beyond providing mass information about a compound, demonstrates much greater sensitivity than NMR. MS analysis, in its most basic form, is represented by analytes being ionised (either 'hard' ionisation, where the analytes undergo fragmentation into smaller, detectable fragments, or 'soft' ionisation, where a single proton is added or removed, depending on the polarity of the ionisation method); the resulting ions are then detected within the instrument itself, producing mass information about the ions.

Mass spectrometers come in many different forms, with the detection method of each instrument providing a specific use-function for a given application; for example, the superior resolution of an FT-ICR (Fourier Transform Ion Cyclotron Resonance) instrument is hindered by its expense and need for liquid nitrogen cooling. During the course of this project, a qToF (quadrupole-Time of Flight) instrument — which couples a quadrupole to a time of flight detector — was utilised; in this instance, detection is achieved by ionising analytes using the 'soft' electrospray ionisation (ESI), followed by time-of-flight analysis, whereby ions are accelerated towards a detector using electromagnetic fields — the time taken to reach the detector is dependent on the size of the ion, such that smaller ions arrive at the detector faster and vice versa. The use of the qToF also allows for tandem mass spectrometry (also known as MS/MS), where individual precursor analyte ions can be preselected for fragmentation; here, the quadrupole mass-selects the precursor ion through the use of electromagnetic fields, followed by fragmentation by collision-induced dissociation (CID) as the precursor gains energy through collisions with an inert gas such as nitrogen or argon; these fragments are then detected through the ToF detector, as with the full scan experiment. The fragments produced during MS/MS are useful in structural elucidation, both for determining the structure of an unknown and for confirming the identity once annotated, since a compound will reliably produce the same fragments when the same mass spectrometer and method is used, making comparison with an authentic reference standard facile.

1.6 Aims and objectives

Despite its reproducibility, ease of use, and quantitative data generation, the limited sensitivity of NMR bottlenecks metabolite identification; low-concentration compounds become indistinguishable from noise, and peak overlap by concentrated compounds regularly obscures signals from the less concentrated. Previous SPE-NMR experiments have utilised standard or widely available methods using classic reversed phase sorbents, but have not altered conditions such as pH — or utilised ion exchange sorbents — to generate novel retention profiles.

The ability to use untargeted SPE methods in order to selectively retain metabolites based on their compound class or other functional groups promises greater control over the annotation and identification processes, which are essential to understanding the human metabolome. Hence this approach is highly relevant to the detection and characterisation of unknowns in complex biological mixtures — an integral aspect of metabolic profiling. Here, we aim to develop methods to selectively retain compounds within complex natural mixtures — such as biofluids, as described here — based on their taxonomy. These methods could consequently be used in an automated, high-throughput, and 'untargeted' manner, in order to generate hypotheses and further research questions, and to improve understanding of the metabolome.

Starting with SPE sorbents with strong retentive capacity, as well as sorbents utilising unique retention mechanisms, SPE experiments would be devised such that their retentive capacity could be determined, qualitatively and quantitatively; this would also result in a number of unknown peaks, which could be identified using NMR and other analytical tools. The SPE methods developed were then translated into automated methods and validated; one of these methods was then used to attempt to selectively retain glucuronide metabolites from natural urine. By demonstrating this selective retention, the utility of the automated methods in selectively retaining metabolites based on their compound class could be demostrated.

With many different SPE sorbents and cartridges available, it became clear that selection criteria would be initially required in order to determine which cartridges would be worth using in future SPE-NMR methods. This was done by determining the breakthrough volume of compounds within a stock of natural human urine through a variety of different SPE cartridges.

Chapter 2

Experimental determination of breakthrough volumes

2.1   Background

When developing solid phase extraction methods, it is important to understand the extent to which different compounds will retain on sorbents with different chemistries. The concept of breakthrough volume (V_B) has been defined in multiple terms — one such definition from Bielicka-Daszkievwicz *et al.* describes "the sample volume which can be loaded on the sorbent bed [of an SPE cartridge] without the loss of the analytes"[85]. As a sample is run continuously through the cartridge, the analyte or analytes are retained via varying intermolecular forces — eventually the retention sites become saturated, causing additional sample to 'bleed' through the cartridge without being retained.

Breakthrough volumes differ depending on factors such as sorbent surface area, particle size, pore size, and pore volume. Experimentally, breakthrough volume can be determined in solid phase extraction using frontal chromatography[85,86] — cartridges are conditioned with an appropriate solvent and equilibrated if necessary (while not being allowed to run dry); sample is then applied continuously to the cartridge with fractions collected at set intervals. These fractions can then be analysed using spectroscopic methods — such as NMR — in order to see precisely when analytes start to filter through, and hence the breakthrough volume. As this

continues, a breakthrough curve can be constructed (fig 2.1):
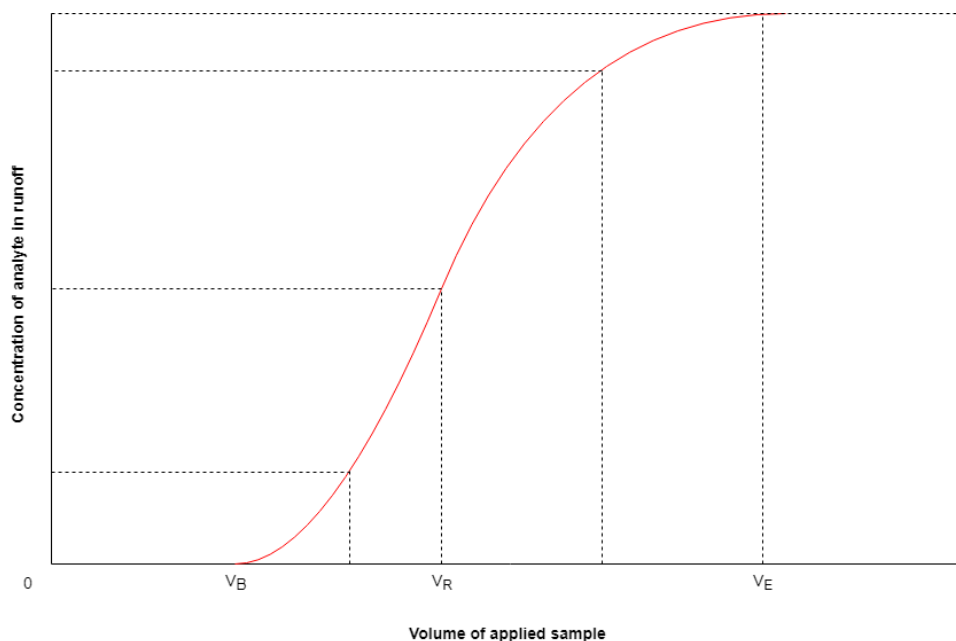
Figure 2.1: Idealised breakthrough graph depicting the change of concentration of analyte in runoff over time. Adapted from Gelencsér *et al.*, 1995[87]

Where V_R = retention volume and V_E = equilibrium volume. For targeted SPE experiments, there is a choice of two different methods when designing experiments — the first usually involves applying sample in excess of V_E to maximise retention, whereas the alternative applies sample below V_B in order to minimise sample waste.

Within the confines of untargeted SPE, experiment design becomes more complicated. Complex mixtures like urine can be formed naturally from a large number of compounds, between which there are great differences in chemical structure and concentration. As such, different compounds will have dramatically different breakthrough and retention volumes within the same sample — for example, in a urine sample, urea will elute through a C_{18} cartridge almost immediately (very low V_B); comparatively, compounds such as dimethylamine (DMA) may not elute for over 10 mL of applied sample (very high V_B). Compounds may also compete for retention sites, leading to different breakthrough volumes being estimated compared to if an analyte were being fed through the SPE cartridge in pure form. There are hence competing issues which require often imperfect solutions — for example, while wanting to obtain the most information about breakthrough volume across all compounds present, time limitations relating

to the desire to identify and quantify all compounds in a spectra multiple times limits what can be reasonably achieved.

In order to generally describe breakthrough for untargeted SPE of complex natural mixtures using NMR, two different approaches were devised. First, one can take a set of representative molecules and use it to generalise the retention behaviour of a range of compounds. This provides the most quantitative data (as only one set of peaks needs to be tracked) — however, representative compounds cannot inherently represent entire compound classes perfectly, and the complexity of urine makes it inevitable that many compounds may not be represented well, or even at all. Alternatively, spectra can be viewed qualitatively, with notable regions analysed by eye. While this is not suitable alone for estimating breakthrough volume, it can give intuitive information about retention for a broader range of compounds, individually (such as for hippurate, fig 2.2). For these examples, there may be minor variation in reading integrations due to inherent error in TSP buffer concentration, shimming of samples, increase in concentration of overlapped resonances, etc. From these approaches, we can reduce a larger number of sorbents down to a more manageable set for developing untargeted methods.
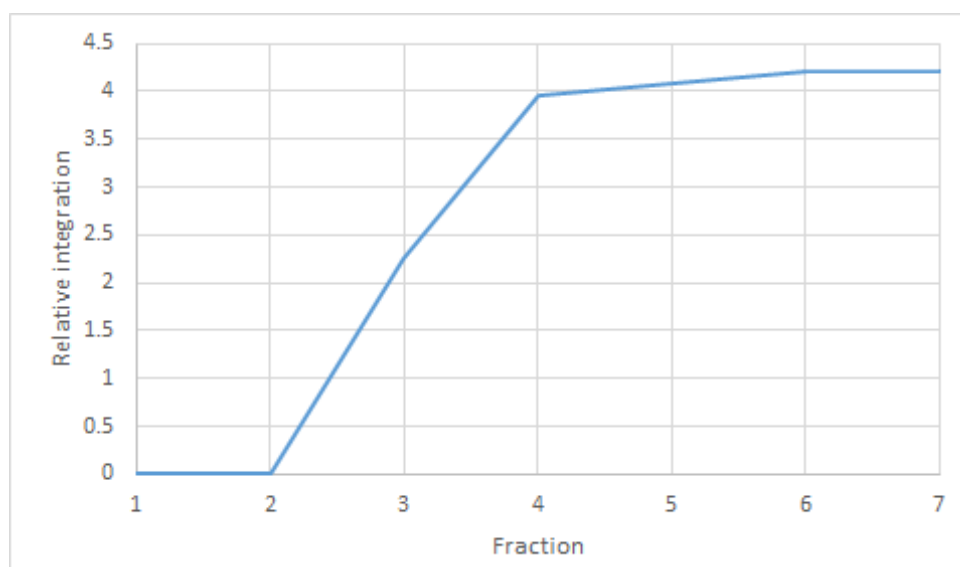
Figure 2.2: Determination of the breakthrough volume for hippurate using a 6 mL 500 mg bed weight phenyl column, based on integration of a NMR resonance at 3.97 ppm (d). The resemblance to the previous ideal breakthrough graph is notable; the breakthrough volume can be easily ascertained as the first fraction where the compound is present.

2.2   Methodology

Sample collection

Urine was collected from a healthy male volunteer in a 500 mL Corning tube pre-rinsed with ultrapure water. The sample was stored in a 4°C refrigerator when not in use, for up to one month, and subsequently disposed of. Buffer containing trimethylsilylpropionate (TSP) as a chemical shift reference standard was added to 540 µL of the sample, as described by Dona *et al.*[88]. 580 µL of the manually vortexed sample was then transferred into a 5mm NMR tube. The sample was analysed using a Bruker Avance III 600 MHz spectrometer equipped with either a BBI room temperature probe or a CryoTCI triple resonance CryoProbe, in order to check for PolyEthyleneGlycol (PEG) contamination, acquiring a one-dimensional water-suppressed ^{1}H NOESY and JRES spectra at 300K.

SPE methods

6 mL capacity, 500 mg sorbent weight, 40-60 µm particle size SPE cartridges were obtained from the following suppliers:

- C_8, C_{18}, Phenyl, Diol cartridges: Thermofisher (Hypersep)

- HLB: Waters (Oasis)

- PBA, WCX, SAX, SCX: Macherey-Nagel (Chromabond)

- NH_2 (WAX), C_{18}-ec, CN: Biotage (Isolute)

Reversed phase cartridges (C_{18}, C_{18}-ec, HLB, phenyl, CN) were conditioned with 6 mL methanol and equilibrated with 6 mL water. Normal phase cartridges diol, CN, and NH_2 were equilibrated with hexane. The ion exchange WAX and SCX cartridge samples were pre-treated with an equal volume (6 mL) v/v with sodium acetate/acetic acid buffer (0.1 M, pH 5); the acidified SCX sample was pre-treated with HCl solution to pH 2. Samples for WCX and SAX cartridges were pre-treated with 5% ammonium hydroxide; the basified SAX sample was pre-treated with NaOH solution to pH 10. Samples to be used in PBA cartridges were

pre-treated with an equal amount (6 mL) v/v of glycine/hydrochloric acid buffer (0.2 M, pH 8.5).

After pre-treatment, conditioning, and equilibration, 30 mL of sample — chosen to guarantee breakthrough of virtually all compounds — was loaded onto each cartridge, with 1 mL fractions collected at appropriate intervals. These 1 mL fractions were dried under nitrogen and reconstituted in 600 µL ultrapure water. Buffer containing trimethylsilylpropionate (TSP) as a reference was then added to 540 µL of the reconstituted fractions, as described by Dona *et al.*[88]. 580 µL of this mixture was then transferred to 5mm NMR tubes for analysis.

NMR acquisition

Samples were analysed using a Bruker Avance III 600 MHz spectrometer equipped with either a BBI room temperature probe or a CryoTCI triple resonance CryoProbe. One-dimensional water-suppressed ^{1}H NOESY spectra, as well as J-resolved spectra, were acquired at 300 K. After acquisition and processing, spectra were normalised relative to the TSP signal at 0 ppm.

2.3   Cartridge selection

In the Human Urine Metabolome DataBase (HMDB, http://hmdb.ca)[89], chemicals are assigned — first to classes, then to subclasses — using the ChemOnt automated taxonomy, categorised by ClassyFire[90]. This class structure can be utilised in order to identify representative molecules. The representative compounds (table 2.1) were chosen for their prevalence in urine, their minimal overlap with other peaks, and their diversity of functional groups. Peaks were annotated through cross-referencing of the obtained ^{1}H NMR spectra with pre-obtained 800 MHz JRES, COSY, TOCSY, and HSQC spectra of a urine sample supplied by the National Phenome Centre (NPC). Both were compared with literature values from the HMDB to confirm assignments to a level 2 (putative identification) standard.

Quantification of metabolites in NMR is typically done using peak integrals with reference to a signal of constant intensity — for these experiments, breakthrough was determined in this way by comparison to TSP. However, complex natural mixtures are sensitive to peak overlap from other compounds, which can interfere with the accuracy of the quantification process —

hence, for future experiments retention capacity was measured with reference to peak height, rather than to peak integration. The breakthrough volume for each representative compound — here, denoted as the first fraction in which the compound can be observed by NMR — is listed in table 2.2.

Table 2.1: Selected representative compounds

Compound	Class/sub-class	^{1}H NMR assignment
Glycine	Alpha amino acid	3.57 (s)
Formate	Carboxylic acid	8.46 (s)
Histidine	Histidine and derivatives (Carboxylate, imidazole)	$\approx$ 8.04 (s)
Trigonelline	Alkaloid (Pyridine, carboxylate)	9.13 (s)
Citrate	Alpha-hydroxy acids and derivatives	2.67 (d, J $\approx$ 15.0 Hz)
2-PY*	Pyridines and derivatives	6.67 (d, J $\approx$ 9.0 Hz)
Dimethylamine (DMA)	Amines (Aliphatic amine)	2.71 (s)
Gamma-aminobutyric acid	Carboxylic acids and derivatives (Gamma amino acid)	2.27 (t, J $\approx$ 8.0 Hz)
& Galactose	Carbohydrates and carbohydrate conjugates (Di-/monosaccharides)	5.26 (d, J $\approx$ 3.8 Hz), 5.24 (d, J $\approx$ 3.8 Hz)
Hippurate	Benzoic acid (Carboxylate)	3.96 (d, J $\approx$ 5.7 Hz)

*N-methyl-2-pyridone-5-carboxamide.

Table 2.2: Breakthrough volume by representative compound (mL)

Cartridge	Glycine	Formate	Histidine	Trigonelline	Citrate	DMA	GABA	2-PY	Sugars	Hippurate	Urea
C_{18}	1	1	2	3	1	2	2	10	2	5	1
C_{18}-ec	1	1	1	1	1	1	2	2	1	1	1
C_8	1	2	2	3	1	2	3	8	2	2	1
HLB	1	1	1	1	1	1	7	5	1	6	1
Phenyl	1	2	2	3	1	2	3	8	2	3	1
CN (RP)	1	2	1	2	1	1	1	2	1	1	1
Diol	1	1	1	1	1	1	1	2	1	1	1
PBA	-	1	2	3	1	11	1	3	2	1	1
SAX (pH 7)	1	1	1	2	2	1	1	1	1	1	1
SAX (pH 8)	-	2	2	2	3	2	2	2	-	1	1
SAX (pH 10)	1	2	1	2	2	1	1	2	2	1	1
SCX (pH 2)	3	-	3	3	1	-	2	2	2	2	1
SCX (pH 5)	1	1	7	2	1	14	1	2	2	1	1
SCX (pH 7)	6	-	>15	12	1	13	2	2	-	2	1
WAX (pH 7)	1	-	7	2	7	1	2	3	3	1	1
WCX (pH 7)	1	1	1	2	2	>15	2	2	2	2	1

where - indicates that breakthrough volume could not be determined due to use of compounds in methodology.

With these results, it became possible to critically evaluate each cartridge for further experimentation. As previously mentioned, this meant qualitatively describing retention broadly across regions, and quantitatively describing the breakthrough volume of the representative molecules.

C_{18}

The C_{18} cartridge retained peaks across the entire scannable range — most visibly, in the aromatic region. A comparison between the final fraction and the original, unfiltered sample shows identical traces for urea (5.79 ppm, s); urea is consistently unretained regardless of cartridge used. Visual comparison in the aromatics region shows that while compounds have been eluted almost fully, there remain some resonances which are marginally less intense than in the original spectrum, or even still retained (such as in the doublet at 8.8ppm (fig 2.3). Similarly, in the aliphatic region, there were significant differences in retention between the untreated 'raw' sample and the final fractions.
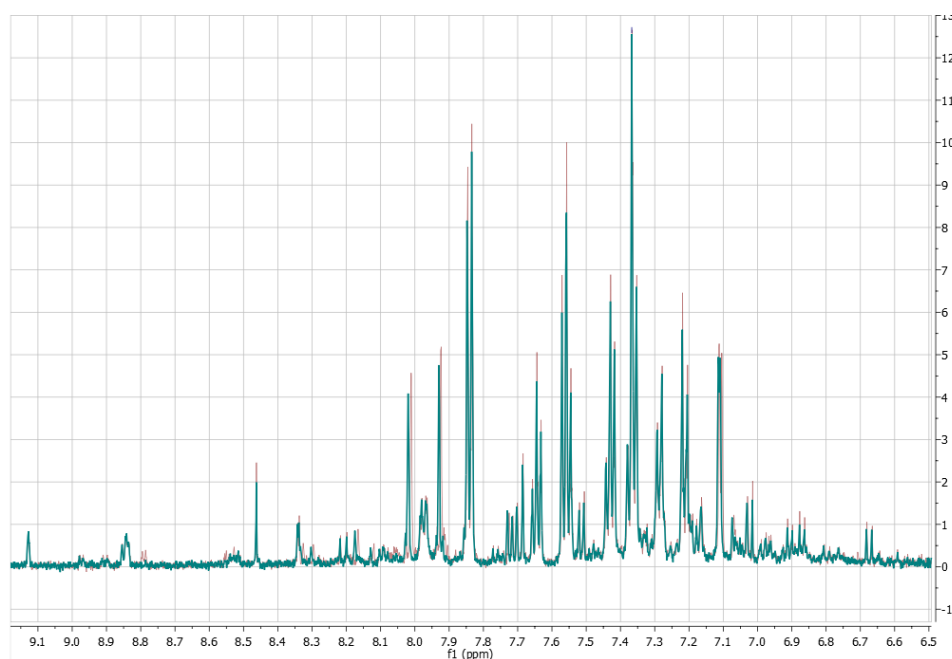
Figure 2.3: Superimposed spectra demonstrating minimal loss between the untreated 'raw' urine (red) and the final C_{18} fraction spectra (blue, bold) in the aromatic region.

The representative molecules show wide-ranging retention. *N*-methyl-2-pyridone-5-carboxamide (2-PY) is the most well retained (requiring 10 mL before breakthrough), with hippuric acid (5

mL) and trigonelline (3 mL) also being well retained. It is notable that all of these molecules are aromatic, and that they are (hence) all retained to different extents. It is also notable that the trigonelline is less well retained than hippuric acid and 2-PY, possibly due to its aromatic system being based on pyridine, rather than on benzene, contributing some polar character to the molecule; this polar molecule would then find retention in the silanol groups, which act to give the sorbent some ion exchange character. It is possible that GABA is also retained to some extent – despite only apparently having breakthrough of 2 mL, the region itself overlaps with other peaks; additionally, we only see the characteristic breakthrough curve begin at around 10 mL. The retention of GABA would fit in with the C_{18} hydrophobic retention profile — despite GABA being a zwitterion at pH 7, it may have enough hydrophobic character from its butyrate 'backbone' to be retained. We additionally see some limited retention of histidine (imidazole-based) and dimethylamine (aliphatic amine). Glycine, formic acid, citric acid, and the sugars do not show retention in C_{18}. Finally, when evaluating cartridges for future work, it is worth considering the extent to which the sorbent is utilised across the literature — in this case, C_{18} cartridges are extremely common, and is often the first line of choice when developing targeted SPE methods for its wide-ranging retention. For this reason, as well as for its excellent retention capacity, C_{18} was selected for further study.

C_{18}-ec

C_{18}-ec cartridges are typically used for the retention of very non-polar compounds, due to the 'elimination' of the silanol polar character via the endcapping process. This elimination of the polar character is unlikely to be total (endcapping will not react with every silanol, simply due to the sheer number of silanols present) — as a result, it might be expected that the endcapped cartridge wouldn't have an overwhelmingly dissimilar retention profile to its non-endcapped equivalent. In practice, however, this doesn't appear to be the case, as every representative compound bar 2-PY and GABA breaks through immediately. This stands to reason in comparison with the C_{18} results; the compounds retained in C_{18} (2-PY, hippuric acid, trigonelline, GABA) either contained charged or polar constituents (trigonelline, GABA), or contained electron-directing heteroatoms which influenced the electron density around the

aromatic ring (hippuric acid, 2-PY). With the loss of the silanol polar character, the ability to retain these compounds decreases. Human urine has very few (if any) very non-polar compounds — by function, it is predominantly composed of wastes soluble in water, and hence excludes most lipids, etc., which are immiscible. As a result of the loss of its polar character, and due to the lack of very lipophilic metabolites in human urine, C_{18}-ec was not considered generally useful for further study during this project.

C_8

The intermolecular forces involved in the retention of metabolites for C_8 are virtually the same as in a C18 cartridge, although C_8 does exhibit slightly less hydrophobic character (and, consequently, more polar character) than C_{18}; the substitution of octadecyl for octyl chains reduces the number of possible hydrophobic forces, and consequently brings the silanol groups into closer proximity. This translates well into the findings of the experiment — as a result of the reduced hydrophobic forces, the breakthrough volumes for 2-PY and hippuric acid are decreased. At the same time, the increased polar character results in marginally increased breakthrough volumes for formic acid, GABA, and the sugars. Since the retention profiles are virtually the same — with smaller GABA/hippuric acid breakthrough volumes for C_8 — it makes little sense to study both C_{18} and C_8, as larger volumes of sample can be used with no noticeable disadvantage on a C_8 cartridge. As such, research on C_8 was not taken further.

HLB

HLB is 'a hydrophilic-lipophilic-balanced, water-wettable, reversed-phase sorbent'[91] which aims to provide retention of polar compounds in a manner comparable to a C_{18} cartridge — this is achieved through the use of a dual-monomer system consisting of divinylbenzene and pyrrolidinone moieties. The breakthrough volumes for the representative molecules, however, do not immediately appear to agree with this. 2-PY, GABA, and Hippuric acid are very strongly retained — but other polar molecules, such as trigonelline and histidine, elute immediately. The retention profile appears similar to the C_{18} retention profile, although with some significant differences; for example, GABA is strongly retained on HLB (breakthrough occurs at 7 mL), whereas it was only weakly retained on C18 (2 mL). Similarly, 2-PY is moderately retained on

HLB (5 mL), yet was very strongly retained on C18 (10 mL).

With this information alone, it might be suggested that HLB be eliminated due to a retention profile very similar to C_{18}. However, visual inspection of the fraction spectra reveals significant retention of compounds not retained as well by C_{18} from across the spectral width, especially the aromatic region (fig 2.4); on top of this, the ability of HLB cartridges to function without conditioning and equilibration steps — and their tolerance to drying during the SPE process — warranted their further study.

Phenyl

In theory, phenyl modifications utilise a unique mechanism for retention: π-stacking is a weak (0.05-40 kJ/mol) intermolecular attraction between aromatic and conjugated systems. The weakness of this interaction demonstrates itself in the spectra — while it exhibits a slightly different retention profile to that of C_{18}, the breakthrough volumes are significantly smaller. The similarities between the C_{18} and phenyl cartridges suggest that a dual mechanism of retention is present; on top of the π-stacking, it's likely that significant hydrophobic interactions are exhibited due to the homogeneity of the electron distribution around the benzene ring. This dual mechanism would also suggest some form of competition between different compounds classes; the finite number of available retention sites would mean that the more strongly attracted compounds are hence more likely to be retained. As a result of this, only the compounds exhibiting π-stacking interactions of comparable or higher strength than the hydrophobic forces will be notably retained. While the retention profile for the phenyl cartridge was not promising in terms of its uniqueness, the possibility for the unique intermolecular forces to play a larger role under different conditions lead it to be chosen for future study.
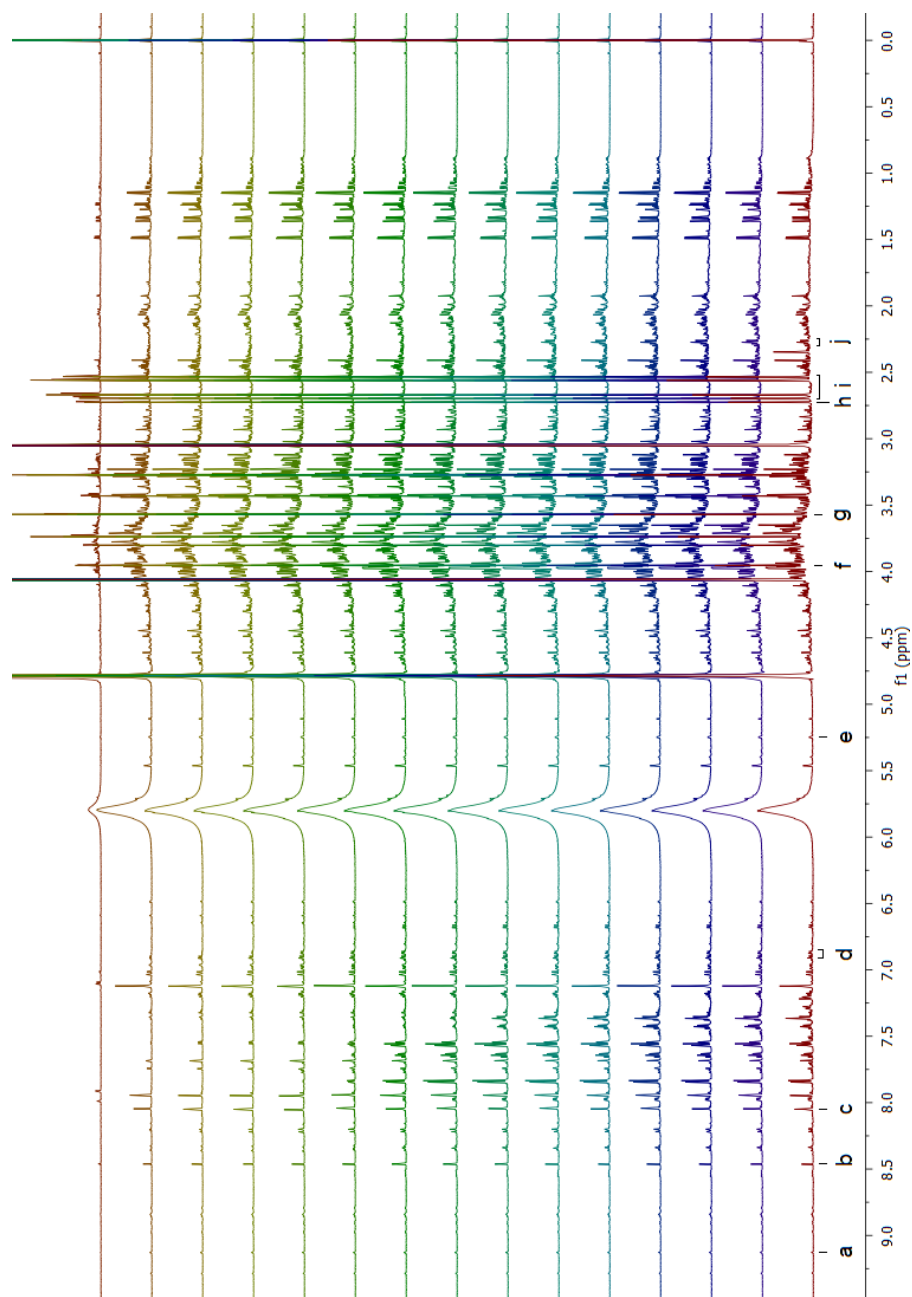
Figure 2.4: Stacked spectra demonstrating HLB breakthrough (From top: fractions 1 through 14; bottom, untreated 'raw' urine). a = trigonelline, b = formate, c = histidine, d = 2-PY, e = lactose/galactose, f = hippurate, g = glycine, h = DMA, i = citrate, j = gamma-aminobutyrate.

CN (reversed phase)

The CN cartridge is typically used in the normal phase in order to retain polar compounds — in theory it can also be used in reverse phase applications in order to retain compounds which would otherwise be irreversibly retained by C_{18}. When compared to C_{18}, however, there appeared to be very little retention in the reversed phase — it is possible that the intermolecular forces involved are weak enough to drastically reduce the breakthrough volumes to virtually nothing. As such, there was little reason to study the CN cartridge in the reverse phase further, as the breakthrough volumes are too small to be workable, despite any possible unique retention profile.

Normal phase sorbents: CN, diol, NH$_2$

The use of normal phase sorbents for the purpose of studying urine immediately suggests that they would produce few results of value, as the solvents in use are not likely to interact particularly strongly with charged and polar solutes dissolved aqueously. The sorbent, itself polar, would be likely to retain not only compounds from across the metabolome, but also potentially water itself. As expected, the retention profile of the CN, diol, NH$_2$ cartridges using urine is virtually non-existent, with every metabolite studied breaking through almost immediately (fig 2.5).
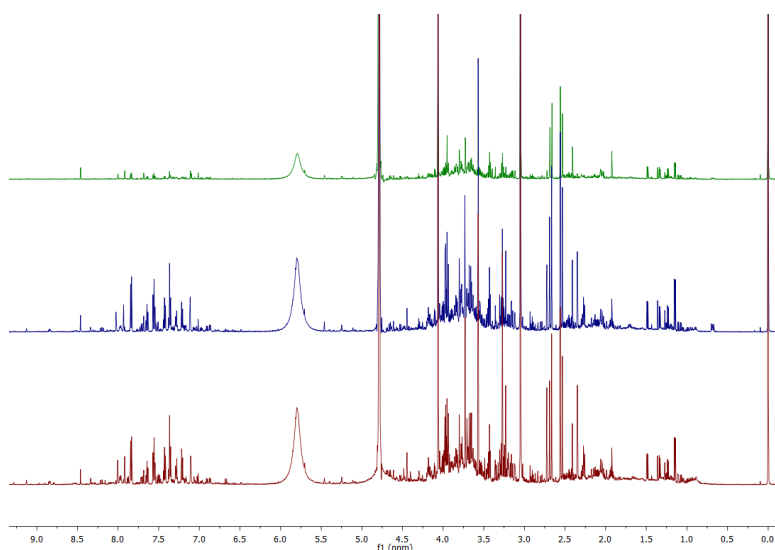
Figure 2.5: Stacked spectra demonstrating diol breakthrough (From top: fraction 1, fraction 2, untreated 'raw' urine).

Visual inspection of the NMR spectra over time reveal no retention across all datasets, with the single exception of 2-PY, which elutes in the second fraction when using diol cartridges. Due to these results, no significant reasons to continue studying normal phase cartridges using human urine were clear, and hence research was discontinued.

PBA

The PBA cartridge utilises a unique covalent retention mechanism, and is expected to selectively capture cis-diols, α–hydroxy acids, aromatic ortho-hydroxy acids, aminoalcohols, and diketones, amongst other similar compounds, usually with two vicinial heteroatoms. Interestingly, these predictions are only partially validated by the results obtained. Lactose and galactose both contain cis-diols, and as a result achieve breakthrough after 2 mL applied; similarly, the α-amino acid histidine is retained, and breaks through after 2 mL. These results do not demonstrate particularly impressive retention capability of cis-diols. By contract, the compound most well retained is DMA — a secondary amine – which does not achieve breakthrough until 11 mL; similarly, 2-PY is retained until 3 mL of sample has been applied, despite being a para-aromatic compound. Citric acid, a tri-carboxylic acid and -hydroxyl, is not retained at all — which suggests that the steric requirements for retention are particularly strict. The retention of glycine could not be studied due to the use of glycine buffer in this instance — future work made use of a phosphate buffer as an alternative. It was worth questioning the effect of the buffer on the experiment, since as an α-amino acid, glycine — like histidine — is likely to be retained at least partially, allowing it to compete for retention sites. Even with these considerations, the unique retention mechanism — and, hence, the demonstrated unique retention profile — of the PBA sorbent recommended its further study.

SAX and WAX

The strong anionic exchange cartridge was trialled under three different conditions: sample pre-treated 1:1 with 2% formic acid, sample buffered to pH 8 with tris buffer, and sample basified to pH 10 with NaOH. Tris was selected due to its excellent buffering capacity — however, the use of tris (an organic molecule) resulted in significant resonances being introduced to the spectra. The experiment would be later repeated with a pH 8 inorganic phosphate buffer, which

gave comparable results to the tris buffer without the additional NMR signals.

Ion exchange cartridges selectively retain charged molecules — as a result, it is necessary to control pH of urine for metabolties to be retained. In this instance, strong anionic exchange involves the retention of all negatively charged species, whereas weak anionic exchange will retain only strong acids. This distinction can be demonstrated by the differences in breakthrough volumes between SAX cartridges buffered to pH 8 and SAX cartridges without buffering; while most compounds — bar trigonelline and citric acid — are eluted immediately in the bufferless sample, there is a much wider retention of metabolites when samples are buffered to pH 8. This is caused by the deprotonation of acids at higher pHs. Basification of the sample to pH 10 decreases the retentive capability of the cartridge.

The WAX cartridge — formed of protonated NH_2 active groups (as opposed to the permanently charged N^+Me_3 SAX group) is designed to retain strong acids only, due to its relative pH sensitivity. As such, the WAX cartridge selectively retains histidine and citric acid very strongly, and only glycine and hippuric acid are eluted immediately; all other representative molecules only achieve breakthrough after 2 mL or more of applied sample. As such, there are significant differences between SAX and WAX; of the two, the SAX cartridge was selected for future experiments, with the understanding that the pH must be neutral or even lowered.

SCX and WCX

The strong cationic exchange cartridge was trialled under three different conditions: sample pre-treated 1:1 with 2% formic acid, sample buffered to pH 5 with sodium acetate buffer, and sample acidified to pH 2 with HCl. SCX cartridges are expected to retain all positively charged species due to the presence of a pseudo-permanent negative charge on the sulphate modification. As with the SAX cartridges, the metabolites must first be charged.

The results for both buffered pH 5 and unbuffered pH 7 experiments are quite comparable — both show a very strong retention for DMA and histidine. The unbuffered sample also shows a strong retention of trigonelline and glycine: when buffered with sodium acetate, neither are retained. It is unclear why this disparity in retention profiles is caused.

As with the WAX cartridge, the breakthrough results for the WCX sorbent generally align

with expectations — while the very strong retention of DMA is unchanged, citric acid is no longer retained quite as strongly, breaking through at 2 mL of applied urine. Glycine and formic acid are eluted immediately — however, histidine is also eluted immediately, despite possibly having a retention capacity of more than 15 mL in SCX. This clear retention advantage over WCX made SCX the preferred choice for the next steps in the project.

Some cartridges — most notable among the ion exchange sorbents — demonstrated some level of sorbent 'bleed', possibly caused by disintegration of the silica modifications (fig 2.6). For example, WCX bleed caused distinctive peaks at 1.65 (m) and 2.22 (t, J $\approx$ 7.35 Hz); this can be diagnosed through simple comparison of the untreated 'raw' urine with the breakthrough fractions. In the WCX breakthrough experiments, the resonances can be seen to decrease as the fractions increase, suggesting that this may be a temporary problem — the reuse of the cartridge may hence remove this problem altogether while maintaining retention capacity. However, as the WCX cartridge was not selected for further study, this was not confirmed. Sorbent bleed was also observed in the WAX, SCX, and SAX cartridges — however, the resonances caused by the bleed were much more consistent in the strong ion exchange and did not diminish over time.
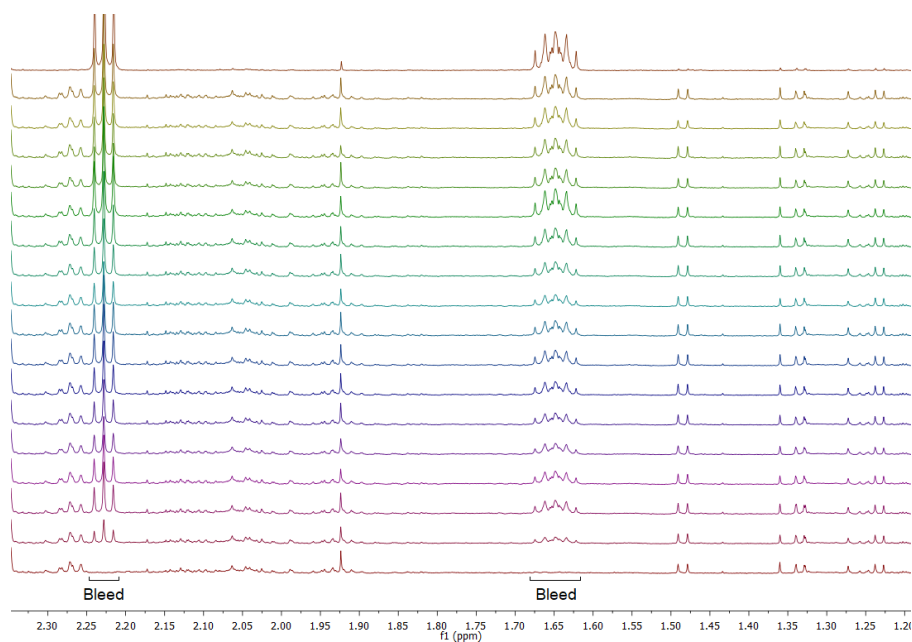
Figure 2.6: Demonstration of cartridge 'bleed' in WCX breakthrough experiment, present in breakthrough fractions (1mL each, from top), but not in original sample (bottom spectrum).

2.4   Conclusion

Breakthrough experiments were undertaken using human urine on a range of SPE sorbents in order to narrow down a number of sorbents for further study with artificial and natural samples. Sorbents requiring different solvent phases were utilised, with reversed phase sorbents being the most reliable (and most studied in the literature). Normal phase sorbents demonstrated no significant breakthrough, and research into reversed phase sorbents such as C_{18}-ec and CN was discontinued due to minimal or unremarkably different retention profiles. The cartridges selected for further study included C_{18}, HLB, phenyl, SCX, SAX, and PBA; for these, a 'compromise' volume of 3mL of applied sample on 6mL, 500mg bed weight cartridges appeared to give a good balance between compound breakthrough and metabolite concentration. These cartridges would hence be taken forward for further experiments, in order to gauge their retention profiles.

Chapter 3

Development of SPE-NMR protocols

The publication 'Application of novel solid phase extraction-NMR protocols for metabolic profiling of human urine' (Appendix A) is directly related to work presented in this chapter — content from the paper is hence present in part.

The experiments described were run with both artificial and natural human urine samples. The use of an artificial mixture was intended to simplify annotation efforts and solidify observed trends, whereas the complexity of natural urine allowed for greater insight into broader retention trends.

3.1   Methodology

3.1.1   Artificial urine formulation

As previously detailed, natural urine can contain thousands of small molecules, with compounds of higher concentration often obscuring compounds of lower concentration in NMR spectra. The development of a stock of artificial urine for the project aimed to produce insight into the retention capabilities of the different cartridge chemistries with a less complex mixture than that of natural urine, using a representative mixture of around 50 organic compounds.

The list of included metabolites and their concentrations for the artificial urine was based on previous formulations by Takis *et al.*[92], which varied concentrations of ions present in several standard mixtures in order to measure the chemical shift changes caused by those inorganic

compounds =. Additional metabolites such as benzoic acid were included from the author's 'spiked metabolite' list in order to broaden the range of metabolites with aromatic functional groups. Furthermore, inorganic salts were added in order to better simulate natural urine. A quantity of the metabolites equal to their median concentration in human urine were dissolved in 1L deionised water with stirring, with the resulting solution (table 3.1) stored at 4°C.

Table 3.1: List of compounds included in artificial urine formulation

Compound	Concentration*[7]	Assigned subclass	Actual mass (mg)
3-aminoisobutanoic acid	2.2-140.0	Beta amino acids and derivatives	74.25
3-methylhistidine	2.8-59.8	Histidine and derivatives	134.99
Acetic acid	2.5-106.0	Carboxylic acids	88.56
trans-Aconitic acid	1.8-25.1	Tricarboxylic acids and derivatives	149.72
L-alanine	7.1-43.0	Alanine and derivatives	46.34
Allantoin	4.9-29.3	Imidazoles	45.22
L-aspartic acid	3.5-21.8	Aspartic acid and derivatives	48.44
Benzoic acid	1.2-10.5	Benzoic acid and derivatives	179.21
Betaine	2.7-24.7	Alpha amino acids and derivatives	30.47
Citric acid	49-600	Tricarboxylic acids and derivatives	1445
Creatine	3-448	Alpha amino acids and derivatives	940.05
L-Cystine	5.0-24.6	L-cystine-S-conjugates	203.55
Dimethyl sulfone	1.3-49	Sulfones	59.58
Dimethylamine	20.3-59.2	Organonitrogen compounds	33.36
Erythritol	6.8-64	Carbohydrates and carbohydrate conjugates	98.61
Ethanolamine	24.8-56.2	Amines	42.76
Formic acid	6.9-120.9	Carboxylic acids	67.53
D-Glucose	12.5-58.4	Carbohydrates and carbohydrate conjugates	85.06

L-Glutamic acid	3.3-18.4	Glutamic acid and derivatives	47.08
L-Glutamine	19.1-77.9	Alpha amino acids and derivatives	41.13
Glycerol	4.0-19.0	Carbohydrates and carbohydrate conjugates	≈ 47.63
Glycine	44-300	Alpha amino acids and derivatives	64.70
Glycolic acid	3.7-122.0	Alpha hydroxy acids and derivatives	117.31
Guanidoacetic acid	10.6-97.3	Alpha amino acids and derivatives	27.27
Hippuric acid	19.0-622.0	Benzoic acids and derivatives	1381
L-Histidine	17-90	Histidine and derivatives	176.99
L-Lactic acid	3.5-29.3	Alpha hydroxy acids and derivatives	43.51
L-Lysine	3.7-51.3	Alpha amino acids and derivatives	96.92
Methanol	10-117	Alcohols and polyols	46.78
Myoinositol	7.9-36.1	Alcohols and polyols	80.33
L-Phenylalanine	3.5-11.2	Phenylalanine and derivatives	109.16
p-Hydroxyphenylacetic acid	1.4-44.3	Phenols	105.73
Propylene glycol	1.4-44.3	Phenols	≈ 134.50
L-Serine	11.6-53.4	Serine and derivatives	77.56
Succinic acid	0.3-33.3	Dicarboxylic acid and derivatives	47.12
Tartaric acid	2.6-64.4	Carbohydrates and carbohydrate conjugates	145.22
Taurine	13-251	Organosulfonic acids and derivatives	111.34
L-Threonine	6.4-25.2	Alpha amino acids and derivatives	32.11
Trigonelline	5.5-109.3	Alkaloids and derivatives	36.52
Trimethylamine-N-oxide	4.8-509	Aminoxides	472
Urea	174-49,097	Ureas	10175

*(uM/mM creatinine)

3.1.2   Natural urine sample collection

Urine was collected from 12 healthy volunteers in 500 mL Corning tubes pre-rinsed with ultrapure water — each volunteer provided informed consent in writing. Each urine sample was individually analysed by a member of the NMR team using NMR spectroscopy to check for polyethylene glycol — one sample was subsequently discarded due to significant contamination. The remaining samples were pooled into a pre-rinsed polypropylene carboy with a stir bar. The pooled sample was homogenised by stirring for five minutes, after which 15 mL aliquots of the pooled sample were dispensed into 20 mL Sterilin sample tubes. Samples were labelled sequentially and stored at -80 °C; samples to be used were subsequently transferred to a 4°C fridge for thawing.

For pH-altered urine samples, concentrated HCl or NaOH solution was added dropwise to 250 mL of pooled urine until the desired pH was achieved. The acidified (pH 2 and pH 5) and basified (pH 11 and pH 9) samples were then stored at 4°C.

3.1.3   Solid phase extraction protocols

All solid phase extraction cartridges obtained had 6 mL capacity, 500 mg bed weight, and 40-60μm particle size. C_{18} cartridges were acquired from Thermofisher (Hypersep™) and Agilent (Bond Elut™). C_{18}-ec, NH2 (Weak Anion eXchange — 'WAX'), CN, and SO_3 (Strong Cation eXchange — 'SCX') cartridges were acquired from Biotage (Isolute™). C_8, phenyl, and diol cartridges were acquired from Thermofisher (Hypersep™). Hydrophilic/Lipophilic Balance (HLB) cartridges were acquired from Waters (Oasis™). PhenylBoronic Acid (PBA), NMe_4^- (Strong Anion eXchange — 'SAX'), and CH_2COO^- (Weak Cation eXchange — 'WCX') cartridges were acquired from Macherey-Nagel (Chromabond™). Sample pretreatment, conditioning, equilibration, wash, and elution steps were tailored for the cartridge and method. Different methods utilised a variety of pH levels and solvent systems (table 3.2). The use of different conditions altered the retention profiles exhibited by the different cartridges, resulting in elutions comprising different metabolites.

Table 3.2: Descriptions of individual SPE methods

Cartridge	Method name	Modification
C_{18}	Neutral	No modification
	Formic acid	2% formic acid added to conditioning, equilibration, wash, and elution steps
	Weakly acidified	Sample acidified to pH 5
	Strongly acidified	Sample acidified to pH 2
HLB	Neutral	No modification
	Formic acid	2% formic acid added to conditioning, equilibration, wash, and elution steps
	Weakly acidified	Sample acidified to pH 5
	Strongly acidified	Sample acidified to pH 2
Phenyl	Neutral	No modification
	Formic acid	2% formic acid added to conditioning, equilibration, wash, and elution steps
	Weakly acidified	Sample acidified to pH 5
	Strongly acidified	Sample acidified to pH 2
SCX	Neutral	No modification
	Formic acid	2% formic acid added to conditioning, equilibration, wash, and elution steps
	Weakly acidified	Sample acidified to pH 5
	Strongly acidified	Sample acidified to pH 2
SAX	Neutral	No modification
	Formic acid	2% formic acid added to conditioning, equilibration, wash, and elution steps
	Weakly acidified	Sample acidified to pH 5

	Strongly acidified	Sample acidified to pH 2
PBA	Glycine buffer	Glycine-based buffer
	Phosphate buffer	Sodium phosphate-based buffer

Reversed phase cartridge (C_{18}, HLB, and phenyl), neutral pH sample

A Thermofisher Hypersep C_{18} cartridge (6 mL capacity and 500 mg bed weight) was conditioned with methanol (6 mL), then equilibrated with water (6 mL). Pooled urine (3 mL) was loaded onto the cartridge, which was then washed with water (6 mL) to eliminate interferences. The retained metabolites were then eluted with methanol (6 mL).

Reversed phase cartridge (C_{18}, HLB, and phenyl), 2% formic acid (all steps), acetonitrile elution

A Thermofisher Hypersep C_{18} cartridge (6 mL capacity and 500 mg bed weight) was conditioned with 2% formic acid in acetonitrile (6 mL), then equilibrated with 2% formic acid in water (6 mL). Pooled urine (3 mL) was loaded onto the cartridge, which was then washed with 2% formic acid in water (6 mL) to eliminate interferences. The retained metabolites were then eluted with 2% formic acid in acetonitrile (6 mL).

Strong cation exchange, neutral pH sample

A Macherey-Nagel Chromabond SCX cartridge (6 mL capacity and 500 mg bed weight) was conditioned with methanol (6 mL), then equilibrated with water (6 mL). Pooled urine (3 mL) was loaded onto the cartridge, which was then washed with 2% formic acid solution (6 mL). Methanol (6 mL) was used to elute the first set of metabolites, followed by 5% NH_4OH in methanol (6 mL) for the second elution.

Strong cation exchange, 2% formic acid

A Macherey-Nagel Chromabond SCX cartridge (6 mL capacity and 500 mg bed weight) was conditioned with 2% formic acid in acetonitrile (6 mL), then equilibrated with 2% formic acid in water (6 mL). Pooled urine (3 mL) was loaded onto the cartridge, which was then washed with 2% formic acid solution (6 mL). Acetonitrile (6 mL) was used to elute the first set of

CHAPTER 3. DEVELOPMENT OF SPE-NMR PROTOCOLS

metabolites, followed by 5% NH_4OH in methanol (6 mL) for the second elution.

Strong anion exchange, neutral pH sample

A Macherey-Nagel Chromabond SCX cartridge (6 mL capacity and 500 mg bed weight) was conditioned with acetonitrile (6 mL), then equilibrated with water (6 mL). Pooled urine (3 mL) was loaded onto the cartridge, which was then washed with 5% NH_4OH solution (6 mL). Acetonitrile (6 mL) was used to elute the first set of metabolites, followed by 2% formic acid in acetonitrile (6 mL) for the second elution.

Phenylboronic acid, sodium phosphate buffer

A Macherey-Nagel Chromabond PBA cartridge (6 mL capacity and 500 mg bed weight) was conditioned with a solution of 1% HCl in 70:30 water:acetonitrile (6 mL), then equilibrated with sodium phosphate buffer basified to pH 10 with sodium hydroxide (6 mL). Pooled urine (3 mL) was loaded onto the cartridge, which was then washed with sodium phosphate buffer basified to pH 8.5 with sodium hydroxide (6 mL). Water (6 mL) was used to elute the first set of metabolites, followed by a solution of 1% HCl in 70:30 water:acetonitrile (6 mL) for the second elution.

NMR sample preparation

Washes and elutions were dried under nitrogen and reconstituted in ultrapure water (3 mL). Buffer containing trimethylsilylpropionate (TSP) as a chemical shift reference standard was added to 540 µL of reconstituted sample, as described by Dona *et al.*[88]. 580 µL of the manually vortexed sample was then transferred into 5mm SampleJet NMR racks.

Samples which required additional 2D NMR experiments were dried under nitrogen and reconstituted in D_2O (3 mL). TSP phosphate buffer (60 µL) was added to 540 µL of reconstituted sample, and 580 µL of the resulting manually vortexed sample was transferred into 5mm NMR tubes.

NMR data acquisition

All 1D experiments were run using a Bruker Avance III 600 MHz spectrometer equipped with a BBI room temperature probe and SampleJet. Samples were analysed using one-dimensional

water-suppressed ^{1}H NOESY experiments at 300 K.

Additional ^{1}H–^{1}H J-resolved experiments, and 2D-NMR experiments including ^{1}H–^{1}H Total Correlation Spectroscopy (TOCSY), ^{1}H–^{1}H Correlation Spectroscopy (COSY), and ^{1}H–^{13}C Heteronuclear Single Quantum Coherence spectroscopy (HSQC), were utilised for metabolite annotation. The data from the 2D NMR experiments was acquired using a Bruker Avance III 600 MHz spectrometer equipped with a CryoTCI triple resonance CryoProbe.

Data analysis

NMR datasets were imported into MatLab using the Imperial Metabolic Profiling and Chemometrics Toolbox (IMPaCTS)[93]. Water (4.26—5.50 ppm) and formate (8.25—8.63 ppm) peak regions were removed from the spectra to eliminate interferences; the spectra were then normalised against the TSP region (-0.5 to 0.5 ppm) using a probabilistic quotient normalisation function[94]. Principal Component Analysis (PCA) plots were subsequently constructed with 5 principal components.

3.2 Peak annotation

Experiments were run isocratically, as any protocols developed 'manually' would require transcription to an automated system limited to three elutions (including wash). Gradient elution experiments were hence covered later into the project, in order to further demonstrate the ability of SPE cartridges to aid the annotation process.

Different cartridge chemistries were utilised in order to produce unique retention profiles for different compound classes (as demonstrated in fig 3.1). All samples utilised a sample load incorporating 3 mL of urine as a compromise between substrate retention capacity and spectral resolution. Each elution demonstrated differing retention profiles for each method — replicates of the same method, however, had little difference between spectra. Hence the reproducibility of the SPE methods outlined here can be guaranteed.
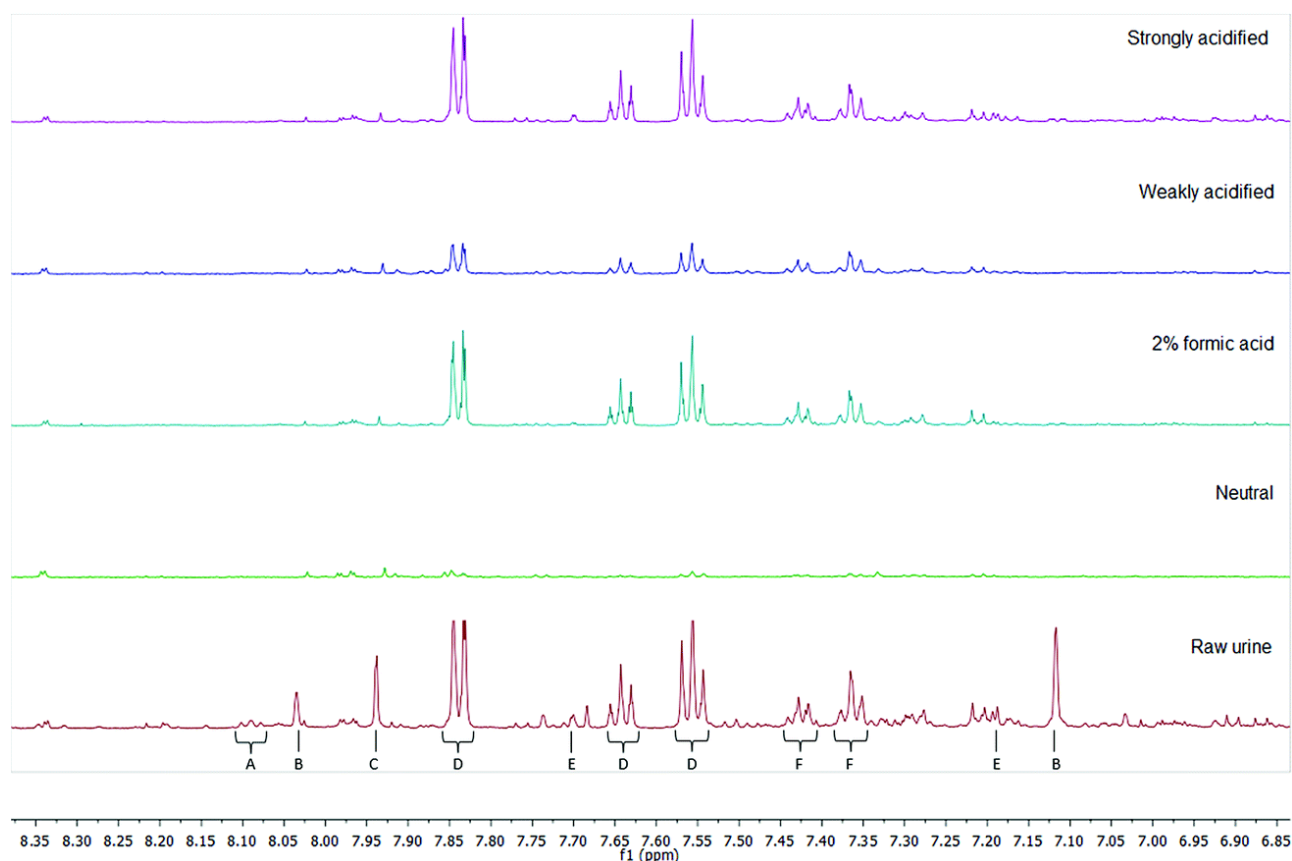
Figure 3.1: A comparison of the aromatic regions of 600 MHz NMR spectra of elutions from different C_{18} SPE methods on natural human urine. From top: sample acidified to pH 2, sample acidified to pH 5, 2% formic acid added to all steps, neutral sample, and pooled urine before SPE treatment. Select metabolites labelled: (A) trigonelline, (B) 3-methylhistidine, (C) histidine, (D) hippurate, (E) 2-furoylglycine, and (F) phenylalanine.

3.2.1   Representative metabolite analysis

Quantifying the extent of retention of different classes of compound can be done using molecules representative of a chemical class. In the Human Urine Metabolome database, chemicals are assigned taxonomically — first to classes, then to subclasses — using the ChemOnt automated taxonomy. As the taxonomy is automated, its class structure can be utilised in order to demonstrate retention profiles for given methods.

Two separate lists of metabolites were utilised to generate a list of representative compounds. One set of metabolites was examined and ranked according to their frequency of occurrence in the human urinary metabolome[7], such that the metabolites being examined would have a significant chance of being characterised in pooled urine samples. The second set was generated from the

previously described method for producing artificial urine[92]. The two lists were combined, and the resulting set of metabolites (table 3.3) was then characterised by their assigned subclass from the ChemOnt automated taxonomy.

Table 3.3: List of representative compounds

Compound	Occurrence (%)[7]	Assigned subclass	Representative peaks (ppm)
trans-Aconitic acid	55	Tricarboxylic acids and derivatives	6.59 (s)
Tartaric acid	82	Carbohydrates and carbohydrate conjugates	4.40 (s)
Succinic acid	91	Dicarboxylic acid and derivatives	2.41 (s)
L-asparagine	95	Asparagine acid and derivatives	2.88 (dd)
L-aspartic acid	95	Aspartic acid and derivatives	2.82 (dd)
2-furoylglycine	100	*N*-acyl-alpha amino acids	6.64 (q)
3-aminoisobutanoic acid	100	Beta amino acids and derivatives	1.20 (d)
3-methylhistidine	100	Histidine and derivatives	8.04 (s)
Acetic acid	100	Carboxylic acids	1.93 (s)
Allantoin	100	Imidazoles	5.40 (s)
L-alanine	100	Alanine and derivatives	1.49 (d)
Betaine	100	Alpha amino acids and derivatives	3.27 (s)
Citric acid	100	Tricarboxylic acids and derivatives	2.69 (d)
Creatine	100	Alpha amino acids and derivatives	3.05 (s)
Creatinine	100	Alpha amino acids and derivatives	4.06 (s)
L-Cysteine	100	Cysteine and derivatives	3.93 (s)
L-Cystine	100	L-cystine-S-conjugates	3.92 (s)
Dimethyl sulfone	100	Sulfones	3.14 (s)
Dimethylamine	100	Organonitrogen compounds	2.72 (s)
Erythritol	100	Carbohydrates and carbohydrate conjugates	3.70 (m)
Ethanolamine	100	Amines	3.83 (d)

Formic acid	100	Carboxylic acids	8.46 (s)
D-Glucose	100	Carbohydrates and carbohydrate conjugates	5.21 (d)
L-Glutamic acid	100	Glutamic acid and derivatives	2.08 (m)
L-Glutamine	100	Alpha amino acids and derivatives	2.46 (m)
Glycine	100	Alpha amino acids and derivatives	3.57 (s)
Glycolic acid	100	Alpha hydroxy acids and derivatives	3.96 (s)
Guanidoacetic acid	100	Alpha amino acids and derivatives	3.78 (s)
Hippuric acid	100	Benzoic acids and derivatives	7.56 (t)
L-Histidine	100	Histidine and derivatives	7.95 (s)
L-Lactic acid	100	Alpha hydroxy acids and derivatives	1.34 (d)
L-Lysine	100	Alpha amino acids and derivatives	1.45 (m)
Methanol	100	Alcohols and polyols	3.37 (s)
Myoinositol	100	Alcohols and polyols	3.55 (dd)
L-Phenylalanine	100	Phenylalanine and derivatives	7.37 (m)
p-Hydroxyphenylacetic acid	100	Phenols	3.45 (s)
Propylene glycol	100	Phenols	1.15 (d)
L-Serine	10	Serine and derivatives	1.33 (d)
Taurine	100	Organosulfonic acids and derivatives	3.43 (t)
L-Threonine	100	Alpha amino acids and derivatives	4.25 (m)
Trigonelline	100	Alkaloids and derivatives	9.13 (s)
Trimethylamine-N-oxide	100	Aminoxides	3.27 (s)
Urea	100	Ureas	5.80 (s, br)

The intensities of specific peaks corresponding to these metabolites in the elutions were compared to the intensities of the same peaks in the raw pooled urine samples. From this, the percentage retention per compound class could be estimated — hence, insight into which methods can selectively retain different compound classes can be achieved. For example, the peak intensity of retained creatinine (belonging to the subclass 'alpha amino acids and derivatives') from a method utilising a C_{18} cartridge at neutral conditions was measured at 257. When

compared to the intensity of the same peak in the 'raw' urine sample (5027), this gives an estimated retention capacity of creatinine using this method of 5%. After averaging with the other members of that subclass, C_{18} under neutral conditions has an overall retention of α-amino acids and derivatives of 0.75%.

The compound class retentions per method were then summed to produce a measure of the total retention capacity out of 100 — where methods ranking 0 would retain nothing, while methods ranking 100 would retain everything (table 3.4). It can be expected that the sum of the retention capacity estimates of the elutions and the washes for a given method is equal to 100.

Table 3.4: Total retention capacity estimates for each SPE method

SPE method	Total retention capacity (%)
C_{18}, neutral conditions	4.57
C_{18}, 2% formic acid	12.91
C_{18}, weakly acidified sample	10.65
C_{18}, strongly acidified sample	15.69
HLB, neutral conditions	8.87
HLB, 2% formic acid	17.40
HLB, weakly acidified sample	13.77
HLB, strongly acidified sample	22.13
Phenyl, neutral conditions	1.68
Phenyl, 2% formic acid	2.05
Phenyl, weakly acidified sample	3.41
Phenyl, strongly acidified sample	5.67
SCX, neutral conditions	0
SCX, 2% formic acid	0
SCX, weakly acidified sample	2.57
SCX, strongly acidified sample	14.67
SAX, neutral conditions	12.11

SAX, 2% formic acid	5.28
SAX, weakly basified sample	2.24
SAX, strongly basified sample	0.35
PBA, phosphate buffer	0.93

These results demonstrate that the method with the most wide-ranging retention utilises a strong acidified (pH 2) sample with an HLB cartridge, whereas the least retaining methods utilise SCX with no sample acidification.

3.2.2 Annotation of key retained metabolites

An alternative approach to quantifying the retention profiles of each SPE cartridge can be achieved using PCA. PCA is a statistical tool that can be used to reduce the dimensionality of data and produce visual representations of correlations between datasets — for NMR spectra, it allows for clustering and trends between experiments to be demonstrated, as well as for discovering potential outliers. A PCA structure built with the datasets from all elution methods utilising natural urine demonstrated that the elutions from acidified ion exchange methods were clearly separated from the other chemistries (fig 3.2). Table 3.5 lists the NMR signals visible in the PCA loadings plot, as well as their correlation coefficient and their tentative assignment — all assignments are made with comparison to the reported values in the literature, and hence can be considered annotated to a level 2 standard[19]. It demonstrates that a small number of peaks (creatinine, histidine, creatine, trimethylamine-N-oxide, and 3-methylhistidine being the most prominent) contributed to 69.19% of the difference between elutions. These metabolites were all retained by SCX cartridges under acidic conditions, and their spectral peaks are sensitive to pH changes, causing significant chemical shifts even in buffered samples.
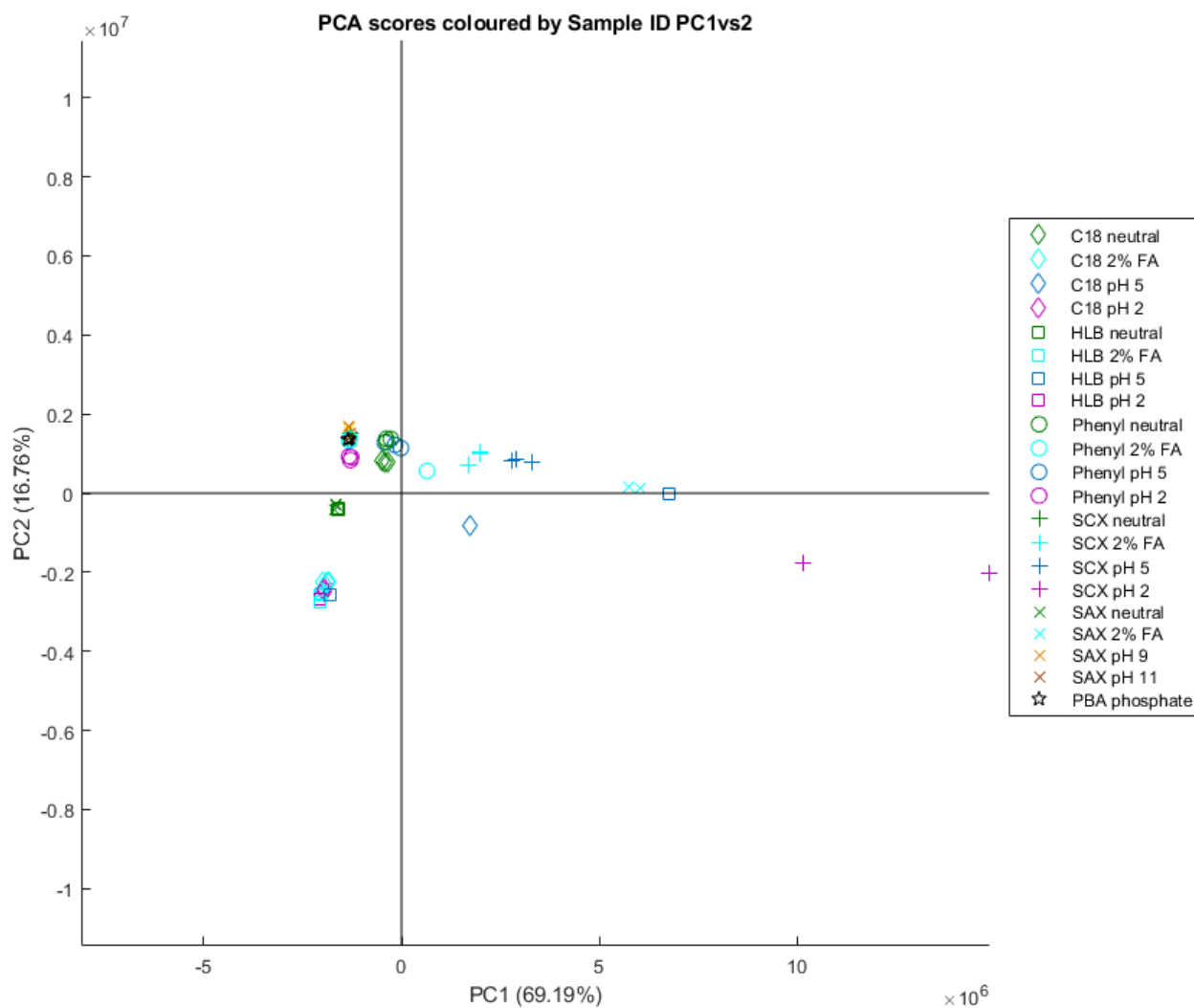
Figure 3.2: PCA scores plot built using NMR data from all SPE elutions of natural urine, PC1 (69.19%) vs PC2 (16.76%).

Table 3.5: Natural urine all elutions, assignments from PC1 loadings plot

Peak (ppm)	Assignment	Pearson's correlation
7.96 (s)	3-Methylhistidine	0.64
7.90 (s)	Histidine	0.89
7.70 (s)	1-Methylhistidine	0.81
7.10 (s)	Histidine	0.95
7.08 (s)	3-Methylhistidine	0.90
7.02 (s)	1-Methylhistidine	0.92

4.06 (s)	Creatinine	1.00
4.05 (d)	Unknown A	0.79
4.00 (dd)	Histidine	0.90
3.72 (s)	3-Methylhistidine	0.89
3.70 (s)	3-Methylhistidine	0.80
3.29 (s)	**Unknown B**	0.93
3.27 (s)	TMAO	0.96
3.25 (s)	Unknown B	0.94
3.24 (s^a)	Unknown C	0.94
3.23 (d^a)	Unknown D	0.95
3.21 (s)	Unknown E	0.84
3.18 (s)	Unknown F	0.87
3.16 (s)	Unknown G	0.93
3.14 (s)	Unknown H	0.89
3.05 (s)	Creatinine	0.99
3.04 (s)	Creatine	0.83

Unknowns in bold have both ^{1}H and ^{13}C signals identifiable, but could not be matched to compounds in metabolite databases. ^{13}C signals could not be identified for unbolded unknowns.

The ion exchange method most separated from other datasets is the one which produces elutions from strongly acidified urine using an SCX cartridge. However, there is also significant differentiation with the SAX elutions with 2% formic acid — under these conditions, the SAX cartridge begins to retain compounds like creatinine and histidine where it otherwise wouldn't under neutral or basic conditions. It is not clear why the SAX retention profile would begin to resemble that of SCX; one possibility could be that the silanol groups from the silica on which the SO_3^- and ammonium modifications are based are more able to retain these compounds under acidic conditions. However, silanols are usually protonated under acidic conditions, and hence would not express this ionic character; the explanation additionally doesn't account for why C_{18}

with 2% formic acid — which, similarly, contains silanol groups — does not retain creatinine or histidine.

Besides the separation between ion exchange and reversed phase methods, there is additional separation along the secondary principal component axis between phenyl, C_{18} neutral, SAX neutral, and SCX neutral elutions, and C_{18} acidified and HLB acidified elutions — with HLB neutral elutions found in between the two clusters (fig 3.3). Removing elutions from ion exchange cartridges and reconstructing the PCA affords a similar differentiation with more clarity. The clusters suggest that C_{18} neutral elutions have more in common with phenyl elutions than with acidified C_{18} or even HLB neutral elutions.
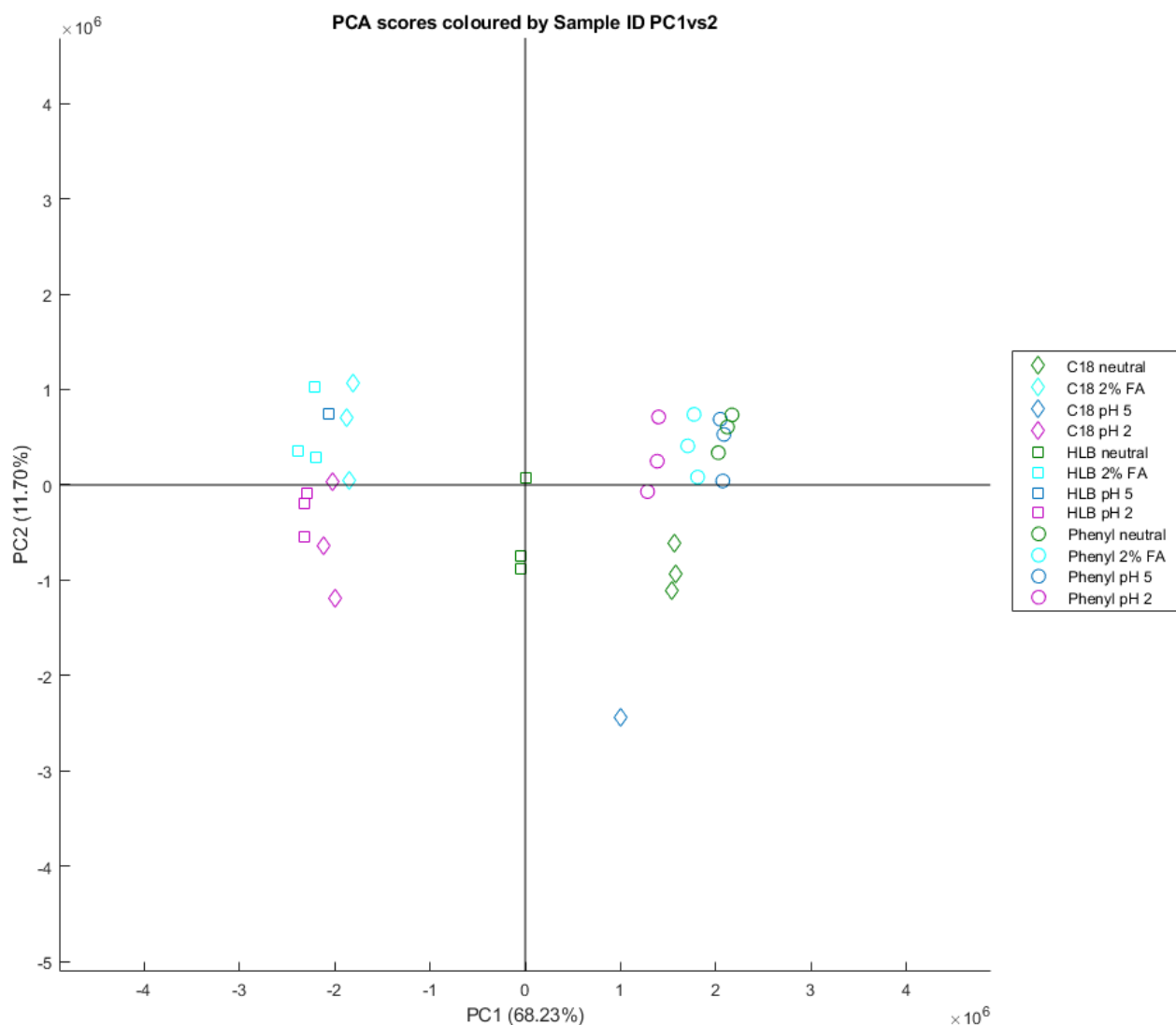
Figure 3.3: A PCA scores plot built using NMR data from reversed-phase SPE elutions of natural urine, PC1 (68.23%) vs PC2 (11.70%).

The NMR signals responsible for the separation between reversed phase elutions are tabulated in table 3.6 — it also notes whether the annotated metabolites or unknowns are positively correlated with methods utilising phenyl cartridges (and hence better retained by phenyl), or are negatively correlated (and hence better retained by C_{18}/HLB).

Signals marked with an asterisk * are not visible in the loadings plot (as the area is removed in pre-processing), but can be observed by manual inspection of the spectra. Compounds in italics were confirmed by NMR spike-in experiment — the methods detailing the identification process for selected compounds are described in chapter 4.

Table 3.6: Natural urine reversed-phase elutions, assignments from PC1 loadings plot

Peak (ppm)	Assignment	Pearson's correlation	Phenyl correlation
9.13 (s)	*Trigonelline*	0.78	Positive
8.84 (t)	*Trigonelline*	0.77	Positive
[8.45 (d)]	Quinolinate	*	Negative
[8.34 (d)]	N-*Methyl-2-pyridone-5-carboxamide*	*	Negative
8.09 (t)	*Trigonelline*	0.76	Positive
8.03 (s)	3-Methylxanthine	0.84	Negative
8.00 (s)	*Phenylacetylglutamine*[95]	0.91	Negative
7.97 (dd)	N-*Methyl-2-pyridone-5-carboxamide*	0.98	Negative
7.94 (s)	**Unknown I**	0.77	Negative
7.91 (d[a])	Unknown J	0.93	Negative
7.87 (d)	Unknown K	0.90	Negative
7.84 (dd)	Hippurate	0.98	Negative
7.77 (d[a])	4-Hydroxyhippurate[a]	0.92	Negative
7.70 (dd)	2-Furoylglycine	0.92	Negative
7.64 (tt)	Hippurate	0.98	Negative
7.56 (t)	Hippurate	0.99	Negative

7.46 (dd)	Quinolate	0.93	Negative
7.43 (m)	*Phenylacetylglutamine*	0.97	Negative
7.37 (m)	*Phenylacetylglutamine*	0.97	Negative
7.30 (t)	*3-hydroxyhippurate*	0.98	Negative
7.30 (s)	*4-cresol sulfate*	0.97	Negative
7.28 (s)	*4-cresol sulfate*	0.79	Negative
7.24 (d)	Unknown M	0.96	Negative
7.2 2(s)	*4-cresol sulfate*	0.97	Negative
7.20 (s)	*4-cresol sulfate*	0.82	Negative
7.19 (d)	2-Furoylglycine	0.92	Negative
7.17 (d)	*4-Hydroxyphenylacetate*	0.94	Negative
7.12 (ddd)	*3-hydroxyhippurate*	0.96	Negative
7.07 (m)	**Unknown O**	0.98	Negative
6.98 (d)	*HPHPA*	0.94	Negative
6.93 (t)	*HPHPA*	0.89	Negative
6.87 (d)	*4-Hydroxyphenyl-acetate*	0.95	Negative
6.85 (dd)	*HPHPA*	0.92	Negative
6.68 (d)	N-*Methyl-2-pyridone-5-carboxamide*	0.97	Negative
6.64 (dd)	2-Furoylglycine	0.86	Negative
4.19 (td)	*Phenylacetyl-glutamine*	0.98	Negative
4.06 (s)	Creatinine	0.55	Positive
4.01 (dd)	Phenylalanine	0.86	Negative
3.97 (d)	Hippurate	0.98	Negative
3.96 (s)	*3-Hydroxyhippurate*	0.97	Negative
3.94 (s^a)	Unknown Q	0.99	Negative
3.93 (s)	2-Furoylglycine	0.93	Negative
3.92 (s^a)	Unknown R	0.96	Negative

3.87 (s/t^a)	Unknown S	0.96	Negative
3.70 (s)	**Unknown T**	0.96	Negative
3.67 (d)	*Phenylacetylglutamine*	0.97	Negative
3.65 (s)	N-*Methyl-2-pyridone-5-carboxamide*	0.98	Negative
3.63 (s)	**Unknown T**	0.98	Negative
3.54 (s)	Unknown V	0.92	Negative
3.53 (s)	**Unknown W**	0.99	Negative
3.49 (s)	**Unknown X**	0.88	Negative
3.48 (s)	**Unknown Y**	0.96	Negative
3.45 (s)	Unknown Z	0.98	Negative
3.34 (s)	**Unknown AA**	0.96	Negative
3.32 (s)	**Unknown AB**	0.96	Negative
3.30 (s)	Unknown AC	0.97	Negative
3.27 (s)	Unknown AD	0.67	Positive
3.17 (s)	**Unknown AE**	0.98	Negative
3.11 (s)	**Unknown AF**	0.79	Positive
3.05 (s)	Creatinine	0.55	Positive
3.00 (s)	Unknown AG	0.91	Negative
2.68 (m)	*HPHPA*	0.89	Negative
2.63 (d)	Unknown T	0.92	Negative
2.35 (s)	*4-cresol sulfate*	0.78	Negative
2.27 (t^a)	*Phenylacetyl-glutamine*	0.98	Negative
2.12 (m^a)	*Phenylacetyl-glutamine*	0.98	Negative
1.93 (m^a)	*Phenylacetyl-glutamine*	0.98	Negative
1.31 (m)	Unknown BX	0.82	Negative
0.94 (m^a)	Unknown AN	0.92	Negative
0.89 (t)	Unknown BY	0.92	Negative

The PC3 vs. PC4 plot in the all-elutions structure also demonstrated clustering of the PBA elutions (fig 3.4), positively correlated with the PC4 dimension. The PC4 loadings hence closely resemble the averaged spectra from the PBA elutions — the metabolites (table 3.7) being mostly represented by mannitol and N-methylnicotinamide.
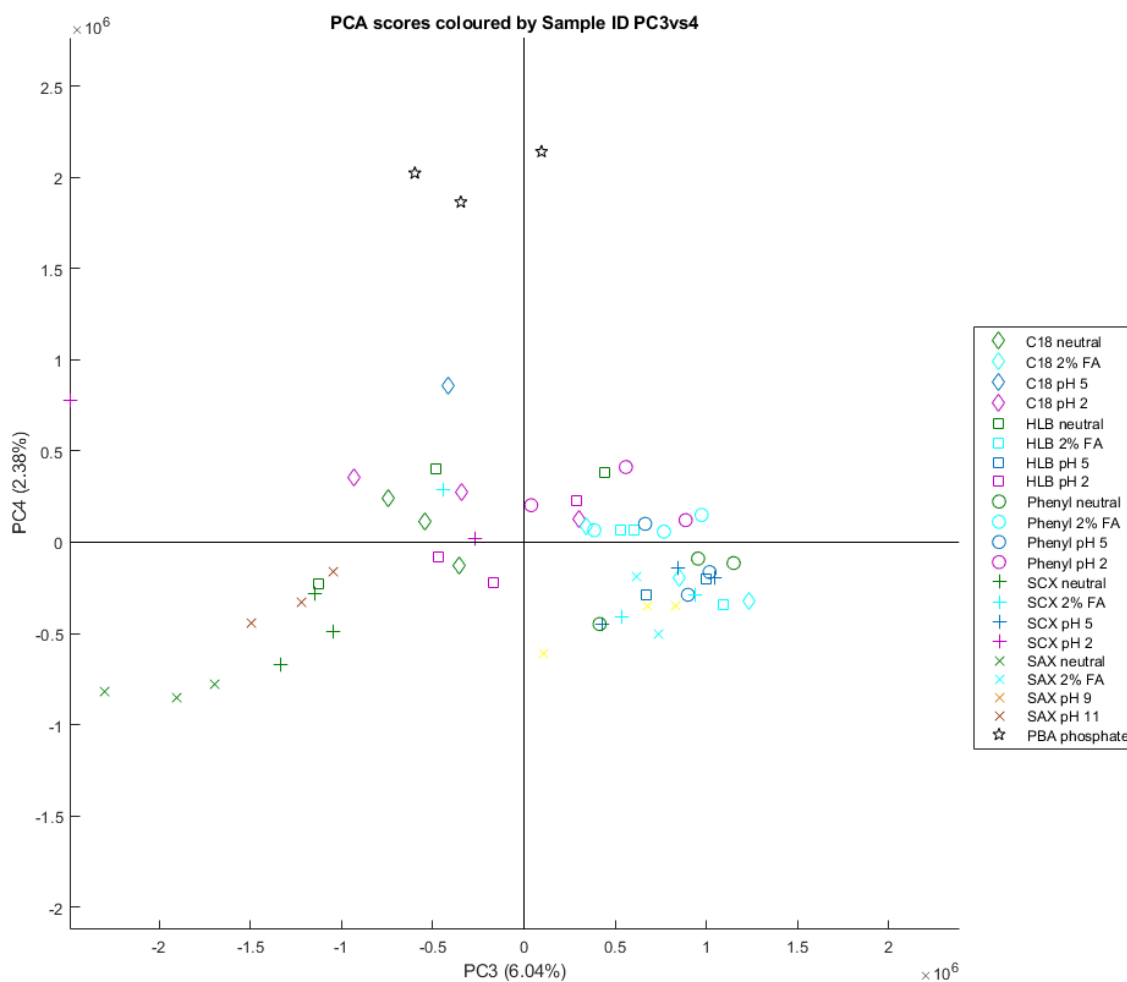
Figure 3.4: A PCA scores plot built using NMR data from all SPE elutions of natural urine, PC3 (6.04%) vs PC4 (2.38%).

Table 3.7: Natural urine all elutions, assignments from PC4 loadings plot

Peak (ppm)	Assignment	Pearson's correlation
9.29 (s)	*1-Methylnicotinamide*	0.82
8.97 (d)	*1-Methylnicotinamide*	0.72
8.90 (d)	*1-Methylnicotinamide*	0.71

8.79 (d)	**Unknown BQ**	0.69
[8.53 (t)]	**Unknown BQ**	*
8.06 (t^a)	**Unknown BQ**	0.63
7.68 (s)	**Unknown BR**	0.88
5.85 (d)	Unknown BS	0.49
[4.31 (t)]	Unknown BS	*
4.16 (t)	**Unknown BS**	0.53
4.11 (s)	Unknown BT	0.68
4.02 (t)	Unknown BU	0.73
3.89 (d)	*Mannitol*	0.86
3.87 (d)	*Mannitol*	0.87
3.82 (s)	*Mannitol*	0.86
3.80 (s)	*Mannitol*	0.84
3.77 (m)	*Mannitol*	0.87
3.67 (d^a)	Unknown BV	0.87
3.69 (dd)	*Mannitol*	0.86
3.20 (s)	Unknown BW	0.79
2.76 (t)	Unknown BZ	0.76
2.72 (s)	DMA	0.85
2.01 (s)	*Acetamide*	0.82

A similar PCA can be constructed for reversed phase elutions utilising artificial urine — similarly, there is separation across the first component, demonstrating a notable difference between phenyl and C_{18}/HLB retention profiles. As with the natural urine elutions, the loadings (table 3.8) can be annotated to demonstrate the most important spectral differences separating different methods — where the correlation with the phenyl datasets are positive the compound is more likely to appear in SPE elutions utilising phenyl cartridges, and vice versa.

Table 3.8: Artificial urine reversed-phase elutions, assignments from PC1 loadings plot

Peak (ppm)	Assignment	Pearson's correlation	Phenyl correlation
9.13 (s)	Trigonelline	0.95	Positive
8.84 (t)	Trigonelline	0.97	Positive
8.09 (t)	Trigonelline	0.92	Positive
7.97 (d)	Hippurate	0.97	Negative
7.88 (m)	Benzoate	1.00	Negative
7.84 (dt)	Hippurate	1.00	Negative
7.78 (t)	Hippurate[a]	0.97	Negative
7.70 (t)	Hippurate[a]	0.97	Negative
7.64 (tt)	Hippurate	1.00	Negative
7.56 (m)	Hippurate	1.00	Negative
7.49 (m)	Benzoate	1.00	Negative
7.44 (tt)	Phenylalanine	0.99	Negative
7.38 (tt)	Phenylalanine	0.99	Negative
7.34 (m)	Phenylalanine	0.99	Negative
7.17 (dt)	4-Hydroxyphenyl-acetate	0.93	Negative
6.87 (dt)	4-Hydroxyphenyl-acetate	0.93	Negative
6.59 (s)	trans-Aconitate	0.85	Negative
4.00 (q)	Phenylalanine	0.96	Negative
3.97 (d)	Hippurate	0.99	Negative
3.30 (d)	Phenylalanine[a]	0.97	Negative
3.27 (s)	TMAO	0.91	Positive
3.13 (m)	Phenylalanine	0.99	Negative
3.05 (s)	Creatine	0.93	Positive
2.66 (d)	Citrate	0.91	Negative
2.55 (d)	Citrate	0.94	Negative

2.41 (s)	Succinate	0.84	Negative

Where ^{a} indicates a tentative assignment.

Here, artificial urine was used to demonstrate that a mixture of representative compounds can be used to estimate the retention capacity of cartridges without using natural urine — as with the natural urine reversed phase elutions, metabolites such as hippurate can be shown to be retained on C_{18}/HLB, but not on phenyl cartridges; similarly, trigonelline can be shown to be retained on phenyl, but not on C_{18}/HLB. This allows for greater control over future SPE experiments aimed at characterising retention profiles of SPE cartridges.

Clustering can be observed forming an almost linear scale for C_{18} and HLB cartridges (fig 3.5) — with C_{18} neutral elutions at one end, acidified elutions at the other, and HLB neutral elutions in between.
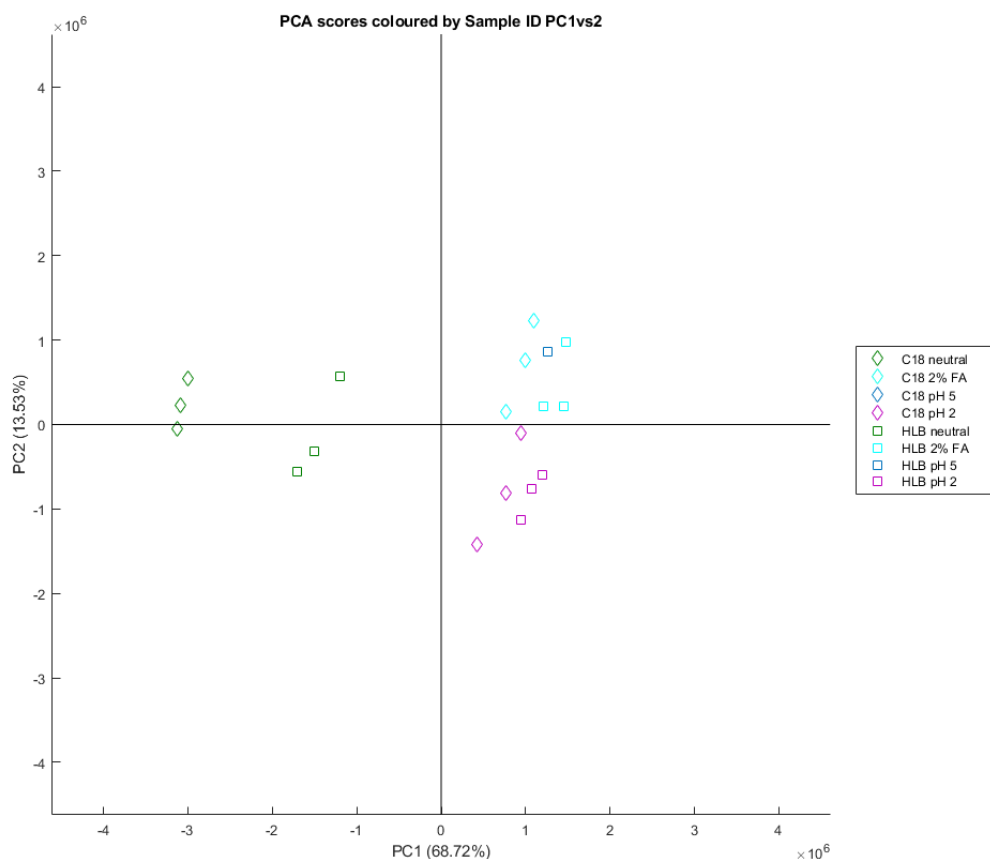
Figure 3.5: A PCA scores plot built using NMR data from C_{18} and HLB elutions of natural urine, PC1 (68.72%) vs PC2 (13.53%).

Many of the assigned peaks (table 3.9) in the aromatic region are caused by differences in

chemical shift between identical compounds (likely due to pH differences) — for example, *N*-methyl-2-pyridone-5-carboxamide (2-PY) is significantly correlated both positively and negatively with phenyl elutions (fig 3.6), as its spectral peaks undergo chemical shifting due to pH differences in different experiments.
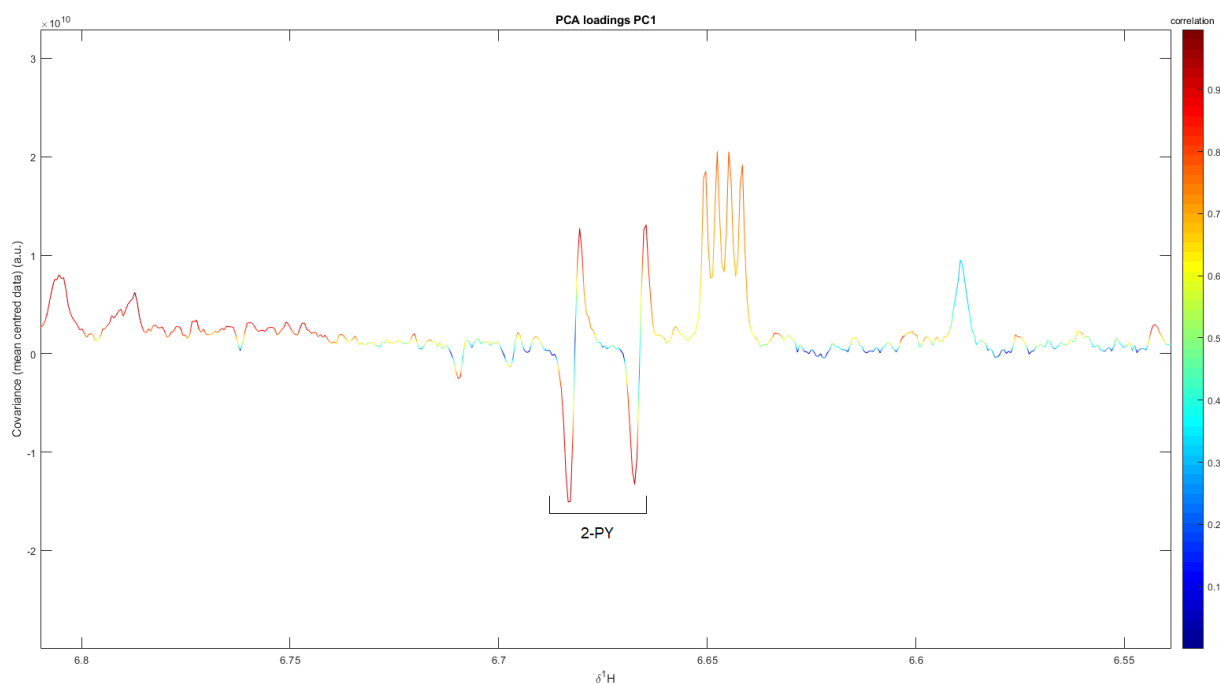
Figure 3.6: The PC1 loadings plot for C_{18}/HLB elutions, demonstrating both positive and negative correlation of 2PY.

Table 3.9: Natural urine C_{18} and HLB elutions, assignments from PC1 loadings plot

Peak (ppm)	Assignment	Pearson's correlation	HLB correlation
8.05 (m[a])	Unknown AN	0.87	Positive
7.97 (br s[a])	Unknown AO	0.91	Positive
7.97 (dd)	N-*Methyl-2-pyridone-5-carboxamide*	0.91	Negative
7.96 (dd)	N-*Methyl-2-pyridone-5-carboxamide*	0.96	Positive
7.94 (s)	**Unknown I**	0.70	Positive
7.93 (s)	Unknown AP	0.83	Negative
7.92 (s)	Unknown AQ	0.82	Negative
7.91 (s)	Unknown AR	0.91	Positive

7.91 (s)	Unknown AS	0.61	Negative
7.90 (s)	Unknown AT	0.84	Positive
7.87 (d)	Unknown K	0.83	Positive
7.84 (dd)	Hippurate	0.91	Positive
7.76 (d[a])	4-Hydroxyhippurate[a]	0.88	Positive
7.70 (dd)	2-Furoylglycine	0.81	Positive
7.64 (tt)	Hippurate	0.99	Positive
7.56 (t)	Hippurate	0.99	Positive
7.43 (m)	*Phenylacetylglutamine*	0.95	Positive
7.37 (m)	*Phenylacetylglutamine*	0.95	Positive
7.31 (s)	Unknown AU	0.83	Positive
7.30 (t)	*3-hydroxyhippurate*	0.92	Positive
7.19 (d)	2-Furoylglycine	0.80	Positive
7.17 (d)	4-Hydroxyphenyl-acetate	0.85	Positive
7.12 (ddd)	*3-hydroxyhippurate*	0.96	Positive
6.98 (d)	4-Hydroxyhippurate[a]	0.85	Positive
6.87 (d)	*4-Hydroxyphenyl-acetate*	0.90	Positive
6.81 (s/t[a])	Unknown AV	0.92	Positive
6.68 (d)	N-*Methyl-2-pyridone-5-carboxamide*	0.92	Positive
6.67 (d)	N-*Methyl-2-pyridone-5-carboxamide*	0.91	Negative
4.19 (td)	*Phenylacetylglutamine*	0.96	Positive
4.06 (s)	Creatinine	0.80	Negative
3.97 (d)	Hippurate	0.99	Positive
3.96 (s)	*3-hydroxyhippurate*	0.94	Positive
3.94 (d[a])	Unknown AX	0.92	Negative
3.93 (d[a])	Unknown AY	0.95	Positive
3.93 (s)	Unknown AZ	0.88	Positive

3.87 (s/t^a)	**Unknown S**	0.98	Positive
3.70 (s)	**Unknown T**	0.98	Positive
3.67 (d^a)	**Unknown U**	0.95	Positive
3.65 (s)	N-*Methyl-2-pyridone-5-carboxamide*	0.92	Negative
3.65 (s)	N-*Methyl-2-pyridone-5-carboxamide*	0.97	Positive
3.63 (s)	Unknown T	0.97	Positive
3.53 (s)	Unknown BC	0.79	Negative
3.53 (s)	3-Methylxanthine^a	0.95	Positive
3.49 (s)	**Unknown X**	0.81	Negative
3.49 (s)	Unknown BE	0.84	Positive
3.48 (s)	Unknown BF	0.90	Positive
3.46 (s)	Unknown BG	0.93	Positive
3.45 (s)	Unknown BH	0.88	Positive
3.36 (s)	Unknown BI	0.87	Negative
3.35 (s^a)	Unknown BJ	0.86	Negative
3.34 (s^a)	**Unknown AA**	0.95	Positive
3.33 (s^a)	Unknown BL	0.83	Negative
3.32 (s^a)	Unknown BM	0.93	Positive
3.30 (s^a)	Unknown BN	0.96	Positive
3.27 (s)	TMAO	0.80	Negative
3.17 (s)	**Unknown AE**	0.88	Negative
3.17 (s)	Unknown BP	0.95	Positive
3.05 (s)	Creatinine	0.80	Negative
2.27 (t)	*Phenylacetyl-glutamine*	0.97	Positive
2.12 (m)	*Phenylacetyl-glutamine*	0.97	Positive
0.93 (m)	*Phenylacetyl-glutamine*	0.97	Positive

The unknown compounds with both ^{1}H and ^{13}C signals associated which weren't able to be identified by database search are hence listed in table 3.10.

Table 3.10: Unknown compounds with ^{1}H and ^{13}C signals

Identifier	^{1}H Peaks (ppm)	^{13}C Peaks	Elution
Unknown B	3.29 (s)	27.57	HLB
Unknown I	7.94 (s)	141.11	HLB
Unknown O	7.07 (m)	116.68	HLB
Unknown S	3.87 (s/t[a])	55.92	HLB
Unknown T	3.70 (s), 3.63 (s)	45.3, 71.58	HLB
Unknown U	3.67 (d[a])	45.3	HLB
Unknown W	3.53 (s)	32.08	HLB
Unknown X	3.49 (s)	31.80	HLB
Unknown Y	3.48 (s)	31.80	HLB
Unknown AA	3.34 (s)	30.60	HLB
Unknown AB	3.32 (s)	30.22	HLB
Unknown AE	3.17 (s)	40.21	HLB
Unknown AF	3.11 (s)	48.7[a]	C$_{18}$
Unknown BQ	8.79 (d), 8.53 (t), 8.06 (t[a])	148.07, 148.63, 130.80	PBA
Unknown BR	7.68 (s)	144.32	PBA
Unknown BS	4.16 (t)	73.50	PBA

3.3 Discussion

The guiding philosophy behind this use of SPE-NMR suggests that not only each cartridge, but each pH and solvent system utilised in a given experiment, would result in different retention profiles. These retention profiles can be classified either through annotation of a selection of common metabolites from different compound classes — or, using a more holistic approach, determining the compounds more likely to be retained under different conditions through data treatment. Selective use of methods can then be utilised to reduce peak overlap and aid metabolite identification. One example of this is displayed in fig 3.7; interferences in the 'raw' pooled urine sample are removed by the use of a PBA-based method to clearly reveal mannitol.
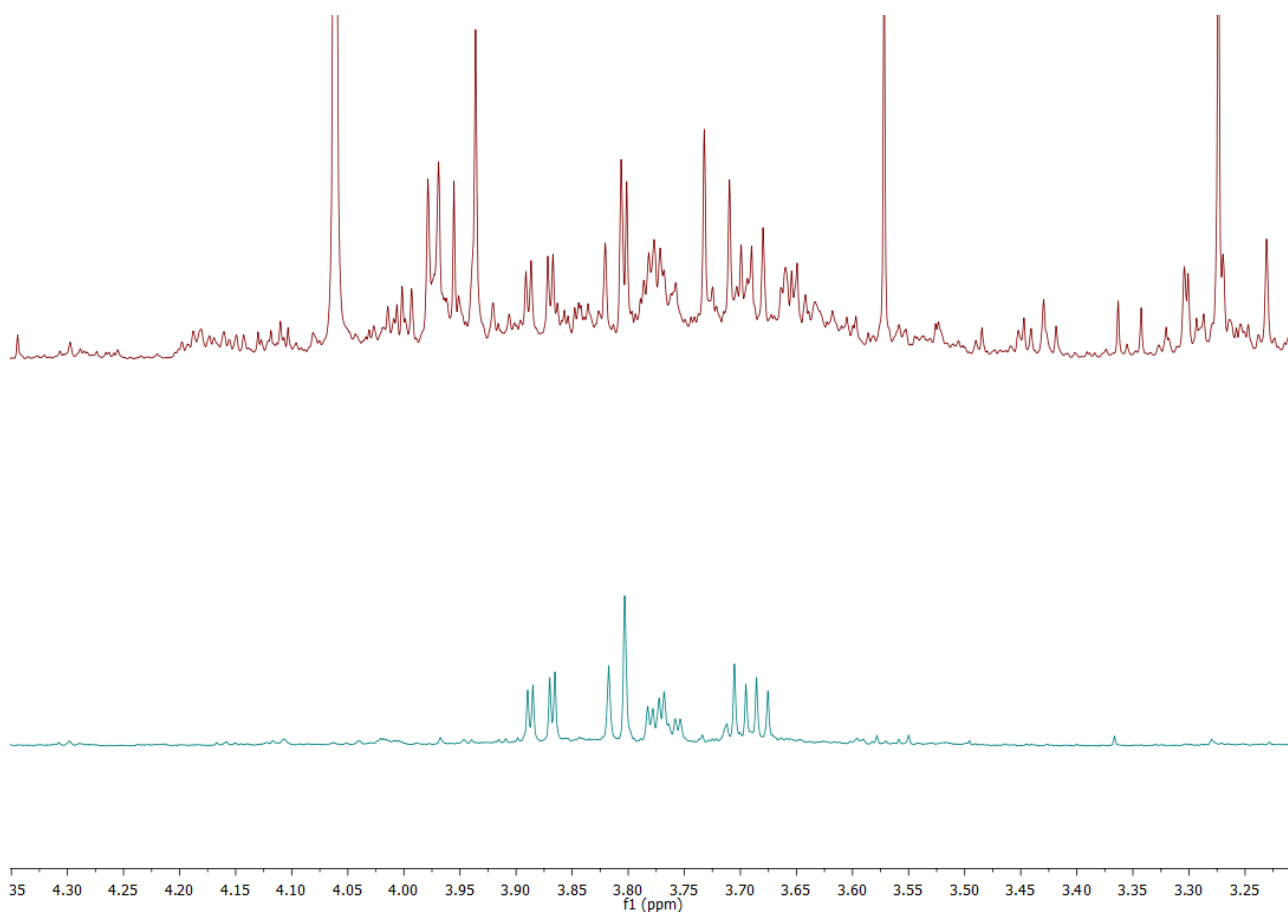
Figure 3.7: Comparison of the 3.30–4.30 ppm region of the 600 MHz ^{1}H NMR spectra of the pooled urine sample (top) and PBA SPE-treated urine (bottom), the latter revealing only mannitol peaks.

Annotation of a selection of common metabolites is facile and provides immediate and useful information of individual compounds. Ideally, this could be done using a list of metabolites

representative of the compound classes generally found in human urine; unfortunately, the partial identification of the human urinary metabolome hinders the creation of a fully representative sample. Additionally, the natural rate of occurrence of metabolites may make a representation of an 'average' sample difficult or even impossible. Peak intensities can also be impacted by NMR shimming, peak overlap, and pH changes – all of which can affect the intensity recorded. Despite these shortcomings, general trends can be established by considering functional groups and structural commonalities between compound classes.

Clustering in PCA plots can be used to demonstrate substantial differences between datasets. The Strong Cation Exchange (SCX) cartridge — utilising the pseudo-permanently charged phenylsulfonic derivative ($pK_a \approx 2.1$) — provides the greatest separation between clusters when included in PCA structures, due to the ion exchange mechanisms not present in reversed phase chromatography. Ion exchange retention profiles rely heavily on pH control, since all compounds must have at least one positively charged atom in order to have sufficient attraction to the sorbent to be retained; hence, compounds that do not have a positive charge at physiological pH must be in acidic solution for retention to occur. This intrusive sample adjustment will naturally affect the chemistry of the biofluid; using acidified (or basified) conditions is, then, necessarily a trade-off between greater insight into the metabolome through retention and authenticity of the sample itself. It is additionally feasible that compounds not normally present in the sample may be formed and retained due to the change of conditions, although this was not noted during the course of the experiments.

The importance of pH control is reflected in the retention capacity of SCX cartridges; the neutral pH retention profile is one of the least retaining methods with an estimated retention capacity of 0 — at pH 2, its retention capacity (14.67%) is comparable to a more widely recognised reversed phase method, such as HLB with 2% formic acid in all steps (17.40%). The compounds best retained on SCX under acidic conditions were predominantly histidine-based — with histidine, 3-methylhistidine, and 1-methylhistidine being well retained at pH 2. Creatinine and TMAO were also present in the elutions. A cationic nitrogen atom, possibly stabilised by electron-donating groups through hyperconjugation, may serve as the most important characteristic uniting these compounds. Other compounds without nitrogen-containing functional groups were

generally not present in the acidified SCX elutions, although the presence of a nitrogen-containing functional group did not necessarily result in retention — for example, of the proteinogenic amino acids that were retained, only histidine was retained in any significant quantity. It may be notable that histidine has a pK_a of 6.04 (pyrrolic nitrogen), far lower than the other positively charged amino acids, arginine (pK_a 12.10) and lysine (pK_a 10.67).

Removing ion exchange elutions from the dataset and reconstructing a PCA plot demonstrates additional separation between C_{18}/HLB and phenyl elutions and allows for further probing into the differences between the reversed phase methods. Previous uses of phenyl cartridges in the literature have remarked on their similar retention capabilities to C_{18}, with slightly better retention for polycyclic aromatic compounds[96], but slightly worse retention for other hydrocarbons[97]. Phenyl cartridges utilise π-stacking on top of hydrophobic forces in order to provide additional retention for aromatic compounds — however, the strength of π–π interactions tends to be bound between around 8–12 kJ mol^{-1} for benzene dimers[46]. For comparison, hydrophobic forces may be up to 4 times stronger[44]; hence, despite having an additional mechanism of action, the actual retention capacity for phenyl cartridges is significantly lower than that of C_{18} or HLB cartridges across all methods due to a weaker hydrophobic retention mechanism.

The two major compounds that were retained selectively by phenyl (but not by C_{18} or HLB) were trigonelline and creatinine, both nitrogen-containing heterocycles. Other compounds with phenyl functional groups — such as phenylalanine or hippurate — did not experience greater retention using phenyl cartridges, and in fact retained much less, if at all. Conversely, the cyclic metabolites retained under acidic conditions by C_{18}/HLB — but not by phenyl — are predominantly aromatic, with both heterocycles (2-furoylglycine, quinolinate, and N-methyl-2-pyridone-5-carboxamine), and hydrocarbon rings (hippurate, phenylacetylglutamine, and 4-hydroxyphenylacetate) present. The aromatic heterocycles here do not have the ability to form cationic nitrogen in the rings themselves, unlike the aromatic compounds retained in phenyl elutions, such as trigonelline. Short chain fatty acids such as valeric acid were also retained under non-neutral conditions. It is unclear why charged metabolites like trigonelline (fig 3.8) would be better retained on phenyl cartridges — although, as the charge is positive, it

is hypothetically possible that the π-electron clouds located above and below the benzene rings are able to electrostatically attract these metabolites with enough strength that they can be retained.
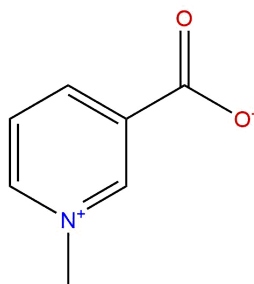
Figure 3.8: Chemical structure of trigonelline.

Out of the remaining reversed phase sorbents, the C_{18} and HLB cartridges are known to have similar retention profiles to each other[83], with HLB cartridges often being preferred for their tolerance to drying and the possibility for elimination of conditioning and equilibration steps. The differences between the two can be demonstrated through comparison of elutions — while the two have comparable retention, HLB cartridges tend to retain a slightly larger range of compounds in greater quantities. This is especially true under neutral conditions, where C_{18} cartridges retain relatively little. Generally speaking, as with the SCX cartridges, more acidic conditions result in greater retention — possibly due to the deionisation of silanol groups on the surface of the sorbent. However, this again comes with the trade-off of authenticity, as the acidic conditions may cause signal suppression, unwanted reactions between metabolites, or general degradation of the sample itself. Use of 2% formic acid in all steps allows for a balance between the two – while the retention does not extend as deeply as that under pH 2 conditions, the higher pH environment should not be as destructive to sample authenticity, and chemical shifts caused by drastic pH changes should be absent. HLB and C_{18} under acidic conditions are powerful analytical methods that can reveal signals that are otherwise not visible — for example, 3-hydroxyhippurate, which displays a doublet of doublet of doublets (ddd) signal at 7.12 ppm, is normally obscured by a 3-methylhistidine peak (fig 3.9). The use of reversed phase methods can hence provide additional information about the human urinary metabolome.

C_{18} and HLB cartridges themselves can also be differentiated from each other. On top of the obvious differences in sorbent structure (a hydrocarbon chain, compared to a polymer containing

divinylbenzene), HLB cartridges are not silica-based — hence, silanol groups present in C_{18} cartridges are not present in HLB. These silanol groups produce secondary interactions with metabolites, commonly expressed as a weak cation exchanger, which can influence retention. This is reflected in the differences between C_{18} elutions under neutral and acidified conditions — at lower pH, the silanols are generally protonated; at neutral pH, at least some silanols are deprotonated, giving the cartridge the ability to selectively retain some cations. Indeed, observing the PCA results shows that compounds such as creatinine and TMAO are retained under neutral conditions — these compounds also being retained by the SCX cartridge under the appropriate conditions.
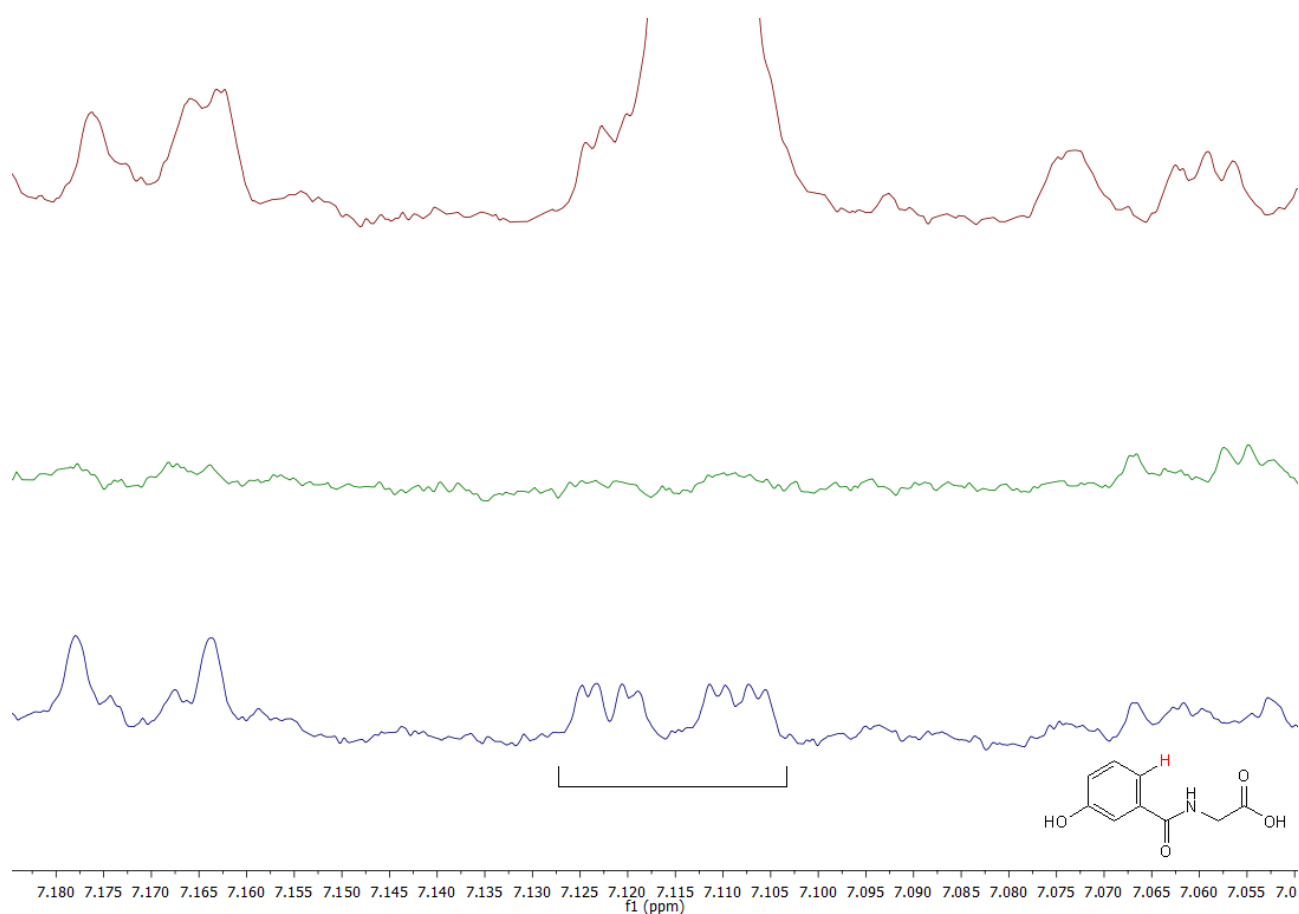
Figure 3.9: The ^{1}H NMR multiplet of 3-hydroxyhippurate (ddd), normally obscured by that of 3-methylhistidine (top), is revealed after HLB SPE treatment under acidic conditions (bottom), but not under neutral conditions (middle).

The final elutions to consider were those afforded from methods utilising phenylboronic acid (PBA) cartridges. PBA cartridges utilise a unique covalent bonding mechanism in order to selectively retain diols, -hydroxy ketones, or any other functional groups where two unsubstituted

heteroatoms are separated by at least one carbon[73,98]. It is not clear whether the diols must adopt a specific isomerism for retention to occur: mannitol is heavily retained in the elutions, but contains both R and S carbon centres, as well as terminal hydroxyls which can rotate to become a given conformer — its retention hence does not give additional insight. Other compounds retained include acetate, acetamide, and *N*-methylnicotinamide, a metabolite of niacin. The presence of adjacent heteroatoms does not guarantee good retention: for example, citric acid is poorly retained, despite having three carboxylate groups. There is also some retention of dimethylamine in both artificial and natural urine samples, despite it not being a diol — however, it could hypothetically be retained through a single substitution of water at the boronate, rather than through a double-substitution, as is normally the case.

3.4   Conclusion

Having profiled the retentive capacity of several SPE methods, there remained a significant number of peaks requiring annotation and identification; in some cases, this required the use of orthogonal analytical tools to complement the information ascertained via NMR. This would give a further understanding of the compounds being retained — and hence a fuller picture of the retention profiles of the methods — before they could be translated into automated protocols.

Chapter 4

Use of SPE-NMR for metabolite annotation and structural elucidation

4.1   SPE methods utilising gradient elutions

The SPE protocols developed in chapter 3 were necessarily isocratic, as the automated system they were to be transferred to (chapter 5) has limited capacity for additional solvents. Nevertheless, the use of gradient techniques with SPE for fractionation promises greater insight into sample composition.

Solid phase extraction is typically coupled with either on-line or off-line liquid chromatography as part of a broader hyphenated technique, when applied to complex mixtures — this might be prior to chromatography, increasing the concentration of target compounds (SPE-HPLC)[99]; alternatively, the SPE cartridge might be placed such as to concentrate eluted compounds, remove matrix effects, or enable solvent change (HPLC-SPE)[100,101].

The use of gradient elutions in solid phase extraction enables further fractionation of samples — in contrast with the retained/not retained binary of SPE demonstrated hitherto, the use of multiple solvents of increasing eluotropic strength allows for greater insight into metabolite retention in biofluids such as urine. This fractionation also allows for even greater reduction of spectral peak overlap, further increasing understanding of the urinary metabolome.

During the course of the project, gradient elution methods on C_{18} and HLB cartridges were

applied to the pooled urine sample. Initially this utilised only neutral conditions — however, the consequent disappointing retention capacity of the neutral methods, and the superiority of the HLB cartridge over over reversed-phase methods, led to a greater focus on the HLB gradient method utilising 2% formic acid in the wash step. As previously discussed, HLB cartridges carry the unique benefit of not requiring conditioning and equilibration steps for retention.

4.1.1 Methodology

Solid phase extraction

Pooled urine (3 mL) was loaded onto a Waters Oasis HLB cartridge (6 mL capacity and 500 mg bed weight). This was then washed with 2% formic acid in water (6 mL) to eliminate interferences. The retained metabolites were then eluted stepwise with 6mL respectively of methanol-water mixtures v/v (20:80, 40:60, 60:40, 80:20), and finally with 6mL 100% methanol.

NMR sample preparation

Washes and elutions were dried under nitrogen and reconstituted in ultrapure water (3 mL). Buffer containing trimethylsilylpropionate (TSP) as a chemical shift reference standard was added to 540 µL of reconstituted sample, as described by Dona *et al.*[88]. 580 µL of the manually vortexed sample was then transferred into 5mm SampleJet NMR racks.

Samples which required additional 2D NMR experiments were dried under nitrogen and reconstituted in D_2O (3 mL). TSP phosphate buffer (60 µL) was added to 540 µL of reconstituted sample, and 580 µL of the resulting manually vortexed sample was transferred into 5mm NMR tubes.

NMR data acquisition

All 1D experiments were run using a Bruker Avance III 600 MHz spectrometer equipped with a BBI room temperature probe and SampleJet. Samples were analysed using one-dimensional water-suppressed ^{1}H NOESY experiments at 300 K.

Additional ^{1}H–^{1}H J-resolved experiments, and 2D-NMR experiments including ^{1}H–^{1}H Total Correlation Spectroscopy (TOCSY), ^{1}H–^{1}H Correlation Spectroscopy (COSY), and ^{1}H–^{13}C

Heteronuclear Single Quantum Coherence spectroscopy (HSQC), were utilised for metabolite annotation. The data from the 2D NMR experiments was acquired using a Bruker Avance III 600 MHz spectrometer equipped with a CryoTCI triple resonance CryoProbe.

4.1.2 Evaluation of gradient method

The use of gradient elutions — and consequent simplification of the sample mixture — not only reduces overlap of metabolite resonances in NMR spectra generally, but also consequently allows for improved resolution when interpreting 2D NMR spectra. This was noticeable during metabolite annotation, as underlying metabolites often frustrated structural elucidation efforts by demonstrating signals not associated with the target molecule — for example, comparison between the ^{1}H NOESY and 2D COSY spectra of the HLB 2% FA isocratic elution, and those of the HLB 2% FA gradient 80% methanol elution, demonstrated a considerable simplification (fig. 4.1) — this led to a significant increase in peak resolution.
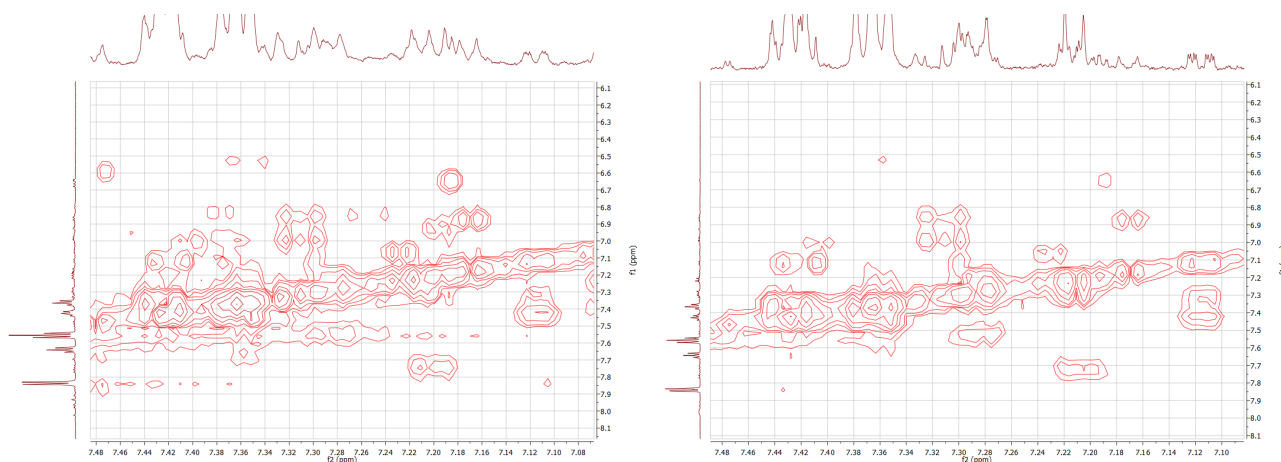
Figure 4.1: The COSY spectrum of the HLB isocratic elution with 2% formic acid elution in D_2O (left), compared with the COSY spectrum of the 80% methanol elution from the gradient method (right). The resolutions of the ^{1}H trace and crosspeaks have been improved, noise has been reduced, and confounding metabolites have been removed.

The formic acid added to the wash is not retained by reversed-phase cartridges, and is hence eliminated in the first (20% methanol) elution. This first elution also predominantly eliminates aliphatics and other non-aromatic metabolites only weakly retained on HLB; aromatic compounds are not present in the spectra until the 60% methanol elution.

Since reversed-phase cartridges already selectively retain hydrophobic compounds, a stepwise linear gradient (as demonstrated here) tends to concentrate most of the retained compounds in the methanol-rich elutions, with relatively few metabolites being eluted earlier on. As a result, the 80% methanol elution appears to contain the bulk of retained metabolites in all regions of the spectra (fig.4.2). The concentration of metabolites in later fractions isn't necessarily detrimental (the presence of more polar metabolites in earlier fractions suggests structural information about those compounds, and subsequently aids structural elucidation), but is worth considering when planning gradient SPE experiments.
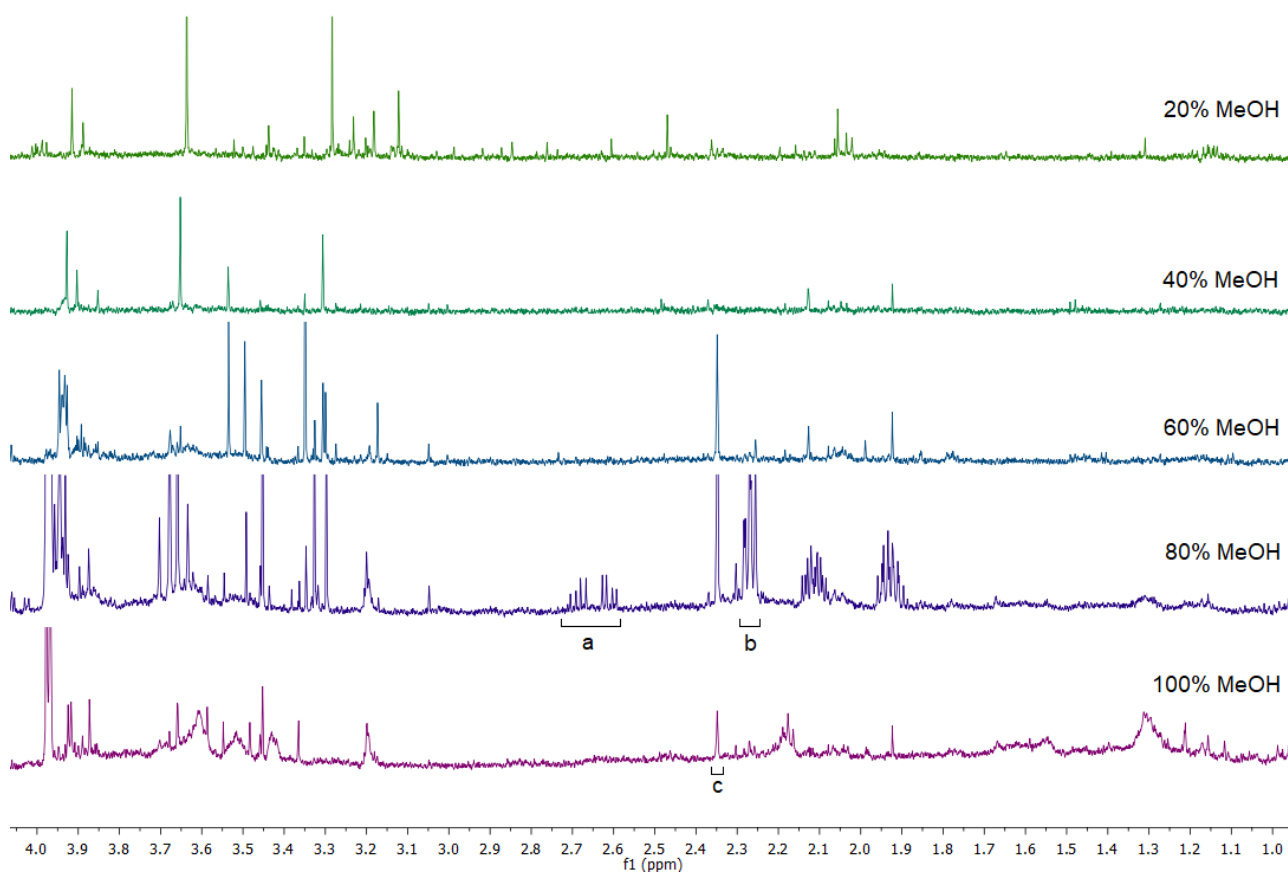
Figure 4.2: Comparison of HLB 80% methanol elution (violet) with other HLB elutions (light green, dark green, light blue, purple). Across the entire spectrum, this elution appears to contain the greatest number of peaks, and likely contains the largest number of metabolites. a = 3-(3-hydroxyphenyl)-3-Hydroxypropanoic acid, b = N-acetylglutamate, c = 4-cresol sulfate

Despite the benefits of this method of fractionation, there are two primary drawbacks of the gradient elution. First, the multiple solvents required to run the gradient experiment are

not supported by the Bruker SamplePro automated SPE system, hence requiring a significant base time and energy investment compared to the automated isocratic methods. Second, and perhaps more importantly, a gradient elution requires further iterations (and, consequently, additional time investment) to achieve useful results.

The isocratic methods described in chapter 3 are designed to be untargeted — the retention profiles of the sorbents are based on the intermolecular forces exhibited by compound classes or specific functional groups, and hence require minimal setup and little knowledge of the constituent compounds in the complex mixture in order to produce useful results. By comparison, the benefits of the gradient elution are dependent on the exact nature of the gradient itself; a shallower gradient at higher methanol concentrations this could result in even greater separation, but the time and effort required to fine-tune the experiment might render some of the rationale for running an untargeted SPE method obsolete.

Having run a simple gradient elution on the HLB cartridge, useful information about specific unknowns noted during the isocratic method development was able to be obtained from the acquired NMR spectra. The results demonstrated here would become vital in identifying some of the retained metabolites.

4.2   Annotation of significant retained metabolites

The results from the SPE urine experiments (Chapter 3) required the use of annotation and identification efforts covering a range of analytical techniques. While a significant number of peaks could be annotated through searching reference databases (such as those provided by the HMDB), several others were not able to be annotated in this manner, despite having attached HSQC and TOCSY data. Below, the analyses detailing the routes towards the identification of several 'unknown' peaks with a metabolite are described; each individual example demonstrates the unique method through which the identity of the metabolite was found. In each successive case, the information required — and, hence, the level of analysis required to definitively confirm the identity of the annotated peaks — increases (fig. 4.3). In this way, a wide range of analytical techniques are demonstrated for several compounds across different compound classes.
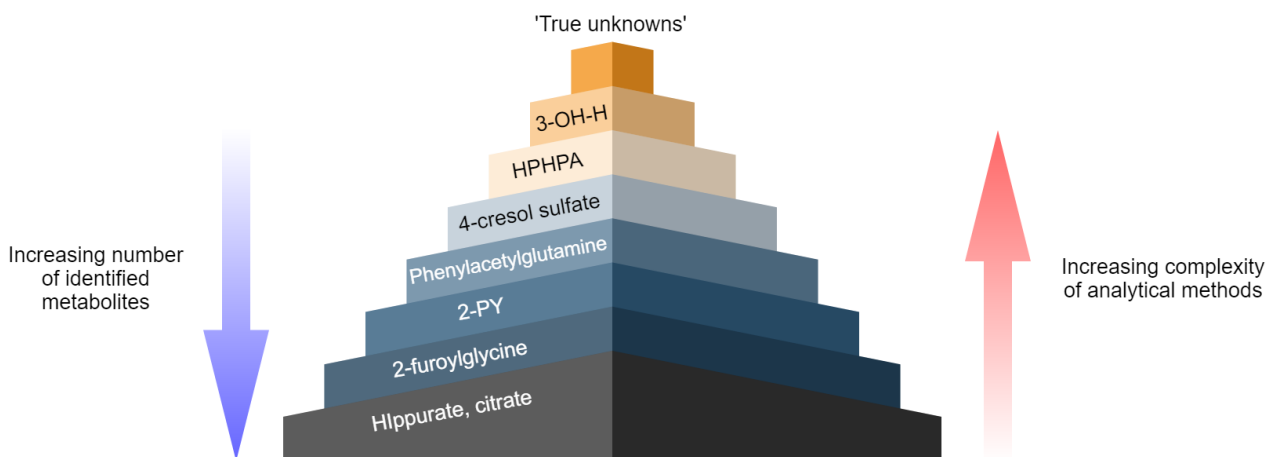
Figure 4.3: Graphical representation of the relationship between the number of identified metabolites and the complexity required to elucidate those metabolites in the peak annotation and identification process. 3-OH-H = 3-hydroxyhippurate; HPHPA = 3-(3-hydroxyphenyl)-3-hydroxypropanoic acid; 2-PY = N-methyl-2-pyridone-5-carboxamide.

4.2.1   2-furoylglycine
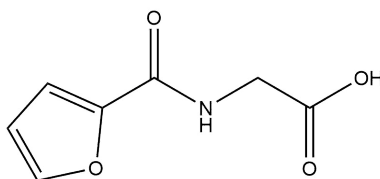
Figure 4.4: Structure of 2-furoylglycine.

Initial analysis of the reversed-phase elutions can be undertaken using basic techniques such as side-by-side comparison of wash and elution. In this instance, the compound 2-furoylglycine (fig 4.4) — an exogenous metabolite[102] — demonstrates a distinctive and usually unobstructed multiplet at 6.64 ppm (dd, J = 1.74 Hz), which can be easily identified through search of databases like the Human Metabolome Database (HMDB). The use of PCA analysis can also be used to confirm the retention pattern of this compound — the PC1 loadings plot for all reversed-phase SPE elutions (chapter 3, fig 3.3) demonstrates that 2-furoylglycine is commonly retained by C_{18} and HLB cartridges under acidic conditions (fig 4.5).
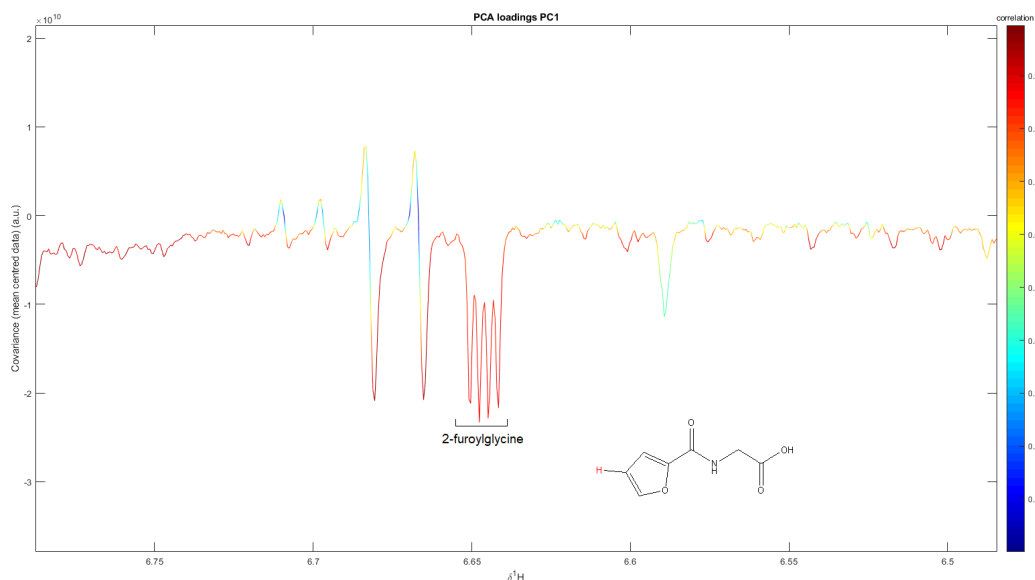
Figure 4.5: PC1 loading of reversed-phase elutions showing a doublet of doublets at 6.64 ppm corresponding to 2-furoylglycine, with Pearson's correlation = 0.87.

4.2.2 *N*-methyl-2-pyridone-5-carboxamide (2-PY) and phenylacetyl-glutamine (PAG)

During inspection of the PC1 loading for natural urine reversed phase elutions, several highly-correlated peaks of interest were not able to be annotated using a spectral library. This included the multiplet 7.97 (dd, J ≈ 9, 3 Hz, fig 4.6); the strongly downfield aromatic signal suggested a heterocyclic compound with ortho- and meta-couplings.
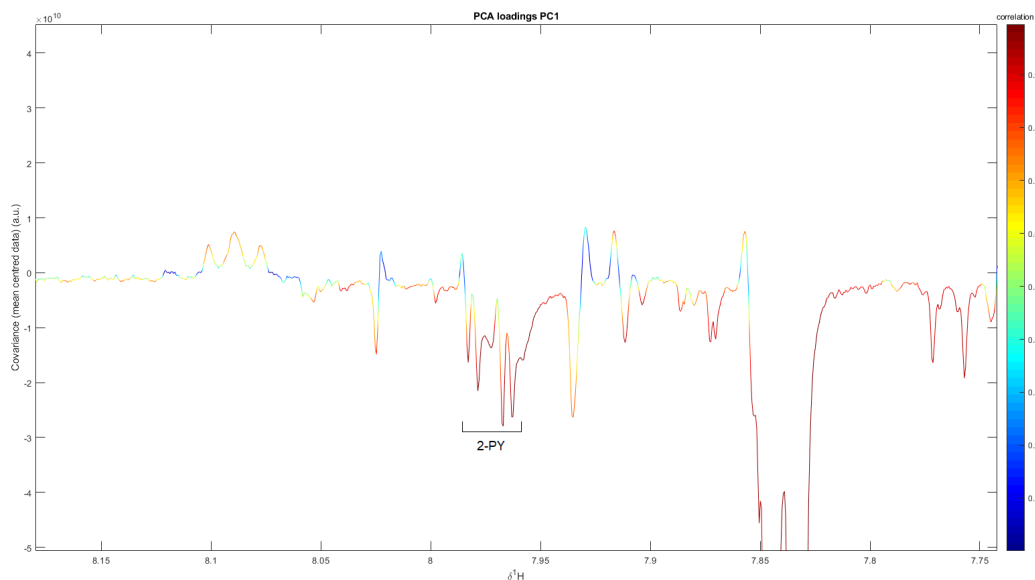
Figure 4.6: PC1 loading showing a doublet of doublets at 7.97 ppm corresponding to
2-PY, with correlation = 0.98. In this instance the signals are overlapped with other
resonances, but are able to be confirmed as one multiplet through comparison of
correlations and referral to raw spectra.

Database searches for a multiplet at 7.97 with a ^{13}C signal of 142.3 ppm returned no results. However, manual referencing with online search engines and compiled spreadsheet databases containing reported metabolites afforded previous literature describing the NMR assignments of the compound N-methyl-2-pyridone-5-carboxamide (2-PY), a metabolite of niacin found in the urine of coffee drinkers[103]. This identity was putatively confirmed through annotation of other relevant peaks in the PC1 loadings, and definitively confirmed through an NMR spike-in experiment using an authentic reference standard.
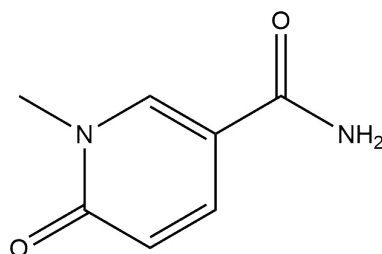
Figure 4.7: Structure of N-methyl-2-pyridone-5-carboxamide.

Several other peaks found in the PC1 loading for natural urine reversed phase elutions were unable to be identified, including a multiplet at 4.19 ppm (m, fig 4.8):
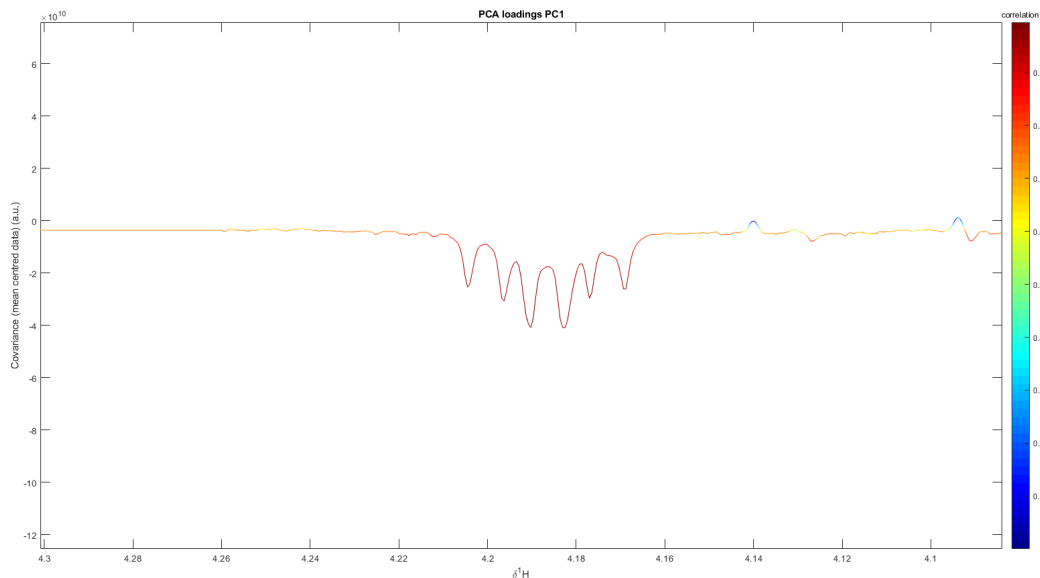
Figure 4.8: PC1 loading showing a multiplet at 4.19 ppm corresponding to PAG,
with correlation = 0.98.

As with the identification efforts for 2-PY, no database matches were able to be made, despite additional HSQC data. However, TOCSY correlations with peaks at 2.27 (t, J ≈ 8.4 Hz), 2.12 (m), and 1.93 (m) ppm — which were also present in the PC1 loading — enabled the discovery of literature describing phenylacetylglutamine (fig 4.9), a liver conjugate of phenylacetate and potential biomarker of chronic kidney disease[104]. The identity of the unknown resonances as PAG was confirmed through NMR spike-in experiment.
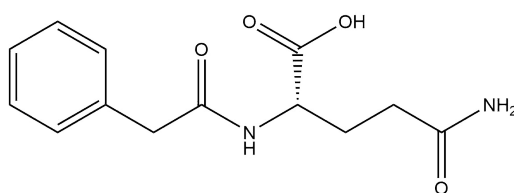
Figure 4.9: Structure of phenylacetylglutamine.

While 2-PY and PAG NMR assignments had been reported in the literature, and both are recorded on databases such as HMDB, NMR data is not present for those entries. It is hence worth remembering that greater numbers of metabolites recorded as present in database statistics do not necessarily translate as greater understanding of the human urinary metabolome; similarly, greater understanding of a metabolome is not necessarily useful if that understanding is not readily accessible to researchers and analysts.

Regardless, the use of SPE here has enabled the simple analysis of spectra through fraction-
ation — whereas in untreated 'raw' urine the driving multiplet was obscured by other signals,
they appear without significant overlap in the 80% methanol HLB gradient elution (fig 4.10).
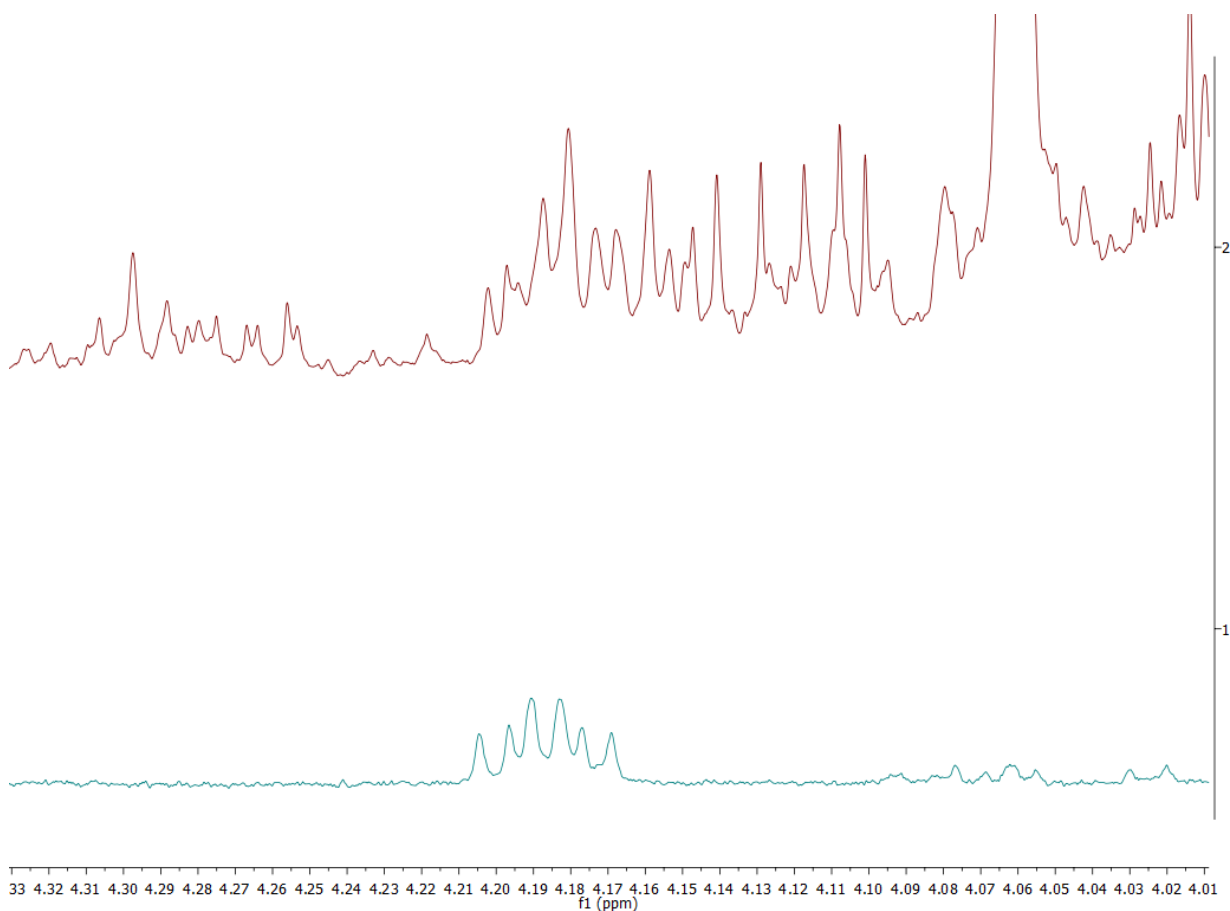
Figure 4.10: Stacked spectra demonstrating revealed multiplet at 4.19 ppm in HLB
gradient elution 5 (blue), previously obscured in untreated 'raw' urine (red).

4.2.3 4-cresol sulfate and 3-(3-Hydroxyphenyl)-3-Hydroxypropanoic acid (HPHPA)
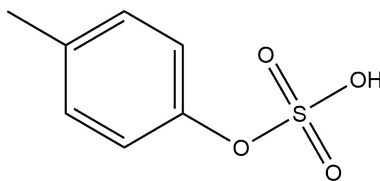
Figure 4.11: Structure of 4-cresol sulfate.

During analysis of gradient elution HLB-treated spectra (chapter 3), several resonances in later elutions were revealed where they had previously been overlapped. A singlet at 2.35 ppm (^{13}C 22.73 ppm), typical of a methyl group substituted on a phenyl ring, was found in the fourth, fifth, and sixth elution with TOCSY correlations to peaks at at 8.22, 7.30 and 7.21 ppm. These aromatic shifts are indicative of a para-substituted benzene group, particularly para-cresol (1-hydroxy-4-methylbenzene) — the resonance at 8.22 ppm is not present in the ^{1}H NMR, characteristic of a hydroxyl group. However, a database search for 4-cresol demonstrated NMR resonances which, while indeed similar, were shifted noticeably upfield compared to the target resonances. The potentially deshielded nature of metabolite resonances — along with the fact that 4-cresol is typically conjugated in human urine — implied that the conjugation product 4-cresol sulfate best fit the identity of the target compound[105]. An NMR spike-in experiment confirmed the identity of the unknown as 4-cresol sulfate. In this instance database entries for 4-cresol sulfate existed and ^{1}H NMR spectra were attached, but the spectra were significantly different to those measured experimentally.
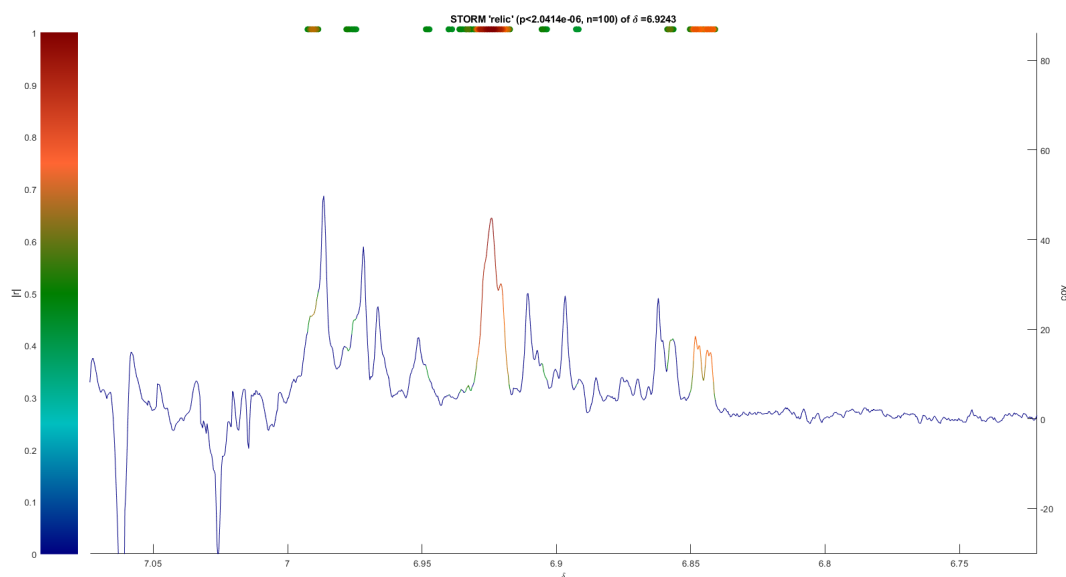
Figure 4.12: STORM graph demonstrating relics correlating with the driver region 6.92 - 6.93 ppm, taken from the 100 most similar spectra to the average of the total dataset.

Similarly, another resonance with a distinctive signal at 6.93ppm (t, J = 2.15 Hz) returned no results through database searching. This signal demonstrated TOCSY couplings with signals

at 7.31, 6.98, and 6.85 ppm. Utilising Statistical TOtal Correlational SpectroscopY (STOCSY) in conjunction with the Airwave spectral library[106], statistical correlations were determined with signals at 6.99, 6.85, 5.02, and 2.63 ppm, as well as with signals corresponding to hippurate and trigonelline. The use of SubseT Optimisation by Reference Matching (STORM) with the 100 most similar spectra to the average of the total Airwave dataset similarly suggested 'relics' at 6.99, 6.85, and 2.63 ppm, among others (fig 4.12).

Requiring more information, spectra derived from the 'Urinemap' project — fractions of urine samples separated first using reversed-phase columns, then by HydrophILc Interaction Chromatography (HILIC) — were studied for greater insight into the unknown. TOCSY, COSY, HSQC, HMBC, and selective TOCSY experiments were run on a Urinemap fraction with the unknown present in order to gain additional information: it became clear that the aromatic signal multiplicities, as well as the broadening of the 6.92 and 6.98 signals (resonances with this chemical shift being indicative of a hydroxyl group), implied an asymmetrical meta-substituted phenolic compound (fig 4.13).
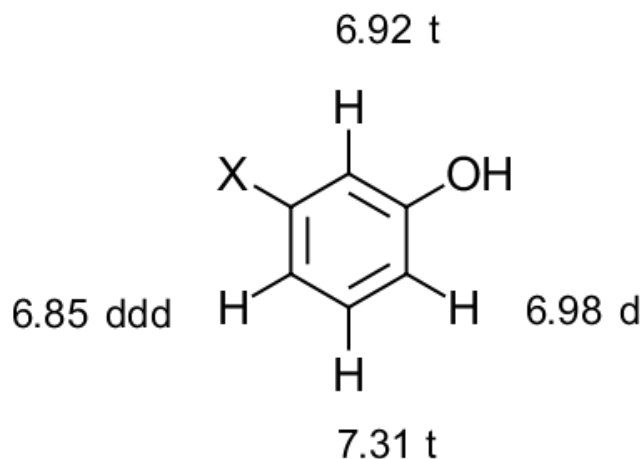
Figure 4.13: Initial tentative assignments for unknown resonances discovered through 2D NMR spectroscopy, with X representing a then-unknown sidechain.

The doublet of doublet of doublets at 6.85 had J-couplings of 8.1, 2.6, and 1.0 Hz; a J-value of 2.6 Hz (the couplings at 8.1 and 1.0 corresponding to ortho-H3 and meta-H6, respectively) suggested a long-range four-bond alyllic, acetylenic coupling, or an ortho and two non-identical meta couplings. Returning to the STOCSY correlations, peaks at 5.02 ppm (1H, dd, J ≈ 8.2, 5.9) — likely a CH proton with both phenyl and either OH or OMe connections — and

2.63 ppm (2H, AB) were determined to be part of the side chain. A resonance at 2.63 ppm is indicative of an adjacent -COOH or -CONH$_2$ group, and the AB two-spin system was likely to be a CH$_2$, with both protons coupling individually to the vicinial CH. A structural search of the potential structure with an OH connection and an adjacent -COOH compound revealed 3-(3-HydroxyPhenyl)-3-HydroxyPropanoic Acid (HPHPA), a previously reported metabolite[107] (fig 4.14); the identity of the unknown as HPHPA was confirmed by NMR spike-in experiment and MS/MS analysis.
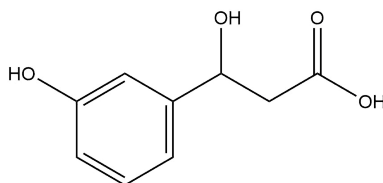
Figure 4.14: Structure of 3-(3-Hydroxyphenyl)-3-Hydroxypropanoic acid.

In this instance, both ^{1}H and ^{13}C data had been recorded on HMDB, but not only was this data not searchable, one resonance (at 6.85 ppm) did not match experimental results — with the database recording a three-peaked multiplet, instead of a doublet of doublet of doublets. Similarly, the ^{13}C signals bore little resemblance to the experimental results, with one peak registering a difference of over 7 ppm between the database value (130.5 ppm) and the actual recorded value (137.9 ppm); while this could be caused by pH effects or similar, these effects are not considered by the search algorithm for the database, hindering identification efforts.

4.2.4 3-hydroxyhippurate

One of the resonances noted in the reversed phase elutions PCA (chapter 3) is a distinctive doublet of doublet of doublets (J $\approx$ 8.0, 2.5, 1.0 Hz) — in untreated urine, this multiplet is typically overlapped by histidine (fig 4.15). Attempts to use STOCSY and STORM to elucidate more structural information produced no useful results.

The unknown was also present in the fifth elution (80% methanol/20% water) of the HLB gradient experiment. A battery of COSY, TOCSY, HSQC, and HMBC spectra were acquired on this elution, allowing for the associated ^{13}C signal for the driver peak at 7.12 ppm to be annotated at 122.0 ppm.
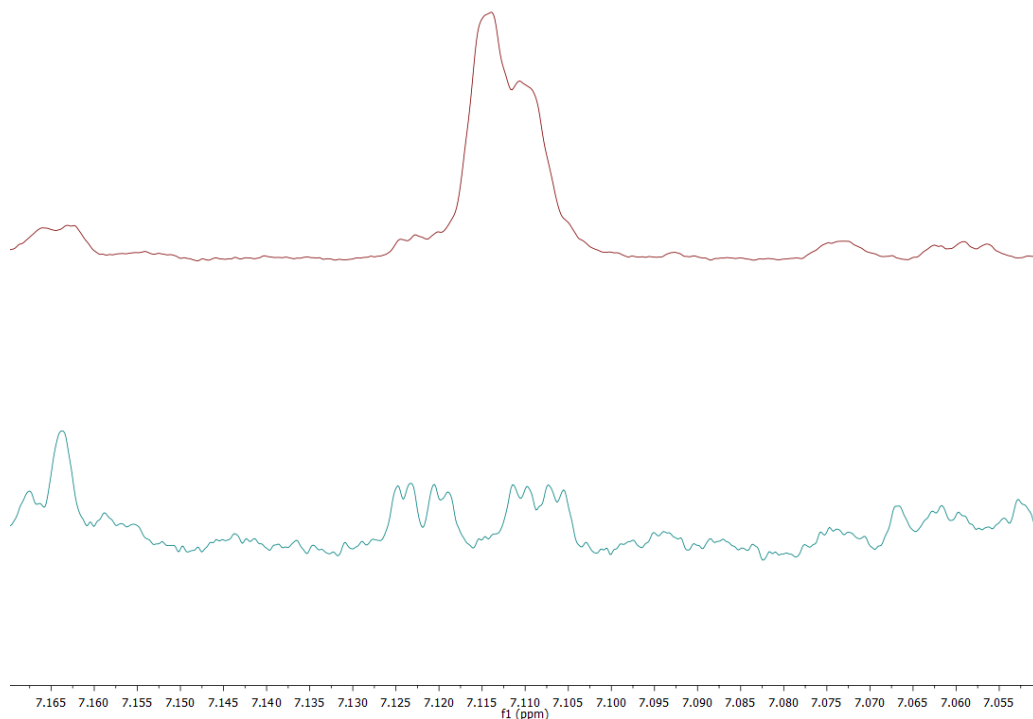
Figure 4.15: NMR spectra demonstrating overlap of the ddd resonance at 7.12 ppm.
After treatment with the HLB 2% formic acid method, the multiplet is revealed
(bottom); in untreated urine, it is obscured (top).
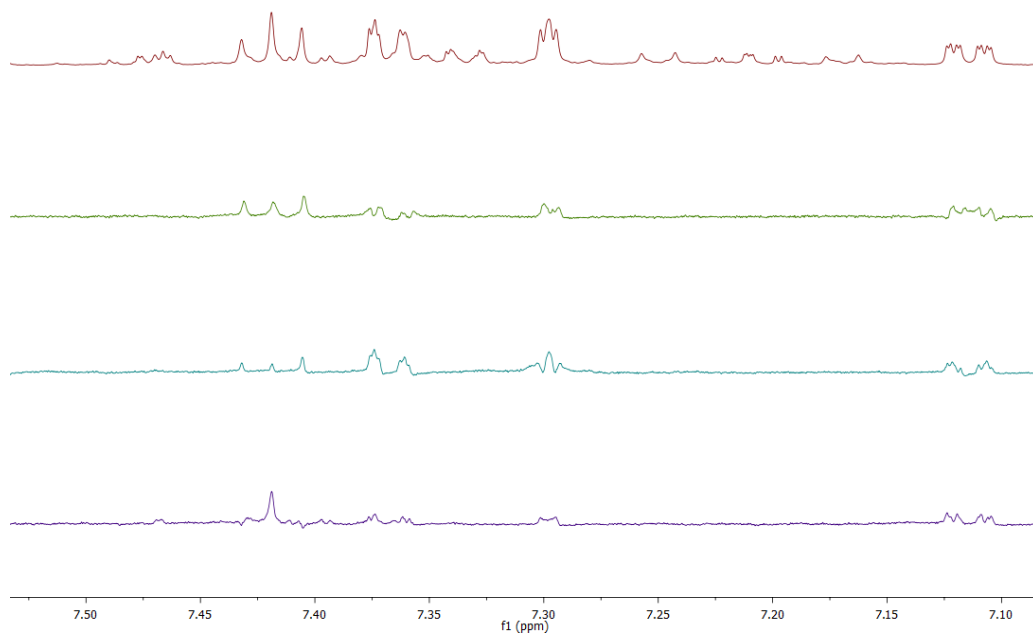
Figure 4.16: Stacked NMR spectra demonstrating the use of selective 1D TOCSY
experiments to identify other resonances within the spin system of the molecule.
From top: standard 1D NOESY, driver peak 7.12 ppm (ddd), driver peak 7.30 ppm
(t), driver peak 7.42 ppm (t).

Due to overlap even within this gradient elution, reliable ^{1}H correlations could not be established using standard COSY/TOCSY alone; the use of selective 1D TOCSY experiments allowed for the driver peak at 7.12 to be conclusively correlated with 7.42 ppm (t, J ≈ 7.9), 7.37 ppm (dt, J ≈ 7.7, 1.4), and 7.30 ppm (t, J ≈ 2.2). These resonances are demonstrated in fig. 4.16.

Without mass spectral data, further attempts at structural elucidation would be futile — however, the SPE elution was still complex enough that simple LC-MS approaches would likely result in ion suppression, and the consequent loss of sample signal. A fractionation system was hence employed in order to further isolate the target compound.

Fractionation methodology

Using a HLB gradient elution method with 2% formic acid added at all steps (as described in chapter 3), four replicates of the fifth elution — comprised of 80% methanol, 20% water — were acquired using 3 mL pooled urine. These replicates were dried under nitrogen; each was reconstituted in 1 mL ultrapure water, and these samples pooled together to achieve 4 mL of the pooled elution. This combined elution was then divided into four sample vials of 1mL each; the contents of three of the vials were injected through a reversed-phase 4.6 mm x 150 mm Atlantis T3 column, utilising a method previously described by Whiley et al.[108]. Fractions were continuously collected from the column at a rate of 1mL/min for a total of 90 fractions; this method was repeated for each of the three samples.

The acquired fractions were transferred to 96-well plates and dried under nitrogen. One of these plates, containing 15 µL of each fraction, was utilised for LC-MS experiments; the contents of the remaining three plates, previously containing 5850 µL per well in total, were redissolved in 750 µL deionised water per well. The fractions were then prepared for ^{1}H NMR acquisition with 1D NOESY and J-resolved spectra acquired for each fraction, as previously described.

LC-MS experiments were run using a 2.1×150 mm HSS T3 column at 45°C. A gradient elution was used at a flow rate of 0.6 mL/min starting from 99% water with 1% acetonitrile (both with the addition of 0.1% formic acid), continuing with a linear gradient until 45% water with 55% acetonitrile at 9.9 minutes, finished with a final ramp to 0% water within 0.7 minutes

at a rate of 1 mL/min[109].

Fractionation results

The original fractionation method described by Whiley *et al.* utilised untreated urine samples — in this instance, by comparison, the method utilised urine which had already been subject to a prior SPE fractionation. Despite the use of reversed-phase chromatography not being strictly orthogonal to the gradient SPE approach utilised previously, the LC fractionation lead to significant purification of the target compound.

1D NOESY experiments demonstrated that fraction 36 contained a significant concentration of the unknown, as well as a marginally smaller amount of a second compound. Integration of the unknown peaks allowed for reliable integration of the aromatic multiplets, all of which had a relative area of one proton. From these results, it was deduced that the compound likely contained a disubstituted benzene ring, with the substituents being ortho- or meta- substituted.

Comparisons to literature suggested that the second compound was either paraxathine or theophylline, both metabolites of caffeine with very similar structures and NMR spectra — as paraxanthine is involved in the main pathway for caffeine metabolism in humans, it is more likely to be present in greater quantities than theophylline[110], and hence is more likely to fit the identity of the impurity. Separately, HPHPA was identified by NMR in fraction 38; the negative mode mass spectrum demonstrated a peak eluting at 2.92 mins with a mass corresponding to $[C_9H_{10}O_4 - H]^-$, confirming the presence of HPHPA.

LC-MS was run on fractions 34-39 using a Waters Acquity UPLC system coupled to a Waters Xevo G2 QToF mass spectrometer. For fraction 36 in positive mode, only one significant peak appears in the chromatogram at the relevant time point, at 2.77 mins: this peak registers a mass of 181.0743, equivalent to $[C_7H_8N_4O_2 + H]^+$, confirming the presence of paraxanthine. In negative mode, two peaks appear — a larger peak at 2.77 mins, followed by a smaller peak at 2.80 mins with a mass corresponding to paraxanthine. The mass reading for the larger peak at 2.77 mins corresponded to a compound at 194.0450 Da (fig 4.17).
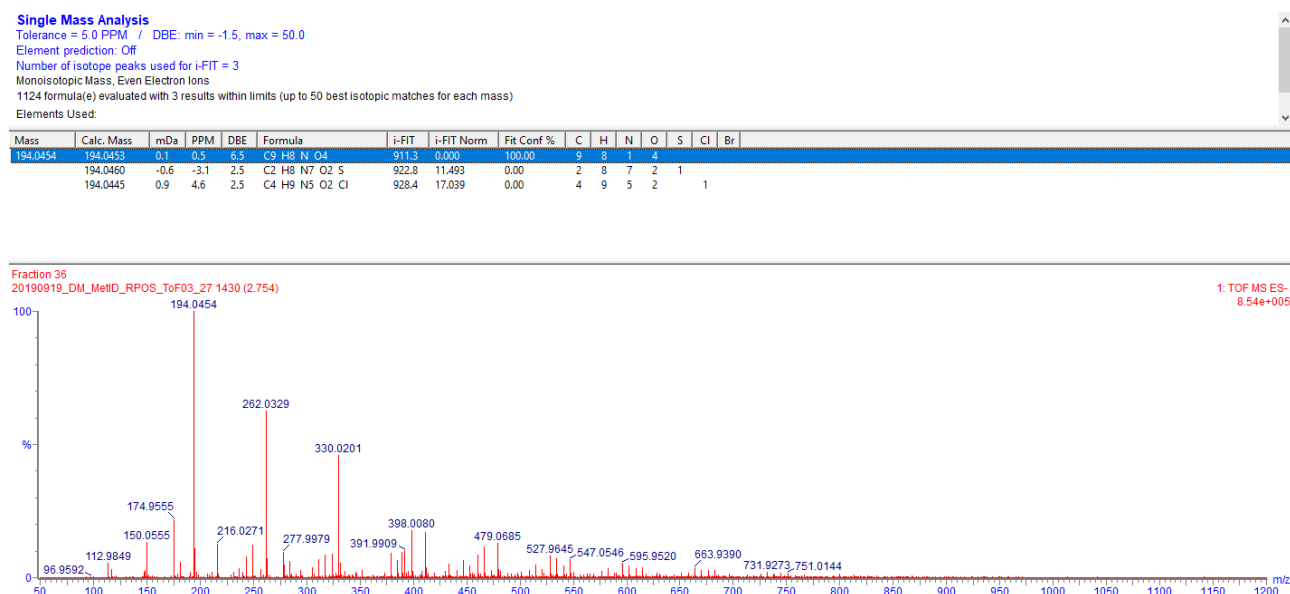
Figure 4.17: Elemental composition analysis of the peak eluting at 2.77 mins revealed the likely chemical formula of the unknown compound.

Elemental composition analysis suggested that this molecular weight was equivalent to $[C_9H_9NO_4 - H]^-$ with a 100% fit confidence; an HMDB formula search revealed 11 potential compounds, of which two — the structural isomers 2-hydroxyhippurate and 3-hydroxyhippurate — were suitable candidates. NMR spectral data for 2-hydroxyhippurate present on the database revealed several inconsistencies in chemical shift compared to the acquired fraction spectra. By contrast, no data (by NMR or LC-MS) for 3-hydroxyhippurate was present on the database; a review of the literature described several cases where 3-hydroxyhippurate had been analysed by NMR, but where no assignments had been provided. After NMR spike-in with an authentic reference standard, the identity of the unknown as 3-hydroxyhippurate was confirmed (table 4.1); this was further confirmed by MS/MS analysis (fig 4.18).

Table 4.1: NMR spectral assignments of 3-hydroxyhippurate

ID	^{1}H ppm	^{13}C ppm	H	Type	Multiplicity, J (Hz)	COSY	TOCSY	HMBC
1	-	158.8	-	C	-	-	-	-
2	7.30	117.0	1	CH	t (2.2)	7.11	7.12, 7.44	122.0, 158.8
3	-	138.0	-	C	-	-	-	-
4	7.37	122.0	1	CH	dt (7.7, 1.4)	7.11	7.14, 7.45	116.8, 122.0

5	7.42	133.0	1	CH	t (7.9)	7.11	7.14, 7.30	116.8, 122.0, 138.0, 158.8
6	7.12	122.0	1	CH	ddd (8.0, 2.5, 1.0)	7.28, 7.42	7.30, 7.41	116.8, 122.0, 158.8
9	3.95	46.9	2	CH$_2$	d (5.62)	8.48	8.50	-
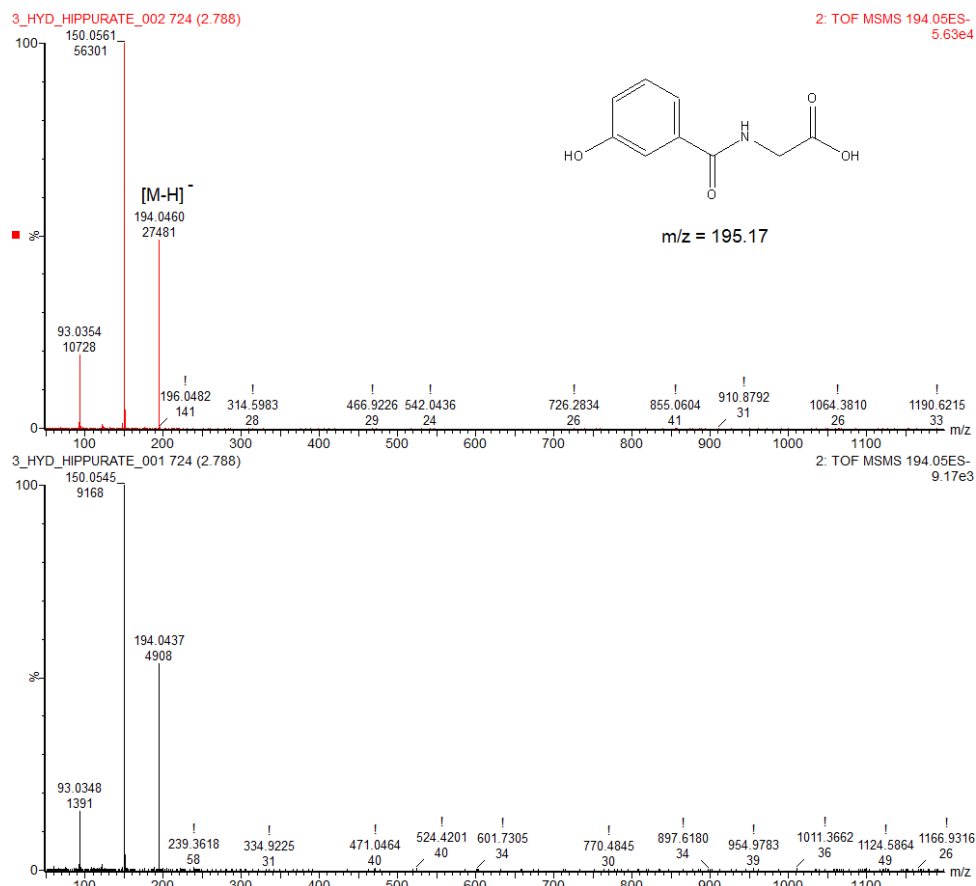
Figure 4.18: MS/MS spectra demonstrating identical fragmentation pattern for
reference standard (top) and fraction sample (bottom).

In this example, multiple analytical approaches were utilised in order to elucidate the structure
of a compound with assignments and other NMR spectral data previously unreported in the
literature. The metabolite has been linked to the consumption of blackcurrant juice and 'colonic
degradation of polyphenols'[111]; while it is normally present in urine, its expression in NMR
spectra is often suppressed by overlapping histidine signals. Utilising SPE-NMR techniques which
retain the compound — but not histidine — the compound can be qualitatively, and possibly
quantitatively, identified. The same result can be achieved through the use of complementary
methods which selectively retain histidine, but not 3-hydroxyhippurate, such as pH-controlled

SCX approaches (fig 4.19).
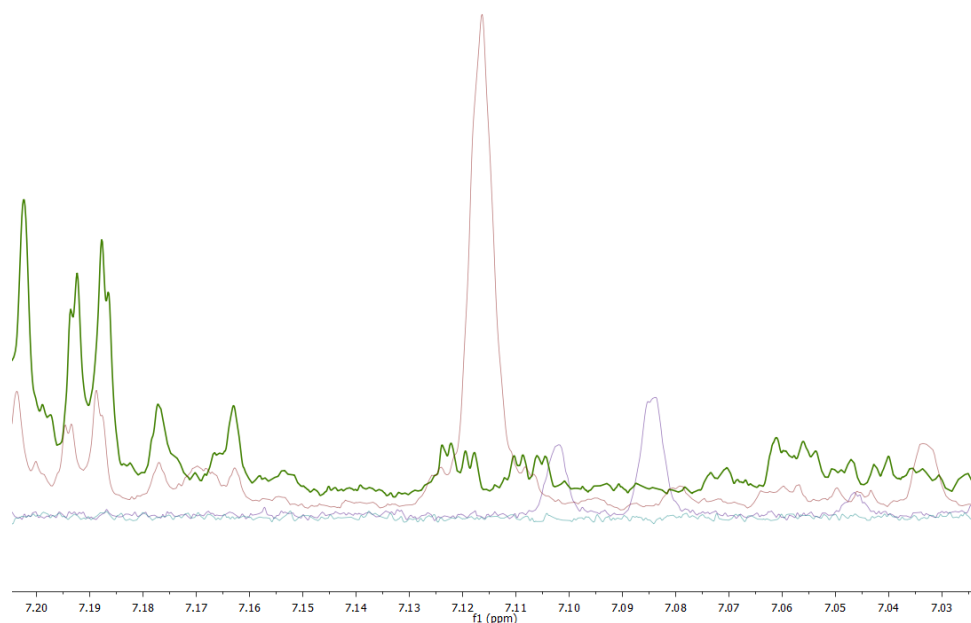
Figure 4.19: Overlayed NMR spectra revealing 3-hydroxyhippurate in the wash
(green), obscured in the raw sample (red), and absent from the elutions (blue,
purple) of an SCX SPE method.

4.3   Conclusion

By using NMR analysis coupled with PCA and STOCSY, significant insight into the retention
profiles of the methods taken forward from chapter 2 had been achieved; however, many peaks
remained 'unknown' despite having both 1D and 2D NMR data associated. The work in this
chapter demonstrates a case study in metabolite identification, and how increasingly complex
analysis utilising multiple approaches must be used in order to accurately profile metabolites
within biofluids. The efforts to identify these unknown peaks also demonstrate the more
'targeted' aspect of metabolic profiling, as spectral information noted during the course of
'untargeted' experiments in chapter 3 are subject to further scrutiny. In total, the 'tiered'
system of metabolite identification described in this chapter provides a workflow for future
identification efforts promoted by the use of the untargeted SPE-NMR methods, which had now
been characterised enough to be translated onto an automated SPE system.

Chapter 5

Automation and application of SPE-NMR protocols

The methods and submethods described within are based, in whole or in part, on routines distributed by the Bruker Corporation as part of the normal functioning of the SamplePro SPE system. These methods have been heavily edited by the author in order to enable the robot to support the running of the untargeted SPE methods developed in previous chapters. The results of this chapter have been transcribed into a standard protocol, which is described in appendix B.

5.1   Introduction

The use of untargeted solid phase extraction methods on complex natural mixtures broadens the capacity for metabolite annotation and identification, and for structural elucidation — this was demonstrated in chapters 3 and 4, where SPE-treated urine was analysed in order to reveal previously obscured metabolites and aid structural elucidation efforts. Being able to automate these SPE methods would entail a significant reduction in labour and time, as the protocols could be run from start to finish without further input from the analyst.

5.1.1 Description of machine features and changes to default methods

The Bruker SamplePro SPE system utilises robotics in order to perform SPE experiments; new routines can be designed, and existing routines modified, through the use of a visual scripting tool. Originally, the system aimed to run SPE methods and consequently transfer the resulting elutions to SampleJet racks for immediate NMR analysis — this required strict control of solvents, with the use of a deuterated methanol/buffer solution usually taking the place of the final elution. While this approach is considered very high-throughput, untargeted methods often require analysis of both wash and elution(s) — in addition, the solvent restrictions limited the methods able to be utilised on the machine.
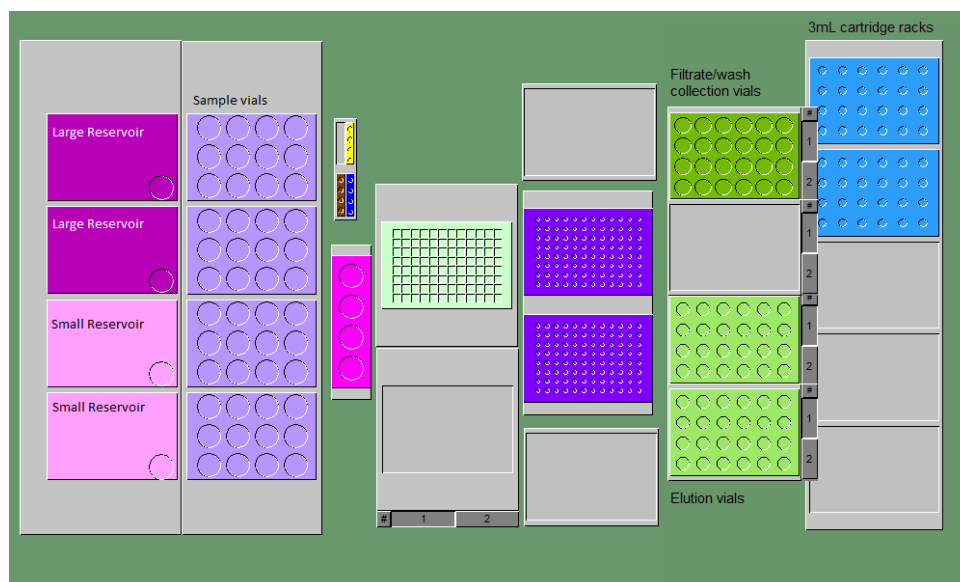
Figure 5.1: Visualisation of SPE robot layout, from layout editor software. This layout is used for experiments utilising 3 mL cartridges.

The primary moving parts of the robot are the rack gripper, the pipette arm (with four individual pipettes mounted), and the solvent system, with four standard inlet tubes which feed into any solvent bottle and which are capable of dispensing solvent through any of the four mounted pipettes. On the station, there are additionally two large solvent reservoirs, two small solvent reservoirs, a pipette washing station, and the SPE experiment station (with waste inlet). SPE cartridges used in the robot are kept in 4x6 cartridge racks, with accompanying 4x6 collection vial racks able to collect both the wash and the elutions; this entails one wash vial

rack with cover, one first elution vial rack with cover, and one second elution vial rack with cover. During normal operation, the gripper arm is used to move racks to and from the SPE experiment station — where the solvent runoff from the cartridges is not collected (for example, during the conditioning and equilibration steps), it runs directly to waste. By comparison, in the wash and elution steps, the vials are placed underneath the cartridges such as to collect the runoff. This layout is demonstrated in the visualisation for 3 mL cartridges (fig 5.1) and the photos of the machine (fig 5.2); for methods utilising 6 mL cartridges, the racks are the only variable.
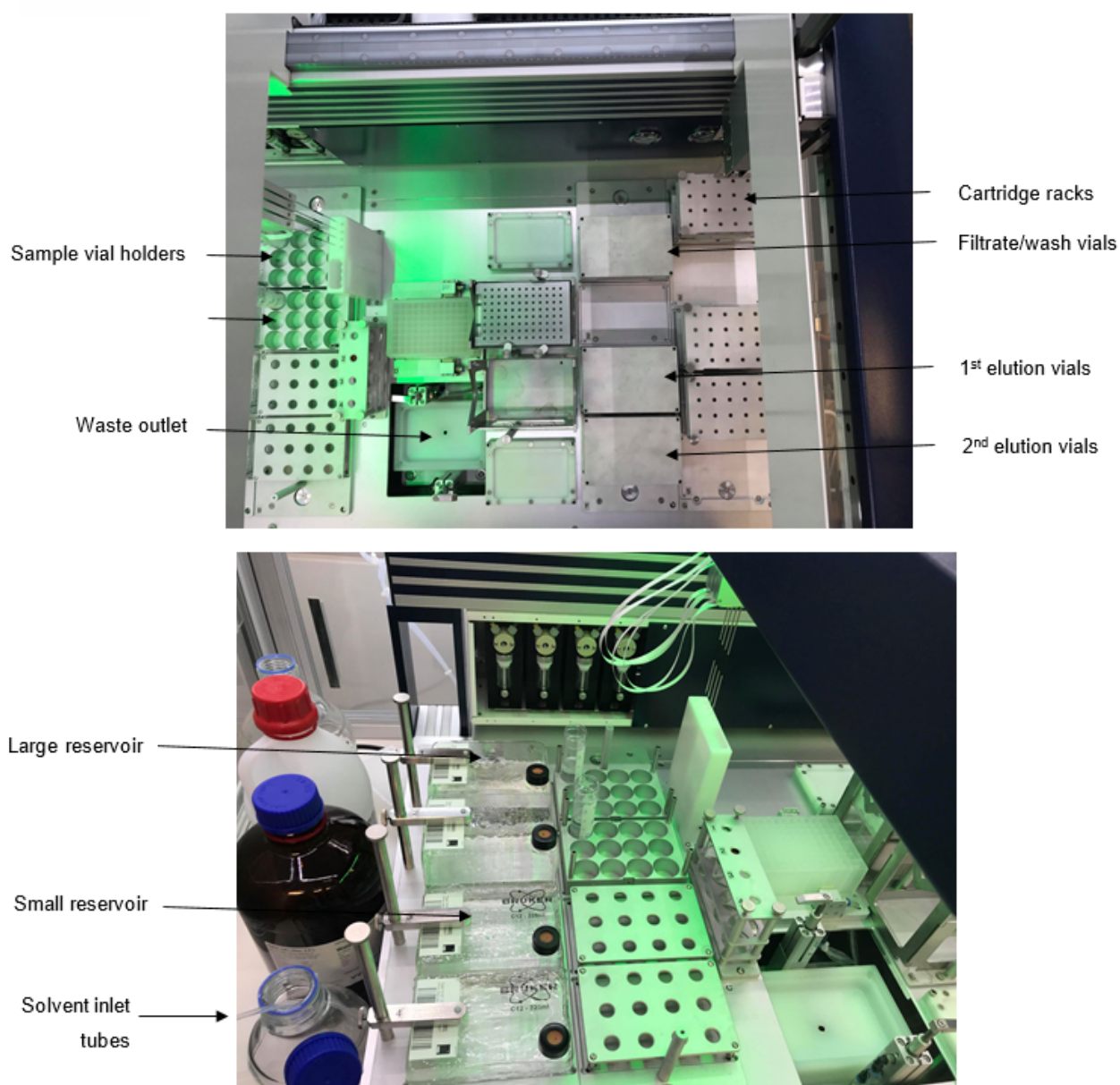
Figure 5.2: SPE robot layout.

The SamplePro system was originally designed such that the elutions from isocratic SPE experiments were eluted using specific solvents — usually containing deuterated methanol and buffer — which could then be automatically transferred into SampleJet racks. While this promises a very high-throughput workflow, the necessity of the deuterated solvent restricts flexibility in experimental design more broadly, especially for ion exchange cartridges. As a result, it became more convenient in the long term to disable the automatic SampleJet preparation in order to allow for greater control over the solvent systems in use — the resulting washes and elutions from the use of the robot can then be dried under nitrogen and reconstituted in an appropriate solvent (such as water or D_2O) for analysis.

The default 'master' method supplied with the machine contained support only for 3 mL capacity SPE cartridges, where those cartridges were filled with reversed-phase materials such as Merck Milipore LiChrolut EN, a polymer-based sorbent. This limited the ability to run the untargeted SPE methods developed, which utilise 6 mL (500 mg bed weight) cartridges with both reversed-phase and ion exchange sorbents. Learning to use the visual scripting tool and adapting the methods and subroutines was a necessary step towards the machine supporting 6 mL SPE cartridges and ion exchange methods.

Several problems were noted during the development of the automated methods due to machine limitations, which predominantly became apparent during the construction of subroutines for 6 mL SPE cartridges. One such problem involved the 'push-through' subroutine, where liquid loaded onto the cartridge in a given step is supposed to experience enough back-pressure from the needle to allow the solvent (or sample) to flow through; without such a step, it could take several hours for the liquid to drain through. With the default 3mL method provided the back-pressure is accounted for such that the liquid drains rapidly, but the fluid dynamics of 6mL cartridges require additional steps to be taken to clear the sorbent. In this instance, pressurised nitrogen was utilised in order to provide the necessary back-pressure for the liquid in 6 mL methods to drain in a timely manner.

5.2 SOP construction

'Master' methods can be initiated by the user in order to begin the method called; by comparison, subroutines can only be started when called by a master method. Both methods and the subroutines called within them can themselves call individual events or actions in order to produce a desired effect. These events may include moving the injection apparatus to a specific X, Y, and Z coordinate in space; the machine can also aspirate or dispense liquid from solvent bottles, reservoirs, or cartridges, move cartridge or vial racks around the workspace, and provide a nitrogen flow through the dispensing needle.

5.2.1 Master methods

Four startable 'master' methods were created, based on a default method originally provided with the machine. These master methods are 'TargetedMethod', 'TargetedMethod_3mL', 'UntargetedMethod', and 'UntargetedMethod_3mL'. One additional method, 'Reset' (alg. 1), was also developed for troubleshooting purposes — in the event of a failure, the Reset method will shut off nitrogen pressure and liquid dispensing (if applicable) and unlock the fixations holding cartridge racks in place.

Algorithm 1 Startable Method: Reset

1: **Write To Memory Device** Device SPEPressureValueSet13, Write: 0
2: **Write To Memory Device** Device SPEPressureValueSet24, Write: 0
3: **Switch device** Device LPipArmValveTip1, Switch to OFF
4: **Switch device** Device LPipArmValveTip2, Switch to OFF
5: **Switch device** Device LPipArmValveTip3, Switch to OFF
6: **Switch device** Device LPipArmValveTip4, Switch to OFF
7: **Switch device** Device PIPStationFixation, Switch to OFF
8: **Switch device** Device SPEStationFixation, Switch to OFF

The four master methods are functionally similar, being based on the default 'MasterMethod' shipped with the machine. Notable differences are the default parameters — which are specified by users during setup — and the push-through subroutine utilised; for 3 mL cartridges this is handled by subroutine 'zzPushThru_3mL', whereas for 6 mL cartridges this is handled by subroutine 'zzPushThru'. Each subroutine is associated with a specific SPE step or automatable action — this is demonstrated in alg. 2, as each subroutine is called in turn.

Algorithm 2 Startable Method: UntargetedMethod

 1: *Define dependent variables*
 2: **Store** _$0NofSamples := _$NofSamples
 3: **Store** _$0CondVolume := _$CondVolume
 4: **Store** _$0EquiVolume := _$EquiVolume
 5: **Store** _$0LoadSampleVolume := _$LoadSampleVolume
 6: **Store** _$0CollectVolume := _$CollectVolume
 7: **Store** _$0WashVolume := _$WashVolume
 8: **Store** _$0DryDelay := _$DryDelay
 9: **Store** _$0Elut1Volume := _$Elut1Volume
10: **Store** _$0Elut2Volume := _$Elut2Volume
11: **Flush** Diluent: System, 1 flush cycle
12:
13: *Steps in batch mode*
14: *Conditioning step*
15: **Show info** Text: Conditioning
16: **Call subroutine** z1Conditioning, _$CondVolume=_$CondVolume, _$NofSamples=_$NofSamples
17:
18: *Equilibration step*
19: **Show info**Text: Equilibration
20: **Call subroutine** z2Equilibration, _$EquiVolume=_$EquiVolume, _$NofSamples=_$NofSamples
21:
22: *Load step*
23: **Show info** Text: Load samples
24: **Call subroutine** z3LoadSample, _$LoadSampleVolume=_$LoadSampleVolume,
    _$NofSamples=_$NofSamples
25: **Call subroutine** zzParkAllRacks
26:
27: *Wash step*
28: **for** _$0CurrentSample := 1 to _$NofSamples (Repeat _$NofSamples times) **do**
29:     **Show info** Text: Washing
30:     **Call subroutine** z4WashAndCollectCartridges, _$CurrentSample=_$0CurrentSample,
    _$WashVolume=_$WashVolume
31: **end for**
32: **Call subroutine** zzParkAllRacks
33:
34: *Elution1 step*
35: **for** _$0CurrentSample := 1 to _$NofSamples (Repeat _$NofSamples times) **do**
36:     **Show info** Text: Elution 1
37:     **Call subroutine** z6Elution1, _$CurrentSample=_$0CurrentSample, _$DryDelay=_$DryDelay,
    _$Elut1Volume=_$Elut1Volume
38: **end for**
39: **Call subroutine** zzParkAllRacks
40:
41: *Elution2 step (only if Elution2 volume > 0)*
42: **for** _$0CurrentSample := 1 to _$NofSamples (Repeat _$NofSamples times) **do**
43:     **Show info** Text: Elution 2
44:     **Call subroutine** z7Elution2, _$CurrentSample=_$0CurrentSample, _$Elut2Volume=_$Elut2Volume
45: **end for**
46: **Call subroutine** zzParkAllRacks
47:
48: **Flush** Diluent: System, 1 flush cycle

5.2.2   Major subroutines

Each subroutine executes a relevant step of the SPE process. Separate subroutines exist for both 6 mL and 3 mL methods — in each instance, the primary difference between the two is the 'PushThru' subroutine, which is dependent on the cartridge size being used as previously detailed.

Conditioning

The conditioning subroutine draws solvent from solvent tubes 3 and 4. The robot picks up the cartridge rack and places it on top of the waste inlet. Solvent (up to 3000 µL) is aspirated from the solvent bottles; this is then dispensed onto the cartridges through tips 3 and 4, and the liquid pushed through using nitrogen. This is repeated until all conditioning solvent has been dispensed onto the cartridges and pushed through (alg 3).

Algorithm 3 Subroutine: z1Conditioning

```
 1: Set conditioning values
 2: Store _$0CondVolume := _$CondVolume
 3: Store _$0SourceRack := System 3 4
 4:
 5: Used Tips
 6: Store _$0TipBit := 12
 7:
 8: Set pushthrough values
 9: Store _$0PushThruVolume := _$CondVolume
10: Store _$0PreventDropsVolume := 0
11:
12: Prepare racks
13: Store _$0DRNumber := 1
14: Call subroutine zzPlaceCartridgeRack: _$DRNumber=_$0DRNumber, _$Slot=1
15:
16: for _$0CurrentSample := 1 to _$NofSamples (Repeat _$NofSamples times) do
17:     Calculate pipetting positions
18:     Store _$0DestPos := (_$c1 * 2) - 1
19:     Store _$0DRNumber := 1
20:
21:     if _$0DestPost > 24 then
22:         Store _$0DestPos := (_$c1 * 2) - 25
23:         Store _$0DRNumber := 2
24:     end if
```

25:     *Transfer liquid to cartridges*
26:       **Store** _$CalcPipVolume := _$0CondVolume
27:       **while** _$0CalcPipVolume > 0 **do**
28:         **if** _$0CalcPipVolume > 3000 **then**
29:           **Store** _$AspVol := 3000
30:           **Pipette action - Block begin**
31:           **Pipette action** Systemliq. System → Cartridges_&_$0DRNumber, _$0DestPos - _$0DestPos + 1, Dilution Vol: _$0AspVol, Liq. Param: Conditioning, Device LPipArm, Used Tips _$0TipBit
32:           **Pipette action - Block end and execute**
33:           **Store** _$DryPressure := 6000
34:           **Store** _$PushTime := 2
35:           **Call subroutine** zzPushThru: _$DestPos = _$0DestPos, _$DRNumber = _$0DRNumber, _$DryPressure = _$0DryPressure, _$PreventDropsVolume = _$0PreventDropsVolume, _$PushThruVolume = _$0PushThruVolume, _$PushTime = _$0PushTime, _$TipBit = 3
36:         **end if**
37:
38:         **if** _$0CalcPipVolume < 3000 **then**
39:           **Store** _$AspVol := _$0CalcPipVolume
40:           **Pipette action - Block begin**
41:           **Pipette action** Systemliq. System → Cartridges_&_$0DRNumber, _$0DestPos - _$0DestPos 1, Dilution Vol: _$0AspVol, Liq. Param: Conditioning, Device LPipArm, Used Tips _$0TipBit
42:           **Pipette action - Block end and execute**
43:           **Store** _$DryPressure := 6000
44:           **Store** _$PushTime := _$0CondVolume/750
45:           **Call subroutine** zzPushThru: _$DestPos = _$0DestPos, _$DRNumber = _$0DRNumber, _$DryPressure = _$0DryPressure, _$PreventDropsVolume = _$0PreventDropsVolume, _$PushThruVolume = _$0PushThruVolume, _$PushTime = _$0PushTime, _$TipBit = 3
46:         **end if**
47:
48:         **Store** _$0CalcPipVolume = _$0CalcPipeVolume _$0AspVol
49:       **end while**
50: **end for**
51:
52: **Switch device** Device SPEStationFixation, Switch to OFF
53: **Flush** Diluent: System, 1000 µL
54:
55: *Only if two racks in process*
56: **for** _$0DRNumber == 2 **do**
57:   **Call subroutine** zzChangeCartridgeRack: _$DRNumber = _$0DRNumber
58: **end for**

Equilibration

Following the conditioning step, the equilibration subroutine draws solvent from solvent tubes 1 and 2. Solvent (up to 3000 µL) is aspirated from the solvent bottles; this is then dispensed onto the cartridges through tips 1 and 2, and the liquid pushed through using nitrogen. This is repeated until all equilibration solvent has been dispensed onto the cartridges and pushed through; the cartridge rack is then returned to its home position (alg 4).

Algorithm 4 Subroutine: z2Equilibration

1: *Set equilibration values*
2: **Store** _$0EquiVolume := _$EquiVolume
3: **Store** _$0SourceRack := System 1 2
4:
5: *Used Tips*
6: **Store** _$0TipBit := 3
7:
8: *Set pushthrough values*
9: **Store** _$0PushThruVolume := _$EquiVolume*2
10: **Store** _$0PreventDropsVolume := 200
11:
12: *Prepare racks*
13: **Store** _$0DRNumber := 1
14: **Call subroutine** zzPlaceCartridgeRack: _$DRNumber=_$0DRNumber, _$Slot=1
15:
16: **for** _$0CurrentSample := 1 to _$NofSamples (Repeat _$NofSamples times) **do**
17:     *Calculate pipetting positions*
18:     **Store** _$0DestPos := (_$c1 * 2) - 1
19:     **Store** _$0DRNumber := 1
20:
21:     **if** _$0DestPost > 24 **then**
22:         **Store** _$0DestPos := (_$c1 * 2) - 25
23:         **Store** _$0DRNumber := 2
24:     **end if**
25:     *Change racks if DestPos = 25*
26:     **if** _$0DRNumber == 2 **then**
27:         **Switch device** Device SPEStationFixation, Switch to OFF
28:         **Move Rack** RGripArm, Rack Cartridges_1, Destination Carrier: 5Pos, Slot: 1
29:         **Call subroutine** zzPlaceCartridgeRack: _$DRNumber = _$0DRNumber, _$Slot = 1
30:     **end if**
31:
32:     *Transfer liquid to cartridges*
33:     **Store** _$CalcPipVolume := _$0EquiVolume
34:     **while** _$0CalcPipVolume > 0 **do**
35:         **if** _$0CalcPipVolume > 3000 **then**
36:             **Store** _$AspVol := 3000
37:             **Pipette action - Block begin**
38:             **Pipette action** Systemliq. System → Cartridges_&_$0DRNumber, _$0DestPos - _$0DestPos + 1, Dilution Vol: _$0AspVol, Liq. Param: Conditioning, Device LPipArm, Used Tips _$0TipBit
39:             **Pipette action - Block end and execute**
40:             **Store** _$DryPressure := 7500
41:             **Store** _$PushTime := 2
42:             **Call subroutine** zzPushThru: _$DestPos = _$0DestPos, _$DRNumber = _$0DRNumber, _$DryPressure = _$0DryPressure, _$PreventDropsVolume = _$0PreventDropsVolume, _$PushThruVolume = _$0PushThruVolume, _$PushTime = _$0PushTime, _$TipBit = 3
43:         **end if**

44:     **if** _$0CalcPipVolume < 3000 **then**
45:         **Store** _$AspVol := _$0CalcPipVolume
46:         **Pipette action - Block begin**
47:         **Pipette action** Systemliq. System → Cartridges_&_$0DRNumber, _$0DestPos - _$0DestPos 1, Dilution Vol: _$0AspVol, Liq. Param: Conditioning, Device LPipArm, Used Tips _$0TipBit
48:         **Pipette action - Block end and execute**
49:         **Store** _$DryPressure := 7500
50:         **Store** _$PushTime := _$0EquiVolume/1500
51:         **Call subroutine** zzPushThru: _$DestPos = _$0DestPos, _$DRNumber = _$0DRNumber, _$DryPressure = _$0DryPressure, _$PreventDropsVolume = _$0PreventDropsVolume, _$PushThruVolume = _$0PushThruVolume, _$PushTime = _$0PushTime, _$TipBit = 3
52:     **end if**
53:
54:     **Store** _$0CalcPipVolume = _$0CalcPipeVolume _$0AspVol
55:   **end while**
56: **end for**
57:
58: **Switch device** Device SPEStationFixation, Switch to OFF
59: **Flush** Diluent: System, 1000 µL
60:
61: **Call subroutine** zzParkCartridgePack: _$DRNumber = _$0DRNumber
62: *Only if two racks in process*
63: **for** _$0DRNumber == 2 **do**
64:   **Call subroutine** zzChangeCartridgeRack: _$DRNumber = _$0DRNumber
65:
66: **end for**

Sample load

If the filtrate is to be collected, wash vials are first placed to catch the filtrate from the load, with the cartridge rack placed above; if the filtrate isn't collected, the filtrate runs through to waste. The load subroutine draws sample from two sample vials simultaneously (up to 3000 µL), which is deposited onto the cartridges through tips 1 and 2 — the sample is then pushed through with nitrogen. This is repeated until the sample has been dispensed onto the cartridges and pushed through (alg 5).

Algorithm 5 Subroutine: z3LoadSample

1: *Set LoadSamples values*
2: **Store** _$0LoadSampleVolume := _$LoadSampleVolume
3: **Store** _$0Collect := _$Collect
4:
5: *Used Tips*
6: **Store** _$0TipBit := 3

7: *Set pushthrough values*
8: **Store** _$0PushThruVolume := _$LoadSampleVolume*2
9: **Store** _$0PreventDropsVolume := 200
10:
11: **for** _$0C1 := 1 to _$NofSamples (Repeat _$NofSamples times) **do**
12:     *Calculate pipetting positions*
13:     **Store** _$0DestPos := (_$c1 * 2) - 1
14:     **Store** _$0SourceRack1 := Sample_1
15:     **Store** _$0SourceRack2 := Sample_2
16:     **Store** _$0SourcePos := _$c1
17:     **if** _$c1 > 12 **then**
18:         **Store** _$0DRNumber := 2
19:         **Store** _$0DestPos := (_$c1 * 2) - 25
20:         **Store** _$0SourceRack1 := Sample_3
21:         **Store** _$0SourceRack2 := Sample_4
22:         **Store** _$0SourcePos := _$c1 - 12
23:     **end if**
24:     *Change racks if DestPos = 25*
25:     **if** _$0DRNumber == 2 **then**
26:         **Store** _$0ChangeRackNumber := _$0DRNumber -1
27:         **Call subroutine** zzChangeRacksCollect: _$CoverType=SampleCover_&_$0ChangeRackNumber, _$DRNumber = _$0ChangeRackNumber, _$ExtractRack=SampleExtract_&_$0ChangeRackNumber
28:     **end if**
29:     **if** _$0Collect == 0 **then**
30:         **Call subroutine** zzPlaceCartridgeRack: _$DRNumber=_$0DRNumber, _$Slot=1
31:     **end if**
32:     **if** _$0Collect == 1 **then**
33:         **Call subroutine** zzPlaceRacksCollect: _$CoverType=SampleCover_&_$0DRNumber, _$DRNumber=_$0DRNumber, _$ExtractRack=SampleExtract_&_$0DRNumber
34:     **end if**
35:
36:     *Transfer liquid to cartridges*
37:     **Store** _$CalcPipVolume := _$0LoadSampleVolume
38:     **while** _$0CalcPipVolume > 0 **do**
39:         **if** _$0CalcPipVolume > 3000 **then**
40:             **Store** _$AspVol := 3000
41:             **Pipette action - Block begin**
42:             **Pipette action** Systemliq. System → Cartridges_&_$0DRNumber, _$0DestPos - _$0DestPos + 1, Dilution Vol: _$0AspVol, Liq. Param: Conditioning, Device LPipArm, Used Tips _$0TipBit
43:             **Pipette action - Block end and execute**
44:             **Store** _$DryPressure := 2500
45:             **Store** _$PushTime := 3
46:             **Call subroutine** zzPushThru: _$DestPos = _$0DestPos, _$DRNumber = _$0DRNumber, _$DryPressure = _$0DryPressure, _$PreventDropsVolume = _$0PreventDropsVolume, _$PushThruVolume = _$0PushThruVolume, _$PushTime = _$0PushTime, _$TipBit = 3
47:         **end if**

48:         **if** _$0CalcPipVolume < 3000 **then**
49:           **Store** _$AspVol := _$0CalcPipVolume
50:           **Pipette action - Block begin**
51:           **Pipette action** Systemliq. System → Cartridges _&_ _$0DRNumber, _$0DestPos - _$0DestPos 1, Dilution Vol: _$0AspVol, Liq. Param: Conditioning, Device LPipArm, Used Tips _$0TipBit
52:           **Pipette action - Block end and execute**
53:           **Store** _$DryPressure := 2500
54:           **Store** _$PushTime := _$0LoadSampleVolume/1500
55:           **Call subroutine** zzPushThru: _$DestPos = _$0DestPos, _$DRNumber = _$0DRNumber, _$DryPressure = _$0DryPressure, _$PreventDropsVolume = _$0PreventDropsVolume, _$PushThruVolume = _$0PushThruVolume, _$PushTime = _$0PushTime, _$TipBit = 3
56:         **end if**
57:
58:         **Store** _$0CalcPipVolume = _$0CalcPipeVolume _$0AspVol
59:       **end while**
60: **end for**
61:
62: **if** _$0Collect == 1 **then**
63:     **Call subroutine** zzChangeRacksCollect: _$DRNumber=_$0DRNumber
64: **end if**
65: **if** _$0Collect == 0 **then**
66:     **if** _$0DRNumber == 2 **then**
67:       **Call subroutine** zzChangeCartridgeRack: _$DRNumber = _$0DRNumber
68:     **end if**
69: **end if**

Wash

If the filtrate was not collected in the load step, the cartridge rack is replaced in its home position, the wash vials are placed in the dispensing area and the cartridge rack placed on top. The wash subroutine draws solvent from solvent tubes 1 and 2. Solvent (up to 3000 μL) is aspirated from the solvent bottles; this is then dispensed onto the cartridges through tips 1 and 2, and the liquid pushed through using nitrogen. This is repeated until all wash solvent has been dispensed onto the cartridges and pushed through; the cartridge rack is then returned to its home position (alg 6).

Algorithm 6 Subroutine: z4WashAndCollectSample

1: *Set wash values*
2: **Store** _$0WashVolume := _$WashVolume
3:
4: **if** _$WashVolume > 0 **then**
5:     **Store** _$0SourceRack := System 1 2
6:
7:     *Set pushthrough values*
8:     **Store** _$0PushThruVolume := _$WashVolume*2
9:     **Store** _$0PreventDropsVolume := 300

```
10:     Used Tips
11:        Store _$0TipBit := 3
12:
13:     Calculate pipetting positions
14:        Store _$0CurrentSample := _$CurrentSample
15:        Store _$0DestPos := (_$0CurrentSample * 2) - 1
16:        Store _$0DRNumber := 1
17:        Call subroutine zzPlaceRacksCollect: _$CoverType=SampleCover_&_$0DRNumber,
   _$DRNumber=_$0DRNumber, _$ExtractRack=SampleExtract_&_$0DRNumber
18:     Transfer liquid to cartridges
19:        Store _$CalcPipVolume := _$0WashVolume
20:        while _$0CalcPipVolume > 0 do
21:           if _$0CalcPipVolume > 3000 then
22:              Store _$AspVol := 3000
23:              Pipette action - Block begin
24:              Pipette action Systemliq. System → Cartridges_&_$0DRNumber, _$0DestPos - _$0DestPos
   + 1, Dilution Vol: _$0AspVol, Liq. Param: Conditioning, Device LPipArm, Used Tips _$0TipBit
25:              Pipette action - Block end and execute
26:              Store _$DryPressure := 5000
27:              Store _$PushTime := 5
28:              Call subroutine zzPushThru: _$DestPos = _$0DestPos, _$DRNumber = _$0DRNumber,
   _$DryPressure = _$0DryPressure, _$PreventDropsVolume = _$0PreventDropsVolume,
   _$PushThruVolume = _$0PushThruVolume, _$PushTime = _$0PushTime, _$TipBit = 3
29:           end if
30:           if _$0CalcPipVolume < 3000 then
31:              Store _$AspVol := _$0CalcPipVolume
32:              Pipette action - Block begin
33:              Pipette action Systemliq. System → Cartridges_&_$0DRNumber, _$0DestPos - _$0DestPos
   1, Dilution Vol: _$0AspVol, Liq. Param: Conditioning, Device LPipArm, Used Tips _$0TipBit
34:              Pipette action - Block end and execute
35:              Store _$DryPressure := 5000
36:              Store _$PushTime := _$WashVolume/100
37:              Call subroutine zzPushThru: _$DestPos = _$0DestPos, _$DRNumber = _$0DRNumber,
   _$DryPressure = _$0DryPressure, _$PreventDropsVolume = _$0PreventDropsVolume,
   _$PushThruVolume = _$0PushThruVolume, _$PushTime = _$0PushTime, _$TipBit = 3
38:              Flush Diluent: System, 5000 µL
39:           end if
40:
41:           Store _$0CalcPipVolume = _$0CalcPipeVolume _$0AspVol
42:        end while
43: end if
```

Elution 1

The first elution subroutine also incorporates the drying of cartridges — tips 3 and 4 are emptied of solvent and nitrogen is blown through the cartridges for a given amount of time. The wash vials are then replaced to their storage position, and draws solvent from solvent tubes 3 and 4. Solvent (up to 3000 µL) is aspirated from the solvent bottles; this is then dispensed

onto the cartridges through tips 3 and 4, and the liquid pushed through using nitrogen. This is repeated until all elution solvent has been dispensed onto the cartridges and pushed through; the cartridge rack and elution 1 vials are then returned to their home position (alg 7).

Algorithm 7 Subroutine: z6Elution1

 1: *Set Elution1 values*
 2: **Store** _$0Elut1Volume := _$Elut1Volume
 3: **Store** _$0SourceRack := System 3 4
 4:
 5: **Store** _$0DryCartridges := 1
 6: **Store** _$0DryDelay := _$DryDelay * 60
 7: *Set pushthrough values*
 8: **Store** _$0PushThruVolume := _$Elut1Volume
 9: **Store** _$0PreventDropsVolume := 500
10:
11: *Used Tips*
12: **Store** _$0TipBit := 12
13:
14: *Calculate pipetting positions*
15: **Store** _$0CurrentSample := _$CurrentSample
16: **Store** _$0DestPos := (_$0CurrentSample * 2) - 1
17: **Store** _$0DRNumber := 1
18: **if** _$0DestPost > 24 **then**
19:     **Store** _$0DestPos := (_$0CurrentSample * 2) - 1
20:     **Store** _$0DRNumber := 2
21:     **Switch device** Device PIPStationFixation, Switch to OFF
22:     **Move Rack** RGripArm, Rack Cartridges_1, Destination Carrier: 5Pos, Slot: 1
23: **end if**
24: *Dry cartridges*
25: **if** _$0DryCartridges==1 **then**
26:     *Empty lines*
27:     **if** _$0CurrentSample == 1 **then**
28:         **Call subroutine** zzEmptyTips: _$TipBit=3
29:     **end if**
30:     *Prepare racks*
31:     **Store** _$0DryPressure := 22000
32:     **Call subroutine** zzPlaceCartridgeRack: _$DRNumber=_$0DRNumber, _$Slot=1
33:     **Call subroutine** zzDryCartridges: _$DestPos=_$0DestPos, _$DRNumber=_$0DRNumber, _$DryDelay=_$0DryDelay, _$DryPressure=_$0DryPressure, _$TipBit=3
34:     **Call subroutine** zzParkCartridgeRack: _$DRNumber=_$0DRNumber
35:     **Flush** Diluent: System, 1 flush cycle
36: **end if**
37: *Prepare racks*
38: **Call subroutine** zzPlaceRacksCollect: _$CoverTpye=ElutCover_&_$0DRNumber, _$DRNumber=_$0DRNumber, _$ExtractRack=ElutExtract_&_$0DRNumber
39:

40: *Transfer liquid to cartridges*
41: **Store** _$CalcPipVolume := _$0Elut1Volume
42: **while** _$0CalcPipVolume > 0 **do**
43:     **if** _$0CalcPipVolume > 3000 **then**
44:         **Store** _$AspVol := 3000
45:         **Pipette action - Block begin**
46:         **Pipette action** Systemliq. System → Cartridges_&_$0DRNumber, _$0DestPos - _$0DestPos + 1, Dilution Vol: _$0AspVol, Liq. Param: Conditioning, Device LPipArm, Used Tips _$0TipBit
47:         **Pipette action - Block end and execute**
48:         **Store** _$DryPressure := 2500
49:         **Store** _$PushTime := 3
50:         **Call subroutine** zzPushThru: _$DestPos = _$0DestPos, _$DRNumber = _$0DRNumber, _$DryPressure = _$0DryPressure, _$PreventDropsVolume = _$0PreventDropsVolume, _$PushThruVolume = _$0PushThruVolume, _$PushTime = _$0PushTime, _$TipBit = 12
51:     **end if**
52:     **if** _$0CalcPipVolume < 3000 **then**
53:         **Store** _$AspVol := _$0CalcPipVolume
54:         **Pipette action - Block begin**
55:         **Pipette action** Systemliq. System → Cartridges_&_$0DRNumber, _$0DestPos - _$0DestPos 1, Dilution Vol: _$0AspVol, Liq. Param: Conditioning, Device LPipArm, Used Tips _$0TipBit
56:         **Pipette action - Block end and execute**
57:         **Store** _$DryPressure := 2000
58:         **Store** _$PushTime := _$0Elut1Volume/100
59:         **Call subroutine** zzPushThru: _$DestPos = _$0DestPos, _$DRNumber = _$0DRNumber, _$DryPressure = _$0DryPressure, _$PreventDropsVolume = _$0PreventDropsVolume, _$PushThruVolume = _$0PushThruVolume, _$PushTime = _$0PushTime, _$TipBit = 12
60:         **Flush** Diluent: System, 500 μ L
61:         **Call subroutine** zzParkAllRacks
62:     **end if**
63:
64:     **Store** _$0CalcPipVolume = _$0CalcPipeVolume _$0AspVol
65: **end while**

Elution 2

In the second elution, solvent is drawn from the small solvent reservoirs. This solvent (up to 3000 µL) is then dispensed onto the cartridges through tips 3 and 4, and the liquid pushed through using nitrogen. This is repeated until all elution solvent has been dispensed onto the cartridges and pushed through; the cartridge rack and elution 2 vials are then returned to their home position (alg 8).

Algorithm 8 Subroutine: z7Elution2

1: *Set Elution2 values*
2: **Store** _$0Elut2Volume := _$Elut2Volume
3: **if** _$0Elut2Volume > 0 **then**
4:     **Store** _$0SourceRack := Solvent_3
5:     **Store** _$0SourceRack := Solvent_4

6:      *Set pushthrough values*
7:      **Store** _$0PushThruVolume := _$Elut2Volume
8:      **Store** _$0PreventDropsVolume := 500
9:
10:     *Used Tips*
11:     **Store** _$0TipBit := 12
12:
13:     *Calculate pipetting positions*
14:     **Store** _$0CurrentSample := _$CurrentSample
15:     **Store** _$0DestPos := (_$0CurrentSample * 2) - 1
16:     **Store** _$0DRNumber := 1
17:
18:     *Prepare racks*
19:     **Call subroutine** zzPlaceRacksCollect: _$CoverTpye=ElutCover_&_$0DRNumber, _$DRNumber=_$0DRNumber, _$ExtractRack=ElutExtract_&_$0DRNumber
20:     *Transfer liquid to cartridges*
21:     **Store** _$CalcPipVolume := _$0Elut2Volume
22:     **while** _$0CalcPipVolume > 0 **do**
23:        **if** _$0CalcPipVolume > 3000 **then**
24:           **Store** _$AspVol := 3000
25:           **Pipette action - Block begin**
26:           **Pipette action** Systemliq. System → Cartridges_&_$0DRNumber, _$0DestPos - _$0DestPos + 1, Dilution Vol: _$0AspVol, Liq. Param: Conditioning, Device LPipArm, Used Tips _$0TipBit
27:           **Pipette action - Block end and execute**
28:           **Store** _$DryPressure := 2500
29:           **Store** _$PushTime := 3
30:           **Call subroutine** zzPushThru: _$DestPos = _$0DestPos, _$DRNumber = _$0DRNumber, _$DryPressure = _$0DryPressure, _$PreventDropsVolume = _$0PreventDropsVolume, _$PushThruVolume = _$0PushThruVolume, _$PushTime = _$0PushTime, _$TipBit = 12
31:        **end if**
32:        **if** _$0CalcPipVolume < 3000 **then**
33:           **Store** _$AspVol := _$0CalcPipVolume
34:           **Pipette action - Block begin**
35:           **Pipette action** Systemliq. System → Cartridges_&_$0DRNumber, _$0DestPos - _$0DestPos 1, Dilution Vol: _$0AspVol, Liq. Param: Conditioning, Device LPipArm, Used Tips _$0TipBit
36:           **Pipette action - Block end and execute**
37:           **Store** _$DryPressure := 2000
38:           **Store** _$PushTime := _$0Elut2Volume/100
39:           **Call subroutine** zzPushThru: _$DestPos = _$0DestPos, _$DRNumber = _$0DRNumber, _$DryPressure = _$0DryPressure, _$PreventDropsVolume = _$0PreventDropsVolume, _$PushThruVolume = _$0PushThruVolume, _$PushTime = _$0PushTime, _$TipBit = 12
40:           **Flush** Diluent: System, 500 μ L
41:           **Call subroutine** zzParkAllRacks
42:        **end if**
43:
44:        **Store** _$0CalcPipVolume = _$0CalcPipeVolume _$0AspVol
45:     **end while**
46: **end if**

5.2.3 Minor subroutines

Subroutines 'zzPushThru' and 'zzDryCartridges' form an integral part of the overall method, with the former being significantly edited from the original. The remaining subroutines are included for completeness.

Pushthrough

The tips are emptied in preparation for the nitrogen; nitrogen is then pushed through at a predefined pressure for a predefined amount of time (dependent on the method calling the subroutine). After the time has elapsed the nitrogen pressure slowly ramps off and the system is flushed (alg 9).

Algorithm 9 Subroutine: zzPushThru

1: **Store** _$0DryPressure := _$DryPressure
2: **Store** _$0PushThruVolume := _$PushThruVolume
3: **Store** _$0PreventDropsVolume := _$PreventDropsVolume
4: **Store** _$0DestPos := _$DestPos
5: **Store** _$0DRNumber := _$DRNumber
6: **Store** _$0TipBit := _$TipBit
7: **Store** _$0PushTime := _$PushTime
8:
9: **Call subroutine** zzEmptyTips
10: *Pipette*
11: **Write to memory device** Device SPETime, Write:200
12: *Set up pressure, start slowly*
13: **Write to memory device** Device SPEPressureValueSet13, Write:533
14: **Write to memory device** Device SPEPressureValueSet24, Write:533
15: *Move to rack and blow slowly*
16: **XY movement to rack position** Cartridges_&_$0DRNumber, _$0DestPost - _$0DestPos + 1, Device LPipArmPip, Used Tips _$0TipBit
17: **Z movement at current position** Device LPipArmPip, UsedTips _$0TipBit: Go to ZSCAN, Submerge 15 mm, Speed 100
18: **if** _$TipBit <= 3 **then**
19:     **Switch device** Device LPipArmValveTip1, Switch to ON
20:     **Switch device** Device LPipArmValveTip2, Switch to ON
21: **end if**
22: **if** _$TipBit > 3 **then**
23:     **Switch device** Device LPipArmValveTip3, Switch to ON
24:     **Switch device** Device LPipArmValveTip4, Switch to ON
25: **end if**
26: **Delay** Delay of 5 secs

27: **Switch device** Device SPESpressureValueSet13, Write: _$0DryPressure
28: **Switch device** Device SPESpressureValueSet24, Write: _$0DryPressure
29: **Delay** Delay of _$Pushtime secs
30: **Switch device** Device SPESpressureValueSet13, Write: _$0DryPressure / 2
31: **Switch device** Device SPESpressureValueSet24, Write: _$0DryPressure / 2
32: **Delay** Delay of 5 secs
33: **Switch device** Device SPESpressureValueSet13, Write: _$0DryPressure / 5
34: **Switch device** Device SPESpressureValueSet24, Write: _$0DryPressure / 5
35: **Delay** Delay of 5 secs
36: **Switch device** Device SPESpressureValueSet13, Write: _$0DryPressure / 10
37: **Switch device** Device SPESpressureValueSet24, Write: _$0DryPressure / 10
38: **Delay** Delay of 5 secs
39: **Switch device** Device SPESpressureValueSet13, Write: 0
40: **Switch device** Device SPESpressureValueSet24, Write: 0
41: **Delay** Delay of 5 secs
42: **Switch device** Device LPipArmValveTip1, Switch to OFF
43: **Switch device** Device LPipArmValveTip2, Switch to OFF
44: **Switch device** Device LPipArmValveTip3, Switch to OFF
45: **Switch device** Device LPipArmValveTip4, Switch to OFF
46: **Z movement at current position** Device LPipArmPip, UsedTips _$0TipBit: Go to ZSCAN, Submerge mm, Speed 100
47: **Z movement at current position** Device LPipArmPip, UsedTips _$0TipBit: Go to GLOBALZTRAVEL, Submerge 15 mm, Speed
48: **Flush** Diluent: System, 1 flush cycle

Dry cartridges

Similar to the pushthrough subroutine, drying the cartridges involves the pushing through of nitrogen at a predefined pressure for a predefined amount of time, after which it ramps off (alg 10).

Algorithm 10 Subroutine: zzDryCartridges

1: **Store** _$0DryPressure := 22000
2: **Store** _$0DryDelay := _$DryDelay
3: **Store** _$0DestPos := _$DestPos
4: **Store** _$0DRNumber := _$DRNumber
5: **Store** _$0TipBit := _$TipBit
6: **Write to memory device** Device SPETime, Write:_$0DryDelay + 200
7: **Write to memory device** Device SPEPressureValueSet13, Write:533
8: **Write to memory device** Device SPEPressureValueSet24, Write:533
9: **XY movement to rack position** Cartridges_&_$0DRNumber, _$0DestPost - _$0DestPos + 1, Device LPipArmPip, Used Tips _$0TipBit
10: **Z movement at current position** Device LPipArmPip, UsedTips _$0TipBit: Go to ZSCAN, Submerge 15 mm, Speed 100
11: **if** _$TipBit <= 3 **then**
12:     **Switch device** Device LPipArmValveTip1, Switch to ON
13:     **Switch device** Device LPipArmValveTip2, Switch to ON
14: **end if**

15: **if** _$TipBit > 3 **then**
16:     **Switch device** Device LPipArmValveTip3, Switch to ON
17:     **Switch device** Device LPipArmValveTip4, Switch to ON
18: **end if**
19: **Delay** Delay of 5 secs
20: **Switch device** Device SPESpressureValueSet13, Write: _$0DryPressure
21: **Switch device** Device SPESpressureValueSet24, Write: _$0DryPressure
22: **Delay** Delay of _$DryDelay secs
23: **Switch device** Device SPESpressureValueSet13, Write: _$0DryPressure / 2
24: **Switch device** Device SPESpressureValueSet24, Write: _$0DryPressure / 2
25: **Delay** Delay of 5 secs
26: **Switch device** Device SPESpressureValueSet13, Write: _$0DryPressure / 5
27: **Switch device** Device SPESpressureValueSet24, Write: _$0DryPressure / 5
28: **Delay** Delay of 5 secs
29: **Switch device** Device SPESpressureValueSet13, Write: _$0DryPressure / 10
30: **Switch device** Device SPESpressureValueSet24, Write: _$0DryPressure / 10
31: **Delay** Delay of 5 secs
32: **Switch device** Device SPESpressureValueSet13, Write: 0
33: **Switch device** Device SPESpressureValueSet24, Write: 0
34: **Delay** Delay of 5 secs
35: **Switch device** Device LPipArmValveTip1, Switch to OFF
36: **Switch device** Device LPipArmValveTip2, Switch to OFF
37: **Switch device** Device LPipArmValveTip3, Switch to OFF
38: **Switch device** Device LPipArmValveTip4, Switch to OFF
39: **Z movement at current position** Device LPipArmPip, UsedTips _$0TipBit: Go to ZSCAN, Submerge mm, Speed 100
40: **Z movement at current position** Device LPipArmPip, UsedTips _$0TipBit: Go to GLOBALZTRAVEL, Submerge 15 mm, Speed
41: **Flush** Diluent: System, 1 flush cycle

Empty tips

Emptying the tips (alg 11, normally used for transferring liquids) is done by withdrawing any solvent from the tips and applying a short stream of nitrogen over a waste inlet — this prepares the station for nitrogen drying and pushthrough.

Algorithm 11 Subroutine: zzEmptyTips
1: **Store** _$0TipBit := _$TipBit
2:
3: **XY movement to rack position** DRY, 1 - 4, Device LPipArmPip, Used Tips 15
4: **Simple Aspirate with PipPump** Device LPipArmPip, Used Tips _$0TipBit, Pump: Asp volume 4000 µL, Execute
5: *Set up pressure, start slowly*
6: **Write to memory device** Device SPEPressureValueSet13, Write:1500
7: **Write to memory device** Device SPEPressureValueSet24, Write:1500
8: **Switch device** Device VACUUMPUMP, Switch to ON
9: **Z movement at current position** Device LPipArmPip, UsedTips _$0TipBit: Go to ZMAX, Submerge mm, Speed

10: **if** _$TipBit <= 3 **then**
11:     **Switch device** Device LPipArmValveTip1, Switch to ON
12:     **Switch device** Device LPipArmValveTip2, Switch to ON
13: **end if**
14: **if** _$TipBit > 3 **then**
15:     **Switch device** Device LPipArmValveTip3, Switch to ON
16:     **Switch device** Device LPipArmValveTip4, Switch to ON
17: **end if**
18: **Switch device** Device DRYVALVE, switch to ON
19: **Delay** Delay of 5 secs
20: **Switch device** Device SPESpressureValueSet13, Write: 16000
21: **Switch device** Device SPESpressureValueSet24, Write: 16000
22: **Delay** Delay of 5 secs
23: **Switch device** Device SPESpressureValueSet13, Write: 8000
24: **Switch device** Device SPESpressureValueSet24, Write: 8000
25: **Delay** Delay of 5 secs
26: **Switch device** Device SPESpressureValueSet13, Write: 4000
27: **Switch device** Device SPESpressureValueSet24, Write: 4000
28: **Delay** Delay of 5 secs
29: **Switch device** Device SPESpressureValueSet13, Write: 0
30: **Switch device** Device SPESpressureValueSet24, Write: 0
31: **Delay** Delay of 5 secs
32: *Close all valves*
33: **Switch device** Device DRYVALVE, switch to OFF
34: **Switch device** Device VACUUMPUMP, Switch to OFF
35: **Switch device** Device LPipArmValveTip1, Switch to OFF
36: **Switch device** Device LPipArmValveTip2, Switch to OFF
37: **Switch device** Device LPipArmValveTip3, Switch to OFF
38: **Switch device** Device LPipArmValveTip4, Switch to OFF
39: **Z movement at current position** Device LPipArmPip, UsedTips _$0TipBit: Go to GLOBALZTRAVEL, Submerge 15 mm, Speed

Rack movement subroutines

The 'place cartridge rack' subroutine (alg 12) moves the appropriate cartridge rack to the appropriate position for the method in the SPE station.

Algorithm 12 Subroutine: zzPlaceCartridgeRack
1: **Switch device** Device SPEStationFixation, Switch to OFF
2: **Move Rack** RGripArm, Rack Cartridges_&_$DRNumber, Destination Carrier: SPEStation, Slot: _$Slot
3: **Switch device** Device SPEStationFixation, Switch to ON

The 'change cartridge rack' subroutine (alg 13) disables the SPE station fixation and swaps out the cartridge rack when multiple experiments are running, necessitating multiple racks.

Algorithm 13 Subroutine: zzChangeCartridgeRack
1: **Switch device** Device SPEStationFixation, Switch to OFF
2: **Move Rack** RGripArm, Rack Cartridges_&_$DRNumber, Destination Carrier: 5Pos, Slot: _$DRNumber

The 'place collection racks' subroutine (alg 14) moves the collection vial and cartridge racks to the SPE station, dependent on the method being run.

Algorithm 14 Subroutine: zzPlaceRacksCollect

1: **Switch device** Device SPEStationFixation, Switch to OFF
2: **Move Rack** RGripArm, Rack _$0CoverType, Destination Carrier: Park_Cover, Slot: 1
3: **Move Rack** RGripArm, Rack _$0ExtractRack, Destination Carrier: SPEStation, Slot: 1
4: **Move Rack** RGripArm, Rack Cartridges_&_$DRNumber, Destination Carrier: SPEStation, Slot: 2
5: **Switch device** Device SPEStationFixation, Switch to ON

The 'change collection rack' subroutine (alg 15) changes the collection vial rack when multiple experiments are running, necessitating multiple racks.

Algorithm 15 Subroutine: zzChangeRackCollect

1: **Switch device** Device SPEStationFixation, Switch to OFF
2: **Move Rack** RGripArm, Rack Cartridges_&_$DRNumber, Destination Carrier: 5Pos, Slot: _$DRNumber
3: **Move Rack** RGripArm, Rack ElutExtract_&_$DRNumber, Destination Carrier: ElutExtract_&_$DRNumber, Slot: 1
4: **Move Rack** RGripArm, Rack ElutCover_&_$DRNumber, Destination Carrier: ElutExtract_&_$DRNumber, Slot: 2
5: **Move Rack** RGripArm, Rack SampleExtract_&_$DRNumber, Destination Carrier: SampleExtract_&_$DRNumber, Slot: 1
6: **Move Rack** RGripArm, Rack SampleCover_&_$DRNumber, Destination Carrier: SampleExtract_&_$DRNumber, Slot: 2

The 'park cartridge rack' subroutine (alg 16) returns the cartridge rack to its home position.

Algorithm 16 Subroutine: zzParkAllRacks

1: **Switch device** Device SPEStationFixation, Switch to OFF
2: **Move Rack** RGripArm, Rack Cartridges_&_$DRNumber, Destination Carrier: Park_Cartridge, Slot: 1

The 'park all racks' subroutine (alg 17) replaces all cartridge and vial racks in their home positions.

Algorithm 17 Subroutine: zzParkAllRacks

1: **Switch device** Device SPEStationFixation, Switch to OFF
2: **Move Rack** RGripArm, Rack Cartridges_1, Destination Carrier: 5Pos, Slot: 1
3: **Move Rack** RGripArm, Rack ElutExtract_1, Destination Carrier: ElutExtract_1, Slot: 1
4: **Move Rack** RGripArm, Rack ElutCover_1, Destination Carrier: ElutExtract_1, Slot: 2
5: **Move Rack** RGripArm, Rack ElutExtract_2, Destination Carrier: ElutExtract_2, Slot: 1
6: **Move Rack** RGripArm, Rack ElutCover_2, Destination Carrier: ElutExtract_2, Slot: 2
7: **Move Rack** RGripArm, Rack SampleExtract_1, Destination Carrier: SampleExtract_1, Slot: 1
8: **Move Rack** RGripArm, Rack SampleCover_1, Destination Carrier: SampleExtract_1, Slot: 2

5.3   Validation of methods

With the SPE protocols having been automated, the final step before their broader use required that they be demonstrated to function in an identical manner to manual SPE methods. In chapter 3, the use of principal component analysis was fruitful in demonstrating differences between the retention protocols of the manual SPE methods developed; here, however, the same techniques can be used in order to demonstrate similarity between datasets.

5.3.1   Methodology

SPE experiments

Utilising the automation protocol written for the Bruker SamplePro SPE robotic system developed in this chapter, a Waters Oasis HLB cartridge (6 mL capacity and 500 mg bed weight) was conditioned with methanol (6 mL), then equilibrated with water (6 mL). Pooled urine (3 mL) was loaded onto the cartridge, which was then washed with water (6 mL) to eliminate interferences. The cartridge was then dried with nitrogen for 10 minutes; the retained metabolites were then eluted with methanol (6 mL). This experiment was replicated in quadruplicate.

The remaining validation methods are tabulated in table 5.1.

NMR sample preparation

Washes and elutions were dried under nitrogen and reconstituted in ultrapure water (3 mL). Buffer containing trimethylsilylpropionate (TSP) as a chemical shift reference standard was added to 540 µL of reconstituted sample, as described by Dona *et al.*[88]. 580 µL of the manually vortexed sample was then transferred into 5mm SampleJet NMR racks.

Samples which required additional 2D NMR experiments were dried under nitrogen and reconstituted in D_2O (3 mL). TSP phosphate buffer (60 µL) was added to 540 µL of reconstituted sample, and 580 µL of the resulting manually vortexed sample was transferred into 5mm NMR tubes.

Table 5.1: Conditions utilised for each automated SPE method

SPE Cartridge	Conditioning	Equilibration	Sample	Wash	Elution 1	Elution 2
HLB	2% FA in MeOH	2% FA in water	Urine	2% FA in water	2% FA in MeOH	-
HLB	MeOH	Water	pH 5 Urine	Water	MeOH	-
HLB	MeOH	Water	pH 2 Urine	Water	MeOH	-
SCX	ACN	Water	Urine	2% FA in water	ACN	5% NH$_4$OH in MeOH
SCX	ACN	Water	pH 5 Urine	2% FA in water	ACN	5% NH$_4$OH in MeOH
SCX	ACN	Water	pH 2 Urine	2% FA in water	ACN	5% NH$_4$OH in MeOH

NMR data acquisition

All 1D experiments were run using a Bruker Avance III 600 MHz spectrometer equipped with a BBI room temperature probe and SampleJet. Samples were analysed using one-dimensional water-suppressed ^{1}H NOESY experiments at 300 K.

Data analysis

NMR datasets were imported into MatLab using the Imperial Metabolic Profiling and Chemometrics Toolbox (IMPaCTS)[93]. Water (4.26-5.50 ppm) and formate (8.25–8.63 ppm) peak regions were removed from the spectra to eliminate interferences; the spectra were then normalised against the TSP region (-0.5 to 0.5 ppm) using a probabilistic quotient normalisation function[94]. Principal Component Analysis (PCA) plots were subsequently constructed with 5 principal components.

5.3.2   Results

If automated methods are to replace manual methods, the spectral data produced by the two must be indistinguishable. Principal component analysis allows for the dimensionality of data to be reduced down to a set of orthogonal components which aim to represent as much variability within the data as possible. A PCA was hence constructed from datasets consisting of the elutions from HLB and SCX methods (fig 5.3).

The distribution of elution datasets here is — naturally — comparable to the distribution of C_{18}, HLB, phenyl, SCX, SAX, and PBA elutions as demonstrated in chapter 3 (fig 3.2). In this instance, the bulk of the differentiation is caused by the natural differences between reversed-phase and ion exchange retention methods; by contrast, the differences in PCA scores between manual and automated methods are minor for all pairwise comparisons, and are of the same order of magnitude as differences between replicates.
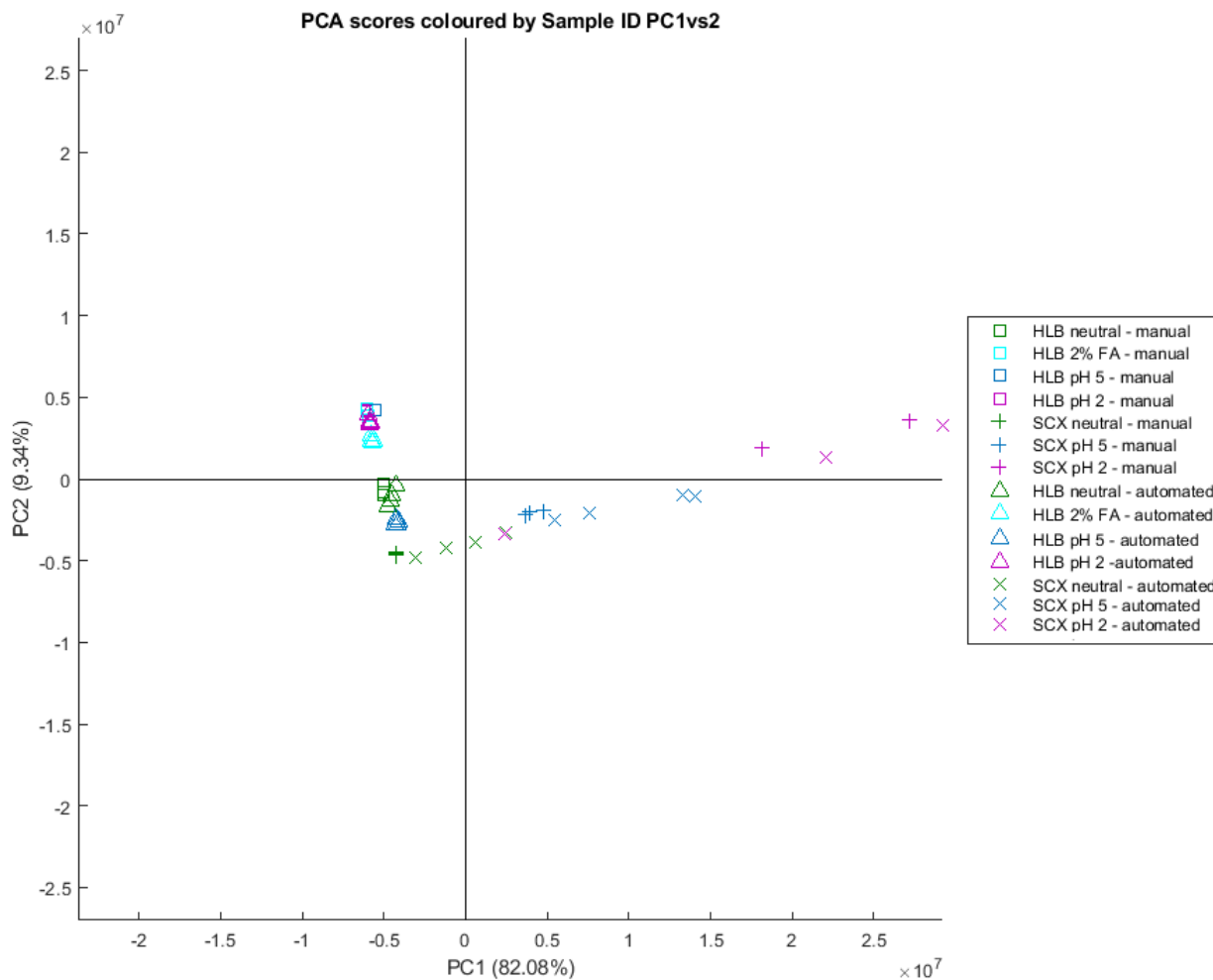
Figure 5.3: PCA scores plot built using NMR data from all automated and manual HLB and SCX SPE elutions of natural urine, PC1 (82.08%) vs PC2 (9.34%).

The analysis becomes marginally more nuanced when inspecting the PC3v4 scores plot for the same datasets (fig 5.4); in this instance, there is significant clustering along PC3 between automated and manual methods. The only significant peak in the loadings plot is at 0.00 ppm, suggesting that this differentiation is caused by shimming (fig 5.5).
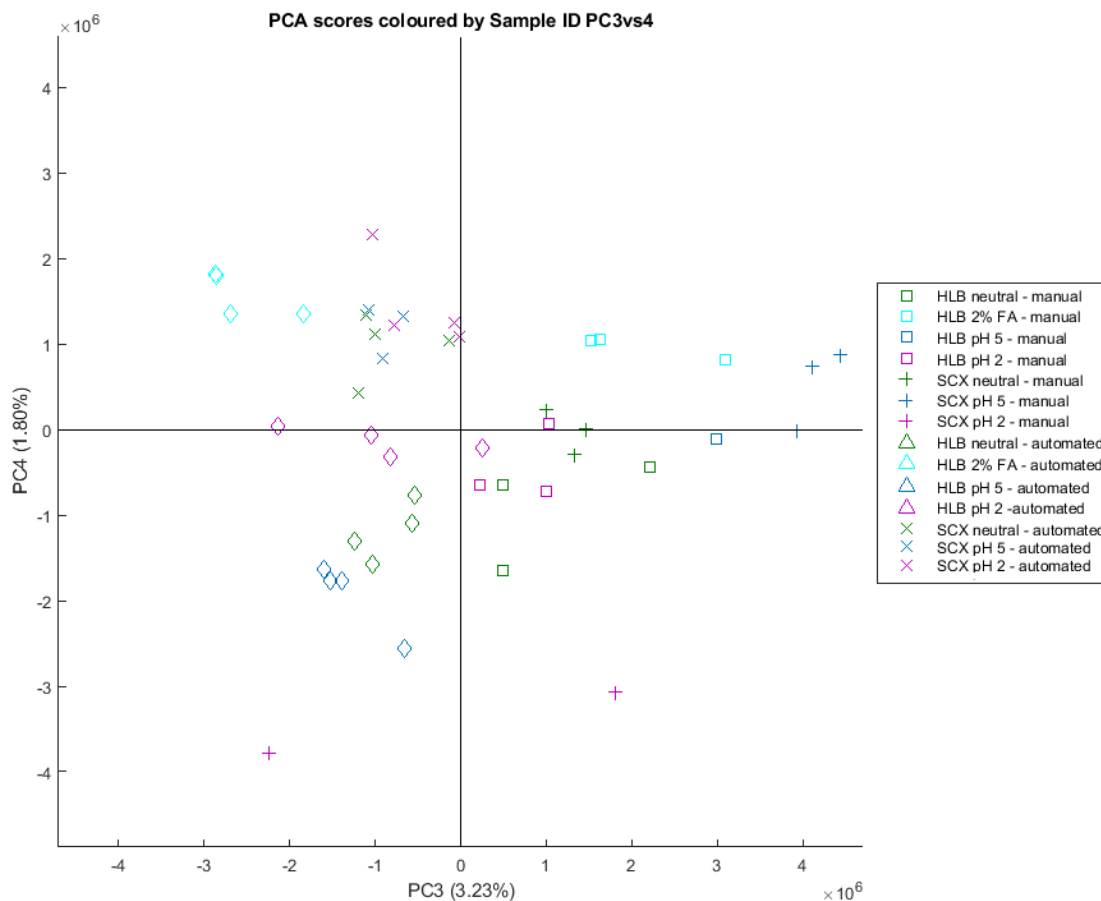
Figure 5.4: PCA scores plot built using NMR data from all automated and manual HLB and SCX SPE elutions of natural urine, PC3 (3.23%) vs PC4 (1.80%).
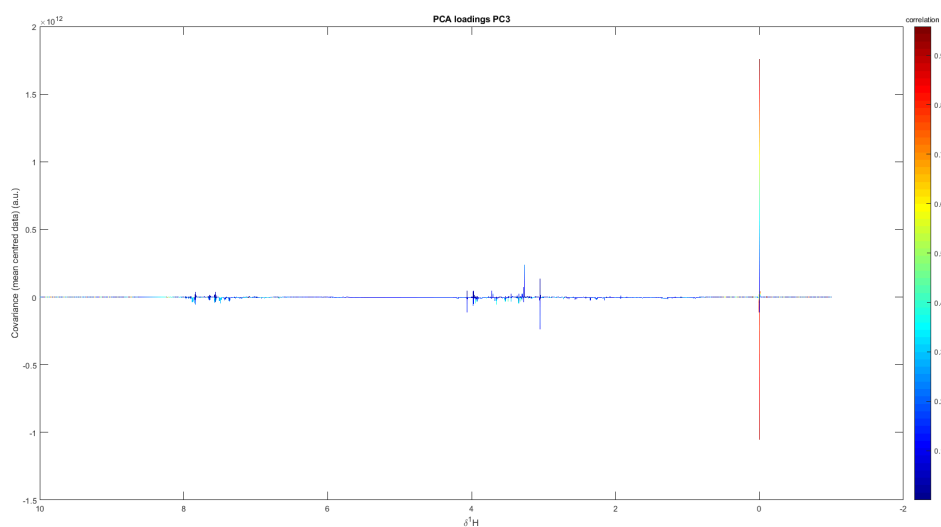
Figure 5.5: PC3 loadings plot from all automated and manual HLB and SCX SPE elutions of natural urine, demonstrating that the overwhelming contribution to differences between spectra being TSP peak intensity.

When running untargeted SPE experiments, it is also important to be able to analyse the wash component. The PCA constructed for the HLB raw samples, washes, and elutions demonstrated differentiation based only on method, with only minor differences between manual and automated methods of comparable size to replicate differences (fig 5.6).
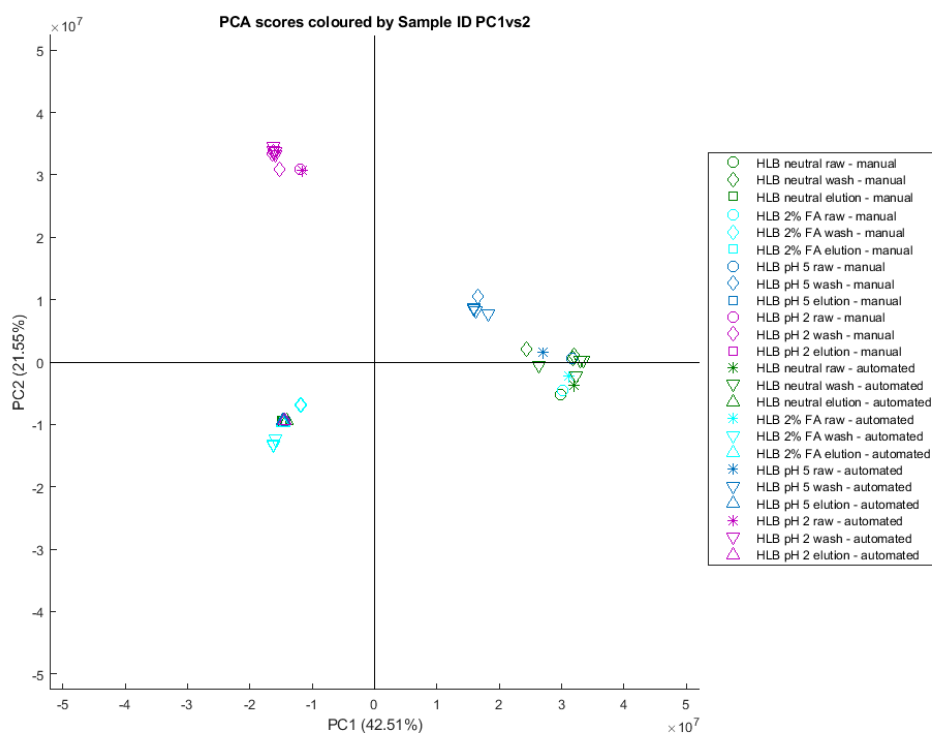
Figure 5.6: PCA scores plot built using NMR data from all automated and manual HLB SPE raw samples, washes, and elutions of natural urine, PC1 (59.45%) vs PC2 (20.01%).

Finally, a PCA plot was constructed using datasets from HLB elutions alone (fig 5.7). In this instance, there is some differentiation along PC2 — apparently being driven predominantly by the HLB 2% FA elutions.
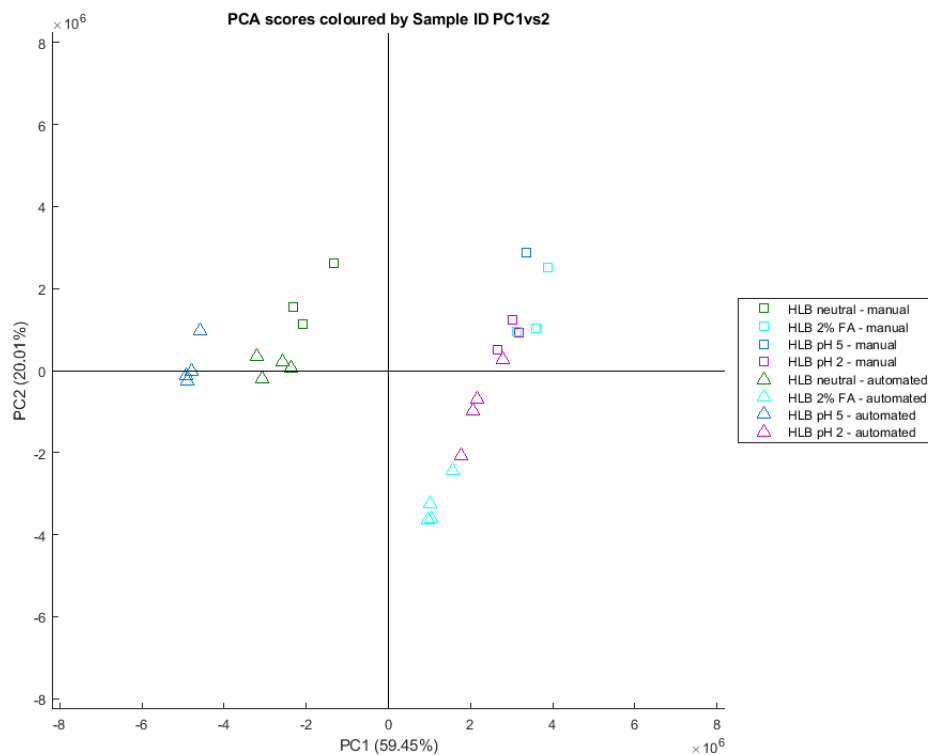
Figure 5.7: PCA scores plot built using NMR data from all automated and manual HLB SPE raw samples, washes, and elutions of natural urine, PC1 (59.45%) vs PC2 (20.01%).

The loadings plot (fig 5.8) demonstrates that the majority of this differentiation appears to be caused partly by shimming, but also partly by marginally greater retention of a specific compound producing spectral peaks at 7.29 (d), 7.21 (d), and 2.35 (s) — reference to previous annotation efforts implies that this metabolite is 4-cresol sulfate. Comparing the raw untreated urine samples together does demonstrate a small increase in intensity for the 4-cresol sulfate peaks. A subsequent literature search described the metabolic process of protein-binding of 4-cresol sulfate to plasma albumin, with the resulting protein complex then being filtered into urine[112]. Since the complex itself would not demonstrate the same peaks in NMR as the free molecule, it is hence possible that this binding is interrupted as urine stocks slowly degrade, causing a release of the free molecule and hence appearing to increase the concentration. As a result, this does not appear to affect the mechanism between automated and manual SPE experiments and hence is no impediment to replacement of manual methods with automated equivalents.

144
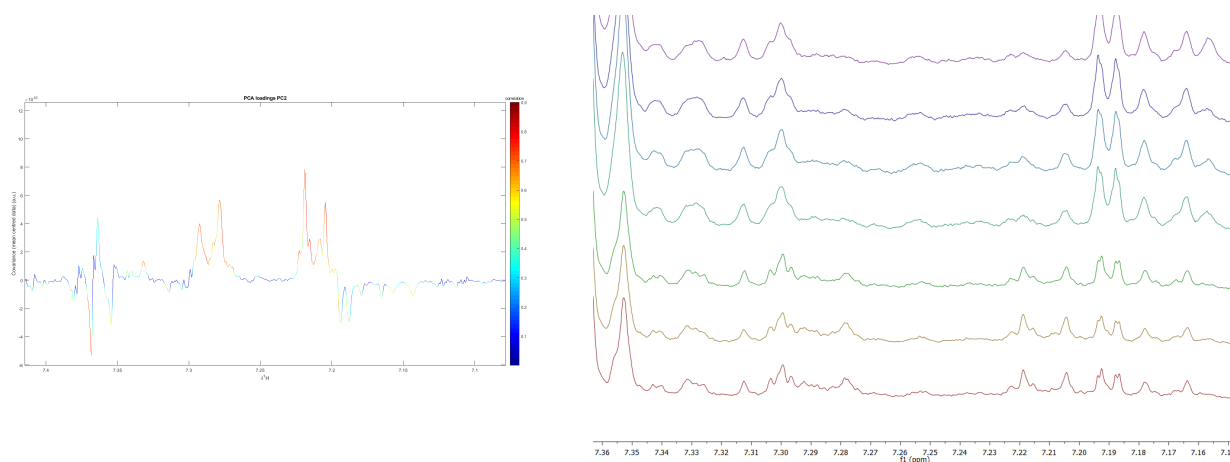
Figure 5.8: PC2 loadings plot of the HLB elutions (left) demonstrating characteristic peaks of 4-cresol sulfate; stacked spectra (right) demonstrating greater retention of the metabolite at 7.28 and 7.22 ppm in the three manual 1D spectra (bottom) compared to the four automated 1D spectra (top).

5.3.3 Discussion

Having developed manual SPE-NMR methods for metabolite annotation, it became apparent that the automation of these methods would significantly reduce the time needed to achieve fractionation and concentration, as well as increasing the reproducibility of data through use of consistent and precise steps. Learning the visual script editor associated with the Bruker SamplePro SPE robot allowed for greater control over automated method development, expanding support to 6 mL and ion exchange SPE cartridges. The validation of these methods revealed no significant differences between the automated protocols and their manual equivalents, with any notable variability being caused by unrelated phenomena. The automated protocols would go on to be used in the selective retention of glucuronides with a method incorporating HLB cartridges with 2% formic acid.

5.4 Selective retention of glucuronides through use of automated untargeted SPE methods

5.4.1 Introduction

Having developed untargeted SPE protocols (chapter 3), it became feasible to run automated experiments which aimed to retain metabolites based on a specific compound class or functional group. Exometabolites — such as drugs and pharmaceuticals — are commonly conjugated with glucuronic acid in the body during phase II metabolism to reduce toxicity and aid elimination[113]. As a result, glucuronides are usually water-soluble and pharmacologically inactive, although some continue to demonstrate activity — for example, morphine-6-glucuronide is well known to have analgesic properties more potent than unmetabolised morphine[114]. Glucuronides are more easily diagnosable in NMR spectra than comparable phase II metabolites, such as sulphates, due to the presence of peaks from the glucuronide moiety (a doublet at 5.12ppm being diagnostic of a glucuronide metabolite), and are ubiquitous in the human urinary metabolome.

In-born errors of metabolism may qualitatively or quantitatively affect the profile of a biofluid, leading to metabolite levels differing significantly from those of a comparable healthy person — the rare disease cerebrotendinous xanthomatosis, for example, causes an increase in bile alcohol glucuronides in the bile, urine, and stool of patients, as a consequence of a deficiency of the enzyme sterol 27-hydroxylase restricting the metabolic pathways available[115]. Beyond diagnostics, some glucuronides — such as fluorouracil–glucuronide and the aforementioned morphine-6-glucuronide — demonstrate increased therapeutic potential, making these compounds, among others, potential drug candidates[116]. As a result, an automated method able to selectively retain and/or eliminate glucuronides in a sample would have significant importance in metabolic profiling in the screening of biological samples.

When selecting an automated method with which to attempt to selectively retain glucuronides in urine, two SPE sorbents were considered viable. The first, a method utilising HLB cartridges with 2% formic acid in all steps, had retained a significant number of metabolites when tested on natural urine; the hydrocarbon backbone of the sugar presented a possible point of retention

on a reversed-phase cartridge, although the hydrophobic character of that moeity had not been demonstrated. With limited time available, the broad retentive capacity of this cartridge appeared a good candidate for attempting to selectively retain glucuronides.

The second method — utilising a PBA cartridge — appeared ideal for glucuronide retention, as PBA cartridges are nominally able to retain compounds containing a vicinal diol through their unique covalent retention mechanism. However, two issues advised against the use of the PBA cartridge in this situation: the first being, simply, that the manual experiment described in chapter 3 requires the use of five separate solvents, whereas the automated system supports only up to four solvents, prohibiting its automation. More importantly, however, was the contradiction between the nominal retention targets of the PBA cartridge with the reality achieved through manual experimentation.

Mannitol — a linear sugar alcohol — was noticeably retained through the use of the PBA method, as expected; in urine, mannitol is expected to be present in a concentration of between 1-20µmol/mmol creatinine[7]. However, the closely related isomer sorbitol has also been recorded as having a concentration in urine of between 2.5-18.7µmol/mmol creatinine[7], yet is not present in the PBA elutions. The precise reason for this is unclear; this could range from steric effects interfering with the ability of the PBA sorbent to retain sorbitol, to limited retention sites leading to the more sterically favoured isomer — mannitol — being selectively retained to the detriment of other metabolites. With this in mind, methods utilising PBA cartridges are, for now, apparently unlikely to result in broad retention of glucuronides, even in the instance where PBA methods are able to be automated.

A literature search revealed similar experiments run by Tugnait *et al.*, where both C_{18} and PBA cartridges were used to attempt to selectively retained glucuronides. Both were moderately successful in isolating glucuronides (phenophthalein glcuruonide, 6-bromo-2-naphthyl-β-D-glucuronide, α-Naphthyl-β-D-glucuronide, and p-Nitrophyl-β-D-glucuronide), although the authors describe the C_{18} phase as having low specificity despite being more generally useful, whereas the PBA cartridge demonstrated higher selectivity for a 'more limited range of compounds'. In practice, this meant that the aglycone structure likely also performed an important, complementary role in the context of retention on SPE sorbents.

While HLB with 2% formic acid was hence considered the more appropriate method to use in the limited time available, the desired retention of glucuronides was established almost entirely on the hydrophobic character exhibited by the sugar moeity; however, none of the methods developed — including methods utilising PBA — resulted in any significant retention of compounds assigned to the 'carbohydrates and carbohydrate conjugates', incorporating erythritol, glucose, tartaric acid, and glycerol. While this doesn't necessarily eliminate the possibility of retention of glucuronide conjugates — especially since the chemical character of the aglycone will contribute to retention, as previously demonstrated by Wilson et al. with the retention of paracetamol and ibuprofen glucuronides[54] — experimental results are required to measure the extent of retention, if present at all.

In total, five conjugates were targeted for retention — ibuprofen glucuronide, paracetamol glucuronide, pregnanediol-3-glucuronide, bolasterone glucuronide, and ethyl glucuronide. Metabolites derived from the NSAIDS ibuprofen and paracetamol are easily generated *in situ* and contain benzene rings, which provide hydrophobic character needed for retention on reversed-phase cartridges. Pregnanediol-3-glucuronide and bolasterone glucuronide are non-aromatic endogenous steroid metabolites with a high non-polar surface area, which can demonstrate the importance — or lack of importance — of aromatic substituents in retention on reversed-phase cartridges. Finally, ethyl glucuronide is one of the simplest exogenous glucuronides, suggesting that its retention — if applicable — should be predominantly guided by intermolecular forces acting on the glucuronide moiety.

5.4.2   Methodology

Pregnanediol, bolasterone, and ethyl glucuronides

10 mg pregnanediol from the Australian National Measurement Institute was dissolved in 2 mL ultrapure water. 175 µL of this solution was spiked into 4 mL pooled urine.

10 mg bolasterone glucuronide from the Australian National Measurement Institute was dissolved in 2 mL ultrapure water. 700 µL of this solution was spiked into 4 mL pooled urine.

10 mg ethyl glucuronide from Enzo life sciences was dissolved in 2mL ultrapure water. 175

µL of this solution was spiked into 4 mL pooled urine.

Spike-in experiments

All samples were analysed using one-dimensional water-suppressed ^{1}H NOESY experiments at 300 K.

A sample of pooled urine was analysed by NMR spectroscopy to produce a ^{1}H NOESY spectrum. To this sample, 2 µL of 5 mg/mL pregnanediol-3-glucuronide, ethyl glucuronide, or bolasterone glucuronide was added and a ^{1}H NOESY spectrum acquired. This was repeated with doubled quantities of pregnanediol-3-glucuronide (4 µL, 8 µL, 16 µL, etc) until the peaks of the glucuronide became clear — for pregnanediol-3-glucuronide this was after 30 µL had been applied; for ethyl glucuronide, after 126 30 µL; and for bolasterone glucuronide, after 30 µL.

Ibuprofen and paracetamol glucuronides

A healthy adult male was dosed with 400 mg ibuprofen. After 4 hours a urine sample was collected into 500 mL Corning tubes pre-rinsed with ultrapure water, and stored at 4°C.

A healthy adult male was dosed with 500 mg paracetamol. After 4 hours a urine sample was collected into 500 mL Corning tubes pre-rinsed with ultrapure water, and stored at 4°C.

A healthy adult male was dosed with 400 mg ibuprofen and 500 mg paracetamol. After 4 hours urine samples were collected into 500 mL Corning tubes pre-rinsed with ultrapure water, and stored at 4°C.

SPE protocol

Using the automation protocol written for the Bruker SamplePro SPE robotic system developed in this chapter, a Waters Oasis HLB cartridge (6 mL capacity and 500 mg bed weight) was conditioned with 2% formic acid in acetonitrile (6 mL), then equilibrated with 2% formic acid in water (6 mL). Glucuronide-spiked urine (3 mL) was loaded onto the cartridge, which was then washed with 2% formic acid in water (6 mL) and dried under nitrogen for 10 minutes. The retained metabolites were then eluted with 2% formic acid in acetonitrile (6 mL).

All glucuronide SPE experiments were run in quadruplicate.

NMR sample preparation

Washes and elutions were dried under nitrogen and reconstituted in ultrapure water (3 mL). Buffer containing trimethylsilylpropionate (TSP) as a chemical shift reference standard was added to 540 µL of reconstituted sample, as described by Dona *et al.*[88]. 580 µL of the manually vortexed sample was then transferred into 5mm SampleJet NMR racks.

NMR data acquisition

All 1D experiments were run using a Bruker Avance III 600 MHz spectrometer equipped with a BBI room temperature probe and a SampleJet. Samples were analysed using one-dimensional water-suppressed ^{1}H NOESY experiments at 300 K.

Additional ^{1}H–^{1}H J-resolved experiments, and 2D-NMR experiments, including ^{1}H–^{1}H Total Correlation Spectroscopy (TOCSY), ^{1}H–^{1}H Correlation Spectroscopy (COSY), and ^{1}H–^{13}C Heteronuclear Single Quantum Coherence spectroscopy (HSQC), were utilised for metabolite annotation. The data from the 2D NMR experiments was acquired using a Bruker Avance III 600 MHz spectrometer equipped with a CryoTCI triple resonance CryoProbe.

5.4.3 Evaluation of glucuronide retention on HLB cartridge

While NMR assignments for ibuprofen and paracetamol glucuronide are known in the literature, comparisons between 1D NMR spectra and literature can only achieve a limited level of confidence. Hence, to achieve a high level of confidence, the generation of ibuprofen- and paracetamol glucuronide requires a control urine sample (in this instance, previously acquired pooled urine), urine containing paracetamol glucuronide, urine containing ibuprofen glucuronide, and urine containing both glucuronides (fig 5.9). Thankfully, collection of these samples is facile as they are generated in significant concentrations as part of typical phase II metabolism in healthy adults. Ethyl glucuronide can similarly be generated *in situ* — with assignments from this method having been reported by Kim *et al.*[117] — but this (as with dosing healthy adults with bolasterone and pregnanediol) was considered unsuitable for the purposes of the project.
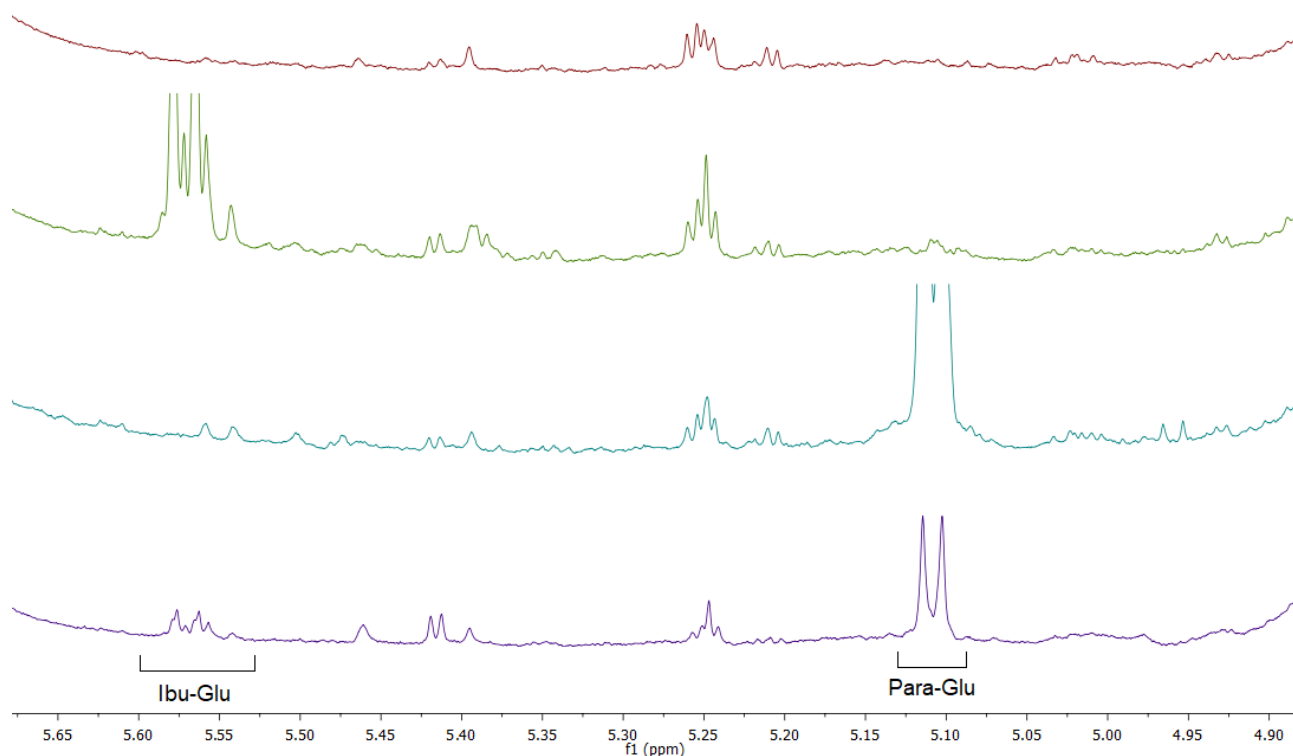
Figure 5.9: Spectral region 5.75-5.05 ppm demonstrating ibuprofen glucuronide multiplet at 5.57 ppm in ibuprofen (green) and ibuprofen/paracetamol (purple) dosed urine, and paracetamol glucuronide doublet at 5.11 in paracetamol (blue) and ibuprofen/paracetamol (purple) dosed urine.

The spike-in experiments and the collection of ibuprofen/paracetamol dosed urine allowed for representative peaks, unlikely to be overlapped by other peaks, to be recorded from the spectra (table 5.2) — pKa and log P were predicted using PerkinElmer Chemdraw Professional 18.

Table 5.2: Representative ^{1}H peaks for drug metabolites, assigned through spike-in experiment

Metabolite	NMR assignments (δ)	Predicted pKa	Predicted log P
Bolasterone glucuronide	0.81 (s)	17.06	1.78
Ethyl glucuronide	4.48 (d)	17.05	-1.72
Ibuprofen glucuronide	0.88 (d), 1.07 (d), 1.21 (s), 5.57 (m)	12.77	1.87

Paracetamol glucuronide	2.17 (d), 5.11 (d)	13.12	-1.33
Pregnanediol-3-glucuronide	0.63 (s), 0.93 (s)	17.05	2.52
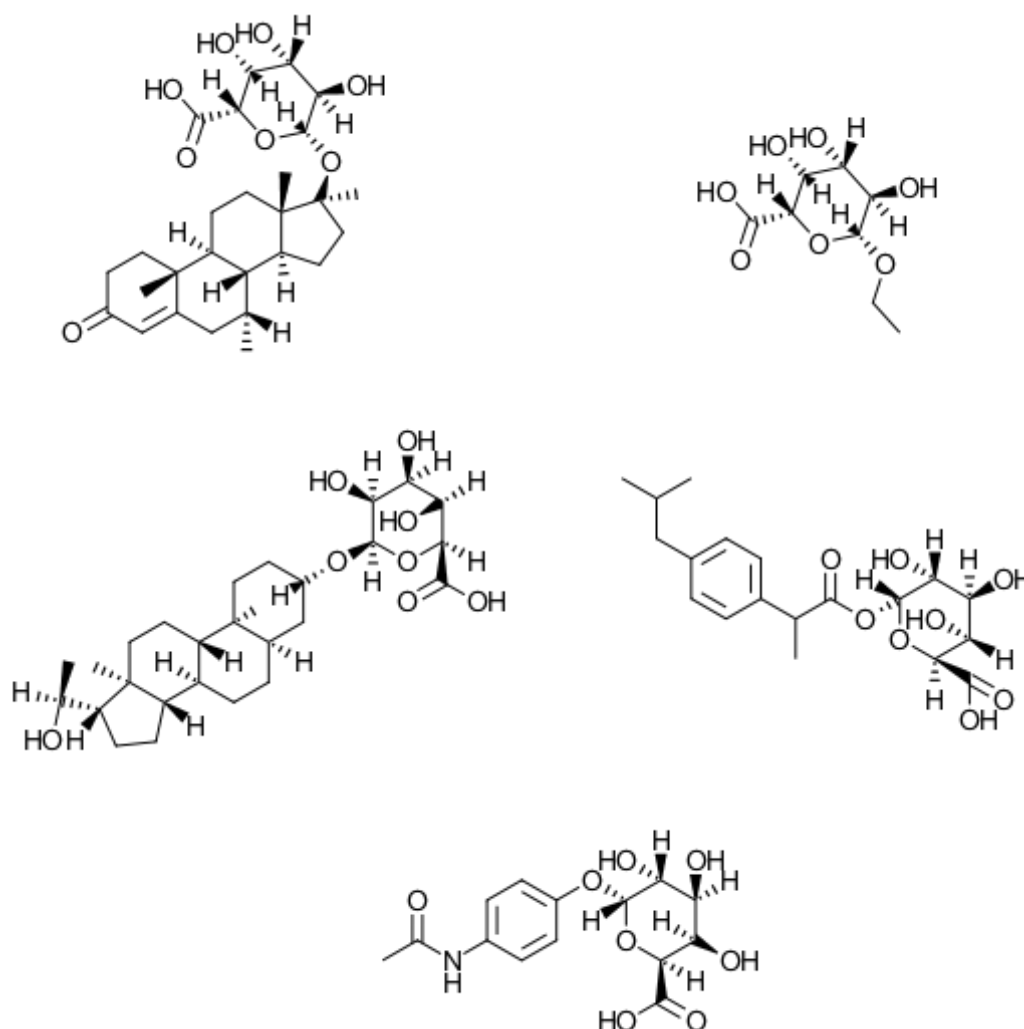
Figure 5.10: Drug metabolites, clockwise from top left: Bolasterone glucuronide, ethyl glucuronide, ibuprofen glucuronide, paracetamol glucuronde, pregnanediol-3-glucuronide.

The spectra generated from the SPE elutions require minimal processing to afford results demonstrating retention. Using the HLB 2% formic acid method, both paracetamol- and ibuprofen glucuronide were significantly retained. Through measuring peak area by comparison with the constant TSP concentration, it was determined that 60% of paracetamol glucuronide was retained, with the remaining 40% expressed in the wash (possibly due to saturation of retention sites on the sorbent); by comparison, 100% of ibuprofen glucuronide was retained (fig

5.11).
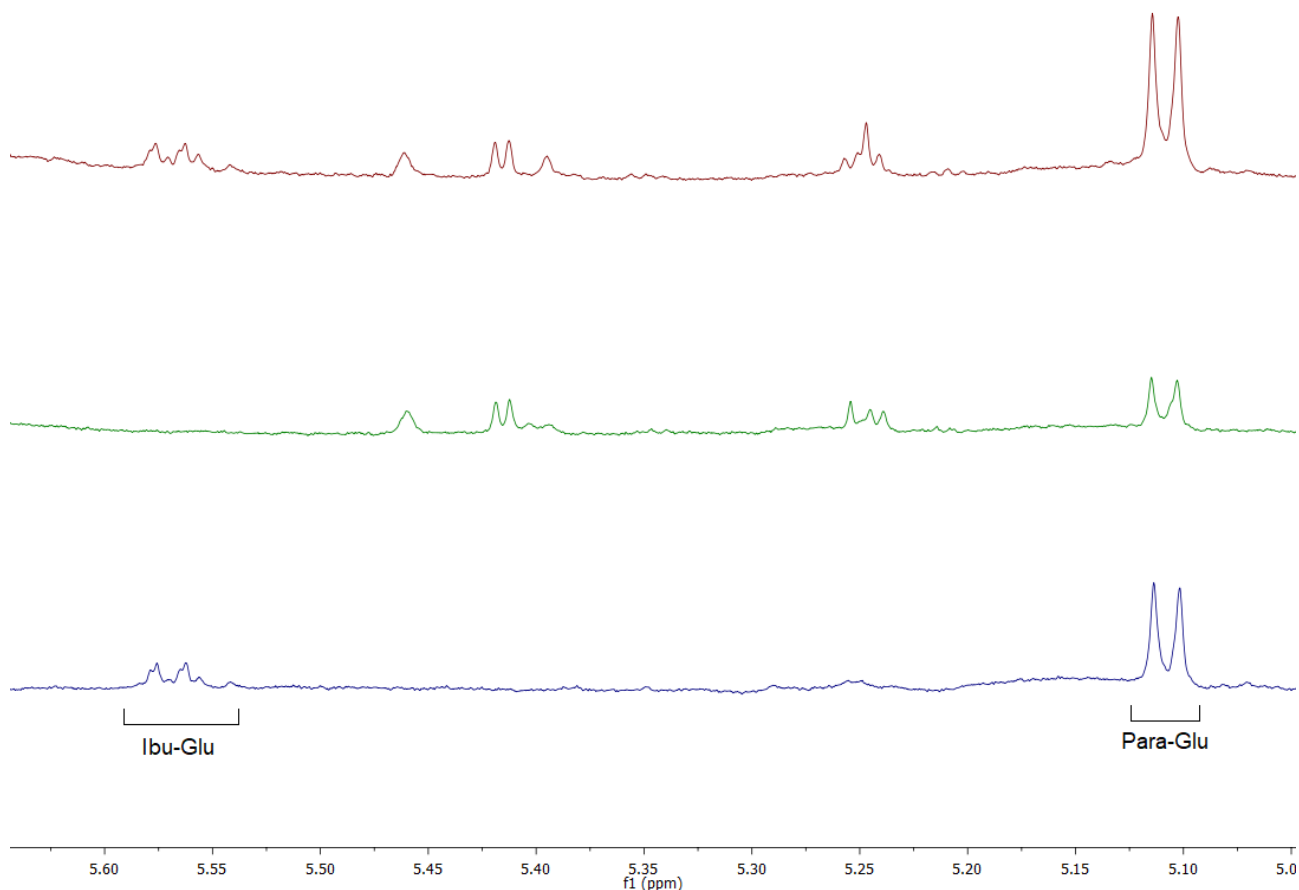
Figure 5.11: NMR spectra demonstrating ibuprofen- and paracetamol glucuronide peaks in the raw sample (red) and elution (blue). Paracetamol glucuronide is also present in the wash (green), but ibuprofen glucuronide is not.

Pregnanediol-3-glucuronide was similarly retained in three out of four elution fractions (although, curiously, absent in both wash and elution in the first). Comparatively, bolasterone glucuronide — an isomer of pregnanediol-3-glucuronide — is retained in all four elutions and absent in all four washes (fig 5.12).
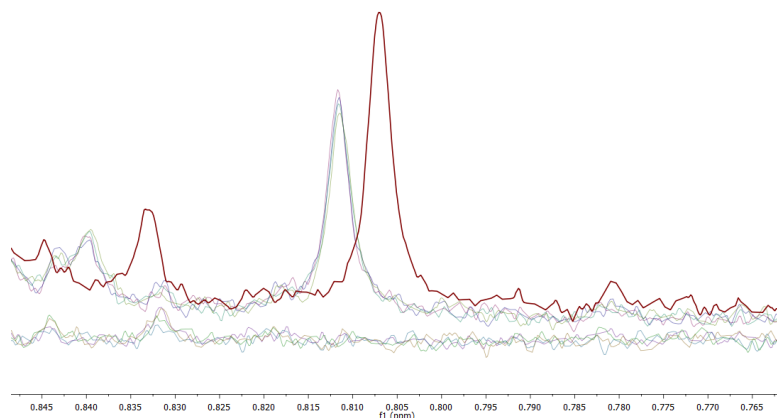
Figure 5.12: The bolasterone glucuronide peak at 0.81 is present in the raw, untreated urine (red) and the four elutions — albeit slightly shifted. By contrast, it is totally absent from the four washes.

Ethyl glucuronide, however, does not appear to demonstrate retention; while the expected doublet at 4.48 ppm does not appear in the wash, the region where the signal is anticipated to appear is empty in the elution (fig 5.13).
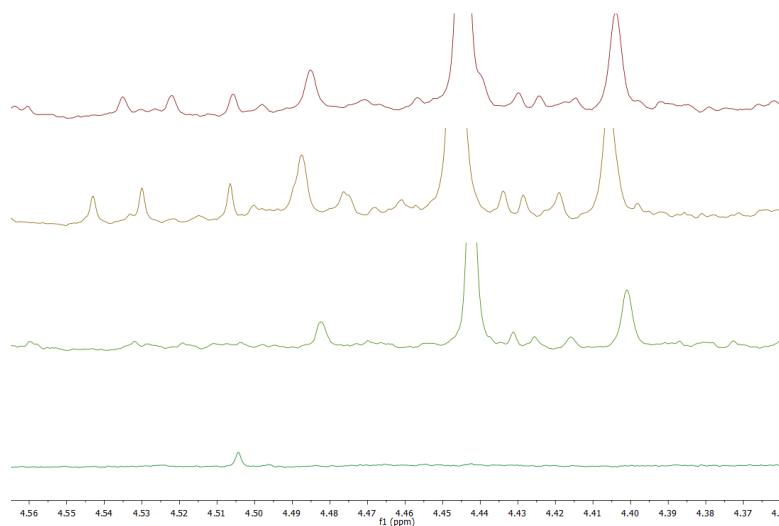
Figure 5.13: Stacked spectra from the ethyl glucuronide SPE experiments demonstrating a peak at 4.47 in the spiked raw sample (yellow) not present in the original, unspiked sample (red). A peak possibly corresponding to the glucuronide is present at 4.48 ppm in the wash (light green), but no peaks are present in the elution (dark green).

5.5   Conclusion

The translation of the manual protocols developed in chapter 3 into automated methods proved successful, with no significant variation between the manual and automated datasets being observed. Having developed the automated methods, one of them — utilising the HLB cartridge with 2$ formic acid at all stages — was used to attempt to selectively retain glucuronides, a common class of metabolites found in biofluids like urine. Unfortunately, the non-retention of ethyl glucuronide using HLB cartridges and 2% formic acid at all stages suggests that the retention of glucuronides is not solely dictated through intermolecular forces acting on the sugar. It is possible that the glucuronide does aid retention, but the results demonstrated here still suggest that metabolites require a significant hydrophobic component for retention to occur; it is hence unclear what effect the gluruonide has on retention, if any.

It is worth noting that a significant number of drug metabolites do possess hydrophobic character, such as from benzene rings; in fact, finding a glucuronide standard without a significant hydrocarbon skeleton proved difficult until ethyl glucuronide was sourced. Having demonstrated that a number of these (both aromatic and non-aromatic) can be retained, this SPE method may still be useful in selectively retaining some glucuronides. An inherent limitation of using SPE methods to gauge retention based on functional groups or compound classes is the extent to which results can be generalised to that entire compound class — while the confidence for this would be higher if ethyl glucuronide were retained, some degree of confidence can still be ascertained. Further experimentation using a broader range of glucuronides would be required in order to increase confidence further.

Chapter 6

General discussion and conclusion

During the course of this project, we have demonstrated and compared the use of different SPE methods for the retention of different compound classes within the human urinary metabolome. Different retention profiles can give unique insight into the metabolome by revealing metabolite peaks in NMR spectroscopy that had previously been obscured by peak overlap — SPE can hence be used to either remove the suppressing metabolite(s), or to isolate the unknown metabolite itself. These retention profiles can be differentiated based not only on their capacity for individual metabolites, but also for broader subclasses of compounds united by their structure commonalities, including shared functional groups. Hence, different methods can be utilised in order to give greater control over the annotation process.

6.1 Breakthrough volume experiments

The breakthrough experiments described in chapter 2 allowed each of the retention profiles of the SPE cartridges available to be studied through use of frontal chromatography. Retention here was gauged by referencing a restricted set of representative molecules with a broad range of functional groups, rather than the much larger key molecule tables used later; this enabled a preliminary sense of how useful each chemical modification might be when applied to biofluids, without requiring the significant annotation and quantification efforts which the subsequent method development steps would need. It also allowed for estimates of an ideal loading volume — while low loading volumes would minimise the loss of analytes, this would come at the cost of

spectral resolution. By contrast, a high loading volume would mean greater resolution (due to higher concentration in solution) but a greater loss of analytes.

The results from these experiments demonstrated that while normal phase sorbents broadly exhibited little useful retention (as anticipated), the reversed-phase and ion exchange sorbents showed far more promise. Of these, six cartridges — C_{18}, HLB, phenyl, SCX, SAX, and PBA were considered for further work. While some of these cartridges (most notably C_{18} and HLB) were selected for their broad-ranging retention capability, others (such as PBA) were selected for the unique set of compounds they retained. At this stage, it became clear that sorbents capable of broad-ranging retention are not necessarily preferable to sorbents demonstrating retention of specific compound classes or functional groups — indeed, broader retention may limit the removal of overlapping peaks.

6.2 Untargeted SPE method development

While SPE sub-profiling on complex mixtures has been demonstrated previously, the development efforts in chapter 3 aimed to drastically broaden the SPE methods for fractionation and concentration available to researchers — this resulted in what was described as 'untargeted SPE', which aims to increase concentration (and hence signal-noise) and reduce overlap in spectra as an aid to metabolite profiling, and where the target for retention isn't a single compound but instead broad classes of compounds in a complex mixture. The use of pH control proved to be exceptionally important, not only with ion exchange sorbents where pH control dictates the main retentive capacity of the sorbents, but even with reversed-phase methods, where more acidic conditions often resulted in better retention at the cost of sample authenticity (basified samples were also run, but were not included in chapter 3 due to retention being negligable).

Analysing the elutions resulting from the methods also re-contextualised the scale of the knowledge gap facing researchers — the number of easily annotated spectral features was far outnumbered by the number of NMR peaks where little information, such as correlated ^{13}C data, was available. The reduction of noise and peak overlap helped with this problem, as peaks from low-concentration compounds previously suppressed by high-concentration compounds became

visible — this was most notable for sorbents like PBA, where compounds such as mannitol was very clearly retained. However, the understanding of the retention mechanisms remains poor; for example, it remains unclear why mannitol was very well retained on PBA, whereas isomers of mannitol also usually present at a comparable concentration in urine — such as sorbitol — saw no retention.

During this manual stage there were several ideas generally considered outside the scope of this project — of metabolite annotation — which were briefly tested but which produced inconclusive results. One such idea involved the use of custom mixed-mode SPE cartridges containing C_{18}, SAX, and SCX sorbents in order to retain virtually everything in the biofluid. Given that the behaviour of adsorbed compounds is often different to the behaviour of the same compounds in solution, possibly including ramifications for the stability of the constitutent metabolites. One early attempt at 'stacking' cartridges contaning the three sorbents in series achieved inconclusive and non-replicable results — future efforts must figure out the optimal character and quantities of these sorbents for maximum retention.

Similarly, the use of '2D' SPE did not form a serious area of research — naturally, a second dimension (for example, following a reversed-phase method with an ion-exchange method) would result in greater separation compared to a single elution method. However, the number of possible combinations from running a second dimension on wash and elution on a variety of cartridges would have quickly become unmanageable, and furthermore would not be automatable in one method regardless. Further research into 2D SPE will require strict discipline over which datasets are to be studied, lest the researcher become overwhelmed by a considerable amount of generated data.

Overall, the manual SPE methods developed in this chapter demonstrated significant retentive capacity, with some specific methods able to selectively retain compounds based on their taxonomic class or on specific functional groups. In some cases, the mechanisms for retention were clear; in others, such as with PBA and mannitol/sorbitol, they were less clear. While the basic theory behind solid phase extraction — and chromatography more broadly — is well established, the exact mechanisms and forces determining retention on SPE sorbents is far less understood, especially with lesser studied sorbents like PBA. Further research into these

sorbents would help inform solvent conditions and provide greater control over the retention profiles of each method.

6.3   Metabolite identification and structural elucidation

The manual methods developed in chapter 3 revealed several peaks which were not easily annotated by reference to database entries or literature. The process of annotating and identifying the major revealed peaks required the use of multiple analytical methods, ranging from 2D NMR spectroscopy to LC-MS on further fractionated samples; this was aided in part by the use of gradient SPE.

Gradient SPE — while not currently automatable — proved to be significantly helpful in the identification of several of the unknown peaks detailed during method development. In this instance, a linear gradient from 0% to 100% methanol was used — however, due to the inherent propensity of reversed-phase cartridges to retain significantly hydrophobic compounds, this naturally concentrated most of the metabolites in the latter fractions. The development of non-linear gradients for SPE was not considered during the course of the project due to the fact that the amount of time required to fine-tune such a gradient would be excessive for this context, but in a broader capacity such methods would allow for greater control over the isolation of compounds by class and functional groups.

The most common theme during annotation efforts was the limitations of metabolite databases — even for a highly replicable analytical platform like NMR, spectra in the most used databases were often either missing, incomplete, vulnerable to solvent or pH effects, or even wrong. These databases are vital tools for metabolite annotation and provide quick and simple lookup values for unknown peaks found during metabolic profiling; where gaps in the knowledge of the human metabolome are known, they are reflected in the databases. The importance that database entries be kept up to date — and that researchers make efforts to submit unknowns to these databases — hence cannot be overstated.

Two of the named compounds in this chapter required an additional, orthogonal set of data for structural elucidation and confirmation of identity. While LC-SPE-NMR-MS systems have

been previously reported[118], in this instance the analytical method would perhaps be better described as gSPE (gradient SPE)-NMR-LC-MS. In this configuration, the sample was first reduced to a relatively small number of fractions and analysed by NMR; after unknown peaks are identified, they was then further fractionated and subsequently analysed using LC-MS methods. While this method was not high-throughput in implementation, it proved very effective in identifying the unknowns being targeted; as demonstrated by Whiley *et al.* (2019)[108], it has also proven a valuable technique when annotating large sections of the urine metabolome. The use of fractionation may also prove useful in studying other biofluids, such as tissue extractions.

The holistic aim of the project is to aid metabolite annotation and identification efforts, and consequently to increase understanding of the human urinary metabolome. With this in mind, it's worth noting that despite the successes in annotation and identification described in chapter 4, many of the peaks — often singlets in overcrowded spectra with no attached ^{13}C data — are still without assignments; indeed, it is unclear whether these peaks even correspond to known compounds or otherwise. The future work described for chapter 3, if able to provide greater isolation of metabolites, would expand the profiling 'toolbox' available to researchers, and hence increase the understanding of the metabolome.

Finally, the automated methods developed were then applied to glucuronide-spiked urine; using a 2% formic acid HLB method, we aimed to selectively retain the glucuronide metabolites based on the sugar moiety. In the choice between HLB and PBA, previous experience suggested that PBA would suffer from the same problems demonstrated previously, where compounds would inexplicably fail to be retained. As a result, while the HLB method did successfully retain 4/5 of the metabolites, it is likely that this retention was at least aided by the core, non-glucuronide moeity. Further research — testing different PBA methods on the spiked glucuronide samples, altering the HLB methods to attempt to increase retention capability, and using a broader range of compounds to find the specific and key features of metabolites needed for retention would hopefully lead to more control over retention. Further, if the PBA methods can be successfully adjusted to selectively retain glucuronides, the use of 2D SPE on the reversed-phase elutions could minimise peak overlap even further.

6.4 Development and application of automated SPE methods

The ultimate stage of this project involved the translation of the manual protocols developed in chapter 3 into automated methods for general use. This required learning the commands associated with the visual scripting tool in order to create new routines and subroutines suited to the requirements of the manual methods — most notably including support for ion exchange and 6 mL SPE cartridges. The methods were then successfully validated against the manual methods, demonstrating no significant difference from hand-run SPE but promising greater speed and consistency with minimal effort.

The biggest obstructions to unlocking the full potential of automated SPE are the physical limitations of the machine itself and the commands supported by the software. The use of gradient SPE as described in chapter 4 was a great aid in annotating several unknown spectral peaks, but as gradient SPE is not supported by the automated system its impact is limited. There are some remaining leads which might provide limited support for gradient methods — for example, the automated system currently supports up to four solvents, but the solvent system may be able to be 'tricked' into aspirating two separate solvents in one subroutine. Similarly, the pushthrough subroutine for 6 mL cartridge — currently utilising a nitrogen stream to create backpressure — may have room for streamlining if a better alternative can be found; this would significantly cut down experiment run times, on top of removing the need for a nitrogen stream since the cartridge drying step is ultimately optional.

It should be noted that the system is entirely off-line; while it is unlikely that the machine as it currently exists can be significantly modified to bring it online without disproportionate effort and expenditure, future machine designs might consider in-line elution and wash drying and reconstitution, on top of the NMR sample preparation capabilities which currently exist. However, it is important that any additional features be optional, since the flexibility of SPE is a strength which must be preserved if it is to preserve its position as a useful tool in the toolbox of researchers studying metabolic profiling.

6.5   Conclusion

During the course of this project, we have aimed to expand knowledge of the metabolome. Use of NMR spectroscopy is widespread, since the instrument is able to produce replicable and quantitative structural information about the contents within a complex, multicomponent mixture. The high complexity of biofluids often results in peak overlap and 'drowning out' of signals, which is a hindrance to annotation efforts. By developing untargeted SPE methods aiming to retain metabolites based on their compound class or specific functional groups — which can be run on an automated system with minimal effort. We have aided this process; the use of the developed methods has allowed for the annotation of several metabolites (of which some did not have reliable NMR data recorded), and generated areas for further research and development. Despite the shortcomings of solid phase extraction, we are confident that its uses as demonstrated here should preserve its place within the analytical toolbox for the foreseeable future.

Bibliography

(*1*)   Nicholson, J. K., Lindon, J. C., and Holmes, E. (1999). 'Metabonomics': understanding the metabolic responses of living systems to pathophysiological stimuli via multivariate statistical analysis of biological NMR spectroscopic data. *Xenobiotica 29*, 1181–1189.

(*2*)   Ser, Z., Liu, X., Tang, N. N., and Locasale, J. W. (2015). Extraction parameters for metabolomics from cultured cells. *Analytical biochemistry 475*, 22–8.

(*3*)   Sapcariu, S. C., Kanashova, T., Weindl, D., Ghelfi, J., Dittmar, G., and Hiller, K. (2014). Simultaneous extraction of proteins and metabolites from cells in culture. *MethodsX 1*, 74–80.

(*4*)   Risa, Ø., Melø, T. M., Sonnew, and Ald, U. (2009). Quantification of amounts and 13C content of metabolites in brain tissue using high-resolution magic angle spinning 13C NMR spectroscopy. *NMR in Biomedicine 22*, 266–271.

(*5*)   Huang, W., Serra, O., Dastmalchi, K., Jin, L., Yang, L., and Stark, R. E. (2017). Comprehensive MS and Solid-State NMR Metabolomic Profiling Reveals Molecular Variations in Native Periderms from Four Solanum tuberosum Potato Cultivars. *Journal of Agricultural and Food Chemistry 65*, 2258–2274.

(*6*)   Amann, A., Costello, B. d. L., Miekisch, W., Schubert, J., Buszewski, B., Pleil, J., Ratcliffe, N., and Risby, T. (2014). The human volatilome: volatile organic compounds (VOCs) in exhaled breath, skin emanations, urine, feces and saliva. *Journal of Breath Research 8*, 034001.

(*7*)   Bouatra, S., Aziat, F., Mandal, R., Guo, A. C., Wilson, M. R., Knox, C., Bjorndahl, T. C., Krishnamurthy, R., Saleem, F., Liu, P., Dame, Z. T., Poelzer, J., Huynh, J.,

Yallou, F. S., Psychogios, N., Dong, E., Bogumil, R., Roehring, C., and Wishart, D. S. (2013). The Human Urine Metabolome. *PLoS ONE 8*, ed. by Dzeja, P., e73076.

(*8*)   Nicholson, J. K., Foxall, P. J. D., Spraul, M., Farrant, R. D., and Lindon, J. C. (1995). 750 MHz 1H and 1H-13C NMR Spectroscopy of Human Blood Plasma. *Analytical Chemistry 67*, 793–811.

(*9*)   Hügle, T., Kovacs, H., Heijnen, I. A. F. M., Daikeler, T., Baisch, U., Hicks, J. M., and Valderrabano, V. Synovial fluid metabolomics in different forms of arthritis assessed by nuclear magnetic resonance spectroscopy. *Clinical and experimental rheumatology 30*, 240–5.

(*10*)   Schrimpe-Rutledge, A. C., Codreanu, S. G., Sherrod, S. D., and McLean, J. A. (2016). Untargeted Metabolomics Strategies-Challenges and Emerging Directions. *Journal of the American Society for Mass Spectrometry 27*, 1897–1905.

(*11*)   Cloarec, O., Dumas, M. E., Craig, A., Barton, R. H., Trygg, J., Hudson, J., Blancher, C., Gauguier, D., Lindon, J. C., Holmes, E., and Nicholson, J. (2005). Statistical total correlation spectroscopy: An exploratory approach for latent biomarker identification from metabolic 1H NMR data sets. *Analytical Chemistry 77*, 1282–1289.

(*12*)   Posma, J. M., Garcia-Perez, I., De Iorio, M., Lindon, J. C., Elliott, P., Holmes, E., Ebbels, T. M., and Nicholson, J. K. (2012). Subset optimization by reference matching (STORM): An optimized statistical approach for recovery of metabolic biomarker structural information from 1H NMR spectra of biofluids. *Analytical Chemistry 84*, 10694–10701.

(*13*)   Robinette, S. L., Veselkov, K. A., Bohus, E., Coen, M., Keun, H. C., Ebbels, T. M., Beckonert, O., Holmes, E. C., Lindon, J. C., and Nicholson, J. K. (2009). Cluster analysis statistical spectroscopy using nuclear magnetic resonance generated metabolic data sets from perturbed biological systems. *Analytical Chemistry 81*, 6581–6589.

(*14*)   Crockford, D. J., Holmes, E., Lindon, J. C., Plumb, R. S., Zirah, S., Bruce, S. J., Rainville, P., Stumpf, C. L., and Nicholson, J. K. (2006). Statistical heterospectroscopy,

an approach to the integrated analysis of NMR and UPLC-MS data sets: Application in metabonomic toxicology studies. *Analytical Chemistry 78*, 363–371.

(*15*) Zhang, B., Xie, M., Bruschweiler-Li, L., Bingol, K., and Brüschweiler, R. (2015). Use of Charged Nanoparticles in NMR-Based Metabolomics for Spectral Simplification and Improved Metabolite Identification. *Analytical Chemistry 87*, 7211–7217.

(*16*) Keun, H. C., and Athersuch, T. J. In *Methods in molecular biology (Clifton, N.J.)* 2011; Vol. 708, pp 321–334.

(*17*) Dettmer, K., Aronov, P. A., and Hammock, B. D. (2007). Mass spectrometry-based metabolomics. *Mass Spectrometry Reviews 26*, 51–78.

(*18*) Zhou, B., Xiao, J. F., Tuli, L., and Ressom, H. W. (2012). LC-MS-based metabolomics. *Molecular bioSystems 8*, 470–81.

(*19*) Sumner, L. W., Alexander, A. E., Ae, A., Ae, D. B., Ae, M. H. B., Beger, R., Daykin, C. A., Teresa, A. E., Fan, W.-M., Oliver, A. E., Ae, F., Goodacre, R., Julian, A. E., Griffin, L., Thomas, A. E., Ae, H., Hardy, N., James, A. E., Ae, H., Higashi, R., Joachim, A. E., Ae, K., Ae, A. N. L., Lindon, J. C., Philip, A. E., Ae, M., Ae, A. W. N., Reily, M. D., Ae, J. J. T., Viant, M. R., Sumner, L. W., Amberg, A., Barrett, D., Beale, M. H., Beger, R., Daykin, C. A., Higashi, Á. R., Higashi, R., Fiehn, O., and Goodacre, R. (2007). Proposed minimum reporting standards for chemical analysis Chemical Analysis Working Group (CAWG) Metabolomics Standards Initiative (MSI). *Metabolomics 3*, 211–221.

(*20*) Creek, D. J., Dunn, W. B., Fiehn, O., Griffin, J. L., Hall, R. D., Lei, Z., Mistrik, R., Neumann, S., Schymanski, E. L., Sumner, L. W., Trengove, R., and Wolfender, J.-L. (2014). Metabolite identification: are you sure? And how do your peers gauge your confidence? *Metabolomics 10*, 350–353.

(*21*) Sumner, L. W., Lei, Z., Nikolau, B. J., Saito, K., Roessner, U., and Trengove, R. (2014). Proposed quantitative and alphanumeric metabolite identification metrics. *Metabolomics 10*, 1047–1049.

(*22*)   Everett, J. R. (2015). A New Paradigm for Known Metabolite Identification in Metabo-nomics/Metabolomics: Metabolite Identification Efficiency. *Computational and Structural Biotechnology Journal 13*, 131–144.

(*23*)   Edmands, W. M. B., Beckonert, O. P., Stella, C., Campbell, A., Lake, B. G., Lindon, J. C., Holmes, E., and Gooderham, N. J. (2011). Identification of Human Urinary Biomarkers of Cruciferous Vegetable Consumption by Metabonomic Profiling. *J. Proteome Res 10*, 4513–4521.

(*24*)   Stella, C., Beckwith-Hall, B., Cloarec, O., Holmes, E., Lindon, J. C., Powell, J., Van Der Ouderaa, F., Bingham, S., Cross, A. J., and Nicholson, J. K. (2006). Susceptibility of human metabolic phenotypes to dietary modulation. *Journal of Proteome Research 5*, 2780–2788.

(*25*)   Bingol, K., and Brüschweiler, R. (2015). NMR/MS Translator for the Enhanced Simultaneous Analysis of Metabolomics Mixtures by NMR Spectroscopy and Mass Spectrometry: Application to Human Urine. *Journal of Proteome Research 14*, 2642–2648.

(*26*)   Claridge, T. D., *High-Resolution NMR Techniques in Organic Chemistry: Third Edition*, 2016, pp 1–541.

(*27*)   Mckay, R. T. (2011). How the 1D-NOESY suppresses solvent signal in metabonomics NMR spectroscopy: An examination of the pulse sequence components and evolution. *Concepts in Magnetic Resonance Part A 38A*, 197–220.

(*28*)   Patt, S. L., and Sykes, B. D. (1972). Water Eliminated Fourier Transform NMR Spectroscopy. *The Journal of Chemical Physics 56*, 3182–3184.

(*29*)   Beckonert, O., Keun, H. C., Ebbels, T. M. D., Bundy, J., Holmes, E., Lindon, J. C., and Nicholson, J. K. (2007). Metabolic profiling, metabolomic and metabonomic procedures for NMR spectroscopy of urine, plasma, serum and tissue extracts. *Nature Protocols 2*, 2692–2703.

(*30*)   Ludwig, C., and Viant, M. R. (2010). Two-dimensional J-resolved NMR spectroscopy: review of a key methodology in the metabolomics toolbox. *Phytochemical Analysis 21*, 22–32.

(*31*)   Braunschweiler, L., and Ernst, R. (1983). Coherence transfer by isotropic mixing: Application to proton correlation spectroscopy. *Journal of Magnetic Resonance (1969) 53*, 521–528.

(*32*)   Aue, W. P., Bartholdi, E., and Ernst, R. R. (1976). Two-dimensional spectroscopy. Application to nuclear magnetic resonance. *The Journal of Chemical Physics 64*, 2229–2246.

(*33*)   Bodenhausen, G., and Ruben, D. J. (1980). Natural abundance nitrogen-15 NMR by enhanced heteronuclear spectroscopy. *Chemical Physics Letters 69*, 185–189.

(*34*)   Bax, A., and Summers, M. F. (1986). Proton and carbon-13 assignments from sensitivity-enhanced detection of heteronuclear multiple-bond connectivity by 2D multiple quantum NMR. *Journal of the American Chemical Society 108*, 2093–2094.

(*35*)   Olejniczak, E. T., and Eaton, H. L. (1990). Extrapolation of time-domain data with linear prediction increases resolution and sensitivity. *Journal of Magnetic Resonance (1969) 87*, 628–632.

(*36*)   Parella, T. (1996). High-Quality 1D Spectra by Implementing Pulsed-Field Gradients as the Coherence Pathway Selection Procedure. *Magnetic Resonance in Chemistry 34*, 329–347.

(*37*)   Walker, J. M., and Ralph, R., *Molecular biomethods handbook*; 9, 2009; Vol. 74, pp 1057–1057.

(*38*)   Pearson, K. (1901). LIII. On lines and planes of closest fit to systems of points in space. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science 2*, 559–572.

(*39*)   Skov, T., Honoré, A. H., Jensen, H. M., Næs, T., and Engelsen, S. B. (2014). Chemometrics in foodomics: Handling data structures from multiple analytical platforms. *TrAC Trends in Analytical Chemistry 60*, 71–79.

(*40*) Dona, A. C., Kyriakides, M., Scott, F., Shephard, E. A., Varshavi, D., Veselkov, K., and Everett, J. R. (2016). A guide to the identification of metabolites in NMR-based metabonomics/metabolomics experiments. *Computational and Structural Biotechnology Journal 14*, 135–153.

(*41*) Poole, C. F., and Poole, S. K. In *Solid-Phase Extraction: Principles, Techniques, and Applications*, 2000, pp 183–226.

(*42*) Vervoort, R. J. M., Debets, A. J. J., Claessens, H. A., Cramers, C. A., and De Jong, G. J. (2000). Optimisation and characterisation of silica-based reversed-phase liquid chromatographic systems for the analysis of basic pharmaceuticals. *Journal of Chromatography A 897*, 1–22.

(*43*) Sigma-Aldrich Guide to Solid Phase Extraction `http://www.sigmaaldrich.com/Graphics/Supelco/objects/4600/4538.pdf`.

(*44*) Petrucci, R. H., Herring, F. G., Madura, J. D., and Bissonnette, C., *General Chemistry: Principles and Modern Applications*, 2016, pp 602–664.

(*45*) Privalovl, P. L., and Gill, S. J. (1989). The hydrophobic effect: a reappraisal. *Pure & Appl. Chem 61*, 1097–1104.

(*46*) Sinnokrot, M. O., Valeev, E. F., and Sherrill, C. D. (2002). Estimates of the ab initio limit for π-π interactions: The benzene dimer. *Journal of the American Chemical Society 124*, 10887–10893.

(*47*) Moloney, M. G., *Structure and reactivity in organic chemistry*, p 306.

(*48*) Arunan, E., Desiraju, G. R., Klein, R. A., Sadlej, J., Scheiner, S., Alkorta, I., Clary, D. C., Crabtree, R. H., Dannenberg, J. J., Hobza, P., Kjaergaard, H. G., Legon, A. C., Mennucci, B., and Nesbitt, D. J. (2011). Defining the hydrogen bond: An account (IUPAC Technical Report). *Pure Appl. Chem 83*, 1619–1636.

(*49*) Blahová, E., and Brandšteterová, E. (2004). Approaches in Sample Handling before HPLC Analysis of Complex Matrices. *Analysis 58*, 362–373.

(50) Agilent Phenylboronic Acid (PBA) Solid Phase Extraction Mechanisms and Applications technical overview `http://cn.agilent.com/cs/library/technicaloverviews/public/SI-02442.pdf`.

(51) Roberts, L. D., Souza, A. L., Gerszten, R. E., and Clish, C. B. (2012). Targeted metabolomics. *Current protocols in molecular biology Chapter 30*, 1–24.

(52) Ruane, R., and Wilson, I. (1987). The use of C18 bonded silica in the solid phase extraction of basic drugs - possible role for ionic interactions with residual silanols. *Journal of Pharmaceutical and Biomedical Analysis 5*, 723–727.

(53) Wilson, I. D., and Nicholson, J. K. (1987). Solid-phase extraction chromatography and nuclear magnetic resonance spectrometry for the identification and isolation of drug metabolites in urine. *Analytical Chemistry 59*, 2830–2832.

(54) Wilson, I. D., and Nicholson, J. K. (1988). Solid phase extraction chromatography and NMR spectroscopy (SPEC-NMR) for the rapid identification of drug metabolites in urine. *Journal of Pharmaceutical and Biomedical Analysis 6*, 151–165.

(55) Michopoulos, F., Gika, H., Palachanis, D., Theodoridis, G., and Wilson, I. D. (2015). Solid phase extraction methodology for UPLC-MS based metabolic profiling of urine samples. *Electrophoresis 36*, 2170–2178.

(56) Huang, Z., and Zhang, S. (2003). Confirmation of amphetamine, methamphetamine, MDA and MDMA in urine samples using disk solid-phase extraction and gas chromatography–mass spectrometry after immunoassay screening. *Journal of Chromatography B 792*, 241–247.

(57) Kumazawa, T., Hasegawa, C., Lee, X.-P., Hara, K., Seno, H., Suzuki, O., and Sato, K. (2007). Simultaneous determination of methamphetamine and amphetamine in human urine using pipette tip solid-phase extraction and gas chromatography-mass spectrometry. *Journal of Pharmaceutical and Biomedical Analysis 44*, 602–607.

(58) Hegstad, S., Øiestad, E. L., Johansen, U., and Christophersen, A. S. (2006). Determination of benzodiazepines in human urine using solid-phase extraction and high-

performance liquid chromatography-electrospray ionization tandem mass spectrometry. *Journal of analytical toxicology 30*, 31–7.

(59) Black, D. A., Clark, G. D., Haver, V. M., Garbin, J. A., and Saxon, A. J. (1994). Analysis of urinary benzodiazepines using solid-phase extraction and gas chromatography-mass spectrometry. *Journal of analytical toxicology 18*, 185–8.

(60) Casas, M., Berrueta, L., Gallo, B., and Vicente, F. (1993). Solid-phase extraction of 1,4-benzodiazepines from biological fluids. *Journal of Pharmaceutical and Biomedical Analysis 11*, 277–284.

(61) Borrey, D., Meyer, E., Lambert, W., Van Peteghem, C., and De Leenheer, A. (2001). Simultaneous determination of fifteen low-dosed benzodiazepines in human urine by solid-phase extraction and gas chromatography-mass spectrometry. *Journal of Chromatography B: Biomedical Sciences and Applications 765*, 187–197.

(62) Georga, K. A., Samanidou, V. F., and Papadoyannis, I. N. (2001). Use of novel solid-phase extraction sorbent materials for high-performance liquid chromatography quantitation of caffeine metabolism products methylxanthines and methyluric acids in samples of biological origin. *Journal of Chromatography B: Biomedical Sciences and Applications 759*, 209–218.

(63) Lillsunde, P., and Korte, T. (1991). Comprehensive Drug Screening in Urine Using Solid-Phase Extraction and Combined TLC and GC/MS Identification. *Journal of Analytical Toxicology 15*, 71–81.

(64) Gustafson, R. A., Moolchan, E. T., Barnes, A., Levine, B., and Huestis, M. A. (2003). Validated method for the simultaneous determination of Δ9-tetrahydrocannabinol (THC), 11-hydroxy-THC and 11-nor-9-carboxy-THC in human plasma using solid phase extraction and gas chromatography-mass spectrometry with positive chemical ionization. *Journal of Chromatography B 798*, 145–154.

(65) Teixeira, H., Verstraete, A., Proença, P., Corte-Real, F., Monsanto, P., and Vieira, D. N. (2007). Validated method for the simultaneous determination of Δ9-THC and Δ9-THC-COOH in oral fluid, urine and whole blood using solid-phase extraction and

liquid chromatography-mass spectrometry with electrospray ionization. *Forensic Science International 170*, 148–155.

(*66*)   Taylor, R. W., Jain, N. C., and George, M. P. (1987). Simultaneous Identification of Cocaine and Benzoylecgonine Using Solid Phase Extraction and Gas Chromatography/Mass Spectrometry. *Journal of Analytical Toxicology 11*, 233–234.

(*67*)   De Giovanni, N., and Rossi, S. S. (1994). Simultaneous detection of cocaine and heroin metabolites in urine by solid-phase extraction and gas chromatography-mass spectrometry. *Journal of Chromatography B: Biomedical Sciences and Applications 658*, 69–73.

(*68*)   Wang, W.-L., Darwin, W. D., and Cone, E. J. (1994). Simultaneous assay of cocaine, heroin and metabolites in hair, plasma, saliva and urine by gas chromatography—mass spectrometry. *Journal of Chromatography B: Biomedical Sciences and Applications 660*, 279–290.

(*69*)   Chang, S.-Y., Whigan, D. B., Vachharajani, N. N., and Patel, R. (1997). High-performance liquid chromatographic assay for the quantitation of irbesartan (SR 47436 / BMS-186295) in human plasma and urine. *Journal of Chromatography B 702*, 149–155.

(*70*)   Caro, E., Marcé, R. M., Cormack, P. A., Sherrington, D. C., and Borrull, F. (2006). Novel enrofloxacin imprinted polymer applied to the solid-phase extraction of fluorinated quinolones from urine and tissue samples. *Analytica Chimica Acta 562*, 145–151.

(*71*)   Di, B., Su, M.-X., Yu, F., Qu, L.-J., Zhao, L.-P., Cheng, M.-C., and He, L.-P. (2008). Solid-phase extraction and liquid chromatography/tandem mass spectrometry assay for the determination of pitavastatin in human plasma and urine for application to Phase I clinical pharmacokinetic studies. *Journal of Chromatography B 868*, 95–101.

(*72*)   Zhao, J. J., Xie, I. H., Yang, A. Y., Roadcap, B. A., and Rogers, J. D. (2000). Quantitation of simvastatin and its beta-hydroxy acid in human plasma by liquid-liquid cartridge extraction and liquid chromatography/tandem mass spectrometry. *Journal of Mass Spectrometry 35*, 1133–1143.

(*73*)   Tugnait, M., Ghauri, F., Wilson, I., and Nicholson, J. (1991). NMR-monitored solid-phase extraction of phenolphthalein glucuronide on phenylboronic acid and C18 bonded phases. *Journal of Pharmaceutical and Biomedical Analysis 9*, 895–899.

(*74*)   Logan, B. K., Stafford, D. T., Tebbett, I. R., and Moore, C. M. (1990). Rapid Screening for 100 Basic Drugs and Metabolites in Urine Using Cation Exchange Solid - Phase Extraction and High - Performance Liquid Chromatography with Diode Array Detection. *Journal of Analytical Toxicology 14*, 154–159.

(*75*)   Kato, K., Shoda, S., Takahashi, M., Doi, N., Yoshimura, Y., and Nakazawa, H. (2003). Determination of three phthalate metabolites in human urine using on-line solid-phase extraction–liquid chromatography–tandem mass spectrometry. *Journal of Chromatography B 788*, 407–411.

(*76*)   Silva, C. L., Passos, M., and Câmara, J. S. (2011). Investigation of urinary volatile organic metabolites as potential cancer biomarkers by solid-phase microextraction in combination with gas chromatography-mass spectrometry. *British Journal of Cancer 105*, 1894–1904.

(*77*)   Gonzalo-Lumbreras, R., Pimentel-Trapero, D., and Izqierdo-Hornillos, R. (2001). Solvent and solid-phase extraction of natural and synthetic anabolic steroids in human urine. *Journal of Chromatography B: Biomedical Sciences and Applications 754*, 419–425.

(*78*)   Cho, Y.-D., and Choi, M.-H. (2006). Alternative Sample Preparation Techniques in Gas Chromatographic-Mass Spectrometric Analysis of Urinary Androgenic Steroids. *Bulletin of the Korean Chemical Society 27*, 1315–1322.

(*79*)   Guo, F., Shao, J., Liu, Q., Shi, J.-b., and Jiang, G.-b. (2014). Automated and sensitive determination of four anabolic androgenic steroids in urine by online turbulent flow solid-phase extraction coupled with liquid chromatography–tandem mass spectrometry: A novel approach for clinical monitoring and doping control. *Talanta 125*, 432–438.

(*80*)   Kumazawa, T., Seno, H., Lee, X.-P., Ishii, A., Watanabe-Suzuki, K., Sato, K., and Suzuki, O. (1999). Extraction of methylxanthines from human body fluids by solid-phase microextraction. *Analytica Chimica Acta 387*, 53–60.

(*81*) Posma, J. M., Garcia-Perez, I., Heaton, J. C., Burdisso, P., Mathers, J. C., Draper, J., Lewis, M., Lindon, J. C., Frost, G., Holmes, E., and Nicholson, J. K. (2017). Integrated Analytical and Statistical Two-Dimensional Spectroscopy Strategy for Metabolite Identification: Application to Dietary Biomarkers. *Analytical Chemistry 89*, 3300–3309.

(*82*) Yang, W. J., Wang, Y. W., Zhou, Q. F., and Tang, H. R. (2008). Analysis of human urine metabolites using SPE and NMR spectroscopy. *Science in China, Series B: Chemistry 51*, 218–225.

(*83*) Jacobs, D. M., Spiesser, L., Garnier, M., De Roo, N., Van Dorsten, F., Hollebrands, B., Van Velzen, E., Draijer, R., and Van Duynhoven, J. (2012). SPE-NMR metabolite sub-profiling of urine. *Analytical and Bioanalytical Chemistry 404*, 2349–2361.

(*84*) de Hoffman, E., and Stroobant, V., *Mass Spectrometry: Principles and Applications*, 2007.

(*85*) Bielicka-Daszkiewicz, K., and Voelkel, A. (2009). Theoretical and experimental methods of determination of the breakthrough volume of SPE sorbents. *Talanta 80*, 614–621.

(*86*) Hennio, M.-C. (1999). Solid-phase extraction: method development, sorbents, and coupling with liquid chromatography. *Journal of Chromatography A 856*, 3–54.

(*87*) Gelencsér, A., Kiss, G., Krivácsy, Z., Varga-Puchony, Z., and Hlavay, J. (1995). A simple method for the determination of capacity factor on solid-phase extraction cartridges. I. *Journal of Chromatography A 693*, 217–225.

(*88*) Dona, A. C., Jiménez, B., Schäfer, H., Humpfer, E., Spraul, M., Lewis, M. R., Pearce, J. T. M., Holmes, E., Lindon, J. C., and Nicholson, J. K. (2014). Precision High-Throughput Proton NMR Spectroscopy of Human Urine, Serum, and Plasma for Large-Scale Metabolic Phenotyping. *Analytical Chemistry 86*, 9887–9894.

(*89*) Wishart, D. S., Tzur, D., Knox, C., Eisner, R., Guo, A. C., Young, N., Cheng, D., Jewell, K., Arndt, D., Sawhney, S., Fung, C., Nikolai, L., Lewis, M., Coutouly, M. A., Forsythe, I., Tang, P., Shrivastava, S., Jeroncic, K., Stothard, P., Amegbey, G., Block, D., Hau, D. D., Wagner, J., Miniaci, J., Clements, M., Gebremedhin, M., Guo, N., Zhang, Y., Duggan, G. E., MacInnis, G. D., Weljie, A. M., Dowlatabadi, R., Bamforth, F., Clive, D.,

Greiner, R., Li, L., Marrie, T., Sykes, B. D., Vogel, H. J., and Querengesser, L. (2007). HMDB: The human metabolome database. *Nucleic Acids Research 35*, D521–D526.

(*90*) Djoumbou Feunang, Y., Eisner, R., Knox, C., Chepelev, L., Hastings, J., Owen, G., Fahy, E., Steinbeck, C., Subramanian, S., Bolton, E., Greiner, R., and Wishart, D. S. (2016). ClassyFire: automated chemical classification with a comprehensive, computable taxonomy. *Journal of Cheminformatics 8*, 61.

(*91*) Waters Oasis Care and Use Manual `https://www.waters.com/webassets/cms/category/docs/720002816EN%20care%20and%20use%20manual.pdf`.

(*92*) Takis, P. G., Schäfer, H., Spraul, M., and Luchinat, C. (2017). Deconvoluting inter-relationships between concentrations and chemical shifts in urine provides a powerful analysis tool. *Nature Communications 8*, 1662.

(*93*) Imperial Metabolic Profiling and Chemometrics Toolbox (IMPaCTS) `http://doi.org/10.5281/zenodo.803330`.

(*94*) Dieterle, F., Ross, A., Schlotterbeck, G., and Senn, H. (2006). Probabilistic quotient normalization as robust method to account for dilution of complex biological mixtures. Application in 1H NMR metabonomics. *Analytical Chemistry 78*, 4281–4290.

(*95*) Barton, W., Penney, N. C., Cronin, O., Garcia-Perez, I., Molloy, M. G., Holmes, E., Shanahan, F., Cotter, P. D., and O'Sullivan, O. (2018). The microbiome of professional athletes differs from that of more sedentary subjects in composition and particularly at the functional metabolic level. *Gut 67*, 625–633.

(*96*) Marcé, R., and Borrull, F. (2000). Solid-phase extraction of polycyclic aromatic compounds. *Journal of Chromatography A 885*, 273–290.

(*97*) Hennion, M.-C. (1999). Solid-phase extraction: method development, sorbents, and coupling with liquid chromatography. *Journal of Chromatography A 856*, 3–54.

(*98*) Martin, P., Leadbetter, B., and Wilson, I. (1993). Immobilized phenylboronic acids for the selective extraction of β-blocking drugs from aqueous solution and plasma. *Journal of Pharmaceutical and Biomedical Analysis 11*, 307–312.

(*99*) Godejohann, M., Preiss, A., Mügge, C., and Wünsch, G. (1997). Application of On-Line HPLC-1H NMR to Environmental Samples: Analysis of Groundwater near Former Ammunition Plants. *Analytical Chemistry 69*, 3832–3837.

(*100*) Griffiths, L., and Horton, R. (1998). Optimization of LC-NMR. III-Increased signal-to-noise ratio through column trapping. *Magnetic Resonance in Chemistry 36*, 104–109.

(*101*) Nyberg, N. T., Baumann, H., and Kenne, L. (2001). Application of solid-phase extraction coupled to an NMR flow-probe in the analysis of HPLC fractions. *Magnetic Resonance in Chemistry 39*, 236–240.

(*102*) Pettersen, J. E., and Jellum, E. (1972). The identification and metabolic origin of 2-furoylglycine and 2,5-furandicarboxylic acid in human urine. *Clinica Chimica Acta 41*, 199–207.

(*103*) Wong, P., Bachki, A., Banerjee, K., and Leyland-Jones, B. (2002). Identification of N1-methyl-2-pyridone-5-carboxamide and N1-methyl-4-pyridone-5-carboxamide as components in urine extracts of individuals consuming coffee. *Journal of Pharmaceutical and Biomedical Analysis 30*, 773–780.

(*104*) Posada-Ayala, M., Zubiri, I., Martin-Lorenzo, M., Sanz-Maroto, A., Molero, D., Gonzalez-Calero, L., Fernandez-Fernandez, B., de la Cuesta, F., Laborde, C. M., Barderas, M. G., Ortiz, A., Vivanco, F., and Alvarez-Llamas, G. (2014). Identification of a urine metabolomic signature in patients with advanced-stage chronic kidney disease. *Kidney International 85*, 103–111.

(*105*) Vanholder, R., De Smet, R., and Lesaffer, G. (1999). p-Cresol: a toxin revealing many neglected but relevant aspects of uraemic toxicity. *Nephrology Dialysis Transplantation 14*, 2813–2815.

(*106*) Elliott, P., Vergnaud, A.-C., Singh, D., Neasham, D., Spear, J., and Heard, A. (2014). The Airwave Health Monitoring Study of police officers and staff in Great Britain: Rationale, design and methods. *Environmental Research 134*, 280–285.

(*107*) Wishart, D. S., Knox, C., Guo, A. C., Eisner, R., Young, N., Gautam, B., Hau, D. D., Psychogios, N., Dong, E., Bouatra, S., Mandal, R., Sinelnikov, I., Xia, J., Jia, L., Cruz, J. A., Lim, E., Sobsey, C. A., Shrivastava, S., Huang, P., Liu, P., Fang, L., Peng, J., Fradette, R., Cheng, D., Tzur, D., Clements, M., Lewis, A., de souza, A., Zuniga, A., Dawe, M., Xiong, Y., Clive, D., Greiner, R., Nazyrova, A., Shaykhutdinov, R., Li, L., Vogel, H. J., and Forsythei, I. (2009). HMDB: A knowledgebase for the human metabolome. *Nucleic Acids Research 37*, 603–610.

(*108*) Whiley, L., Chekmeneva, E., Berry, D. J., Jiménez, B., Yuen, A. H. Y., Salam, A., Hussain, H., Witt, M., Takats, Z., Nicholson, J., and Lewis, M. R. (2019). Systematic Isolation and Structure Elucidation of Urinary Metabolites Optimized for the Analytical-Scale Molecular Profiling Laboratory. *Analytical Chemistry 91*, 8873–8882.

(*109*) Lewis, M. R., Pearce, J. T. M., Spagou, K., Green, M., Dona, A. C., Yuen, A. H. Y., David, M., Berry, D. J., Chappell, K., Horneffer-van der Sluis, V., Shaw, R., Lovestone, S., Elliott, P., Shockcor, J., Lindon, J. C., Cloarec, O., Takats, Z., Holmes, E., and Nicholson, J. K. (2016). Development and Application of Ultra-Performance Liquid Chromatography-TOF MS for Precision Large Scale Urinary Metabolic Phenotyping. *Analytical Chemistry 88*, 9004–9013.

(*110*) Lelo, A., Miners, J. O., Robson, R. A., and Birkett, D. J. (1986). Quantitative assessment of caffeine partial clearances in man. *British journal of clinical pharmacology 22*, 183–6.

(*111*) Rechner, A. R., Kuhnle, G., Hu, H., Roedig-Penman, A., van den Braak, M. H., Moore, K. P., and Rice-Evans, C. A. (2002). The Metabolism of Dietary Polyphenols and the Relevance to Circulating Levels of Conjugated Metabolites. *Free Radical Research 36*, 1229–1241.

(*112*) Gryp, T., Vanholder, R., Vaneechoutte, M., and Glorieux, G. (2017). P-Cresyl Sulfate. *Toxins 9*, 52.

(*113*) Kroemer, H. K., and Klotz, U. (1992). Glucuronidation of Drugs. *Clinical Pharmacokinetics 23*, 292–310.

(114) Van Dorp, E. L., Romberg, R., Sarton, E., Bovill, J. G., and Dahan, A. (2006). Morphine-6-glucuronide: Morphine's successor for postoperative pain relief? *Anesthesia and Analgesia 102*, 1789–1797.

(115) Sundaram, S. S., Bove, K. E., Lovell, M. A., and Sokol, R. J. (2008). Mechanisms of disease: Inborn errors of bile acid synthesis. *Nature clinical practice. Gastroenterology & hepatology 5*, 456–68.

(116) Shipkova, M., Armstrong, V. W., Oellerich, M., and Wieland, E. (2003). Acyl Glucuronide Drug Metabolites: Toxicological and Analytical Implications. *Therapeutic Drug Monitoring 25*, 1–16.

(117) Kim, S., Lee, M., Yoon, D., Lee, D.-K., Choi, H.-J., and Kim, S. (2013). 1D Proton NMR Spectroscopic Determination of Ethanol and Ethyl Glucuronide in Human Urine. *Bulletin of the Korean Chemical Society 34*, 2413–2418.

(118) Schlotterbeck, G., and Ceccarelli, S. M. (2009). LC-SPE-NMR-MS: A total analysis system for bioanalysis. *Bioanalysis 1*, 549–559.

Appendix A

Faraday Discussions article

Application of novel solid phase extraction-NMR protocols for metabolic profiling of human urine

Daniel McGill,[*a] Elena Chekmeneva,[a] John C. Lindon,[a] Zoltan Takats,[a] and Jeremy Nicholson,[*b]

[*] Corresponding authors

[a] Division of Computational and Systems Medicine, Department of Surgery and Cancer, Faculty of Medicine, Imperial College London, South Kensington Campus, London SW7 2AZ, UK E-mail: d.mcgill16@imperial.ac.uk

[b] The Australian National Phenome Center, Research and Innovation, Murdoch University, 50 South Street Murdoch, Perth WA6150, Australia

ABSTRACT

Metabolite identification and annotation procedures are necessary for the discovery of biomarkers indicative of phenotypes or disease states, but these processes can be bottlenecked by the sheer complexity of biofluids containing thousands of different compounds. Here we describe low-cost novel SPE-NMR protocols utilising different cartridges and conditions, on both natural and artificial urine mixtures, which produce unique retention profiles useful for metabolic profiling. We find that different SPE methods applied to biofluids such as urine can be used to selectively retain metabolites based on compound taxonomy or other key functional groups, reducing peak overlap through concentration and fractionation of unknowns and hence promising greater control over the metabolite annotation/identification process.

Appendix B

Use of the SamplePro Solid Phase Extraction (SPE) Robot SOP

B.1   Purpose

This Standard Operating Procedure (SOP) covers the methods for targeted and untargeted solid phase extraction (SPE) utilising the Bruker SamplePro SPE system.

The aim of this SOP is to allow individuals to use the SPE robot in an efficient manner for targeted and untargeted treatment of samples using reversed phase and ion exchange cartridges. The robot is generally safe as any movement of the device is prohibited while the safety guards are not aligned. However, misuse of the robot by an untrained individual may lead to delays or even damage to the machine. This SOP hence aims to minimise these accidents.

B.2   Scope

The Bruker SamplePro SPE system can be utilised to produce SPE-treated samples, which can be subject to further analytical methods such as NMR or MS.

B.3   Materials

- Appropriate solvents for the method – commonly water (deionised) and methanol or acetonitrile.

- 6mL or 3mL capacity SPE cartridges

  - Size-appropriate gilson caps

- Glass vials

  - 25mL sample vials

  - 12mL wash vials

  - 5mL sample elution vials

B.4   Procedures

B.4.1   Setup

To be done before running any SPE method

1. The well plates and glass vials should be situated in the robot as shown in the figure below (Fig. B.1). Ensure that the glass vials are empty. SPE cartridges can be reused if desired, but must have a Gilson cap attached, and must be fixed to the SPE cartridge rack.
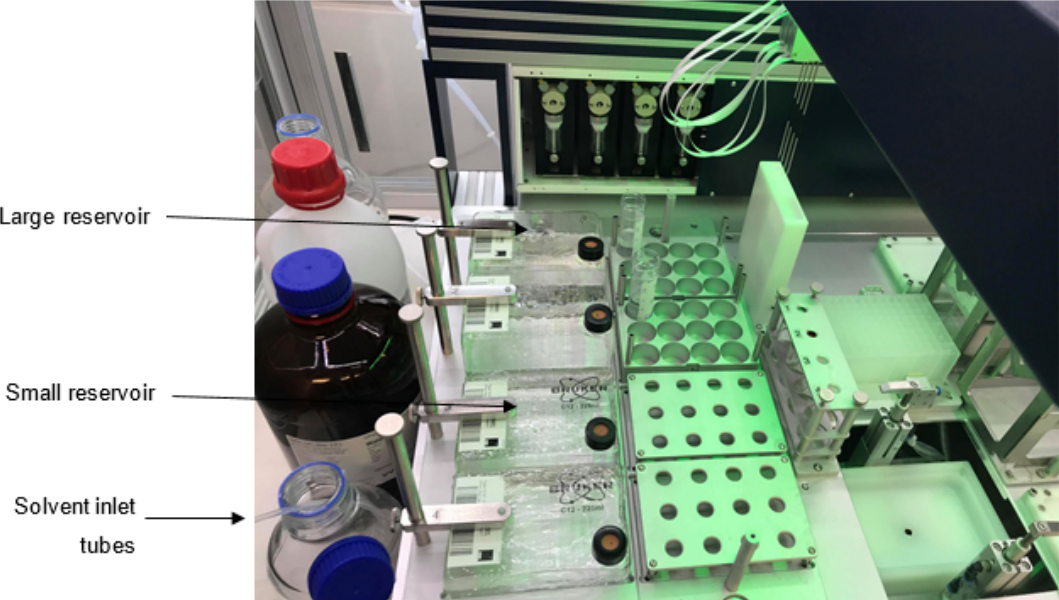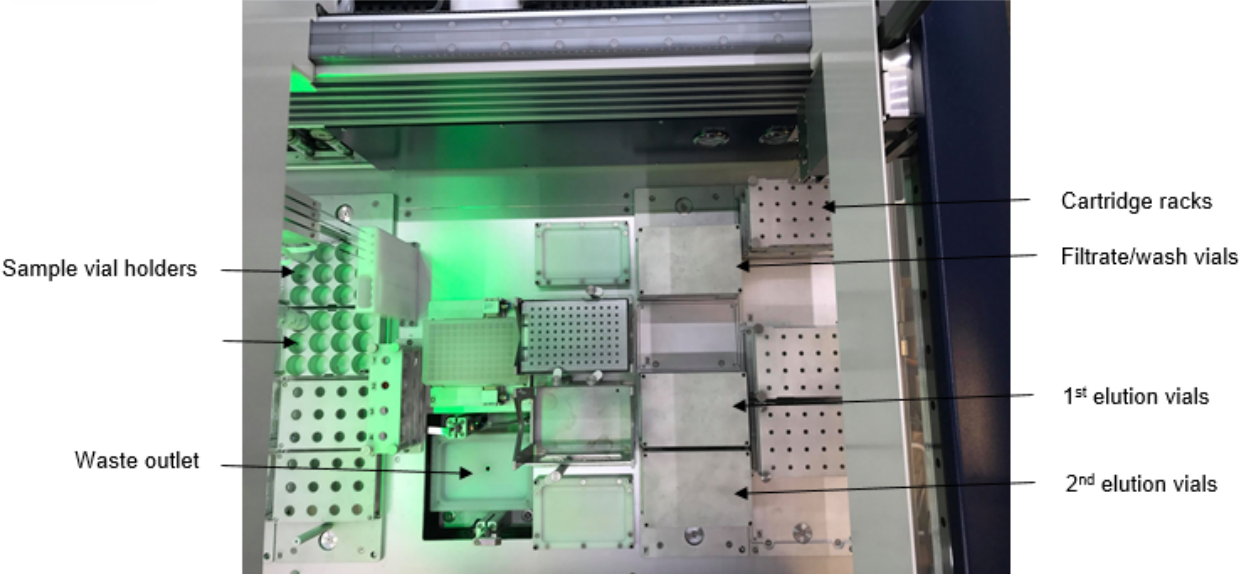
Figure B.1: SPE robot layout.

For 3 mL cartridge methods, set up the instrument as below (fig B.2)
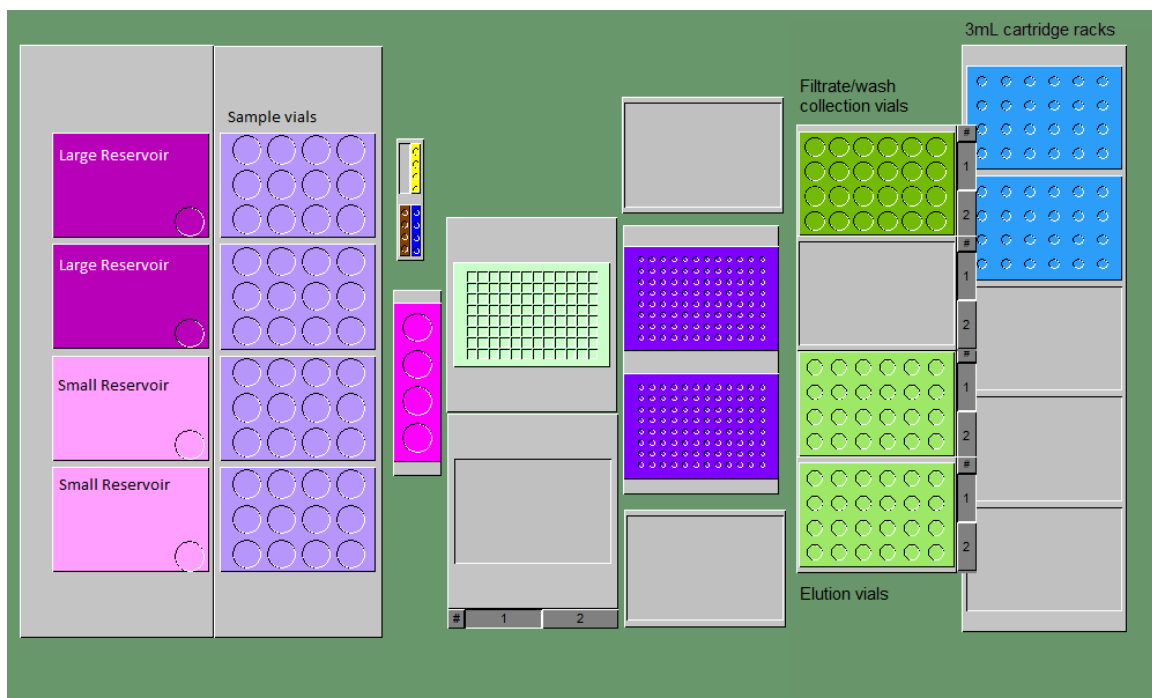

Figure B.2: 3 mL cartridge experiment layout.

For 6mL cartridge methods, the racks should be replaced with their 6mL equivalents, as below (fig 4):
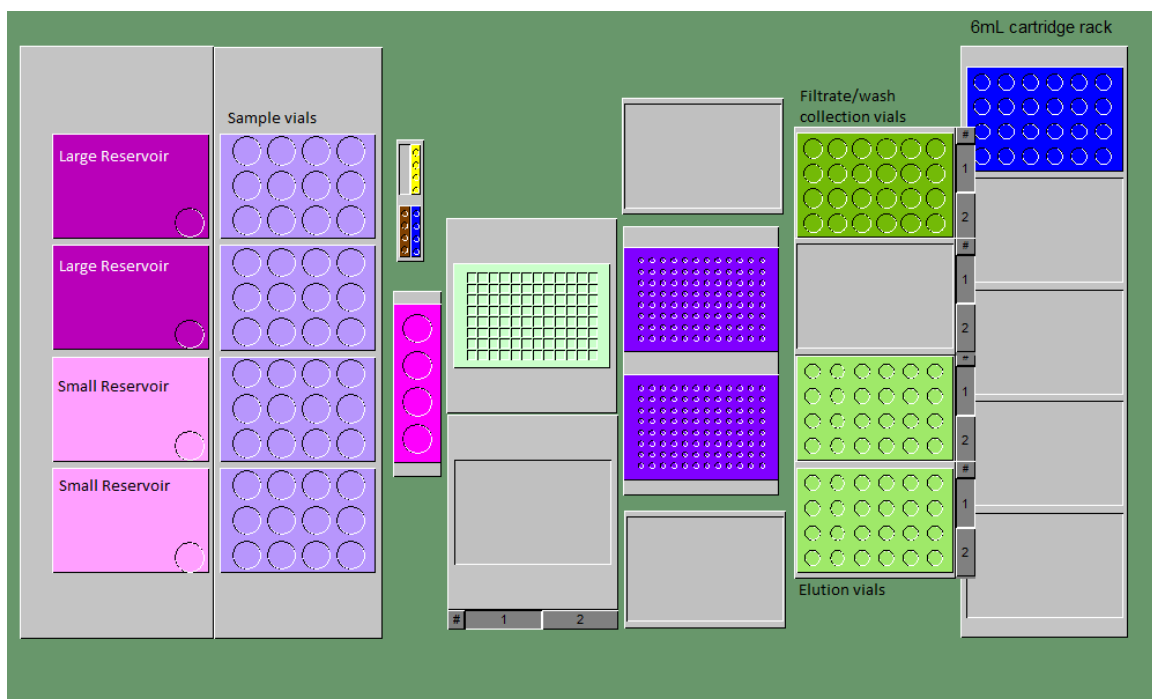

Figure B.3: 6 mL cartridge experiment layout.

Adding samples to the machine

The machine has a specific sample loading pattern which is necessary to understand for successful experiments.

- A **sample** is one vial or container, placed in the appropriate position in the sample vial holders. There must always be an even number of samples such that the two sample vial holders are 'balanced'; if you have an odd number of samples, it will be necessary to also run one blank of water.

- During operation, the robot will run two samples simultaneously — one from each holder — in one **sample load**. If you have only one sample, there **MUST** be a blank in the corresponding opposite holder (see fig. B.4). Generally speaking, the number of sample loads is equal to the total number of samples (including blanks) divided by 2.

The sample vials should have a minimum of 1mL and a maximum of 20mL per vial. When setting up samples for experiments, you may wish to add blanks (water) in the first positions in order to measure any possible impurities and/or contaminants from the process. Regardless, racks must be filled sequentially with samples as follows (fig. B.4):
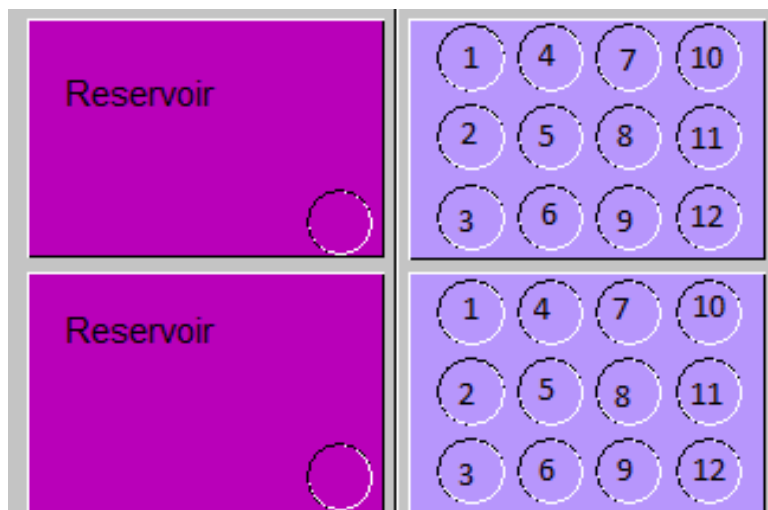

Figure B.4: Sample vial loading order.

An equivalent number of empty filtrate/wash and elution vials should be added to their respective holders. SPE cartridges can be interchanged from the rack by unscrewing the four screws and replacing cartridges (with attached Gilson caps) where necessary.

Step description

Each step causes the robot to perform a unique action (for 6mL SPE experiments, each is followed by the use of nitrogen to 'push through' the liquid in the cartridge). These steps are described below:

- Conditioning: The robot picks up the cartridge rack and moves it above the waste outlet. It then injects solvent from tubes 3 and 4 onto the first two cartridges.

- Equilibration: With the cartridge rack remaining above the waste outlet, the robot will inject solvent from tubes 1 and 2 onto the first two cartridges. The cartridge rack is then replaced in its original position.

- Load: If the experiment is set to collect filtrate, the wash vials will be moved and the cartridges placed above them. The robot will then inject the samples onto the cartridge. The wash vials and the cartridge rack are then replaced. If filtrate is not being collected, the cartridge rack will simply be placed above the waste outlet.

- Wash: The wash vials and cartridge rack are placed as before. The cartridges are injected with solvent from tubes 1 and 2 onto the first two cartridges. Both are then replaced.

- Drying: The cartridge rack is placed above the waste outlet and pressurized nitrogen is passed over the used cartridges. This step is optional but helps to eliminate remaining water, speeding up the drying-down process.

- Elution 1: The first elution vials are placed with the cartridge rack above them, and solvent from tubes 3 and 4 are injected onto the cartridges. Both are then replaced.

- Elution 2: The second elution vials are placed with the cartridge rack above them, and solvent from the two small reservoirs are injected onto the cartridges. Both are then replaced.

For a 6mL experiment with 10 mins of drying time, in all the process will take roughly one hour per sample load.

B.4.2   Targeted SPE

Runs and collects elutions for each sample. Targeted SPE methods tend to be used when a compound is known to be present in a sample, and aim to selectively extract that compound in as great a concentration as possible.

1. The SPE instrument should be set up as described in **Setup**. Solvent tubes 1 and 2 will be used for the equilibration and wash steps, while tubes 3 and 4 will be used in the conditioning and elution 1 steps; the tubes should hence be inserted into a bottle of appropriate solvent. Elution 2 solvent comes from the small reservoirs (fig B.1). See table for suggested solvents:

Table B.1: Targeted SPE recommended solvents.

-	Conditioning	Equilibration/Wash	Elution 1	Elution 2
RP	Methanol	Water (deionised)	Methanol	-
SCX	Methanol	2% formic acid solution	Methanol	5% NH_4OH in MeOH
SAX	Acetonitrile	5% NH_4OH solution	Acetonitrile	2% formic acid in ACN

2. On the PC desktop, open **Method runner**.

3. Select **Tools**.

4. Select **Flush system**, ensure all four lines are selected and that the number of flushes is set to 4 — this ensures that the lines are completely flushed through, and all air is removed. If there are still bubbles, repeat the flush until the lines are filled with solvent.

5. After the system has flushed, double click on double click on **TargetedMethod** for methods utilizing 6 mL capacity SPE cartridges or **TargetedMethod_3mL** for methods utilizing 3 mL capacity SPE cartridges, as appropriate. Press **Next** and fill out the parameter set (fig B.5) as desired.

Figure B.5: Parameters for targeted SPE method.

- *Number of sample loads*: The number of sample loads the program will run. Remember: samples are run in duplicate, so your number of loads is equal to the number of samples (including blanks) divided by 2. Minimum 1, maximum 12.

- *Conditioning volume* [µL]: Amount of solvent (methanol) to use in the conditioning step. Recommended/default 5000 µL, maximum 10000 µL.

- *Equilibration volume* [µL]: Amount of solvent (water) to use in the equilibration step. Recommended/default 5000 µL, maximum 10000 µL.

- *Sample load volume* [µL]: Amount of sample to load onto the cartridge. For targeted SPE, the aim is to load as much of your target compound onto the cartridge as possible without wasting sample. For 6 mL cartridges, suggested volumes for targeted SPE is 10000 µL. Default 10000 µL.

- *Collect sample load filtrate?*: If set to 1, the robot will collect the filtrate which runs out from the cartridge while loading the sample – this can be useful to observe breakthrough and potential contaminants from the cartridge. If set to 0, the filtrate will go to waste. Default '0'.

- *Wash volume* [µL]: Amount of solvent (water) to use in the wash step. This will go to waste. Recommended/default 5000 µL, maximum 10000 µL.

- *Cartridge drying time* [min]: Number of minutes to dry the sorbent bed with nitrogen, eliminating water (and improving drying down time) in the elution. Recommended/default 15 mins.

- Elution 1 volume [µL]: Amount of solvent (methanol) to use in elution step. This will be collected in vials. Recommended/default 5000 µL, max 5000 µL.

- Elution 2 volume [µL]: Amount of solvent to use in the second elution step. This will be collected in vials. Default 0 µL, max 5000 µL.

6. Press **OK**. The machine LEDs will turn white and begin to run your samples. Should the machine error at any point through the run, the light will turn red and the run will need to be aborted. After resolution of the error, the run will need to be manually set up again — see **Resuming an interrupted run**.

7. Once the robot has finished the preparation the light should turn green. A system flush should be carried out as described in step 4 to clean the lines.

8. Remove the elutions from the rack. These can be transferred to falcon tubes and dried in the nitrogen dryer to remove the solvents.

9. Remove the SPE cartridges and samples from the machine. Take off the Gilson caps from the cartridges and set them aside. The cartridges can be kept or discarded as appropriate.

B.4.3   Untargeted SPE

Runs and collects both wash and elution for each sample. Untargeted SPE methods tend to be used when trying to extract a compound or compound class from a sample in order to aid metabolite identification and structural elucidation efforts.

1. The SPE instrument should be set up as described in **Setup**. Solvent tubes 1 and 2 will be used for the equilibration and wash steps, while tubes 3 and 4 will be used in the conditioning and elution 1 steps; the tubes should hence be inserted into a bottle of appropriate solvent. Elution 2 solvent comes from the small reservoirs (fig B.2). See table for suggested solvents:

Table B.2: Untargeted SPE recommended solvents.

-	Conditioning	Equilibration/Wash	Elution 1	Elution 2
RP	Methanol	Water (deionised)	Methanol	-
SCX	Methanol	2% formic acid solution	Methanol	5% NH_4OH in MeOH
SAX	Acetonitrile	5% NH_4OH solution	Acetonitrile	2% formic acid in ACN

2. On the PC desktop, open **Method runner**.

3. Select **Tools**.

4. Select **Flush system**, ensure all four lines are selected and that the number of flushes is set to 4 — this ensures that the lines are completely flushed through, and all air is removed. If there are still bubbles, repeat the flush until the lines are filled with solvent.

5. After the system has flushed, double click on double click on **UntargetedMethod** for methods utilizing 6 mL capacity SPE cartridges or **UntargetedMethod_3mL** for methods utilizing 3 mL capacity SPE cartridges, as appropriate. Press **Next** and fill out the parameter set (fig B.5) as desired.

Figure B.6: Parameters for untargeted SPE method.

- *Number of sample loads*: The number of sample loads the program will run. Remember: samples are run in duplicate, so your number of loads is equal to the number of samples (including blanks) divided by 2. Minimum 1, maximum 12.

- *Conditioning volume* [µL]: Amount of solvent (methanol) to use in the conditioning step. Recommended/default 5000 µL, maximum 10000 µL.

- *Equilibration volume* [µL]: Amount of solvent (water) to use in the equilibration step. Recommended/default 5000 µL, maximum 10000 µL.

- *Sample load volume* [µL]: Amount of sample to load onto the cartridge. For untargeted SPE, there is an antagonism between higher sample volumes leading to better concentration (and hence better resolution), and lower sample volumes leading to less breakthrough of compounds. For 6mL cartridges, suggested volumes for untargeted SPE is 3000 µL. Recommended/default 3000 µL.

- *Collect sample load filtrate?*: If set to 1, the robot will collect the filtrate which runs out from the cartridge while loading the sample – this will be mixed into the wash vial. If set to 0, the filtrate will go to waste. Default '1'.

189

- *Wash volume* [µL]: Amount of solvent (water) to use in the wash step. This will go to waste. Recommended/default 5000 µL, maximum 10000 µL.

- *Cartridge drying time* [min]: Number of minutes to dry the sorbent bed with nitrogen, eliminating water (and improving drying down time) in the elution. Recommended minimum 1 minute in order to push through remaining water from wash step. Recommended/default 15 mins.

- Elution 1 volume [µL]: Amount of solvent (methanol) to use in elution step. This will be collected in vials. Recommended/default 5000 µL, max 5000 µL.

- Elution 2 volume [µL]: Amount of solvent to use in the second elution step, if desired. This will be collected in vials. Default 0 µL, max 5000 µL.

6. Press **OK**. The machine LEDs will turn white and begin to run your samples. Should the machine error at any point through the run, the light will turn red and the run will need to be aborted. After resolution of the error, the run will need to be manually set up again — see **Resuming an interrupted run**.

7. Once the robot has finished the preparation the light should turn green. A system flush should be carried out as described in step 4 to clean the lines.

8. Remove the elutions from the rack. These can be transferred to falcon tubes and dried in the nitrogen dryer to remove the solvents.

9. Remove the SPE cartridges and samples from the machine. Take off the Gilson caps from the cartridges and set them aside. The cartridges can be kept or discarded as appropriate.

B.5 Troubleshooting

B.5.1 The gripper attempts to move into a space already occupied by an object

1. Note which step the operation failed at. Press **Abort** to end the run.

2. Run the **Reset** method.

3. Replace the racks in their default positions, as detailed in **Setup**.

4. If the run failed before the conditioning step, follow the SOP for your method as previously described in **Procedures**. If the run failed after the conditioning step, see **Resuming an interrupted run**.

B.5.2 Resuming an interrupted run

If the method fails at the beginning of a run, resume the method from the start as appropriate. If the method is interrupted partway through a run, follow these instructions to resolve the issue.

1. Resolve the error on screen by pressing **Abort** to end the run. Close the **Runner**.

2. Open **Designer** from the desktop.

3. Expand the '**Methods**' folder. Open the method you were running by double clicking.

4. Select the first line containing the steps you want to skip with a left click. While holding shift, click the last line of the step you want to skip, then right click. In the below example (fig B.7), the conditioning and equilibration steps of the method 'MasterMethod' will be skipped.
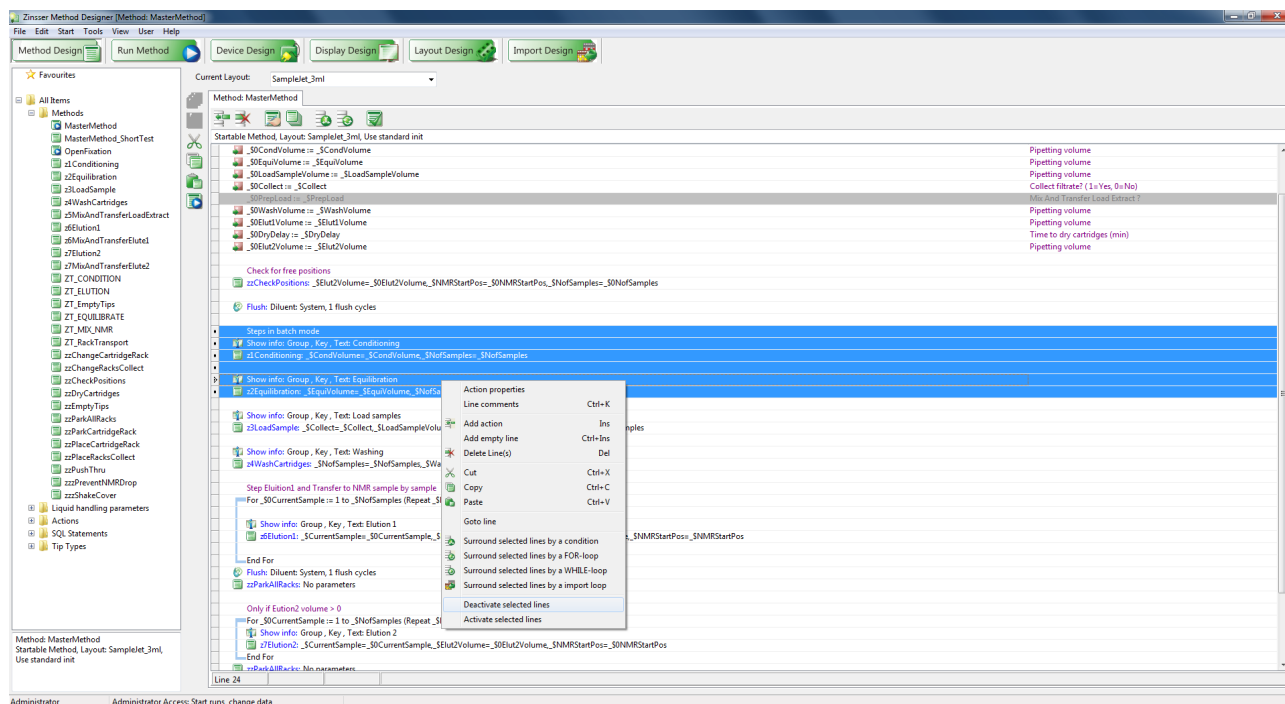
Figure B.7: Muting specific actions in the SPE method designer.

5. Click **Deactivate selected lines**, then click the **Save** icon.

6. Close the **Designer** and open the **Runner**. Double click the method and fill out the parameter set as before. Click **Ok** to start the run.

7. At the end of the run, open the **Designer**, open the edited method, re-select the deactivated lines as before, then right click and select **Activate selected lines**. Then click the **Save** icon and close the **Designer**.

B.5.3 The waste bins are full

Waste solvent from the experiments collect inside solvent bins inside the central cupboard of the robot, locked with a key. This solvent should be emptied according to section disposal protocols. After emptying, the container can be replaced with the nozzle from the waste outlet placed inside the bin.

'What did we learn, Palmer?'

'I don't know, sir.'

'I don't f—ing know either. I guess we learned not to do it again.'

'Yes, sir.'

'F—ed if I know what we did.'

'Yes, sir, it's hard to say.'

- Burn After Reading (2008)