UNIVERSITY OF
LIVERPOOL

Biochemical Investigations in the Rare Disease Alkaptonuria: Studies on the Metabolome and the Nature of Ochronotic Pigment

Thesis submitted in accordance with the requirements of the University of Liverpool for the degree of Doctor of Philosophy
by **Brendan Paul Norman**

September 2019

ACKNOWLEDGEMENTS

ABSTRACT

Alkaptonuria (AKU) is a rare inherited disorder of tyrosine metabolism caused by lack of the enzyme homogentisate 1,2-dioxygenase (HGD). The primary biochemical consequence of HGD deficiency is increased circulating concentration of homogentisic acid (HGA); this is the central cause of the devastating multi-systemic damage observed in AKU. One of the most striking pathophysiological features of AKU is accumulation of ochronotic pigment derived from HGA in tissues throughout the body. HGA has an affinity for cartilaginous tissue. Presence of ochronotic pigment in cartilage of load-bearing joints causes severe early-onset osteoarthropathy; the greatest cause of morbidity in AKU.

This thesis addresses two major unanswered questions in AKU. First is the wider metabolic consequences both of HGD deficiency and those observed following treatment with the promising HGA-lowering agent nitisinone. A metabolomics approach was employed to address these questions, with the aim of discovering biomarkers with potential clinical value in monitoring disease in AKU and in assessing the wider metabolic consequences of nitisinone treatment. Second, this thesis explores the nature of ochronotic pigment derived from HGA. The chemical structure of ochronotic pigment is not known, nor are the mechanisms by which HGA interacts with and binds into the cartilage extracellular matrix. Various analytical chemistry techniques were employed to study the chemical properties of ochronotic pigment and pigmented cartilage in patients and mice with AKU.

The data from chemical analyses presented in this thesis provide new insights into the nature of ochronotic pigment. The chemical structure of ochronotic pigment remains unknown, but the pigment had long been widely cited as a polymer in the literature. First, these data fundamentally challenge the idea that ochronotic pigment is a polymeric species. Through study of ochronotic tissue samples, the data also indicate for the first time a specific alteration to the cartilage matrix that enables binding of HGA-derived species.

An LC-QTOF-MS metabolic phenotyping strategy was developed, enabling identification of hundreds of metabolites simultaneously, based on chemical properties of accurate mass and chromatographic retention time. Application of this technique to AKU showed that the biochemical alterations following a) targeted deletion of the *HGD* gene in mice to cause AKU, and b) nitisinone treatment, reflect a complex, interconnected network of metabolite changes. The data not only stimulate myriad further lines of research into AKU pathophysiology but also reveal previously uncharacterised association between metabolites at a network level. These studies provide a prime example of how rare diseases such as inborn errors of metabolism can offer unique windows into metabolism and human physiology more generally.

ABBREVIATIONS

^{18}F-NaF PET	^{18}F-NaF positron emission tomography
3-MT	3-methoxytyramine
5-HIAA	5-hydroxyindoleacetic acid
AKU	Alkaptonuria
AM	Accurate mass
AMRT	Accurate mass - retention time
APCI	Atmospheric-pressure chemical ionisation
BDI-II	Beck's Depression Inventory-II
BQA	Benzoquinone acetic acid
CEF	Compound exchange file
CID	Collision-induced dissociation
CNS	Central nervous system
CP	Cross polarisation
CSF	Cerebrospinal fluid
DHMA	3,4-Dihydroxymandelaldehyde
DNP	Dynamic nuclear polarisation
ENU	*N*-ethyl-*N*-nitrosourea
ESI	Electrospray ionisation
FC	Fold change
FDA	U.S. Food and Drug Administration
FSLG	Frequency-switched Lee-Goldburg (NMR pulse seq.)
FTICR	Fourier-transform ion cyclon resonance
GAG	Glycosaminoglycan
GC	Gas chromatography
HETCOR	HETeronuclear CORrelation spectroscopy
HGA	Homogentisic acid
HGD	Homogentisate 1,2-dioxygenase
HPLA	3-(4-hydroxyphenyl)lactic acid
HPPD	4-hydroxyphenylpyruvic acid dioxygenase
HSQC	Heteronuclear Single Quantum Coherence spectroscopy
HT-I	Hereditary tyrosinaemia type-I
HVA	Homovanillic acid
INEPT	Insensitive Nuclei Enhanced by Polarisation Transfer
LC	Liquid chromatography
LC-QTOF-MS	Liquid chromatography quadrupole time-of flight mass spectrometry
MA	Metadrenaline
MALDI	Matrix-assisted laser desorption/ionisation
MAS	Magic angle spinning
MHPG	3-methoxy-4-hydroxyphenylglycol
MPP	Mass Profiler Professional
MS	Mass spectrometry

NAC	National Alkaptonuria Centre
NIH	National Institutes of Health
NMA	Normetadrenaline
NMDA	*N*-methyl-D-aspartate
NMR	Nuclear magnetic resonance
NTFA	2-Nitro-4-trifluoromethylbenzoic acid
OA	Osteoarthritis
PBS	Phosphate buffered saline
PCA	Principal components analysis
PG	Proteoglycan
PLS	Partial least squares
ppm	Parts per million
QA	Quality assurance
QC	Quality control
RA	Rheumatoid arthritis
RF	Radiofrequency
RT	Retention time
ssNMR	Solid-state nuclear magnetic resonance
TIC	Total ion chromatogram
TOCSY	TOtal Correlation SpectroscopY
TOF	Time-of-flight

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

1.0   GENERAL INTRODUCTION

Published work resulting from content included in this chapter (see Supporting Material):

Ranganath LR, Norman BP, Gallagher JA. Ochronotic pigmentation is caused by homogentisic acid and is the key event in Alkaptonuria leading to the destructive consequences of the disease – a review. J Inherit Metab Dis. 2019;42:776-92.

BPN contributed to the literature review and manuscript writing for this publication. The publication includes key background information on AKU (sections 1.1.1 – 1.1.8 in this thesis) and a literature review on the chemistry of ochronotic pigment (section 1.3 in this thesis).

1.1   Alkaptonuria (AKU)

1.1.1 AKU is an historic disease

AKU is an iconic disease. It holds a unique place in medical history as the first disease shown to follow Mendelian inheritance. Archibald Garrod (1857-1936), an English physician widely regarded as the father of inborn errors of metabolism, studied four disorders: AKU, pentosuria, cystinuria and albinism. Pentosuria (OMIM #260800) is a defect in L-xylulose reductase, necessary for xylitol metabolism, leading to overproduction of pentose sugars and pentosuria, but is otherwise harmless as it does not accumulate in the body or produce a disease process of its own (1). This is unlike AKU (OMIM #203500), in which disease is attributable to accumulation of the metabolite homogentisic acid (HGA) in tissues throughout the body (2–4).

Garrod's close study of families with AKU led to his observation that AKU was more likely to occur in the children of first cousins. Working with William Bateson, the chief proponent of Mendel's principles of heredity at the time, Garrod established that AKU followed a characteristic pattern of inheritance in families that is now referred to as autosomal-recessive. In Garrod's seminal 1902 paper on AKU, he refers to the disease as a "study of chemical individuality" (2). In this regard, Garrod alluded to the notion that inherited diseases are attributable to specific biochemical causes, despite the absence of knowledge at that time about particular molecules such as DNA and the concept of genes. Garrod's pioneering work on AKU paved the way for subsequent discoveries that many other diseases adhere to Mendelian modes of inheritance.

1.1.2 The genetic basis of AKU

It is now known that the specific genetic cause of AKU is bi-allelic mutations in *HGD.* The *HGD* gene encodes the enzyme homogentisate 1,2-dioxygenase (HGD; E.C.1.12.11.5.), an enzyme of phenylalanine and tyrosine metabolism (5). *HGD* is located on chromosome 3 (region: 3q21-q23), comprises 14 exons and encodes a protein of 445 amino acids (6). The HGD enzyme is responsible for metabolising ingested phenylalanine and tyrosine that is surplus to daily needs. For tyrosine, a non-essential amino acid, this also includes that produced during protein turnover. In humans without AKU, significant HGD enzyme expression has only been shown in the liver and kidney (https://www.proteinatlas.org/ENSG00000113924-HGD/tissue) (7).

As of January 2020, 213 unique human *HGD* mutations have been identified (8) (http://hgddatabase.cvtisr.sk/). The general worldwide prevalence of AKU is estimated to be 1 in 100,000-250,000. AKU hotspots have been identified in Dominican Republic, India, Jordan and Slovakia, with the greatest incidence in North West Slovakia; estimated to be 1 in 19,000. These hotspots are attributed to limited gene pools associated with isolated communities. Increased incidence in Dominican Republic is thought to be a classic founder effect, whereas the high incidence in Slovakia is a surprising and somewhat unexplained result of at least 12 different mutations (6). The discovery of AKU hotspots in recent years reflects heightened interest in the disease through the Findacure (www.findacure.org.uk) and DevelopAKUre (www.developakure.eu/) consortia. It is likely that there is a large worldwide pool of undiagnosed AKU, and further hotspots could be identified particularly with heightened awareness of the disease in developing countries (9).

1.1.3 AKU is a disease of tyrosine metabolism

The metabolic defect in AKU, as a result of *HGD* mutations, is a lack of the HGD enzyme. This results in over-production of HGA (2,5-dihydroxyphenylacetic acid or homogentisic acid) (10), the hallmark biochemical feature of the disease (Figure 1.1A). In AKU the increased HGA is excreted in the urine in daily gram quantities, and HGA concentration in the urine is 1000-fold greater than that in the circulation (11). Despite such efficient renal excretion of HGA, the circulating concentration of HGA remains elevated; mean HGA serum concentration in untreated AKU is 30 µmol/L (range 11.9-75.2 µmol/L) (12), compared with <1.5 µmol/L in healthy subjects (13,14). The elevated HGA directly causes a disease process referred to as ochronosis, whereby a striking black-yellow 'ochronotic' pigment accumulates in tissues throughout the body (3,10). It is proposed that ochronotic pigment is formed by the oxidative polymerisation of HGA. In the classic scheme, HGA is oxidised to form the intermediary benzoquinone acetic acid (BQA), which subsequently polymerises to result in the final pigment product (15) (Figure 1.01B).

The most debilitating feature of AKU is a severe, early-onset osteoarthropathy which is directly attributable to ochronosis. Connective tissues, and especially cartilage, are particularly susceptible to ochronotic pigmentation (Figure 1.02). In joint cartilage this leads to a form of osteoarthritis (OA) due to presence of ochronotic pigment, which causes the cartilage to become brittle and break down.

A

Phenylalanine

Phenylalanine hydroxylase

Tyrosine ⟶ Thyroid hormones / Catecholamines / Melanin

Tyrosine aminotransferase

3-(4-Hydroxyphenyl)pyruvic acid ⇌ 3-(4-Hydroxyphenyl)lactic acid

4-Hydroxyphenylpyruvic acid dioxygenase

⊢------- Nitisinone

Homogentisic acid ⟶ Benzoquinone intermediates ⟶ Connective tissue ochronotic pigment & **osteoarthropathy**

Homogentisate 1,2-dioxygenase

Alkaptonuria

Maleylacetoacetic acid ⟶ Succinylacetoacetic acid

Fumarylacetoacetic acid

Fumarylacetoacetate hydrolase

Hereditary Tyrosinaemia Type-1

Succinylacetone

Fumaric acid + Acetoacetic acid

B

Homogentisic acid → *oxidation* → Benzoquinone acetic acid → *polymerisation* → Ochronotic pigment (?)

Figure 1.01. AKU, the tyrosine pathway and ochronosis. A: the phenylalanine-tyrosine metabolic pathway and its defect in AKU. In AKU, lack of the enzyme homogentisate 1,2-dioxygenase causes an increase in homogentisic acid (HGA) in the circulation, which accumulates in connective tissue as ochronotic pigment. Nitisinone is used for reduction of HGA concentration in AKU due to its inhibition of the enzyme 4-hydroxyphenylpyruvic acid dioxygenase. The tyrosine pathway in full is only present in liver and kidney. **B:** the classic scheme of ochronotic pigment formation in AKU. HGA oxidises to benzoquinone (BQA) intermediate, which undergoes subsequent self-polymerisation to form ochronotic pigment (10).

1.1.4 Clinical presentation of AKU

An interesting feature of AKU is the slowly progressive nature of the disease. The genetic defect is present from birth, but often the only sign of the disease in the first decades of life is darkening of urine when left to stand, due to high HGA concentrations (Figure 1.03A). Besides this, the overt clinical manifestations of the disease do not usually appear until the third-fourth decades of life (11).

Ochronosis is central to the pathophysiological features of AKU, which are multisystemic (16). In addition to the aforementioned early-onset osteoarthropathy, the debilitating features of the disease include renal and prostate stones, cardiac valve disease, osteopenia, fractures and rupture of muscle, tendon and ligament (3,17,18).

Ochronosis becomes apparent from the 3rd decade onwards; this is with joint involvement but with ochronotic pigment visible externally such as in the pinna of the ear (Figure 1.03B), followed by the sclera of the eye (Figure 1.03C). Rate of ochronosis and disease progression rapidly increases in the 5th and 6th decades of life (19). The worsening of symptoms with age is potentially due to a number of factors, including decrease in urinary excretion of HGA due to age-related impairment in kidney function (18,20), the cumulative effect of HGA exposure to tissues over time, and degradation of joint cartilage occurring as part of the natural ageing process (9). Renal failure markedly accelerates ochronosis and morbidity in AKU (21,22). Interestingly, there is no correlation between particular *HGD* mutations and HGA excretion or disease severity (18).

Ochronotic osteoarthropathy predominantly affects axial and load-bearing joints; particularly the spinal column, hips and knees (Figure 1.02) (23). Often lower back pain is one of the earliest presenting clinical symptoms (24). Spinal symptoms worsen from the fourth decade onwards, leading to progressive kyphoscoliosis and impaired spinal and thoracic mobility (25). Ochronotic osteoarthropathy is generally degenerative as opposed to inflammatory, and the consequential chronic joint pain almost inevitably requires multiple joint replacement surgeries over the course of lifespan.

In terms of the non-joint features of AKU, cardiovascular manifestations are common (Figure 1.03D-F). Aortic valve disease is a particularly frequent feature, with 40% prevalence by the fifth decade of life (26). Aortic valve stenosis and sclerosis is attributable to the deposition of ochronotic pigment in connective tissue, which leads to dystrophic calcification. Increased prevalence of kidney stones in AKU has been attributed to the extremely high levels of urinary HGA excretion (18); a recent study estimated kidney HGA clearance rates by glomerular filtration and net renal tubular secretion to exceed the theoretical maximum renal plasma flow of 600 mL/min (20).

Aside from the physical features of the disease, psychopathological features such as depression and sleep disturbance are also common in AKU (17). These features are likely to relate to living with a chronic disease and its associated painful and disabling symptoms more generally.
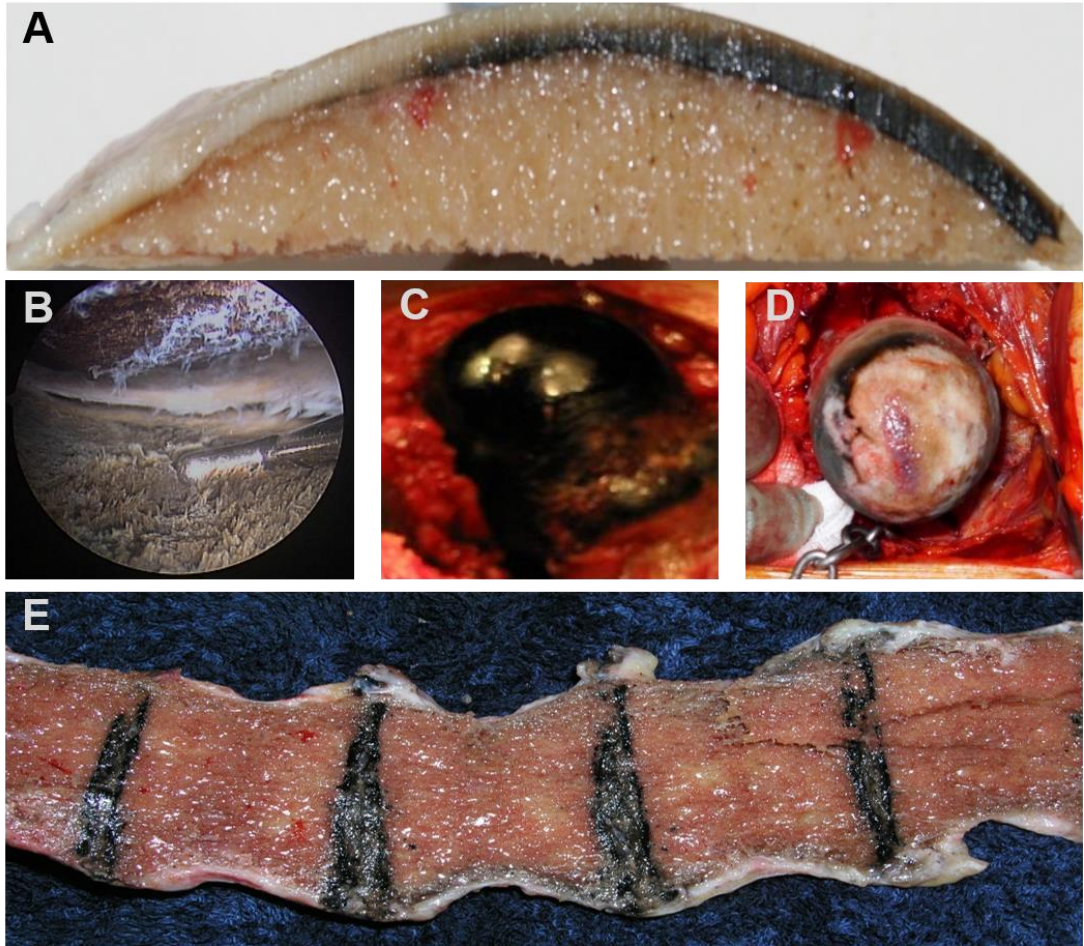
Figure 1.02. Features of ochronosis in human joint tissue. A: unstained cut section of femur condyle showing little pigment on left side progressing to full pigmentation on the right side. **B:** dark pigmentation seen in arthroscopy of knee joint showing fibrillar blackened cartilage. **C:** head of femur showing uniform pigmentation. **D:** hip joint showing rim of cartilage and marked cartilage loss with exposure of underlying bone. **E:** spine seen from posterior or dorsal aspect showing marked ochronosis of cartilaginous intervertebral disks. Figure adapted from (16).
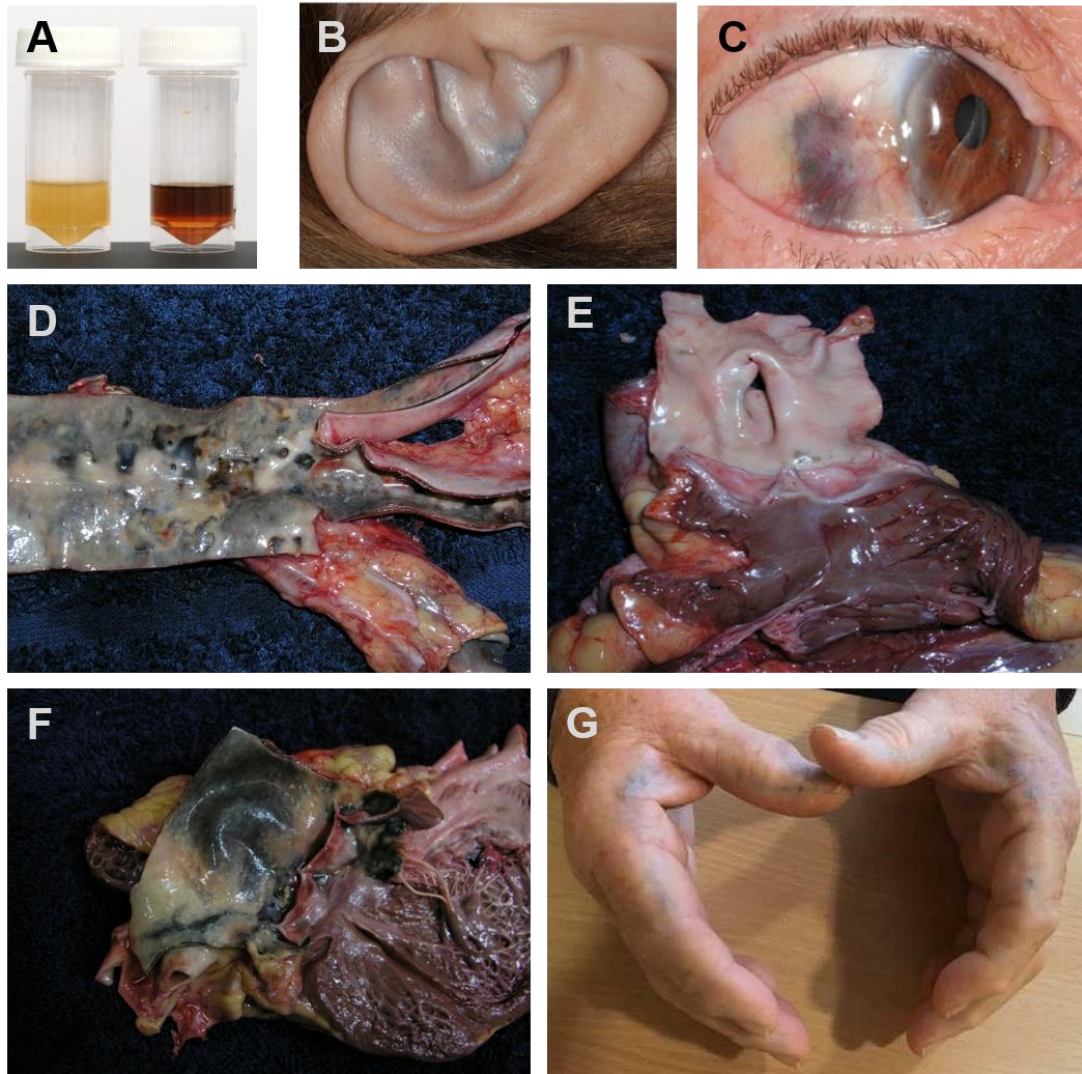
Figure 1.03. Features of ochronosis in human non-joint tissue. A: dark AKU urine (right) compared with non-AKU urine (left). **B:** external ear cartilage pigmentation. **C:** dark pigmentation of temporal aspect of sclera in right eye. **D:** longitudinal cut section of abdominal aorta and common iliac bifurcation, showing more pigment at bifurcation and branch point orifices. **E:** low-pressure pulmonary trunk and valve showing little pigment. **F:** high-pressure aortic root and valve with marked pigmentation. **G:** pigment at junction of palmer and dorsal skin of hands. Figure adapted from (16).

1.1.5 Distribution of ochronosis is not uniform

At post-mortem, pigment is patchy and present in areas of stress/damage in non-cartilaginous tissues, suggesting that damaged tissue undergoes pigmentation whereas undamaged tissue is resistant to pigmentation. Cartilage is highly pigmented, but this is also uneven across the cartilage (23). Possible factors involved in damaging tissues and producing pigment are shown in Table 1.01. Despite the constant exposure of all body tissues to HGA in AKU, most tissues are resistant to pigmentation.

Table 1.01. Factors influencing tissue ochronotic pigmentation.

Tissue	Damaging factor
Spinal intervertebral disks	Weight-bearing stress
Joints	Weight-bearing stress, movement damage
Tendons	Tensile stresses
Ligaments	Tensile stresses
Aortic valve	Systemic blood pressure, expansile stress
Aortic root	Systemic blood pressure, expansile stress
Arterial tree branch points	Shear stress, Bernoulli effect
Cartilage of airways	Expansion, lengthening and contracting
Ear cartilage	Pressure on ears during sleep
Sclera and conjunctiva	UV light damage, stress arising from muscle contraction

Cartilage especially susceptible to ochronosis includes highly loaded tissues such as articular cartilage, fibrocartilage (intervertebral discs, pubic symphysis), costochondral, as well as less-stressed cartilage in the pinna and nose, and the trachea-bronchial system (23). Highly stressed tissues such as joints, spine, tendons and ligament are pigmented. Ochronotic pigment has been noted in the skin, especially at the interface of the palmar and dorsal skin of the hands (Figure 1.03G); possibly sites of greater stress (27). Pigmentation of the nails of the hands and feet has been described. Secretory glands in the axilla, groin and eyelids pigment (28). Ocular tissues also pigment, including the conjunctivum, cornea and sclera. Ochronosis has been observed in the tympanic membrane and ear wax. Cardiac valves are affected by ochronotic pigment more in left side of heart, and aortic more than mitral valve (23); the pulmonary vascular system and the right side of the heart is much less pigmented. Pigmentation can be observed in the arterial system, mostly around branch points and tributaries; areas of greater stress (23). There is currently no information on pigmentation in the venous system. Renal parenchyma, especially the medulla and pyramidal tissues, show ochronosis. The periosteum has been shown to pigment but not the bone itself, raising the possibility of bone mineral binding to collagen preventing binding of HGA-derived molecules. Teeth enamel has been said to pigment, but needs more evidence given that bone does not pigment (29). It is generally held that muscle, liver, lung (excluding the bronchial system) and brain are not affected by ochronotic pigment. There is no convincing description of gastrointestinal and genital pigmentation. There is no description of pigment in pancreas, an organ where alkaline secretions are produced unlike the salivary glands. It is

11

not known if there is biliary excretion of HGA, even though pigmented gall stones make this likely. Paucicellular tissues such as cartilage, tendon and ligament appear to pigment easily compared to highly cellular tissues such as liver.

It is likely that HGA is increased in the cerebrospinal fluid (CSF), tears and saliva although direct evidence is lacking. Circulating HGA is well characterised but tissue HGA, *i.e.* intracellular HGA, has not been directly assessed. It is likely HGA (and other metabolite acids) is protein-bound but this requires characterisation.

1.1.6   Effect of ochronosis

The effect of ochronosis on affected tissues is summarised in Table 1.02. The deposition of ochronotic in tissue including cartilage alters its material properties leading to the tissue becoming hard and brittle. The Young's modulus, a mechanical measure of the stiffness of a solid material, is altered depending upon the degree of pigmentation in AKU. Our group has proposed a model of joint failure based on initiation of ochronosis in calcified cartilage before progressing to involve the entire cartilage and spiralling into joint failure, requiring joint replacement (30).

This process of ochronotic stiffening also compromises intervertebral discs and adjoining bone. Ochronosis of intervertebral disks leads to severe spinal disease characterised by severe pain as well as kyphosis and scoliosis; spinal cord compression by involved discs can require decompressive spinal surgery. Spinal fusion can ensue in the latter stages resulting in loss of mobility and

flexibility in all parts of the spine, but especially in the cervical and lumbar regions.

Although all cardiac valves can show ochronosis, it is the aortic valve and the aortic root that are subject to more pigmentation resulting in severe aortic stenosis requiring valve replacement surgery. Valve replacement surgery is often hazardous due to the friable ochronotic aortic root and valve.

Scleral ochronosis can distort the corneal curvature and result in astigmatism (31,32). Ochronotic ear cartilage may be associated with pain in pinna of the ear (31). Hearing loss especially to high frequency is a feature of AKU (33,34). Rigid articular cartilage can lead to subchondral osteopenia. Generalised osteoporosis is also noted in AKU and associated with increased fracture (35). The failure of ochronotic connective tissue results in tendon ruptures, especially of the Achilles tendon, but other tendon ruptures have also been observed such as flexor and extensor foot, patellar, bicipital and gluteal regions (36,37). Similarly ruptures of ochronotic ligaments have been described. Muscle rupture has also been reported more frequently in AKU. Stone formation featuring ochronotic pigment has been found in kidney, prostate, gall bladder and salivary glands, resulting in symptomatic obstruction of these organs (38). Renal failure can ensue both due to obstruction and renal parenchymal ochronosis, sometimes leading to fatal intractable haemolytic anaemia (39,40).

Table 1.02. Effect of ochronosis on affected tissues.

Tissue with ochronosis	Effect
Articular cartilage	Resorption of calcified articular cartilage and subchondral bone and joint failure
Articular cartilage	Osteopenia due to resorption activated by stiff pigment
Intervertebral discs	Spondylosis, fracture, osteopenia, scoliosis, kyphosis, cord compression, radiculopathy
Ligaments	Rupture
Aortic valve and root	Aortic stenosis and aortic rigidity
Tendons (*e.g.* Achilles, foot flexor and extensor, patellar, gluteal, biceps)	Rupture
Sclera	Distortion in corneal curvature; astigmatism, glaucoma
Ear cartilage	Possible pain in ear
Middle/inner ear	Hearing loss, ear ossicles disorder leading to conductive deafness and high frequency hearing loss

1.1.7 Treatment of AKU

It is reasonable to assume that antioxidant therapy should be efficacious in treating AKU by preventing oxidation of HGA to BQA, the proposed initial step in ochronotic pigment formation. For reasons that are unclear, antioxidant ascorbic acid therapy has no clear benefit in AKU, a condition with a proposed oxidant-damage hypothesis in terms of ochronosis formation (41,42). This is analogous to the lack of efficacy of antioxidant strategies in coronary artery disease despite the well-validated oxidised low-density lipoprotein theory of atherogenesis (17).

Reduced protein intake should decrease flux down the tyrosine catabolic pathway and, in theory, reduce ochronosis and lower morbidity in AKU. However, except for occasional case reports in childhood, there is no evidence that restricting dietary protein decreases ochronosis. Anecdotally, vegetarians and vegans are noted to have less ochronosis and lower morbidity but no systematic evidence exists to support low protein diet in AKU (31). All other approved therapies used in clinical practice are supportive and palliative and do not address the elevated HGA and its effects (17).

A more promising treatment strategy in AKU is pharmacological inhibition of 4-hydroxyphenylpyruvic acid dioxygenase (HPPD), the enzyme responsible for formation of HGA and its downstream metabolites (Figure 1.01A). Nitisinone (2-(2-nitro-4-trifluoromethylbenzoyl)-cyclohexane-1,3-dione), a reversible competitive inhibitor of HPPD, has revolutionised the treatment of hereditary tyrosinaemia type-1 (HT-1; OMIM #276700) (43,44). HT-1 is another rare inherited disorder of tyrosine metabolism, in which defective activity of the enzyme fumarylacetoacetate hydrolase causes accumulation of toxic metabolites that are downstream of HGA in the tyrosine pathway (Figure 1.01A). Increased fumaryl- and maleyl-acetoacetic acid results in hepatorenal damage and increases in derivatives of succinylacetone leads to a porphyria-like neurological crisis. These severe symptoms are present from birth if untreated, therefore nitisinone, a licensed drug in HT-1, is routinely administered as soon as HT-1 diagnosis is confirmed or suspected. In HT-1, nitisinone's inhibition of HPPD controls liver failure in >90% of cases, improves chronic liver and kidney disease and usually abolishes neurological crises (45–47).

Nitisinone was first suggested as a treatment for HGA in AKU in 1998 (48), and is now shown to be highly effective in reducing HGA despite not being approved in AKU (49). Nitisinone not only decreases HGA but also arrests ochronosis in mice (50). A recent publication shows photographic evidence of partial reversal of ochronosis in sclera and ear cartilage in AKU patients (49).

The National Institutes of Health (NIH) carried out a nitisinone interventional study in AKU between 2005 and 2009 and showed a sustained and marked decrease in urine HGA over the 3-year duration of the study (51). However, their agreement with the regulatory agency, the U.S. Food and Drug Administration (FDA), to approve nitisinone for AKU, required them to show a difference in range of motion at the hip between the nitisinone-treated and untreated groups. Statistical significance was not found for the range of motion comparison and the study reported inconclusive.

Since 2012, nitisinone has been used off-label to treat patients attending the United Kingdom National AKU Centre (NAC) in Liverpool, funded by NHS England Highly Specialised Services; 2 mg nitisinone daily, which is the same dose given in the NIH study. The data from the NAC patient cohort were recently reported (49); these data confirmed the biochemical efficacy of nitisinone in AKU in terms of urinary HGA, but also showed slower progression of morbidity. Furthermore, the NAC data showed overt partial reversal of ochronosis, the primary pathogenetic event in AKU.

It is important to note that nitisinone, while shown to be a highly effective therapy for preventing ochronosis, is not a perfect therapy in AKU. It is well-recognised that nitisinone treatment results in significant hyper-tyrosinaemia

in AKU (18) and HT-I (43,44); tyrosine is upstream of HGA and HPPD, the enzyme inhibited by nitisinone (Figure 1.01A). The consequences of hyper-tyrosinaemia are not fully understood in AKU, but concerns have been raised that it may contribute to the neurodevelopmental delay observed in infants with HT-I on nitisinone therapy (44). Recent analyses did not find changes to monoamine neurotransmitters in brain tissue from nitisinone-treated mice (52), although more data are required to fully ascertain the impact of hyper-tyrosinaemia on central nervous system (CNS) homoeostasis. Other potential future therapies for AKU may employ approaches to directly restore HGD activity, for example by enzyme replacement therapy and gene therapy. However, research into these approaches is still in its very early stages, and they are also not without their own concerns and challenges, for example potential off-target effects.

1.1.8 Model systems of AKU

Model systems have been developed for investigating various effects of HGA exposure under controlled conditions *in* vitro, *in vivo* and *ex vivo*. *In vitro* models include human serum (53,54) and osteoblastic (cell line) and chondrocytic (cell line and isolated from human cartilage) cell cultures (55–62). Incubation of cell cultures with HGA leads to ochronotic pigment formation by four weeks. The amount of ochronotic pigment formed is proportionate to the concentrations of HGA in the medium (62). *Ex vivo* organotypic approaches have modelled AKU by studying the effect of HGA on human cartilage explants (55,63,64). In these studies, pericellular pigmentation of cartilage explants is visible after two months of incubation with HGA (63). Development of these model systems has not only been useful in studying the

development of ochronotic pigment under controlled conditions, but also a number of other pathophysiological events closely associated with ochronosis including secondary amyloidosis (55,60,61), perturbed redox homeostasis (53,54,56,57,59,60,65) and disorder of the cartilage extracellular matrix, including proteoglycan loss and collagen fibril re-arrangement (59,61,64).

Several *HGD*-knockout mouse models of AKU have also been developed. These mice were previously only considered to be models of AKU biochemistry as they present with confounding pathology. Despite a marked elevation in plasma and urinary HGA, mice do not develop the striking macroscopic ochronosis observed in adult humans, except in the kidneys after around 13 months (66). Several potential explanations for the lack of widespread macroscopic pigmentation in AKU mice are: a) reduced joint-loading of quadrupedal mice, b) faster cellular turnover for removal of pigment and c) that the shorter lifespan of mice compared with humans might be insufficient for gross pigment deposition (50).

Closer observations revealed that AKU mice do show pigment in cartilage in a way that mirrors the early stages of human ochronosis (50,66). Ochronotic pigment can be observed microscopically in chondrocytes within the articular cartilage at around 15 weeks (50) and this pigmentation increases progressively over the lifespan. This pattern of pigmentation is similar to early human AKU, in which pigment deposition begins within single chondrocytes of the calcified cartilage and eventually results in widespread extracellular pigmentation (30). An exact serum HGA concentration threshold at which ochronotic pigment develops is not known, but a threshold range of 40-60 µmol/L has been reported in mice (67).

Nitisinone is also shown to be an efficacious therapy in AKU mice, as it decreases plasma and urine HGA concentrations, as with patients. Marked hyper-tyrosinaemia also occurs in AKU mice treated with nitisinone (50,68–70). These factors support the use of *HGD*-knockout mice as a model of biochemistry in both untreated and nitisinone-treated AKU. Studies in mice have also enabled the important observation that nitisinone completely arrests the process of ochronosis. Lifetime treatment with nitisinone also completely prevented the deposition of ochronotic pigment in cartilage. Nitisinone treatment could not, however, reverse ochronosis by eliminating existing pigmentation (50,69).

More recently, a mouse model of AKU has been developed in our laboratory by targeted knockout-first disruption of the *HGD* gene (70); the AKU mouse data described in previous studies (50,66–69) was from mice with *HGD* disruption by *N*-ethyl-*N*-nitrosurea (ENU) mutagenesis. In this regard, the targeted *HGD*-knockout mouse is a preferred model of AKU, as ENU mutagenesis causes a high frequency of genomic mutations, which can be potentially confounding (71). However, the AKU phenotypes of these mice, which were both studied as part of this thesis, are almost identical (70).

Additional data from the recent report of the new targeted *HGD*-knockout mouse also further support the indispensable contribution of the liver to clearance of HGA. Conditional liver-specific deletion of HGD (20% of liver HGD mRNA remaining), but not kidney, in non-AKU mice resulted in elevated plasma and urine HGA and ochronotic pigment deposition in cartilage (70).

1.2 Metabolomics

1.2.1 The metabolome and metabolomics

Metabolomics is the comprehensive qualitative or quantitative study of the metabolome, defined as the entire complement of metabolites within a biological system, which includes a cell, tissue, biofluid, organ or organism (72). Metabolites are small molecules, typically <1500 Da, which comprise the precursors, intermediates and products of cellular processes. Metabolites are part of a complex, inter-connected and regulated metabolic network which also includes enzymes and cofactors (73). Metabolic networks are dynamic; they are subject to constant activity, or flux, as a result of the complex interplay between genome and environment (74). The collective metabolite profile or 'fingerprint' is therefore considered to provide a snapshot of metabolism, or a direct readout of the physiological state of an organism at a given point in time (75–77).

Metabolomics is a relatively new field compared with more established 'omics approaches, such as genomics, transcriptomics and proteomics. However the basic notion of metabolomics, in that biological fluids convey key information about the phenotype, can be traced back to ancient China (2000-1500 BC), where sweet-tasting urine was used as an indication of a disease now recognised as diabetes mellitus (78). The idea that a metabolite profile can describe the makeup of a biological fluid for different individuals was first formally recorded in the work of Roger Williams in the 1940's, in which he coined the term 'biochemical individuality' (79). Williams and colleagues

showed that characteristic metabolic profiles in saliva and urine were associated with pathologies such as alcoholism and schizophrenia.

Metabolomics is complementary to the other omics fields noted above but can be considered to confer some advantages as a biomarker discovery approach. Metabolites are downstream of genes, transcripts and proteins and therefore changes in them are amplified relative to the transcriptome or proteome. Second, as metabolites are the end-products of biochemical interactions, the metabolome can be thought of as the closest biochemical representation of the phenotype (76).

The past 20 years has seen a burgeoning of interest in metabolomics, with myriad applications in diverse research fields including medical science, pharmacology, plant biology, synthetic biology, food and environmental analysis and microbiology (80). Metabolomics is also becoming a major part of systems biology approaches (81–83), which aim to understand biological systems as a whole by means of integrating data from different omics modalities.

1.2.2 *Non-targeted metabolomics: opportunities and challenges*

Metabolomics experiments can be performed using targeted or non-targeted approaches (75). In targeted experiments, a pre-defined list of metabolites is measured. The specified metabolites of interest are usually from a specific biochemical hypothesis or part of a common metabolite network or pathway. The advantage of targeted approaches is increased specificity and the ability to absolutely quantify metabolite concentrations. Early metabolomics studies employed targeted metabolic profiling of biofluids using GC-MS; gas

chromatography (GC) coupled with mass spectrometry (MS). These pioneering studies led to the discovery of numerous diseases caused by single gene defects (84,85).

The appeal of non-targeted analyses is the global nature of their scope, as the data acquisition is not limited to a pre-specified set of metabolites. Non-targeted analyses therefore enable less-biased characterisation of the metabolome and comprehensive tracking of the biochemical reactions in a system. The rapid growth in metabolomics in the past two decades is largely attributable to advances in instrumentation and technology which support a non-targeted approach. Improvements in MS and chromatographic techniques (86–90), but also informatics have particularly driven this development, making it possible to separate and measure thousands of metabolites simultaneously from minimal biological material.

While the value of non-targeted metabolomic analyses continues to be borne out (83), the interpretation of the resulting datasets remains a challenge. Metabolite structure identification and data interpretation are two particular challenges and are discussed in the following sections.

1.2.2.1 Metabolite identification

Chemical structure identification still represents a well-documented bottleneck in metabolomics workflows (91,92). There is far greater variation in the chemical configuration of metabolites compared to genes and proteins. Genes and proteins comprise linear combinations of nucleotides (4-letter codes) and amino acids (20-letter codes) respectively, meaning that they can now be identified in a comparatively simple way by sequencing. It should be

acknowledged, however, that chemical identification of post-translationally modified proteins is considerably more challenging. For metabolites, the number of possible atomic and chemical sub-structure arrangements is vast, therefore alternative identification strategies are required which are more complex than sequencing (93). The metabolite identification challenges specific to MS, the analytical platform primarily used for metabolomic analyses presented in this thesis, are described by Dunn *et al.* (94). As MS-based compound identification is based on accurate mass (AM), discrimination between species with the same monoisotopic mass but different chemical structures, such as isomers and chiral molecules, presents a particular challenge. These compounds have the same empirical formula but differ in the structural arrangement of their atoms. Another challenge is the multiplicity observed in mass spectra, whereby more than one ion relates to the parent compound; for example isotopes and adducts formed during the ionisation process. This multiplicity adds complexity to the task of assigning the ions observed to a given structure.

Generally, metabolites can be identified by reference to spectral databases/libraries or by a *de novo* approach (94). Spectral libraries are indispensable tools for metabolite identification amongst metabolomics researchers (83), as discussed in Chapter 2. *De novo* compound identification procedures are usually employed when no candidate compound match is obtained in established databases. Commonly-employed *de novo* structural identification approaches in metabolomics include collision-induced compound fragmentation, also known as MS/MS or MS^2 in mass-spectrometry (see section 1.4), and two-dimensional NMR (see section 1.52).

1.2.2.2 Data interpretation

A second challenge relates to extracting meaningful information from the vast datasets generated from metabolomics experiments. As discussed further in Chapter 3, the high-throughput nature of these experiments means that in the resulting datasets, there are usually more metabolites detected than the number of samples analysed. It is therefore not feasible to interrogate the datasets by traditional univariate approaches based on reviewing the observed values of each variable, or metabolite, on an individual basis. Instead, multivariate chemometric statistical techniques are employed, which consider the values of multiple metabolites across the samples simultaneously (95–97). The overall aim in metabolomics is to extract from a complex background of signals the most relevant features of the data that relate to the biological question at hand (95). Generally, multivariate analyses frequently applied to metabolomic datasets fall under two categories; unsupervised and supervised techniques. These two approaches have the same aim of visualising and identifying the metabolites that discriminate between different sample classes, for example patients with a particular condition *versus* unaffected controls.

Unsupervised approaches are used to visualise the structure of the data in a non-biased way, as the generated models do not consider sample class membership. Principal components analysis (PCA) is one of the most-employed unsupervised method in metabolomics (see Chapter 3, section 3.6.1). PCA can be employed to assess the overall structure of the dataset with the aim of identifying the metabolites that have the greatest contribution to the overall variation present in the data. Supervised approaches are often

used to confirm the unsupervised analyses and to further interrogate the differences in the data for the purposes of biomarker discovery. Supervised methods aim to identify the variables with the greatest contribution to the separation between the pre-specified groupings within the data. In other words, this approach aims to identify the metabolites that predict sample group classification. Data modelling techniques based on partial least squares (PLS) methods are some of the most-widely employed supervised approaches in metabolomics (98), although new techniques are constantly emerging, such as predictive models based on advanced machine-learning algorithms (99).

The plethora of computational tools and multivariate statistical approaches for analysis of metabolomics data reflects the ever-increasing interest in the field, but consequently, selection of the most appropriate is not trivial. Although the field has reached some consensus on favoured tools (*e.g.* multivariate approaches using techniques such as PCA, PLS), the need for harmonisation in data analysis strategies employed across the community has been highlighted (100). There is also recognition that some approaches might be favoured out of convenience rather than for being the best suited to the analysis (101).

1.2.2.2   Standards in metabolomics research

Compared with established omics fields, metabolomics is still relatively in its infancy. Despite the rapid advancements in the field in recent years, consensus guidelines for performing metabolomics experiments are still evolving (102). In 2007, the Metabolomics Standards Initiative published a series of guidelines covering all stages of a metabolomics experiment (103–

107). With reference to the challenges discussed above, these guidelines included minimum reporting standards for analysis of metabolomics data (106) and metabolite identification (107).

In the minimum reporting standards guidelines, 4 levels of metabolite identification were described (107) (Table 1.03). Level 1, the highest level of metabolite identification confidence, requires two matched orthogonal chemical properties against an authentic chemical standard analysed under identical analytical conditions. In LC-MS data, orthogonal chemical properties could refer to measured AM plus isotope pattern, chromatographic retention time (RT) or MS/MS fragmentation spectra. 'Putative' annotation (levels 2 and 3) refer to one or two matched properties only against data that were not acquired under identical analytical conditions; these often include databases or spectral libraries generated in other laboratories using different analytical parameters. Level 4 identification refers to unknown metabolites for which no data currently exists.

Encouraging recent trends in metabolomics are the drive for freely-available, open source computational tools (108), public repositories of metabolomics data (109,110) and open data analysis platforms in which the methods and results from metabolomics analyses are freely disseminated (111).

Table 1.03. Levels of metabolite identification in metabolomics. The 4 levels of metabolite identification are from the 2007 proposed minimum reporting standards of the Chemical Analysis Working Group, as part of the Metabolomics Standards Initiative (107). Modified from (94).

Level	Confidence of identity	Level of evidence
1	Confidently identified compounds	Comparison of two or more orthogonal properties with an authentic chemical standard analysed under identical analytical conditions
2	Putatively annotated compounds	Based upon physicochemical properties and/or spectral similarity with public/commercial spectral libraries, without reference to authentic chemical standards
3	Putatively annotated compound classes	Based upon characteristic physicochemical properties of a chemical class of compounds, or by spectral similarity to known compounds of a chemical class
4	Unknown	Although unidentified and unclassified, these metabolites can still be differentiated and quantified based upon spectral data

1.2.3 LC-QTOF-MS: application to metabolomics

The wide range of analytical techniques employed in metabolomics reflects the diversity of applications within the field. Analytical platforms include MS coupled to chromatographic techniques such as liquid chromatography (LC) and GC, nuclear magnetic resonance (NMR) spectroscopy and vibrational spectroscopy (*e.g.* infrared spectroscopy, Fourier transform-infrared (FTIR) spectroscopy and Raman spectroscopy) (112). Along with NMR, LC-MS is one of the most widely employed techniques in metabolomics (see Table 1.04

which compares the attributes of LC-MS and NMR for metabolomics), owing to its high throughput, soft ionisation and broad coverage of metabolites (113).

Developments in MS and chromatographic techniques have been key factors in the rise in popularity of metabolomics over recent years (86–90). Such advances include the development of mass analysers with improved mass resolution. High-resolution instruments have the ability to separate closely-spaced masses in complex mixtures, for example compounds with the same nominal mass (*i.e.* mass expressed in integer Da units) but with different elemental compositions, and therefore different monoisotopic AM (*i.e.* mass expressed in mDa) (114). Mass analysers such as time-of-flight (TOF), Fourier transform ion cyclotron resonance (FTICR) and Orbitrap are the highest resolution instruments currently available (mass resolution >10,000). Other advances include improvement to electronic components of the hardware, such as the detector, supporting increased signal sensitivity and attenuation of signal saturation. Development of improved software programs has also enabled more user-friendly operation (115).

Table 1.04 Comparison of NMR and MS analytical platforms for metabolomics. Information from (116).

	NMR	MS
Experimental aspects		
Techniques for biofluid samples	1D ^{1}H and 1D ^{31}P 2D CPMG 2D TOCSY 2D COSY	GC/LC/CE-MS ESI-Q-TOF ESI-Q-Orbitrap MALDI-TOF
Techniques for tissue samples	HRMAS	MALDI-TOF
Average metabolite detection	<200	500+
Equipment		
Cost	Very high	Moderate-high
Maintenance cost	High	Moderate
Data automation	Yes	Yes
Experimental time	1D – 5 min; 2D - up to 20 min	Up to 60 min with chromatography
Advantages	Non-destructive Intrinsically quantitative Minimal sample preparation Robust and reproducible Can use for any sample	Widespread High sensitivity (pg-ng) High resolution Low metabolite detection Can be combined with chromatography
Limitations	Moderate resolution (signal overlap) Low sensitivity (µg) Low metabolite detection Very expensive	Destructive method Moderate reproducibility with salty mixtures Polarity bias +ve/-ve

Liquid-chromatography quadrupole time-of-flight mass spectrometry (LC-QTOF-MS) is the analytical platform employed in the metabolomics experiments described in this thesis. First introduced in 1948 (117), TOF mass analysers were originally used for *de novo* peptide sequencing (118). LC-QTOF-MS is a powerful analytical tool which enables rapid measurements with high sensitivity and resolution. TOFs have a theoretically infinite mass range which enables analysis of a wide range of small and large molecules; mass-to-charge ratio (*m/z*), not mass, is detected, and ion species can acquire multiple charges.

The basic principle of TOF is the separation of ionised molecules over a given path by time-of-flight. The 'flight' is the time between the application of a high voltage pulse to the back plate of an ion pulser and when the ion strikes the detector. The flight time for each mass is unique; determined by the energy of its acceleration, distance to travel and its *m/z*. For a given energy and distance, mass equals the square of the ion's time-of-flight (*t*). As TOF analysers are specifically designed to keep the values for energy and distance constant, accurate measurement of time of flight enables the accurate calculation of mass (*m*) by the following formula, in which *A* refers to energy and distance combined:

$m = At^2$

With distance and energy held constant in the TOF flight tube, the result is that ions with lower mass arrive at the detector earlier. This is because ions with smaller mass have increased velocity (119).

Figure 1.04 demonstrates the time-of-flight principle in a schematic representation of the 6550 QTOF-MS (Agilent), the specific instrument used in the metabolomics studies described in this thesis. In this instrument, ions from the sample are drawn into the MS via a hexabore capillary. The ions are focused into a beam by passing through 2 funnels followed by an octupole ion guide. A quadrupole mass filter provides the option to select for specific ions for targeted analyses (see MS/MS analysis, section 1.4). The ions are accelerated to high kinetic energy into the collision cell (left open in the metabolomics experiments here; no collision energy applied). Upon exiting the collision cell, the ions are focused into an optimally focused beam and then pass through a pair of slits into the final vacuum stage. The parallel beam of ions then enters the ion pulser. A high voltage pulse is applied to the backplate, which accelerates the ions through the stack of pulser plates into the flight tube. Once at the top of the flight tube, the direction of the ions is reversed back down the flight tube towards the detector by a two-stage electrostatic mirror. Ions arriving at the detector strike a microchannel plate, causing the release of electrons. Each channel serves as an electron multiplier. The electronic signal results in the emission of photons, which are focused onto a photomultiplier. The photomultiplier amplifies the number of photons, converting to an electrical signal proportional to the number of photons.
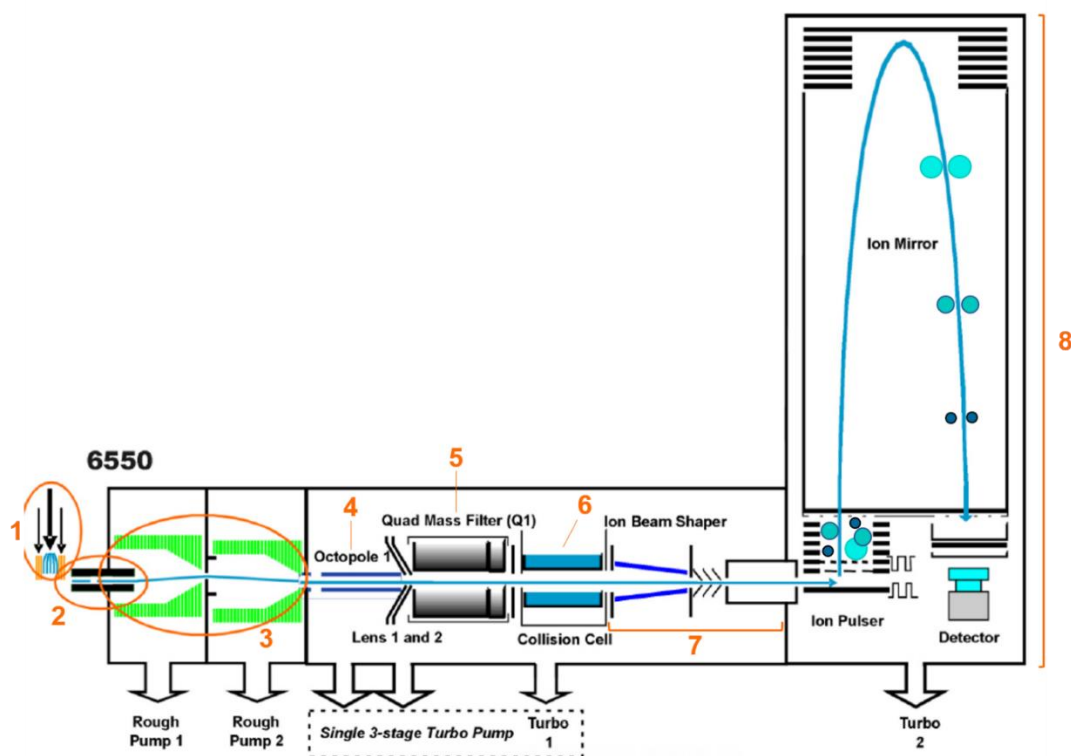
Figure 1.04. Schematic representation of a QTOF mass spectrometer. Model shown here is an Agilent 6550 QTOF. Numbering corresponds to the following hardware features. 1) ion source, 2) capillary inlet to mass spectrometer, 3) iFunnel (x2), 4) octapole ion guide, 5) quadrupole mass filter, 6) collision cell, 7) ion transfer region (ion optics), 8) flight tube comprising ion pulser, two-stage ion mirror, microchannel plate detector connected to photomultiplier. Flight tube (2 m in total), showing schematic representation of molecules separated by time-of-flight according to mass-to-charge ratio (*m/z*). Adapted from (120).

MS is particularly powerful due to the interfacing with ionisation techniques such as electrospray ionisation (ESI) (121) and matrix assisted laser desorption ionisation (MALDI) (122). ESI has become one of the most important techniques used in the coupling of MS to LC. ESI facilitates ionisation of a wide range of molecules (small molecules to polymers and proteins), operates at atmospheric pressure and moderate temperatures and is considered a soft ionisation technique (123). The basic principle of ESI is

that solutes in a solution are transformed into gas phase ions before their introduction to the MS. Figure 1.05A shows a representation of an ESI source attached to an MS inlet. The LC eluent is sprayed, or nebulised, into a chamber at atmospheric pressure. Within the chamber, the droplets are subject to a high-voltage electrostatic field, which results in the ions of one polarity migrating to the surface of the droplet. Consequently, the sample is then simultaneously charged and dispersed into smaller charged droplets (Figure 1.05B) (120). The mechanism of ESI is not fully understood but this 'ion evaporation model' is the most widely accepted for small molecules (124). The ions formed by ESI can be positively charged by addition of hydrogen (or other cations such as sodium or ammonium from mobile phase modifiers), expressed as $[M+H]^+$, where (M) is the molecular ion. Alternatively, ions can be negatively charge by removal of a hydrogen nucleus, expressed as $[M-H]^-$, or possible adduction of other anions (*e.g.* formate from mobile phase modifiers). Efficiency of the generation of gas phase ions by ESI depends on a number of factors, including chemical structure, solvent type and parameters applied to the source (125). Modifiers such as formic acid are often added to solvents to enhance ionisation. These decrease the initial droplet size (by increasing the conductivity) and provide a source of protons to aid the ionisation process.
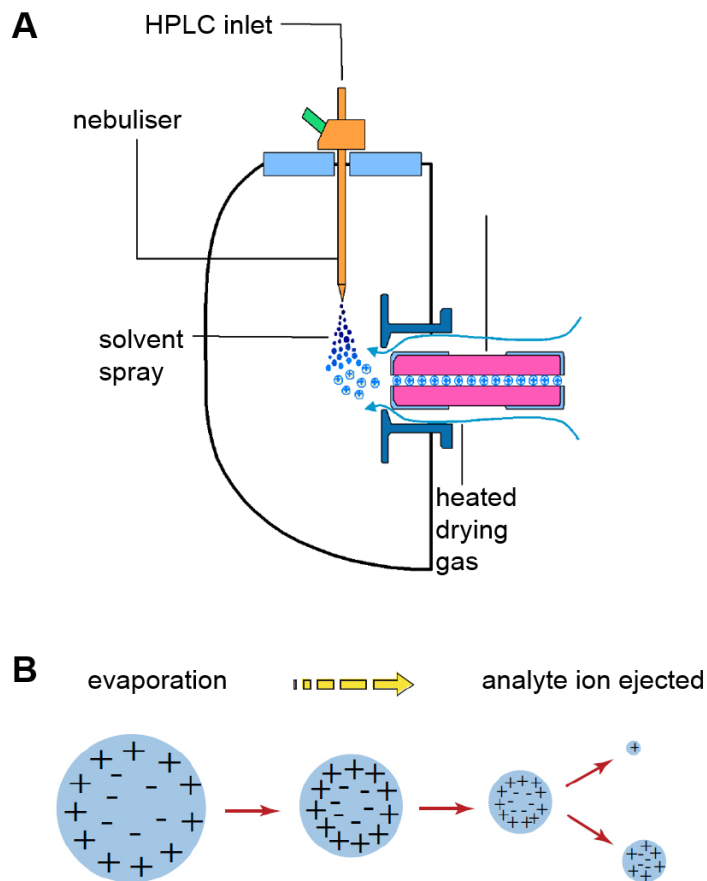
Figure 1.05. Electrospray ionisation (ESI). A: electrospray ionisation source. Note that dual ESI was the method employed in the following studies; a second nebuliser is used to introduce reference mass ions for accurate mass calibration. **B:** desorption of ions from solution in ESI. Adapted from (120).

ESI ionises molecules in the liquid phase, which makes it compatible with LC. Liquid-phase separations have become indispensable to a wide range of MS-based applications in metabolomics. LC enables separation of many components in complex mixtures by RT, the time taken for a given analyte to elute from the LC column. LC separation considerably increases the number of metabolites that can be detected, therefore widening coverage of the metabolome.

LC separation of sample analytes prior to ionisation & MS detection is important for a number of reasons. First, in analysis of complex mixtures, or matrices, presence of other non-volatile or less volatile molecules can interfere with analyte ionisation, a phenomenon known as ion suppression (126). Second, LC offers a second orthogonal dimension of separation to that observed in the MS (based on time-of-flight or *m/z* in a TOF analyser). The TOF separation occurs over the time-frame of milliseconds (*i.e.* faster than the acquisition rate of mass spectra). Without separation by chromatographic RT, which occurs over the course of seconds-minutes, all mass peaks obtained from the sample ions would appear in the same crowded spectrum, presenting a significant challenge in downstream data processing.

Separation in LC is based on the chemical properties of analytes; their differential affinities for the mobile phase, a liquid solvent passed through the column, compared with the column stationary phase, the material within the LC column referred to as the packing. The relative affinity of a given analyte for the mobile *versus* stationary phase is determined by the analyte's charge and polarity. These properties result in the formation of bands for different analytes within the column which elute at different times (Figure 1.06A). There is a vast array of LC column specifications available which are designed to selectively retain and separate different groups of compounds. Broadly speaking, there are two main types of LC; normal-phase and reversed-phase (Figure 1.06B). Reversed-phase LC is more suited to separation of non-polar analytes, which are more strongly retained by the non-polar stationary phase compared with the polar mobile phase. Conversely, normal-phase LC

combines a column that has a polar stationary phase with a non-polar mobile phase, resulting in retention and separation of polar metabolites.

Peak capacity is the term used to define efficiency of the separation in gradient elution LC, defined as the number of theoretical peaks that can be separated within a given column under specified analytical conditions. A number of factors affect peak capacity in LC, including column length, mobile phase flow rate, temperature, make-up of the stationary phase (composition and particle size) and require optimisation for a given system (127,128).
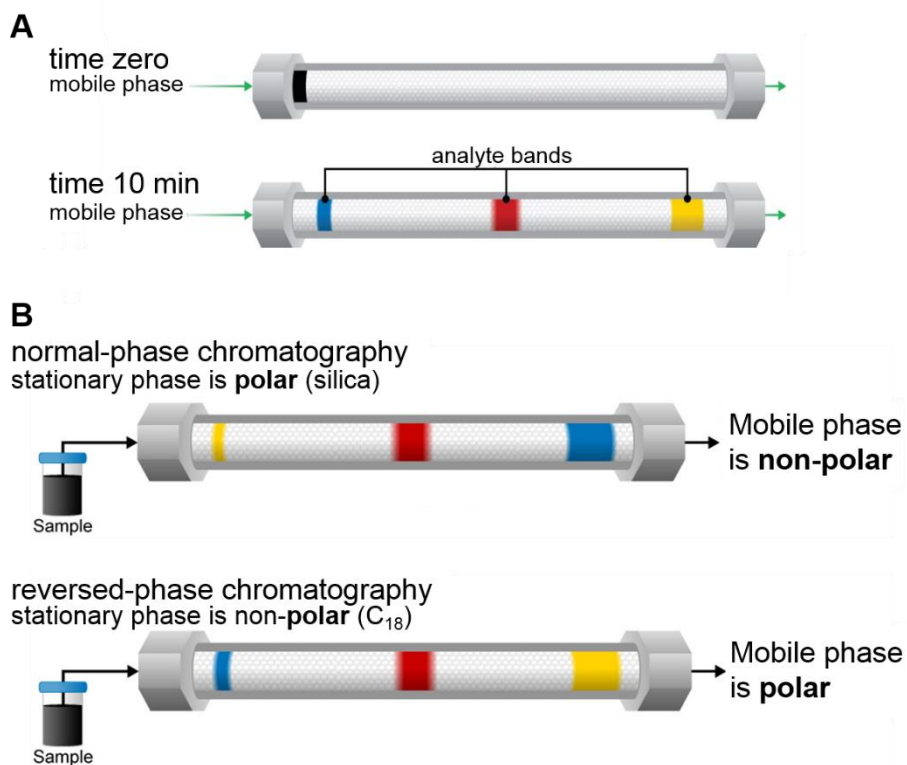
Figure 1.06. Analyte separation by liquid chromatography. A: at time zero the black band represents the injected sample. At 10 min separation of the sample has occurred into different analyte bands, represented by the different coloured dyes. Note that the yellow band is the earliest eluting analyte; it moves the fastest and has a broader band moving faster in the column. The blue band is more strongly retained; it has a narrower, focused analyte band and moves slowest in the column. **B:** differential chromatographic retention of analytes by normal-phase and reversed-phase chromatography. Yellow, red and blue bands represent polar, moderately polar and non-polar analytes respectively. In normal-phase chromatography polar analytes are more strongly retained. In reversed-phase chromatography non-polar analytes are more strongly retained. Gradient elution is commonly employed to overcome the phenomenon whereby well-retained peaks are adequately separated (narrow, focused bands) but less well-retained analytes are subject to band broadening in the column, resulting in broader chromatographic peaks with poor separation. In gradient elution, the solvent composition (water *versus* organic solvent) is altered systematically over time. Adapted from (129).

1.2.4 Metabolomics in the clinical laboratory and its potential in AKU

One application of metabolomics in which its vast potential is becoming more widely recognised is in the medical field. Recent years have seen major interest in metabolomics as an approach to discover new biomarkers for a range of clinical purposes, including those that can predict drug efficacy, disease progression (prognosis) and drug toxicity (77,130–133). In coming years, metabolomics is expected to become a major component of a new strategy to provide precision medicine approaches which aim to tailor medical care based on individual patient metabolome profiles (134).

Metabolomics has clear potential in the clinical management of inborn errors of metabolism (135,136). Despite significant advances in the diagnosis and treatment of some inherited metabolic diseases over recent years, monitoring of disease progression and susceptibility is challenging. Routine clinical biochemistry assays for monitoring inborn errors of metabolism are generally based on the measurement of individual metabolites. With reference to AKU, our laboratory has developed assays using triple-quadrupole MS for measuring key metabolites of the tyrosine metabolic pathway, including HGA and tyrosine, in serum (137) and urine (138) with high quantitative accuracy and precision. The assays have been employed in various clinical trials of nitisinone in AKU (11,139) and continue to be indispensable for monitoring disease progression and response to nitisinone (*i.e.* monitoring concentration of the consequently increased tyrosine) in patients attending the NAC (12,140). A limitation of these methods is that they are limited to specific metabolites within the tyrosine pathway. Metabolomics can offer a wider assessment of metabolism in AKU. This approach will enable investigation of

the wider metabolic consequences of the disease process, including ochronosis, and response to nitisinone.

Metabolomics in AKU has the potential to assess the biochemical impact of ochronosis. Given the early-onset and inevitable OA in AKU, there is a clear opportunity to identify novel metabolite markers of ochronotic osteoarthropathy (see section 1.3.3), for example early markers of cartilage matrix degradation breakdown that could be detected in urine as with bone resorption products. Currently there are no biomarkers that accurately predict onset of OA.

Metabolomic analysis also enables wider assessment of the metabolic impact of nitisinone. This is important for several reasons. First is that the wider consequences of hyper-tyrosinaemia following nitisinone treatment are not known (141). There are several reports of AKU patients on nitisinone presenting with dermal toxicity and ocular dendritiform corneal keratopathy attributable to tyrosine deposition (14,142–144). In these cases, nitisinone treatment had to be withdrawn. The factors leading to corneal keratopathy are unknown, and it is important to identify which patients are particularly at risk of this side effect of nitisinone. Second, metabolomics enables assessment of potential 'off-target' effects of nitisinone treatment beyond its desired 'targeted' effect of inhibiting HGD. The potential for nitisinone to act on other enzymes is not known, and the off-target effects of any drug are essential when considering its suitability in treating a patient (145).

Table 1.04. Rationale for metabolomics in AKU.

Paradigm	Aim of metabolomic analyses
Untreated AKU	Reveal mechanisms of disease
	Correlation with other clinical phenotyping data
	Detect early biomarker of OA
	Wider biochemical consequences of HGD deficiency
	Discovery of AKU sub-phenotypes
Nitisinone-treated AKU	Wider consequences of hyper-tyrosinaemia
	Potential off-target effects of nitisinone
	Biomarker for predicting corneal keratopathy

1.3   The nature of ochronotic pigment in AKU

1.3.1 Structure of ochronotic pigment: classical view

A fundamental question in understanding the process of ochronosis and the development of therapeutic interventions aimed at its prevention or potential reversal concerns the structure of ochronotic pigment itself. The prevailing view on the mechanism by which HGA produces ochronotic pigment largely comes from work carried out by Zannoni and colleagues in the 1960's. These authors stated that guinea pig cartilage and skin contain enzymes, namely HGA polyphenol oxidases, which were shown to catalyse the *in vitro* oxidation of HGA, which is colourless, into a dark, ochronotic-like pigment (146). BQA was demonstrated as an intermediate in the *in vitro* enzymatic oxidation of HGA, and it was proposed that polymerisation of BQA forms ochronotic

pigment (Figure 1.01B) (146,147). However, it is important to note the lack of evidence that polyphenol oxidase enzymes are expressed in human or other mammals (148). Enzymes that can oxidise HGA, HGA-oxidases, are observed in various species of bacteria known to produce pyomelanin pigment derived from HGA (42,149). In these species, HGA-derived pyomelanin is thought to serve various adaptive functions including resistance to environmental stress such as UV light and oxidising agents.

Subsequent research showed that HGA oxidation can occur non-enzymatically between pH 6.8 and 9.5 in the presence of oxygen (150). This phenomenon is observed in the classic sign of AKU; slow, spontaneous darkening of urine. Adding alkali to urine instantly turns urine black, and subsequent acidification does not alter the colour, suggesting a potentially irreversible change. The presence of BQA in oxidized HGA solutions has been confirmed by chromatography (151) and UV-vis spectroscopy (152).

It is still widely-stated in the literature that oxidation of HGA to benzoquinone intermediates occurs spontaneously in AKU and results in a polymeric pigment structure.

1.3.2 Structure of ochronotic pigment: emerging view

Given the general assumption in the literature that ochronotic pigment is a product of HGA oxidative polymerisation, it seems pertinent here to define the term 'polymer'. The accepted definition of a polymer is a large molecule, or macromolecule, composed of multiple repeating subunits of a relatively lower molecular weight monomer (153). The classic concept of a polymer is that the monomer subunits are covalently bound (154).

41

Roberts *et al.* (42) question the widely-held assumption that ochronotic pigment is polymeric. These authors cite the lack of conclusive evidence for this in the literature and make the point that a polymeric structure is not necessarily required to produce the visual properties of a dark pigment, as there are numerous examples of low-molecular weight biological pigments. More recent analysis of synthetically-derived ochronotic pigment solutions using size exclusion chromatography suggested a molecular weight greater than HGA, as indicated by a peak at shorter RT. However, visually pigmented solutions could be formed from HGA (over a shorter period of time; 10 days as opposed to 2 years) without evidence of the peak corresponding to ochronotic pigment and no observed decrease in the HGA peak (155). This suggests that the visual darkening of the solution due to increased pH could be due to presence of the low-molecular weight oxidised form of HGA over a relatively shorter period.

Ochronotic pigment has been referred to as 'melanin-like' in the literature (42); largely because melanin is another dark biological pigment derived from tyrosine and classically considered to be formed by polymerisation. However, there is a growing body of evidence that melanin and ochronotic pigment are not covalently bound structures, and that their macromolecules do not comprise regularly repeating monomers. A recent study reported data from physicochemical analyses on the pyomelanin pigment produced in the bacteria *Rubrivivax benzoatilyticus* (strain JA2) (156). This bacteria mirrors the conditions of ochronotic pigment production in AKU; absence of the *HGD* gene causes accumulation of HGA, resulting in a brown pigment under aerobic conditions and in the presence of phenylalanine (157). FTIR spectroscopy

showed a range of band stretching vibrations indicating various chemical groups (aromatic and aliphatic C-H stretches, phenolic C-O stretches, aromatic ring C=C bonds, and C=O stretches due to –COOH groups) characteristic of a pigment structure derived from HGA. X-ray diffraction spectra of the pigment showed similar characteristics to that of melanin, with broad diffraction indicating an amorphous compound structure. The absorbance spectra showed a broad band at wavelengths 280-350 nm, with increased general absorbance across the UV-visual range without distinct peaks. Similar absorbance properties have been reported previously for ochronotic pigment (42,155) and also melanin (eumelanin derived from 5,6,-dihydroxyindole-2-carboxylic acid); (158), and are thought to reflect a chemically heterogeneous structure of oligomers formed by a range of different bonding mechanisms, also referred to as chemical disorder (159). These absorbance properties might also account for the physical appearance of some pigmented substances, which in the case of melanin is thought to provide its physiologically important optical characteristics, *i.e.* its ability to absorb UV light (42,160). In contrast, specific, well-defined chemical signatures have been obtained for BQA, the proposed low molecular weight oxidation product of HGA. Specific visual-range absorbance peaks corresponding to BQA have been reported in a series of publications by Tokuhara and colleagues. These peaks were observed at 406 and 430 nm from analysis of solutions of HGA or AKU urine following alkalinisation (161,162), although other researchers have been unable to replicate these findings (42). Specific BQA signals were also recently reported by Tokuhara *et al.* (162) from LC-QTOF-MS and NMR analyses.

Computational analyses of eumelanin support a structure formed by stacked eumelanin protomolecules with random-like arrangement; in other words loosely-bound aggregates as opposed to a covalent polymer (160). This aggregate structure of melanin is further supported by mass spectrometric and spectrophotometric analyses indicating a possible formation mechanism by self-aggregation of L-dopa by a combination of non-covalent mechanisms including hydrogen bonds, π- π stacking and ionic bonds (163). The same aggregation was observed for other structurally similar catecholamines which, like HGA, are derived from tyrosine. Together, these more recent data are inconsistent with the idea that ochronotic pigment and melanin are polymeric structures comprised of regularly repeating units with distinct chemical signatures. The term polymer therefore does not appear to accurately describe the chemical nature of ochronotic pigment or melanin, as currently understood.

1.3.3 *Molecular interaction of ochronotic pigment with cartilage matrix*

Aside from the chemical structure of ochronotic pigment itself, the molecular mechanism through which HGA interacts with the cartilage matrix is also unknown. It is also not known, for example, whether the initial binding occurs as HGA, the oxidised intermediate BQA or ochronotic pigment, or if there is a specific binding site within individual proteins or the extracellular matrix. HGA is an inherently water-soluble substance, and should in principle be easily removed from body tissues simply by circulation, unless it is specifically retained by the tissue. It is currently believed that existing tissue ochronosis is not fully reversible. There are indications of partial reversal of pigment following long-term nitisinone therapy (140), but the mechanism of such reversal is currently unknown. It is not known if the ochronotic process is

dynamic (*i.e.* formation and removal of pigment co-existing), although macrophages with pigment has been described in case reports (164–166).

Collagen, specifically type-II, the predominant protein in cartilage, is known to have a highly hierarchical fibrillar structure. First, three chains of type-II collagen form a staggered, hydrogen-bonded homotrimer; the characteristic collagen triple helix. After secretion into the extracellular space, the triple helices self-assemble and crosslink with the aid of other extracellular matrix components (167,168) to give fibrils with a distinctive 67-nm banding pattern. In cartilage, these fibrils further twist together to form aligned fibres.

Ultrastructural examination of pigmented cartilage showed that initial pigmentation is closely associated with the periodicity of collagen fibres. A periodic banding pattern of ochronotic pigment was observed on individual collagen fibres, with very early pigmentation appearing as small granules on the surface of fibres (Figure 1.07A) (169). These findings suggest that a nucleation-like event underlies ochronosis, where initial granular deposits on individual collagen fibres is followed by further rapid pigment deposition (9). The data also suggest that collagen fibrils provide specific binding sites for pigment.

Close observation of tissues obtained from AKU patients (30,66) and mice (50,66) and *in vitro* (62) models of AKU show that cartilage is initially resistant to pigmentation. It is proposed that biomechanical and biochemical changes, such as those that occur in cartilage as part of the natural ageing process, render tissues susceptible to ochronosis. The exposed collagen hypothesis theorises that binding sites become available for HGA following the loss of

protective molecules such as proteoglycans and glycosaminoglycans (GAGs), as illustrated in Figure 1.07B (9). Consistent with this hypothesis, AKU cartilage is shown to have lower levels of extractable GAGs and oligomeric matrix protein than osteoarthritic and non-osteoarthritic cartilage (170). There is also evidence that the structure and maturity of the collagen matrix can influence pigment deposition. Newly-synthesised aberrant matrix proteins in scar tissue also appear to pigment rapidly, as reported in the case of a mediastinal mass from an AKU patient (171). Solid-state NMR (ssNMR) analysis of AKU articular joint cartilage also observed spectra indicative of marked collagen disorder at the atomic level (172), further supporting the idea that disruption to the collagen matrix supports ochronotic deposition. In support of this idea, wider structural damage of collagen organisation has been observed in cartilage collagen, in OA, for example the aligned fibres of type-II collagen have been shown to split apart into smaller diameters in osteoarthritic knee and hip joints (173).
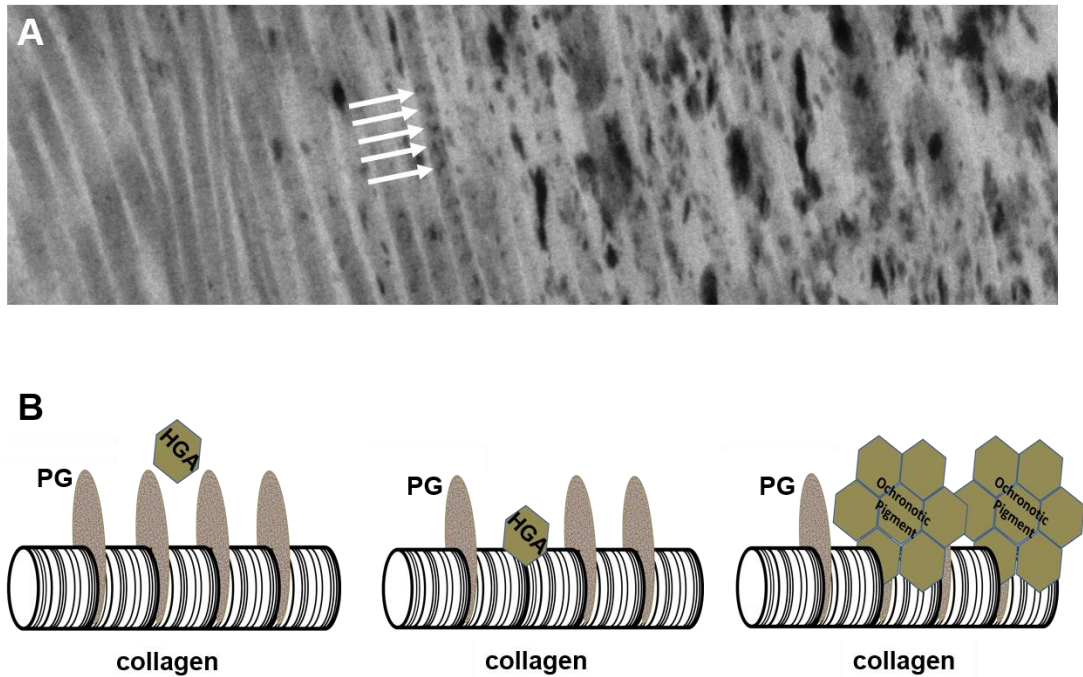
Figure 1.07. Interaction between ochronotic pigment and cartilage collagen. A: transmission electron microscopy image of ochronotic ligamentous capsule. Collagen fibres in longitudinal section show a distinct electron-dense pigment on their surface. Numerous pigment shards can be seen on single fibres. Not all fibres present with pigment deposition. Gradient of pigmentation can be seen running left (no pigment on fibres) to right (large electron-dense shards replacing fibres). Arrows indicate a distinct periodic binding pattern associated with pigment granules on a single fibre. Source: (16). **B:** schematic representation of the exposed collagen hypothesis of ochronosis. Initially, collagen fibrils are resistant to binding of homogentisic acid (HGA) or its derivatives. Following loss or breakdown of specific protective matrix constituents including proteoglycans (PG), binding sites become available for HGA-associated compounds. It is proposed that the initial binding event initiates ochronosis and that the process of widespread joint pigmentation occurs over time. Figure adapted from (9).

In terms of the exact HGA-associated species that initially binds to cartilage, it is thought that the oxidised BQA intermediate is more reactive than HGA (174). Coloured solutions of oxidised HGA are known to show more irreversible binding to hide powder collagen compared to colourless solutions of non-oxidised HGA (175,176). Chemical investigations of BQA using spectrophotometric methods has shown that it binds irreversibly with free amino groups (177,178). Additionally, irreversible binding of BQA occurs with albumin and homogenates of skin and cartilage (179). The current hypotheses of the binding of collagen to HGA or derived species therefore involve BQA to a large degree, though it has been shown that oxidized (coloured) and fresh (colourless) HGA solutions show highly similar infrared spectra (180), suggesting strongly that the conversion of HGA to BQA is incomplete, and that the HGA hydroquinone functionality remains substantial even in coloured solutions of oxidized HGA. While HGA is colourless, the air oxidation of urine from AKU patients and the joint cartilage degradation are both accompanied by a striking colour change from colourless/white to brownish black. The presence of BQA or related oxidised species is inferred from the colour of ochronotic cartilage, though whether the degree of conversion is complete in these solutions and *in vivo* is currently unknown.

1.3.4 Detection of ochronosis

Gross ochronosis is easily visible as dark pigment in tissues such as the cartilage of the ears and sclera of the eyes. Photographs of the eyes and ears have been used to follow the pigmentation process, both in terms of understanding the natural history of AKU as well as to study the effect of

reducing HGA concentrations on ochronotic pigment. This may be the most efficacious way to follow pigment change (140).

However, to detect small amounts of pigment, more sensitive techniques are required. Such techniques can employ the property of HGA as a reducing agent and can be used in tissues *in vitro* and *ex vivo*; one such approach employs Schmorl's stain, a ferricyanide reduction method (Figure 1.08) to detect microscopic ochronotic pigment (62); this supports the idea that ochronotic pigment originates from HGA. Biopsy of tissues such as ear cartilage reveals ochronotic pigment even when it is not visible externally visually through the intact skin (27).

Investigations such as arthroscopy and bronchoscopy can reveal ochronosis and suggest diagnosis of AKU for the first time. Diagnosis has also been made following the observation of black aorta at open heart surgery for aortic valve replacement (181). Direct detection of ochronotic pigment *in vivo* is possible by Raman spectroscopy. Such a technique has been validated in *ex vivo* tissue (182) and is being adapted as an *in vivo* technique using ear cartilage and Achilles tendon (183).
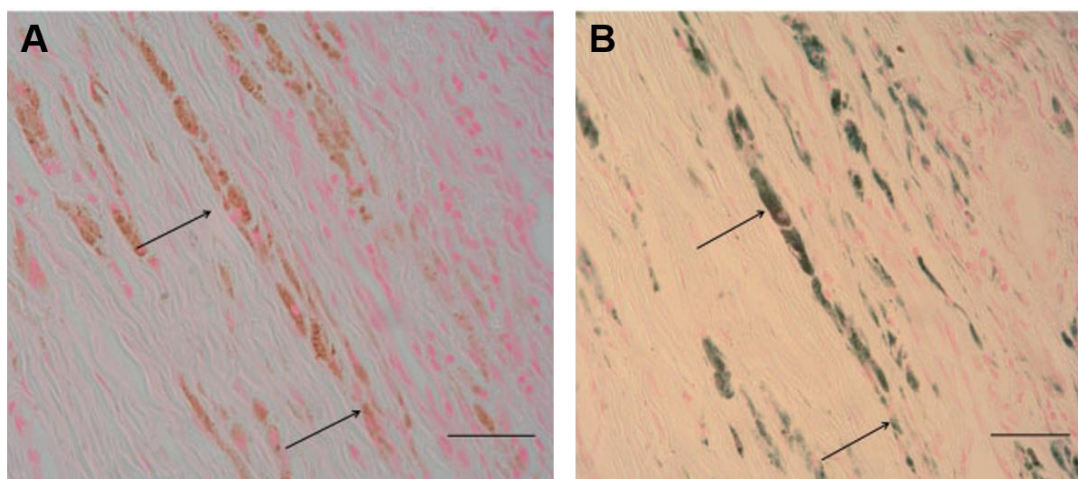
Figure 1.08. Histological detection of ochronotic pigment. A: ochronotic pigment in the joint capsule of the knee of a patient with AKU. Pigment deposits are seen as brown granules in the cartilage extracellular matrix and within individual fibroblasts. Section stained with nuclear fast red. **B:** near serial section of the joint capsule with Schmorl's stain. Ochronotic deposits are stained green. Bar: 50 µm. Source:

Ochronosis can be observed in *ex vivo* samples, especially from joints, and these can vary in the extent and degree of pigmentation. The earliest pigment is articular cartilage is seen in calcified cartilage cells at the junction between the calcified cartilage and subchondral bone (184). Similarly, ochronotic pigment is found in intervertebral discs and adjacent articular cartilage (23).

It would be an advantage to quantify ochronotic pigment *in vivo* but such a technique is not available at present. Availability of techniques to monitor changes in whole body pigment, increase or decrease or no change over time, would be highly informative.

1.4 Analytical techniques for studying ochronotic pigment: mass spectrometry

In addition to metabolomics, MS platforms such as LC-QTOF-MS can be applied for identifying the chemical structures of unknown molecules such as ochronotic pigment. Aside from the more general attributes of the technology that are discussed above (*i.e.* high sensitivity, high-resolution mass accuracy, non-targeted data acquisition), this technique provides additional structural information on compounds by generation of compound fragments, or 'product' ions from the precursor ion which are recorded and combined into a spectrum (MS/MS or MS^2). In MS/MS analyses, a specified collision energy is applied to ions. 'Collision-induced dissociation' (CID) of compounds (this occurs in the collision cell component of a QTOF; Figure 1.04) in this way enables the chemical structure to be deciphered based on the compound fragments or sub-structures generated. MS/MS can be performed by a targeted (Figure 1.09) or non-targeted approach. In targeted MS/MS analyses a specific precursor ion (or ions) is selected for fragmentation (based on AM with/without RT specified – selection by quadrupole filter in QTOF; Figure 1.04), whereas in non-targeted analyses fragmentation can be performed on any precursor ion that exceeds a given abundance threshold within a specified time frame.

Compound identification in MS/MS analyses is often performed by matching the observed fragmentation spectra with data from compound spectral libraries (92,107,185). When there are no data for compounds in spectral databases or libraries, as is most likely the case for ochronotic pigment, this approach has to be performed *de novo*. The *de novo* approach is more challenging, but important clues to the compound's sub-structure can still be gained from the

fragment ions generated. Furthermore, a range of computational tools is available to aid the *de novo* identification of unknown compound fragment ions (186,187).
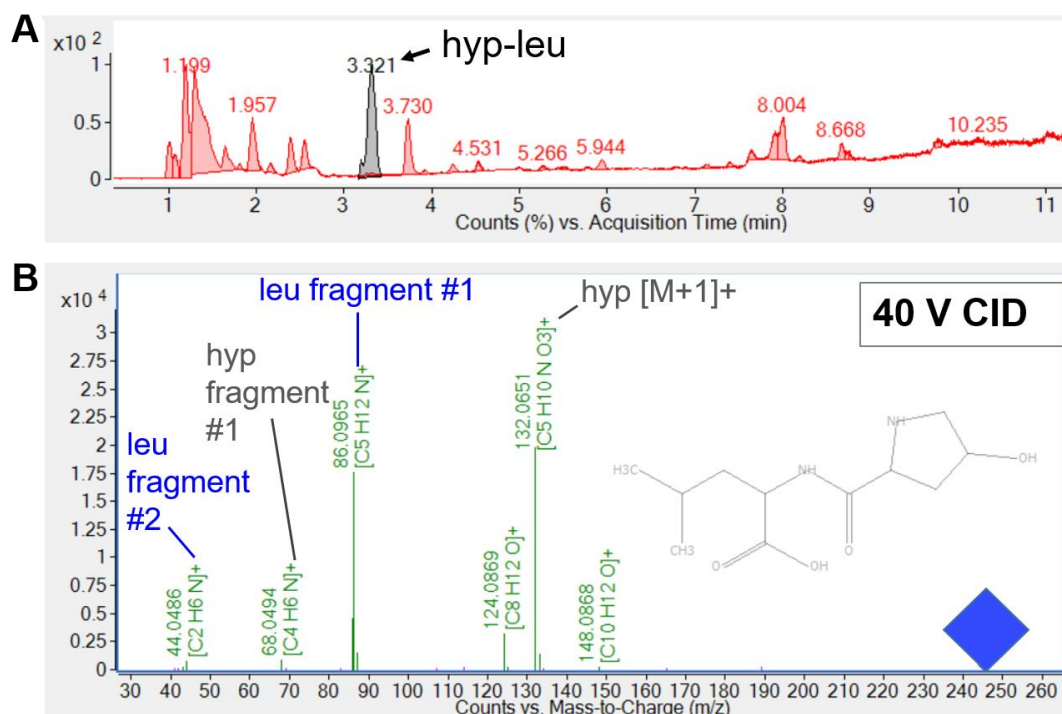
Figure 1.09. MS/MS for chemical structure elucidation. A: selection of precursor ion by accurate mass and RT; dipeptide hydroxyprolyl-leucine (hyp-leu) in this example of targeted MS/MS. **B:** spectrum acquired from collision induced dissociation (CID) at 40 V. Blue diamond represents the fragmented precursor ion. Generation of the product fragment ions confirms identification of this compound as hyp-leu. The annotated product ions are hyp (dissociated from the dipeptide and known fragments of hyp and leu). Source: unpublished data acquired by LC-QTOF-MS analysis of culture media from an *in vitro* osteoclast bone resorption assay. The hyp-leu dipeptide is a previously uncharacterised product of bone type-I collagen degradation.

1.5 Analytical techniques for studying ochronotic pigment: NMR spectroscopy

NMR spectroscopy is an analytical technique widely used in chemistry, structural biology and material science to study structure at the molecular level, by observing local magnetic fields around atomic nuclei. Not all nuclei can be observed by NMR; in biological applications, frequently observed NMR-sensitive nuclei are ^{1}H, ^{13}C, ^{15}N, ^{19}F and ^{31}P. While ^{1}H, ^{19}F and ^{31}P are all highly abundant (>99%) on earth, ^{13}C and ^{15}N are only present at approximately 1% and 0.4% (188), respectively. In the NMR experiments employed to study ochronotic pigment and cartilage samples here, analyses were based on ^{1}H and ^{13}C nuclei (Chapters 6 and 7). ^{1}H is also referred to as a proton NMR as the hydrogen nucleus consists of one proton only.

1.5.1 The vector model

The vector model is useful for explaining the basic concepts of NMR. The fundamental idea of this model is the net alignment of NMR-sensitive nuclei in a given direction, referred to as the magnetisation vector. The magnetisation vector is zero when the sample is not subject to an external magnetic field, but aligned with the applied field in an external magnetic field, such as that generated by the NMR spectrometer magnet (represented by the *z*-axis in Figure 1.10A). In NMR experiments, the magnetisation vector generally needs to be tipped away from the applied magnetic field. Such tipping is generally achieved by applying a resonant radiofrequency (RF) pulse perpendicular to the external magnetic field. Figure 1.10B shows how applying an RF pulse along the *x*-axis moves the alignment of the magnetisation vector away from

the direction of the magnetic field (z-axis). As a result, the magnetisation vector rotates, or precesses, around the direction of the magnetic field at an angle (β) in a cone-like motion (189).

The precession frequency is characteristic of a given nucleus and fundamental to the final NMR spectrum obtained. For example, at a field strength of 9.4 T, ^{13}C precesses at 100 MHz while ^{1}H precesses at 400 MHz. At a higher field strength of 18.8 T, ^{13}C precesses at 200 MHz while ^{1}H precesses at 800 MHz.

NMR data detection is enabled by placing a coil of wire around the sample, as shown in the direction of the x axis in Figure 1.10C. The magnetisation vector cuts the coil and induces a current which is then amplified and recorded as NMR data. The precession of the magnetisation results in a current that oscillates at the Larmor frequency over time; this voltage-vs-time NMR data is the free induction decay (FID). Application of a mathematical process known as Fourier transformation results in the conversion of the FID data from the time domain into the frequency domain (189). Plotting the transformed output (now in units of frequency, *e.g.* Hz) then gives us the NMR spectrum (Figure 1.10D).
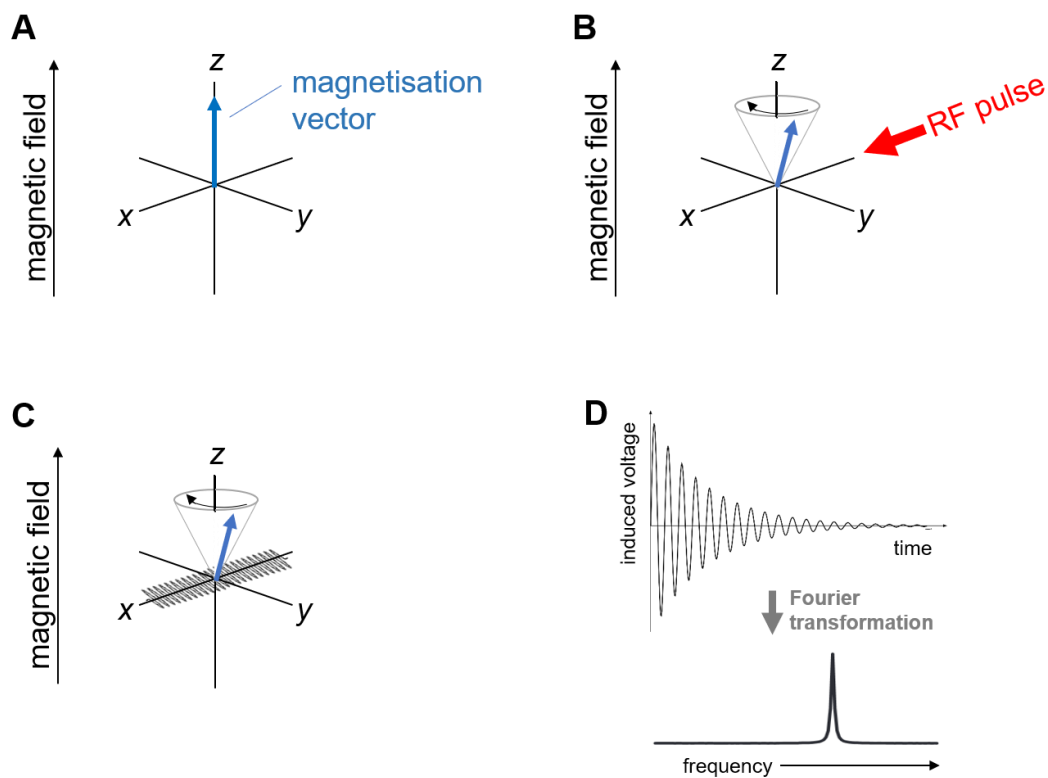
Figure 1.10. Fundamental concepts of NMR spectroscopy according to the vector model. A: at equilibrium a sample has net magnetisation aligned in the direction of an external magnetic field; here the magnetisation vector is aligned with the *z*-axis. **B:** upon application of a radiofrequency (RF) pulse, the magnetisation vector is tilted away from the *z*-axis and precesses around the external magnetic field at an angle in a cone-like motion. The precessional frequency is characteristic of a given NMR-sensitive nuclei. **C:** the precessing magnetisation is detected by a coil, shown here along the *x*-axis. **D:** the free induction decay (FID) NMR signal is converted from the time domain to the frequency domain by Fourier transformation, resulting in the final NMR spectrum. Figures adapted from (189).

The resulting NMR spectrum is a series of chemical shift signals, conventionally plotted in parts per million (ppm). ^1H and ^{13}C chemical shift values are indicative of particular chemical structures, therefore the NMR spectra can be used for structural determination of a molecule. This is illustrated in Figures 1.11A and 1.11B, for ^1H and ^{13}C chemical shifts, respectively.

A

carboxylic acid · aromatic · vinylic · ether · benzylic · 3' alkyl

aldehyde · phenolic · alkyl halide · ketone · 2' alkyl · allylic

amide · vinylic · alcohol · acetyl-inic · 1' alkyl

^1H chemical shift (δ, ppm)

alcohol hydroxyl (0.5-5.0) or amino (1.0-4.0) · TMS

13.0 12.5 12.0 11.5 11.0 10.5 10.0 9.5 9.0 8.5 8.0 7.5 7.0 6.5 6.0 5.5 5.0 4.5 4.0 3.5 3.0 2.5 2.0 1.5 1.0 0.5 0

B

ketone · carboxylic acid · aromatic carbons · alkyne · 1' alkyl

aldehyde · ester · nitrile · alcohol · 2' alkyl

amide · alkene · ether or ester · 3' alkyl

^{13}C chemical shift (δ, ppm)

acyl chloride · nitro compound · alkyl halide · TMS

220 210 200 190 180 170 160 150 140 130 120 110 100 90 80 70 60 50 40 30 20 10 0

Figure 1.11. Typical ^1H (A) and ^{13}C (B) NMR chemical shift values for various chemical structures. Note that these are typical values only; exact values depend on experimental parameters such as solvent composition, temperature and presence of other functional groups. Note that TMS (trimethylsilane) has ^1H and ^{13}C chemical shift values of 0; all other chemical shifts are determined relative to it. Figures adapted from (190).

1.5.2 Two-dimensional NMR

Two-dimensional (2D) NMR is an approach used to probe molecular structures in greater detail. One-dimensional spectra show one frequency dimension and determine the intensity of one type of nuclei in the sample. 2D NMR provides a second frequency dimension, which enables different nuclei to be correlated to each other. The resulting signals in the final 2D spectra represent intensity values, not for single nuclei as in 1D NMR, but for pairs of frequency variables arising from nuclei whose spins are in some way related to each other. 2D NMR spectra are conventionally plotted in the form of contour plots, in which intensity is indicated by contour lines, analogous to the indication of height relief in a topographical map (Figure 1.12). The advantage of the second frequency dimension is that it can make it possible to separate signals that overlap in conventional 1D NMR; see peaks indicated in green in Figure 1.12 (189).

HETCOR (HETeronuclear CORrelation coupling) is a specific 2D NMR experiment discussed in Chapter 7. This experiment is used to detect carbons that are in close proximity (within 5 Å) of hydrogens. These ^{13}C – ^{1}H couplings are represented by the peaks plotted on a 2D spectrum in which the ^{13}C and ^{1}H chemical shifts are plotted against each other, as shown by the schematic representation in Figure 1.12. In ssNMR, the strong ^{1}H – ^{1}H couplings can be mitigated with the frequency-switched Lee-Goldburg (FSLG) sequence to improve resolution in the ^{1}H dimension (191).
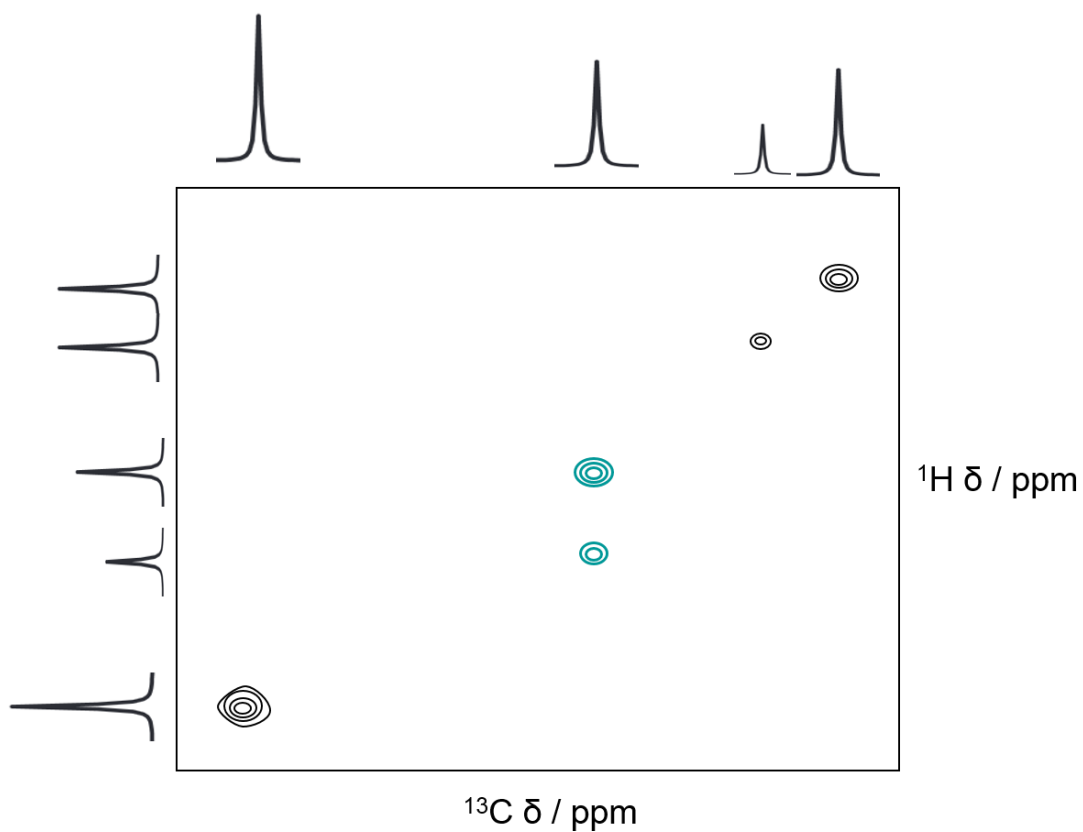
Figure 1.12. Two-dimensional (2D) NMR. Schematic representation of an NMR spectrum obtained from a 2D HETCOR (HETeronuclear CORrelation coupling) experiment. Signals are plotted as a function of two frequency domains, the ^{13}C and ^{1}H chemical shifts (δ), and represent the coupling between the spins of carbons and proximal hydrogens within a molecule. Highlighted in green are two peaks that would overlap in the ^{13}C dimension but are resolved in the ^{1}H dimension, indicating ^{1}H coupling to different ^{13}C nuclei within the sample.

1.5.3 Solid-state NMR and magic angle spinning

In contrast to conventional NMR in which samples are analysed in the solution state, ssNMR deals with the analysis of samples in the solid state. This NMR approach is particularly attractive for chemical analysis of ochronotic tissues in their native, intact form. However, in ssNMR additional measures are required to achieve high-resolution spectra. In solution-state NMR, the many random orientations interconvert rapidly due to 'tumbling'. In ssNMR, the random orientations of the molecular system remain unchanged throughout the NMR experiment in the static state. These random molecular orientations give different spectral frequencies, resulting in much broader lines in the final NMR spectra ("chemical shift anisotropy") and many overlapping signals. Moreover, heteronuclear dipolar couplings (^{1}H-^{13}C) are often several orders of magnitude larger in the solid-state compared to the solution-state, again due to the absence of tumbling and averaging. Magic angle spinning (MAS) is a technique routinely employed in ssNMR to improve spectral resolution by alleviating and even removing the effects of line broadening attributable to chemical shift anisotropy and heteronuclear dipolar coupling (192). In MAS, the sample is spun rapidly in a cylindrical rotor at the 'magic angle' (54.74°) relative to the direction of the applied magnetic field. The rapid rotation enables the undesirable interactions to be averaged out, giving rise to higher resolution ssNMR signals.

1.5.4 Cross polarisation

Cross polarisation (CP) is another useful technique in ssNMR spectroscopy. CP can be employed to enhance the signal of relatively low-abundance nuclei

such as ^{13}C. This approach is particularly applicable in NMR analyses of non-isotopically enriched material, as the natural abundance of ^{13}C is 1.1%. CP is achieved by a specific pulse sequence designed to transfer the larger magnetisation of ^1H to ^{13}C (^1H has greater magnetic sensitivity). Another advantage of this experiment is that it can be repeated more frequently (usually every few seconds) within a given time frame, which is important for signal enhancement. ^{13}C has low abundance, and in non-CP NMR experiments it is necessary to wait potentially minutes for the ^{13}C spin magnetisation to recover before the experiment can be repeated. ^1H has high abundance and is highly coupled to other ^1H nuclei, which means that their spin relaxes more efficiently. Using ^1H to generate the magnetisation means that the relaxation of ^1H dominates over that of ^{13}C and the delay times between experiments can be reduced (192).

1.5.5   *Dynamic nuclear polarisation*

Dynamic nuclear polarisation (DNP) can be used to enhance sensitivity of non-isotopically enriched organic and biological samples in MAS ssNMR experiments (193). In DNP, signal enhancement of one to two orders of magnitude can be achieved by transfer of the much larger polarisation of electron spins to ^1H within the sample. The electrons come from a radical solution which is applied to the sample. Transfer of magnetisation by DNP requires microwave irradiation, which is generated by a gyrotron. To ensure efficient transfer, DNP experiments are carried out at low temperature (*i.e.* 100 K in the experiments detailed in Chapter 7).

1.6 Aims and objectives

This thesis aims to explore two major gaps in the current understanding of AKU. First are the consequences that treatment with nitisinone have on the disease in terms of wider metabolism, and how these relate to the metabolic profile of untreated AKU. Second is the chemical nature of ochronotic pigment; its chemical structure and how its incorporation into cartilage results in severe early-onset osteoarthropathy. The specific objectives are as follows:

1. Establishment of a comprehensive strategy for metabolomic analysis by LC-QTOF-MS

2. Employment of the metabolomic analytical technique developed to explore:

   a) the consequences of nitisinone treatment on wider metabolism in patients and mice with AKU

   b) the consequences of *HGD* deficiency on wider metabolism in untreated mice from a newly-established model of AKU

3. Structural analyses of ochronotic pigment formed *in vitro* (derived from HGA standard) and *in vivo* (ochronotic cartilage) using a range of chemical analytical techniques including MS and NMR

2.0 DEVELOPMENT OF A COMPREHENSIVE LC-QTOF-MS STRATEGY FOR METABOLIC PROFILING AND TARGETED COMPOUND IDENTIFICATION

Published work resulting from content included in this chapter (see Supporting Material):

Norman BP, Davison AS, Ross GA, Milan AM, Hughes AT, Sutherland H, Jarvis JC, Roberts NB, Gallagher JA, Ranganath LR. A comprehensive LC-QTOF-MS metabolic phenotyping strategy: Application to alkaptonuria. Clin Chem. 2019;65(4):530–9.

Norman BP, Davison A, Ross GA, Milan AM, Hughes AT, Sutherland H, Jarvis JC, Roberts NB, Gallagher JA, Ranganath L. Three accurate mass retention time (AMRT) databases generated from IROA Technologies Metabolite Library of Standards by LC-QTOF-MS analysis. figshare. Collection. 2019. Available from:
http://doi.org/10.6084/m9.figshare.c.4378235.v2

The above works report the metabolic profiling strategy developed and validated by BPN, and described in this chapter, in addition to the resulting AMRT databases. BPN was the main contributor to the laboratory analyses performed and the writing of manuscripts/web page content for these publications.

Davison AS, Norman BP, Ross GA, Hughes AT, Khedr M, Milan AM, Gallagher JA, Ranganath LR. Evaluation of the serum metabolome of patients with alkaptonuria before and after two years of treatment with nitisinone using LC-QTOF-MS. JIMD Rep. 2019;48(1):67–74.

The above works resulted from the application of the metabolic profiling strategy developed by BPN and reported in this chapter. BPN contributed to the experiment design, laboratory analyses and manuscript writing for these publications.

2.1 Introduction

One strategy for identification of unknown chemical entities generated by LC-MS is by reference to publicly available spectral libraries, which have become generally accepted as indispensable resources to researchers in the field of metabolomics (194–198). Although the human metabolome, for example, is not yet fully characterised, significant efforts are underway to catalogue analytical data acquired from tens of thousands of biologically relevant chemical standards according to mass spectra and MS/MS spectral fragmentation patterns in addition to data from coupled techniques such as RT (LC), retention indexes (GC) and collisional cross sections (ion mobility). For example, the recently updated 2018 version of the Human Metabolome Database (199) now contains entries for more than 115,000 compounds, containing information on human metabolites including their biological function, tissue and biospecimen location, physiological concentration, disease associations, metabolic pathways, experimental spectra and predicted MS/MS spectra. A range of other mass spectral libraries of small molecule metabolites is commonly employed by researchers in metabolomics; as of 2016 the most widely used spectral libraries were METLIN, MassBank, KEGG, NIST MS, ChemSpider, LipidMaps and PubChem (200). Together, these databases cover a broad range of volatile and non-volatile low molecular weight compounds from different chemical classes including both those of endogenous and exogenous origin.

A major limitation of compound identification using publicly available compound databases is that the data were acquired by different laboratories employing different analytical parameters and techniques. These include

different MS ionisation techniques [*e.g.* ESI, MALDI, atmospheric-pressure chemical ionisation (APCI)] and different analytical parameters applied to coupled techniques such as liquid and GC, capillary electrophoresis *etc.* To resolve this issue, an in-house AM / RT (AMRT) database was generated from the analysis of 619 small molecule metabolite standards with LC-QTOF-MS for the purpose of predicting the identity of unknown entities acquired under identical analytical conditions based on orthogonal chemical properties (*i.e.* AM and RT). It is shown that coupling three separate but complementary LC techniques to MS maximised coverage of the metabolite standards, which spanned a range in molecular weight (45-1354 Da) and chemical classes including carboxylic acids, amino acids, biogenic amines, nucleotides, carbohydrates, fatty acids, lipids, steroids and hormones; overall covering a broad range of primary metabolism.

It is then shown that the targeted LC-QTOF-MS strategy can be applied to identify unknown compounds present in complex biological matrices such as urine.

2.2 LC-QTOF-MS analysis of metabolite standards from the Sigma IROA MSMLS suite

2.2.1 *Metabolite standards library preparation*

619 metabolite standards (molecular weight range: 45-1354 Da) were from the IROA MS metabolite library of standards (MSMLS: Sigma Aldrich, UK) contained in seven 1.2 ml polypropylene deepwell plates, with each individual compound at 5 µg in an individual well (>95% purity). Plates were stored at

-80°C. MSMLS compounds are accessible via the following link: https://www.sigmaaldrich.com/catalog/product/sigma/msmls.

Prior to analysis, plates were thawed at room temperature and compounds reconstituted in 15 µL methanol (LC/MS grade, Sigma Aldrich) followed by 285 µL deionised water (purified in-house via DIRECT-Q 3UV Millipore water purification system). Plates 1-5 were left to stand for at least 1 h at room temperature following addition of water and methanol. Plates 6 and 7, which contained primarily lipid-like compounds, were left to stand for 2 h at room temperature following addition of methanol to ensure solubilisation; water was then added after this period. Plates were then agitated on a plate shaker (MTS 2/4m IKA, Germany) at 600 rpm for 10 minutes. Compounds were pooled across the rows of each plate (12 wells per row). 20 µL was taken from each well and pooled into one well so that 12 compounds were analysed per injection. Fifty-six injections were performed in positive and negative ionisation polarity across all plates; in total 112 injections for each LC-QTOF-MS method (Figure 2.01). Each pool had a total volume of 240 µL with each compound at 0.14 mg/mL (1-31 µmol/L).

It was noted that some rows contained isobaric metabolites; these compounds were added to new pools containing 3-6 compounds of different molecular weight. D-frucotse, D-galactose, L-sorbose, alpha-D-glucose, melibiose and palatinose were injected individually; re-pooling was impractical as they share the same molecular weight. All isobaric compounds were also analysed at 0.14 mg/mL.
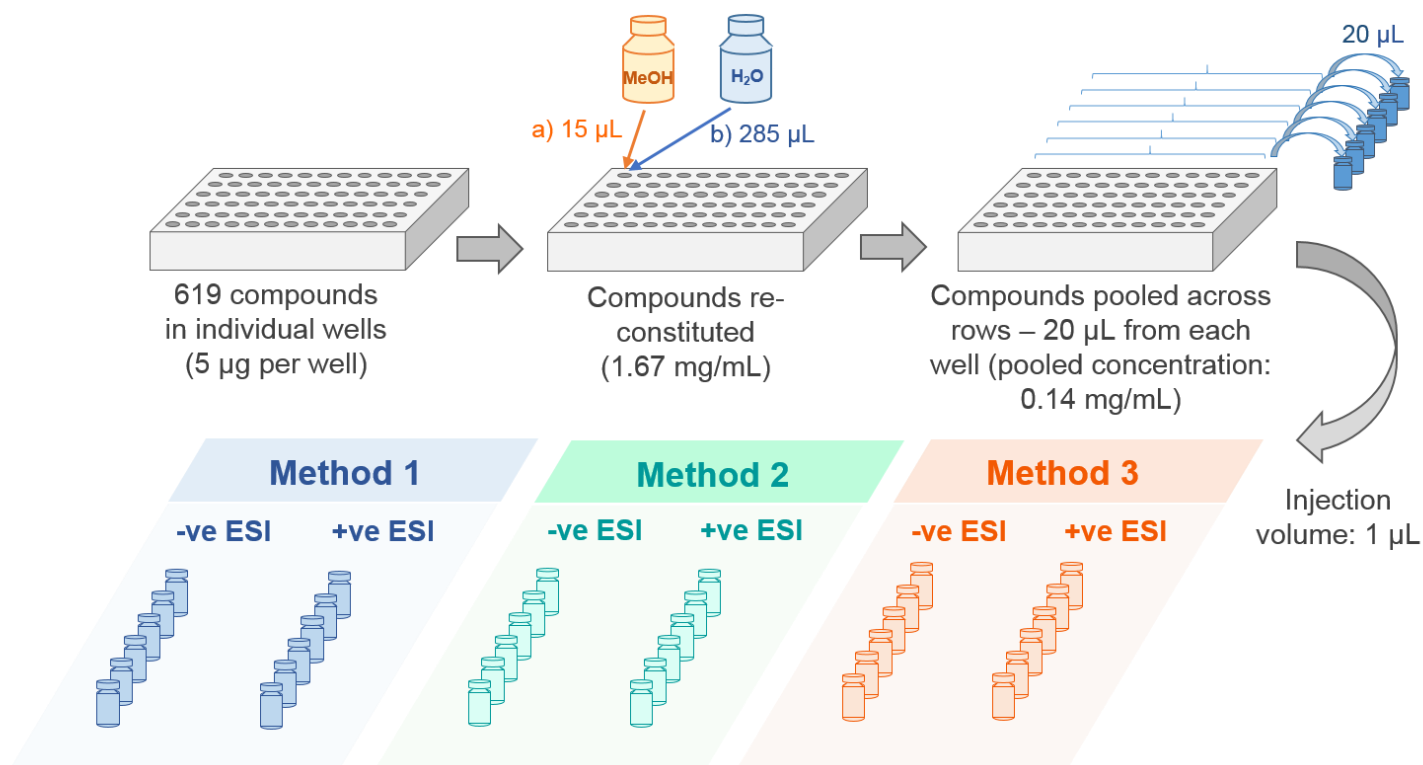
Figure 2.01. Summary of experiment workflow for preparation and LC-QTOF-MS analysis of the 619 metabolite standards. Compounds were pooled across rows (A1-A12, B1-12 *etc.*) in each of seven 96-well plates. Analysis comprised 112 injections of compound pools incorporating positive and negative electrospray ionisation (ESI) polarity for methods 1, 2 and 3.

2.2.2   Analytical conditions

Chemicals

Water for mobiles phases was purified in-house as above. Methanol, acetonitrile, isopropanol (Sigma Aldrich), formic acid (Biosolve, Netherlands) and ammonium formate (Fisher Scientific, Germany) were LC/MS grade.

Equipment

Sample analysis was performed on an Agilent 1290 Infinity UHPLC system coupled to an Agilent 6550 QTOF-MS equipped with a dual AJS ESI source (Agilent Technologies UK Ltd., Cheadle, UK).

Chromatographic conditions

Three chromatographic methods were designed to separate different compound classes. Method 1: non-polar compounds. Method 2: a range of polar and non-polar compounds. Method 3: polar compounds.

Method 1

A Zorbax Eclipse Plus C_{18} column (2.1 x 100 mm, 1.8µm, Agilent) was maintained at 60°C with a flow rate of 0.4 mL/minute. Mobile phases were (A) water and (B) methanol both containing 5 mmol/L ammonium formate and 0.1% formic acid. The elution gradient started at 5% mobile phase B at 0-1 minutes increasing linearly to 100% B by 12 minutes, held at 100% B until 14 minutes, returning to 95% A for 5 minutes to recondition the column.

Method 2

Method 2 employed the same conditions and elution gradient as method A, but with an Atlantis dC$_{18}$ column (3.0 x 100 mm, 3 μm, Waters, UK).

Method 3

A BEH Amide column (3.0 x 150 mm, 1.7 μm, Waters) was maintained at 40°C with a flow rate 0.6 mL/minute. Mobile phases were (A) water and (B) acetonitrile (both contained 0.1% formic acid). The elution gradient started at 99% mobile phase B decreasing linearly to 30% from 1-12 minutes, held at 30% mobile phase B until 12.6 minutes, returning to 99% mobile phase B for 3.4 minutes to recondition the column.

For all methods, sample injection volume was 1 μL for metabolite standards. The autosampler was maintained at 4°C and the autosampler needle was washed with a solution of methanol:isopropanol:water (45:45:10 v/v) between sampling.

QTOF-MS conditions

The mass spectrometer was tuned and calibrated according to protocols recommended by the manufacturer. Acquisition was performed in 2 GHz mode, positive and negative ionisation polarity and mass range 50-1700. The capillary voltage was 4000 V and fragmentor voltage 380 V. The desolvation gas temperature was 200°C with flow rate at 15 L/min. The sheath gas temperature was 300°C with flow rate at 12 L/min. The nebulizer pressure was 40 psig and nozzle voltage 1000 V (± for positive and negative ionisation modes). The acquisition rate was 3 spectra/second.

A reference mass correction solution was prepared in 95:5 methanol:water containing 5 mmol/L purine ($C_5H_4N_4$, CAS No.: 120-73-0), 100 mmol/L trifluoroacetic acid ammonium salt (TFA; $CF_3CO_2NH_4$, CAS No.: 3336-58-1) and 2.5 mmol/L hexakis(1H, 1H, 3H-tetrafluoropropoxy)phosphazine (HP-0921; $C_{18}H_{18}F_{24}N_3O_6P_3$, CAS No.: 58943-98-9) (Agilent). The solution was continually infused at a flow rate of 0.5 mL/min by a separate isocratic pump for constant mass correction [positive ionisation: purine (*m/z* 121.0509), HP-0921 (*m/z* 922.0098); negative ionisation: TFA (*m/z* 112.9856), purine (*m/z* 119.0363), HP-0921 (HP-0921 + formate adduct: *m/z* 966.0007)].

2.2.3 *Acquisition and processing of data from metabolite standards*

Data were acquired and processed using the following software from the MassHunter package (Agilent). Data were acquired with Data Acquisition (build 06.00). Quality checks and processing of raw data files (Agilent '.d' files) were performed with Qualitative Analysis (build 07.00). Extracted ion chromatograms (EICs) of reference masses were performed to check mass accuracy remained <5 ppm throughout the run. Binary pump pressure curves for injections across each analytical sequence were overlaid to check chromatographic reproducibility.

From data acquired from analysis of metabolite standards, compound signals were extracted in Qualitative Analysis (build 07.00) based on the molecular formulae of compounds in each injection using the find by formula algorithm; with a mass window of theoretical AM (calculated from molecular formula) ±5 ppm (Figure 2.02). Allowed species for positive ionisation mode were: H^+ and Na^+, with the addition of NH_4^+ for methods 1 and 2 (mobile phases contained

ammonium formate); and for negative ionisation mode H⁻ with the addition of

CHO_2^- for methods 1 and 2. Dimers were allowed for both ionisation polarities.

Charge state range was 1-2.
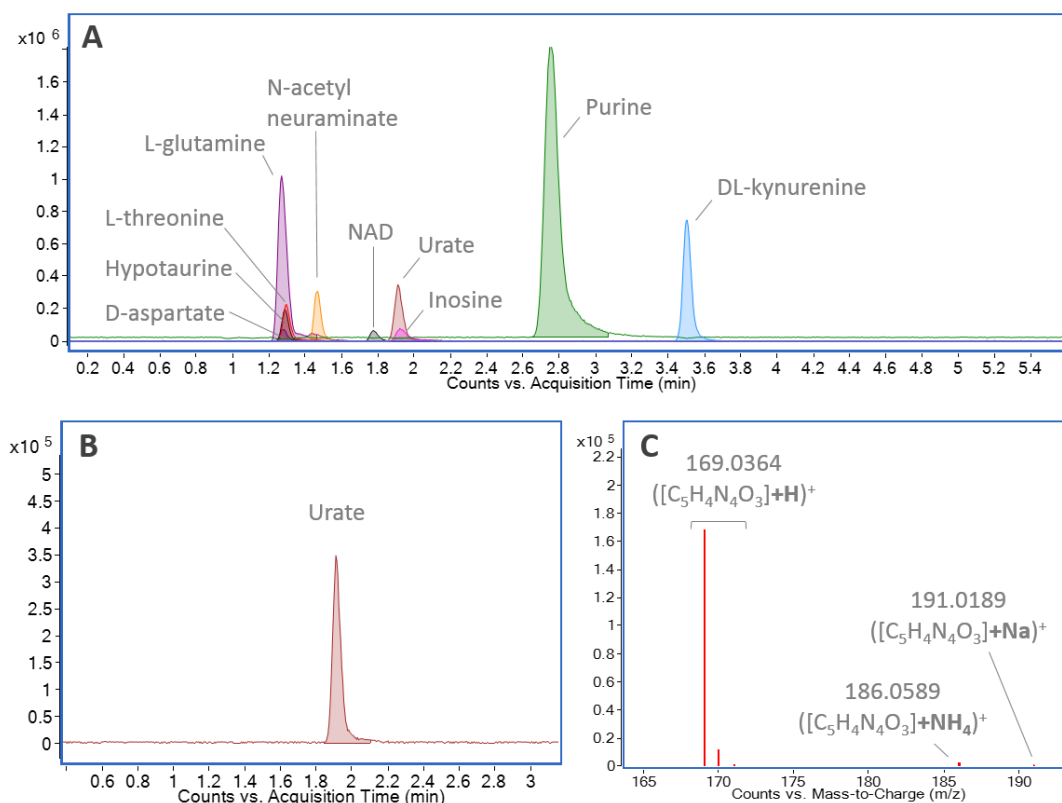
Figure 2.02. Representative example of the data processing output from extraction of signals for metabolite standards. A: combined chromatograms representing the signals for ten standards successfully extracted from the data. The data displayed is from one injection of twelve metabolite standards pooled and analysed by method 2 in positive polarity. **B**: example individual extracted ion chromatogram for urate. **C**: example cleaned mass spectrum for urate; the displayed ions represent the +H (*m/z* 169.0364, 170.0391 and 171.0411, corresponding to ^{12}C, ^{13}C and 2x ^{13}C isotopologues respectively), +NH₄ (*m/z* 186.0589) and +Na (*m/z* 191.0189) species.

2.2.4 Data from LC-QTOF-MS analysis of metabolite standards

Compound signals were extracted as described above. Only standards with RT >0.3 min after the initial column void were considered. Of the 619 standards analysed, 519 (83.8%) were retained sufficiently to be detected by at least one of the three LC-QTOF-MS methods. 177/619 (28.6%) of compounds were detected by all three methods. A considerable proportion of compounds were retained and detected by one (116/619 [18.7%]) or two (226/619 [36.5%]) methods only, demonstrating the utility of combining data from the three chromatographic methods to increase coverage of the metabolome (Figure 2.03A). Figure 2.03 (B-D) shows the distribution in mass and RT for compounds detected by the three methods. AM, RT and charge state of detected compounds were entered into an in-house AMRT metabolite database for each method for matching unknowns against. The databases could then be used to identify unknown peaks in data from experimental analyses.

Figure 2.03 (B-D) also highlights the differences in selectivity between the methods for three chemical groups: carbohydrates, amino acids and lipids. Method 3 retains and separates highly polar, water-soluble compounds such as carbohydrates (20/26 detected, RT range: 5-9.5 minutes). In methods 1 and 2, the same carbohydrates were weakly retained; for method 1 all carbohydrates detected eluted within the column void volume (24/26 detected, RT range: 0.58-0.68 minutes) and for method 2 all carbohydrates detected eluted close to the void volume (23/26 detected, RT range: 1.32-1.44 minutes). Method 1 showed enhanced suitability for analysis of lipid-like compounds (11/11 detected, RT range: 7.69-13.72 minutes) compared to methods 2 (8/11

detected, RT range: 9.27-14.3 minutes) and 3 (7/11 detected, RT range: 1.5-8.26 minutes) which detected fewer of these compounds within a narrower RT range overall. All three methods showed detection and a degree of chromatographic separation for the amino acids, with method 3 appearing to show the most useful chromatographic resolution and retention (19/21 detected, RT range: 5.42-8.1 minutes) compared to methods 1 (21/21 detected, RT range: 0.55-2.82 minutes) and 2 (21/21 detected, RT range: 1.07-4.6 minutes).

Separate AMRT databases were created for methods 1-3, detailing the theoretical AM and RT of each detected metabolite standard, as described below. The databases have been made publicly available to the metabolomics community as a resource via the following link https://doi.org/10.6084/m9.figshare.c.4378235.v2 (201).
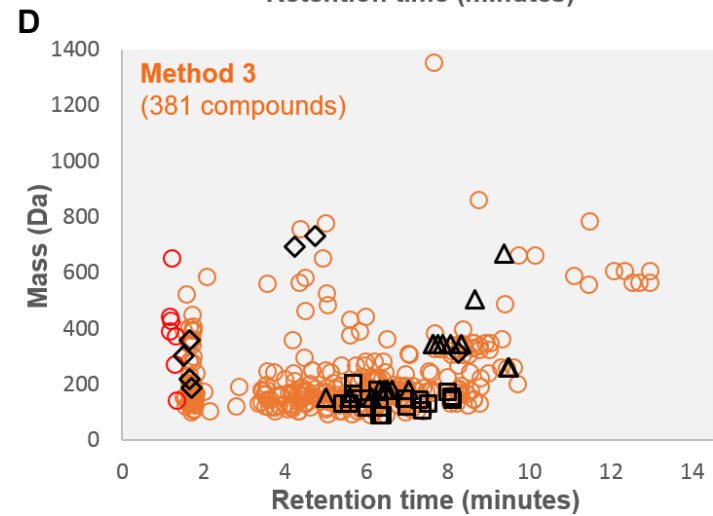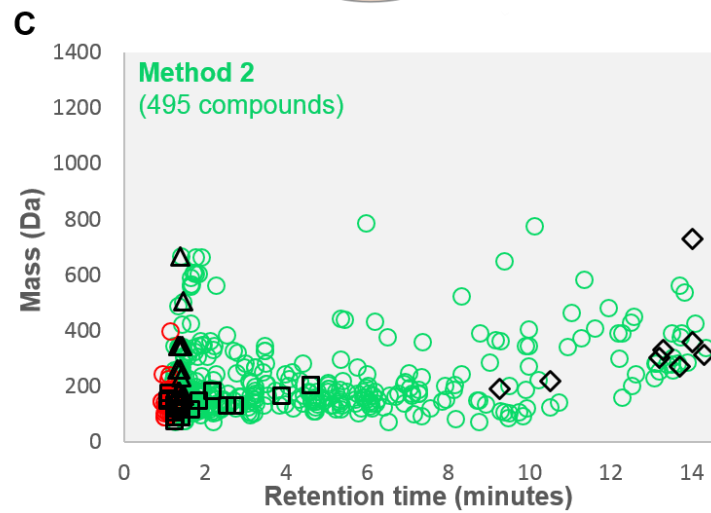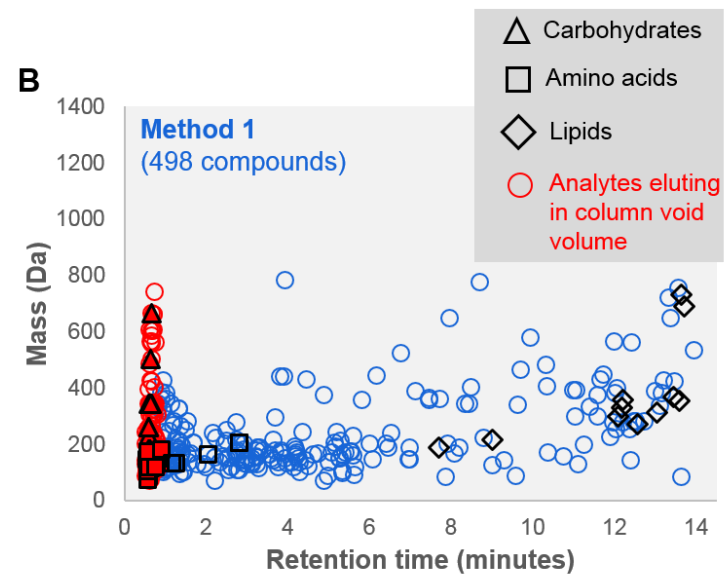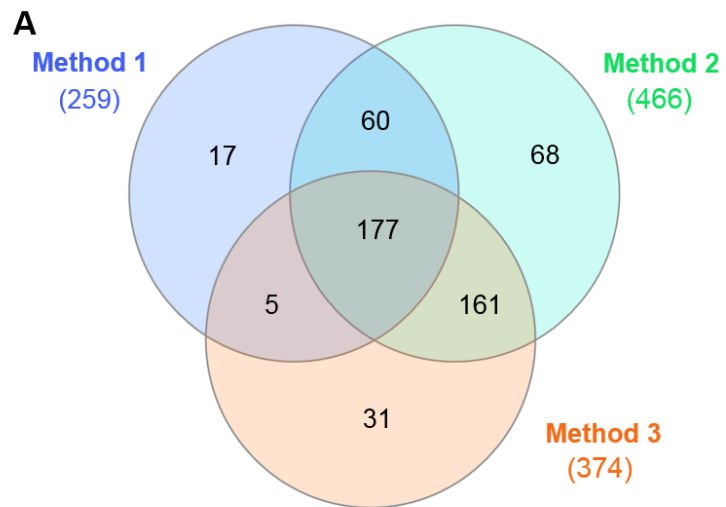
Figure 2.03. Summary of data acquired from analysis of metabolite standards. A: Venn diagram summarising the number of compounds retained and detected by methods 1-3. 177 compounds were retained and detected by all methods. Non-overlapping sections show that an additional 118 compounds were only detected by one method. This shows that coverage of the metabolome can be enhanced by combining data from acquired using methods 1-3. (B-D) Mass vs RT scatter plots for metabolite standards detected with the three analytical methods. **B**: Method 1 is more suited for separation of non-polar lipid-like compounds. **C**: Method 2 provided separation of both amino acids and lipids and **D**: Method 3 is optimal for chromatographic separation of polar compounds such as carbohydrates. Squares, triangles and diamonds represent amino acids, selected carbohydrates and selected lipids respectively. Red markers indicate compounds eluting within the column void volume (defined here as the earliest eluting compound for each method +0.3 minutes). Methods 1 and 2 employed reversed-phase chromatography with Zorbax Eclipse Plus C18 (2.1 x 100 mm, 1.8 µm, Agilent Technologies, UK) and Atlantis dC18 (3 x 100 mm, 3 µm, Waters, UK) columns respectively. Method 3 employed normal-phase chromatography with a BEH amide column (3.0 x 150 mm, 1.7 µm, Waters, UK).

2.2.5 Generation of AMRT databases from metabolite standards

Three separate AMRT databases were created for the metabolite standards analysed by methods 1-3 using PCDL Manager (build 08.00, Agilent). The MassHunter METLIN metabolites PCD/PCDL accurate mass database (Agilent, build 07.00) was used as the basis for the AMRT databases. The version of the METLIN metabolites database employed here contains entries for 30,232 small molecules, including information such as chemical name, empirical formula, monoisotopic AM and deposited spectral and collisional cross section (ion mobility) data. Unique chemical identifiers are also provided

for each compound, including CAS, ChemSpider, METLIN, KEGG and HMDB unique accession numbers. Attaching such accession numbers is important not only for providing an unambiguous chemical identifier, and therefore avoiding potential confusion arising from the use of many different synonyms, which is common for metabolites, but also to enable mapping of metabolites of interest to pathways using tools such as Agilent Pathways Architect; the use of which is described in subsequent chapters. Once each metabolite from the Sigma IROA MSMLS library had been matched with an entry from the METLIN database, all of the selected entries, together with their METLIN information and data, were appended to separate blank PCDL databases for analytical methods 1-3. The Sigma compound name (*i.e.* the synonym used by Sigma), and RT, as provided from the present analysis, of each detected compound were entered into the respective AMRT database for each method (see Figure 2.04).

Figure 2.04. An example of compound entries from the AMRT databases created from analysis of metabolite standards and fully integrated with compound information from the METLIN database. The screenshot shows compound entries in PCDL Manager for the AMRT database created for method 2. From left-to-right, column headings refer to: 1) Name - the name used in the METLIN database, 2) Synonyms - the name used by Sigma-Aldrich, 3) Formula - empirical formula, 4) Mass - theoretical accurate mass, 5) RT – RT measured by the analytical method described above, 6) - 9) chemical database accession numbers, 10) NumSpectra - number of deposited experimental spectra (including positive and negative ionisation and different collision energies), 11) CCS Count - collisional cross section data from ion mobility MS analysis.

2.2.6 Discussion of LC-QTOF-MS analysis of metabolite standards and generation of AMRT libraries

The data from analysis of metabolite standards demonstrate the enhanced coverage achieved from combining the three complementary LC-QTOF-MS techniques as detailed above for non-targeted metabolic profiling. Of the 619 standards, 118 additional compounds were detected by combining data from all three methods. The mass vs RT distributions of compounds that were sufficiently retained to be detected (Figure 2.03) highlight the complementary selectivities of the three methods for chromatographically resolving metabolites from different biochemical groups. Methods 1 and 2 were more successful in resolving non-polar, lipid-like compounds. Method 3 was more successful in resolving highly polar compounds such as polar amino acids and carbohydrates. Method 2 was also selected because the chromatographic column (Atlantis dC18 3.0 x 100 mm, 3 µm, Waters, UK) continues to be used for LC-MS/MS quantitation of nitisinone, HGA and other tyrosine-related metabolites in serum (137) and urine (138) at the NAC .

AM, charge state, RT and MS/MS fragmentation spectra were entered into in-house compound databases to provide unique method-specific chemical signatures of the compounds under study. The following sections detail the application of the databases for identification of unknown compounds acquired from LC-QTOF-MS analysis of human biofluid samples (urine).

2.3 Application of developed AMRT databases for targeted metabolite identification in urine

Generation of the AMRT databases from analysis of metabolite standards, as described above, formed the basis of a targeted data extraction and metabolite identification strategy which was applied to biological samples such as urine. It is important to clarify that the datasets discussed in the following sections were acquired using a non-targeted analytical method, as no precursor ion selection was performed during sample analysis. The targeted aspect of this approach refers to the data analysis only, in which AMRT databases for each of methods 1-3 provided the targets for compound signals to extract from the data by matched AMRT. Further details on this targeted extraction approach and complementary non-targeted extraction are provided in Chapter 3 (section 3.3.1).

2.3.1 Urine used to evaluate targeted LC-QTOF-MS strategy

24-h urine was obtained from 25 AKU patients attending the NAC in Liverpool (12 male, 13 female, mean age (±SD) 51(±15) years) as part of a wider metabolic profiling experiment (see Chapter 4). Samples were collected from the same patients before commencement of nitisinone therapy then at 3 (2 mg every other day), 12 and 24 months (2 mg daily) on nitisinone. An overall pooled sample was created by pooling 20 µL of each individual urine sample. Pooled urine samples were diluted 1:3 urine:deionised water prior to analysis. Data from this pooled sample was used to evaluate the number of AMRT matches obtained using methods 1-3.

2.3.2 AKU-relevant compounds spiked into water and urine matrices

HGA, phenylalanine, metadrenaline (MA), normetadrenaline (NMA), 3-methoxytyramine (3-MT) and tyrosine standards were purchased from Sigma Aldrich, UK. Nitisinone was purchased from APL, Sweden. All dilutions and sample preparation were performed in glass.

Stock solutions of standards were prepared to concentrations of 10 mmol/L. Tyrosine was dissolved in 1 mL sulfuric acid before addition of 9 mL deionised water. HGA and phenylalanine were dissolved in 10 mL deionised water. Nitisinone was dissolved in 5 mL acetonitrile before addition of 5 mL deionised water. MA, NMA and 3-MT were prepared in 10 mL (0.1 mmol/L) hydrochloric acid.

Stock solutions were stored at -80°C. Combined standard solutions of HGA, phenylalanine, tyrosine and nitisinone were prepared in water and urine matrices to 0.1-500 µmol/L via serial dilution with human urine (non-AKU) diluted 1:3 urine: deionised water.

Combined MA, NMA and 3-MT standard solutions were prepared in deionised water to 0.1-25 µmol/L. MA, NMA and 3-MT were focused on in the urine matrix as these are direct metabolites of the parent catecholamine compounds (adrenaline, noradrenaline and dopamine analysed in Sigma library) and are present at higher concentrations in urine. Commercial urine calibration standards (Chromsystems, Munich) were prepared at concentrations of 0-24.8 µmol/L for MA; 0-46.8 µmol/L for NMA; and 0-27.7 µmol/L for 3-MT. These urine calibration standards were also diluted 1:3 urine: deionised water before injection.

2.3.3 Acquisition and processing of data from analysis of urine by LC-QTOF-MS

Urine samples for assessment of the LC-QTOF-MS strategy were analysed under identical analytical conditions as detailed above for metabolite standards. Pooled urine samples from AKU patients were analysed by all three LC-QTOF-MS methods in positive and negative ionisation, mass range 50-1700. Unknown chemical entities from these analyses were matched against the respective AMRT database with mass and RT matches required; theoretical AM ± 10ppm and database RT ± 0.3 minutes. This RT tolerance was chosen in order to allow for a reasonable degree of RT shift for metabolites present in urine matrix compared to the measured RT of their corresponding chemical standards, which were analysed in blank solvent matrix. Allowed ion species were the same as described above for metabolite standards (section 2.2.3). This data processing was performed in Profinder (build 08.00).

For AKU-relevant standards spiked into water and urine, HGA, nitisinone, phenylalanine and tyrosine were analysed by method 2; MA, NMA and 3-MT were analysed by method 3. These methods were selected as they gave the most optimal chromatographic separation of the compounds under study. HGA and nitisinone were analysed in negative ionisation mode as this showed the greatest signal. Tyrosine, phenylalanine, MA, NMA and 3-MT were analysed in positive ionisation mode. Signals for these compounds were quantified using MassHunter Q-TOF Quantitative Analysis (build 06.00), based on total peak area.

2.3.4 Identification of unknown urinary metabolites by AMRT

Unknown chemical entities detected from analysis of pooled urine from AKU patients in positive or negative ionisation polarity were matched against the respective AMRT database generated for each method. Here, urine analysis comprised two replicate injections of the pooled sample from the AKU urine metabolomics experiment; *i.e.* the first (post-equilibration) and last injection of an entire urine profiling experiment comprising 205 injections in total. Only matches obtained from both replicate injections in positive or negative ionisation polarity were analysed further, and only unknowns with single AMRT compound matches were considered identified in this analysis. Several additional QC filters were then applied to AMRT matches to ensure reproducibility across the analytical run; CV <25% for peak area and RT between the two replicate injections. Table 2.01 summarises the number of matches retained for each method following each of these filtering steps. Importantly, no compounds were filtered out due to RT CV >25%; the maximum RT CV% change between replicate injections of the overall pooled QC across all analytical runs was 4.1%.

Figure 2.05 (A) shows that 203 unique compound matches were obtained from urine in at least one polarity by at least one method (see Supporting materials, Appendix 1 for details of the matches obtained by methods 1-3). Fourteen AMRT matches were obtained by all three methods. An additional 61 AMRT matches were obtained by two methods, and 128 matches by only one of the methods, further supporting the increased coverage achieved by the three methods in combination. Scatter plots in Figure 2.05 (B-D) show the mass and RT distribution of the AMRT matches.

Table 2.01 shows the number of AMRT matches obtained when narrower AMRT matching windows were applied. For each method, >50% of matches obtained at ±10 ppm and ±0.3 min post-QC remained when the window was reduced to ±5 ppm and ±0.15 min.
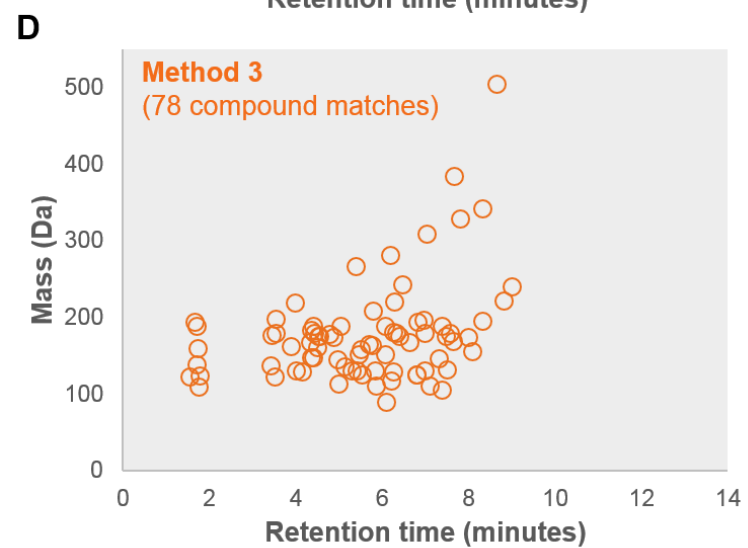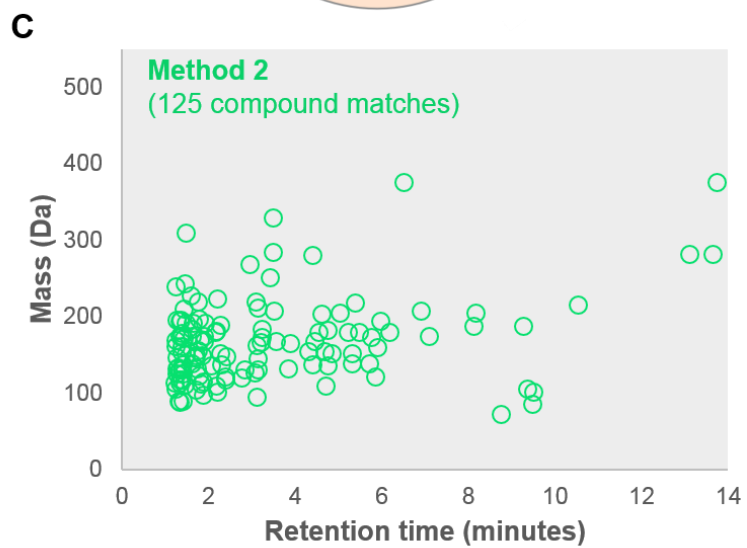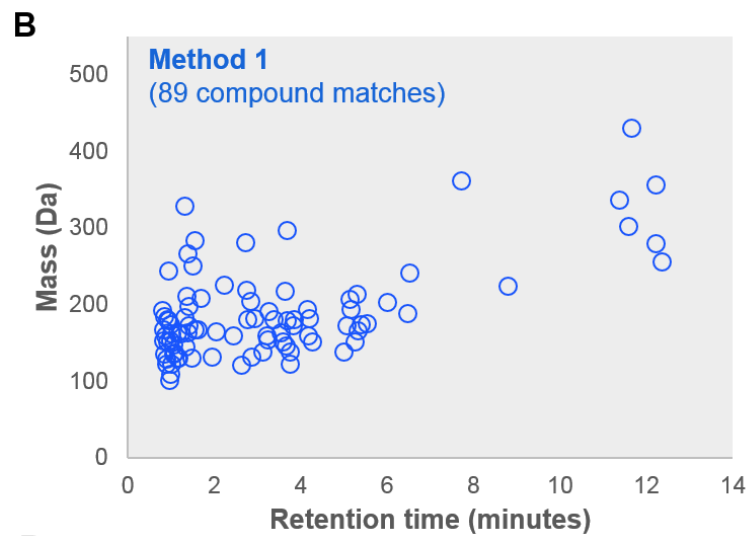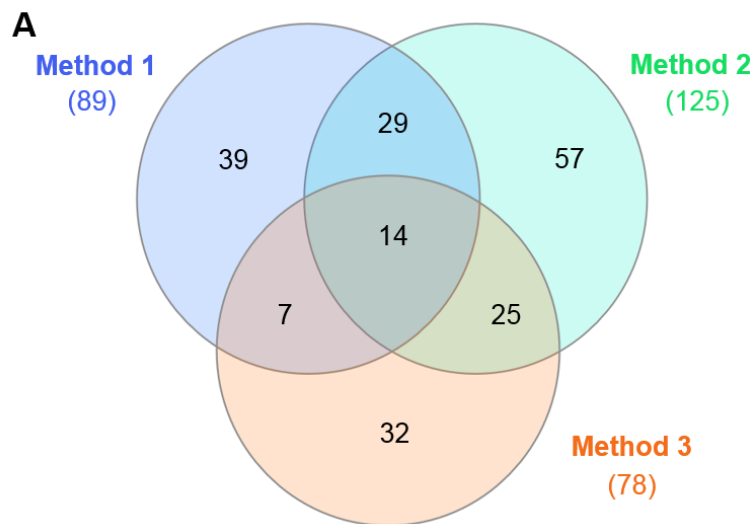
Figure 2.05. Summary of compounds identified from analysis of urine by AMRT. A: Venn diagram summarising the number of urine AMRT matches obtained for methods 1-3. Combining data across the three analytical methods increased the number of AMRT matches, again demonstrating the enhanced coverage of the metabolome achieved by the strategy. **B-D:** mass vs RT scatter plots for AMRT compound matches obtained from LC-QTOF-MS analysis of urine by (B) Method 1, (C) Method 2 and (D) Method 3. These data were obtained from analysis of a pooled human urine sample consisting of equal aliquots of each individual urine sample (n=100) used to assess the effect of nitisinone.

Table 2.01. Number of AMRT matches obtained by the three methods and retained following quality control filtering and decreasing AMRT window size during feature extraction. (+) and (-) refer to positive and negative polarity respectively.

Feature extraction window	Filtering step	Number of urine AMRT matches					
		Method 1		Method 2		Method 3	
		(+)	(-)	(+)	(-)	(+)	(-)
Accurate mass 10ppm, RT 0.3min	None	274	153	248	105	111	74
	(1) Unique AMRTs only, manual curation	80	38	121	43	70	37
	(2) Abundance QC: peak area CV <25% between replicates	75	36	107	35	59	30
	(3) RT shift QC: RT CV <25% between replicates	75	36	107	35	59	30
Accurate mass 5ppm, RT 0.15min	Steps (1) - (3)	44	31	65	25	38	22
Accurate mass 2.5ppm, RT 0.075min	Steps (1) - (3)	17	21	22	12	14	13

2.3.5 Data from analysis of AKU-relevant compounds spiked into water and urine matrices

Sensitivity and response linearity

HGA, NTBC, phenylalanine and tyrosine were analysed by method 1. 3-MT, MA and NMA were analysed by method 3 (methods described in section 2.2.2). Both LC-QTOF-MS methods successfully resolved the compounds under study. R^2 values for HGA, nitisinone, phenylalanine and tyrosine showed linearity in the spiked concentration ranges tested in water and urine matrices (0.1–500 µmol/L for HGA and nitisinone, 0.1 –100 µmol/L for phenylalanine and tyrosine) (Table 2.02). As shown by lower limit of detection (LLOD) values in Table 2.02, a response was recorded for the lowest concentration standard of each compound. LLOD values did not account for HGA, tyrosine and phenylalanine native to the urine sample, however there was no base nitisinone in the urine sample; the donor was not on nitisinone therapy. MA, NMA and 3-MT eluted between 3.7-4.4 minutes. The linear range for these compounds was 0.1-25 µmol/L (Table 2.02).

Table 2.02. *m/z*, retention times (RTs), lower limits of detection (LLOD) and linearity of response for standard compounds across the concentration range. R^2 values for spiked concentration ranges 0.1-500 μmol/L for HGA and NTBC; 0.5-100 μmol/L for tyrosine and phenylalanine (phe); 0.1-25 μmol/L for NMA, MA and 3-MT. HGA, NTBC, phenylalanine and tyrosine were analysed by method 1. 3-MT, MA and NMA were analysed by method 3 (methods described in section 2.2.2).

Compound	*m/z* (polarity)	RT	Water		Urine	
			LLOD (μmol/L)	R^2	LLOD (μmol/L)	R^2
HGA	167.035 (-)	3.5	0.1	0.99	0.1*	0.99
NTBC	328.0438 (-)	10.0	0.1	0.98	0.1	0.96
Phe	166.0863 (+)	3.9	0.1	0.99	0.1*	0.99
Tyrosine	182.0812 (+)	2.2	0.1	0.99	0.1*	0.99
3-MT	151.0768 (+)	3.7	0.1	0.99	0.1	0.99
MA	198.1137 (+)	3.8	0.1	0.99	0.1	0.99
NMA	166.0848 (+)	4.4	0.1	0.98	0.1	0.97

*LLOD is not strictly valid for tyrosine and phenylalanine in urine here as these compounds are native to the urine sample used. The NMA, MA and 3-MT values were from urine commercial calibration standards containing exact concentrations of these compounds

Analysis of matrix effects

Potential matrix effects of ion suppression were assessed by monitoring the signal of isocratically-delivered reference ions throughout the chromatographic run. Components of complex matrices such as serum and urine can prevent ionisation of sample constituents (126), which should be reflected by loss of ionisation of compounds from the reference mass solution given the close proximity of the sample and reference ion sources. As shown in Figure 2.06, low molecular weight reference ions of *m/z* 119.036 (negative ionisation mode) and 121.0509 (positive ionisation mode) showed a relatively consistent signal across the run for both chromatographic methods for urine matrix compared with blank water or methanol injections (remaining within the 10 ppm extraction window). There was some limited evidence of dropout in reference ion signal, particularly in urine spiked to the highest concentrations at regions of greatest elution. Despite this drop-out, the signal remained detectable for reference ions throughout the run, ensuring AM calibration.
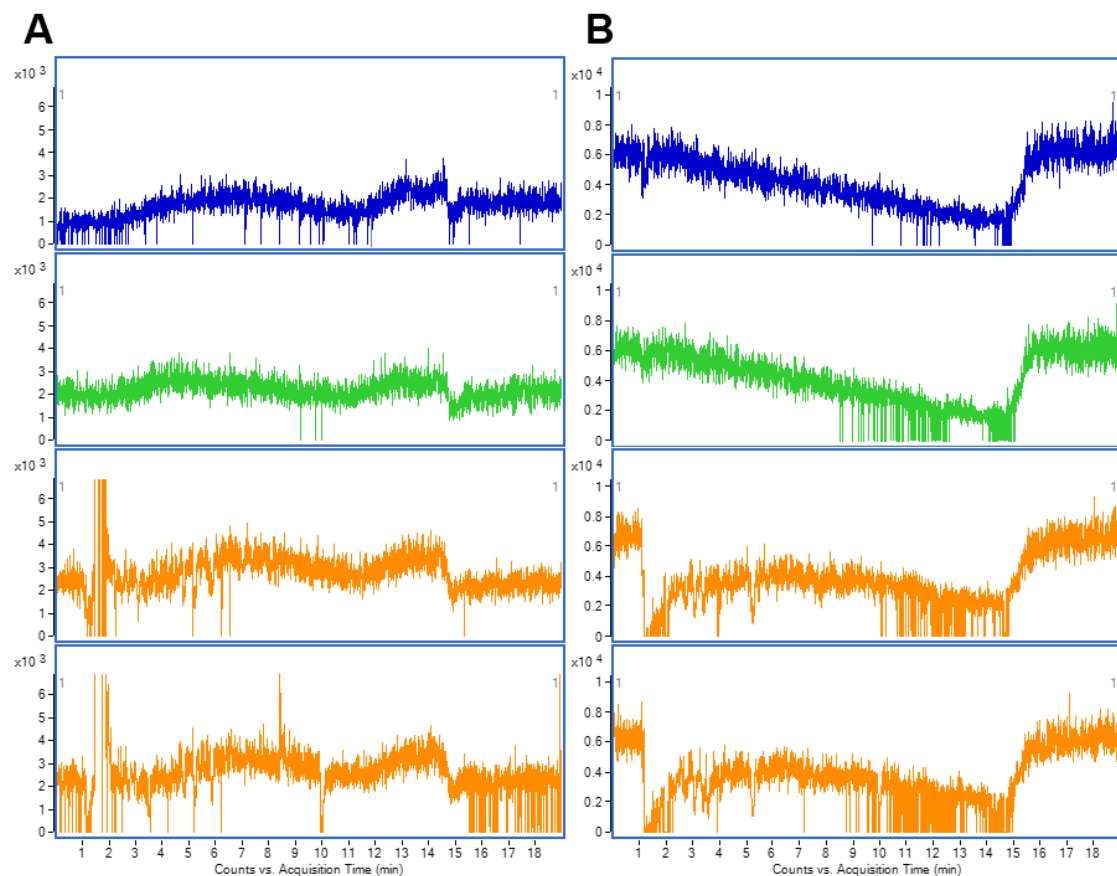
Figure 2.06. Analysis of matrix effects on ion suppression through monitoring signal for reference ions. A: reference ion for negative polarity; *m/z* 119.036, **B**: reference ion for positive polarity; *m/z* 121.0509. Traces represent the signal for these reference ions during analysis of difference matrices; water (blue), methanol (green) and urine diluted 1 in 4 (orange) blank matrix (upper trace) and matrix spiked with HGA, nitisinone, phenylalanine and tyrosine, standards at the highest concentration analysed (lower trace). Data extraction was performed with 10 ppm error. Reference ions were detected within 10 ppm for all sample types. There was some evidence of reference ion drop-out for urine matrices compared with injections of blank water or methanol, but this was generally short-lived and did not appear to affect AM calibration throughout the run.

2.4   Discussion

As compound identification remains a major challenge to the interpretation of metabolomic data, a strategy for comprehensive LC-QTOF-MS metabolic profiling was developed, with compound structure identification by three AMRT databases generated from analysis of 619 metabolite standards. The standards span a range in molecular weight (45-1354 Da) and chemical classes including carboxylic acids, amino acids, biogenic amines, nucleotides, carbohydrates, fatty acids, lipids, steroids and hormones, representing a broad range of primary metabolism. Application of this strategy enabled: a) enhanced coverage of the metabolite standards; b) identification of unknown chemical entities in complex biological matrix (urine) using the appropriate AMRT database; c) chromatographic separation, sub-micromolar detection and linearity across the concentration range for tyrosine-related metabolites of established relevance to AKU when spiked into urine.

It is widely recognised in the literature that different analytical techniques are required according to different metabolite properties (86,112). The distribution in mass and RT of compounds detected by each of the three methods in this study shows their different suitabilities for separating different classes of biological molecules. Method 1 was most successful in resolving non-polar, lipid-like compounds. Method 3 was more successful in resolving highly polar compounds such as polar amino acids and carbohydrates. Method 2 provided the best overall chromatographic separation across the entire suite of standards.

Methods 1 and 2 showed some similarity in analyte retention (Figure 2.03 B-D). Method 2 was selected as the Atlantis dC_{18} column was used to chromatographically resolve metabolites of the tyrosine pathway for the quantitative analysis of serum (137) and urine (138) from AKU patients. This column also provides greater overall retention of polar metabolites over a regular C_{18} column due to endcapping of bi-functionally bonded C_{18} stationary phase material. Method 1 was included here to improve peak capacity owing to the greater number of theoretical plates provided by the smaller particle size of this column (2.1 µm) *versus* the column for method 2 (3.0 µm). Moreover, methods 1 and 2 yielded 39 and 57 (Figure 2.05 A) AMRT matches respectively that were exclusive to each technique. Although it is demonstrated here that use of three separate chromatographic techniques improves coverage of the metabolome, this may not always be feasible. For large-scale studies, for example, it may be pragmatic to use two of the methods to reduce analytical and processing time.

Detection of compounds of known relevance to AKU and/or its treatment with nitisinone spiked into urine at sub-micromolar concentrations shows that the LC-QTOF-MS methods reported are highly sensitive even when applied to low-abundance metabolites within complex biological matrices. Ion suppression is a major concern in MS, particularly using ESI (202), caused by the presence of other compounds that interfere with the ionisation of analytes of interest (126,203,204) and therefore particularly problematic for complex biological matrices such as biofluids which typically contain hundreds of different compounds at high concentrations, ranging from small, simple inorganic salts to

large proteins. Ion suppression can potentially occur in the solution phase or gas phase during the ESI ionisation process and studies on biological sample extracts suggests that a major cause is the presence of non-volatile or less volatile matrix compounds. The exact mechanism by which the matrix interferes with ionisation of a given analyte of interest is not fully understood, but could be attributable to decreased efficiency of either spray droplet formation or evaporation, thereby decreasing the amount of charged ion in the gas phase for a given compound and therefore final MS signal recorded by the detector (202). Various measures can be taken to address ion suppression, for example extensive sample clean-up (*e.g.* solid-phase extraction), signal correction using internal standards and modification of coupled chromatographic separations (126). The lower limits of detection reported here for analysis of AKU-relevant standards supports that ion suppression was not a major concern for the standards analysed using the chromatographic and mass spectrometric conditions applied. Moreover, the monitoring of signal for reference ions for methods 1-3 (see Figure 2.06) showed minimal ion suppression by the introduction of urine sample matrix. Reference ion signal is a useful indicator of ion suppression, as reference ions of a known mass were continually delivered to the ion source throughout the analytical run from a separate probe adjacent to the sample probe, and they therefore compete for ionisation with compounds originating from the sample. Figure 2.06 shows some limited evidence of ion suppression for urine compared to blank solvent matrix, as the reference ion signal decreased at times of high co-elution; *i.e.* around at or close to the initial void for each method, as a number of non-retained compounds

92

elute together and enter the mass spectrometer together (see Figure 2.03B-D; compounds eluting in initial void highlighted in red). The development of three complementary chromatographic techniques however, means that a particular compound of interest will most likely be retained, and therefore subject to more limited ion suppression, by at least one of methods 1-3.

A major advantage of compound identification by an in-house LC-MS compound database compared to public databases (185,199,200,205,206) is that the data were acquired under identical analytical conditions. Data from public databases are typically acquired by different laboratories using different analytical parameters and techniques. Theoretical AM can be calculated for a given compound given the molecular formula, however compound data from public databases such as spectral data and particularly chromatographic RT are largely determined by the technologies and analytical parameters applied. The Metabolomics Standards Initiative has established levels of confidence for metabolite identification (94,107). Only identification strategies, such as the one reported here, that compare two or more orthogonal chemical properties (*e.g.* AMRT) with an authentic chemical standard under identical analytical conditions, achieve the highest level of identification confidence. Identifications by reference to public databases in which data were acquired under different analytical conditions can only achieve the second level of confidence, designated 'putative identification', even when matches are obtained for two orthogonal chemical properties.

In conclusion, an LC-QTOF-MS strategy was developed for enhanced coverage of the metabolome with targeted compound structure identification by reference to three AMRT databases, which have been made publicly available as a resource to the metabolomics community. The strategy and databases developed will support metabolic profiling experiments to address important questions relating to wider metabolism in AKU beyond the tyrosine pathway, and as a metabolic phenotyping strategy more generally. More specifically, this approach will help to streamline and simplify data processing and metabolite identification, which remain the major challenges in metabolomic analyses.

Acknowledgements

3.0 DEVELOPMENT OF A WORKFLOW FOR ANALYSIS OF LC-QTOF-MS METABOLOMIC DATA

3.1   Introduction

In metabolomics, as with other 'omics approaches, the number of observations or chemical entities present in the data often greatly exceeds the number of samples analysed. The high-throughput nature of these experiments is integral to their appeal, but the resulting challenge of extracting meaningful information from the data is substantial (81,207).

First, there is a considerable proportion of redundant data. This redundancy partly reflects background signals from compounds not native to the sample, for example contaminants introduced throughout the various stages of sample collection and preparation that do not convey biologically meaningful information. Data redundancy also refers to the high degree of correlated signals within the data (208); in raw, unprocessed LC-MS metabolomics data, signals present as *m/z*-RT pairs, and there are typically multiple signals representing the same metabolite, for example associated isotopes, adducts and multimers. Furthermore, where the aim of a metabolomics study is typically to discover a biomarker that can reliably discriminate between different groups of samples (*e.g.* cases *versus* controls in a classic clinical metabolomics study), there is a high likelihood of obtaining false-positives, or spurious associations between a given entity and a particular sample group (209). The high tendency for false positives is a reflection of the complexity of metabolomics datasets and may be attributable to a host of potential factors throughout the experimental workflow, including those related to experiment design, sample analysis or from data processing

artefacts introduced during feature extraction, for example false alignment of features across samples.

These challenges and potential pitfalls highlight the need for robust and reproducible data processing strategies in metabolomics studies. This chapter details the data processing workflow employed in the following chapters of this thesis that discuss metabolomics data (*i.e.* Chapters 4 and 5). The workflow incorporates data quality assurance (QA), downstream data processing, QC and subsequent multivariate statistical analyses. The workflow primarily employs vendor software made available through a collaboration with Agilent Technologies (Cheadle, UK), as these computational tools were considered the most compatible with the raw data files generated by an Agilent LC-QTOF-MS. The design of the workflow was guided by the proposed minimum reporting standards for data analysis in metabolomics (106). Note that the data presented in this chapter are for illustrative purposes only. These data were acquired as part of a wider, at present unpublished, LC-QTOF-MS profiling experiment of mouse urine. Also, this chapter concerns only the review of metabolomics data post-acquisition. The methods for design of sample analyses (*i.e.* the analytical sequence) are outlined in Chapter 4 (section 4.2.4), the first results chapter of metabolomics analysis.

3.2 Data Quality Assurance (QA)

The initial step in this workflow is to perform several QA checks on the raw LC-MS data. QA was performed in MassHunter Qualitative Analysis (Agilent, build

07.00) before any subsequent feature extraction on the data. The purpose of this step is to ascertain that the raw data are of sufficient quality to enable valid conclusions following downstream data processing and statistical analyses. The following sections consider two attributes that are fundamental to data integrity in LC-MS and downstream feature extraction and alignment; chromatographic reproducibility and mass accuracy.

3.2.1 Chromatographic reproducibility

The first QA step was to show that the LC conditions remained consistent throughout the analytical run. In LC, the RT of a given compound is determined by pressure, flow rate, column stationary phase composition (material and particle size), mobile phase solvent composition, pH and column temperature. Modern HPLC systems are designed to strictly control these parameters in order to make the analytical conditions consistent for each sample. However, longer analytical runs, such as those spanning across several days, are particularly susceptible to loss of chromatographic reproducibility. This variation could be systematic, for example relating to gradual shifts in chromatographic conditions over time. Variation might also be due to a specific event, for example a sudden increase in pressure resulting from blockage of the analytical column, or changes in pressure due to a leak within the LC system.

An important indicator of chromatographic reproducibility is the binary pump pressure curves. The Agilent 1290 Infinity II system employed in all subsequent metabolomics analyses described has two identical pumps for high-pressure

solvent mixing to create binary chromatographic gradients. The binary pump contains a system pressure sensor so that pump pressure can be measured during the analysis. Figure 3.01A shows the individual LC binary pump pressure curves for each sample overlaid across an analytical sequence comprising 104 injections. This figure shows the characteristic changes in pressure across each sample acquisition as a result of changes in mobile phase composition during the elution gradient (see mobile phase composition and elution gradient details for Method 2, as described in Chapter 2, section 2.2.2). The overlaid pressure curves show the high reproducibility of the chromatographic conditions across the entire analytical sequence, which lasted approximately 33 h in total.

In addition to the reproducibility of binary pump pressure curves, another QA procedure employed was to check the reproducibility of the total ion chromatograms (TICs). This was done for the replicate injections of pooled samples, which were interspersed throughout the analytical sequence for QC purposes, as described in Chapter 4 (section 4.2.4). Figure 3.01B shows a representative example of overlaid TICs for replicate injections of each pooled group sample, which are represented by different colours. The traces for each replicate overlay tightly for each pooled group, further evidencing the chromatographic reproducibility across the analysis.
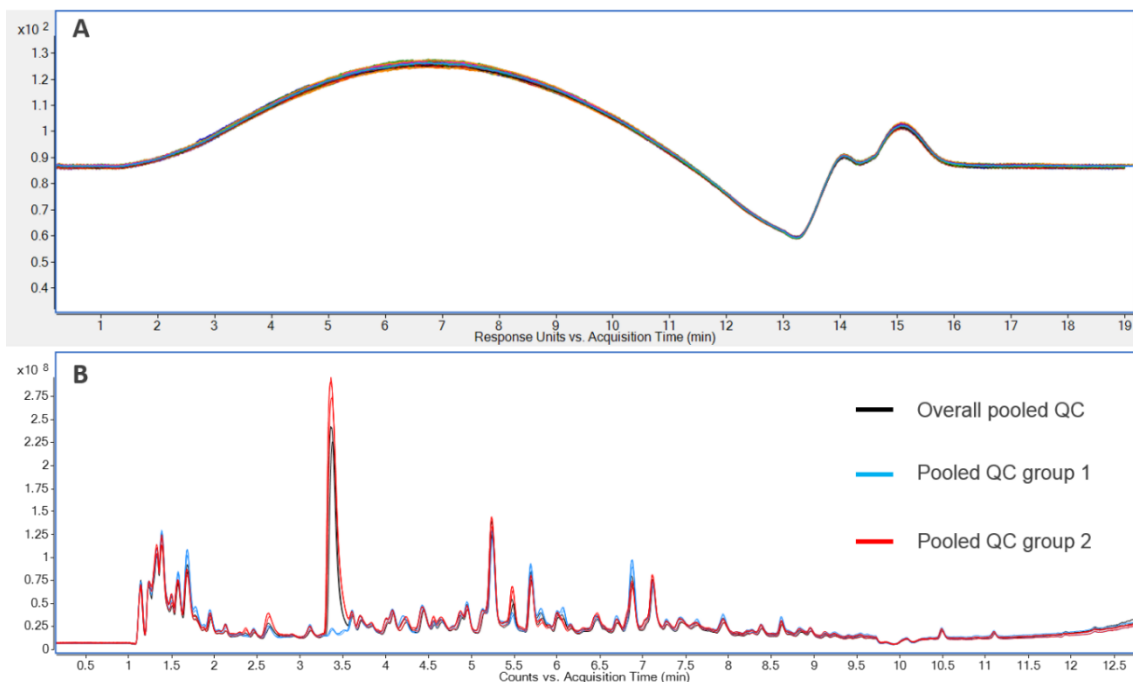
Figure 3.01. Example of chromatographic reproducibility QA procedures performed on data from LC-QTOF-MS analysis of mouse urine. These data were acquired using the chromatographic method (Method 2) described in Chapter 2 (section 2.2.2); a reversed-phase method with an elution gradient of 1-12 minutes. **A**: overlaid binary pump pressure curves of 104 consecutive injections of mouse urine. Note that pressure curves overlay tightly. **B**: overlaid total ion chromatograms (TICs) from replicate injections of three different pooled urine quality control (QC) samples interspersed throughout the analytical run: pooled QC samples for sample groups 1 and 2, and an overall pool comprising all urine samples from the profiling experiment. Note the reproducibility of replicate injections across the analytical run for each of the three QC samples, and also gross differences evident between the pooled samples, for example the peak at 3.5 minutes, which is markedly increased in pooled group 2 and the overall pooled QC. Urine samples were diluted 1:10 urine:water and analysed with a 2 μL injection in negative polarity.

A final check of chromatographic reproducibility was to monitor RT across the run for a signal from a specific compound. Figure 3.02 shows a representative example of this step based on extracting the signal for creatinine (a compound present in urine) from the first and last replicate injections of the overall pooled mouse urine sample from this dataset. Here, extracted ion chromatograms (EICs) were obtained for creatinine using the find by formula option in Masshunter Qualitative Analysis (build 07.00). Figure 3.02A shows the EICs overlaid with the TICs from which they were extracted, and 3.02B (inset) shows just the EICs of the first and last replicate injections overlaid. The EICs show no evidence of an RT shift for this chromatographic peak across the 33-h analysis period. A metabolite standard in blank solvent matrix was also injected as a 'check-out' sample at the start and end of each run (tyrosine and HGA in positive and negative polarity respectively). This enabled an additional assessment of mass accuracy and chromatographic resolution; the latter can be compromised as a result of column stationary phase ageing, typically resulting in broader chromatographic peak shapes.

During this step, TICs from all of the individual samples in an experiment were also overlaid onto one chromatogram. This procedure was used to identify any clear outlier samples based on their TIC. This helped to identify, for example, low abundance TIC profiles due to potential sample preparation errors or issues with the HPLC injector or chromatography.
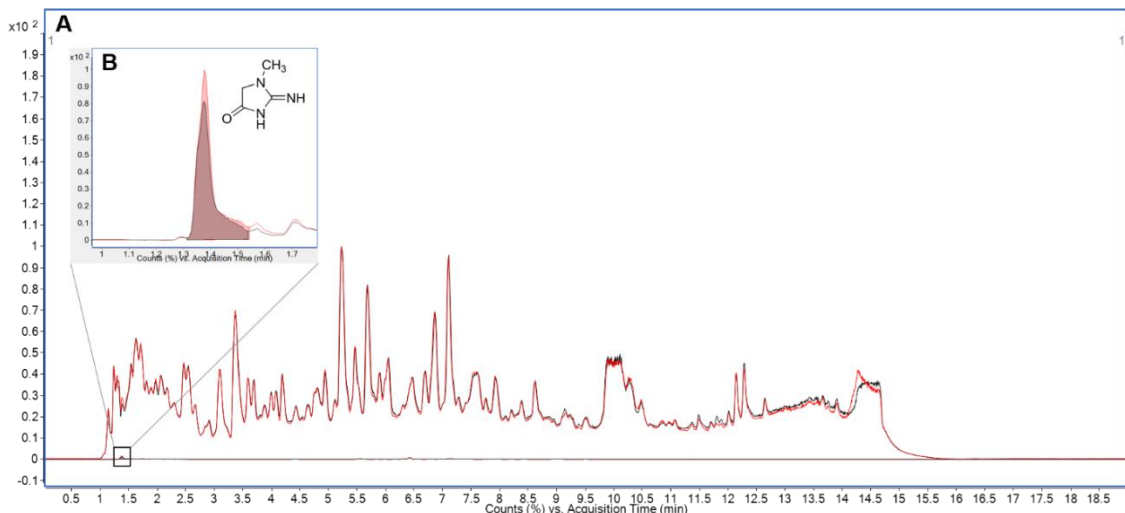
Figure 3.02. Representative data showing retention time (RT) reproducibility between the first and last replicate injections across the run. A: overlaid total ion chromatograms (TICs) from the first (black) and last (red) injections of the same overall pooled urine quality control (QC) sample over an analytical run comprising 104 injections. **B**: overlaid extracted ion chromatograms (EICs) for creatinine for the first (black) and last (red) overall pooled QC injections as above. EICs show minimal RT shift across the run, indicating no need for RT shift corrective measures.

3.2.2 Mass accuracy

In high-resolution mass spectrometric techniques such as QTOF-MS, mass accuracy is typically maintained by calibrating mass values to those obtained from the analysis of various known ions. In this workflow, AM on the 6550 QTOF-MS (Agilent) is maintained by continually infusing a solution of known 'reference' ions via a separate isocratic pump, as described in Chapter 2 (section 2.2.2) and according to manufacturer guidelines. Two reference ions are monitored in positive polarity, and 3 in negative polarity; these are designed to cover a broad

range in mass. The auto-calibration algorithm in the MassHunter Acquisition software (Agilent) automatically calibrates all ion masses to the reference ion signals, provided that the reference ion signals remain within a specified mass window and above a specified abundance threshold. The auto-calibration parameters employed in all LC-QTOF-MS analyses in this thesis were as follows: mass window of 100 ppm and raw abundance threshold of 500 counts. These values were designed for analysis of complex matrices such as urine, in order to ensure that AM auto-calibration was maintained even in the event of reference ion signal drop-out due to ion suppression or at times when there may be momentary increased deviation between the measured and theoretical mass of reference ions.

Reference ion abundance and mass accuracy across the analytical sequence was ascertained by performing EICs of reference masses for all injections within a specified mass window. As shown in Figure 3.03A, all of the reference ions for a given polarity should remain observable within the specified EIC mass window across the entire analytical sequence; the mass window employed here was theoretical AM ±5 ppm. There is some evidence of reference ion dropout at certain points during the elution gradient, particularly in the initial column void (at approx. 1 min). The attenuation of reference ion signal at this point is likely due to ion suppression attributable to the co-elution of many non-retained ion species at this RT. As shown in Figure 3.03B, the reduced reference signal EIC intensity at 1-1.5 min coincides with a sharp increase in the overlaid TIC signal. As the example shown is a reversed-phase method, highly-polar species such as salts

present in urine are likely to elute in the initial void and thus suppress the ionisation of reference ions. However, the momentary drop-out observed is considered acceptable given the complex matrix; it is short-lived and none of the reference ion signals fell below the raw abundance threshold required for mass auto-calibration.
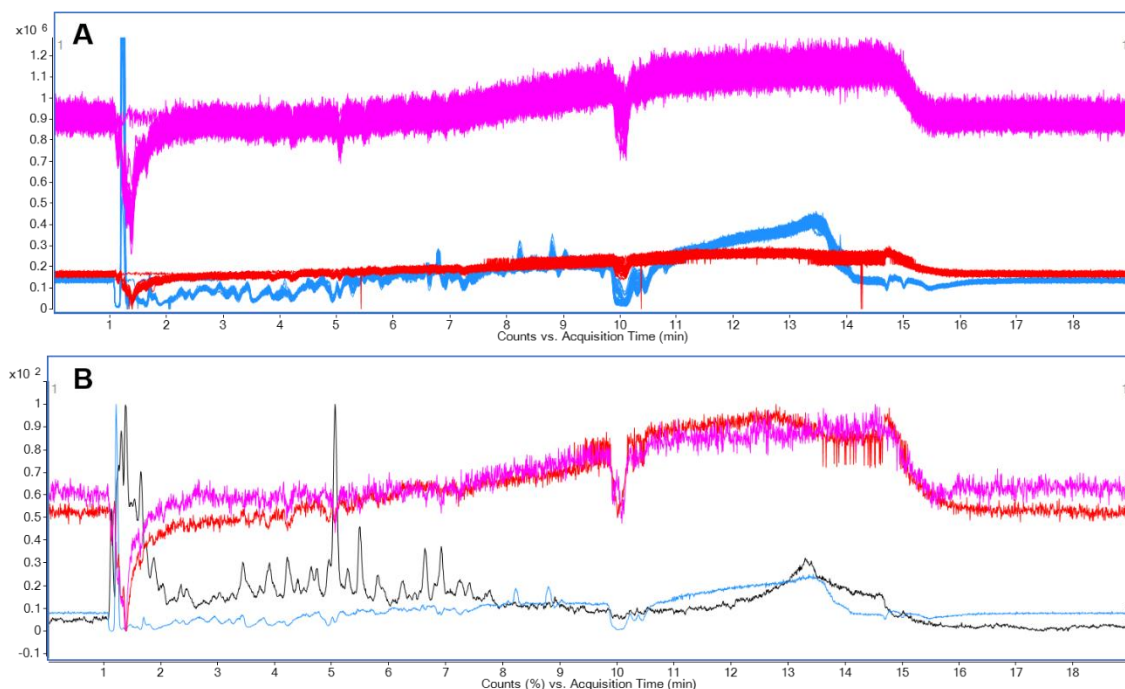
Figure 3.03. Example of mass accuracy quality assurance (QA) based on monitoring the signal for reference ions. A: overlaid extracted ion chromatograms (EICs) for reference ions in negative polarity for the sample injections described in Figure 3.01. Reference ion EICs are represented by blue traces - trifluoroacetic acid ammonium salt (TFA; *m/z* 112.985587); red traces - hexakis(1H, 1H, 3H-tetrafluoropropoxy)phosphazine (HP-0921; *m/z* 966.000725); and pink traces - HP-0921 TFA adduct (*m/z* 1033.988109). Data show that accurate mass calibration based on reference ions was maintained throughout the entire analytical sequence (EICs were obtained with a window of <5 ppm of theoretical reference ion accurate mass). **B**: individual mouse urine total ion chromatogram (TIC; black trace) overlaid with EICs for the reference ions as above. Traces show some limited attenuation of reference ion signal, particularly at times of high elution in the TIC (*e.g.* the initial column void at 1-1.5 min).

3.3 Data pre-processing

Data pre-processing is a critical step in the LC-MS metabolomics data analysis pipeline. In this step, the most important, biologically meaningful features are extracted from the raw data for downstream statistical analyses. A typical pre-processing workflow of an LC-MS metabolic profiling experiment comprises the following stages: noise filtering, peak detection, deisotoping, feature alignment and feature identification (210).

In this workflow, the above actions are performed within the MassHunter Profinder software package (Agilent, build 08.00). In Profinder, these functions are incorporated into the various batch feature extraction workflow algorithms available, which are described in the following section. The exact Profinder feature extraction parameters employed in profiling studies included this thesis are detailed in the respective results chapters.

Raw LC-MS data files contain a list of unique signals, or *m/z* and RT pairings, rather than distinct chemical entities or metabolites. In the raw data, metabolites are represented as a union of chemical signals that should be grouped together during feature extraction (208). A general overview of feature extraction in Profinder is illustrated in Figure 3.04. There are different feature extraction algorithm workflows available in the Profinder software, as discussed below. The general principle behind these approaches is, however, the same; *i.e.* to extract metabolite signals from the raw, unprocessed data and to output this information into a data matrix for statistical analyses.
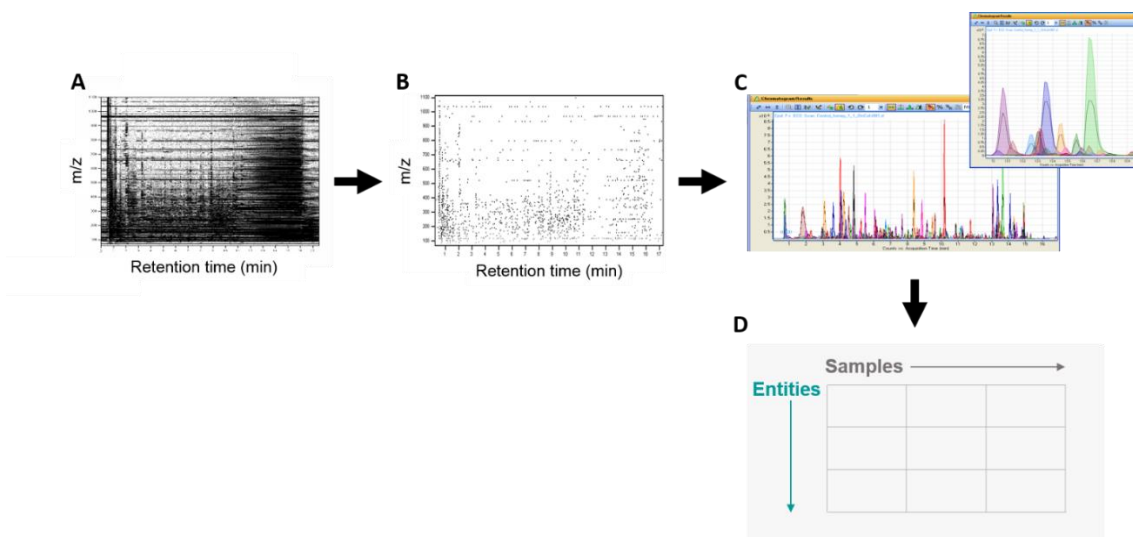
Figure 3.04. Overview of the feature extraction process in pre-processing of untargeted LC-QTOF-MS data. A: two-dimensional *m/z* – RT (retention time) scatter plot of signals in unprocessed raw data. The raw data file contains the total number of signals detected by the spectrometer, which includes signals arising from electronic background noise and ion signals from contaminants within the samples, reagents and laboratory environment, in addition to the signals of interest representing metabolites within the biological sample. **B**: *m/z* – RT scatter plot for the same data file following pre-processing and feature finding to retain the signals representing 'true' chemical signals within the data, which includes metabolites. **C**: overlaid extracted ion chromatograms plotting RT and signal intensity for the chemical entities extracted from the dataset. **D**: exported data matrix comprising abundance values for each extracted chemical entity across all samples in the experiment.

3.3.1 Two complementary feature extraction approaches

A unique characteristic of metabolomics in comparison to other 'omics approaches, such as genomics, transcriptomics and proteomics, is the large number of unidentified chemical structures detected (207). This reflects both the

scope of sensitive, high-resolution profiling technologies such as LC-QTOF-MS and the scale of the metabolome; to date there has been no accepted exhaustive characterisation of the metabolome for any organism (74).

For increased coverage of the metabolome and to ensure that maximal biological information was extracted from the data, two complementary feature extraction algorithms were employed. These extraction workflows are available in Profinder; referred to as 'batch targeted feature extraction' and 'batch recursive feature extraction for small molecules / peptides'. The former algorithm was used to mine datasets against the AMRT databases described in Chapter 2. The advantage of this approach is that the extracted entities are putatively identified by two orthogonal chemical properties, AM and RT. Using this targeted approach, it is manageable to manually review the (typically) resulting several-hundred extracted entities in Profinder, enabling generation of a cleaned and well-curated dataset. As this targeted approach is limited to the IROA MSMLS compounds (Sigma Aldrich, UK), the second feature extraction algorithm stated above was employed to mine the data in using a complementary non-targeted approach. The non-targeted 'recursive' approach typically results in considerably more extracted chemical entities (usually 1000+ entities from urine) compared to AMRT-based targeted feature extraction. The algorithm workflow steps for batch targeted feature extraction and non-targeted batch recursive feature extraction are summarised in Figures 3.05 and 3.06, respectively, and described in more detail below. A more exhaustive explanation of the technical details in each step is beyond the scope of this thesis but available in the Profinder 8 software quick

start   guide   accessible   via

https://www.agilent.com/cs/library/usermanuals/public/G3835-

90027_Profinder_QuickStart.pdf (211) and in the Profinder help manual

accessible through the Profinder (build 08.00) software.

Batch targeted feature extraction employs an algorithm referred to by the vendor

as find by formula. The find by formula algorithm uses the molecular formula

information contained within a compound database (*e.g.* an AMRT database

created using Agilent PCDL Manager; see Chapter 2, section 2.2.5) to calculate

the expected ion masses and isotope distributions and then use these as the

means of extracting features from the data. The targeted feature extraction

workflow comprises 5 steps. In the first step, an input chemical formula source is

specified; *i.e.* the relevant AMRT database in this work, with the 'mass and RT

matches only' option specified. Also specified in this step are the ion species to

consider, for example multimers (monomers to trimers) and expected adducts

based on mobile phase modifiers (*e.g.* NH_4^+ and $COOH^-$) and ions present in the

sample matrix (*e.g.* Na^+ ions in urine), in addition to charge state range. This step

therefore ensures that multiple ion species associated with the same metabolite

are grouped together in the data extraction. In the second step, matching

tolerances and scoring thresholds are specified. These include the size of the

windows for AM and RT matching against the database and several composite

scores employed in the software relating to mass, RT and isotope distribution. In

step 3 the preferred integration algorithm is specified, which refers to the method

for integrating data points across a chromatographic peak; 'agile 2' is the

integration algorithm currently recommended by Agilent and is the default method in Profinder build 08.00. There is an option to smooth EICs during this step, for example using a Gaussian function. Peak abundance thresholds to be applied during integration are also specified during step 3, as is the chromatogram data format (*i.e.* profile or centroid). In step 4, several parameters are specified as to how the spectra are extracted within a chromatographic peak. For example, one single spectrum can be extracted from the apex of the peak only, or alternatively an average of all spectra can be extracted where the corresponding point on the chromatogram is above a specified percentage peak height threshold. In the 5th and final step of the targeted extraction workflow, several filter thresholds can be applied to determine whether a compound should be considered 'found' or otherwise discarded. These filters can be based on absolute or relative height, or a composite score based on the match score against the target chemical formula. Additionally, it is possible to specify that the aforementioned filters are met for a certain number or percentage of samples in the dataset; this can refer to all samples in the experiment or within at least one sample group as specified by the user.
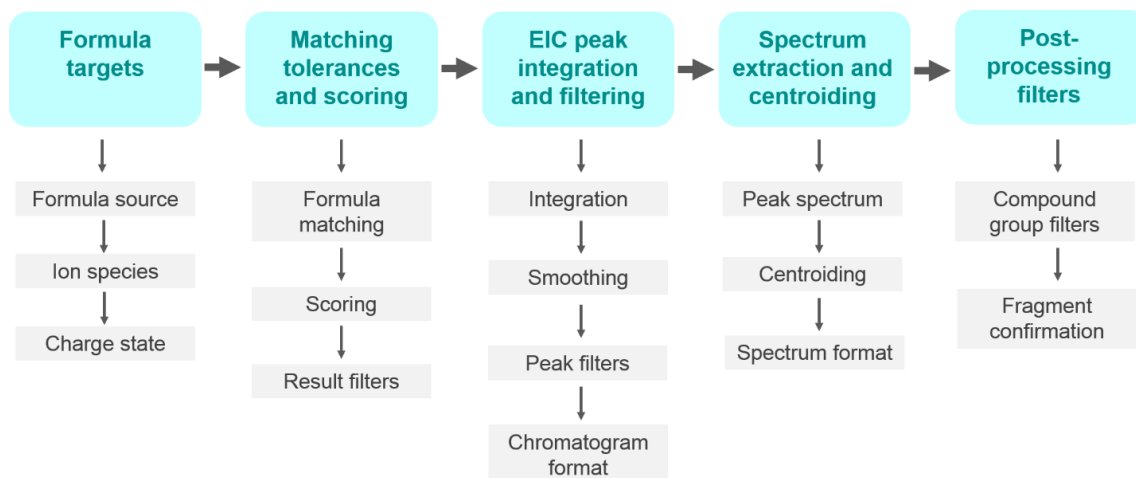
Figure 3.05. Overview of the batch targeted feature extraction algorithm workflow performed in MassHunter Profinder. In the metabolomics studies described in this thesis, batch targeted feature extraction was employed to extract signals from the data that matched the AMRT compound databases that were generated as described in Chapter 2. Figure adapted from (211).

The batch recursive feature extraction workflow for small molecules / peptides includes 8 distinct steps comprising two separate algorithms; molecular feature extraction and find by ion extraction (Figure 3.06). The molecular feature extraction component is similar to the above workflow described for targeted feature extraction, but with the exception that extraction here is not based on specified chemical formula targets. Instead of matching against AMRT targets, the data is binned according to AM and RT in a non-targeted fashion and the extracted entities are aligned across samples according to specified tolerances (see step 3, 'compound binning & alignment' in Figure 3.06). The tolerances for AM are in mDa or ppm, and for RT in minutes or percentage of the expected RT.

The find by ion extraction component of the non-targeted recursive workflow is almost identical to the find by formula algorithm employed in batch targeted feature extraction, as described above, except that the formula sources here are not from a database but the results from the molecular feature extraction. The proposed advantage of this two-component recursive workflow is that in the second stage, the generated features are found in the sample data files in a targeted fashion with greater accuracy. The same post-processing filter options of the extracted entities are available as described above for batch targeted feature extraction.
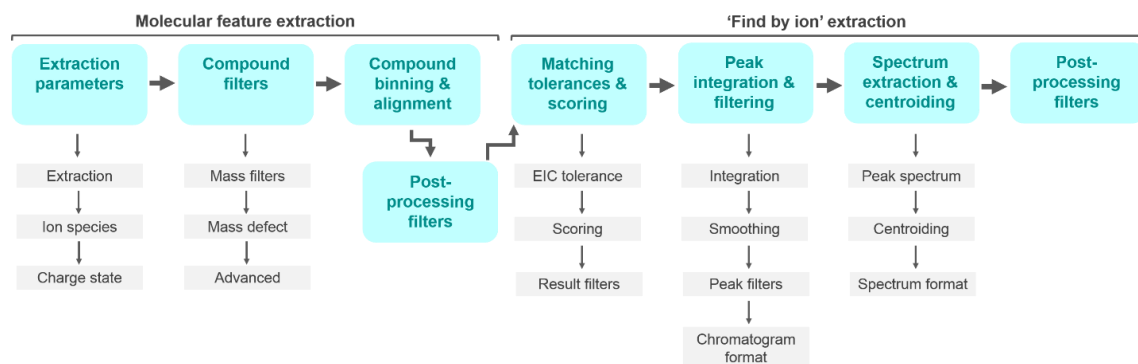
Figure 3.06. Overview of the batch recursive feature extraction for small molecules / peptides algorithm workflow performed in MassHunter Profinder. This algorithm workflow was employed for non-targeted feature extraction. The purpose of this was to extract signals that were not limited to compounds in the AMRT databases described in Chapter 2. The recursive extraction workflow comprises 8 steps. The first 4 steps are part of the untargeted mining for chemical entities, molecular feature extraction, and the following 4 steps are part of the find by ion extraction which performs targeted re-mining of the data to validate the entities generated from molecular feature extraction. Figure adapted from (211).

3.3.2   Manual curation of extracted data

The automated algorithm workflows described above make it possible to perform extraction on large metabolomics datasets. However, it is important to manually review the quality of the results from the extraction, and to correct any errors of integration accordingly. This form of manual curation was performed for the datasets generated from targeted AMRT-based extraction, as these datasets were prioritised, and the relatively fewer extracted chemical entities (typically 300-500), compared with non-targeted extraction (typically >1000) made this task more manageable.

Manual curation was performed in Profinder (build 08.00) by reviewing each found compound individually. As shown in Figure 3.07, for each compound a chromatogram and mass spectrum is displayed in Profinder. It is possible to manually integrate chromatographic peaks for individual samples, groups of samples specified in the experiment or all samples within a dataset by overlaying the extracted chromatograms. In this way, the extracted data were manually curated to correct or optimise the alignment across the samples. It is important to note that this was only performed when considered necessary so as not to inadvertently introduce bias to the data. Figure 3.07A shows a clear example in which manual re-integration was necessary for one sample. For the sample shown in the bottom row, it was evident that chromatographic peak was present that matched the RT in other samples, but that the algorithm had not correctly integrated the peak. Figure 3.07B shows the manually re-integrated data, in which the RT and extracted mass spectrum are now aligned across all samples.
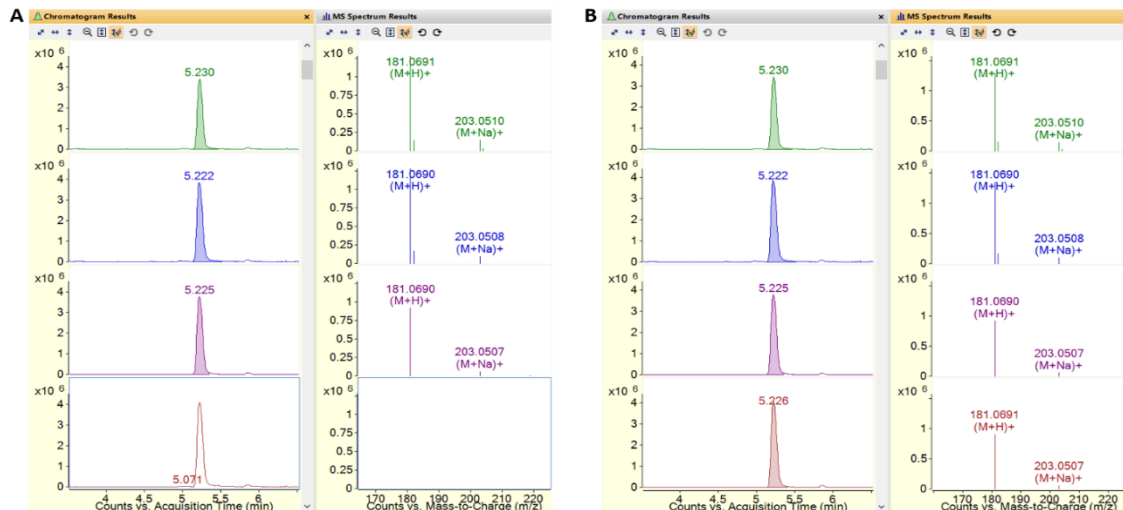
Figure 3.07. Example of manual curation of extracted data. Figure shows the results of AMRT-based feature extraction for xanthine, obtained from human urine in positive polarity. **A:** extracted data before manual curation. It can be seen from the chromatogram results (left panel) for the sample shown in the bottom row that there was a peak present that aligned with the other samples, but this had not been correctly integrated by the algorithm. **B:** results following manual re-integration, showing correct extraction of the aligned chromatographic peak (left panel) with extracted mass spectrum (right panel) for the sample shown in the bottom row.

3.3.3 Export of extracted data

The above data extraction approaches (section 3.1.1) were applied to each experiment dataset in separate Profinder experiments and also separately for positive and negative polarity. This resulted in 4 separate Profinder projects for each dataset comprising targeted and non-targeted extracted data for positive and negative polarity. Each Profinder project was exported as one 'compound exchange format' (CEF) file per sample, which is Agilent's data format for

interchanging between different programs within the MassHunter software suite. This file format encodes a standard metabolomics data matrix, in which each column represents an individual sample, each row a different entity, and abundance values (peak area was employed in the subsequent results chapters) for each entity to populate the matrix (Figure 3.03D). The sample CEF files from each Profinder project were then imported into Mass Profiler Professional (MPP; build 14.5, Agilent), in which the workflow actions in the following sections were performed.

3.4 Data pre-treatment

Data pre-treatment is the process of transforming the extracted, cleaned data to prepare them for statistical analyses. The following workflow actions in this section were performed in MPP (build 14.5) as part of the data import wizard.

3.4.1 Normalisation of urine data based on creatinine

As Chapters 4 and 5 report metabolomic analyses of urine, it is important to consider the need to account for inter-sample differences in urine concentration. Urine is unlike biofluids such as plasma in which solute concentrations are tightly controlled in order to maintain internal homeostasis. Urine volume is greatly influenced by water consumption and other physiological factors, which in turn affects its constituent solute concentrations (212). A routine procedure in urine analyses is therefore to normalise the observed signals for each sample to creatinine concentration; creatinine is a natural waste product of creatine

phosphate in muscle, and is often used for normalisation because it is excreted in urine at a constant, steady rate.

For the patient urine samples studied in this thesis, 24-h creatinine output values were available from the Jaffe creatinine assay on a Cobas 6000 Analyzer (Roche, UK). These raw values were simply entered for each sample data file imported into MPP as an external scalar at the data import stage.

Mouse urine samples were taken on a single-collection basis, therefore it was not possible to obtain 24-h urinary creatinine output for mice. It was also not possible to sub-aliquot part of the mouse samples for Jaffe creatinine analysis due to sample volume constraints. Creatinine was therefore measured for each sample using the internal signal that matched with creatinine by AMRT from the metabolomics data (Figures 3.08A & 3.08B). As creatinine is one of the most abundant compounds in urine (213), it was not surprising to find that the signal from the ^{12}C+H positive polarity ion was saturated, as indicated by an asterisk in Figure 3.08B. However, it can also be seen in Figure 3.08B that the ^{13}C+H ion was not saturated, and this therefore represented a better candidate ion from which to measure the creatinine signal for mice.

Measurement of creatinine by the ^{13}C+H ion signal for mouse urine samples was performed in MassHunter Quantitative Analysis (build 06.00). The raw peak area values from this analysis were used as an external scalar for each sample in the MPP data import stage, as described above for patient samples.
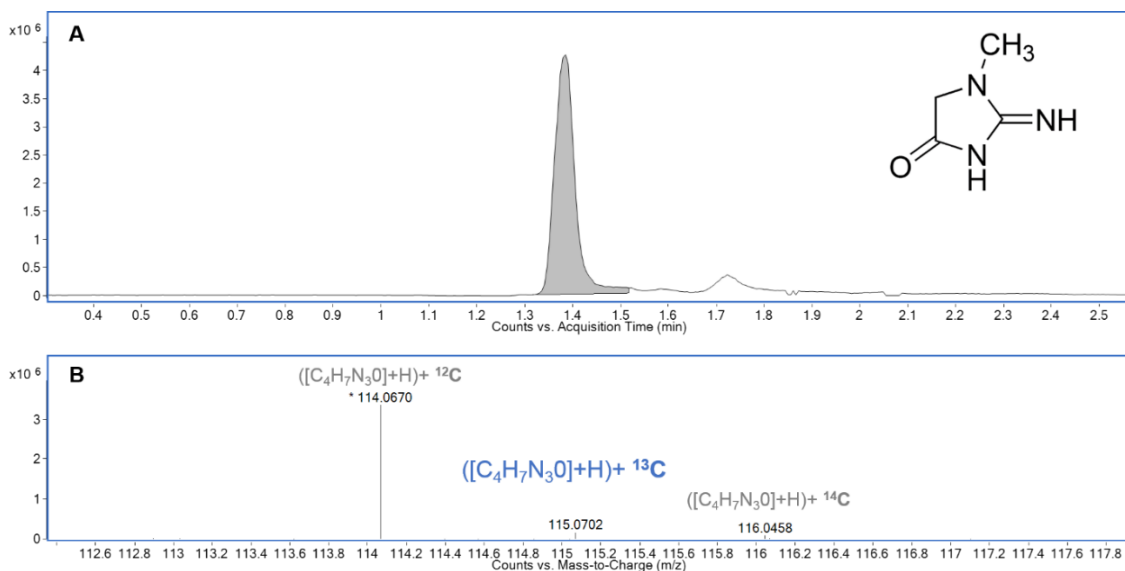
Figure 3.08. Extracted ion chromatogram (A) and mass spectra (B) for creatinine, which was used as an internal scalar for mouse urine data to correct for urine dilution effects. The total peak area representing the ^{13}C +H ion of creatinine (highlighted in blue) in positive polarity was used as an internal scalar for mouse urine as 24-h creatinine values were unavailable. The ^{12}C ion was not used due to saturation of the signal, indicated by an asterisk in the mass spectra.

Currently there is no consensus on the most suitable normalisation strategy for urine metabolomics data. In the method described here, creatinine abundance is used as a proxy for urine concentration, and all observed metabolite abundances are corrected against this. Other normalisation approaches to correct for variation in urine concentration in metabolomics data include those based on urine volume, osmolality and use of data from multiple metabolites or single metabolites other than creatinine (212). It is important to note that urinary creatinine excretion varies, and it is influenced by external stressors such as renal impairment (214). Normalisation based on creatinine was considered appropriate in the

metabolomics studies detailed in subsequent chapters as there is no evidence that urinary creatinine is altered in AKU, neither untreated nor as a consequence of nitisinone therapy. Furthermore, systematic differences in creatinine between the major patient and mouse experimental groups in each dataset (*i.e.* AKU *versus* non-AKU or baseline *versus* nitisinone time points) were checked prior to further metabolomic statistical analyses; all comparisons showed non-significant differences in 24-h creatinine output (patients) or the measured internal creatinine abundance (mice) at the level of $p < 0.05$. This increased confidence that the data were not confounded by creatinine and that the normalisation strategy applied was appropriate.

3.4.2 Data scaling and transformation

Additional procedures performed on the data in this workflow were scaling and transformation. Scaling was performed by 'baselining' each metabolite to its median values across all samples in the experiment. This is referred to as level scaling, which changes the raw abundance values for each metabolite into percentages compared to the median value. The median value was chosen for this scaling as it is considered a more robust alternative to using the mean and focusing on relative response in this way has been previously recommended for identification of biomarkers in metabolomics data (215).

In MPP, imported data are automatically log transformed. Log transformation is a non-linear conversion of data designed to deal with 'skewness' or heteroscedascicity that is typical in bioanalytical data such as measured

metabolomes. These terms refer to the property of metabolomics datasets that the degree of data variation between samples can vastly differ between different metabolites. This unequal variation can be both biological and technical. An example of this phenomenon is the well-documented observation that analytical profiles often show increased variation with increased metabolite intensity, which can lead to a bias towards discovery of statistically significant changes for these more-abundant entities (215,216). The log transformation employed in this workflow is therefore a strategy to reduce the bias towards significant differences for certain chemical entities by means of normalising their variability.

Another possible data processing procedure in LC-MS metabolomics data is RT shift correction. This is designed to account for systematic variations in RT that can sometimes occur across the analytical sequence, for example a gradual drift in RTs over time. It was determined that RT correction was not necessary in this workflow. Figures 3.01 and 3.02 provide representative examples of the high chromatographic reproducibility obtained in the following studies even when sample analysis spanned several days. Furthermore, as shown in Chapter 2 (section 2.3.4), the RT shift of metabolites identified in urine by AMRT by the methods employed in this workflow were minimal; no compound was filtered out due to RT CV <25% across an analytical sequence spanning >48 h. RT correction is perhaps more necessary when analyses using the same method are performed over several batches within the same experiment, but this was not the case for the metabolomics studies described in subsequent chapters.

3.5 Data QC filtering

Data QC filtering was performed using two approaches, first on a per-sample basis and second on a per-entity basis. Here 'entity' refers to a distinct compound signal represented by a group of chemical features (*e.g.* an [M+H]$^+$ or [M-H]$^-$ ion and its associated isotopes and adducts). Entities could be referred to as metabolites, but the term is used cautiously here as non-targeted data extraction could potentially detect low molecular-weight chemical species that do not originate from biological samples but which are introduced artificially during the sample preparation and analysis (*i.e.* chemical contaminants can 'masquerade' as metabolites). Data QC was performed in MPP (build 14.5).

3.5.1 QC on a 'per sample' basis

The primary aim of QC on a per sample basis was to identify clear outlier samples. Samples may become outliers for various biological or technical reasons. These could relate to the subject's metabolism, drug or environmental exposure, or issues with the instrumentation (see identification of outlier samples by TIC; section 3.2.1). The data from such outliers were removed so as not to bias statistical analyses.

PCA was used to identify sample outliers. PCA is a data reduction technique used to characterise the structure of complex, multivariate datasets. Clear sample outliers should be observable in a PCA plot, as these samples would be expected to make a major contribution to the total variance in the dataset, and therefore be separated from the other samples in one of the first principal components. Figure

3.09 shows an example of identifying an outlier sample by PCA. The outlier sample, indicated by an arrow in Figure 3.09A, is separated from the other sample group clusters in principal component 1. The sample also falls outside of the grey ellipse, which represents the Hotteling's T^2 95% confidence interval; Hotteling's T^2 is a measure of the variation within the PCA model. Figure 3.09B shows the same PCA repeated with the outlier sample removed; here all samples now fall within the 95% confidence interval ellipse.
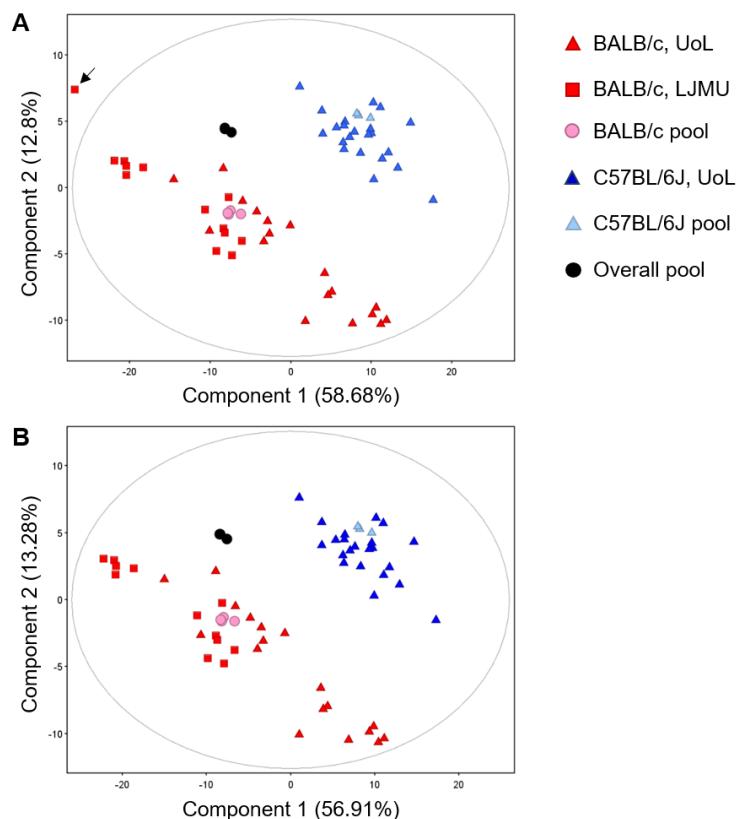
Figure 3.09. Example of principal components analysis (PCA) on LC-QTOF-MS profiling data for quality control (QC) purposes. A: PCA plot of negative polarity urine profiling data from AKU mice. These mice comprised two different background strains; BALB/c and C57BL/6J. The mice were bred from the same laboratory stocks, but 12/18 of the BALB/c mice were housed at Liverpool John Moores University (LJMU); the other mice were housed at the University of Liverpool (UoL). PCA showed clear separation both between BALB/c and C57BL/6J strains, and also between the BALB/c mice housed at UoL compared with LJMU. Note the tight clustering of replicate injections of pooled samples from BALB/c and C57BL/6J strains, which were interspersed throughout the run, showing reproducibility of the analysis. The pooled samples also cluster with the individual samples from the appropriate group, with overall pooled samples clustering between the groups. PCA identified one anomalous sample (indicated by black arrow). **B**: a repeated PCA plot of the above data but with the anomalous sample removed.

3.4.2. QC on a 'per-entity' basis

Pooled QC samples were created for each metabolomics experiment by combining equal aliquots of undiluted individual samples. This gave an overall pooled sample, which contained equal aliquots of all samples included in the experiment, and separate pooled group samples, which contained equal aliquots of samples within a given experiment group. Chapter 4, section 4.2.4 describes how QC samples were incorporated into the design of the analytical sequence.

Per-entity QC was performed in order to filter out non-reproducible entities, which can also introduce bias to downstream statistical analyses. This filtering was based on the data from the replicate injections of pooled QC samples which were interspersed throughout each analytical sequence: *i.e.* the pooled samples for each experiment group, and the overall pooled sample. This was based on the assumption that the pooled samples were representative of each entire experiment dataset, containing all compounds present across the respective sample set. The signals for each entity detected should therefore, in principal, remain consistent across the replicate injections of these pooled QC samples throughout the analytical sequence.

First, entities were filtered based on frequency in pooled QC samples. Entities were retained if they were observed in 100% of replicates from at least one pooled QC sample. This threshold enabled entities that might only have been observed in one pooled QC sample to remain, but only if they were observed consistently across the replicate injections of this sample.

Filtering was then performed based on signal variability across the pooled QC replicates. Entities were retained if their peak area CV was <25% across all replicate injections within each pooled QC sample. Figure 3.10 illustrates a representative example of how this QC step filtered for entities with reproducible profiles across replicate injections of pooled QC samples. As shown in the Figure 3.10A, profiles of a number of chemical entities (represented by individual lines in the profile plot) showed fluctuations in abundance across the replicate injections of each pooled QC. Fluctuations in these entities could be attributable to technical variation potentially introduced at various stages throughout the data acquisition and processing workflow described above but could also relate to the chemical instability or degradation of these compounds over the analytical period. Figure 3.10B shows the same profile plot after application of the aforementioned variability filter; here it is clear that non-reproducible entities across the replicates in each group have been removed.
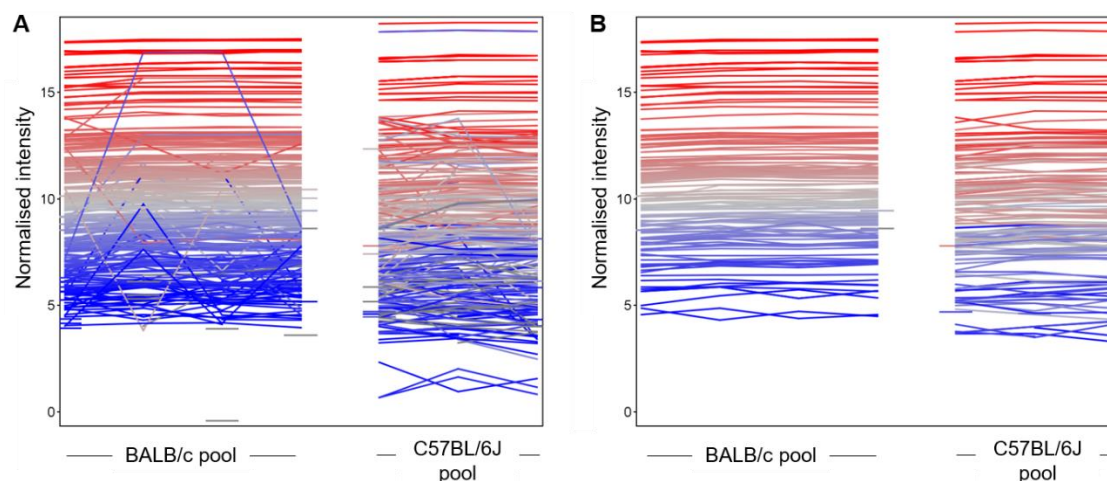
Figure 3.10. The effect of quality control (QC) on profile plots of chemical entities across pooled QC samples. A: profile plot for negative polarity entities from replicate injections of pooled urine samples from BALB/c and C57BL/6J AKU mice interspersed throughout the analytical run pre- data QC. **B**: profile plot for the same data as in (A) following entity QC filtering based on (i) entities detected in 100% of replicate injections for at least one pooled sample group and (ii) entities with CV <25% across all replicate injections of each pooled sample. The profiles in (B) indicate removal of (189/365) less-reproducible entities across replicate injections.

3.6   Statistical analyses: identification of differentially abundant entities

As discussed in Chapter 1, multivariate statistical analyses are required to interpret the complex datasets such as those obtained from metabolomics experiments. This section describes the multivariate statistical techniques employed to investigate potential differences between sample groups in the following experimental chapters. These analyses were also performed in MPP (build 14.5).

125

3.6.1. Principal component analysis (PCA)

PCA was employed to review the overall structure of each dataset and to identify any potential separation between different sample groups based on the first principal components, using 4-component models. This approach effectively reduced the data in each experiment to 4 orthogonal components (*i.e.* each sample was plotted in 4-dimensional space corresponding to these 4 principal components) that explained the most variation in the dataset. Figure 3.11A shows an example of how plotting each sample according to its position/value in the first 2 principal components can identify clear differences between the overall metabolome profiles between different sample groups. In this example, BALB/c mice bred from the same laboratory stocks were separated in principal components 1 and 2 according to whether they were housed at the University of Liverpool (UoL) or Liverpool John Moores (LJMU) laboratories. This difference was attributed to slightly differences in diet and/or water supplied by the two laboratories.

Review of the corresponding PCA loadings plot was performed in order to identify the entities that were contributing most to any separation observed between sample groups. Figure 3.11B shows an example of identifying important entities that contributed to PCA sample separation observed in Figure 3.11A. In Figure 3.11B, these are the labelled entities that have the most extreme loading scores (high or low) on principal components 1 or 2.
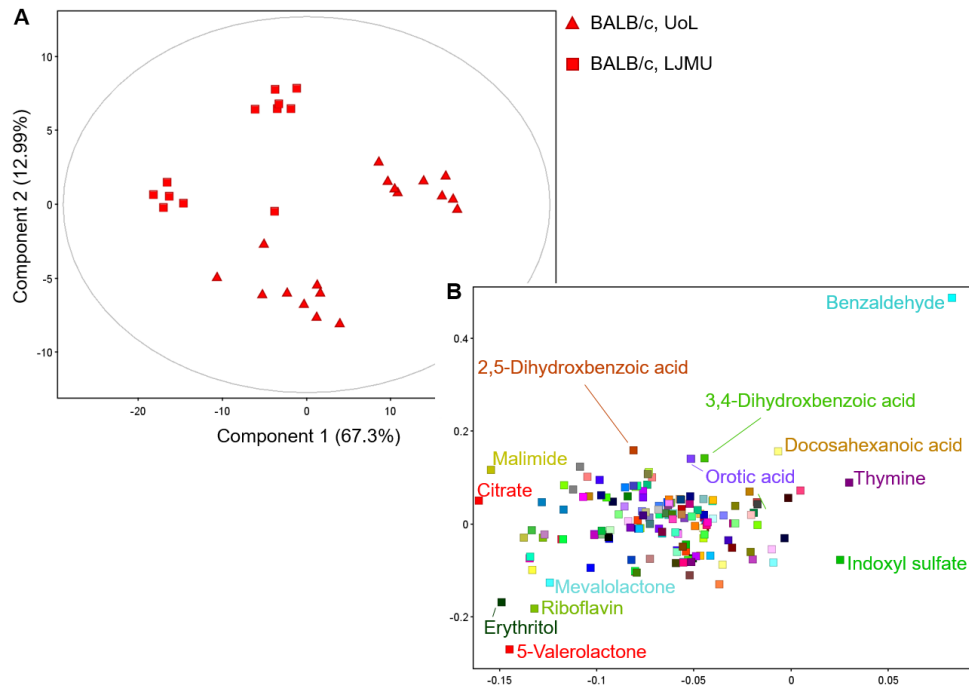
Figure 3.11. Example principal components analysis (PCA) as an approach to identify differentially abundant chemical entities. A: PCA plot showing separation between the urine metabolomes of BALB/c AKU mice housed at the University of Liverpool (UoL) and Liverpool John Moores University (LJMU) (in principal components 1 and 2, represented by *x* and *y* axes respectively). **B**: loadings plot showing the contribution of chemical entities to variation in PCs 1 (*x*-axis) and 2 (*y*-axis). The labelled entities had the greatest loadings on PC 1 and 2, indicating that these made the greatest contribution to the overall difference between mice from the two sites.

3.6.2   Statistical significance analysis

PCA was used as an approach to visualise and explore potential clear separation between the overall metabolome profiles of different sample groups. Statistical significance testing was performed in order to more extensively investigate potential group differences. The specific tests employed for each comparison are

described in the subsequent results chapters (Chapters 4 and 5). In order to mitigate the risk of false-positive results associated with multiple testing, Benjamini-Hochberg false-discovery rate (FDR) p-value adjustment was employed in all significance tests.

Fold change (FC) was also calculated as an additional way to assess the magnitude of change for a given entity between groups. Figure 3.12 shows an example of how volcano plots were employed as a way of incorporating the p-value and FC to visualise the differentially abundant entities between groups. This example also shows how entities that do not necessarily contribute towards clear separation between the sample groups in PCA (Figure 3.11) can be identified as differentially abundant by p-value and FC (*i.e.* α-ketoglutarate, 3-hyroxybenzoate and methyl-β-D-galactoside in Figure 3.12).
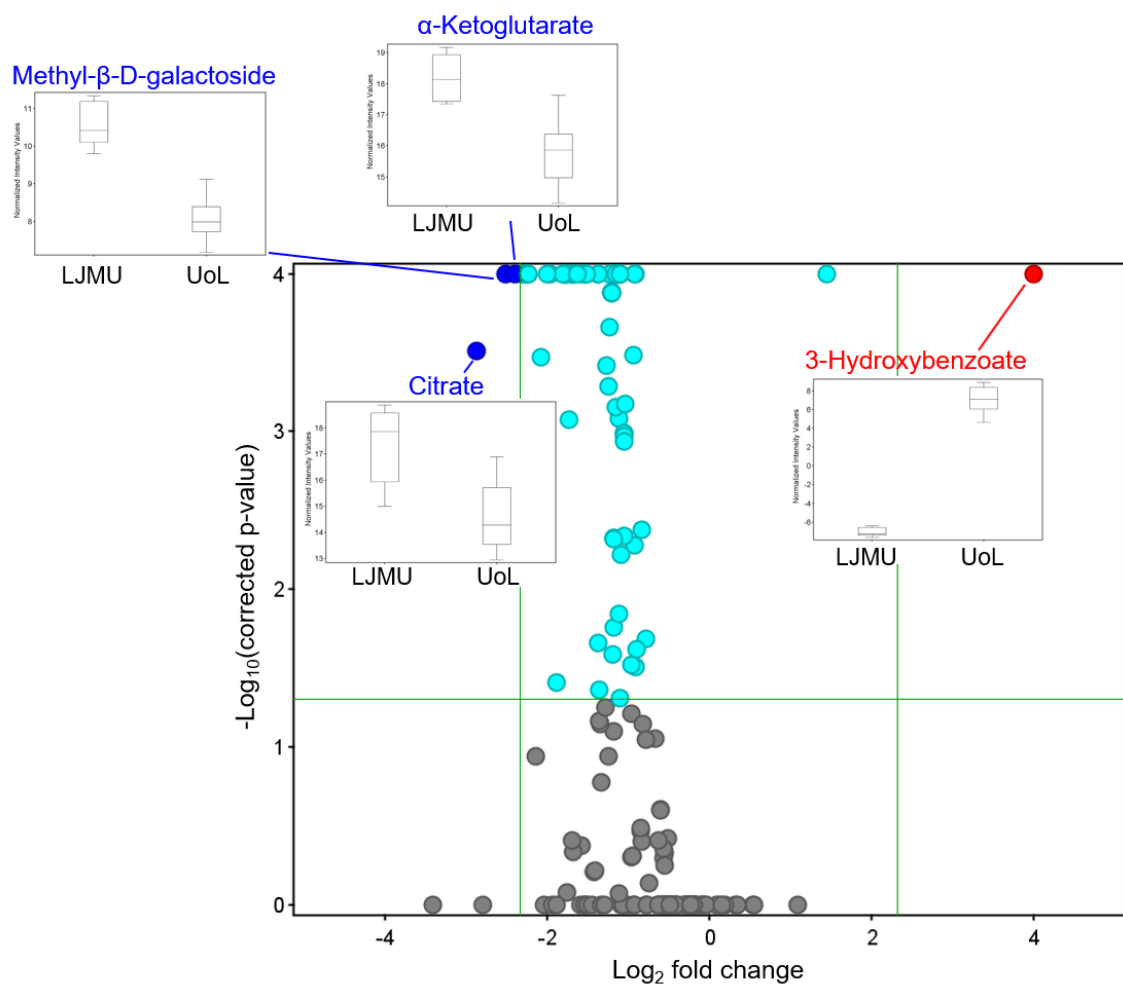
Figure 3.12. Example volcano plot as an approach to identify differentially abundant chemical entities. Entities with p <0.05 (*y*-axis) and fold change >5 (*x*-axis) in peak area between urine from BALB/c AKU mice housed at the University of Liverpool (UoL) compared with Liverpool John Moores University (LJMU) are highlighted in red (up in UoL) and blue (down in UoL).

3.6.3 Data validation

As previously noted (section 3.1), there is an inherent risk for 'type 1' false-positive statistical errors in complex datasets such as those acquired in metabolomics. While the data processing workflow described this far was designed to minimise the FDR, the data from entities identified as important from statistical testing were subjected to further validation to rule out artefacts.

For targeted AMRT-extracted datasets, the profiles across sample groups were assessed visually in MPP by box plots (as shown in Figure 3.12) to confirm the group difference. For these entities, the extracted, untreated profiles were also checked visually in Profinder across the sample groups as additional validation.

For datasets from non-targeted feature extraction, an additional orthogonal data extraction approach was employed by using MassHunter Quantitative analysis (Agilent, build 06.00). This additional validation was performed because these datasets were not manually curated in Profinder, unlike the datasets from targeted AMRT feature extraction. A list of the differentially abundant entities was exported from MPP using the export for recursion option; this exported the measured AM and RT of the entity as a CEF file. The CEF file generated was then imported as a quantitative method in the Quantitative analysis software and used to re-mine the raw, unprocessed data files for these signals. The raw peak areas extracted using this method were then imported to GraphPad Prism (version 6.01; Graphad Software Inc., USA), where the data distribution between the sample groups were reviewed visually by dot plots. Only entities that still showed clear visual trends

(*i.e.* peak area difference) between the groups using this method were considered as validated.

3.7   Chemical structure identification of entities from non-targeted feature extraction

Putative chemical structure identifications were sought for validated differentially abundant entities from non-targeted feature extraction. This was performed using MassHunter ID Browser (Agilent, build 07.00), using the chemical formula sources from the MassHunter METLIN metabolites PCD/PCDL accurate mass database (Agilent, build 07.00). This version of the METLIN library contains formula and AM entries for >32,000 metabolites. The list of differential entities created in MPP was exported directly into ID Browser. The AM matching parameters employed in each experiment are described in subsequent chapters. The output of this procedure contains each exported entity, with potential METLIN metabolite matches based on the parameters and tolerances specified. The matches are ranked by a composite score, based on the degree of closeness between the expected *versus* observed mass and isotope distribution from the database formula.

The METLIN database employed comprises a broad range of metabolites of both endogenous and exogenous origin. As the identification method described was based on formula-matching alone, it is necessary to review the matches. It is often challenging to assign a single chemical structure to a given entity using this approach due to the high number of isomers in comprehensive metabolite databases (compounds with the same chemical formula) and isobaric species

that have a different chemical formula but the same AM. Here, it was considered pragmatic to putatively assign the most likely candidate based on an informed judgement of AKU metabolism and the compounds that were likely to be observed or altered in this context of urine profiling. For example, the matches obtained for exogenous metabolites of a specific drug or environmental exposure that was either not possible for the laboratory mice, or highly unlikely for the patients, were rejected. For other matches however, it was often more difficult to distinguish between multiple hits with similar match scores. While this approach of matching based on formula alone is informative, it is acknowledged that the compound identifications from non-targeted data extraction are putative and require further validation, ideally by MS/MS data and/or analysis of chemical standards (if available). The number of alternative compound matches for putatively identified entities from non-targeted feature extraction are also specified in subsequent chapters.

3.8   Workflow summary and conclusions

The chemometric workflow described provides a method to extract biologically meaningful information from the large, complex datasets generated in LC-QTOF-MS metabolomic analyses. Figure 3.13 provides a schematic overview of the workflow. The workflow is designed to identify differentially abundant chemical entities, or biomarkers in the following metabolomic analyses in AKU. Emphasis was placed on a) maximisation of the information extracted from the raw data, and b) limiting false discovery.

Feature extraction is primarily performed using the AMRT-based strategy described in Chapter 2, but with complementary non-targeted extraction performed in parallel to broaden the extraction beyond the compounds in the AMRT databases generated. The workflow incorporates a series of QA and QC measures to limit the FDR, which involve identification of outlier samples and filtering for reproducible entities based on pooled QC samples. The post-statistical validation of differentially abundant metabolites was performed by review of both the processed and unprocessed extracted data. For data from non-targeted extraction, additional confirmation was performed by re-mining the raw data using orthogonal data extraction (note additional flow arrow back to the raw data in Figure 3.13). These confirmatory procedures limit the potential for false discovery that is caused by analytical or data processing artefacts.
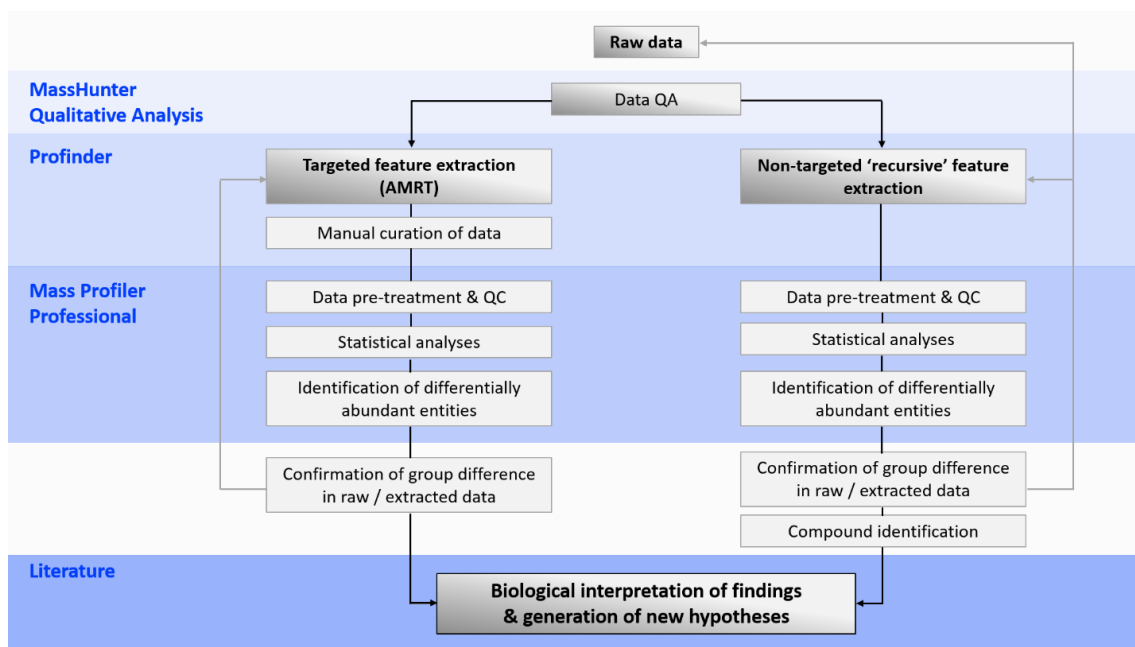
Figure 3.13. Overview of the LC-QTOF-MS chemometric workflow developed. The figure summarises the computational steps taken, starting from raw, unprocessed data and ending with the extrapolation of biologically-relevant information from the data based on the identification of differentially abundant entities/metabolites. The workflow comprises parallel targeted (AMRT-based) and non-targeted feature extraction on each dataset in order to extract maximal information from the dataset. The software programs used in each stage are listed on the left-hand side of the figure.

Metabolomics is a burgeoning field, which is reflected by the ever-increasing development of new computational tools for data analysis. The need for harmonisation in computational approaches employed across the discipline has been highlighted in an international survey of the metabolomics community in 2017 (100). A related but equally important consideration is the most suitable type of QC sample to use in metabolomics experiments. The present workflow uses pooled QC samples, which play a major part of the QA and QC filtering

134

procedures described here. QC filtering of entities based on pooled samples is particularly important in the data from non-targeted feature extraction here, as this type of processing typically resulted in several thousand extracted entities that were therefore beyond the scope of cleaning or curating manually. According to a 2017 survey of the metabolomics community conducted by the Metabolomics Society Data Quality Task Group (102), the majority of researchers in this area employ pooled project material for QC; 83% of the respondents. This approach was more popular than standard reference materials for QC, which were applied by 48% of respondents. There was variation between researchers in the way that pooled samples were used for QC however; 76% of respondents used QC samples for assessing chromatography column integrity, 69% applied replicate injections of QC samples and 15% analysed a single pooled sample. Importantly, the survey highlighted that 70% of respondents did not have access to a protocol for review of metabolomics data quality. This statistic underlines the importance of detailing clear QA and QC procedures to be employed in data processing, as in the workflow described in this chapter.

4.0 INVESTIGATING THE EFFECT OF NITISINONE ON THE URINE METABOLOME IN PATIENTS ATTENDING THE UK NATIONAL ALKAPTONURIA CENTRE (NAC) AND AKU MICE

Published work resulting from content included in this chapter (see Supporting Material):

The above works report the data from metabolic profiling of urine from patients and mice before and after nitisinone treatment by employment of the analytical strategy reported in Chapter 2. BPN contributed to the laboratory analyses performed for sample analysis and was the main contributor to the writing of the manuscript for this publication.

4.1   Introduction

Nitisinone is not currently licensed for AKU but has been used to treat patients attending the NAC in Liverpool, since 2012. Nitisinone works by inhibiting the activity of the enzyme HPPD (E.C. 1.13.11.27) responsible for the formation of HGA. It has been shown to reduce circulating and urinary HGA concentrations (11,12,14,18,217) and to completely arrest ochronosis in an $HGD^{-/-}$ murine model of AKU (50,69) and more recently in AKU patients (140).

In HT-1, in which nitisinone is a licensed treatment, and AKU, one of the major upstream metabolic consequences of nitisinone is that the concentration of circulating tyrosine increases significantly (11,12,14,46,217–220). The fate of the increased tyrosine is unknown. Tyrosine is the precursor for catecholamines, thyroid hormones and melanin, suggesting that increased tyrosine substrate, as in tyrosinaemia, could lead to dysregulation of these pathways. Concerns exist because in HT-1 it is believed that hyper-tyrosinaemia may contribute to neurodevelopmental delay (46,221–224); it is estimated that up to 35% of children treated with nitisinone have learning disabilities (46). Other undesirable side effects of nitisinone which metabolic phenotyping has potential to shed let on are ocular keratopathy and skin rash. These have been reported in a small minority of individual cases in the AKU literature (144), though such patients were not involved in this study as nitisinone treatment is withdrawn in these patients. Nevertheless it would be of clinical benefit to establish predictive markers to identify patients with susceptibility to these symptoms, whose treatment regimen is likely to require adjustment accordingly.

This chapter explores the impact of nitisinone on the urine metabolomes of mice and patients with AKU. Human AKU urine was from patients attending the NAC taken at baseline (when patients had not been administered nitisinone before) then after 3, 12 and 24 months on nitisinone. Mouse urine was from $HGD^{-/-}$ BALB/c mice taken at baseline then at 1 week on nitisinone (Figure 4.01). These treatment durations therefore represent an acute and chronic exposure to nitisinone in mice and humans, respectively.
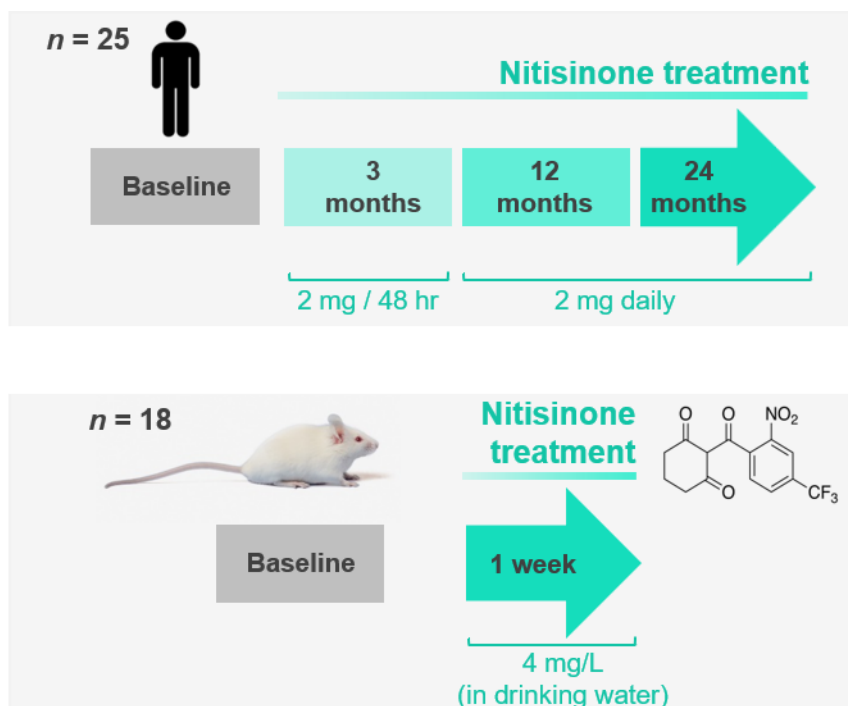
Figure 4.01. Summary of the overall study design comprising patients and mice with AKU. Urine was taken before and after commencement of nitisinone therapy to evaluate the consequences on the metabolome by LC-QTOF-MS metabolic profiling.

4.2 Methods

4.2.1. Chemicals

All chemicals were as described for LC-QTOF-MS analysis in Chapter 2 (section 2.2.2).

4.2.2. Urine collection

24-h urine was obtained from 25 AKU patients attending the NAC (12 male, 13 female; mean age ±SD 51 ±15 years). Samples were collected before nitisinone, then at 3 (2mg every other day), 12 and 24 (2mg daily) months on nitisinone and stored at -80°C until analysis. Samples were diluted 1:3 urine:deionised water and stored at -80°C as three separate aliquots for analysis by methods 1-3 (see Chapter 2, section 2.2.2). Metabolomics analysis was part of the diagnostic service to NAC patients with approval from the Royal Liverpool and Broadgreen University Hospitals Trust Audit Committee (audit no. ACO3836).

Mouse urine was collected from 18 (9 male, 9 female; mean age ±SD 27 ±12 weeks) $HGD^{-/-}$ BALB/c mice (50), bred from laboratory stocks at the University of Liverpool. Samples were obtained on a single-collection basis before treatment then after 1 week on nitisinone, administered in all drinking water (4 mg/L), supplied *ad libitum.* Mouse urine was collected onto cling film, pipetted into sample tubes and stored at -80°C. Analysis was performed on thawed samples, following dilution of 1:9 urine:deionised water (greater dilution than patient urine due to sample volume constraints). Mouse breeding and dosing was authorised under the Animals (Scientific Procedures) Act, UK.

Human and mouse urine samples were pooled separately for QA purposes. For each sampling time point, a representative sample was created by pooling 20 µL of each individual urine sample. An additional overall pool was created separately for human and mouse urine experiments by pooling equal proportions of the pooled urine groups detailed above. Pooled human and mouse samples were stored and treated as above prior to analysis.

4.2.3. LC-QTOF-MS analytical conditions

Non-targeted LC-QTOF-MS analysis was performed on human and mouse urine samples separately using the strategy detailed in Chapter 2. Human urine was analysed by LC methods 1-3. Mouse urine was analysed by LC method 2 only due to limited sample volumes. The mass spectrometer was operated in non-targeted MS mode (*m/z* range 50-1700; see Chapter 2 for detailed parameters).

4.2.4. Design of metabolic profiling experiments

Human and mouse urine were analysed separately. Human samples were analysed in three separate batches for methods 1-3. Mouse samples were analysed in one batch. Each batch comprised analysis in negative polarity followed by positive polarity.

The analytical sequence of each profiling batch was designed according to published guidance (225). Each run commenced with 20 replicate injections of the overall pooled sample to condition the system. The order of individual samples was randomised computationally. Pooled samples were interspersed throughout the analytical sequence; every 10th injection. Injections of each sample group

pool and the overall pooled sample were also placed at the start (post-conditioning) and end of each analytical sequence to assess reproducibility of the analysis for data QA (Figure 4.02).
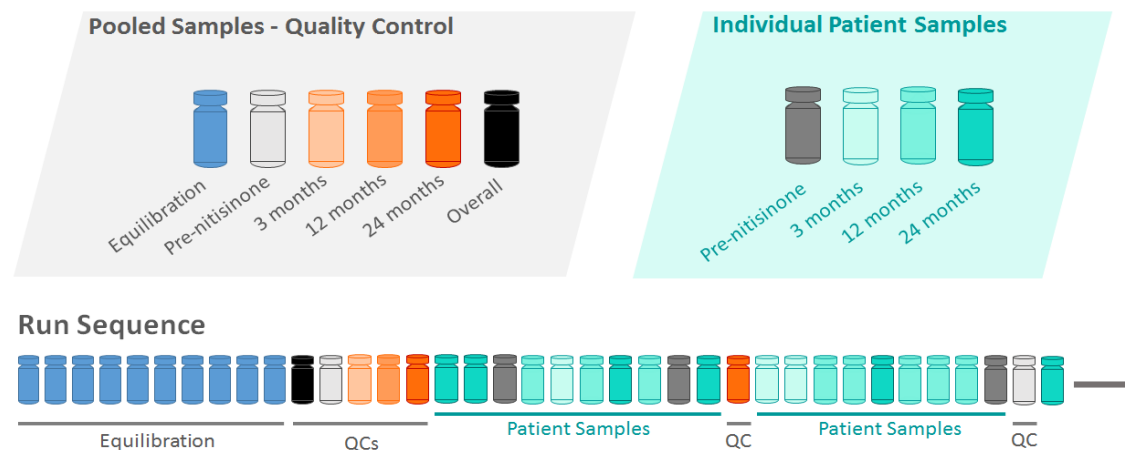
Figure 4.02. Illustration depicting the analytical sequence design for each LC-QTOF-MS profiling batch. The first 20 injections were of the overall pooled sample for system equilibration. Post-equilibration, each run commenced and ended with pooled samples from each sampling time point. Pooled samples were randomly interspersed throughout the sequence at regular intervals for QC purposes. The sequence of individual patient samples from each time point was randomised computationally across each analytical batch.

4.2.5. LC-QTOF-MS data processing

Raw data were mined using two parallel strategies as outlined in Chapter 3 (Figure 3.14). Targeted feature extraction was employed to search for compounds from the AMRT databases generated from the IROA MS metabolite library of standards, as outlined in Chapter 2. Non-targeted feature extraction was

employed to search for compounds not included in the AMRT databases or of known association to nitisinone/AKU metabolism.

4.2.5.1   Targeted data mining based on AMRT databases

Data were mined for AMRT database compounds using targeted feature extraction in Profinder (build 08.00). Feature extraction parameters were the same as described for urine in Chapter 2 (section 2.2.3.) AM (±10 ppm) and database RT (±0.3 min) matches were required.

Data were also mined by targeted feature extraction for the following compounds that were not from the IROA metabolite library AMRT databases, but of particular interest for their predicted role in AKU/nitisinone metabolism. These compounds were associated with a) increased tyrosine: 3-(4-hydroxyphenyl)pyruvic acid, acetyl-L-tyrosine, γ-glutamyl-L-tyrosine (226) and tyramine-sulfate; b) ochronotic pigment: dihydroxybenzaldehyde (see Chapter 6, section 6.3.2), hipposudoric acid and norhipposudoric acid (42); or c) nitisinone metabolism: hydroxy-nitisinone, nitisinone and 2-nitro-4-trifluoromethylbenzoic acid (NTFA) (226). For non-AMRT compounds, the same RT (RT range <0.3 min) was required across samples. These compounds were identified based on theoretical AM ±10 ppm only, except for 3-(4-hydroxyphenyl)pyruvic acid and acetyl-L-tyrosine, which were analysed as standards and appended to the AMRT database for AMRT-based feature extraction as described above.

Allowed ion species for positive ionisation mode were: H^+ and Na^+, with the addition of NH_4^+ for methods 1 and 2 (mobile phases contained ammonium

formate); and for negative ionisation mode H⁻ with the addition of CHO_2^- for methods 1 and 2. Monomers and dimers were allowed for both ionisation polarities. Charge state range was 1-2.

4.2.5.2.   Non-targeted data mining

Non-targeted data mining was performed using recursive feature extraction in Profinder (build 08.00). Chemical entities were aligned across samples with mass and RT windows of 50 ppm and 0.3 min respectively (see Supporting Materials, Appendix 2 for details). Allowed ion species were as above. Entities were retained if observed in at least 2 samples per group (*i.e.* sampling time point) for human urine, or at least 50% of samples per group for mouse urine (*n.b.* stricter parameters for mice as diet and environment were controlled).

4.2.6. NAC patient demographics and clinical data

Data were collected from the patients involved in this study on a number of demographic and clinical parameters, which included gender, age, PET imaging data, *HGD* gene mutation data and scores from self-reported back pain and Beck's Depression Inventory II (BDI-II).

¹⁸F-NaF PET scanning was performed to assess spinal degenerative changes associated with ochronosis (data previously reported by Taylor (227)). With this technique, the injected ¹⁸F-NaF tracer is taken up into hydroxyapatite crystal, enabling quantification and assessment of the spatial distribution of calcified tissue in the spine, including the vertebrae and fibrocartilaginous intervertebral

disks. ^{18}F-NaF uptake is greater in newly formed bone and therefore reflects the amount of actively mineralising bone, or the rate of bone metabolism more generally or blood supply. Increased uptake to bone has been observed in early stage OA, in which it is attributed to the accelerated bone remodeling associated with OA pathophysiology (228). Although ^{18}F-NaF PET is typically used for studying bone as opposed to soft cartilaginous tissue, a previous study of this patient cohort found that ^{18}F-NaF uptake to the lumbar and thoracic intervertebral disks was markedly increased in patients with AKU compared with controls. Increased ^{18}F-NaF uptake in AKU was thought to reflect mineralisation, deposition of calcium hydroxyapatite or calcium pyrophosphate, of intervertebral disks as a consequence of cartilage degeneration associated with ochronosis. In contrast, ^{18}F-NaF uptake to the bony vertebrae, a measure of bone turnover, was comparable between patients with AKU and controls, showing a generalised decline with age in both groups (227). Here, the imaging data studied were mean ^{18}F-NaF uptake values across the intervertebral disks and vertebrae of the lumbar and thoracic spine, as assigned to these patients by Taylor (227). ^{18}F-NaF PET data were available for a sub-set of patients from the following visits to the NAC: baseline (n=15), 12 months on nitisinone (n=14) and 24 months on nitisinone (n=17).

Data on back pain were available for a subset of 16 patients at baseline only; data previously described by Taylor (227). Back pain was self-reported across lumbar and thoracic regions separately, on a scale ranging from 0 (no pain) to 10 (severe

pain). The mean scores across lumbar and thoracic regions for each patient were used for the purpose of the present analyses.

The BDI-II (229,230) is a 21-item, self-completed questionnaire designed to assess depression. Each item has four statements ranked in order of severity of depression (0-3). The questionnaire evaluates a range of moods and emotions linked with depression, including: pessimism, sense of failure, self-dissatisfaction, guilt, punishment, self-dislike, self-accusation, suicidal ideas, crying, irritability, social withdrawal, insomnia and loss of appetite. Total scores are divided into the following cut-offs to provide an overall measure of depression: 0-13 for minimal depression, 14-19 for mild depression, 20-28 for moderate depression and 29-63 for severe depression.

Data on *HGD* mutations were from sequencing of the *HGD* gene (231). Sequencing data were available for all patients included in this metabolomic analysis (*n*=25).

4.2.7. Data QC, statistical analyses and association with patient demographics and clinical data

Additional QC filters were applied to the data, as detailed in Chapter 3 (section 3.4). QC and statistical analysis were performed in MPP (build 14.5). First, urine creatinine values were used to scale metabolite abundances for each sample in order to account for urine dilution effects. 24-h creatinine excretions were used as an external scalar for human urine (average values across each sampling time point used for pooled samples). For mouse urine, the LC-QTOF-MS signal

identified as creatinine by matched AMRT was used as an internal scalar for each sample using the peak area of the ^{13}C [M+H]$^+$ ion (see Chapter 3, section 3.4.1). In short, QC was performed on a per-sample and per-entity basis. Per-sample QC identified anomalous samples using PCA. Per-entity QC was based on the data from pooled samples. Non-reproducible entities across replicate injections of the same pooled sample across each analytical run were removed from the data.

Imputation was not used, as datasets from targeted extraction were carefully curated prior to statistical analyses in order to avoid missing data which could result from analytical or data processing errors, such as incorrect feature alignment across samples. Furthermore, filtering of entities in all data was performed based on QC samples (targeted and non-targeted feature extraction datasets), and additional post-validation of statistically significant entities (Chapter 3, section 3.6.3) helped to minimise the potential for spurious findings attributable to missing data.

Statistical analyses were performed in MPP based on peak area. Human urine profiles were compared at 3, 12 and 24 months on nitisinone with baseline by one-way repeated-measures ANOVA. Mouse urine profiles were compared at baseline vs 1 week on nitisinone by paired t-test. Benjamini-Hochberg FDR adjustment was used in all statistical significance testing. FCs were calculated based on raw peak area. PCAs employing four-component models were also performed on each filtered dataset.

For statistically significant entities for each comparison in the data from non-targeted feature extraction, additional confirmatory quality checks were performed in MassHunter Quantitative Analysis (Agilent, build 09.00), as described in Chapter 3 (section 3.6.3). The results from this additional data extraction were assessed visually by dot plots in order to compare the distribution of peak areas across individual samples in order to validate or reject the results from statistical analyses performed in MPP. This procedure was performed for the data from non-targeted extraction due to the increased number of entities post-QC and therefore the increased risk of false-positives.

4.3 Results

Data in the following sections are divided into those from targeted and non-targeted feature extraction. In assessing the effect of nitisinone treatment on the urine metabolome, for targeted feature extraction data from the three complementary LC techniques were considered, as this was the primary focus of the study. For non-targeted extraction, data from LC method 2 (see Chapter 2, section 2.2.2) were considered only, as this method provided the greatest overall coverage of the metabolome, as shown in Chapter 2.

For studying associations with patient demographics and clinical data, the metabolomic data were from LC method 2 only. This was considered the most pragmatic approach, as it was previously shown that method 2 provided the greatest overall coverage of the metabolome (see Chapter 2, section 2.3.4), and data were acquired from both human and mouse samples using this method.

The total number of compound identifications (from targeted feature extraction) and unidentified chemical entities (from non-targeted feature extraction) obtained across each dataset are summarised in Table 4.01 for human data and Table 4.02 for mouse data. Data from one patient were excluded from each of the datasets from methods 1 and 2 in positive polarity following the identification of at least one of their samples as clear outliers from PCA.

Table 4.01. Summary of the number of chemical entities obtained and retained across human urine datasets and at each subsequent step of quality control (QC) filtering across human urine datasets.

* Note: CV <25% filter was applied across all replicate injections in 4/5 pooled QC sample groups only (all entities removed with the filter applied to 5/5 QC sample groups as used for the other datasets)

| Filtering step | Targeted feature extraction | | | | | | Non-targeted extraction | |
| | Method 1 | | Method 3 | | | | Method 2 | |
	(+)	(-)	(+)	(-)	(+)	(-)	(+)	(-)
Manual curation of entities only	391	162	205	252	286	208	4042	448
Present in 100% of replicate injections of at least one pooled group sample	269	133	174	165	269	133	2869	315
CV <25% across all replicate injections of each pooled group sample	154	64	55	125	133	75	1594	122*
P <0.05 (FDR-adjusted) and FC >2: baseline vs 3, 12 or 24 months on nitisinone (targeted), or baseline vs all time points on nitisinone (non-targeted)	12	9	10	3	3	7	68	14

Table 4.02. Number of entities obtained and retained before and after each subsequent step of quality control (QC) filtering across mouse urine datasets.

Filtering step	Mouse: baseline vs 1 week on nitisinone			
	Targeted		Non-targeted	
	(+)	(-)	(+)	(-)
Manual curation of entities only	399	374	2040	2671
Present in 100% of replicate injections of at least one pooled group sample	363	364	1278	2326
CV <25% across all replicate injections of each pooled group sample	192	185	208	341
P <0.05 (FDR-adjusted) and FC >1.5: baseline vs 1 week on nitisinone	27	21	26	49

4.3.1. Effect of nitisinone treatment

4.3.1.1.   Targeted feature extraction

PCA on the human data from targeted feature extraction showed clear separation between the samples taken at baseline compared with the time points on nitisinone Figure 4.03). This separation was in PCA component 2 (*y*-axis), which explained 12.14% and 14.68% of the total variation in the dataset in negative and positive polarity respectively. The PCA loadings plots in Figure 4.03 (lower section) show that a clear group of metabolites contribute to this separation, as they have the most extreme loadings on component 2 (comprising positive and

negative values) and therefore change with nitisinone treatment. The metabolites with the greatest loadings on component 2 in negative polarity (left plot) were homovanillic acid (HVA), D-(+)-galactosamine, 4-coumarate, phenylacetic acid, 3-(4-hydroxyphenyl)lactic acid (HPLA) and 4-hydroxybenzaldehyde. In positive polarity (right plot), HGA, xanthosine, 3,5-cyclic AMP, L-kynurenine, L-tryptophan, 4-hydroxyphenylacetic acid, tyramine, HPLA and 4-hydroxybenzaldehyde had the greatest loadings on component 2.

PCA on the mouse data also showed clear separation of samples taken before and after nitisinone in component 2 (Figure 4.04). Loadings plots (lower) show that the metabolites with the greatest loadings on component 2 were HPLA and phenylacetic acid in negative polarity (left plot), as for the human data. In positive polarity, HPLA and 4-hydroxyphenylacetic acid were among the group of metabolites with the greatest loading on component 2, as for the human data. However, the positive polarity mouse data (right plot) showed the addition of metabolites identified as 2-hydroxy-4-(methylthiobutyric acid), *N*-acetyl-L-phenylalanine and 3-MT to this cluster of compounds.

Negative polarity

Positive polarity

Baseline ⚫

Nitisinone

3 months ○
12 months ○
24 months ⚫
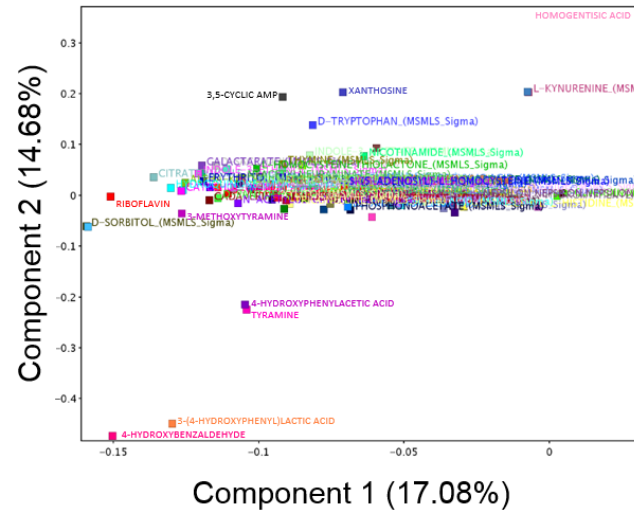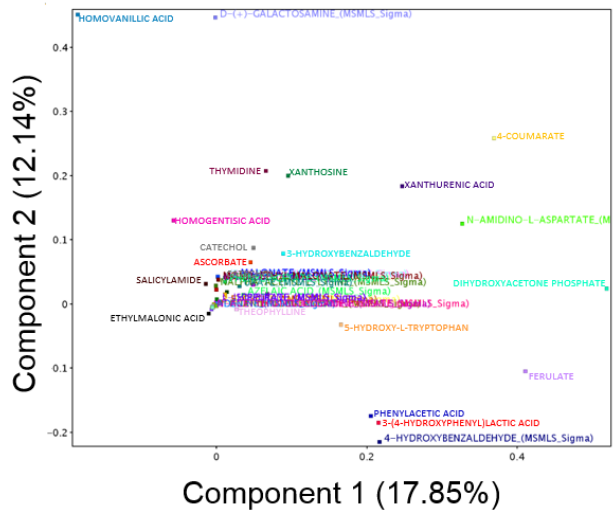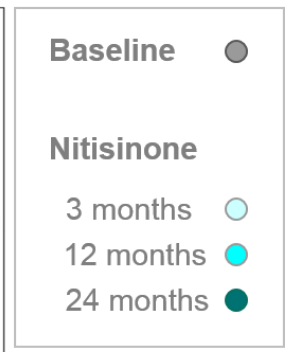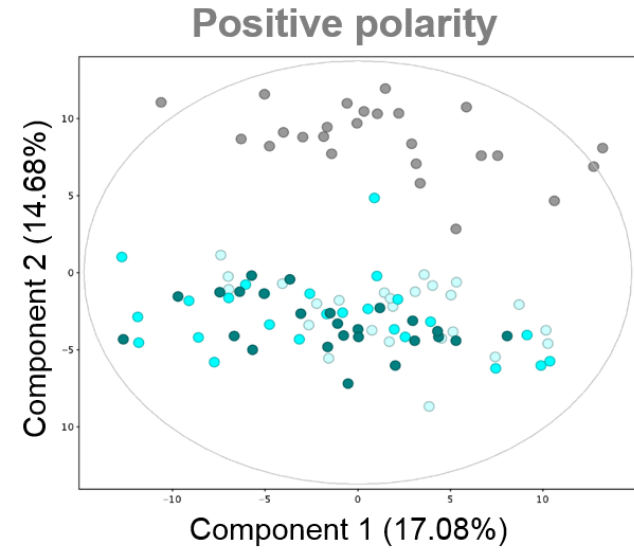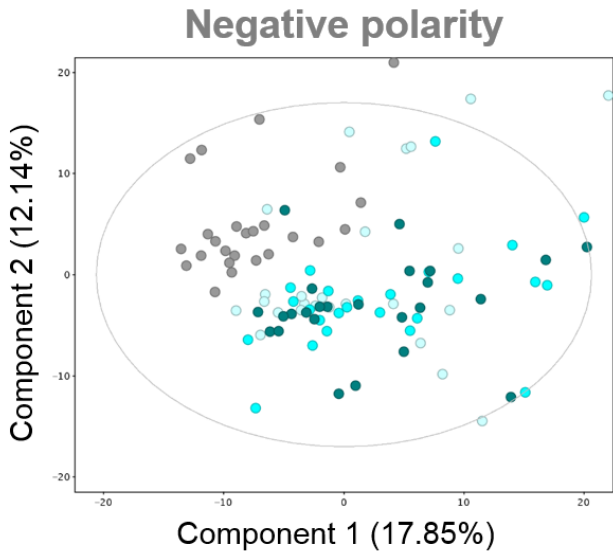
Figure 4.03. PCA to assess the effect of nitisinone treatment on profiles from AKU human urine (targeted feature extraction). PCA plots (upper) in negative (left) and positive polarity (right) both showing clear separation in component 2 between profiles from urine samples taken at baseline compared to those taken at 3, 12 and 24 months on nitisinone. Loadings plots (lower) show the respective contributions of individual metabolites to components 1 and 2.

Negative polarity
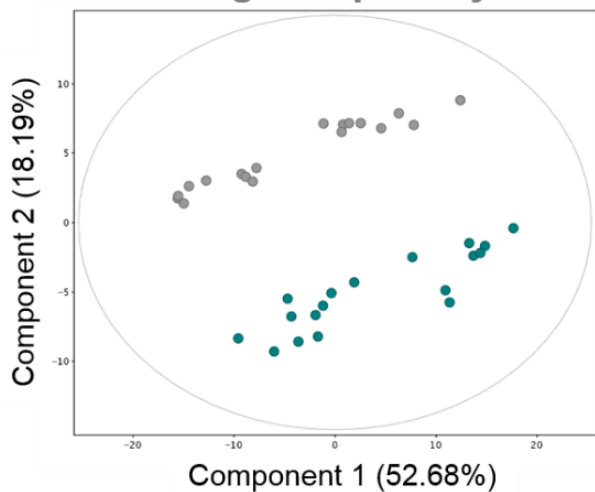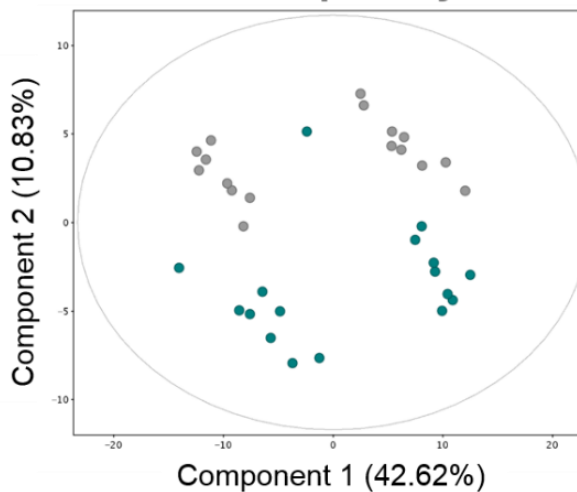
Positive polarity

Baseline

Nitisinone
1 week
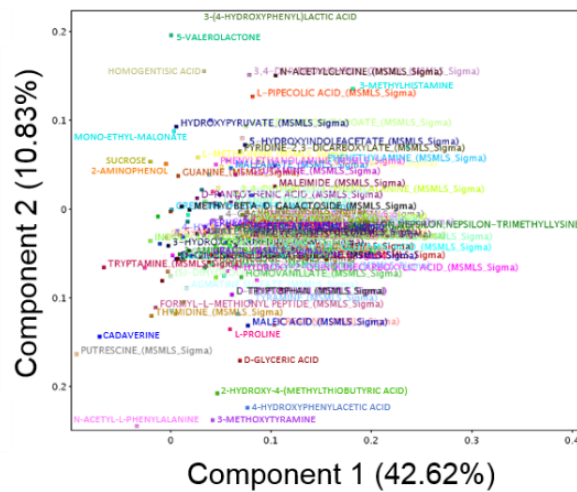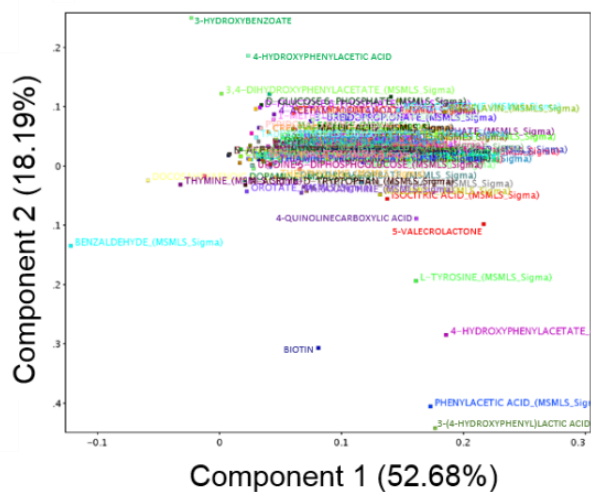
Figure 4.04. PCA to assess the effect of nitisinone treatment in profiling data from AKU mouse urine (targeted feature extraction). PCA plots (upper) in negative (left) and positive polarity (right), both showing clear separation in component 2 between profiles from urine samples taken at baseline compared to those taken at 1 week on nitisinone. Loadings plots (lower) show the respective contributions of individual metabolites to components 1 and 2.

Paired one-way repeated measures ANOVA was performed on each dataset to investigate differentially abundant entities between baseline and the time points on nitisinone. An FC >2 was chosen in order to limit false-discovery and to focus on clear changes. The metabolites that showed statistically significant changes (p <0.05, FC >2) in abundance after 3, 12 or 24 months on nitisinone in humans or 1 week on nitisinone in mice are summarised in Figures 4.05 and 4.06 for AMRT-based feature extraction (n=36) and AM-only feature extraction (n=3), respectively. In patients with AKU, 16 metabolites increased and 15 decreased. In $HGD^{-/-}$ mice, 15 metabolites increased and 7 decreased. Fourteen metabolites changed in the same direction in humans and mice. Importantly, the well-documented changes to urinary HGA, HPLA (both decreased) and tyrosine (increased) on nitisinone treatment were observed, which adds validity to the analytical technique and the data processing workflow. Data from targeted feature extraction also showed alteration (p <0.05, FC >2) to metabolites with a predicted association to AKU/nitisinone metabolism that were not from AMRT databases (Figure 4.06). 2,5-dihydroxybenzaldehyde, probably associated with ochronotic pigment, was decreased on nitisinone. The tyrosine metabolite γ-glutamyl-L-tyrosine and the nitisinone metabolite hydroxy-nitisinone were increased. These changes were observed in human and mouse urine. Interestingly, D-(+)-galactosamine was one of the major contributors to the separation observed in PCA, but it did not show statistical significance and visual inspection of its profile across the samples (in MPP) did not show consistent differences between baseline and on-nitisinone time points.

Together, the majority of metabolite changes observed could be categorised into three main metabolic pathways: those of tyrosine, tryptophan and purine. Increases and decreases are observed on nitisinone in metabolites from each of these pathways, but with notable trends of a greater number of increases in tyrosine pathway metabolites and a greater number of decreases in purine pathway metabolites.

Tyrosine

Increases (red):
- **3-(4-Hydroxyphenyl)lactic acid** (84, p<0.0001)
- **3-(4-Hydroxyphenyl)pyruvic acid*** (16, p<0.0001)
- 3-Methoxytyramine* (3.3, p<0.0001)
- 4-Coumarate (37, p<0.0001)
- 4-Hydroxybenzaldehyde (58, p<0.0001)
- **4-Hydroxyphenylacetic acid** (2.1, p<0.0001)
- **Acetyl-L-tyrosine*** (9.3, p<0.0001)
- **L-Tyrosine*** (100+, p<0.0001)
- Mandelic acid (7, p<0.001)
- **Phenylacetic acid** (16, p<0.0001)
- **Tyramine** (9.1, p<0.0001)

Decreases (blue):
- *3,4-Dihydroxyphenyl glycol* (2.4, p<0.0001)
- **Homogentisic acid*** (5.3, p<0.0001)
- Homovanillic acid (4.2, p<0.0001)

Tryptophan

Increases (red):
- *4-Quinolinecarboxylic acid* (4.8, p<0.0001)
- Xanthurenic acid (2.4, p<0.0001)

Decreases (blue):
- Indoxyl sulfate (11, p<0.0001)
- L-Kynurenine (4.8, p<0.0001)
- L-Tryptophan (2.2, p<0.0001)

Purine

Increases (red):
- Caffeate (2.3, p<0.0001)

Decreases (blue):
- **3,5-Cyclic AMP** (4.3, p<0.0001)
- 3-Ureidopropionate (100+, p<0.0001)
- Adenine (4.6, p<0.0001)
- Allantoin (100+, p<0.0001)
- **Xanthosine** (4.6, p<0.0001)

Other

Increases (red):
- *5-Valeroclactone* (6.6, p<0.0001)
- *Alpha-ketoglutaric acid* (2.2, p=0.0014)
- *Creatine* (2.1, p=0.032)
- Ethylmalonic acid (5.2, p<0.0001)
- *Isocitric acid* (3, p=0.002)
- *Maleimide* (2.8, p=0.0041)

Decreases (blue):
- 2-Hydroxybutyric acid (1.7, p<0.0001)
- **2-Hydroxy-4-(methylthio)butyric acid** (100+, p<0.0001)
- L-Threonine (7.3, p=0.022)
- Methyl vanillate (21, p<0.0001)
- *N-Acetylglycine* (2.4, p<0.0001)

* Finding previously reported

FC's (brackets) are at 24 months nitisinone for patients with AKU unless change observed in mouse only, in which case FC is at 1 week nitisinone in mouse

Bold: Change observed in human and mouse data

Italics: Change observed in mouse data only

Figure 4.05. Summary of urinary metabolite changes in patients and/or mice with AKU following treatment with nitisinone, identified by AMRT-based targeted feature extraction. Red and blue indicate increases and decreases respectively. Fold changes (FCs) are indicated in brackets and were calculated from raw peak area. These compounds are from the Sigma IROA mass spectrometry metabolite library of standards, except 3-(4-hydroxyphenyl)pyruvic acid and acetyl-L-tyrosine, which were subsequently analysed as standards and appended to the AMRT compound database for targeted feature extraction.
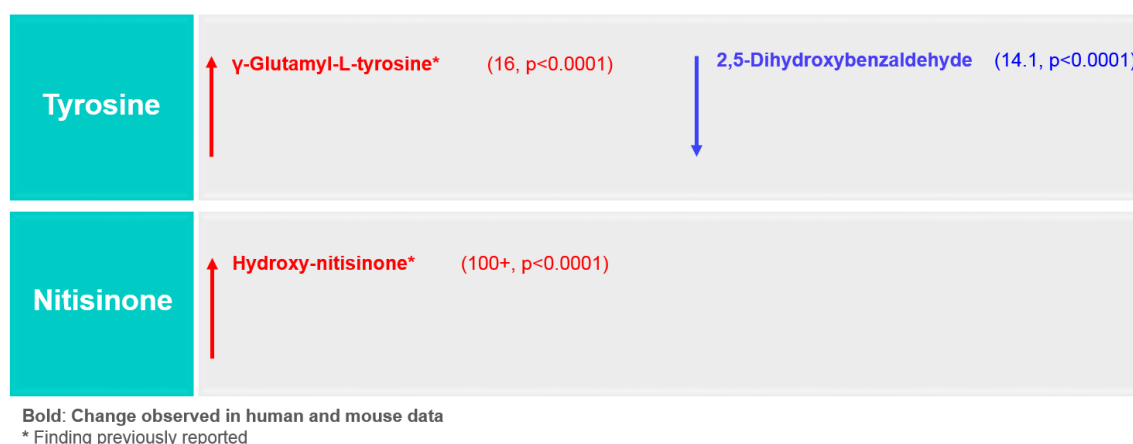
Figure 4.06. Summary of urinary metabolite changes in patients and mice with AKU following treatment with nitisinone, identified by targeted feature extraction (accurate mass only). Red and blue indicate increases and decreases respectively. Fold changes (FCs) are indicated in brackets and were calculated from raw peak area.

4.3.1.2. Non-targeted feature extraction

Non-targeted feature extraction was performed in parallel to targeted feature extraction, only in the raw data acquired using LC method 2. The aim of this approach was to investigate the potential for metabolites/entities to be altered that were not included in the AMRT databases generated, or not previously predicted to have an association with the changes to tyrosine metabolism as a consequence of nitisinone.

For the human data, in total 122 and 1594 entities were retained post-QC from non-targeted feature extraction in negative and positive polarity respectively. The decision was taken to relax the entity filtering parameters for method 2 in negative polarity for non-targeted feature extraction (see Table 4.01). This was because

no entities remained after filtering based on CV <25% across all replicate injections in each (5/5) pooled group QC sample. One-hundred and twenty-two entities remained in the negative polarity dataset after relaxing this filter to CV <25% in 4/5 of the pooled group samples; these entities were retained for subsequent statistical analyses. Figure 4.07 displays the profiles of the remaining entities in the negative polarity non-targeted feature extraction dataset across the pooled sample injection replicates following this relaxation of the CV-based filtering parameter. The 6th (of 7) replicate injection of the pooled baseline sample appears to be the single outlier, showing altered profiles for a subset of metabolites compared with the other replicates. This finding therefore provided confidence that the loss of technical reproducibility was limited to this individual injection in the analytical sequence run, as the same PCA-based sample outlier QC was subsequently performed on the individual samples as with the other datasets.
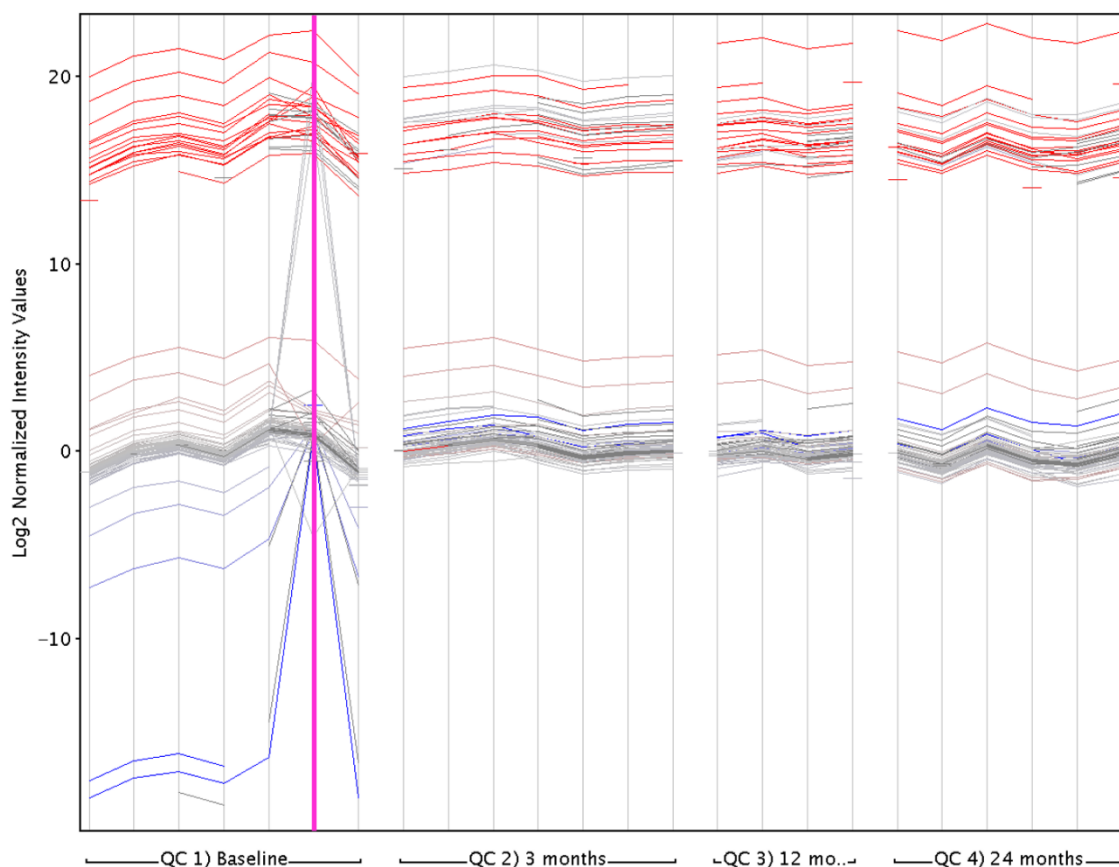
Figure 4.07. Profile plot of remaining chemical entities from non-targeted feature extraction on negative polarity human urine data following relaxation of filtering based on CV across replicate injections of pooled group samples. Filtering was relaxed to retain entities with CV <25% across all replicate injections in 4/5 pooled group QC samples (as no entities remained with a threshold of 5/5 as applied to other datasets). The 6th replicate injection of the baseline pooled sample (highlighted in pink) appears to be discrepant from the other replicates in this group. The clear reproducibility in profiles across the other replicates for each group shows that the compromised technical variation is limited to this one injection and that the relaxed filtering threshold is therefore appropriate.

Figure 4.08 shows the results from PCA on the human data from non-targeted feature extraction. In negative polarity, clear separation was observed between the baseline samples and those from the on-nitisinone time points in principal component 2, as with the targeted feature extraction data. Interestingly, a major separation of samples into two distinct groups is also evident in principal component 1. This separation was observed across all sampling time points and is discussed further in section 4.3.2.3. The corresponding loadings plot (Figure 4.08, lower section) shows the unidentified chemical entities contributing to the separation of samples into baseline vs on-nitisinone time points; those with extreme loadings on principal component 2. In contrast, PCA on the data from positive polarity did not show clear separation in principal components 1-4 between samples taken at baseline and on nitisinone.

PCA on the non-targeted mouse data showed clear separation between samples at baseline and those after 1 week on nitisinone in positive and negative polarity. This separation was in a combination of principal components 1 and 2 (Figure 4.09).
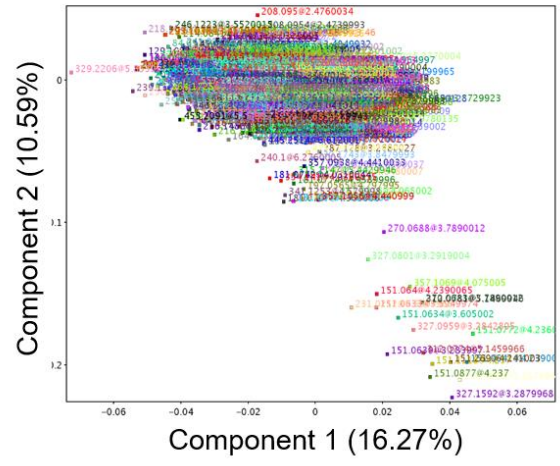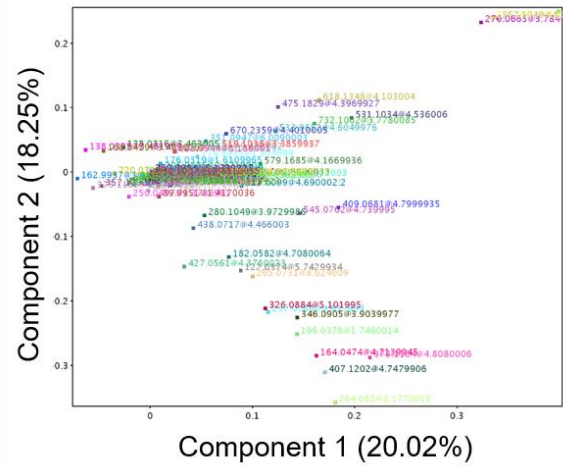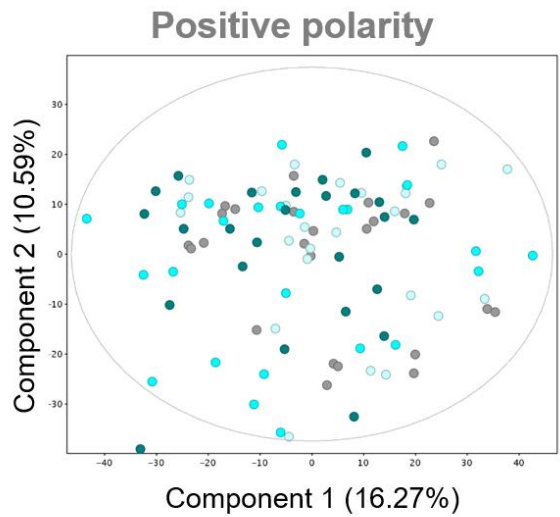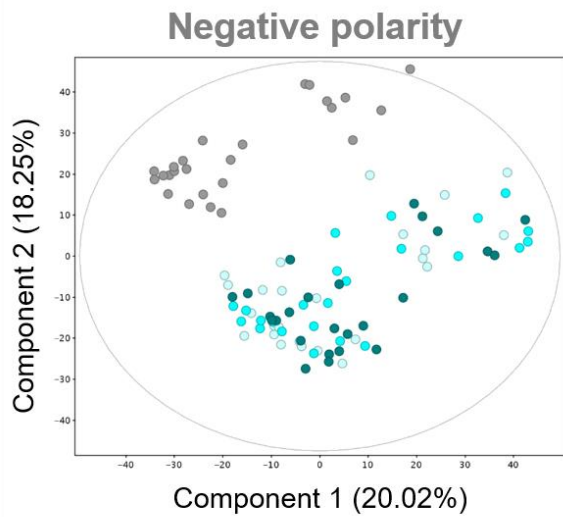
Figure 4.08. PCA to assess the effect of nitisinone treatment in profiling data from AKU human urine (non-targeted feature extraction). PCA plots (upper) in negative (left) and positive polarity (right). In negative polarity, clear separation was observed between profiles from urine samples taken at baseline compared to those taken at 3, 12 and 24 months on nitisinone, as with targeted feature extraction. In positive polarity, separation in components 1 and 2 between samples at baseline and the on-nitisinone time points was less clear. Loadings plots (lower) show the respective contributions of individual unidentified metabolites to components 1 and 2
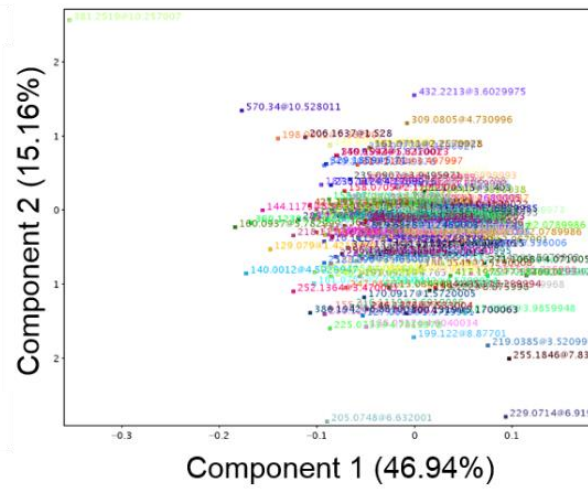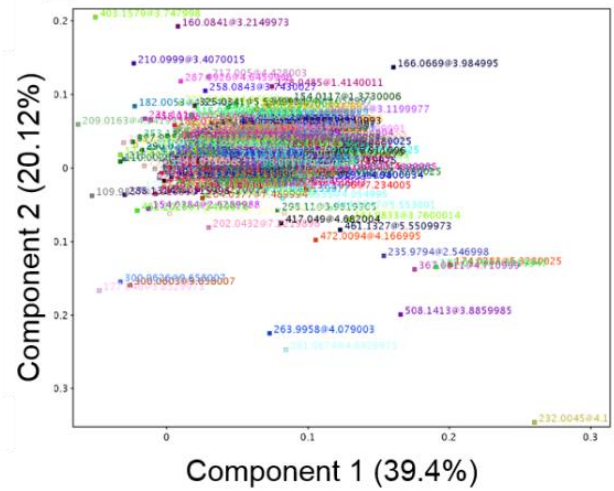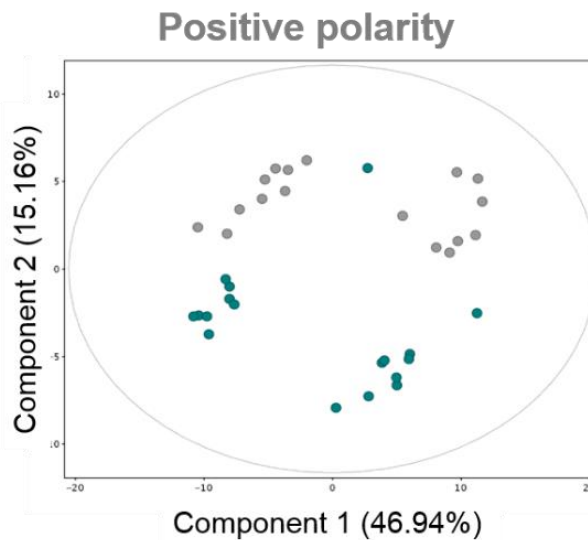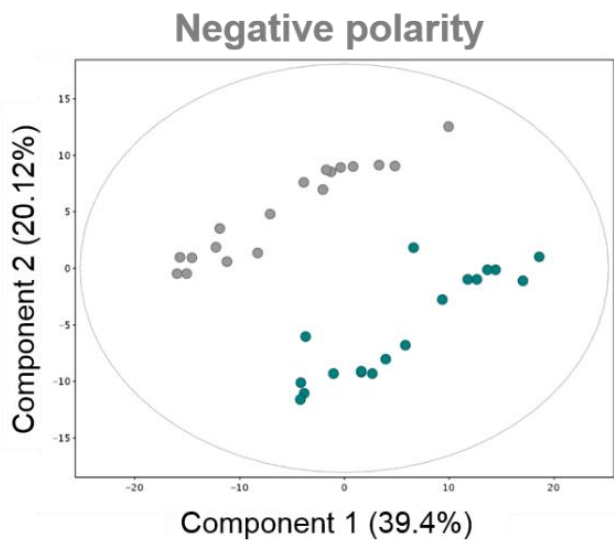
Figure 4.09. PCA to assess the effect of nitisinone treatment in profiling data from AKU mouse urine (non-targeted feature extraction). PCA plots (upper) in negative (left) and positive polarity (right), both showing clear separation between profiles from urine samples taken at baseline compared to those taken at 1 week on nitisinone. Separation was in component 1 and 2 for positive and negative polarities respectively. Loadings plots (lower) show the respective contributions of individual unidentified metabolites to components 1 and 2.

Statistical significance testing was performed on data from non-targeted feature extraction as described above for targeted extraction. In the human data, attention was focused on metabolites that showed altered abundance at all three time points on nitisinone compared with baseline (p <0.05, FC >2). On nitisinone, 42 entities showed altered abundance (29 increased, 13 decreased) in patients, and 65 entities (48 increased, 17 decreased) in mice, both including positive and negative polarity. The altered entities spanned a relatively broad range in RT and were biased towards the lower mass range (upper limit of acquisition *m/z* range was 1700): in the patient data mass range was 115-433 Da, RT range was 1.7-7.1 min; in the mouse data mass range was 56-514 Da, RT range was 1.4-10.4 min.

MassHunter ID Browser (Agilent, build 07.00) was employed for compound identification of these unknown entities of interest. The lists of entities were exported from MPP and into ID Browser. First, unknown entities that matched the AMRT databases generated in-house were identified (AM ±10 ppm, RT ±0.3 min), as these changes had already been noted previously (section 4.3.1.1.). Metabolites identified by this approach and therefore not of further interest given the aims of this complementary non-targeted analysis were: 4-hydroxybenzaldehye, HGA, a known MS/MS fragment of HGA (same RT as HGA, mass = 138.0325, putative formula $C_7H_6O_3$; see Chapter 6, section 6.3.2) and HPLA for patients, and phenylacetic acid for mice. Changes in these AMRT-matched entities on nitisinone were in the same direction as noted above.

For the remaining unknown entities that did not match metabolites in the AMRT databases generated in-house, candidate compound identifications were generated by AM match (±5 ppm) against the MassHunter METLIN metabolites PCD/PCDL accurate mass database (Agilent, build 07.00) and also using ID Browser (Agilent, build 07.00). Tables 4.03 and 4.04 summarise the altered chemical entities from non-targeted extraction that yielded putative compound identifications in human and mouse data respectively. These putative identifications (putative as by AM only (107)) were either the only compound match obtained or were considered the most likely of a list of potential candidates based on predicted association with the metabolite changes observed from targeted feature extraction. The number of other candidate mass matches obtained for these entities is indicated in Tables 4.03 and 4.04. The most notable changes for these putatively identified metabolites were on nitisinone increases in phenylpyruvic acid, dihydroxyphenylpyruvic acid, salicylic acid β-D-glucoside (all FC >100), 3-hydroxymandelic acid (or 3,4-dihydroxymandelaldehyde [DHMA]; both catecholamine metabolites, FC = 16) for patients, and also elevated 4-hydroy-L-glutamic acid (FC = 16) and decreased kynuramine (FC = 6.2) for mice.

Tables S4.01 and S4.02 (Supporting materials, Appendix 3) summarise the altered chemical entities on nitisinone, listed by observed mass and RT, for which no candidate compound identifications were obtained using the approach described.

Table 4.03. Summary of putatively identified chemical entities with differential abundance in urine taken from patients with AKU at baseline vs on-nitisinone, from non-targeted feature extraction. Peak areas for these entities showed statistically significant differences (Benjamini-Hochberg false-discovery rate p <0.05, fold change >2) between baseline and each sampling time point (3, 12 and 24 months) on nitisinone. Putative compound identifications were assigned by reference to the MassHunter METLIN metabolites PCD/PCDL accurate mass database (build 07.00, Agilent), with a match criteria of accurate mass (±5 ppm, unless indicated otherwise by asterisks) only. Where multiple matches were obtained for the same entity, the compound ID shown here was considered the most likely based on a) filtering for endogenous metabolites and b) compounds more likely to be associated with wider tyrosine metabolism or with the metabolites shown to be altered from targeted AMRT-based feature extraction. Fold changes are for baseline vs 24 months on nitisinone time points.

* compound putatively identified by accurate mass match ± 10 ppm (and 'number of compound matches' also with a mass window of ±10 ppm)

Putative compound ID	Formula	Neutral mass	RT (min)	Number of compound matches	Polarity	Direction of alteration on nitisinone	Fold change	p-value
Dihydroxyphenylpyruvic acid	$C_9H_8O_5$	196.038	1.7	2	(-)	↑	100+	<0.0001
Salicylic acid β-D-glucoside*	$C_{13}H_{16}O_8$	346.091	3.9	4	(-)	↑	100+	<0.001
Phenylpyruvic acid	$C_9H_8O_3$	164.047	4.7	10	(-)	↑	100+	<0.0001
3-Hydroxymandelic acid / 3,4-Dihydroxymandelaldehyde	$C_8H_8O_4$	168.042	2.5	19	(-)	↑	16	<0.0001

Homoveratric acid	$C_{10}H_{12}O_4$	196.074	6.1	20	(+)	↑	6.9	<0.0001
Hydroxydopamine*	$C_8H_{11}NO_3$	169.075	5.3	5	(+)	↑	6.8	<0.0001
N-Acetylvanilalanine	$C_{12}H_{15}NO_5$	253.096	5.3	2	(+)	↑	6.1	<0.0001
-	$C_{10}H_9NO_3$	191.058	4.88	4	(+)	↑	5.5	<0.0001
3-Methyleneoxindole	C_9H_7NO	145.053	7	7	(+)	↑	5.1	<0.0001
Aspartyl-phenylalanine	$C_{13}H_{16}N_2O_5$	280.107	3.9	4	(+)	↑	5.1	<0.0001
-	$C_{10}H_{18}N_2O_5$	246.123	4.18	10	(+)	↑	3.4	<0.001
Vanillactic acid*	$C_{10}H_{12}O_5$	212.07	5.2	7	(+)	↑	3.2	<0.001
-	$C_{11}H_8O_3$	188.047	6.6	6	(+)	↑	2.7	<0.001
-	$C_{10}H_9NO_3$	191.059	3.4	4	(+)	↓	5.2	<0.0001
L-Cis-cyclo(aspartylphenylalanyl)	$C_{13}H_{14}N_2O_4$	262.095	6	1	(+)	↓	5	<0.01
4-Hydroxyindole*	C_8H_7NO	133.053	4.1	6	(+)	↓	3.1	<0.001
Indoleacrylic acid*	$C_{11}H_9NO_2$	187.064	4.5	2	(+)	↓	2.4	<0.0001

Table 4.04. Summary of putatively identified chemical entities with differential abundance in urine taken from AKU mice at baseline vs 1 week on nitisinone, from non-targeted feature extraction. Putative compound identifications were assigned with reference to the MassHunter METLIN metabolites PCD/PCDL accurate mass database (build 07.00, Agilent), with a match criteria of accurate mass (±5 ppm, unless indicated otherwise by asterisks) only. Where multiple matches were obtained for the same entity, the compound ID shown here was considered the most likely based on a) filtering for endogenous metabolites and b) compounds more likely to be associated with wider tyrosine metabolism or with the metabolites shown to be altered from targeted AMRT-based feature extraction. P-values are adjusted by Benjamini-Hochberg false-discovery rate.

* compound putatively identified by accurate mass match ±10 ppm (and 'number of compound matches' also with a mass window of ±10 ppm)

Putative compound ID	Formula	Neutral mass	RT (min)	Number of compound matches	Polarity	Direction of alteration on nitisinone	Fold change	p-value
4-Hydroxy-L-glutamic acid	$C_5H_9NO_5$	163.0476	3.5	6	(-)	↑	16	<0.0001
5-Methoxyindoleacetate*	$C_{11}H_{11}NO_3$	205.0748	6.6	5	(+)	↑	7.9	<0.0001
5-Methoxytryptophol	$C_{11}H_{13}NO_2$	237.1096	7.2	4	(-)	↑	2.7	<0.0001
2-Hydroxyphenylacetic acid O-b-D-glucoside	$C_{14}H_{18}O_8$	157.0535	4	6	(-)	↑	2.5	<0.0001
Vanillin 4-sulfate	$C_9H_{10}O_6S$	246.0191	5	1	(-)	↑	2.4	<0.0001

L-glycyl-L-hydroxyproline / *N*-alpha-acetyl-L-glutamine*	$C_7H_{12}N_2O_4$	188.081	2	2	(+)	↑	2.4	<0.001
Valyl-hydroxyproline / Hydroxyprolyl-valine	$C_{10}H_{18}N_2O_4$	230.1279	4.6	2	(+)	↑	2.4	<0.01
γ-glutamyl-γ-aminobutyraldehyde / prolyl-threonine*	$C_9H_{16}N_2O_4$	216.1119	3.6	4	(+)	↑	2.2	<0.001
Vanilloglycine	$C_{10}H_{11}NO_5$	225.0887	4.7	3	(-)	↑	2.1	<0.001
Kynuramine	$C_9H_{12}N_2O$	210.0999	3.4	5	(-)	↓	6.2	<0.0001
N-gamma-acetyldiaminobutyrate	$C_6H_{12}N_2O_3$	160.0841	3.2	6	(-)	↓	4	<0.0001
Isovalerylglutamic acid	$C_{10}H_{17}NO_5$	231.111	3.1	5	(-)	↓	3	<0.001

4.3.2. Associations with clinical data and patient demographics

The clinical and patient demographic data parameters available were: gender, age, ^{18}F-NaF PET imaging, back pain, BDI-II score and *HGD* gene sequencing. These parameters were investigated in the metabolic profiling data acquired using method 2 only. Data from these parameters were imported into each metabolomics dataset in MPP (build 14.5); separate projects for targeted and non-targeted feature extraction and for positive and negative polarity. Separate 'interpretations', or grouping of sample data files, were created in MPP for each individual parameter to assess its relationship with the metabolomic data in each MPP project. As with studying the effect of nitisinone treatment, potential associations between metabolic profiles and clinical data / patient demographics were also studied in the data from non-targeted feature extraction in order to widen the scope of the analyses to compounds that were not included in the AMRT databases that had been generated (see Chapter 2). First, each clinical/demographic parameter was investigated at baseline only, by PCA, in order to investigate its effect in untreated AKU. For each parameter, if PCA revealed separation of metabolome profiles at baseline, the same parameter was additionally investigated in the data from the final 24 months on nitisinone time point in order to check whether the same separation remained after therapy for HGA reduction.

4.3.2.1. Gender

Data from patients with AKU were from 12 males and 13 females, and the mouse data were from 9 males and 9 females.

Targeted feature extraction

PCA on the data from targeted feature extraction showed no clear separation between male and female patients with AKU (Figure 4.10). Unpaired t-tests identified no entities with differentially abundant profiles between male and female patients at baseline (at the level of $p < 0.05$) in positive or negative polarity.
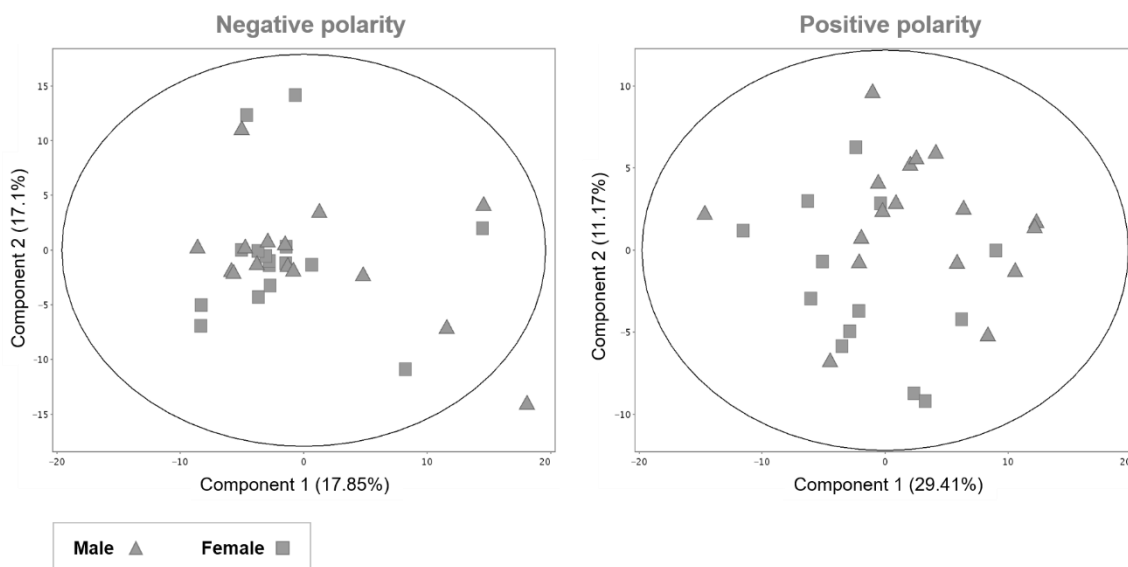
Figure 4.10. PCA to assess the effect of gender in profiling data from AKU human urine taken at baseline (targeted feature extraction). PCA plots in negative (left) and positive polarity (right) show limited separation of profiles by gender in components 1 and 2.

The apparent lack of a significant effect of gender in the human data contrasted with the mouse data. In the mouse data, striking separation was observed between males and females both at baseline and after 1 week on nitisinone; separation in principal component 1 in positive and negative polarity (Figure 4.11). The separation of mouse profiles into baseline *versus* on-nitisinone time points (in principal component 2; as described in section 4.3.1.) was secondary to the separation observed by gender. In positive polarity, histamine was the primary driver of this separation, with the greatest loading on principal component 1 (0.42). Urinary histamine was markedly elevated in female mice ($p < 0.0001$, FC=16). In the negative polarity data, the PCA separation by gender was driven by a larger group of compounds: 4-quinolinecarboxylic acid ($p < 0.0001$, FC=3.15), 4-hydroxyphenylacetate ($p < 0.0001$, FC=4.27), riboflavin ($p < 0.0001$, FC=18.49), 5-valerolactone ($p < 0.0001$, FC=16), erythritol ($p < 0.0001$, FC=8.78), anthranilate ($p < 0.0001$, FC=9.33) (all elevated in females), benzaldehyde ($p < 0.0001$, FC=22.89), and thymine ($p < 0.001$, FC=2.35) (both elevated in males) had the greatest loadings on principal component 1 and showed the most clear differences between male & female mice. The p-values and FCs indicated above are for male vs female mice at baseline. None of these metabolite changes were observed between males and females in the human data ($p > 0.05$ for all compounds).

Figure 4.11. PCA to assess the effect of gender in profiling data from AKU mouse urine, taken at baseline then after 1 week on nitisinone (targeted feature extraction). Unlike the human data, PCA plots from mouse data show clear separation between male and female mice in negative (left plot) and positive (right plot) polarity. Gender appears to contribute to the greatest overall difference in the dataset, with separation between males and females in component 1. Separation by samples taken at baseline vs on nitisinone was in component 2 and therefore secondary to the primary effect of gender.

Non-targeted feature extraction

Non-targeted feature extraction on the data from patients also showed no clear separation between males and females at baseline in principal components 1-4 (Figure 4.12). This agreed with the human data from targeted feature extraction.
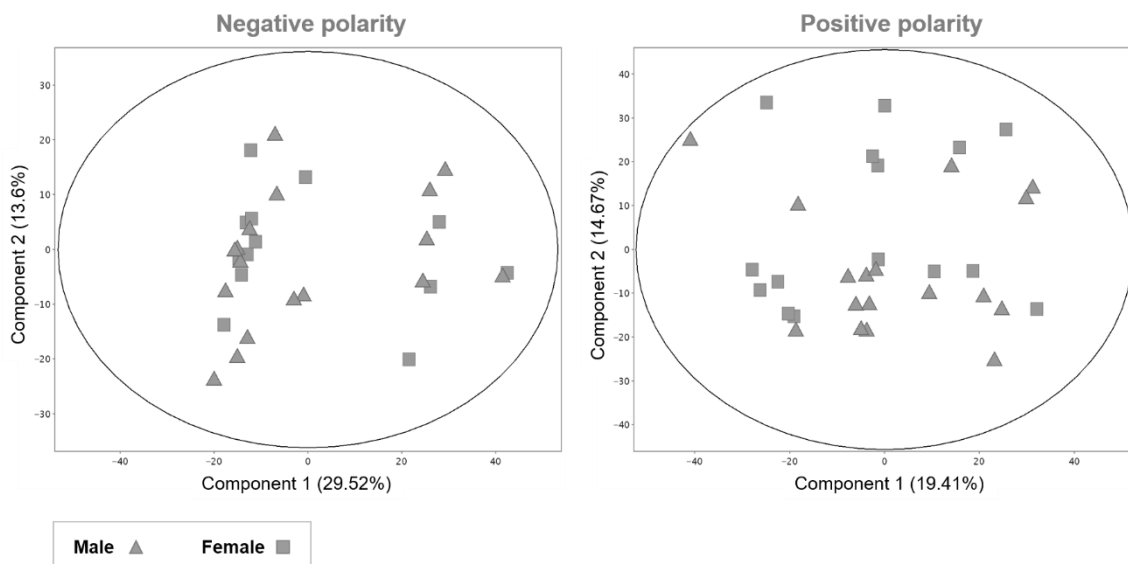
Figure 4.12. PCA to assess the effect of gender in profiling data from AKU human urine taken at baseline (non-targeted feature extraction). PCA plots in negative (left) and positive polarity (right) show limited separation of profiles by gender in components 1 and 2, as with the data from targeted feature extraction.

4.3.2.2.   Age

To study the effect of age on metabolic profiles, patients were separated into the following age groups (age in years): 11-20 (n=2), 21-30 (n=5), 31-40 (n=4), 41-50 (n=6), 51-60 (n=8), 61-70 (n=7), 71-80 (n=1). The effect of age was not investigated in the mouse data; mice of the same approximate age were chosen specifically as part of a tightly controlled experiment with the primary aim to investigate the consequences of nitisinone treatment.

Targeted feature extraction

PCA on the data from targeted feature extraction showed limited separation between the different age groups at baseline (Figure 4.13). The data from negative polarity (left plot, Figure 4.13) showed some evidence of separation by age in component 1, with a tentative trend for positive scores for the profiles of younger patients on the x-axis (representing principal component 1). To identify potential differentially abundant metabolites between younger and older patients, the baseline time point data were separated into two larger age groups; <30 and >40 years. Unpaired t-tests were performed to compare the profiles between these groups but revealed no statistically significant (p <0.05) entities based on peak area in positive or negative polarity.
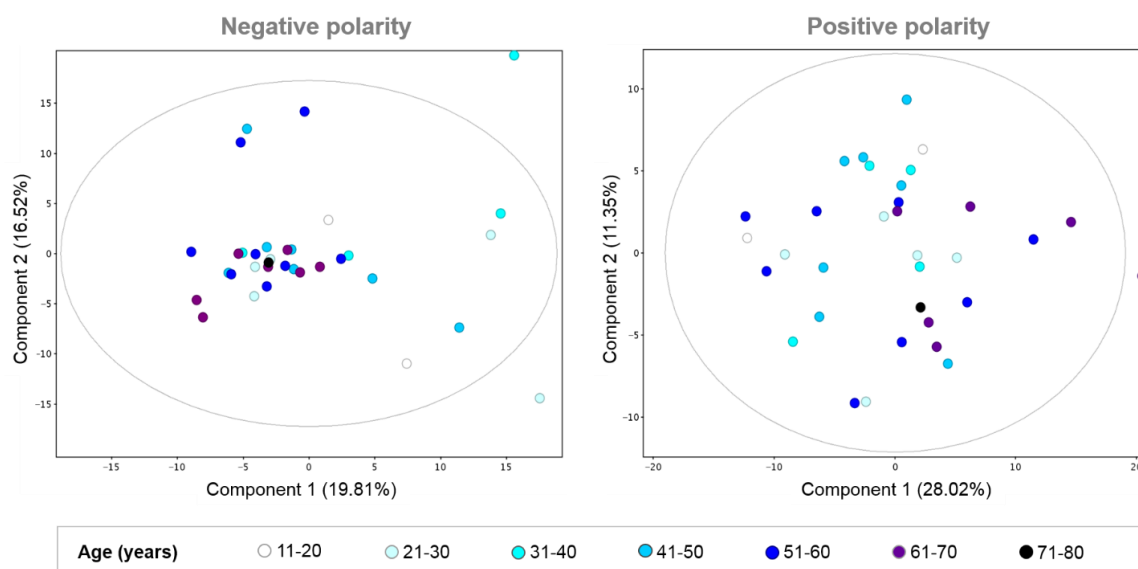
Figure 4.13. PCA to assess the effect of age in profiling data from AKU human urine taken at baseline (targeted feature extraction). PCA plots in negative (left) and positive polarity (right) showing limited separation of profiles by age in components 1 and 2.

Non-targeted feature extraction

PCA on the data from non-targeted feature extraction also did not reveal clear separation at baseline between the profiles of patients from the different age groups (Figure 4.14). Similarly, unpaired t-tests revealed no statistically significant ($p < 0.05$) entities between patients aged <30 years and those >40 years.
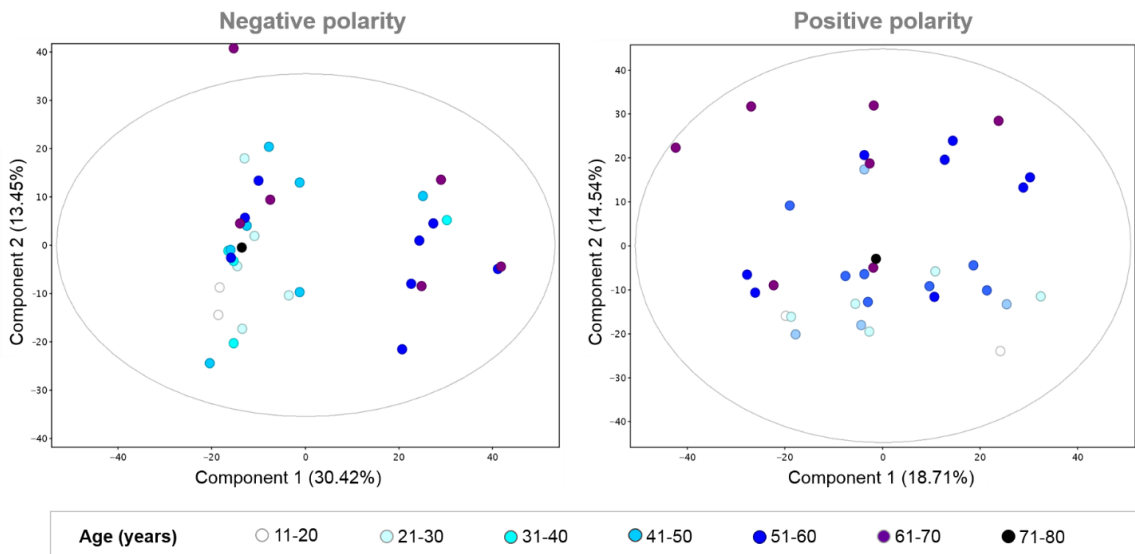
Figure 4.14. PCA to assess the effect of age in profiling data from AKU human urine taken at baseline (non-targeted feature extraction). PCA plots in negative (left) and positive polarity (right) show no clear separation of profiles by age in components 1 or 2, as with the data from targeted feature extraction.

4.3.2.3. ^{18}F-NaF PET imaging data

The data from ^{18}F-NaF PET imaging comprised two measures; ^{18}F-NaF uptake to lumbar/thoracic intervertebral disks and ^{18}F-NaF uptake to lumbar/thoracic vertebrae. For each of these measures, patients were divided into those with

'high' (values >10) *versus* 'low' (values <10) uptake. Increased ^{18}F-NaF uptake to the vertebrae and intervertebral disks is associated with OA pathology, representing accelerated bone turnover and cartilage degeneration (*i.e.* soft tissue mineralisation), respectively (see section 4.2.6). For cartilaginous tissue, although there is no agreed exact ^{18}F-NaF uptake threshold that defines benign or malignant lesions (227,232), one study established that values of 6.4 and above predicted hip OA incidence and progression (228). The cutoff applied here was increased to uptake value of 10 as this threshold explained the clear separation that was observed between patients in PCA (Figures 4.15 and 4.16). The patients in this cohort assigned to the '>10' group for ^{18}F-NaF uptake to intervertebral disks were therefore considered likely to show the most severe spinal OA.

For patients for whom PET data were available, at baseline 6/15 and 4/15 were in the high uptake group for intervertebral disks and vertebrae respectively. At 24 months on nitisinone, 9/17 and 4/17 patients were in the high uptake group for intervertebral disks and vertebrae respectively. Unpaired t-tests were performed to compare the abundance of entities from high *versus* low uptake patients for both measures at baseline and at 24 months on nitisinone.

Targeted feature extraction

PCA on data from targeted feature extraction showed evidence of separation between the metabolomic profiles of patients with high *versus* low ^{18}F-NaF uptake to intervertebral disks. PCA separation between these patients was clearer in

negative polarity; partial separation observed in principal component 1 (Figure 4.15). In positive polarity, separation was observed between these groups of patients in principal component 2, although limited. Across the targeted feature extraction data, 2-deoxyguanosine / adenosine (entity matched both compounds due to identical mass and RT in the AMRT database generated) was the only compound that showed a statistically significant difference in abundance (p=0.013, FC=3.5) between patients with high *versus* low intervertebral disk ^{18}F-NaF uptake (decreased for patients with high ^{18}F-NaF uptake; values >10) (Table 4.05).
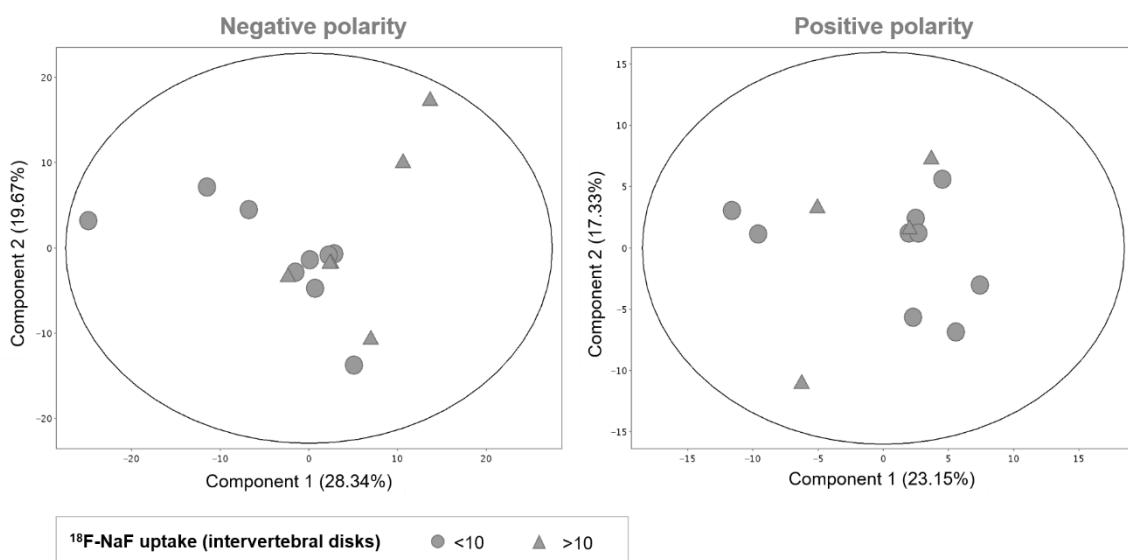
Figure 4.15. PCA to assess the effect of ^{18}F-NaF uptake scores to intervertebral disks in profiling data from AKU human urine taken at baseline (targeted feature extraction). PCA plots in negative (left) and positive polarity (right) show limited separation of profiles from patients with high (score >10) vs low (score <10) ^{18}F-NaF uptake scores.

A degree of PCA separation was also observed between the profiles of patients with high *versus* low ^{18}F-NaF uptake to the vertebrae (Figure 4.16). As with the grouping by uptake to intervertebral disks, separation between the vertebral ^{18}F-NaF uptake groups was clearer in negative polarity (separation in principal component 1). Compounds with differential abundance between patients with high *versus* low vertebral ^{18}F-NaF uptake were L-proline, malonate and mono-methyl glutarate (both decreased in the high uptake group; p <0.0001, FC=16) from targeted feature extraction (Table 4.06). Interestingly, the decrease in L-proline was observed both at baseline and 24 months on nitisinone. The decreases in malonate and mono-methyl glutarate were observed only at baseline and 24 months on nitisinone respectively.
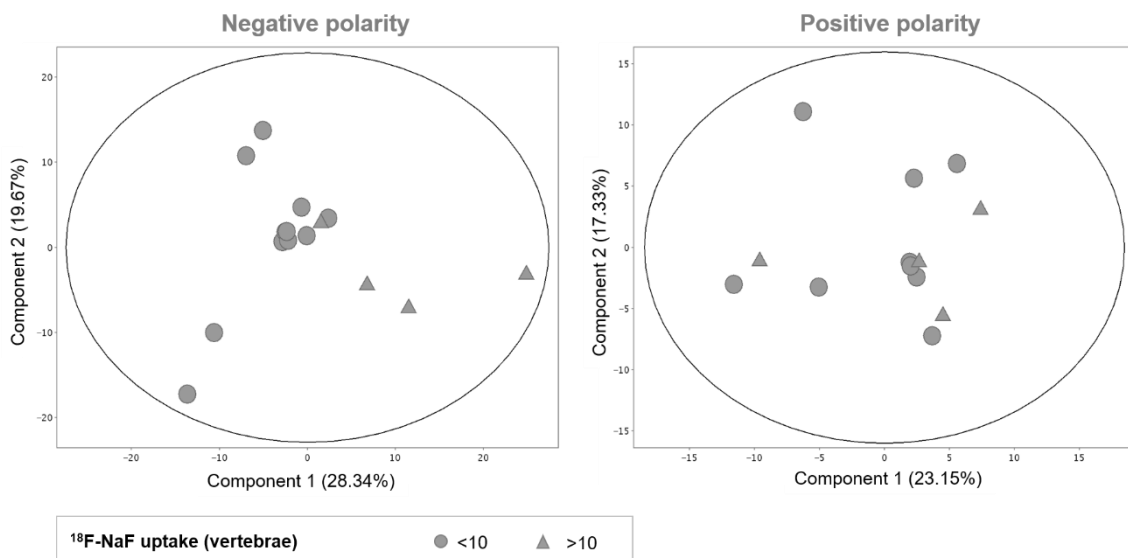
Figure 4.16. PCA to assess the effect of ^{18}F-NaF uptake scores to lumbar/thoracic vertebrae in profiling data from AKU human urine taken at baseline (targeted feature extraction). PCA plots show some separation between profiles from patients with high (score >10) vs low (score <10) ^{18}F-NaF uptake scores in negative polarity, but less clear separation between profiles in positive polarity.

Non-targeted feature extraction

PCA on data from non-targeted feature extraction showed clearer separation between the metabolomic profiles of patients with high *versus* low ^{18}F-NaF uptake to the intervertebral disks in positive and negative polarity (Figure 4.17). In these data, ^{18}F-NaF uptake appears to be the primary driver of the total variation observed in the dataset; separation in principal component 1. Given this degree of separation in the baseline data, the effect of ^{18}F-NaF uptake was additionally visualised in the PCA on data from all time points. As shown in Figure 4.18, addition of all time points further reveals the marked separation of the profiles from patients with high *versus* low ^{18}F-NaF uptake to the intervertebral disks.

Intervertebral ^{18}F-NaF uptake appears to explain the bimodal distribution of profiles across all time points, which is orthogonal to the separation observed between baseline and nitisinone time points. Fourteen entities from non-targeted feature extraction showed statistically significant abundance differences ($p < 0.05$, FC >2) between intervertebral ^{18}F-NaF uptake groups at baseline and/or 24 months on nitisinone (in high uptake group: 8 increased, 6 decreased). Putative AM identifications were obtained for 4/14 of these entities (Table 4.05). The remaining 10 differential entities are reported as unknowns in Table S4.03 (Supporting Materials, Appendix 4). Of the putatively identified entities, the most notable was a metabolite of the painkiller drug acetaminophen (paracetamol), which was increased in the high ^{18}F-NaF uptake group both at baseline and 24 months on nitisinone in positive and negative polarity ($p < 0.0001$, FC=40). Other metabolites showing marked differences were cyclodopa glucoside (increased in the high uptake group at baseline; $p < 0.0001$, FC >100), and caffeic acid (3- or 4) sulfate (decreased in the high uptake group at baseline; $p < 0.0001$, FC=16).
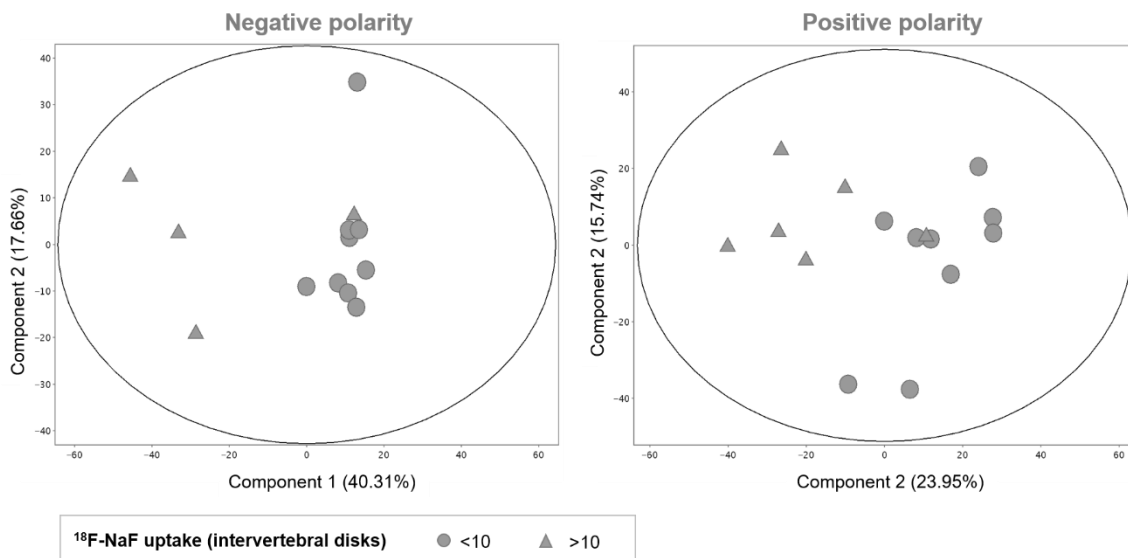
Figure 4.17. PCA to assess the effect of ^{18}**F-NaF uptake scores to intervertebral disks in profiling data from AKU human urine taken at baseline (non-targeted feature extraction).** PCA plots (upper) in negative (left) and positive polarity (right) show separation in component 1 between the urine profiles of patients with high (score >10) vs low (score <10) ^{18}F-NaF uptake scores. Loadings plots (lower) show that a cluster of unidentified entities are contributing to this separation in component 1 in both polarities (extreme negative loading values on component 1, x-axis).
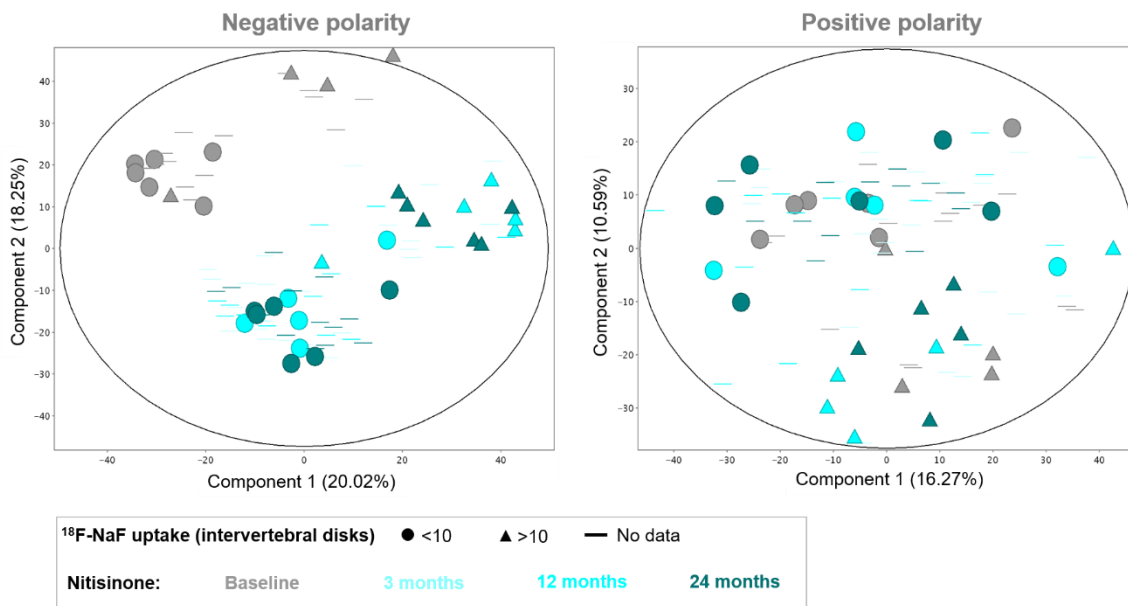
Figure 4.18. PCA to assess the effect of ^{18}F-NaF uptake scores to intervertebral disks in profiling data from AKU human urine (non-targeted feature extraction), with all time points on nitisinone added. Plots show the same separation between the urine profiles of patients with high (score >10) vs low (score <10) ^{18}F-NaF uptake scores both at baseline and on nitisinone treatment. In negative polarity, ^{18}F-NaF uptake appears to account for the major bimodal distribution in the dataset, captured by principal component 1. In the non-targeted data in positive polarity, although no clear separation was observed between profiles at baseline and on nitisinone, profiles are clearly separated into high and low ^{18}F-NaF uptake groups in component 2.

PCA on non-targeted data showed some evidence of separation between patients with high *versus* low ^{18}F-NaF uptake to the vertebrae in positive polarity (partial separation in principal component 1; Figure 4.19). Across positive and negative polarity, 9 entities from non-targeted feature extraction showed statistically significant differences in abundance (p <0.05, FC >2) between patients with high *versus* low ^{18}F-NaF uptake at baseline and/or 24 months on

nitisinone (in high uptake group: 3 entities increased, 6 entities decreased). Putative AM-based compound identifications were obtained for 7 of these entities (see Table 4.06); the other 2 entities are reported as unknowns (see Table S4.04, Appendix 4). The most notable of these putatively identified compounds were iduronic acid (increased in high uptake group), adrenochrome, L-tyrosine methyl ester 4-sulfate, diferulic acid, buchananine and clitidine (decreased in high uptake group); all with p <0.0001 and FC=16.
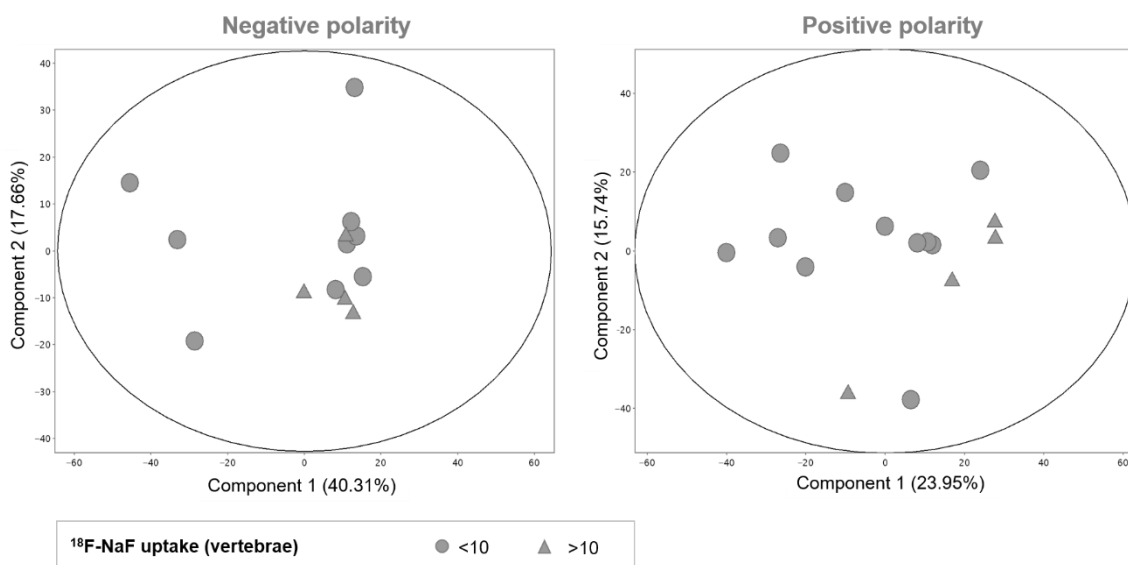
Figure 4.19. PCA to assess the effect of ^{18}F-NaF uptake scores to lumbar/thoracic vertebrae in profiling data from AKU human urine taken at baseline (non-targeted feature extraction). PCA plots indicate separation between profiles from patients with high (score >10) vs low (score <10) ^{18}F-NaF uptake scores in positive polarity, but less clear separation between profiles in negative polarity.

Table 4.05. Summary of putatively identified urinary chemical entities with differential abundance between patients with high (score >10) vs low (score <10) ^{18}F-NaF uptake to intervertebral disks. Compound matches were by AMRT against the metabolite database generated in-house, unless indicated otherwise. Putative compound identifications for entities from non-targeted feature extraction were assigned by reference to the MassHunter METLIN metabolites PCD/PCDL accurate mass database (build 07.00, Agilent), with a match criteria of accurate mass (±5 ppm) only. In the data from non-targeted feature extraction, where multiple matches were obtained for the same entity the compound ID shown here was considered the most likely based on filtering for endogenous metabolites and known medications taken by patients attending the UK National Alkaptonuria Centre. P-values are adjusted by Benjamini-Hochberg false-discovery rate.

* Entity putatively identified by accurate mass only (±5 ppm)

[†] Entity with statistically significant difference between patients with high vs low uptake score <u>at baseline</u>

[‡] Entity with statistically significant difference between patients with high vs low uptake score <u>at 24 months on nitisinone</u>

Putative compound ID	Formula	Neutral mass	RT (min)	Number of compound matches	Polarity	Direction of alteration in 'high' ^{18}F-NaF uptake group	Fold change	p-value
Cyclodopa glucoside *[†]	$C_{15}H_{19}NO_9$	618.1348	4.2	7	(-)	↑	>100	0.026
Acetaminophen *[†‡]	$C_8H_9NO_2$	151.0633	3.6	10	(+/-)	↑	40.1	<0.0001
Caffeic acid 3-sulfate / Caffeic acid 4-sulfate *[†]	$C_9H_8O_7$	259.9984	5.2	1	(-)	↓	16	<0.0001
2-Deoxyguanosine / Adenosine [‡]	$C_5H_9NO_3S$	163.0311	3.4	2	(+)	↓	3.5	0.013

Table 4.06. Summary of putatively identified urinary chemical entities with differential abundance between patients with high (score >10) vs low (score <10) ^{18}F-NaF uptake to lumbar/thoracic vertebrae. Compound matches were by AMRT against the metabolite database generated in-house, unless indicated otherwise. Putative compound identifications for entities from non-targeted feature extraction were assigned by reference to the MassHunter METLIN metabolites PCD/PCDL accurate mass database (build 07.00, Agilent), with a match criteria of accurate mass (±5 ppm) only. In the data from non-targeted feature extraction, where multiple matches were obtained for the same entity the compound ID shown here was considered the most likely based on filtering for endogenous metabolites and known medications taken by patients attending the UK National Alkaptonuria Centre. P-values are adjusted by Benjamini-Hochberg false-discovery rate.

* Entity putatively identified by accurate mass only (±5 ppm)

† Entity with statistically significant difference between patients with high vs low uptake score <u>at baseline</u>

‡ Entity with statistically significant difference between patients with high vs low uptake score <u>at 24 months on nitisinone</u>

Putative compound ID	Formula	Neutral mass	RT (min)	Number of compound matches	Polarity	Direction of alteration in 'high' ^{18}F-NaF uptake group	Fold change	p-value
Iduronic acid *‡	$C_6H_{10}O_7$	194.0419	1.3	20	(-)	↑	16	<0.0001
Octanoic acid *‡	$C_{14}H_{24}O_8$	320.1463	9.3	12	(-)	↑	2.3	<0.0001
L-proline †‡	$C_5H_9NO_2$	115.0632	1.4	1	(-)	↓	16	<0.0001

Malonate [†]	$C_3H_4O_4$	104.0109	1.6	1	(-)	↓	16	<0.0001
Mono-methyl glutarate [‡]	$C_9H_8O_7$	146.0575	5.3	1	(-)	↓	16	<0.0001
Adrenochrome *[‡]	$C_9H_9NO_3$	225.0624	4.8	10	(-)	↓	16	<0.0001
L-tyrosine methyl ester 4-sulfate *[‡]	$C_{10}H_{13}NO_6S$	275.0461	4.5	1	(-)	↓	16	<0.0001
Diferulic acid *[†]	$C_{20}H_{18}O_8$	386.0987	4.6	8	(+)	↓	16	<0.0001
Buchananine *[‡]	$C_{12}H_{15}NO_7$	285.0846	1.3	1	(+)	↓	16	<0.0001
Clitidine *[†]	$C_{11}H_{14}N_2O_6$	270.0828	3.7	1	(+)	↓	3.8	<0.0001

4.3.2.4.   Back pain

Back pain data were mean values for pain across lumbar and thoracic spinal regions, self-reported on a scale from 0 (no pain) to 10 (severe pain). Back pain data were only available at baseline. For the purposes of the following statistical analyses, patients were divided into two groups for 'high' (score >5; n=6) and 'low' (score <5; n=10) pain.
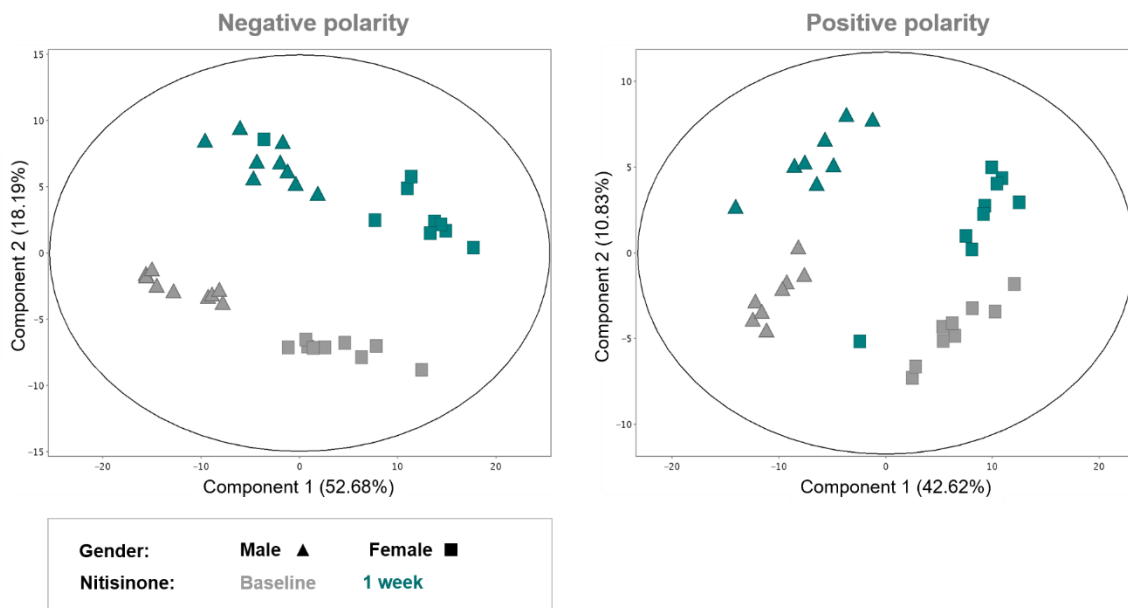
Targeted feature extraction

PCA on data from targeted feature extraction showed no clear separation between patients with high *versus* low back pain (Figure 4.20). Unpaired t-test did not find any entities from targeted AMRT-based extraction with differential abundance between these patient groups (at the level of p <0.05, FC >2).
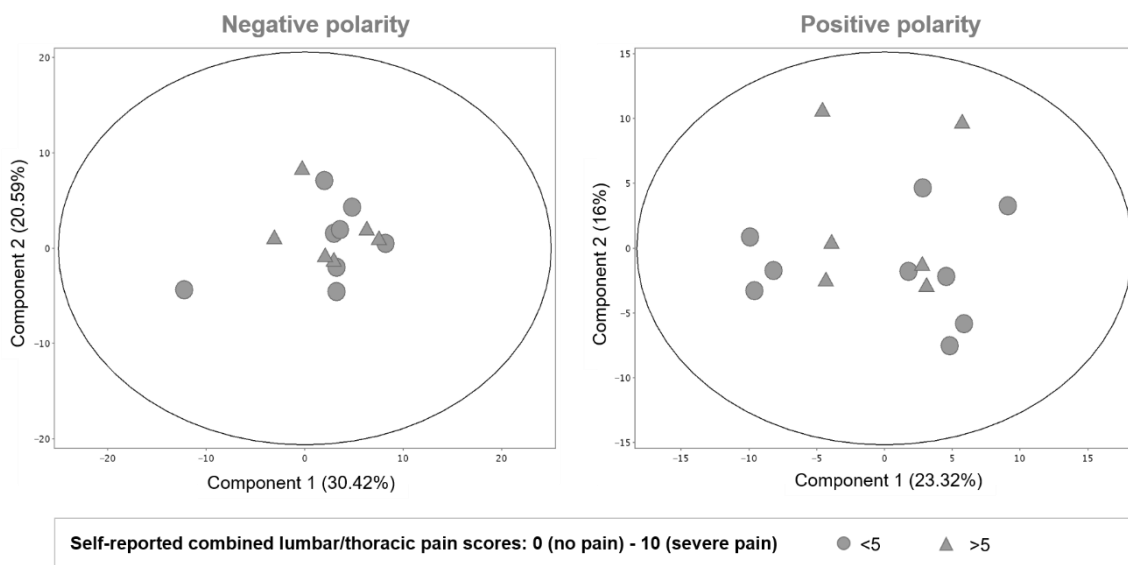
Figure 4.20. PCA to assess the effect of self-reported back pain scores in profiling data from AKU human urine taken at baseline (targeted feature extraction). PCA plots show no clear separation in components 1 and 2 between the profiles of patients with high (score >5) vs low (score <5) pain scores.

Non-targeted feature extraction

PCA on data from non-targeted feature extraction showed clearer separation of patients into high and low back pain groups in negative polarity (separation in principal component 2; Figure 4.21). Unpaired t-test revealed two entities with differential abundance between high and low pain groups; caffeic acid (3- or 4-)-sulfate and an entity putatively identified as a dipeptide (prolyl-threonine or threonyl-proline). These entities were both increased in the high pain group, with p <0.0001 and FC of 2.2 (Table 4.07).
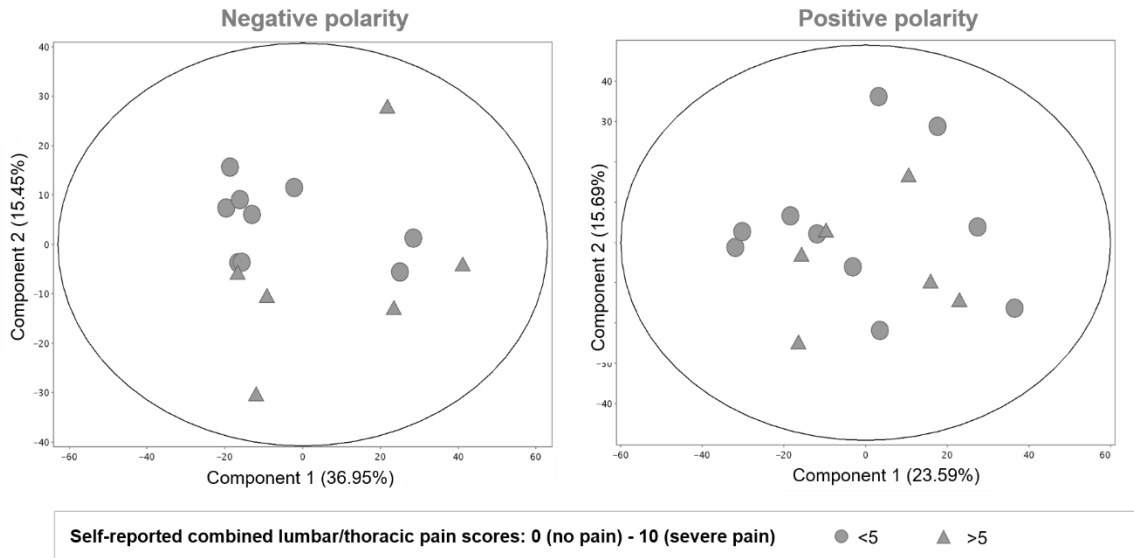
Figure 4.21. PCA to assess the effect of self-reported back pain scores in profiling data from AKU human urine taken at baseline (non-targeted feature extraction). There is some evidence of separation between the profiles of patients with high (score >5) vs low (score <5) pain scores in negative polarity (left plot), but no clear separation in positive polarity.

Table 4.07. Summary of unidentified urinary chemical entities with differential abundance between patients with high (>5/10) vs low (<5/10) self-reported lumbar/thoracic pain scores at baseline. Putative compound identifications were obtained for these compounds using the MassHunter METLIN metabolites PCD/PCDL accurate mass database (build 07.00, Agilent), with match criterion of accurate mass ±5 ppm. P-values are adjusted by Benjamini-Hochberg false-discovery rate.

Putative compound ID	Formula	Neutral mass	RT (min)	Number of compound matches	Polarity	Direction of alteration in 'high' pain group	Fold change	p-value
Caffeic acid 3-sulfate / Caffeic acid 4-sulfate	$C_9H_8O_7$	259.9984	5.2	2	(-)	↑	2.2	<0.0001
Prolyl-threonine / Threonyl-proline	$C_9H_{16}N_2O_4$	216.1109	3.1	8	(-)	↑	2.2	<0.0001

4.3.2.5.   BDI-II score

Patients were divided into the following categories based on level of depression indicated by BDI-II score: minimal (*n*=19), mild (*n*=1), moderate (*n*=3) and severe (*n*=2).

PCA on data from targeted (Figure 4.22) and non-targeted (Figure 4.23) feature extraction showed no clear separation between patient profiles by BDI-II score. Statistical analyses were not performed on the data from these groups due to the low number of patients in 'mild', 'moderate' and 'severe' depression groups.
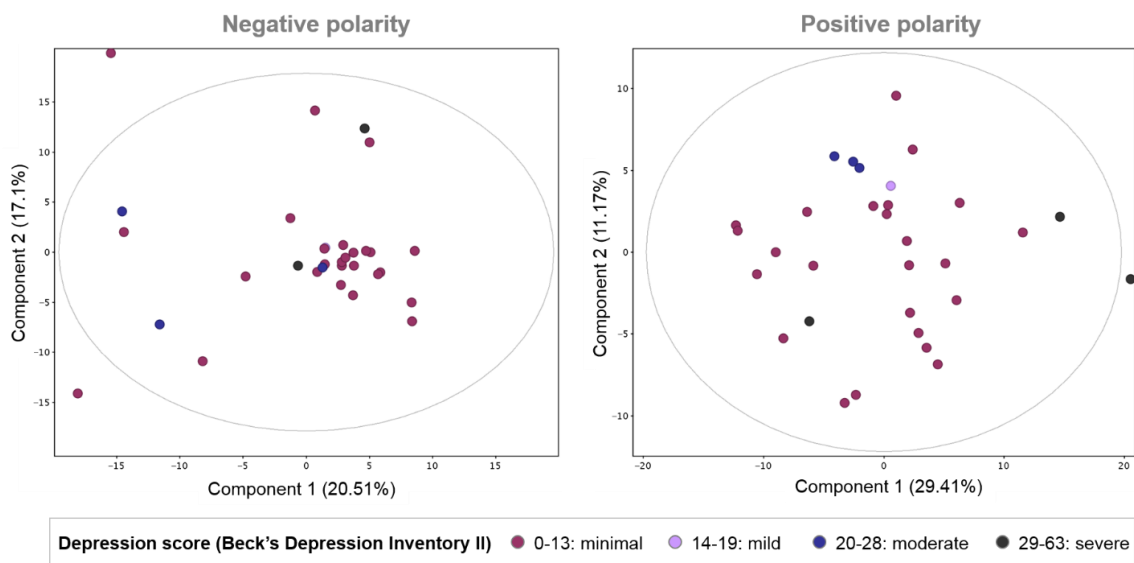
Figure 4.22. PCA to assess the effect of depression (Beck's Depression Inventory II scores) in profiling data from AKU human urine taken at baseline (targeted feature extraction). PCA plots show no clear separation in components 1 and 2 between the profiles of patients from the different depression score categories.
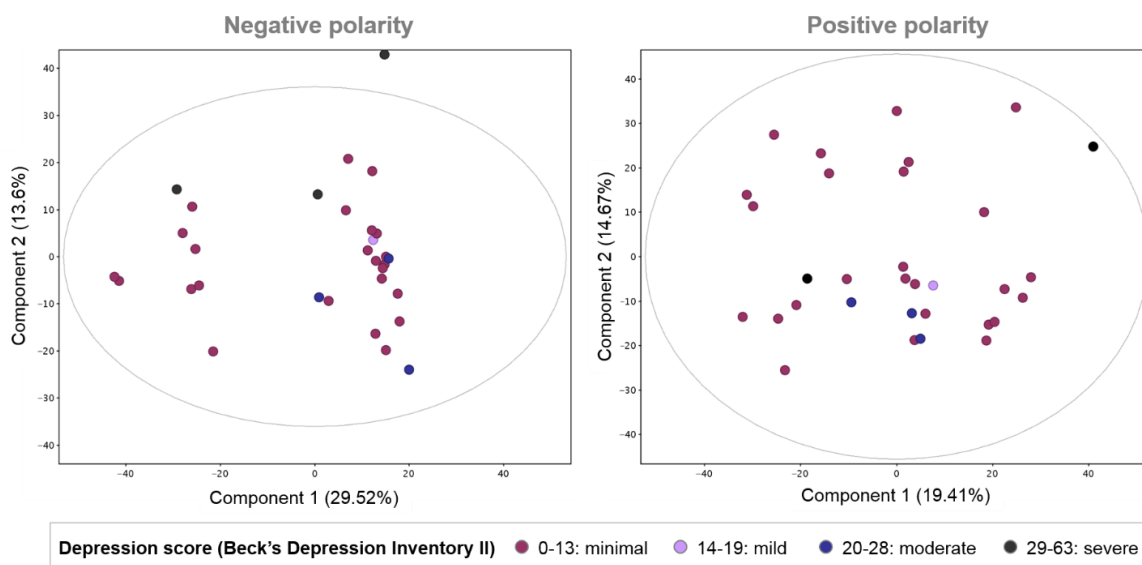
Figure 4.23. PCA plots visualising the effect of depression (Beck's Depression Inventory II scores) in profiling data from AKU human urine taken at baseline (non-targeted feature extraction). PCA plots show no clear separation in components 1 and 2 between the profiles of patients from the different depression score categories, as with targeted feature extraction.

4.3.2.6. HGD gene sequencing data

Based on the data from sequencing of the *HGD* gene, patients were divided into the following categories of mutation: substitution (n=19), nonsense (n=1), indel (n=2), substitution with - aberrant splicing / exon skipping / frameshift mutation (n=2), and substitution with indel and frameshift mutation (n=1) (231). Patients from 'substitution', 'nonsense' and 'indel' mutation categories were all bi-allelic for the same category of mutation but these categories also included patients with compound heterozygous mutations.

Data from targeted (Figure 4.24) and non-targeted (Figure 4.25) feature extraction at the baseline time point showed no clear separation between the profiles of patients from the different *HGD* mutation categories. As with the BDI-II score categories, statistical analyses were not performed to compare profiles between different mutation groups, due to the low number of patients with mutations other than substitutions.
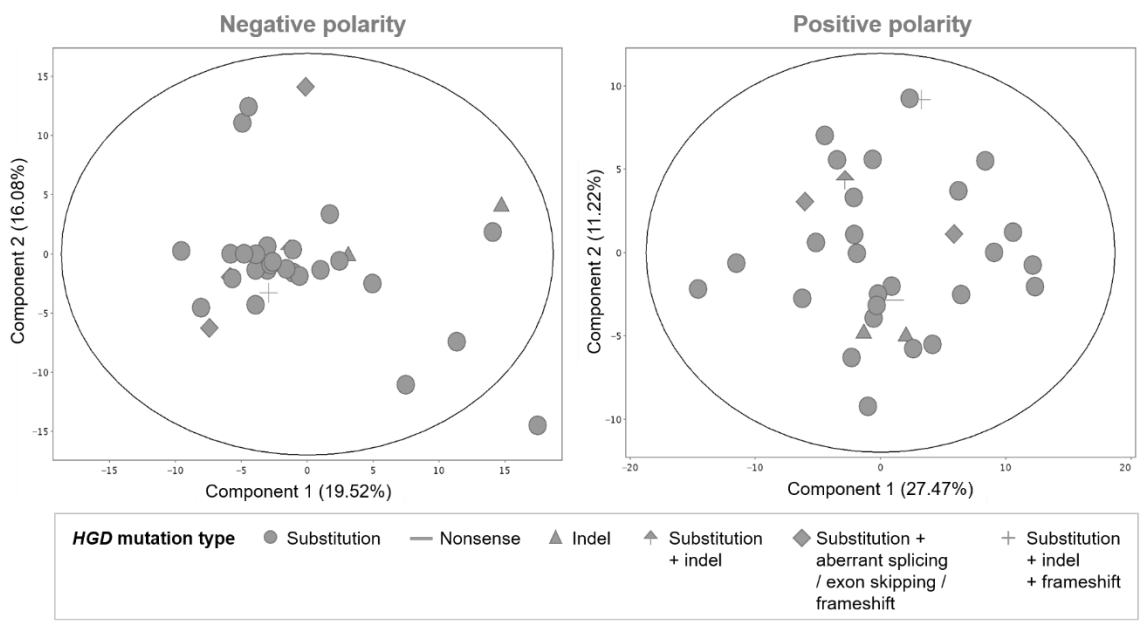
Figure 4.24. PCA to assess the effect of *HGD* mutation type in profiling data from AKU human urine taken at baseline (targeted feature extraction). PCA plots show no clear separation in components 1 and 2 between the profiles of patients with different types of *HGD* mutation.
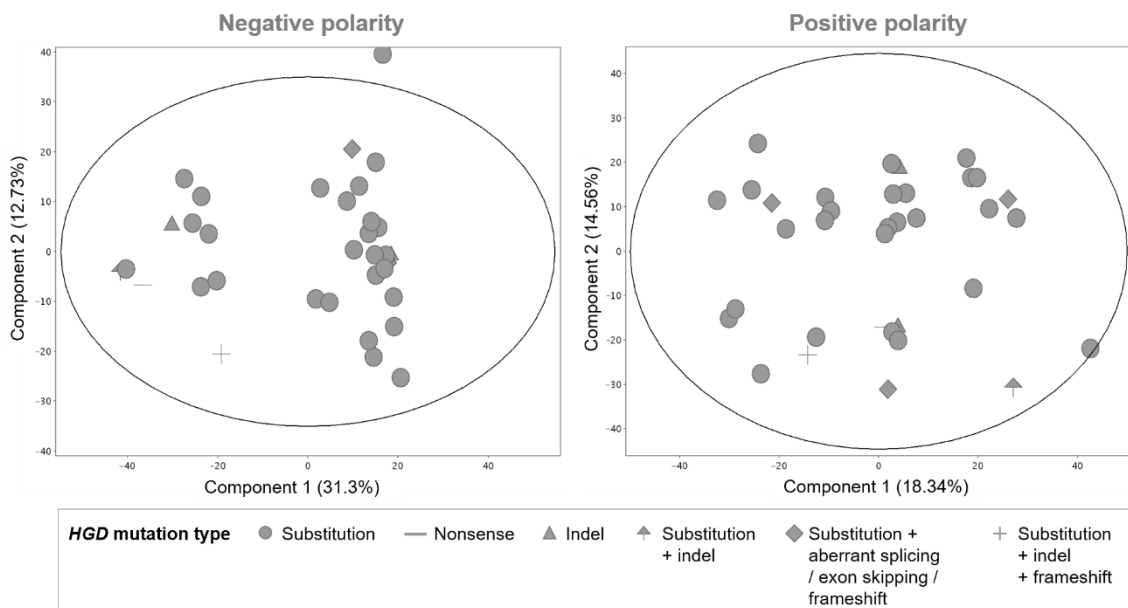
Figure 4.25. PCA to assess the effect of *HGD* mutation type in profiling data from AKU human urine taken at baseline (non-targeted feature extraction). PCA plots show no clear separation in components 1 and 2 between the profiles of patients with different types of *HGD* mutation, as with the data from targeted feature extraction.

4.4   Discussion

4.4.1. Effect of nitisinone treatment

The data described in this chapter show the power of the metabolomics approach developed for studying AKU. The findings extend what is currently known about nitisinone-induced hyper-tyrosinaemia by identifying for the first time the wider metabolic consequences of nitisinone treatment in AKU. In addition to confirming well-documented changes to the metabolites tyrosine, HPPA (both increased) and HGA (decreased) following nitisinone treatment (11,12,220), the data indicate alterations to three main metabolic pathways: tyrosine/dopamine

(Figures 4.26 and 4.27), tryptophan (Figure 4.28) and purine (Figure 4.29) pathways. These alterations are illustrated in the context of metabolic pathways in the aforementioned figures. The identification of the above well-known metabolite changes following nitisinone and the concordance between human and mouse data supports the validity of the data, as the $HGD^{-/-}$ mouse model studied here is shown to accurately recapitulate the early stages of the human AKU phenotype (50,68). Fourteen metabolites showed the same direction of alteration in both AKU human and mouse urine: 4-hydroxyphenylacetic acid, γ-glutamyl-L-tyrosine, acetyl-L-tyrosine, HPLA, HPPA, hydroxy-nitisinone, L-tyrosine, phenylacetic acid and tyramine increased on nitisinone; 2,5-dihydroxybenzaldehye, 2-hydroxy-4-(methylthio)butyric acid, 3,5-cyclic-AMP, HGA and xanthosine decreased on nitisinone.

4.4.1.1.   Hyper-tyrosinaemia and alternative routes of tyrosine clearance

Increases in the tyrosine derivatives acetyl-L-tyrosine, tyramine and 4-coumarate found here in urine could represent important clearance routes separate from the conventional tyrosine transamination pathway. These routes appear to be activated by the superphysiological tyrosine concentrations observed on nitisinone. In humans, tyrosine can be converted to tyramine from decarboxylation by aromatic-L-amino-acid decarboxylase (233); normally a minor metabolic route under physiological conditions. Increased plasma acetyl-L-tyrosine and γ-glutamyl-L-tyrosine have been shown previously with increasing doses of nitisinone in AKU (226,234). Increased tyrosine following nitisinone has been attributed to adverse events such as dermal toxicity and corneal keratopathy

caused by high tyrosine concentrations in the aqueous humour of the eye (14,142–144). Incidence of tyrosine keratopathy in AKU is estimated at 5% on a daily dose of 2mg nitisinone (14). Conversion of tyrosine to the more soluble acetyl-L-tyrosine could help to prevent tyrosine crystallisation in the eye underlying corneal lesions observed in some patients. On-nitisinone increases in the tyrosine metabolites HPPA, HPLA and 4-hydroxyphenylacetic acid are likely to be direct consequences of nitisinone's inhibition of HPPD activity (see Figure 4.26). The changes to HPLA and 4-hydroxyphenylacetic acid have not been previously reported in AKU.
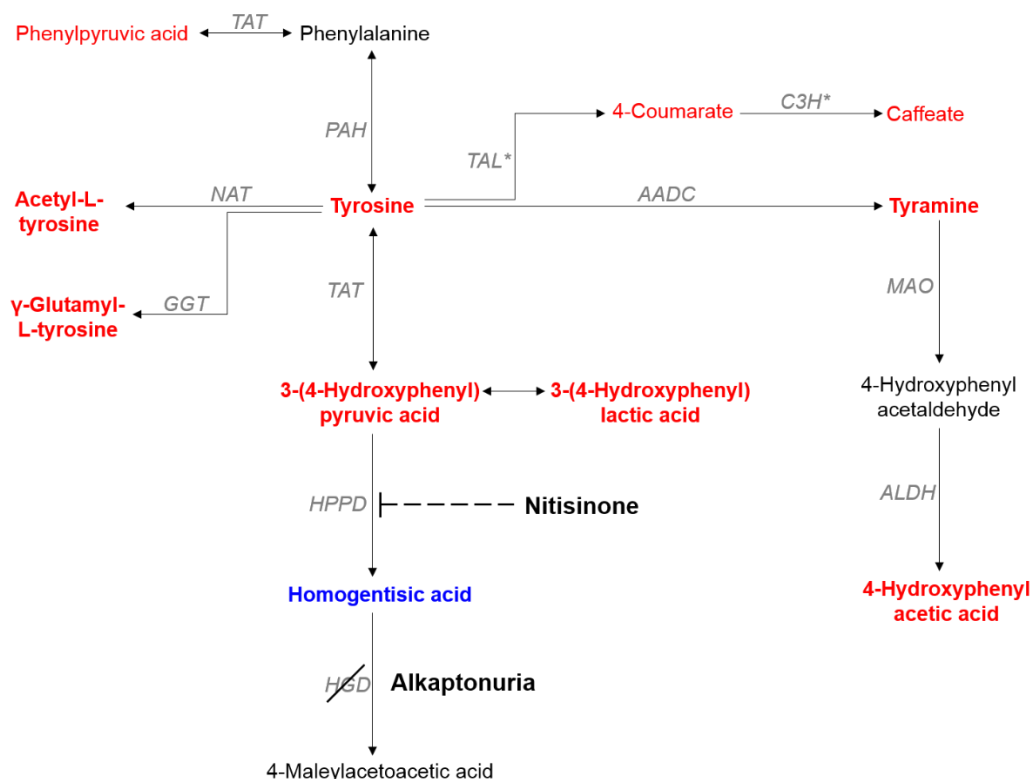
Figure 4.26. Non-conventional routes of tyrosine metabolism are activated as a consequence of nitisinone-induced hyper-tyrosinaemia. Red and blue indicate increased and decreased urinary abundance on nitisinone, respectively. Compounds in bold were altered in patients and mice with AKU. Nitisinone is used in AKU for reduction of HGA by its inhibition of HPPD. Consequently, the metabolites directly upstream of HGA accumulate; 3-(4-hydroxyphenyl)pyruvate 3-(4-hydroxyphenyl)lactate and tyrosine. The present data show that alternative mechanisms are recruited for metabolism of the increased tyrosine apart from the conventional tyrosine trans-aminase pathway. These include its conversion to acetyl-L-tyrosine, γ-glutamyl-L-tyrosine, tyramine and 4-coumarate.

AADC – aromatic amino acid decarboxylase, ALDH – alcohol dehydrogenase, C3H – p-coumarate-3-hydroxylase, GGT – gamma-glutamyl transferase, HGD – homogentisate 1,2-dioxygenase, HPPD – hydroxyphenylpyruvic acid dioxygenase, MAO – monoamine oxygenase, NAT –*N*-acetyl-transferase, PAH – phenylalanine hydroxylase, TAL – tyrosine ammonia-lyase, TAT – tyrosine aminotransferase (*enzyme not expressed in human or mouse).

4.4.1.2.   Altered dopamine and tryptophan-serotonin metabolism

Tyrosine is the precursor for catecholamines, thyroid hormones and melanin, suggesting that increased tyrosine substrate, as in tyrosinaemia, could lead to dysregulation of these pathways. Here, urinary profiles generally indicated increased levels of tyrosine-derived catecholamine metabolites and decreased levels of tryptophan-serotonin neurotransmitter metabolites on nitisinone therapy. Changes to catecholamine metabolites were increased 3-MT, a previously reported finding (235), and mandelic acid, and decreased HVA. These metabolites are derived from tyrosine via L-DOPA and dopamine. 3-MT is also derived directly from dopamine and is then converted to HVA. An increase in urinary 3-MT with decreased HVA therefore could suggest increased activity in the pathway dopamine catabolism from 3-MT, as opposed to alternative routes via noradrenaline and DOPAL. Alternatively, increased 3-MT could be a renal effect, reflecting decarboxylation of excess tyrosine delivered to the kidney to L-DOPA (the tyrosine pathway in full is only present in liver and kidney). Changes to metabolites of tryptophan-serotonin metabolism on nitisinone were decreased L-tryptophan, L-kynurenine and indoxyl sulfate and increased xanthurenic acid, with the addition of increased 4-quinolinecarboxylic acid in mice.

The data reported here corroborate but also add detail to previous reports of altered catecholamine and tryptophan-serotonin metabolism following nitisinone therapy. A previous report of increased urinary levels of the catecholamine metabolite 3-MT in nitisinone-treated AKU was replicated (235). It was also shown here for the first time that L-tryptophan is decreased in urine on nitisinone.

Decreased concentrations of the serotonin metabolite 5-hydroxyindoleacetic acid (5-HIAA) have been reported in HT-1 CSF (236) and in AKU urine (235) following nitisinone treatment, although this was not found here. Also for the first time, it was shown that decreased tryptophan metabolism following nitisinone is associated with alterations to the downstream kynurenine metabolic pathway; decreased kynurenine and increased xanthurenic acid. Additionally, the finding of decreased indoxyl sulfate implicates the indole pathway derived directly from tryptophan. Alterations to indolepyruvate pathway metabolites indole-carboxaldehyde, indole-lactate and indole-pyruvate were reported previously in AKU plasma following nitisinone, although these were increases (237).

The biological significance of the alterations on nitisinone to catecholamine and serotonin metabolism reported here are not fully understood. Much of the literature on nitisinone-induced hyper-tyrosinaemia comes from HT-1, where there are concerns that increased tyrosine may contribute to neurodevelopmental delay (46,221–224). It is estimated that up to 35% of juvenile HT-1 patients treated with nitisinone have intellectual disabilities (46), but it is unknown whether this is attributable to tyrosinaemia, low phenylalanine concentrations, acute liver disease (which children with HT-1 usually present with) or an intrinsic effect of HT-1. Several biochemical mechanisms have been proposed for neurodevelopmental delay associated with tyrosinaemia in HT-1 (236,238–240). Those which directly relate to catecholamine and tryptophan metabolism include increased CNS dopamine, the precursor for catecholamines, or decreased transport of tryptophan, the precursor for serotonin, into the brain.
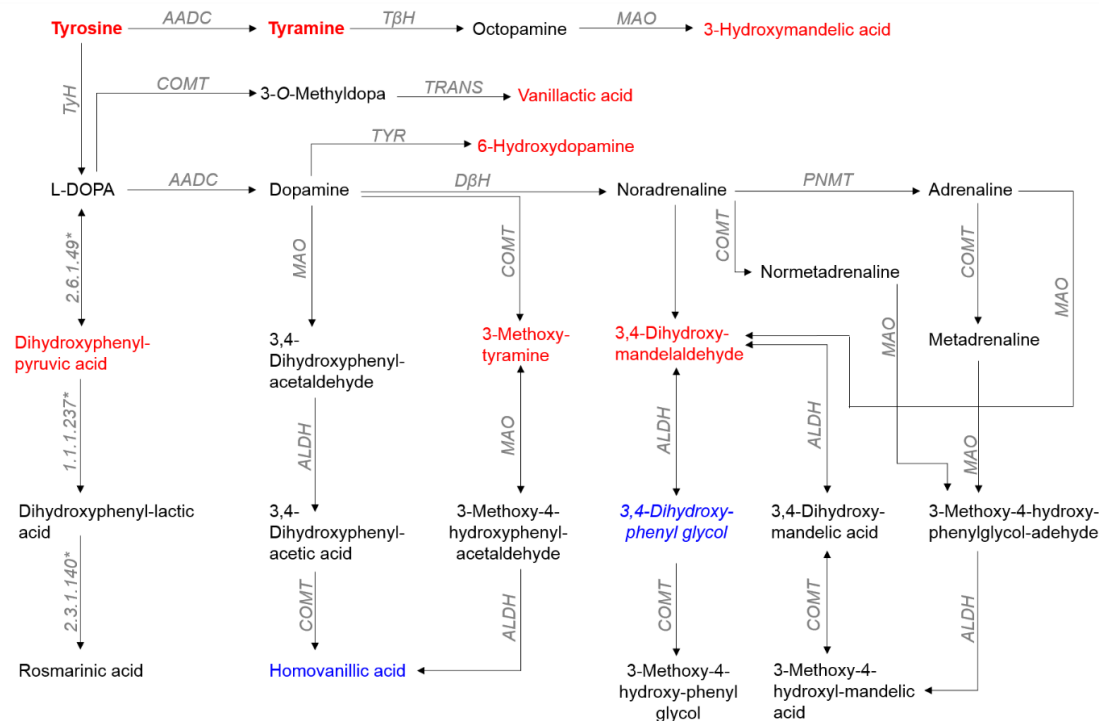
Figure 4.27. Alterations observed to dopamine metabolism as a consequence of nitisinone-induced tyrosinaemia. Red and blue indicate increased and decreased urinary abundance on nitisinone, respectively. Compounds in bold were altered in patients and mice with AKU. Compounds in italics were altered in mice only. Increases were observed on nitisinone to compounds derived directly from L-DOPA (dihydroxyphenylpyruvic acid), dopamine (3-methoxytyramine and 6-hydroxydopamine) and noradrenaline (3,4-dihydroxymandelaldehyde). Increased 3-hydroxymandelic acid was derived from increased tyamine via octopamine. Decreases were observed in homovanillic acid and 3,4-dihydroxyphenyl glycol, metabolites that are downstream of catecholamines and metadrenalines.

AADC – aromatic L-amino acid decarboxylase, ALDH – aldehyde dehydrogenase, COMT – catechol-O-methyltransferase, DβT – dopamine β-hydroxylase, MAO – monoamine oxidase, PNMT - TβH – tyramine β-hydroxylase, TRANS – transferase, TyH – tyrosine hydroxylase, TYR – tyrosinase, EC 2.6.1.49 – dihydroxyphenylalanine transaminase, EC 1.1.1.237 – hydroxypehnylpyruvate reductase, EC 2.3.1.140 – rosmarinate synthase (*enzyme not expressed in human or mouse).
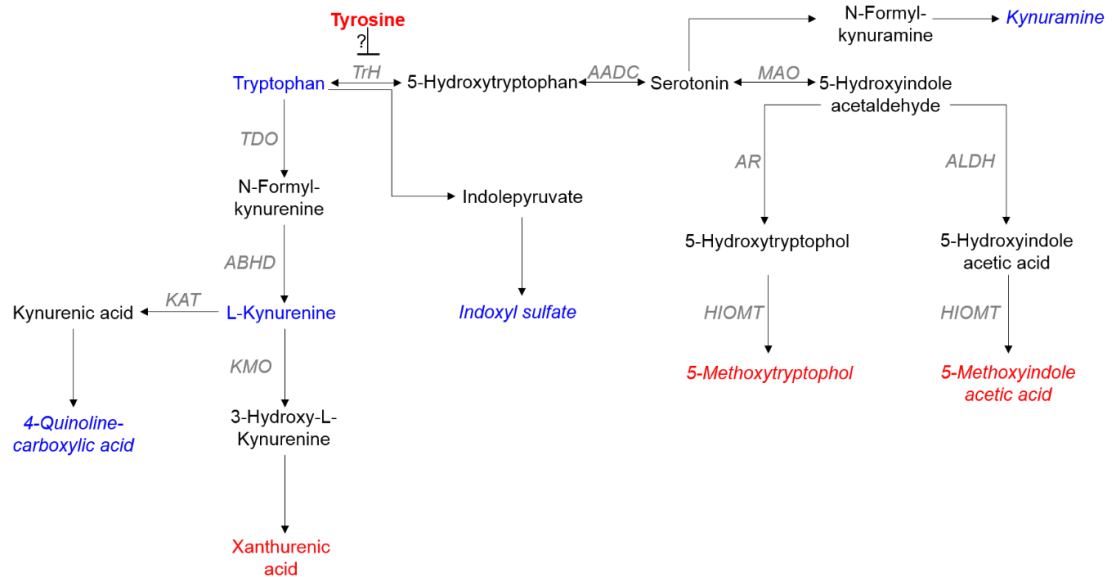
Figure 4.28. Alterations to tryptophan-serotonin metabolism and tryptophan-derived kynurenine and indole pyruvate pathways observed following nitisinone. Red and blue indicate increased and decreased urinary abundance on nitisinone, respectively. Compounds in bold were altered in patients and mice with AKU. Compounds in italics were altered in mice only. Increases were observed in metabolites downstream of serotonin via 5-hydroxyindole acetaldehyde; 5-methoxytryptophol and 5-methoxyindoleacetic acid. Conversely, decreased accumulation was observed in metabolites of the kynurenine and indole pyruvate pathways, which originate from tryptophan, were also observed. Metabolites in these pathways were decreased, with the exception of xanthurenic acid, a metabolite of kynurenine, which was increased. Kynuramine, derived from serotonin, was also decreased. It is previously suggested that metabolic alteration to tryptophan metabolism following nitisinone treatment is attributable to tyrosine's inhibition of tryptophan hydroxylase.

AADC – aromatic L-amino acid decarboxylase, ABDH – abhydrolase, ALDH – aldehyde dehydrogenase, AR – aldehyde reductase, HIOMT – hydroxyindole-O-methyltransferase, KAT – kynurenine aminotransferase, KMO – kynurenine monoxygenase, MAO – monoamine oxidase, TDO – tryptophan dioxygenase, TrH – tryptophan hydroxylase.

4.4.1.3.   Novel association between hyper-tyrosinaemia and alteration to the tryptophan-derived kynurenine pathway

Decreased kynurenine following nitisinone is an interesting and potentially significant finding (Figure 4.28). The kynurenine pathway of tryptophan metabolism has received considerable interest for its involvement in a range of CNS pathologies including depression (241,242), pain, particularly pain associated with migraine (243,244), and neurodegenerative diseases such as Alzheimer's, Parkinson's and Huntington's diseases and Multiple Sclerosis (241,242). Most metabolites of the kynurenine pathway possess neuroactive properties by their ability to act on *N*-methyl-D-aspartate (NMDA) and metabotropic glutamate receptors in the CNS, thereby modulating glutamate release and neuronal transmission (245,246). The kynurenine metabolite quinolinic acid is an NMDA receptor agonist and is thought to be an important contributor to neuronal excitotoxicity underlying neurodegeneration (247,248). The kynurenine metabolites 3-hydroxykynurenine and 3-hydroxyanthranilic acid can induce neuroinflammation by direct generation of free radical species even under physiological conditions (249). In contrast, kynurenic acid is considered to be neuroprotective by its antagonistic action on glutamate receptors (247).

Of particular relevance to AKU, NMDA glutamate receptors also have important roles in peripheral tissue; they are present in osteoblast and osteoclast cells and are thought to regulate bone formation (250–252). Kynurenine metabolism is also linked with musculoskeletal diseases such as rheumatoid arthritis (RA) and osteoporosis. Decreased plasma tryptophan and 3-hydroxyanthranilic acid were

206

reported in osteoporosis and RA compared to healthy controls, with increased kynurenine and xanthurenic acid in RA only (253). Anthranilic acid, a direct metabolite of kynurenine was also increased in osteoporosis at baseline compared to healthy controls, and its decrease to control levels after two years of treatment was associated with clinical improvement indicated by bone densitometry (254). Tryptophan, kynurenine and anthranilic acid were increased in synovial fluid in RA compared with OA, suggesting that tryptophan-kynurenine metabolism is associated with the increased inflammatory or autoimmune pathophysiological component of RA (255). Given the association of kynurenine metabolism with a range of pathophysiological processes both in the CNS and periphery, the overall de-activation of this pathway following nitisinone in AKU, as indicated by urinary metabolite profiles, is potentially beneficial. Decreased bioavailability of kynurenine on nitisinone could reduce the levels of its derivative toxic metabolites 3-hydroxykynurenine, anthranilic acid, and 3-hydroxyanthranilic acid which are shown to induce oxidative stress. However, the protective effects of some kynurenine pathway metabolites, in addition to the diversity and complexity of their regulatory effects on glutamatergic signalling in the CNS and periphery should also be considered. However it is important to consider whether kynurenine pathway alteration is a compensatory effect to mitigate neuronal damage, a direct off-target effect of nitisinone or a consequence of nitisinone alleviating oxidative stress associated with HGA/ochronosis.

4.4.1.4.   Altered purine metabolism

The finding of altered purine metabolism following nitisinone is also novel. Most of the changes in this pathway were decreased metabolite levels (xanthosine, adenine, 3-ureidopropionate and allantoin), except for caffeic acid which was increased. The purine pathway is the metabolic route through which purines are broken down into uric acid. Levels of metabolites in this pathway are affected by endogenous purine production and exogenous purine pools (256). Increased serum uric acid is implicated in a range of diseases including gouty arthritis (gout), type 2 diabetes and cardiovascular and neurological disease (257,258). Inflammation is one of the major pathogenic mechanisms by which uric acid causes disease in these conditions. This is perhaps most recognised in gout, where high uric acid leads to crystal deposition in joints, tendons and tissues, resulting in inflammation, swelling and pain in joints (258,259). Xanthosine and allantoin are also metabolites derived from xanthine that are known to accumulate in oxidative conditions associated with chronic inflammation (260–263). Allantoin is not usually observed in human urine although it can be formed from uric acid non-enzymatically under high levels of reactive oxygen species. It is therefore conceivable that reduced levels of uric acid precursors (such as the purine nucleobase adenine) which are strongly linked with inflammation, could have additional benefit in AKU arthritis. Caffeic acid has been grouped into the xanthine pathway as it, in addition to its oxidation product, is shown to inhibit activity of xanthine oxidase (EC 1.17.3.2) (264,265); the enzyme that converts hypoxanthine to xanthine and then to uric acid by oxidation. Increased caffeic acid

could also be an additional beneficial effect of nitisinone; caffeic acid is known to have a range anti-inflammatory, antioxidant, immunomodulatory and neuroprotective properties (266–269). As summarised in Figure 4.29, the alterations to metabolites of the purine pathway, particularly decreases in xanthine-derived compounds, following nitisinone treatment are possibly a reflection of reduced oxidative stress due to HGA reduction. Oxidative stress is thought to be an important pathological feature of AKU that is a direct consequence of HGA and ochronotic osteoarthropathy (16,41).
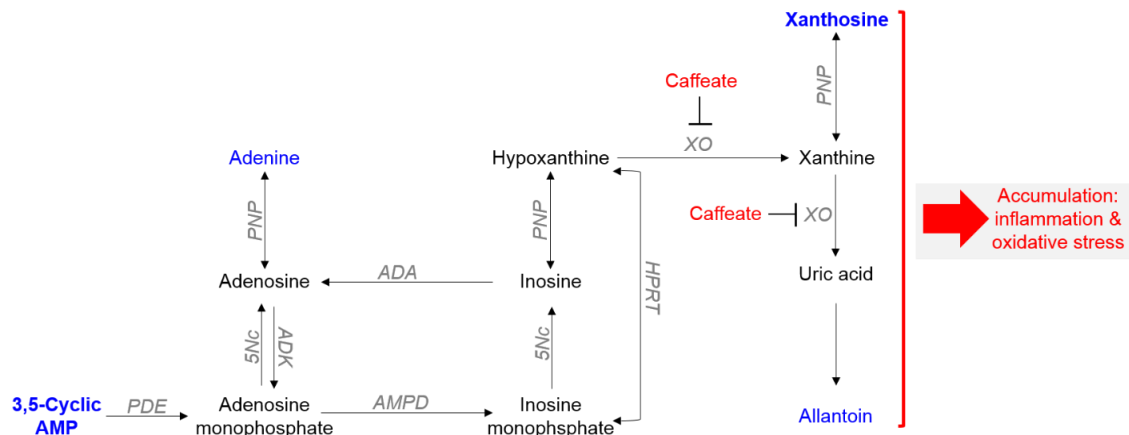
Figure 4.29. Alterations to purine metabolism observed following nitisinone treatment. Red and blue indicate increased and decreased urinary abundance on nitisinone, respectively. Compounds in bold were altered in patients and mice with AKU. Altered compounds not in bold were observed in patients only. The purine pathway metabolites altered were decreased following nitisinone, except for caffeate which increased. Decreases in xanthosine and allantoin, metabolites derived from xanthine, are particularly noteworthy, as increases in these compounds have been linked previously with oxidative stress and inflammation. It is proposed that the decreases in xanthine metabolites are attributable to the inhibition of caffeate on xanthine oxidase activity.

5Nc – 5'-nucleotidase, ADA – adenosine deaminase, ADK – adenosine kinase, AMPD – adenosine monophosphate deaminase, HPRT – hypoxanthine phosphoribosyltransferase, PDE – phosphodiesterase, PNP – purine nucleoside phosphorylase, XO – xanthine oxidase.

4.4.1.5.   Relationship between tyrosine, tryptophan and purine metabolism

Interestingly, other metabolic profiling studies have also found a link between the three altered pathways here of tyrosine, tryptophan and purine metabolism. Analysis of CSF from subjects with mild cognitive impairment or Alzheimer's disease found that these conditions were associated with an overlapping network of metabolic changes within tyrosine, tryptophan and purine pathways (270,271). The ratio of kynurenine : tryptophan was also increased in Alzheimer's disease (270). The relationship between alterations to tyrosine, tryptophan and purine metabolic pathways is not well recognised nor understood, but the concurrent association with Alzheimer's disease and nitisinone-treated AKU has important implications for cognitive function following nitisinone-induced tyrosinaemia which warrants further study.

4.4.1.6.   Miscellaneous metabolite alteration

Changes were also observed for other metabolites on nitisinone that are not easily grouped into common metabolic pathways, therefore making interpretation challenging. 2-Hydroxybutyric acid, methyl vanillate, 3,5-cyclic AMP, 2-hydroxy-4-(methylthio)butyric acid and L-threonine were decreased, and ethylmalonic acid was increased on nitisinone. These findings implicate a diverse range of metabolic alterations as a consequence of nitisinone therapy. Ethylmalonic acid and 2-hydroxy-4-(methylthio)butyric acid are short-chain organic acids. Decreased methyl vanillate could be a dietary change as it is derived from vanillate, a flavouring agent frequently used in foods. L-threonine is an essential

amino acid, which has been previously monitored in AKU serum by our group following nitisinone as part of a wider amino acid profiling study (220). This previous study found a trend of decreased serum threonine after 36 months on nitisinone, but this did not reach statistical significance. 2-Hydroxybutyric acid is produced in the conversion of cystathionine to cysteine in the methionine-glutathione synthesis pathway (272). One could argue that decreased 2-Hydroxybutarate on nitisinone is a marker of decreased oxidative stress; at times of increased oxidative stress more homocysteine is converted to cystathione to increase the production of glutathione via cysteine as a redox homeostasis mechanism (273). Regarding the on-nitisinone decrease in cAMP, the role of cAMP is well-recognised in the musculoskeletal system and in endocrinology more generally. cAMP is a key signal transduction secondary messenger in a range of physiological processes. The role of cAMP has been extensively studied in the cAMP – protein kinase A signalling pathway, which is highly linked to osteogenesis (274). cAMP is also thought to be involved in stimulation of native growth factors which has attracted interest in it as a therapeutic target in the fields of musculoskeletal engineering and regeneration (275). Given the diverse roles of cAMP in signalling cascades in various tissues, the exact significance of the urinary decrease in cAMP observed here on nitisinone is uncertain and requires further investigation.

4.4.1.7. Consideration of potentially confounding factors

The potential effect of diet on some of the metabolic changes observed on nitisinone should also be considered. Dietary proteins contribute significantly to

exogenous purines and tryptophan (256,276,277), so it is possible that general de-activation of these metabolic pathways could relate to diet. However, patients attending the NAC are not specifically advised to restrict their protein intake upon commencement of nitisinone therapy. The within-subjects study design (*i.e.* paired data) should also improve control over inter-individual differences in demographic or lifestyle factors such as diet. In the mouse study it was possible to strictly control potentially confounding factors that could affect metabolism, such as diet and genetic diversity. This increases the likelihood that the observed metabolite changes are attributable to nitisinone therapy.

4.4.1.8. Contribution of complementary non-targeted data extraction

It is worth noting the contribution of the complementary non-targeted feature extraction approach conducted in parallel to AMRT-based extraction. Although the considerable number of chemical entities from non-targeted feature extraction that remained unidentified (by AM match <10 ppm; see Tables S4.01 and S4.02 in Supporting materials, Appendix 3) demonstrates the challenging nature of this approach, its employment helped to elucidate further individual metabolite changes within the pathway-level alterations discussed above. These included further increases in metabolites of the wider tyrosine pathway; dihydroxyphenylpyruvic acid, phenylpyruvic acid, 3-hydroxymandelic acid (DHMA), hydroxydopamine, *N*-acetylvanilalanine, vanillactic acid and vanillin 4-sulfate. Hydroxydopamine and vanillactic acid are catecholamine metabolites deriving from tyrosine via L-DOPA (Figure 4.27). Concentrations of 5- and 6-hydroxydopamine were previously found to be increased in urine from patients

with Parkinson's treated with L-DOPA, and 6-hydroxydopamine is thought to be derived from dopamine by the action of tyrosinase (278). Vanillactic acid is derived from L-DOPA via 3-O-methyldopa (279). 3-Hydroxymandelic acid could be associated with the increased tyramine on nitisinone, as it is a product of dopamine and octopamine; octopamine is derived from tyrosine via tyramine.

Interestingly, non-targeted extraction on the mouse data also revealed further changes to metabolites associated with tryptophan-serotonin metabolism; 5-methoxyindoleacetate and 5-methoxytryptophol increased and kynuramine decreased on nitisinone. Kynuramine is an endogenously-occurring di-amine metabolite, and its decrease on nitisinone appears to be in accordance with that also observed for L-kynurenine. Whereas L-kynurenine is derived from tryptophan, via formyl-kynurenine, kynuramine is thought to be formed via serotonin via formyl-kynuramine (280). 5-Methoxyindoleacetate and 5-methoxytryptophol are metabolites of serotonin, formed via 5-hydroxyindole acetaldehyde. Increases in these metabolites therefore appear to be related, but are perhaps unexpected given previous reports of decreased urinary 5-HIAA in AKU patients on nitisinone (235). Together with the data discussed above, these additional findings could be taken to suggest the following urinary pathway alterations on nitisinone a) accumulation of downstream products of serotonin following its conversion to 5-hydroxyindole acetaldehyde via monoamine oxidase, b) decreases in metabolites of the tryptophan-derived kynurenine pathway, and c) decreases in metabolites of the tryptophan-derived indolepyruvate pathway. These observations are incorporated into the tryptophan-serotonin pathway

summary in Figure 4.28. Further analysis should aim to confirm the compound assignments of entities obtained from non-targeted feature extraction by obtaining chemical standards, where possible, and analysing them under the same conditions to match RT and also generate MS/MS spectra for confirmatory matching of fragmentation spectra.

4.4.2. Associations between the metabolome and demographic or clinical data

4.4.2.1.   Associations with ^{18}F-NaF PET imaging data

In these studies, also investigated for the first time in AKU were potential associations between metabolome-wide profiles and clinical/demographic data. These associations were of interest aside from the effect of nitisinone and were therefore primarily studied in the data at the baseline time point (or additionally at 24 months on nitisinone if an effect was observed at baseline). The clinical variables with the clearest observable impact on the metabolomes of these patients were those related to monitoring the progression of osteoarthropathy; ^{18}F-NaF PET imaging data and self-reported lumbar/thoracic back pain. These clinical data from the NAC patient cohort are described extensively by Taylor (2018). In this form of PET imaging, the injected ^{18}F-NaF tracer binds to tissue actively undergoing mineralisation and is therefore designed to measure the rate of bone remodelling or metabolism (281). Employment of this form of PET imaging has identified early pathological changes in OA before the emergence of clear radiographic evidence of cartilage degeneration. For example, in patients

with early-stage OA it was shown that increased bone remodelling in the hip, as measured by ^{18}F-NaF, correlated with radiographic stage of OA and severity of hip pain (282). ^{18}F-NaF uptake to the spinal musculoskeletal system shows a unique distribution in AKU, with increased uptake to intervertebral disks compared with the bony vertebrae. The inverse of this pattern is usually observed for healthy subjects. Increased ^{18}F-NaF uptake to the intervertebral disks in AKU is thought to reflect calcification of the cartilaginous intervertebral disks associated with ochronotic osteoarthropathy. In this patient cohort it was previously shown that ^{18}F-NaF uptake to intervertebral disks increases with age, reaching maximum values at around 60 years. Mean ^{18}F-NaF uptake to the vertebrae or intervertebral disks did not change significantly across visits over this period for patients attending the NAC. Combined lumbar/thoracic pain scores were previously shown to be positively correlated with age and ^{18}F-NaF uptake to the intervertebral disks, consistent with this form of uptake as a marker of OA progression. Interestingly, ^{18}F-NaF uptake to the vertebrae was found to be negatively correlated with pain and showed a general decline with age in these patients; the opposite pattern to that observed for ^{18}F-NaF uptake to the intervertebral disks.

From the imaging data, ^{18}F-NaF uptake to the intervertebral disks appeared to have the greatest influence on the metabolomic data. This was particularly true in the data from non-targeted feature extraction, in which the major bimodal distribution of the data across all time points observed from PCA corresponded to the grouping of patients by high (value >10) *versus* low (value <10) uptake to

the intervertebral disks (Figure 4.18). Clearer PCA separation was observed in the data from non-targeted feature extraction for 'high' *versus* 'low' groups for both intervertebral disk ^{18}F-NaF uptake and back pain, which supports the parallel approach of mining the data for chemical entities/compounds that were not contained in the AMRT databases generated. The identification of one chemical entity, AM-matched with caffeic acid sulfate, that increased in patients with high intervertebral disk ^{18}F-NaF uptake and back pain was particularly interesting. The biological interpretation of this finding is not clear and warrants further attention. Caffeic acid sulfate is presumably a metabolic product of caffeic acid (caffeate), an antioxidant which was also shown to be increased in these patients while on nitisinone (section 4.3.1.1.).

Decreased urinary adenosine/2-deoxyguanosine was also observed for patients with high ^{18}F-NaF uptake to intervertebral disks. The method employed cannot distinguish between these two possible compound identifications, as both have identical AM and RT using the chromatographic technique employed here (Method 2; Chapter 2, section 2.2.2.). Adenosine is associated with oxidative stress observed in chronic inflammatory conditions such as arthritis (283). Furthermore, methotrexate, an approved therapy for RA, is proposed to act by inhibiting adenosine-mediated inflammation (284). In RA, methotrexate results in an increase in extracellular adenosine, which binds to cell surface adenosine receptors. As a consequence of this binding, lymphocyte proliferation is inhibited, decreasing the levels of inflammatory cytokines and promoting an anti-inflammatory environment (285). There is also evidence that increased adenosine

production induces chondrocyte apoptosis. Adenosine in the presence of adenosine deaminase or adenosine kinase inhibitors induced cell death in a murine chondrocyte cell line, and activity of 5'-nucleotidase, the enzyme that produces adenosine, increased with development of OA in the STR/ort mouse model of spontaneous OA and was particularly elevated close to the sites of OA lesions (286). In addition, concentrations of 5'-nucleotidase have been found to be markedly elevated in synovial fluid from patients with RA, and slightly elevated in OA (287). However, the association between local increases in adenosine in tissues affected by OA and urinary excretion of adenosine is unclear. For example, relatively increased concentrations of urinary 5-aminoimidazole-4-carboxamine, an analogue of adenosine monophosphate, but not adenosine was associated with reduced disease severity in patients with RA treated with methotrexate (285). Nevertheless, the implied association in the present data between decreased urinary adenosine and increased OA and inflammation in AKU, as indicated by intervertebral ^{18}F-NaF uptake, is worthy of further study.

The observation of entities with differential abundance between patients with high *versus* low ^{18}F-NaF uptake to the vertebrae was also of interest. Previous studies have found bone ^{18}F-NaF uptake to increase in OA and in bone tissues subject to greater mechanical stress, such as loading (227,282,288), which are both known to stimulate bone turnover or remodelling. However, it is unclear whether the increased ^{18}F-NaF to the vertebrae in the present AKU cohort is a true reflection of the aberrant bone remodelling observed as a consequence OA. Uptake of ^{18}F-NaF to the vertebrae generally decreased with age in this cohort,

the opposite of the trend for intervertebral disks, possibly reflecting the well-recognised decrease in bone density in older populations. These entities should therefore be considered potential markers of bone remodelling or turnover more generally, but not necessarily as a specific consequence of osteoarthropathy.

The entity identified as iduronic acid, increased in the 'high' vertebral ^{18}F-NaF uptake group, is a particularly interesting potential marker of bone tissue degradation. Iduronic acid is the major uronic acid component of the GAGs dermatan sulfate and heparin (289). GAGs are polysaccharides made from disaccharide repeating units comprising an amino sugar and an uronic sugar or galactose. GAGs are an important component of the extracellular matrix; their high polarity enables them to bind and organise water molecules, which supports the lubricating or shock-absorbing properties that GAGs lend to tissues such as bone and cartilage (290). The increased urinary iduronic acid in these patients could therefore reflect the breakdown of the aforementioned GAGs as part of the process of tissue remodelling. Proline was another entity associated with vertebral ^{18}F-NaF uptake in this study. Urinary proline is a potential marker of bone turnover, as it contributes 12% of the total protein content of type I bone collagen (291). Increased urinary proline would support this interpretation, but in contrast proline was decreased in patients with high vertebral ^{18}F-NaF uptake. Malonic acid was also decreased in patients with high vertebral ^{18}F-NaF uptake. *In vitro*, malonic acid supplementation is shown to select for bone marrow stromal cells that subsequently undergo mineralisation to become osteoprogenitor cells,

which are the precursors for specialised bone cells (292). However, the role of malonic acid in bone metabolism more generally is not clear.

One entity that was increased in patients with high ^{18}F-NaF uptake to intervertebral disks is particularly noteworthy. This entity was initially of high interest as a potential endogenous marker of osteoarthropathy, but was subsequently identified as paracetamol (acetamainophen). This finding therefore highlights the importance of also considering exogenous drug metabolites when attempting to chemically identify unknown potential biomarkers. This finding is not surprising, but can be considered to validate the approach employed to search for novel markers associated with osteoarthropathy and/or pain.

It is also important to note absence of an association between entities attributable to ochronotic pigment itself and the PET imaging and pain data. As ochronosis is central to the osteoarthropathy and joint destruction observed in AKU (16), one might expect the concentrations of compounds attributable to ochronotic pigment to be highly associated with these clinical data. It is therefore somewhat surprising that HGA was not found to correlate with ^{18}F-NaF uptake to the intervertebral disks or vertebrae. In support of this finding, previous studies of a separate cohort of non-treated AKU patients showed no statistically significant correlation between urinary or serum HGA and these ^{18}F-NaF-PET imaging measures in the lumbar and thoracic spine (227). These data therefore suggest that there is not a simple linear relationship between concentrations of HGA in the circulation or urine and the degree of spinal osteoarthropathy as indicated by ^{18}F-NaF-PET. A potential explanation is that compounds related to ochronotic pigment have a

closer relationship with osteoarthropathy than HGA, but that these were not detected here. As discussed in Chapter 7, chemical analysis of ochronotic pigment poses a serious challenge for various reasons, some of which specifically relate to the ESI MS technique employed here. Furthermore, neither the chemical structure nor the formation mechanism of ochronotic pigment are known, therefore mining the data for specific associated chemical entities is challenging. It is possible that some of the chemical entities of interest that remained unknown are derived from aggregation of HGA or ochronotic pigment. However, none of these unidentified entities matched the mass of entities observed in the mass spectra of HGA or BQA standard, or pigment derived from synthetic HGA (see Chapter 7). Further biochemical analyses in larger patient cohorts, such as the recently completed SONIA (Suitability of Nitisinone in Alkatonuria) II clinical trial (139), are encouraged in order to capture the complex pathophysiological factors that mediate progression of ochronotic osteoarthropathy in AKU. Additionally, it is worth pointing out that serum and urinary HGA (and flux down the tyrosine pathway more generally) represent snapshots in time. Lifetime exposure to HGA is probably more relevant and could explain the lack of correlation observed between HGA and osteoarthropathy measured by ^{18}F-NaF-PET.

4.4.2.2.   Associations with gender

There was a marked effect of gender in the metabolomic data (from targeted AMRT-based feature extraction) for mice, but not patients. This prominent gender difference was perhaps only observed for mice due to the strict control over environmental factors with potential to have major influences on the metabolome,

including age, diet and genetic diversity. The most notable gender difference was in a group of metabolites that were clearly elevated in urine from female mice: 4-hydroxyphenylacetate, 4-quinolinecarboxylic acid, 5-valerolactone, anthranilate, erythritol, histamine and riboflavin. Of these metabolites, the greatest increases were for histamine (FC = 16), riboflavin (FC = 18) and 5-valerolactone (16). These findings are potentially relevant to the observation that disease is more severe in males with AKU; symptoms in females tend to lag behind those in males by approximately 10 years (18).

Previous research has shown increased urinary histamine excretion in the rat (293,294), but not the mouse to my knowledge. It is hypothesised that in the rat, the reason for this observation is that males, but not females, possess an additional minor histamine-metabolising pathway (294). Riboflavin (also known as vitamin B_2) is a vitamin of dietary origin. Riboflavin is an essential component of flavin mononucleotide and flavin adenine dinucleotide; coenzymes that play major roles in energy production and other diverse functions such as cellular function, growth and development and metabolism of fats, drugs and steroids (295). Excess riboflavin is excreted in the urine in free form. In humans, urinary riboflavin (μg/g creatinine) was previously shown to be slightly increased for females compared with males in a cohort of 412 adults in South Korea (296), although this could reflect the increased ratio of dietary riboflavin intake to energy intake for females in this study. Anthranilate (anthranilic acid) is an aromatic acid of the tryptophan biosynthesis pathway. Increased urinary anthranilic acid was previously observed in female mice, compared with males, injected with the food

flavouring and fragrance agent cinnamyl anthranilate (297). Increased excretion of anthranilic acid in female mice was proposed to reflect the increased capacity of female mice to metabolise cinnamyl anthranilate, by cytochrome P450-mediated hydrolysis, resulting in formation of anthranilic acid which was excreted in urine. Increased 4-hydroxyphenylacetate in urine of female mice is an interesting finding, as this metabolite was elevated in patients and mice on nitisinone (see section 4.3.1.1) and thought to be associated with tyrosine accumulation. However, the reason for its increase in female mice is unclear; it is unlikely to be associated with gender-related differences in response to nitisinone, as 4-hydroxyphenylacetate was increased in females at baseline, and gender differences were not observed for tyrosine or HGA on nitisinone.

Prominent gender differences were also observed in the metabolites benzaldehyde and thymine; these were increased in urine from male mice. The reason for the marked increase in benzaldehyde (FC = 23) is unclear. Benzaldehyde is a volatile compound used in some cosmetics and flavouring agents and it is perhaps surprising that this compound is found in mouse urine. Thymine is a pyrimidine nucleobase, and increases in urinary thymine is a diagnostic indicator of dihydropyrimidine dehydrogenase deficiency (OMIM 222748), a rare inherited disease caused by a disruption in pyrimidine catabolism. In accordance with the present data, a previous study of 276 Japanese adults found that urinary thymine (and dihydrothymine) was significantly increased in males compared with females, and markedly increased beyond normal reference ranges in patients with dihydropyrimidine dehydrogenase deficiency (298).

4.4.2.3.　　　Lack of clear associations with age, BDI-II score and *HGD* gene sequencing data

Stratification of patients according to age, BDI-II score and *HGD* mutation type at baseline did not show any clear effect of these variables on metabolome profiles. These factors were investigated at baseline in order to assess potential sub-groups of patients based on disease aside from the effect of nitisinone. These analyses were somewhat limited by the low numbers of patients in some categories, specifically the oldest and youngest age categories, BDI-II score categories other than 'minimal' and *HGD* mutations other than bi-allelic substitutions.

4.4.3. Inferring alterations to metabolism from urine

The metabolite changes reported here are urinary, so it is important to consider whether they relate to changes in the circulation and CNS. CNS concentrations are particularly important for the altered tryptophan metabolites given their neuroactive properties.

Urine metabolomics is recognised as a powerful technique for studying novel metabolic markers and biochemical pathways of disease (299). Urine is readily available, easily obtained, and often simpler to analyse than other bodily fluids, such as blood, which typically requires deproteinisation in metabolomics applications. Urine concentrations of metabolites are determined by renal clearance of waste products which serves to achieve homeostasis. The comparison of metabolite concentrations in different bodily fluids is highly

complex as there are numerous sites of synthesis for many metabolites. The catecholamines, for example, as altered here, have a complex pattern of synthesis involving sympathetic neurons of the CNS, the adrenal medulla, the gastrointestinal and mesenteric systems and the liver (300). In addition, some catecholamine metabolites, such as MHPG (3-methoxy-4-hydroxyphenylglycol; derived from VMA) are observed primarily in their conjugated forms in urine; only 3% of MHPG in urine is in its free non-conjugated form, the remainder is present as glucuronide or sulfate conjugates (301).

The source of a given metabolite observed in urine, and therefore its role in a given state or disease, can sometimes be unclear. Nevertheless, some of the urine metabolite alterations reported here agree with changes previously observed in plasma/serum on nitisinone; these include HGA (decreased), tyrosine, *N*-acetyltyrosine and HPPA (all increased). This supports that urinary concentrations of metabolites can be indicative of their concentrations in the circulation or CNS, although further research is required to confirm the indication.

4.5   Chapter 4 conclusion

In conclusion, employment of the developed metabolic profiling strategy, as described in Chapter 2, helped to identify previously unreported urinary metabolite changes following treatment with the HGA-reducing agent nitisinone. In these data, novel associations were identified between nitisinone treatment and alterations to the tryptophan-kynurenine and purine metabolic pathways. Together the findings suggest that nitisinone-induced tyrosinaemia induces a

pattern of altered metabolism to three main metabolic pathways which previously were generally considered to operate independently; namely the pathways of tyrosine, tryptophan and purine metabolism. Use of the profiling strategy developed has therefore significantly advanced our knowledge of metabolism in AKU, and supports its application to other key questions relating to the pathogenesis and treatment of AKU.

Also considered here, for the first time in AKU, were potential associations between metabolomic profiles and patient demographics and clinical data. The parameters that showed the greatest impact on urinary metabolomes were gender (in mice only) and clinical data relating to musculoskeletal disease; ^{18}F-NaF PET imaging and pain. The metabolites/entities with profiles that correlated with these clinical parameters are potential markers of ochronotic osteoarthropathy, the central pathological feature of AKU that is responsible for tissue and joint destruction.

Acknowledgements

5.0   METABOLOMIC CHARACTERISATION OF A NEW TARGETED *HGD*[-/-]

MOUSE MODEL OF ALKAPTONURIA

5.1 Introduction

Metabolic profiling has clear potential in AKU as both a phenotyping and biomarker discovery tool. However, untreated AKU has not been compared with non-AKU at the metabolome level before. Here the aforementioned LC-QTOF-MS profiling strategy was employed to characterise the metabolome of a mouse model of AKU by targeted homozygous knockout of the *HGD* gene (*HGD*$^{-/-}$; see Chapter 1, section 1.1.8), with non-AKU heterozygous knockout (*HGD*$^{+/-}$) mice as controls (Experiment 1; Figure 5.01). The *HGD*$^{-/-}$ mouse accurately recapitulates the human AKU disease phenotype and closely matches that of a previously reported AKU mouse with *HGD* knockout by ENU mutagenesis (50,69), with the hallmark features of elevated plasma and urine HGA, development of ochronosis and its inhibition by nitisinone (302). Validation of metabolite alterations observed was performed by assessing whether the direction of alteration was reversed following nitisinone treatment in *HGD*$^{-/-}$ mice or patients with AKU (Experiment 2, Figure 5.02). Additional validation of the entities that were increased in *HGD*$^{-/-}$ mice was sought by performing a metabolic flux experiment in which mice were injected with stable isotopically-labelled HGA (Experiment 3, Figure 5.01). This HGA flux experiment aimed to establish whether the increased metabolites were directly derived from HGA.

The overall aim of these experiments was to investigate metabolic markers of disease in AKU. Identification of novel metabolite alterations has potential to elucidate novel disease mechanisms and to improve current strategies for monitoring the AKU disease process and its response to treatment. The reason

for studying the $HGD^{-/-}$ mouse was that in these mice HGD disruption is performed in a targeted and controlled way, and that potentially confounding genetic and environmental influences on the metabolome could be carefully controlled; the mice were bred from the same colonies, fed the same diet and housed in identical conditions. In this way, studying homozygous *versus* heterozygous HGD knockout mice is considered to offer the most relevant comparison of the AKU *versus* non-AKU metabolome.
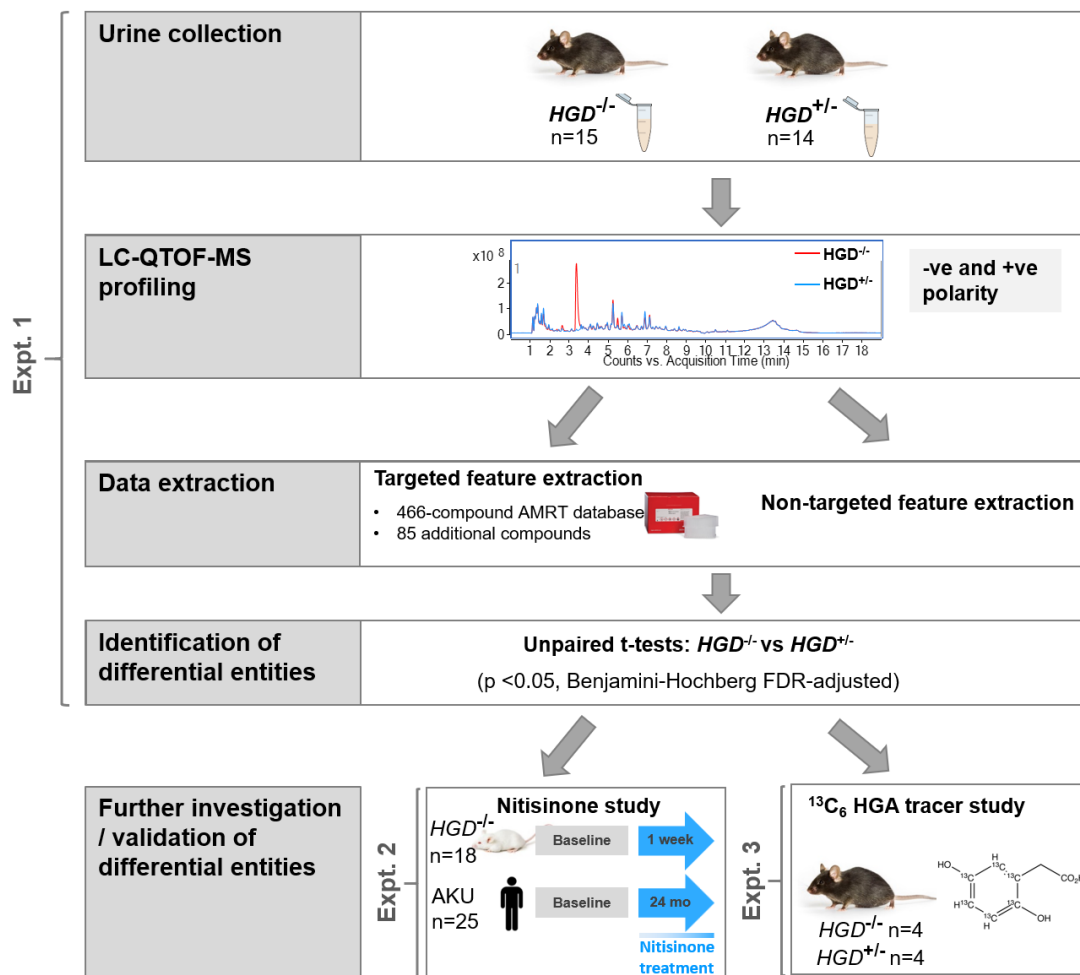
Figure 5.01. Schematic overview of the overall study design, incorporating Experiments 1-3. In Experiment 1, urine was collected from $HGD^{-/-}$ and $HGD^{+/-}$ mice and profiled by LC-QTOF-MS. Targeted and non-targeted feature extraction was performed on the data in parallel and subsequent unpaired t-tests were employed to identify differentially abundant chemical entities between $HGD^{-/-}$ and $HGD^{+/-}$. These entities were then further investigated in LC-QTOF-MS data from two additional datasets; a previously published study examining the effect of nitisinone on the urine metabolome of $HGD^{-/-}$ BALB/c mice and patients with AKU (Experiment 2, see Chapter 4 (303)) and a plasma flux analysis using a $^{13}C_6$ labelled HGA tracer (Experiment 3). These further studies were designed to add validation to the differentially abundant entities between $HGD^{-/-}$ and $HGD^{+/-}$ mice by reversal of their direction of alteration following nitisinone (Experiment 2), or

for entities that were increased in *HGD*$^{-/-}$ mice in Experiment 1, observation in the tracer study (Experiment 3) that they are directly derived from the injected labelled HGA.

5.2   Methods

5.2.1   Chemicals

All chemicals were as described for LC-QTOF-MS analysis in Chapter 2 (section 2.2.2).

5.2.2   Urine collection (Experiment 1)

For metabolomic analysis of the targeted *HGD* knockout phenotype (302), urine was collected from 15 *HGD*$^{-/-}$ (mean age ±SD 12.8 ±0.1 weeks) and 14 *HGD*$^{+/-}$ (mean age 11.6 ±0.3 weeks) male C57BL/6 mice. The ages of *HGD*$^{-/-}$ and *HGD*$^{+/-}$ mice at the time of sampling were purposely designed to be as equivalent as was pragmatically possible, however the age difference of approximately 1 week between these mice was statistically significant (unpaired t-test, $p < 0.0001$). Despite this difference in age, the measured weight of mice at the time of urine collection did not differ significantly between *HGD*$^{-/-}$ (mean weight 28.5 ±1.8 g) and *HGD*$^{+/-}$ (mean weight 28.1 ±0.8 g) groups (unpaired t-test, $p = 0.39$), indicating that both groups of mice had reached equivalent skeletal maturity.

Mouse urine was collected and samples were stored as described in Chapter 4 (section 4.2.2). Samples and group pools for QC were also prepared as described in section 4.2.2. Mouse breeding and dosing was authorised under UK Home Office guidelines.

5.2.3 Investigating the effect of nitisinone on metabolites showing alteration in HGD$^{-/-}$ mice (Experiment 2)

The effect of nitisinone treatment on the abundance of the urinary metabolites shown to be altered in HGD$^{-/-}$ mice (Figure 5.01, Experiment 1) was studied in the raw profiling data from Chapter 4. These data were acquired under identical analytical conditions to those employed here. The data were from urine collected from HGD$^{-/-}$ mice at baseline then after 1 week on nitisinone (in drinking water at 4 mg/L) and from patients with AKU at baseline then after 24 months on nitisinone (2 mg daily). For these mice, HGD disruption was by ENU mutagenesis; the disease phenotypes of HGD$^{-/-}$ mice from targeted knockout and ENU mutagenesis models are identical (302).

5.2.4 Design of in vivo metabolic flux experiment and sample collection (Experiment 3)

Metabolic flux analysis was performed by injecting mice with a stable $^{13}C_6$-labelled HGA tracer and taking blood samples at various intervals over the course of 60 min. Eight C57BL/6 mice were used in this experiment; 4 HGD$^{-/-}$ (3 male, 1 female, mean age ±SD 56 ±2.3 weeks) and 4 HGD$^{+/-}$ (4 female, mean age ±SD 58 ±0 weeks).

A 1.96 mg/mL $^{13}C_6$ HGA tracer solution was prepared in sterile saline. Mice were kept anaesthetised throughout the experiment with isoflurane. $^{13}C_6$ HGA tracer solution was then injected into the tail vein of each mouse. Injection volume was adjusted for each mouse in order to achieve an approximate 1 mmol/L final blood

concentration of the tracer. This calculation was based on the estimated total blood volume of each mouse from its weight, based on the assumption of a total blood volume of 75 mL/Kg for a C57BL/6 mouse (304). Venous tail bleed samples were taken at intervals of 2, 5, 10, 20, 40 and 60 min (or as many of these time points as possible, according to LASA guidelines (304). Blood was collected into Microvette 300 µL lithium heparin capillary tubes (Sarstedt, Germany) and centrifuged at 1500 *g* for 10 min. 25 µL of plasma supernatant was removed and stored at -80 °C prior to analysis. Individual plasma samples were analysed following dilution by 1:9 plasma:deionised water. Protein precipitation was not performed due to sample volume constraints associated with taking consecutive samples from the same mice. In addition, plasma samples were analysed by method 2 only, in which the LC column used is less susceptible to blockage by proteins, due to the larger particle size (3 µm), compared with the columns used in methods 1 (1.8 µm) and 3 (1.7 µm).

5.2.5 LC-QTOF-MS analysis

All samples were analysed as described in Chapter 4 (section 4.2.3) using LC method 2 (Chapter 2, section 2.2.2) The analytical sequence was designed as described in Chapter 4 (section 2.4.4).

5.2.6 LC-QTOF-MS data processing and statistical analyses

Experiment 1: parallel targeted and non-targeted data mining

Parallel targeted and non-targeted feature extraction was performed on the data, according to the strategy outlined in Chapter 3.

Targeted feature extraction was used to extract signals matching the 466-compound AMRT database generated using parameters detailed in Chapter 2 (section 2.2.3.); AM (±10 ppm) and database RT (±0.3 min) matches were required. Eighty-five additional metabolites (Table S5.01, Supporting Material, Appendix 5) were appended to the database for targeted extraction as they were of potential interest in AKU or from wider tyrosine metabolism. Forty-three were from the following pathway databases available from Pathways to PCDL (build 07.00, Agilent, UK): 'citrate degradation', 'noradrenaline and adrenaline degradation' and 'superpathway of phenylalanine, tyrosine and tryptophan biosynthesis'. Six metabolites were added as they were predicted to show potential alteration in $HGD^{-/-}$ due to association with tyrosine conjugation (acetyl-L-tyrosine, and γ-glutamyl-tyrosine) based on the data presented in Chapter 4 (303) or predicted association with ochronotic pigment derived from HGA (2,5-dihydroxybenzaldehyde, di-dehydro-homogentisic acid, hipposudoric acid and norhipposudoric acid). Thirty-six metabolites were from a list of potential biotransformation products directly derived from HGA and compiled using the Biotransformation Mass Defects application (Agilent; see Supplementary Table S5.01, Appendix 5). This tool provides a list of potential metabolic biotransformation products covering both phase I (n=19) and II (n=17) metabolism for a given compound based on its empirical formula (Figure 5.02). The data were mined for these additional metabolites with putative identification by AM (± 5ppm) only.

Non-targeted feature extraction was performed on the data using the parameters described in Chapter 4 (section 4.2.5.2) for mouse urine.

Figure 5.02. Screenshot depicting application of the Biotransformation Mass Defects (Agilent) tool for the generation of accurate mass targets for theoretical biotransformation products derived from HGA. For a given compound empirical formula, the tool can be used to generate predicted products from phase I and II metabolism. For HGA, the example shown here (compound formula; $C_8H_8O_4$), 92 theoretically possible biotransformations were generated, 36 of which had predicted formulae and accurate mass targets; these 36 compound targets were appended to the compound database for targeted data mining.

Experiment 1: data processing, QC and statistical analyses

Creatinine normalisation and QC, using data from pooled samples, was performed as described in Chapter 4 (section 4.2.7). Statistically significant entities were then identified in each dataset by unpaired t-test, comparing $HGD^{-/-}$ vs $HGD^{+/-}$ mice with multiple testing correction performed using Benjamini-Hochberg FDR adjustment. FCs were calculated for differentially abundant entities, also after normalisation and QC; these were $\log_2$-transformed and based on peak area. FC cut-off was not applied for entities with FDR-adjusted p <0.05 however, in order to retain entities that showed consistent although relatively low magnitude abundance differences between comparison groups. PCA was performed on each filtered dataset using four-component models.

Experiment 2: data processing and statistical analyses

Entities with FDR-adjusted p<0.05 in Experiment 1 ($HGD^{-/-}$ vs $HGD^{+/-}$) were mined in data from Experiment 2 ($HGD^{-/-}$ mice at baseline then after 1 week on nitisinone). This was performed by creating a custom AMRT database containing only these compound targets, using PCDL Manager (build 08.00, Agilent). The dataset was mined for these compounds using targeted feature extraction, also in Profinder (build 08.00, Agilent) using the same parameters as for Experiment 1. The profiles of the extracted entities were compared between baseline and 1 week on nitisinone samples with paired t-tests (FDR adjusted) and $\log_2$-transformed FC as above.

Experiment 3: data processing and statistical analyses

Data from plasma samples obtained from Experiment 3 were mined using the batch isotopologue extraction function in Profinder (build 08.00). This algorithm uses an AMRT target database to extract features and their AM isotopologues to perform flux analysis. Here, extraction was performed using the same AMRT database employed in Experiment 2 comprising the differentially abundant metabolites between HGD$^{-/-}$ vs HGD$^{+/-}$ mice. Database hits were investigated for potential association with the $^{13}C_6$ HGA tracer by examining the relative abundances of the M+0 to M+6 isotopologues. Extraction was performed with AM and RT match windows of ±5 ppm and ±0.3 min respectively.

The results from isotopologue extraction on plasma $^{13}C_6$ HGA metabolic flux data were reviewed visually in Profinder for clear evidence of an isotope label likely to be derived from HGA. This data review was performed for compound matches individually, taking into account the predicted number of ^{13}C atoms derived from the HGA tracer based on the compound's chemical structure.

5.3   Results

5.3.1 Identification of differentially abundant urinary metabolites between

HGD$^{-/-}$ vs HGD$^{+/-}$ mice (Experiment 1)

PCA showed clear differences between the urine profiles of *HGD$^{-/-}$* and *HGD$^{+/-}$* mice with targeted (Figure 5.03A & B) and non-targeted feature extraction, with separation in principal component 1. Data from targeted and non-targeted

extraction are presented separately in the following sections. Table 5.01 details the number of entities extracted from each method and the number retained following QC.

Table 5.01. Number of entities retained before and after quality control (QC). QC filtering using pooled group sample data was performed in Mass Profiler Professional. **'**Targeted' and 'non-targeted' refer to feature extraction approaches. Entities from targeted feature extraction are AMRT-matched against a database generated from IROA Technology MS metabolite library of standards (466 compounds) or AM-matched against an appended list of compounds from wider tyrosine metabolism (85 metabolites). Entities from non-targeted feature extraction are unidentified.

| | $HGD^{-/-}$ vs $HGD^{+/-}$ | | | |
| | Targeted | | Non-targeted | |
QC filtering step	(+)	(-)	(+)	(-)
Manual curation of entities only	323	324	662	552
Present in 100% of replicate injections of at least one pooled group sample	323	324	569	525
CV <25% across replicate injections of each pooled group sample	243	250	213	359

Targeted feature extraction

Targeted feature extraction was performed to search for metabolites based on AMRT (AM ±10 ppm, RT ±0.3 min) or AM (±5 ppm) alone. 27/250 and 15/243 metabolites showed abundance differences (FDR-adjusted p <0.05) between $HGD^{-/-}$ and $HGD^{+/-}$ in negative and positive polarity respectively. Table 5.02 shows the altered metabolites ranked in order of FC. PCA loadings plots (Figure 5.03C & D) and volcano plots (Figure 5.03E & F) show that the greatest differences between $HGD^{-/-}$ and $HGD^{+/-}$ urine were in metabolites associated with HGA in negative and positive polarity. HGA and 8 predicted HGA biotransformation products were markedly elevated in $HGD^{-/-}$. Interestingly, HGA-sulfate (FC=9.3, p <0.0001) showed a greater FC increase than HGA (FC=7.9, p <0.0001). Six other HGA products were increased with FC >1.5 and p <0.0001: HGA-glucuronide, HGA-hydroxyl-sulfate, HGA-*N*-acetylcysteine, HGA-hydrate, HGA-(de-alkene)-epoxide and HGA-(de-ethyl)-carboxylate. Acetyl-HGA was elevated in $HGD^{-/-}$ (FC=2.3, p <0.0001) but did not pass QC filtering by CV <25% across replicate injections of pooled QC samples; a decrease in signal across the run indicated a sample stability issue (despite the autosampler being maintained at 4°C).

Twenty-six significantly altered entities, excluding the aforementioned HGA, were AMRT-matched with compounds from the 466-compound library developed in-house. The most significantly altered (p <0.05 and FC >1.5) of these were: thymidine-5'-diphospho-α-D-glucose, DL-3,4-dihydroxymandelic acid, 2-aminophenol, p-hydroxyphenylacetic acid (all increased in $HGD^{-/-}$), malic acid, isocitrate and inosine 5'-monophosphate (all decreased in $HGD^{-/-}$).
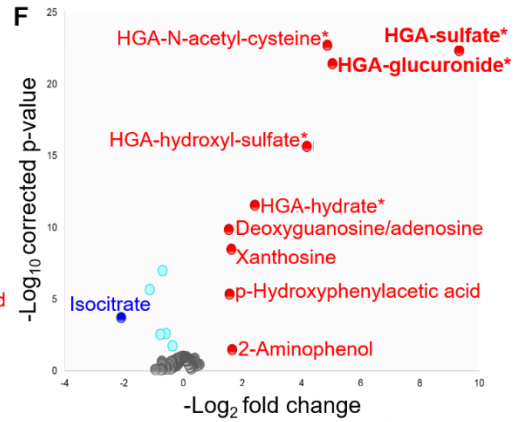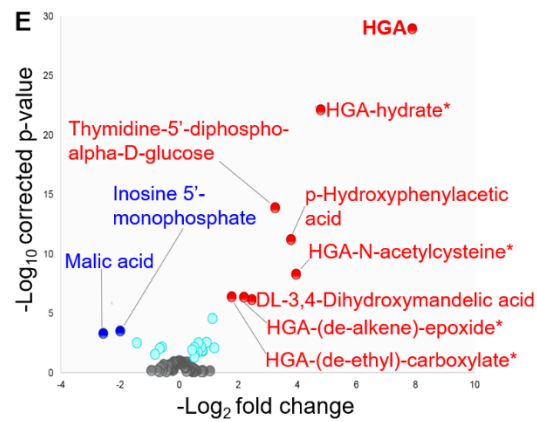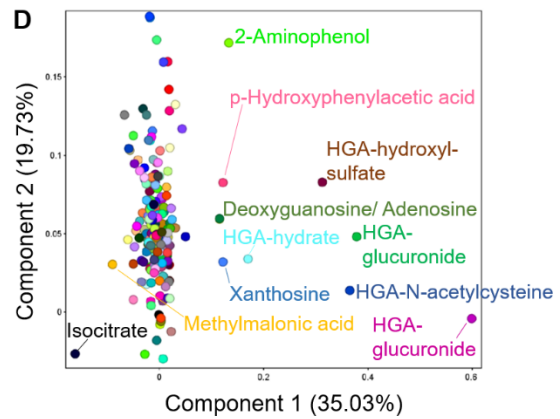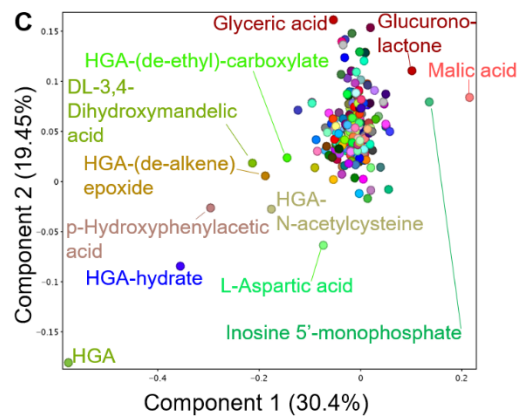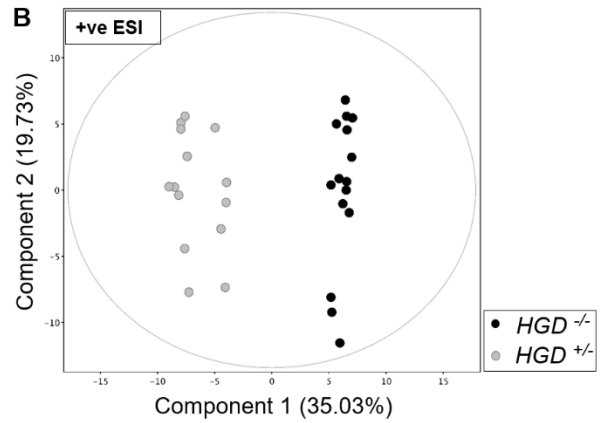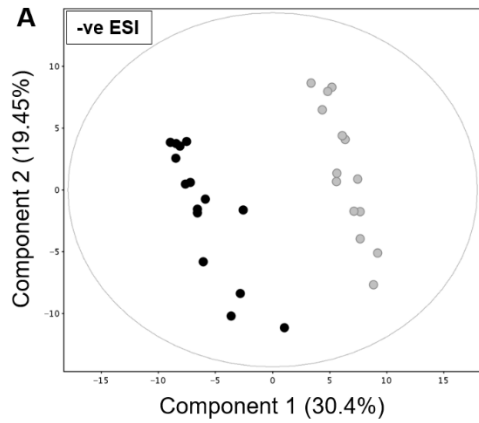
Figure 5.03. Clear differences between the urine metabolomes of HGD$^{-/-}$ and HGD$^{+/-}$ mice. A-D: principal components analysis (PCA) on data from targeted feature extraction, with PCA plots showing separation between *HGD*$^{-/-}$ and *HGD*$^{+/-}$ mice by component 1 in A, negative, and B, positive ionisation polarities. Lower plots show the corresponding PCA loadings of metabolites on components 1 and 2 in C, negative, and D, positive polarity. E-F: volcano plots illustrating selection of statistically significant urinary chemical entities between *HGD*$^{-/-}$ and *HGD*$^{+/-}$ mice based on p-value and fold change (FC). E, negative polarity; F, positive polarity. Entities with p <0.05 (Benjamini-Hochberg FDR adjusted) and FC >1.5 are labelled, with red and blue indicating increased and decreased abundance, respectively, in *HGD*$^{-/-}$. Turquoise indicates adjusted p <0.05 but FC <1.5.

Bold text indicates that the increase observed in HGD$^{-/-}$ was confirmed in mouse plasma following injection with $^{13}C_6$ HGA tracer.

* Compound not previously reported in the literature.

Table 5.02. Summary of identified metabolites or chemical entities showing altered abundance in *HGD$^{-/-}$* mice. Direction of alteration and fold change is indicated in *HGD$^{-/-}$* (relative to *HGD$^{+/-}$*) and following nitisinone treatment in *HGD$^{-/-}$* (paired data; on nitisinone relative to baseline). P-values are false discovery rate adjusted. Where entities were significantly different in positive and negative polarity, the result with the lowest fold change is provided.

* Compound identification by matched accurate mass (±10 ppm) and RT (±0.3 min) (otherwise, identification by accurate mass ±5 ppm only).

† Entity obtained from non-targeted data mining with putative identification by accurate mass (±10 ppm) only

‡ Note acetyl-HGA failed quality control filtering (CV >25% across replicate injections of QC pooled samples; due to suspected compound stability issue over analysis period).

Compound	Neutral mass	RT (min)	Direction of alteration in *HGD$^{-/-}$*	Fold change	p-value
HGA-sulfate	247.9996	2.6	↑	9.3	<0.0001
HGA*	168.0434	3.4	↑	7.9	<0.0001
HGA-glucuronide	344.075	2.4	↑	5	<0.0001
HGA-hydroxyl-sulfate	263.9931	2.8	↑	4.2	<0.0001
HGA-*N*-acetylcysteine	329.057	4.5	↑	4	<0.0001
Thymidine-5'-diphospho-alpha-D-glucose*	564.0758	2.4	↑	3.3	<0.0001
Unknown	268.9662	2.6	↑	2.8	<0.0001
Unknown	367.0289	3.4	↑	2.6	<0.0001

DL-3,4-Dihydroxymandelic acid*	184.037	1.9	↑	2.5	<0.0001
HGA-hydrate	186.053	3.4	↑	2.4	<0.0001
Unknown	507.141	3.4	↑	2.4	<0.0001
Acetyl-HGA[‡]	210.053	6.2	↑	2.3	<0.0001
Unknown	287.9939	6.0	↑	2.3	0.025
HGA-(de-alkene)-epoxide	184.0369	1.8	↑	2.2	<0.0001
2-Amino-6-(benzylthio)purine[†]	257.0758	3.5	↑	2.2	<0.0001
Unknown	257.0391	3.4	↑	2.0	<0.0001
HGA-(de-ethyl)-carboxylate	184.0017	3.4	↑	1.8	<0.0001
2-Aminophenol*	109.0534	1.9	↑	1.7	0.035
Unknown	294.0949	3.6	↑	1.7	<0.001
p-Hydroxyphenylacetic acid*	152.0478	5.5	↑	1.6	<0.0001
Unknown	361.0827	1.7	↑	1.5	<0.001
3-Dehydroshikimic acid*	172.0373	1.8	↑	1.2	<0.01
Hydroxymethyl-HGA	198.0526	4.8	↑	1.1	<0.0001
Aceturic acid*	117.0426	1.7	↑	0.94	<0.01
Nicotinic acid*	123.032	1.8	↑	0.8	0.015
4-Hydroxybenzaldehyde*	122.0368	5.6	↑	0.8	0.014

Vanillylmandelic acid*	198.0528	3.3	↑	0.75	<0.01
N-Acetyl-DL-serine*	147.0532	1.6	↑	0.66	<0.01
N-Acetylaspartate*	175.0481	1.8	↑	0.64	0.018
Xanthosine*	284.0762	3.4	↑	0.53	0.028
N-Alpha-acetyl-l-asparagine*	174.0641	1.6	↑	0.52	<0.01
N-Acetyl-L-phenylalanine*	207.0895	6.8	↑	0.52	0.049
3,5-Cyclic-AMP*	329.0525	3.4	↑	0.44	0.012
Malic acid*	134.0211	1.5	↓	2.6	<0.001
Isocitrate*	192.0278	1.7	↓	2.1	<0.001
Vanilyl pyruvic acid†	256.0582	5.0	↓	2.1	<0.001
Inosine 5'-monophosphate*	348.05	1.8	↓	2	<0.001
Unknown	242.0427	4.0	↓	1.7	<0.0001
Unknown	362.086	9.6	↓	1.6	<0.01
Glucuronolactone*	176.0321	1.6	↓	1.4	<0.01
Methylmalonic acid*	118.0266	2.0	↓	1.1	<0.0001
Monomethyl-glutaric acid*	146.0579	5.1	↓	0.82	0.029
N-Acetyl-L-glutamic acid*	189.0637	2.1	↓	0.76	<0.01
Citramalic acid*	148.039	2.6	↓	0.7	<0.0001

244

L-anserine*	240.1222	1.3	↓	0.59	<0.01
2,3-Cyclic AMP*	329.0525	2.7	↓	0.58	<0.01
O-Succinyl-L-homoserine*	219.0743	2.0	↓	0.35	0.019

Non-targeted feature extraction

Non-targeted feature extraction was performed using recursive feature extraction, resulting in 359 and 213 chemical entities post-QC in negative and positive polarity respectively; mass range = 54-3108 Da and RT range = 1.1-11.5 min. Comparison of *HGD$^{-/-}$* and *HGD$^{+/-}$* revealed entities with clear abundance differences (p <0.05, FC >1.5); 9 entities in negative polarity (6 increased in *HGA$^{-/-}$*, 3 decreased in *HGD$^{-/-}$*) and 2 entities in positive polarity (both increased in *HGD$^{-/-}$*). The mass range of these entities was 242-507 Da, and RT range was 1.7-9.6 min (Table 5.02). The chemical identities of these entities are unknown, but potential candidates were obtained for two entities based on AM match (<10 ppm) with the Masshunter METLIN metabolites PCD/PCDL accurate mass database (build 07.00); 2-amino-6-(benzylthio)purine (increased in *HGD$^{-/-}$*) and vanilyl pyruvic acid (decreased in *HGD$^{-/-}$*). The METLIN database used contains 30,232 compound entries and was accessed through PCDL Manager (build 08.00, Agilent, UK).

5.3.2 The effect of nitisinone treatment on chemical entities altered in HGD$^{-/-}$ mice

To investigate the effect of nitisinone treatment on the chemical entities shown to be altered here in *HGD$^{-/-}$*, they were searched in the data from previous mouse and human urine profiling experiments (Figure 5.01, Experiment 2). These data were acquired from *HGD$^{-/-}$* mice at baseline then after 1 week on nitisinone, and from patients with AKU at baseline then after 24 months on nitisinone. Entities

were extracted by targeted molecular feature extraction, with AM (±10 ppm) and RT (±0.3 min) matches required. Targeted data extraction and subsequent QC parameters were as described above.

Eight entities from targeted feature extraction that were altered in $HGD^{-/-}$ vs $HGD^{+/-}$ mice were significantly altered in the opposite direction in urine from both $HGD^{-/-}$ mice and patients with AKU following nitisinone (based on peak area; baseline vs on nitisinone, Benjamini-Hochberg FDR-adjusted p <0.05; Table 5.03); 7 decreased, and 1 increased. The 7 decreased metabolites were HGA, the HGA biotransformation products HGA-sulfate, HGA-glucuronide, HGA-hydrate and hydroxymethyl-HGA, and also xanthosine and 3,5-cyclic-AMP. The increased metabolite was monomethyl-glutaric acid. Interestingly, p-hydroxyphenylacetic acid and 4-hydroxybenzaldehyde were increased both in $HGD^{-/-}$ vs $HGD^{+/-}$ and also following nitisinone in $HGD^{-/-}$ mice and patients with AKU.

In addition, 19 entities that were altered in $HGD^{-/-}$ vs $HGD^{+/-}$ mice were significantly altered in the opposite direction following nitisinone in urine from either $HGD^{-/-}$ mice (n=14; 9 decreased, 5 increased on nitisinone) or patients (n=5; 4 decreased, 1 increased on nitisinone) with AKU only (Table 5.03). These compounds comprised the remaining HGA biotransformation products, with the exception of HGA-(de-ethyl)-carboxylate, which were all decreased following nitisinone. Following nitisinone, acetyl-HGA was decreased in human urine only, and HGA-hydroxyl-sulfate, HGA-*N*-acetylcysteine and HGA-(de-alkene)-epoxide were decreased in mouse urine only.

Table 5.03. The effect of nitisinone treatment in *HGD^{-/-}* mice and patients with AKU on the abundance of urinary chemical entities altered in *HGD^{-/-}* vs *HGD^{+/-}* mice. The entities that showed differences between *HGD^{-/-}* vs *HGD^{+/-}* mice (Table 5.02) were examined in two additional datasets. Paired t-tests were employed to compare the abundances at baseline *versus* 1 week on nitisinone (4 mg/L, in drinking water) for *HGD^{-/-}* mice and 24 months on 2mg daily nitisinone for patients with AKU. Only entities with false discovery rate adjusted p<0.05 baseline vs on nitisinone in mouse or human are displayed. Direction of alteration and fold change are indicated. Where entities were significantly different in positive and negative polarity, the result with the lowest fold change is provided. Note: no fold change indicated for hydroxymethyl-HGA in humans, as this compound was not detected for any patient on nitisinone.

* Compound identification by matched accurate mass (±10ppm) and RT (±0.3min) (otherwise, identification by accurate mass ±5ppm only).

† Entity obtained from non-targeted data mining with putative identification by accurate mass (±10ppm) only

ND: compound not detected at baseline or on nitisinone.

Compound	Neutral mass	RT (min)	*HGD^{-/-}* mouse urine			Human urine		
			Alteration 1 week on nitisinone	Fold change	p-value	Alteration 24 months on nitisinone	Fold change	p-value
HGA-glucuronide	344.075	2.4	↓	2.4	<0.0001	↓	2.4	<0.0001
HGA-(de-alkene)-epoxide	184.0369	1.8	↓	2.3	<0.0001	-		
HGA-hydroxyl-sulfate	263.9931	2.8	↓	2.3	<0.0001	ND		
2-Aminophenol*	109.0534	1.9	↓	2.2	<0.001	-		
HGA-sulfate	247.9996	2.6	↓	2.2	<0.0001	↓	3.5	<0.0001

Metabolite	m/z	RT						
Unknown	268.9662	2.6	↓	2.0	<0.0001	ND		
DL-3,4-Dihydroxymandelic acid*	184.037	1.9	↓	1.6	<0.0001	-		
HGA*	168.0434	3.4	↓	1.6	<0.0001	↓	3.4	<0.0001
Hydroxymethyl-HGA	198.0526	4.8	↓	1.1	<0.001	↓	-	<0.0001
Aceturic acid*	117.0426	1.7	↓	1.1	<0.0001	-		
Xanthosine*	284.0762	3.4	↓	1	<0.0001	↓	2.2	<0.0001
HGA-hydrate	186.053	3.4	↓	0.8	<0.0001	↓	3.8	<0.0001
HGA-N-acetylcysteine	329.057	4.5	↓	0.7	<0.001	-		
N-Acetyl-DL-serine*	147.0532	1.6	↓	0.5	0.031	-		
Nicotinic acid*	123.032	1.8	↓	0.4	0.014	-		
3,5-Cyclic-AMP*	329.0525	3.4	↓	0.3	0.041	↓	2	<0.0001
Acetyl-HGA	210.053	6.2	-			↓	2.7	<0.0001
Vanillylmandelic acid*	198.0528	3.3	-			↓	1	<0.01
2-Amino-6-(benzylthio)purine†	257.0758	3.5	-			↓	0.9	0.035
N-Alpha-acetyl-l-asparagine*	174.0641	1.6	-			↓	0.4	<0.01
Malic acid*	134.0211	1.5	ND			↑	0.8	0.028
p-Hydroxyphenylacetic acid*	152.0478	5.5	↑	3.3	<0.0001	↑	3	<0.0001

4-Hydroxybenzaldehyde*	122.0368	5.6	↑	2.32	<0.0001	↑	6.6	<0.0001
N-Acetyl-L-phenylalanine*	207.0895	6.8	↑	1.8	<0.0001	-		
Unknown	294.0949	3.6	↑	1.8	<0.0001	↓	0.6	<0.001
Unknown	362.086	9.6	↑	1.4	<0.0001	ND		
Isocitrate*	192.0278	1.7	↑	1.2	<0.01	-		
N-Acetylaspartate*	175.0481	1.8	↑	0.9	0.023	-		
Citramalic acid*	148.039	2.6	↑	0.8	0.023	-		
N-Acetyl-L-glutamic acid*	189.064	2.1	↑	0.8	<0.0001	-		
Monomethyl-glutaric acid*	146.0579	5.1	↑	0.4	<0.01	↑	1.7	0.041

5.3.3. Confirmation of HGA biotransformation products by $^{13}C_6$ HGA metabolic flux analysis (Experiment 3)

Data from isotopologue extraction were compared between plasma collected from the same mice across the time intervals available (2-60 min). The M+6 isotopologue was of particular interest as mice were injected with $^{13}C_6$-labelled HGA (although M+ 0-5 isotopologues were also considered for compounds where the labelled ^{13}C carbons of the benzene ring could potentially be removed through metabolism of HGA).

The M+6 peak was clearly observable for HGA over the sampling time course in *HGD$^{-/-}$* and *HGD$^{+/-}$* mice, although the signal decreased considerably from 10-20 min post-injection (Figure 5.04). As indicated in Figure 5.03 (compounds in bold text), 2 HGA biotransformation products that were increased in *HGD$^{-/-}$* urine were observed with clear M+6 peaks in plasma over the time course; HGA-glucuronide and HGA-sulfate. These data confirm that the compounds are derived from HGA. HGA-sulfate showed a similar time course profile to HGA. For this compound the native M+0 isotopologues were also absent from *HGD$^{+/-}$* plasma at all time points, in contrast to the M+6 isotopologue whose profile appeared to closely follow that of the HGA M+6 peak over the time course in both *HGD$^{-/-}$* and *HGD$^{+/-}$*. For HGA-glucuronide, the native M+0 isotopologue was also absent from *HGD$^{+/-}$* plasma, but the M+6 isotopologue was only observed in *HGD$^{-/-}$*, indicating that glucuronidation of the HGA tracer only occurred in *HGD$^{-/-}$* mice (Figure 5.04).
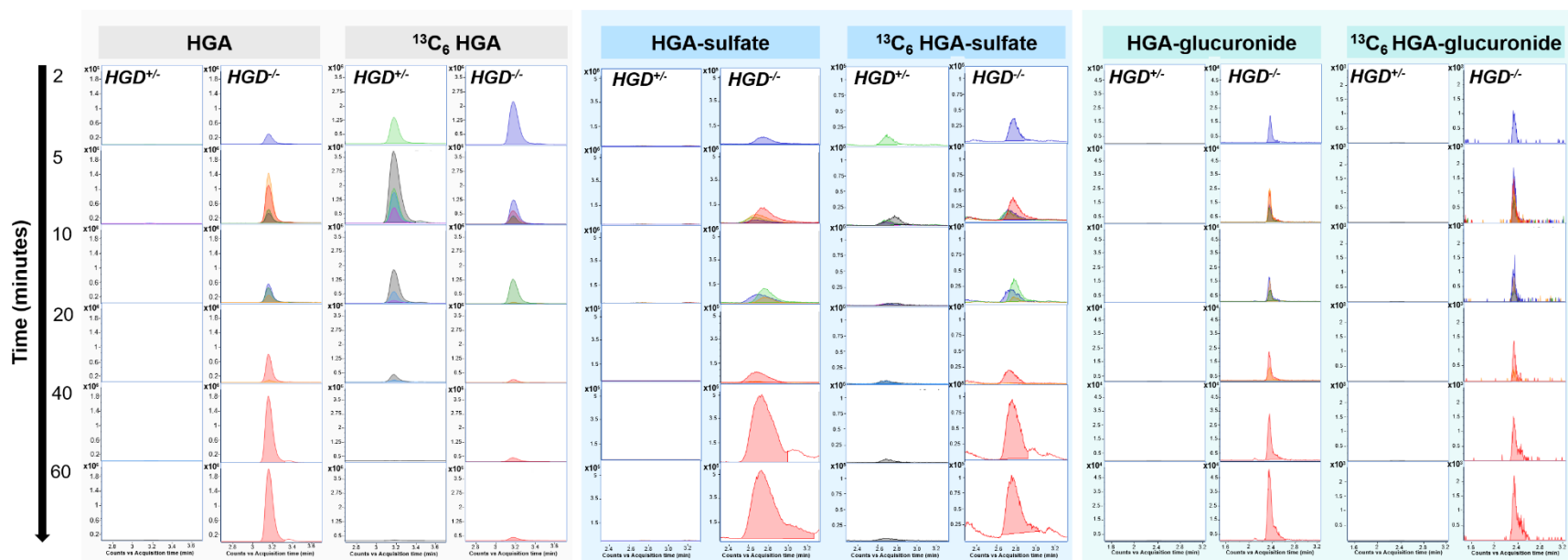
Figure 5.04. Isotopologue extraction results on plasma from the *in vivo* metabolic flux experiment using injected $^{13}C_6$-labelled homogentisic acid (HGA). Data shown are from *HGD$^{-/-}$* and *HGD$^{+/-}$* samples taken at intervals of 2, 5, 10, 20, 40 and 60 min (when possible) after injection. Extracted ion chromatograms (EICs) represent the M+0 (native compound) and M+6 ($^{13}C_6$-labelled form) isotopologue signals for homogentisic acid (HGA), HGA-sulfate and HGA-glucuronide. EICs show clear M+6 peaks for these compounds following injection (but only from *HGD$^{-/-}$* mice for HGA-glucuronide), confirming that they are derived from the labelled HGA.

5.4   Discussion

AKU holds a unique place in history as the first disease shown to follow Mendelian inheritance (305). The genetic cause of AKU has therefore long been understood. Mutations in a single gene, *HGD*, have the clear biochemical consequence of elevated HGA which is deposited in tissues and forms ochronotic pigment; the central pathophysiological event that causes the morbidity observed in AKU (16). Despite this archetypal genotype-phenotype relationship, the data reported here from the first metabolome-wide comparison of non-treated AKU vs non-AKU ($HGD^{-/-}$ vs $HGD^{+/-}$ mice; Experiment 1) reveal that the biochemical consequences of HGD-deficiency are wider-reaching than previously recognised and extend beyond tyrosine metabolism. The data also reveal that previously uncharacterised phase I and II metabolic processes are recruited as detoxification mechanisms in AKU for clearance of the markedly increased HGA. The observation of a reversed direction of alteration for a number of these urinary metabolites following nitisinone treatment in mice and patients with AKU (Experiment 2) indicates that these changes observed in the mouse model of AKU also apply to the human form of the disease.

5.4.1 Phase I and II metabolic biotransformation processes are recruited for clearance of HGA in AKU

The clearest differences between the urine metabolomes of $HGD^{-/-}$ *versus* $HGD^{+/-}$ mice were associated with increased HGA; HGA and 8 previously unreported HGA-derived biotransformation products were increased in $HGD^{-/-}$ (FDR-adjusted

p<0.0001). The HGA-derived products result from a range of phase I and II metabolic processes and represent alternative metabolic routes that appear to be activated at the supraphysiological circulating concentrations of HGA observed in AKU. Figure 5.05 shows the predicted structures for 7 of these novel compounds and the enzyme families responsible for their production. The observation of reversed alteration (decreases) in HGA-sulfate, HGA-glucuronide, HGA-hydrate, acetyl-HGA and hydroxymethyl-HGA following nitisinone treatment in urine from patients with AKU (Experiment 2) adds validation and indicates that these newly-identified HGA biotransformation processes are also recruited in human AKU. Observation of $^{13}C_6$-labelled forms of HGA-glucuronide and HGA-sulfate in plasma following intravenous injection of $^{13}C_6$-HGA in mice (Experiment 3), giving final blood concentrations up to 1 mmol/L, confirmed that these products are directly derived from HGA. $^{13}C_6$-HGA-sulfate was observed for $HGD^{-/-}$ and $HGD^{+/-}$ mice and followed the profile of the $^{13}C_6$-HGA across the sampling time course. Interestingly, $^{13}C_6$-HGA-glucuronide was only observed in $HGD^{-/-}$, suggesting that glucuronidation is an additional reaction recruited only in HGD deficiency and perhaps reflective of prior upregulation of glucuronyltransferase activity for HGA clearance as part of the pre-existing AKU phenotype. HGA-glucuronide (FC=5) and HGA-sulfate (FC=9.3), together with HGA (FC=7.9), showed the greatest overall increases in $HGD^{-/-}$ versus $HGD^{+/-}$. The magnitude of these alterations are of potential high clinical importance in AKU both in monitoring the disease and in the development of rational disease management strategies in AKU, as they reveal for the first time the specific metabolic processes that are major

mechanisms for clearance and excretion of the elevated HGA in $HGD^{-/-}$ mice and patients with AKU. HGA is the culprit molecule in AKU which causes the pathological manifestations of the disease. HGA forms ochronotic pigment, purportedly through its reactive oxidation products, which binds to collagenous tissues and modifies the extracellular matrix (15). Ochronosis of articular cartilage causes it to become stiff and brittle, leading to the mechanical failure that underlies the inevitable and severe osteoarthropathy observed in AKU (306). Oxidative stress is also considered a major consequence of ochronosis. The benzoquinone compounds that are intermediates in ochronotic pigment formation are thought to be highly reactive species capable of inducing oxidative changes to proteins and lipids (41,307). These oxidative alterations are likely to perpetuate the chronic inflammation present in AKU as a result of ochronotic osteoarthropathy and could account for cases of acute fatal metabolic consequences reported in the literature (21). I propose that the phase I and II biotransformations identified are detoxification mechanisms to render HGA chemically inert in order to prevent its participation in these pathological processes.

The HGA biotransformation data further current knowledge of metabolism in AKU considerably, as very little exists in the literature about HGA clearance beside its conversion to maleylacetoacetic acid by the HGD enzyme in the traditional tyrosine degradation pathway. Phase I and II pathways are traditionally considered as routes for drug metabolism with their enzymes primarily located in the liver; and within close proximity to enzymes of the major tyrosine pathway

(174). Phase I metabolism comprises reactions to introduce reactive or polar groups to substrates and includes hydroxylation, oxidation, reduction and hydrolysis. Phase II metabolism involves conjugation reactions to form glucuronide, glutathione, mercapturic acid, amino acid, methyl and acetyl conjugates (308,309). There is heterogeneity in the reactions required for metabolism of different drugs/compounds; phase I and II reactions may occur sequentially, or some compounds may only undergo phase I or II metabolism (310,311). Clearance of HGA in AKU by alternative phase I and II metabolic reactions has been predicted previously by reference to acetylsalicylic acid (aspirin), a structurally analogous weak organic acid whose metabolism is well-characterised (174). The present data confirm these predictions for the first time by revealing that in $HGD^{-/-}$ mice a range of phase I and II mechanisms, alone and in combination, are recruited for HGA clearance. The phase I metabolites observed were formed via hydration (HGA-hydrate), alkene epoxidation (HGA-(de-alkene)-epoxide) and ethyl to carboxylate (HGA-(de-ethyl)-carboxylate) transformations. Phase II metabolite products were formed via conjugation with glucuronic acid (HGA-glucuronide), *N*-acetylcysteine (HGA-*N*-acetylcysteine desaturate) and sulfate (HGA-sulfate) and acetyl (acetyl-HGA) groups. Sequential phase I and II reactions were hydroxylation with subsequent methylation (hydroxymethyl-HGA) or sulfation (HGA-hydroxysulfate).
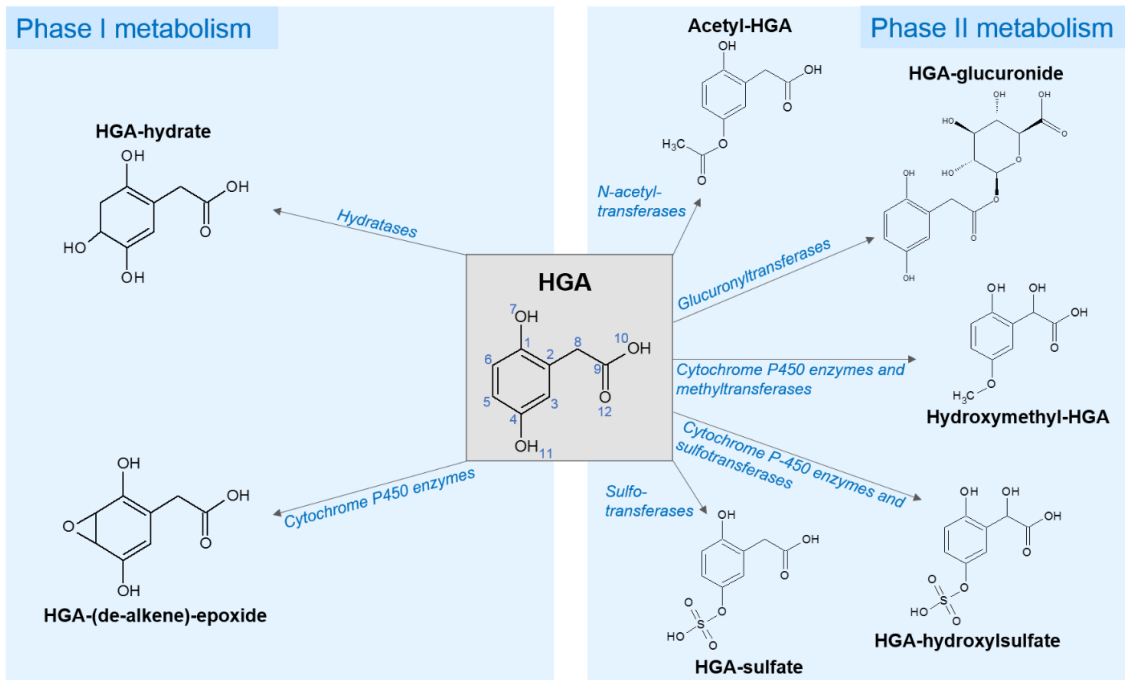
Figure 5.05. Predicted structures of newly-identified homogentisic acid (HGA) clearance products resulting from phase I and II metabolism. The proposed sites for metabolism/conjugation were those with the greatest probability from *in silico* prediction tools Stardrop (Optibrium, UK) and Reacting Atom (Way2Drug, Russia), for phase I and II biotransformations respectively. Note: HGA positions labelled 7 and 11 have equal probability for sulfation (deltaP = 0.8).

The phase I and II processes identified are catalysed by enzymes belonging to supergene families that can act on a broad range of substrates. For example, glucuronidation is quantitatively one of the most important phase II biotransformation reactions and is known to be carried out by 15 UPD-glucuronosyltransferase enzymes in humans (312,313). Identification of the specific enzymes within these families that catalyse these HGA biotransformations could inform future therapeutic interventions aimed at enhancing their activity in AKU. However, potential off-target effects of such

approaches should be considered given the range of substrates that enzymes of phase I and II metabolism can act upon. It is also worth highlighting that not all products of phase I and II metabolism reflect adaptive clearance mechanisms, as some biotransformation reactions produce reactive metabolites. Such 'bioactivation' reactions increase the risk of adverse effects and are therefore a major concern in drug design (314,315). In relation to the present data, HGA-(de-alkene)-epoxide is likely a product of HGA bioactivation, as epoxide metabolites are generally reactive, electrophilic compounds that require further enzymatically-mediated detoxification (316,317).

5.4.2 Associated alteration to tyrosine, purine and TCA cycle metabolism in AKU

Previous metabolomic studies in AKU have focused solely on the impact of nitisinone on the metabolome (220,234,235,303,318) (52), which was also the main focus of Chapter 4. Nitisinone reversibly inhibits HPPD, the enzyme responsible for production of HGA, and it has emerged as a promising therapeutic agent in AKU. Nitisinone reduces plasma and urine HGA concentrations (11,12,14,18,219), completely arrests ochronosis in AKU mice (50,69) and more recently was shown to slow progression of morbidity in AKU patients with access to nitisinone (2mg daily dose) through the NAC in the UK (140). LC-QTOF-MS metabolomics employing the AMRT identification strategy also used here showed for the first time that nitisinone treatment resulted in a host of previously undescribed changes in urine (303) and serum (234) metabolites beyond the tyrosine pathway in patients and mice with AKU. On nitisinone, a complex pattern

of alteration was observed in metabolites of tyrosine, tryptophan, purine and TCA cycle pathways. These findings are of clinical importance in AKU and also HT-1, in which nitisinone is licensed, as there are concerns surrounding the wider metabolic fate of the markedly increased tyrosine levels observed in patients on nitisinone (46).

This study explored the metabolome-wide consequences of HGD deficiency for the first time, revealing that in untreated $HGD^{-/-}$ vs $HGD^{+/-}$ mice there is alteration to tyrosine, purine and TCA cycle urinary metabolites (Figure 5.06). These data add to a growing body of evidence from metabolomic studies that perturbations in tyrosine, purine and energy metabolism are linked in numerous disease states. Two CSF metabolic profiling studies in Alzheimer's disease report an overlapping network of perturbations in tyrosine, tryptophan, purine and tocopherol pathways (270,271). LC-MS metabolomics also revealed altered phenylalanine, tyrosine, purine, TCA cycle, energy and lipid metabolism in serotonin-deficient (tryptophan hydroxylase knockout) mice (319): serotonin deficiency is associated with numerous neuropathologies including Alzheimer's and Parkinson's diseases, schizophrenia and depression. Associated network-level alteration to tyrosine, phenylalanine, purine, pyrimidine and energy metabolism was also reported more recently in a metabolomic study of urine from rats with diabetic nephropathy (261).
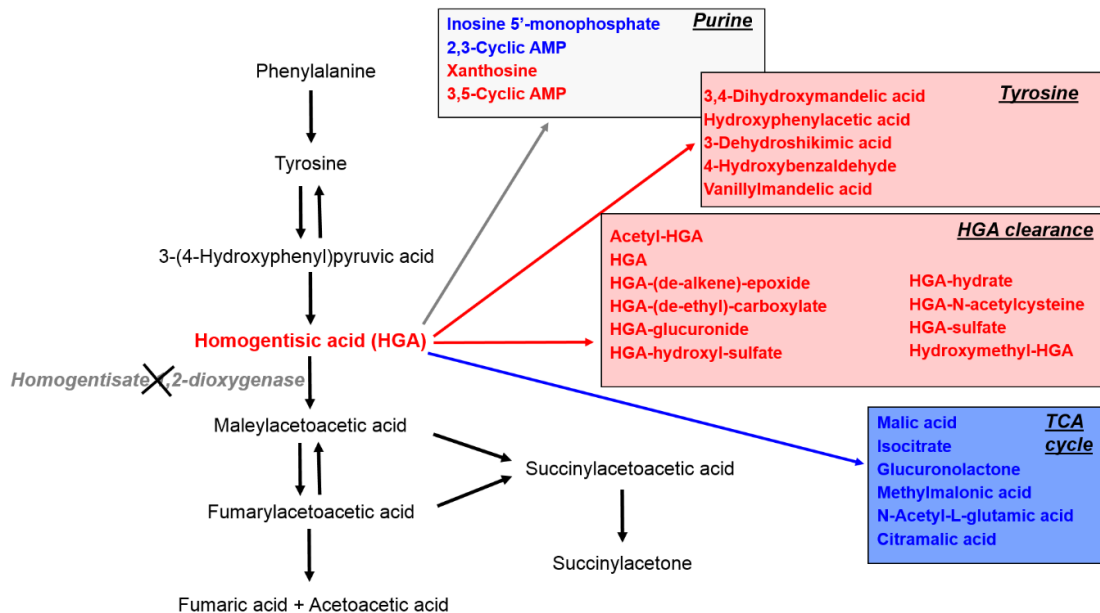
Figure 5.06. Summary of metabolites altered in HGD$^{-/-}$ mouse urine grouped by their associated pathways. Left: the tyrosine degradation pathway showing lack of the enzyme HGD in AKU and the consequential increase in homogentisic acid (HGA). Right (boxes): observed metabolite alterations grouped by pathway; red and blue indicate increased and decreased abundance respectively. Tyrosine metabolites including HGA, and HGA clearance products from phase I and II metabolism were elevated. Metabolites associated with the TCA cycle were decreased. A combination of increased and decreased abundance was observed for purine pathway metabolites.

Alteration to tyrosine metabolism

The observed increases in tyrosine metabolites (excluding HGA-associated clearance products) were not expected in untreated AKU. Increases in tyrosine metabolites previously observed in nitisinone-treated AKU were thought to be a direct consequence of the markedly elevated tyrosine induced by inhibition of HPPD (234,303). Increased p-hydroxyphenylacetic acid and 4-hydroxybenzaldehyde were observed previously in serum (234) and urine (303).

The biological significance of these changes is unknown. p-Hydroxyphenylacetic acid is generated from gut microbiota and its increase following nitisinone was proposed to reflect enhanced microbial metabolism due to reduced oxidative stress upon commencement of treatment (234). 4-Hydroxybenzaldehyde is a naturally occurring metabolite thought to have various therapeutic benefits in numerous diseases (320). Benzaldehydes may also relate to the proposed benzoquinone intermediates involved in the formation of ochronotic pigment from HGA, but the reason for the further increase in 4-hydroxybenzaldehyde following nitisinone is inconsistent with this and requires explanation.

DL-3,4-Dihydroxymandelic acid and vanillylmandelic acid (VMA) were also increased in $HGD^{-/-}$ mice; these are catecholamine metabolites derived from noradrenaline. Noradrenaline is derived from tyrosine via L-DOPA and dopamine, and it can be metabolised to DL-3,4-dihydroxymandelic acid (via DHMA) then to VMA or alternatively to adrenaline or NMA. Increased DL-3,4-dihydroxymandelic acid and VMA indicate increased metabolism of noradrenaline via DHMA as opposed to adrenaline or NMA in untreated AKU, which is also supported by a previous observation of decreased urinary NMA in patients with AKU following nitisinone (235). Neuroblastomas secrete catecholamine and MA metabolites, and the elevation of these compounds in urine forms part of the primary diagnostic criteria (321,322). Increased urinary output of DHMA and VMA is well-recognised in neuroblastoma, although the main markers of interest are usually increased urinary 3-MT, dopamine and HVA.

Together, it is not clear how these unexpected changes to tyrosine and peripheral

neurotransmitter metabolism are related to the increased HGA observed in untreated AKU. Previous metabolomics studies, including the data reported in Chapter 4, found similar alteration to peripheral neurotransmitter metabolism in patients with AKU on nitisinone treatment (234,235,303), and this was thought to result from the well described hyper-tyrosinaemia observed following nitisinone therapy through its action on the HPPD enzyme. However, these studies did not include a non-AKU control group, preventing comparison of the untreated AKU metabolome with that of healthy individuals. The data reported here reveal that even in untreated AKU, HGD deficiency leads to accumulation of tyrosine pathway metabolites upstream of HGA. The mechanism for altered peripheral neurotransmitter metabolism in untreated AKU is unknown. One could speculate that the alteration relates to the increased HGA or an unknown feature of the disease. For example, it is conceivable that HGA could act on other enzymes of neurotransmitter metabolites and therefore alter their activity. Further research is required to explore this potential mechanism. It is also important to determine whether the changes observed in peripheral neurotransmitter metabolism are a true reflection of altered CNS homeostasis. Recent mass spectrometric mouse brain imaging analyses support that in nitisinone-treated AKU, CNS metabolic alteration is limited to tyrosine and tyramine (52). Further analyses of CNS tissues and biofluids (CSF) are required to provide a clearer picture of neurotransmitter metabolism in untreated AKU.

Alteration to purine metabolism

In nitisinone-treated AKU the previously reported purine metabolite changes were generally decreased concentrations. Following nitisinone, 3-ureidoproprionate, adenine and allantoin were decreased in human urine, and 3,5-cyclic AMP and xanthosine were decreased in human and mouse urine (303). The present data indicate a more complex pattern of alteration to purine metabolism in non-treated AKU, with increased 3,5-cyclic AMP and xanthosine, and decreased 2,3-cyclic AMP and inosine 5'-monophosphate.

The link between altered tyrosine metabolism, as observed in nitisinone-treated and untreated AKU, and purine metabolism is unclear. The alteration in purine metabolism perhaps most associated with pathology is elevated uric acid, the classic example being gout, in which elevated blood uric acid leads to formation of crystals in and around joints, causing painful arthritis. Increased urinary uric acid is also a well-known risk factor for renal dysfunction, hypertension, hyperlipidemia, diabetes and obesity (323). Purine catabolism is very important in the homeostatic response to various states of mitochondrial oxidative stress, with shifts occurring to favour breakdown to xanthine and uric acid (324,325). The increased xanthosine and decreased 2,3-cyclic AMP and inosine 5'-monophosphate reported here for $HGD^{-/-}$ mice is consistent with a shift in purine catabolism in response to oxidative conditions, as xanthosine is a direct precursor for xanthine and uric acid. In support of this interpretation, xanthosine is reportedly increased in urine from rats with the inflammatory conditions chronic kidney disease (260) and diabetes nephropathy, and decreased markedly following

administration of nephroprotective therapy (261). It is however, unclear why concomitant increases were not observed for xanthine and uric acid in this study. There is equilibrium between xanthosine and xanthine, but the explanation of an apparent equilibrium shift in favour of xanthosine specifically remains unknown. It is also unclear why accumulation of xanthosine but not allantoin was observed in $HGD^{-/-}$, as allantoin is also derived from xanthine, and formed spontaneously from uric acid under conditions of oxidative stress (see Chapter 4, Figure 4.29).

As discussed in Chapter 4, nitisinone treatment led to decreases in urinary concentrations of the purine metabolites 3,5-cyclic AMP, 3-ureidopropionate, adenine, allantoin and xanthosine. The increased xanthosine in untreated $HGD^{-/-}$ here supports the interpretation that nitisinone alleviates the shift in purine metabolism towards xanthine that is a consequence of the oxidative conditions associated with AKU pathology. Oxidative stress is considered a major pathological feature of AKU that is a consequence of elevated HGA and ochronotic osteoarthropathy (16,41). Significant oxidative modification has been observed in AKU for a range of serum and chondrocyte proteins in addition to lipid peroxidation and depletion of endogenous anti-oxidants (41). These oxidative conditions are thought to induce further inflammation and to initiate a vicious circle of disease in AKU, in which oxidative stress leads to further ochronosis and tissue destruction, which in turn stimulates further inflammation and oxidation.

Alteration to TCA cycle metabolism

Decreases in TCA-related metabolites in *HGD*[-/-] mice is the first indication of perturbed energy metabolism in untreated AKU. Citramalic acid, isocitrate and *N*-acetyl-L-glutamic acid were also increased following nitisinone in *HGD*[-/-], suggesting for the first time that nitisinone is at least partially restorative. Malic acid and isocitrate are major intermediates of the TCA cycle; the primary energy-producing pathway of the cell. Citramalic acid and *N*-acetyl-L-glutamic acid are more peripheral to the TCA cycle but have been included here as they are derived from succinic acid and glutamine respectively. Blood and urine levels of TCA metabolites are generally considered to directly reflect overall TCA cycle activity (326). Inhibited TCA cycle activity could be due to overall mitochondrial biogenesis, decreased expression of genes encoding TCA cycle enzymes or reduced substrate availability. The latter explanation seems the most likely in AKU, in which HGD deficiency prevents further metabolism of HGA to fumaric acid, an intermediate of the TCA cycle. Decreased bioavailability of fumaric acid may then explain the decreases observed for other TCA cycle metabolites. Previous analyses have investigated the serum levels of TCA cycle metabolites in AKU, but these were limited to studying the effect of nitisinone. In contrast to the present urinary data, the TCA metabolites succinic acid and α-ketoglutaric acid were decreased in serum from patients with AKU following nitisinone (234). The significance of these decreases is not known, and further studies should investigate whether the alteration to urinary TCA cycle metabolites reported here also applies to serum in untreated AKU.

Decreased urinary citrate (precursor for and structural isomer of isocitrate) is a well-known risk factor for kidney stone formation. It is possible that the decreased isocitrate reported here for $HGD^{-/-}$ mice is associated with the increased risk of kidney stones in AKU (327). In this analysis isocitrate was the closest match obtained for this particular entity (RT = 1.7 min) based on its AMRT in our database generated from standards (RT = 1.6 min), but citric acid is also a potential match, with the same AM but slightly different database RT (1.9 min). It is important to note that the database RTs were obtained from standards analysed in blank matrix, therefore a degree of RT shift can be expected in urine matrix due to effects such as pH. Hypocitratururia is generally defined as citric acid excretion <320 mg (1.67 mmol) per day in adults (328). The present profiling data are semi-quantitative, therefore targeted quantitative assays are required to determine exact urinary citric acid concentrations in AKU patients and the additional therapeutic value of nitisinone in this regard.

5.4.3 Study strengths and limitations

A major challenge in studies aiming to characterise the metabolome-wide profile of a human disease is the selection of a suitable control group. The metabolome is extremely sensitive to genetic and environmental influences (93) which can bias the interpretation of classic 'case *versus* control' comparisons unless careful measures are taken to control them. Studying gender and age-matched genetically related mice housed in identical living conditions and with targeted disruption to the *HGD* gene (homozygous *versus* heterozygous knockout) limited potentially confounding genetic and environmental variation and increased

confidence that the metabolite alterations observed are associated with the AKU phenotype.

A limitation of urine metabolomics studies in mice is the increased difficulty of 24-h urine collection compared with human subjects, meaning that metabolite abundances obtained here could not be normalised to total creatinine output over this period to account for urine dilution. Attempts were made to overcome this by normalising metabolite signals to that obtained for creatinine in this analysis by matched AMRT against a standard, as described in Chapter 3 (section 3.4.1) (303). A second limitation in Experiment 3 was that the final predicted blood concentration of the injected $^{13}C_6$ HGA (1mmol/L) was in excess of HGA concentrations observed in AKU; 30 ±12.9 µmol/L in untreated AKU patients (12). This concentration was designed to overcome potential issues of analytical sensitivity associated with rapid metabolism of the tracer to downstream products. Further studies employing quantitative assays for the HGA clearance metabolites identified could more accurately assess the contribution of each of the phase I and II metabolic processes at HGA concentrations observed in AKU.

5.5   Chapter 5 conclusion

The metabolome-wide profiles of untreated AKU ($HGD^{-/-}$ mice) *versus* non-AKU ($HGD^{+/-}$ mice) were compared for the first time. The data indicate that targeted homozygous disruption to the HGD gene in mice, which causes a clear AKU phenotype, is associated with a host of previously unknown metabolite changes in a complex network of pathways associated with altered tyrosine metabolism.

Using biotransformation prediction tools it was also revealed for the first time that a range of phase I and II metabolic processes are important for clearance of the elevated HGA that is central to the pathological manifestation of AKU. Decreases in 5/8 HGA biotransformation products were also observed in urine from patients with AKU following nitisinone therapy for HGA reduction, indicating that these HGA clearance mechanisms are also recruited in the human disease. These data also have wider, clinically-significant implications beyond AKU by showing that phase I and II metabolic processes are recruited for detoxification of endogenous metabolites that can accumulate as a consequence of inherited metabolic disease. The findings in this chapter therefore highlight that it is a fundamental misconception to view these ancient, evolved processes solely as mechanisms for drug clearance.

Acknowledgements

6.0   STUDY OF HGA AND OCHRONOTIC PIGMENT FORMED *IN VITRO*

6.1   Introduction

Ochronotic pigment is derived from HGA and central to disease pathology in AKU. Elucidation of the chemical structure of ochronotic pigment and the mechanism by which it interacts with the cartilage matrix is fundamental to understanding the destructive process of ochronosis (329).

One of the simplest approaches to studying ochronotic pigment formation is from pure solutions of HGA *in vitro* in the absence of the complex macromolecular collagen matrix. It is well-known that pigment can be formed non-enzymatically via spontaneous autoxidation of HGA (150). As discussed in Chapter 1 (section 1.3), it has long been generally assumed in the literature that ochronotic pigment formation adheres to the classic scheme, whereby HGA is oxidised to the intermediary BQA, which then undergoes subsequent self-polymerisation to produce ochronotic pigment (Figure 1.01B). However, there is a lack evidence that ochronotic pigment is polymeric in proportion, and even less evidence that ochronotic pigment meets the classic definition of a polymer in that it comprises regularly repeating chemical subunits that are covalently bound.

In the following series of experiments, a range of NMR- and MS-based analytical techniques were employed to study pigment formed in the laboratory under controlled conditions from pure solutions of synthetic HGA.

6.2　Methods

6.2.1　Solutions of HGA, BQA and HGA-derived pigment

The following solutions were prepared from synthetic HGA (≥98% purity; Sigma-Aldrich, UK). All solutions were prepared using deionised water (purified in-house via DIRECT-Q 3UV Millipore water purification system) and in glass. A photograph of solutions 1-5 is shown in Figure 6.01.

1.　HGA (no alkalinisation)

2.　BQA (by nitrous acid method)

3.　HGA – to BQA (by autoxidation): pH 9.5

4.　HGA – to BQA (by autoxidation): pH 11

5.　Pigment from HGA aged for 4 months at 37°C (referred to as 'HGA-derived pigment' from here onwards)

First, 10 mL of a 100 mmol/L stock solution of HGA was prepared from HGA powder. Solution 1 was prepared by 1:49 dilution of the HGA stock in water, resulting in a 2 mmol/L solution of HGA.

Solution 2 was prepared following the nitrous acid method described by Consden *et al.* (152). First, 0.1 mL of a 0.15 mol/L sodium nitrite (Sigma-Aldrich, UK) solution was added to 0.1 mL of the 100 mmol/L HGA stock solution. 2 mL of a 0.01 mmol/L phosphate buffered saline (PBS) solution (pH 6.8) was then added, followed by 0.5 mL of 0.1 mmol/L HCl. The solution was left to stand for 5 min, then 2.3 mL PBS (as above) was added, giving a final BQA concentration of 2 mmol/L. The pH was adjusted to 6.7 using 5 mmol/L NaOH.

271

Solutions 3 and 4 were prepared from dilution of the HGA stock in 0.01 PBS solution (as above) to give HGA solutions of 2 mmol/L. Autoxidation of HGA was initiated by adjusting the pH to 9.5 and 11, for solutions 3 and 4 respectively, using NaOH as above. These solutions are referred to as BQA formed via autoxidation.

For formation of pigment from synthetic HGA (solution 5), the HGA stock solution was diluted 1:9 with water to give 10 mL of a 10 mmol/L aqueous solution of HGA. The solution was incubated in a glass vial at 37°C. A gradual darkening of the solution was observed across the total incubation period of approximately 4 months. The solution was then diluted 1:4 in deionised water for concentration equivalence (2 mmol/L) with the other solutions.

A solution of melanin was prepared for comparison. The melanin was formed synthetically by oxidation of tyrosine with hydrogen peroxide (Sigma-Aldrich, UK). Melanin powder was dissolved in 1 mL dimethyl sulfoxide to give a 10 mmol/L stock solution. The stock solution was diluted 1:4 in water, also for concentration equivalence.

6.2.2 Incubation of HGA and BQA with amino acids

A series of incubations were performed at 37°C between HGA/BQA and several amino acids to assess potential interactions. Incubations were between HGA and arginine, cysteine, glycine or hydroxylysine, and between BQA and glycine or cysteine. Incubations were performed in glass.

HGA and BQA (nitrous acid method) were from the same stock solutions described above. All solutions were prepared at 37°C from 0.01 mmol/L PBS (pH

7.4) that had equilibrated to this temperature. Separate solutions were prepared for each amino acid (Sigma-Aldrich, UK). Two concentrations were prepared for each amino acid (1 and 0.1 mmol/L) while the concentration of HGA or BQA remained constant (33 µmol/L). Control solutions were the amino acid (1 mmol/L) without HGA/BQA, and HGA/BQA (33 µmol/L) without amino acid. All solutions were adjusted to pH 7.4 with NaOH or HCl.

Samples were taken from each solution over the course of the incubation (0–48 h) and stored immediately at -80°C prior to analysis.

6.2.3 Ultrafiltration of HGA-derived pigment

Ultrafiltration was performed on solutions of HGA-derived pigment and melanin. HGA-derived pigment (prepared from 10 mmol/L HGA) and melanin (10 mmol/L) solutions were prepared as described above (section 6.2.1). Ultrafiltration was performed on 4 mL of each solution using Amicon Ultra-4 centrifugal 10 KDa filter units (Merck-Millipore, Ireland); centrifugation at 2500x *g* for 10 min.

6.2.4 Spectrophotometry

Absorbance measurements were taken across the UV-visual range (200-1000 nm) using a Spectrostar Nano absorbance reader (BMG Labtech, UK) and Microvette cuvettes (Sigma-Aldrich, UK).

6.2.5 Mass spectrometry

LC-QTOF-MS analysis of HGA-derived solutions 1-5 (analysed at 10 µmol/L concentrations, injection volume 5 µL) of and samples taken from the HGA/BQA

amino acid incubation experiments (analysed neat, injection volume 1 µL) was performed on an Agilent 6550 QTOF mass spectrometer. Analysis was by Method 2 (reversed-phase LC), as described in Chapter 2 (section 2.2.2).

Additional flow injection analysis (FIA; *i.e.* direct injection of sample into the mass spectrometer without chromatography) was performed for the following samples by QTOF-MS to rule out potential artefacts arising from interaction with the LC column. Sample injection was via a coupled Agilent 1290 Infinity II HPLC system with a zero dead volume connector in place of the LC column. Mobile phase delivery was isocratic (A, water and B, methanol at ratio of 1:1, both with 0.1% formic acid) at flow rate 0.4 mL/min. For HGA-derived pigment (solution 5), melanin and samples taken from the ultrafiltration experiment FIA was performed on an Agilent 6545 QTOF-MS (analysed at 20 µmol/L concentrations, injection volume 10 µL). For samples from the amino acid incubation experiment FIA was performed on an Agilent 6550 QTOF-MS (analysed neat, injection volume 1-2 µL). MS acquisition parameters were identical to those described in section 2.2.2.

In addition to the MS analyses described above which employed ESI (Agilent dual AJS ESI source), HGA-derived pigment was also analysed by QTOF using APCI. APCI analyses were performed on an Agilent 6546 QTOF-MS via FIA, with the dual AJS source replaced with an APCI source (Agilent) (concentration 100 µmol/L, injection volume 1 µL). The MS acquisition parameters were unchanged (except acquisition rate 2 GHz).

Raw MS data files were reviewed in MassHunter Qualitative Analysis (Agilent; build 07.00).

6.2.6 NMR spectroscopy

NMR was performed on HGA-derived pigment and samples from the incubation of HGA with cysteine (48 h at 37°C). Samples were prepared to final concentration of 10% sample with 500 mmol/L phosphate (pH 7.4), and 10% ^2H_2O 200 µL of each prepared sample was then transferred to a 3 mm outer diameter NMR tube (Bruker, UK).

1D ^1H NMR spectra were acquired at 700 MHz using a Bruker Avance III HD 700 MHz NMR spectrometer fitted with a 5 mm triple resonance (TCI) cryoprobe at 25°C. Spectra had spectral widths of 17.16 ppm, 4 s relaxation delay, 32 scans with 4 dummy scans collected into 73 k data points (3.067 s acquisition time).

All spectra were zero filled to 128 k data points with exponential line broadening of 0.3 Hz applied before Fourier transformation, spectra were automatically phased, referenced and baseline corrected in Topspin using vendor supplied automated routines (version 3.5, Bruker, UK).

2D NMR was performed on a 700 MHz spectrometer in two separate experiments known as $^1H^{13}C$ HSQC and $^1H^1H$ TOCSY. Single quantum coherence spectroscopy (HSQC) is a heteronuclear 2D NMR experiment that detects correlations between a carbon and its attached protons (one-bond correlation). The resulting spectrum is plotted with ^1H chemical shift in one dimension and another heteronucleus (in this case ^{13}C) on the other axis, as with the HETCOR

2D NMR spectra presented in Chapter 7 (also see Chapter 1, section 1.5.2). <u>T</u>otal <u>C</u>orrelation <u>S</u>pectroscop<u>Y</u> (TOCSY) is a homonuclear 2D NMR experiment which detects correlations between all protons within a given spin system. All HSQC and TOCSY spectra were acquired at 25°C with 1s relaxation delay ^{1}H spectral width of 12 ppm and 256 increments in the indirect dimension, additionally TOCSY was acquired with 4 scans and HSQC with 16 scans. The spins of these nuclei are interconnected by through-bond interactions, known as 'spin-spin' coupling. In the resulting 2D spectrum, two ^{1}H chemical shift dimensions are plotted against each other and 'cross peaks' represent the correlated proton signals (typically hydrogens/protons between 3 and 6 bonds away from each other).

All spectra were referenced in both ^{1}H and ^{13}C dimensions to selectively (d_4) deuterated trimethylsilylpropanoic acid (Sigma, UK).

6.3   Results

6.3.1 Oxidation of HGA: clear change in colour and absorbance

Formation of BQA and HGA-derived pigment (solutions 2-4) from colourless HGA solution was accompanied by clear change in colour (Figure 6.01). Various degrees of darkening were observed for solutions 2-4. The colour of the HGA-derived pigment sample developed gradually over the 4-month incubation period, whereas immediate colour change was observed for the BQA samples formed by nitrous acid method or alkalinisation. Interestingly, the BQA sample formed by alkalinisation to pH 11 was the darkest in colour.
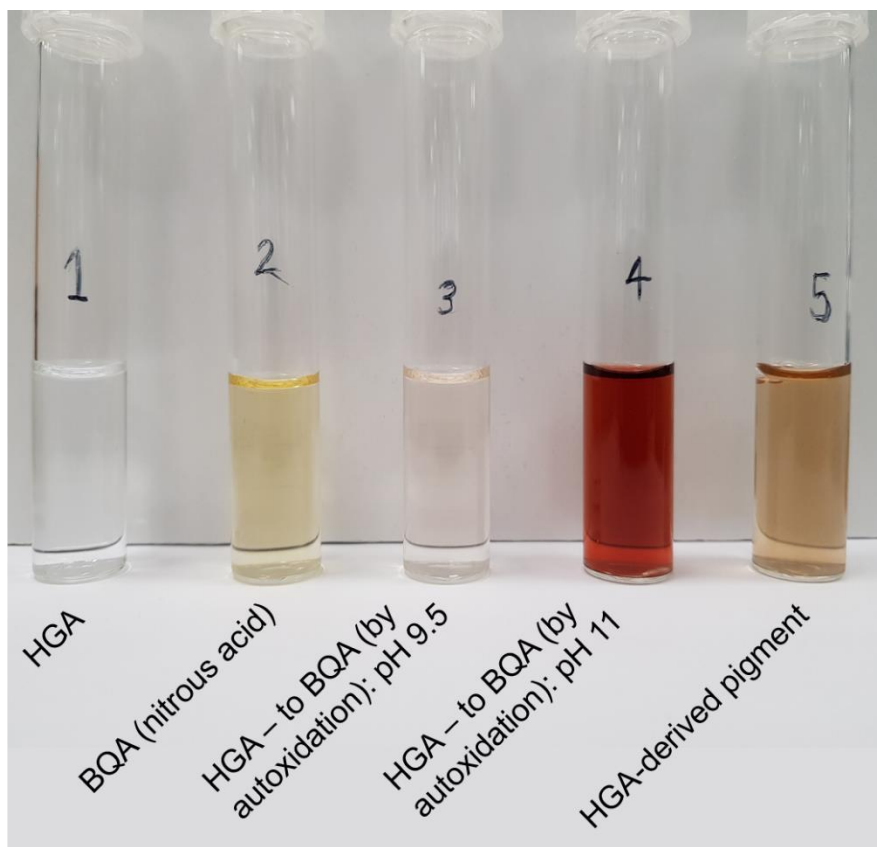
Figure 6.01. Solutions prepared from synthetic HGA for subsequent analysis. Conversion to of HGA (solution 1) to BQA (solution 2) was accompanied by a marked change in UV-visual absorbance (Figure 6.02A). There was a clear shift in the absorbance peak from 290 nm for HGA (hydroquinone) to 250 (benzoquinone) for BQA as previously described (152,330). A series of absorbance measurements were taken at various time points from HGA incubated at 37°C over 48 h (the HGA control solution in the amino acid experiment), as shown in Figure 6.02B. The clear HGA solution at 0 h was visibly pigmented after 48 h (see inset photograph). The UV-visual absorbance traces showed a clear trend across the time series. First, there was a shift in the 290 nm hydroquinone peak to the 250 nm benzoquinone peak (maximum value at 12 h), indicating oxidation of HGA. At 24-28 h there was a decrease in the 250 nm benzoquinone peak accompanied by a more general increase in absorbance across the UV-visual range (200-650 nm), indicating further conformational change beyond oxidation.
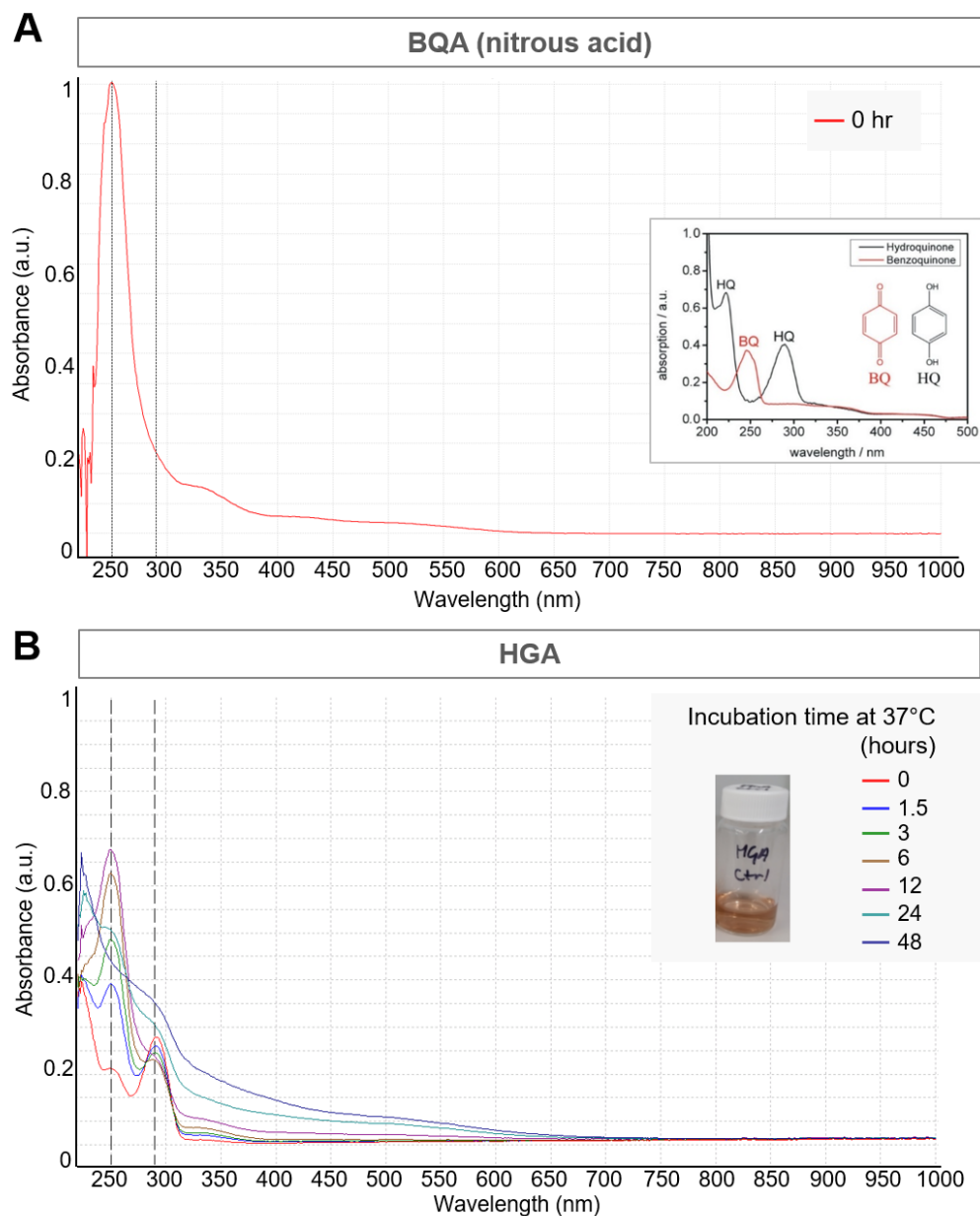
Figure 6.02. Clear change in UV-visual absorbance upon autoxidation of HGA. A: conversion of HGA to BQA by nitrous acid method was accompanied by a shift in the peak at 290 nm (hydroquinone; HQ) to 250 nm (benzoquinone, BQ), as previously described (152). Inset figure adapted from (330). **B:** incubation of HGA over 48 h at 37°C showed gradual development of the 250 peak at 1.5-12 h, indicating oxidation to benzoquinone, followed by a broader increase in absorbance across the UV-visual range at 24-48 h. Inset photograph shows the visibly darkened end product of the HGA incubation at 48 h.

6.3.2 HGA-derived pigment is fundamentally different to synthetic melanin

The LC-QTOF-MS FIA spectra obtained from HGA-derived pigment were first compared with freshly prepared HGA and melanin (Figure 6.03). Melanin was analysed as a reference biological pigment. The data presented are from negative polarity for HGA and HGA-derived pigment and positive polarity for melanin, as these showed the greatest signal.

There was remarkable similarity between the spectra from HGA-derived pigment and HGA. Both spectra were dominated by low molecular weight ions in the *m/z* range 100-400 with no discernible peaks >500 *m/z*. Signal corresponding to HGA (*m/z* 167.035) was also prominent in the pigment spectrum; this was particularly surprising given the long incubation period at 37°C and the disappearance of the sharp 290 nm hydroquinone absorbance peak for HGA within 24 h, as discussed above (Figure 6.02).

The most prominent peaks in the HGA-derived pigment spectrum were ≤167 *m/z*, indicated by arrows in Figure 6.03. These peaks remained after subtracting the background spectra (removing reference ion signals and background noise). The peaks were also present in the spectrum from HGA, except the 137.0245 ion which appeared to be unique to HGA-derived pigment in FIA. This ion was therefore initially of interest, but it was also observed for fresh HGA in subsequent LC-QTOF-MS analyses (see Figure 6.05).
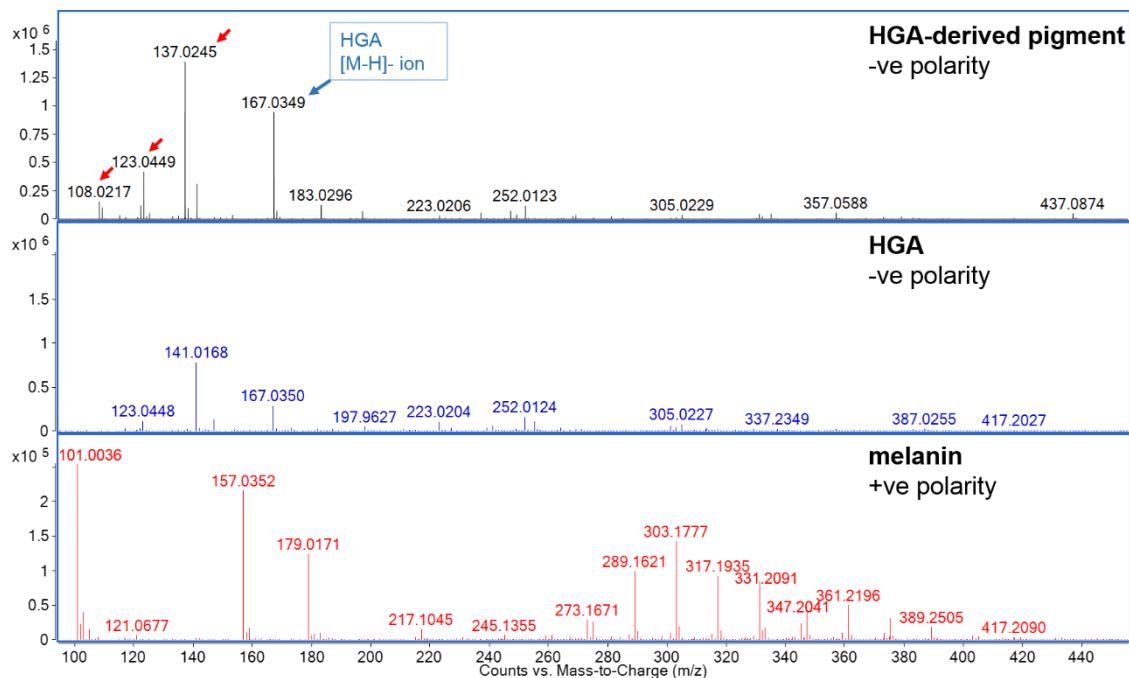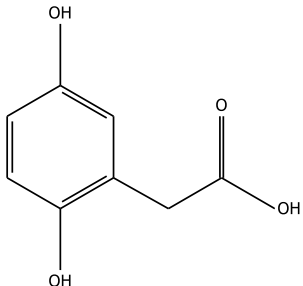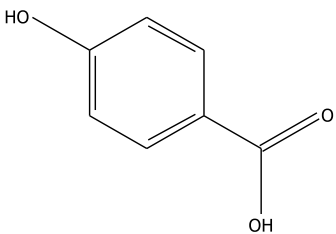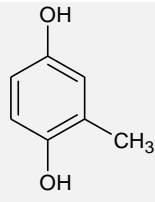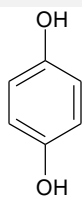
Figure 6.03. Flow injection analysis (FIA) mass spectra from analysis of HGA-derived pigment, HGA and melanin. The spectra for HGA and HGA-derived pigment are remarkably similar and both dominated by low molecular weight ions. Red arrows indicate fragment ions in the pigment that are known to be derived from HGA. In contrast, the spectrum for melanin shows regularly repeating units (in the *m/z* range 270-400) typical of a polymer. Spectra were acquired by FIA with electrospray ionisation.

Targeted MS/MS fragmentation analysis was performed (also by FIA) to identify the chemical structures of the most prominent peaks detected from HGA-derived pigment. Candidate compound/formula matches with the observed fragmentation spectra were obtained using the MassHunter METLIN metabolites PCD/PCDL accurate mass database (Agilent, build 07.00) and Molecular Structure Correlator (Agilent, build 07.00). These tools were used for comparing acquired targeted MS/MS spectra against published MSMS spectra (METLIN) for candidate

compounds and correlating acquired fragmentation patterns against fragmentations predicted *in silico* (Molecular Structure Correlator). Candidate compounds and structures were generated for each entity, except for the 141.0168 *m/z* ion which did not yield any formula hits by either approach (Table 6.01). The ions with *m/z* 167, 123 and 108 were confirmed to be associated with HGA based on the structures identified using this approach. The 167 ion is the monoisotopic [M-H]$^-$ ion for HGA, and 123 and 108 are known fragments of HGA (fragmentation spectra from the METLIN library noted above). The 137 *m/z* ion is not a known fragment of HGA from METLIN but does appear to be associated with HGA. Interestingly, the compound match for this ion was 2,5- or 3,4-dihydroxybenzaldehyde; how this structure relates to HGA and/or pigment formation is, however, unclear. The 141 *m/z* ion is also not a known fragment of HGA but was observed in the spectra of HGA and HGA-derived pigment; no formula matches were obtained for this entity. The presence of these ions in the spectra from MS analysis without application of collision energy (*i.e.* not MS/MS), as shown in Figure 6.03, suggest that they are fragments of HGA formed in the ionisation source; a well-known phenomenon in LC-MS analyses by ESI (331). However, the ions remained after reducing the source fragmentor voltage from 120 V to 80 V in an attempt to prevent in-source fragmentation.

Table 6.01. Compound structure identifications for ions observed from HGA-derived pigment. Structures were identified in negative polarity by matched MS/MS fragmentation spectra using the MassHunter METLIN metabolites PCD/PCDL accurate mass database (Agilent, build 07.00) or Molecular Structure Correlator (Agilent, build 07.00).

m/z	Top candidate formula	Top METLIN compound match (MS/MS spectra)	Predicted structure
167.0349	$C_8H_8O_4$	HGA	
141.0168	-	-	-
137.0245	$C_7H_6O_3$	2,5- or 3,4-dihydroxy-benzaldehyde	
123.0449	$C_7H_8O_2$	HGA fragment	
108.0217	$C_6H_5O_2$	HGA fragment	

The spectrum obtained from melanin was strikingly different to HGA and HGA-derived pigment. Melanin showed a spectral pattern comprising ions spaced at regular mass intervals in the *m/z* range 270-400. This pattern suggests a classic polymeric structure composed of regularly repeating units. Molecular feature extraction was performed on ions in the *m/z* range 270-400 and candidate molecular formulae generated (Table 6.02). The formulae show that the mass intervals observed, 14 and 16 *m/z* units, are due to successive incorporation of methylene (CH_2) groups and oxygen, respectively, into the melanin chemical structure.

Table 6.02. Positive polarity ion species and formula matches for the regularly repeating mass units obtained from melanin. Formulae shown were the top matches generated following molecular feature extraction in MassHunter Qualitative Analysis (Agilent; build 07.00).

m/z	Formula match
273.1673	$C_{10} H_{20} N_6 O_3$
275.1463	$C_9 H_{18} N_6 O_4$
289.1622	$C_{10} H_{20} N_6 O_4$
303.1776	$C_{11} H_{22} N_6 O_4$
317.1937	$C_{12} H_{24} N_6 O_4$
331.173	$C_{12} H_{22} N_6 O_5$
331.2093	$C_{13} H_{26} N_6 O_4$
333.1884	$C_{12} H_{24} N_6 O_5$
345.1884	$C_{13} H_{24} N_6 O_5$
345.2244	$C_{14} H_{28} N_6 O_4$
347.2044	$C_{13} H_{26} N_6 O_5$
361.2199	$C_{14} H_{28} N_6 O_5$
375.2353	$C_{15} H_{30} N_6 O_5$
389.2508	$C_{16} H_{32} N_6 O_5$

Ultrafiltration was performed on HGA-derived pigment and melanin solutions using a 10 KDa filter as an additional simple assessment of molecular weight. First, both solutions were centrifuged prior to filtration. As seen in Figure 6.04A, a small pellet was formed in in the HGA-derived pigment sample by undissolved pigment particulates. However, the supernatant remained heavily pigmented (Figure 6.04B, solution 2). In contrast to melanin, which remained black (Figure

6.04B solutions 5-8), HGA-derived pigment solution became noticeably lighter in colour post-filtration (Figure 6.04B, solution 4). The portion of the solution that had not passed the filter after 10 min of centrifugation appeared slightly darker, indicating retention of pigment by the filter. Together, these findings suggest that HGA-derived pigment, but not melanin, contains a portion of compounds with molecular weight >10 kDa.
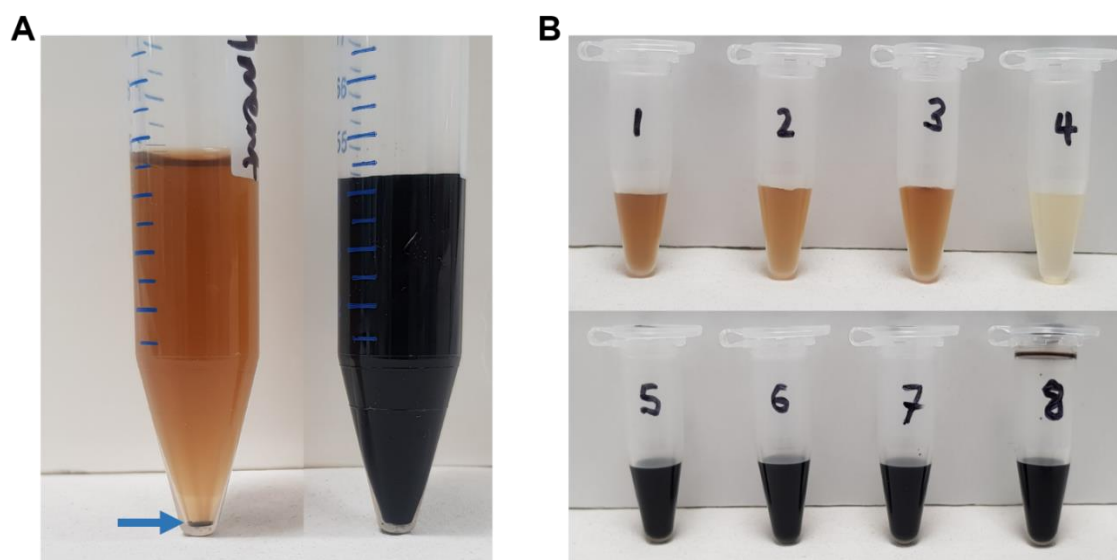
Figure 6.04. Ultrafiltration of HGA-derived pigment and melanin solutions. A: HGA-derived pigment (left) and melanin (right) solutions post-centrifugation. Blue arrow indicates pellet formed by undissolved particulates of HGA-derived pigment. **B:** samples from HGA-derived pigment (top photograph) and melanin (bottom photograph) taken pre- (samples 1 & 5) and post-centrifugation (samples 2 & 6). Samples 3 and 7 are the portions that had not passed the 10 kDa filter after 10 min at 2500x g. Samples 4 and 8 are samples of the filtrate from both solutions. The filtrate was visibly lighter for HGA, indicating some retention of pigment species by the filter.

6.3.3 HGA-derived pigment does not behave as a polymer in mass spectrometry

Given the indication of high molecular weight (>10 kDa) species in HGA-derived pigment from ultrafiltration, it was surprising that the mass spectrum of HGA-derived pigment was dominated by low molecular weight ions (100-300 *m/z*). LC-QTOF-MS analysis was performed to further compare HGA-derived pigment with HGA and BQA intermediates (solutions 1-5, section 6.2.1). TICs showed no clear differences between the solutions. Consistent with the data shown in Figure 6.03, a clear HGA signal (HGA [M-1H]$^-$ and [2M-H]$^-$ ions) was observed for HGA, BQA and HGA-derived pigment, with RT matching that of HGA from the AMRT database described in Chapter 2 (Figure 6.05, left panel). Interestingly, the HGA-associated signal was removed for the HGA – BQA solution formed by raising the pH was to 11 (solution 4), but without clear formation of other ion products.

The extracted mass spectra (Figure 6.05, right panel) were almost identical for solutions 1-3 and 5. The ions indicated by red arrows have been highlighted because here they were previously confirmed as compounds originating from HGA and not HGA-derived pigment (see Table 6.01). Further confirmation of this was provided here as these ions chromatographed (*i.e.* same RT) with HGA.
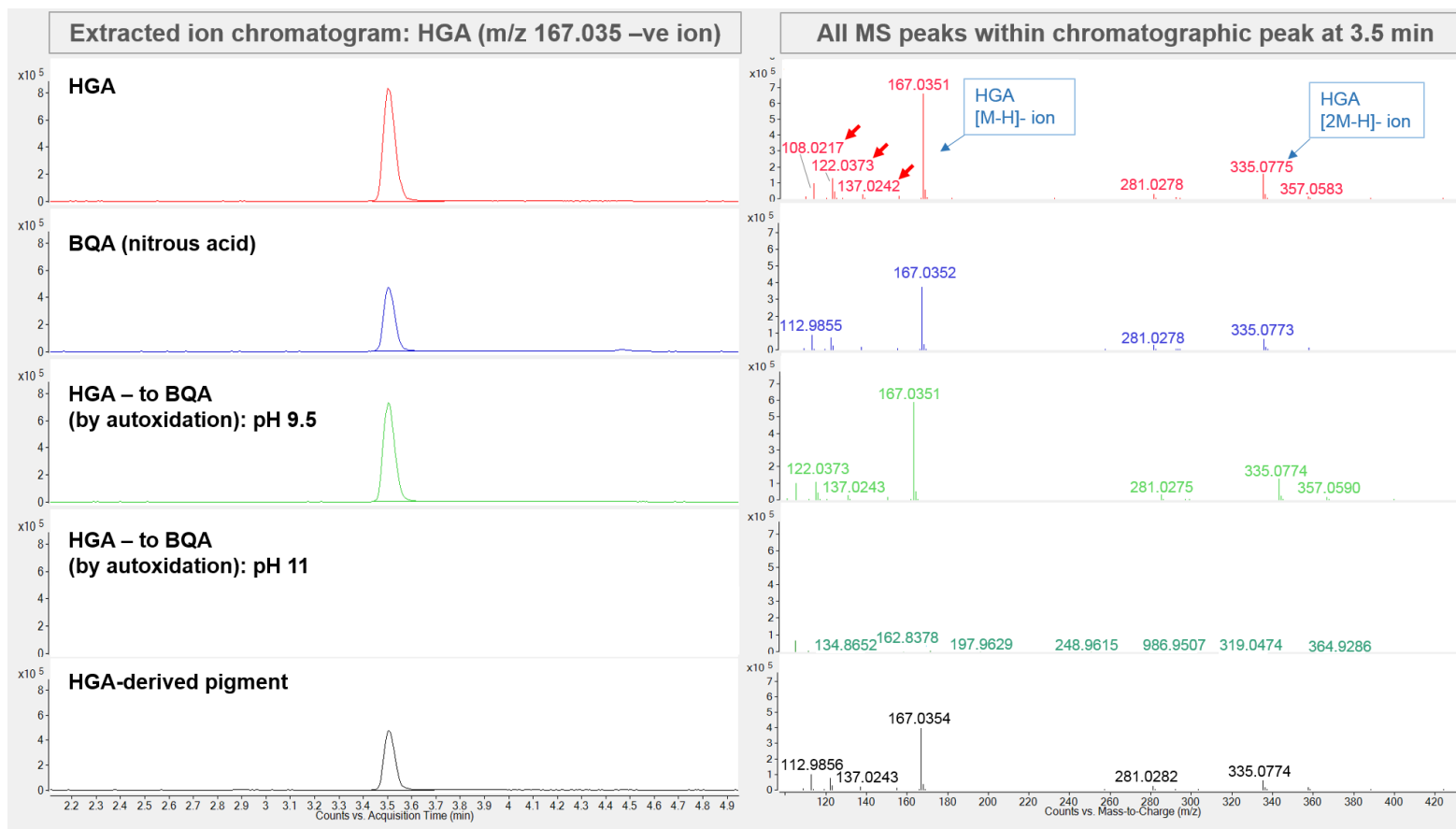
Figure 6.05. Mass spectra from LC-QTOF-MS analysis of HGA, BQA and HGA-derived pigment. Spectra show remarkable similarity, dominated by low molecular weight ions and ion signals known to be associated with HGA (indicated with arrows). No signal for BQA (expected *m/z* 165) was observed for any solution. Solutions analysed by reversed-phase LC method 2 (Chapter 2, section 2.2.2) in electrospray ionisation.

287

Clear observation of HGA-associated ions was surprising for BQA (solution 2) given that the UV-visual absorbance spectrum confirmed conversion to the benzoquinone form (Figure 6.02). The *m/z* of the HGA ion observed was 167 but the expected *m/z* for BQA would be 165; *i.e.* loss of 2 hydrogens from HGA upon oxidation (Figure 6.06). One consideration was that potential redox artefacts were being introduced by the ESI technique. It has been shown previously that the electron-rich environment in the ESI source can result in reduction of a range of quinones (332). In the same way, benzoquinone species in BQA solutions or HGA-derived pigment could be reduced back to HGA during ESI due to gain of electrons; *i.e.* gain of 2 hydrogens taking the mass from 165 back to 167 (Figure 6.06).
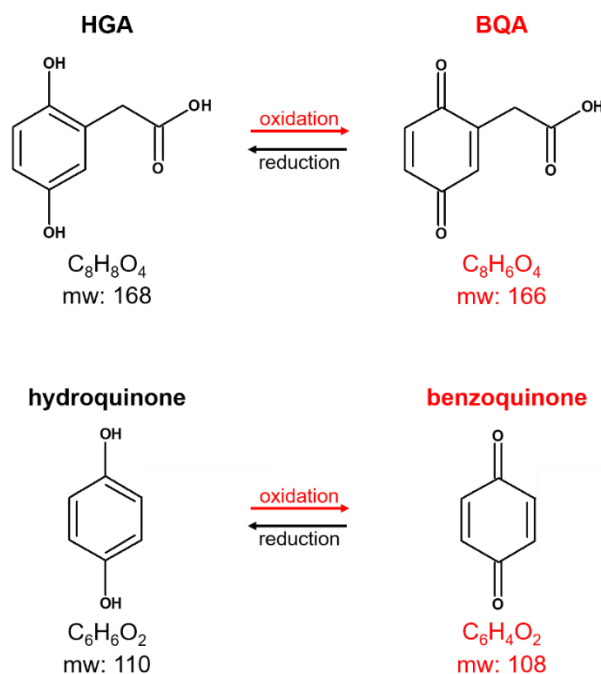
Figure 6.06. Redox transitions of HGA – BQA and hydroquinone - benzoquinone.

In order to rule out potential BQA/HGA-derived pigment reduction artefacts attributable to ESI, pigment was also analysed by an alternative ionisation technique; APCI (via FIA). Figure 6.07 shows the mass spectra obtained from HGA-derived pigment solutions 2-4 from the ultrafiltration experiment (Figure 6.04) by APCI. As with ESI, the spectra from these solutions were dominated by the HGA [M-H]$^-$ signal (167 m/z), without evidence of a 165 m/z ion from BQA. This indicated predominance of the reduced hydroquinone species in the pigment (Figure 6.06). Extending the acquisition mass range to 3200 m/z (from 1700 m/z) did not reveal any discernible high molecular weight signals beyond those in the mass range shown in Figure 6.07. In summary, given the striking dark appearance of HGA-derived pigment solution, there is remarkably minimal difference from fresh HGA in the mass spectrum obtained from MS by ESI and APCI.
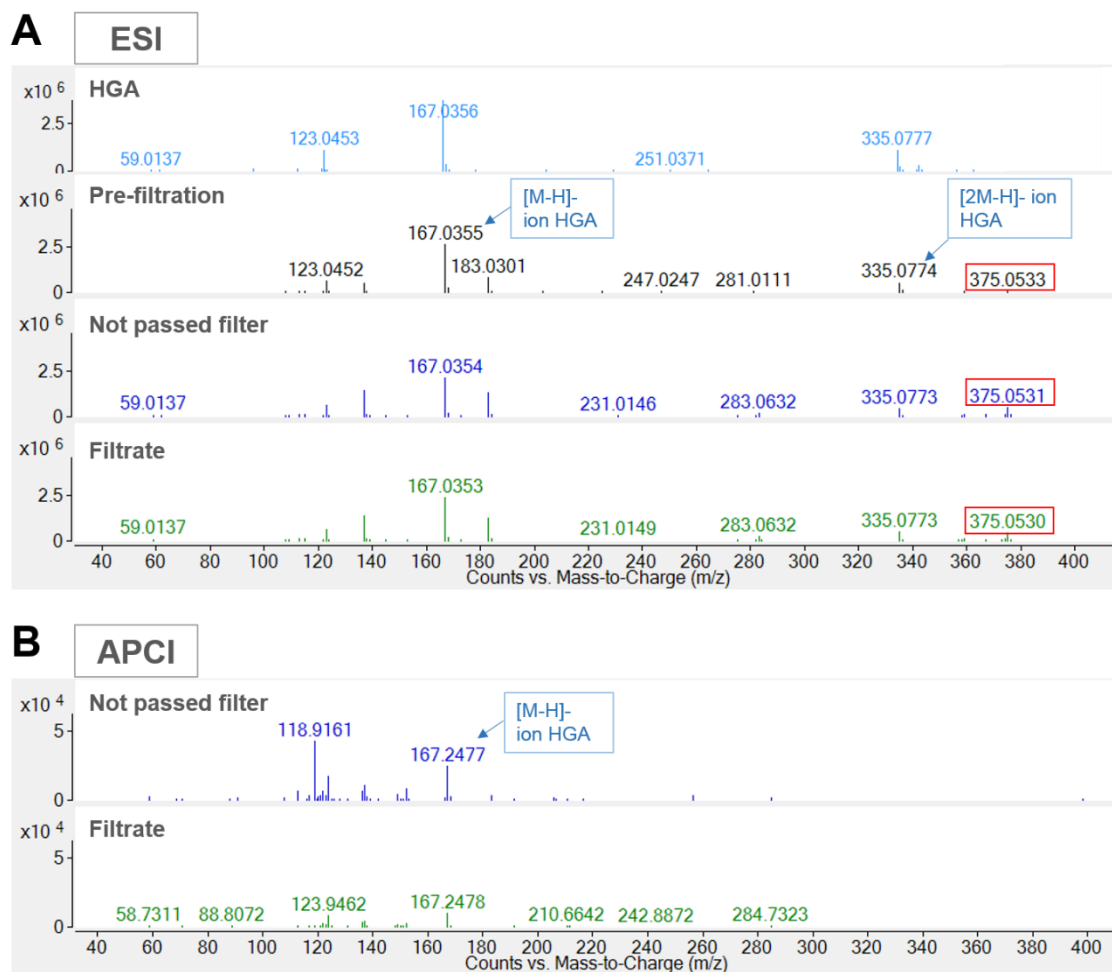
Figure 6.07. HGA signal dominates the mass spectra of HGA-derived pigment in MS by ESI and APCI. The HGA ([M-H]$^-$ ion 167 *m/z*) signal is prominent in solutions of HGA-derived pigment before (darker in colour) and after (lighter in colour) ultrafiltration (see Figure 6.04). No discernible signal was observed corresponding to BQA (expected *m/z* 165) or for ions greater than the mass range shown. ESI revealed a series of ions from pigment samples (but not freshly prepared HGA) with *m/z* greater than that of the HGA [M-H]$^-$ ion, which could potentially represent aggregation products formed by HGA in the process pigment formation. The 375.0531 ion is highlighted (red box); targeted MS/MS was subsequently performed on this ion (see Figure 6.08).

6.3.4 First observation of an aggregate from HGA-derived pigment by mass spectrometry

Closer inspection of the ESI spectra from FIA (shown in Figure 6.07) shows a collection of ions with m/z >167 (HGA [M-H]$^-$ signal), up to 400 m/z, that were observed from the pigment samples but not fresh HGA. These could therefore represent potential candidate aggregate structures resulting from HGA oxidation. Targeted MS/MS analysis was performed on these ions to investigate their structures. As shown in Figure 6.08, fragmentation of the 375 ion (blue diamond) generated a series of peaks, the most prominent being negatively charged ions with m/z 209.0259, 167.035 (HGA), 137.0252, 121.0295, 108.0217 (HGA fragments; see Table 6.01) and 165.0195. The 209 and 165 ions were particularly interesting as they are not known fragments of HGA. The 165 ion appears to represent BQA (AM match = 1.01 ppm). The only formula match generated for the 209 ion was $C_{13}H_6O_3$, but no compound match (mass match <10 ppm) was obtained in the MassHunter METLIN metabolites PCD/PCDL database (Agilent, build 07.00). Together, these data reveal for the first time compound ions that are structural components of the oxidation/aggregation of HGA and BQA in formation of HGA-derived pigment.
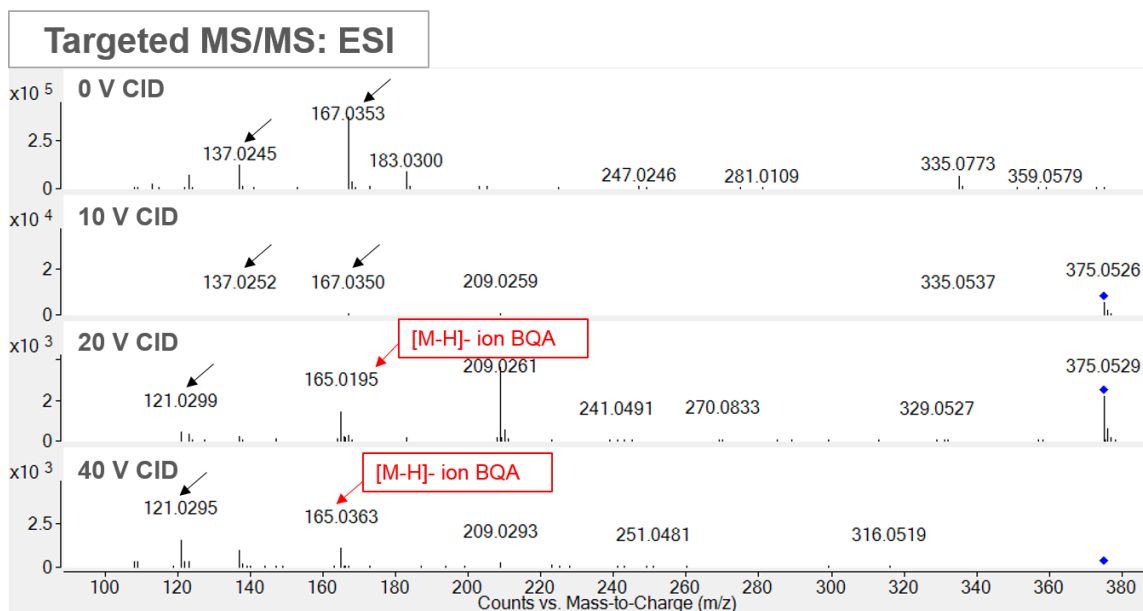
Figure 6.08. Targeted MS/MS analysis identifies constituent fragments of potential HGA aggregation compound. Fragment ions generated from collision-induced dissociation (CID; at 0-40 V) of the 375 ion (present in HGA-derived pigment but not HGA) revealed constituent ion fragments relating to HGA and known sub-structures of HGA (indicated by black arrows). Other generated fragments were an ion of greater mass than HGA (*m/z* 209.0259, formula match $C_{13}H_6O_3$) and BQA (red arrow). The data suggests that these ions are involved in the oxidation and/or aggregation of HGA in forming pigment.

6.3.5 Interaction between HGA or derived species with amino acids in solution

HGA and BQA (by nitrous acid method) were incubated with various amino acids in solution in order to indicate potential interaction/binding sites in the cartilage matrix. The LC-QTOF-MS mass spectra obtained from these solutions after 48-h incubations are shown in Figure 6.09.
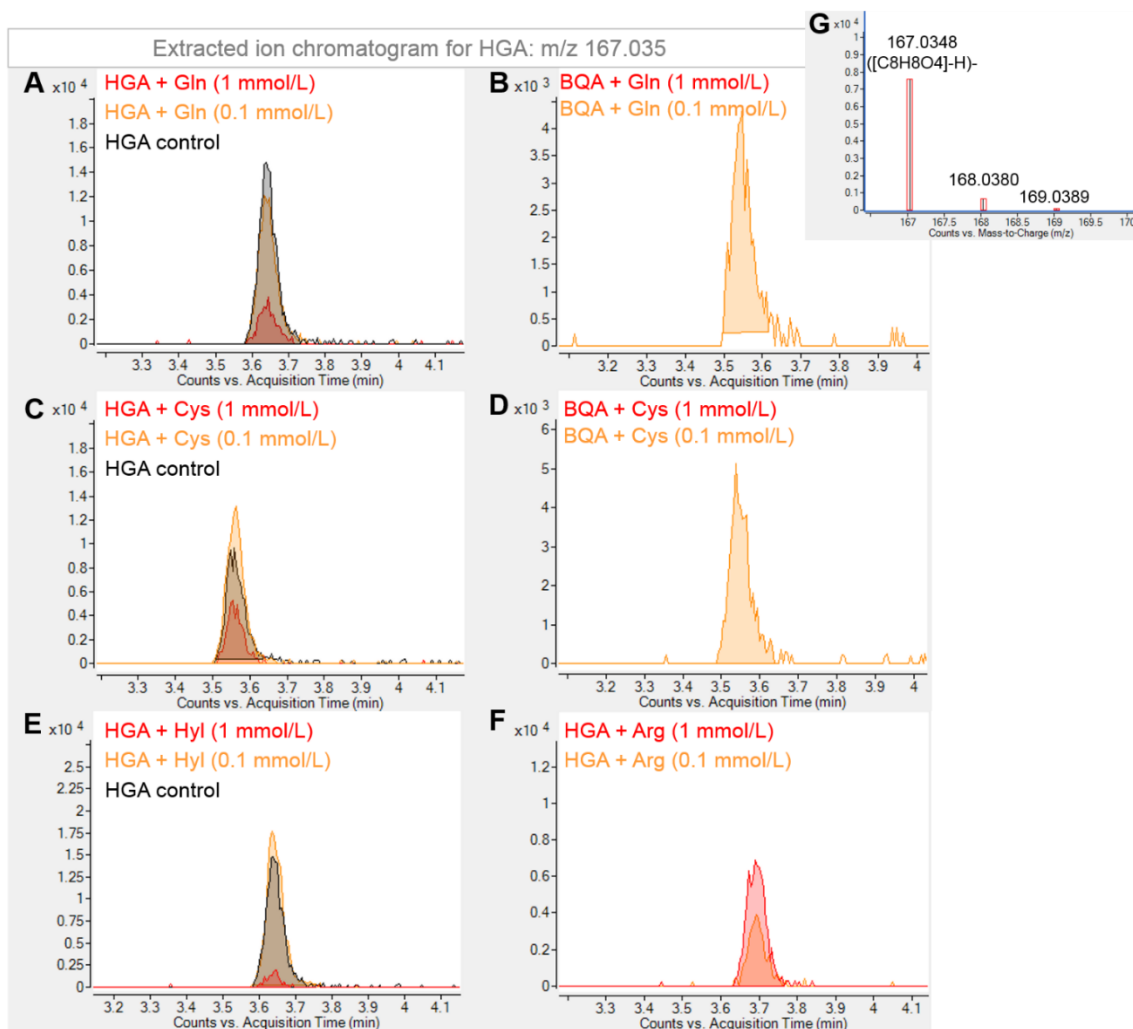
Figure 6.09. Evidence of interaction between HGA/BQA and amino acids.

A-F: extracted ion chromatograms (EICs) for HGA ([M-H]⁻ ion 167.035 *m/z*) obtained from solutions incubated for 48 h at 37°C (pH 7.4). EICs indicated interactions for HGA and BQA with glycine and hydroxyproline, and to a lesser extent HGA with cysteine (HGA signal attenuated by increased amino acid concentration). **G:** representative extracted mass spectrum for HGA (^{12}C – ^{14}C isotope distribution visible). Solutions analysed by reversed-phase LC method 2 (Chapter 2, section 2.2.2) in electrospray ionisation.

Interaction between HGA and glycine and hydroxylysine was indicated by attenuation of the HGA ([M-H]$^-$ ion 167.035 *m/z*]) signal at high (1 mmol/L) *versus* low (0.1 mmol/L) concentrations of these amino acids. For these solutions, HGA incubated with the low concentration of amino acid showed approximately equivalent HGA signal to that of the HGA control (HGA incubated alone over the same 48-h period). There was also evidence of a weaker interaction between HGA and cysteine, but no interaction between HGA and arginine. Interactions were also indicated between BQA and glycine and cysteine; HGA was no longer detectable at the high 1 mmol/L concentrations for these amino acids.

Despite evidence of these interactions, newly formed products were not detected (*i.e.* amino acid + HGA/BQA expected adducts). Potential products were searched in the data based on the predicted amino acid + HGA/BQA adducts; AM calculated from the combination of their formulas (with loss of H_2O). Additional analysis by FIA also did not show emergence of any new peaks indicating formation of new products.

6.3.6 NMR spectroscopy confirms HGA conformational change

Analyses were performed by NMR spectroscopy to confirm that HGA had undergone significant structural change in the formation of HGA-derived pigment. First, a series of 1D ^{1}H experiments were performed on HGA (0.33 mmol/L) that had been incubated at 37°C for 48 h at pH 7.4 with and without 1 mmol/L cysteine (samples from the amino acid incubation experiment; section 6.2.2).

As shown in Figure 6.10, 3 distinct groups of signals ('poses' or 'conformers') were observed in the ^{1}H aromatic region for both samples over the 14-day analysis period. The three poses refer to distinct pairs of ^{1}H resonances which increased and decreased at the same rates. The poses had characteristic profiles over the time period. At the earliest time points, pose 1 (red shading) dominated the spectra for HGA with and without cysteine. For HGA without cysteine, pose 2 (blue shading) dominated at days 2-5, then pose 3 (magenta shading) emerged from day 4 onwards. At day 14, each of these signals were undetectable / attenuated considerably. Interestingly, presence of cysteine appeared to stabilise the peaks observed at day 0 (pose 1). For HGA with cysteine, the 3 poses showed the same temporal profiles across the period but with decreased rate of progression, indicating slower conformational change; an effect presumably attributable to cysteine being a reducing agent and therefore attenuating the rate of oxidation. This interpretation was also supported by the colour of the solutions after the 48-h incubation; darkening was observed for HGA (see Figure 6.02B; inset photograph) and the colouration was visibly lighter, though still darkened, with cysteine present (both solutions were pH 7.4 at the start and end of each incubation).
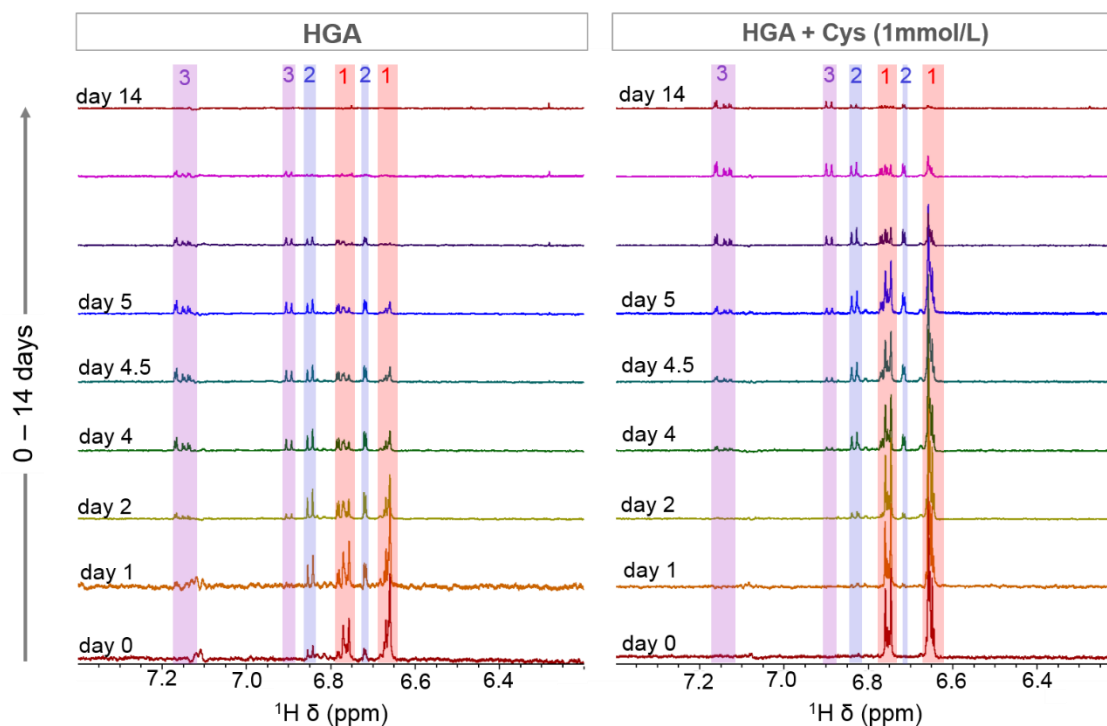
Figure 6.10. 1D ^{1}H NMR spectra show signals associated with HGA turnover over 14 days. A time series of spectra (aromatic ^{1}H range shown) were obtained from solutions of HGA and HGA + cysteine (1 mmol/L); both solutions had been pre-incubated for 48 h at 37°C, pH 7.4). At day 0, pose 1 (red shading) dominated, followed by increase in the peaks representing pose 2 (blue shading). Pose 3 (magenta shading) was the last signal to emerge, remaining longer than poses 1 and 2. Presence of cysteine appeared to stabilise this rate of conformational change, with pose 1 predominating the spectra for a longer period.

Note: Over the 14-day analysis period temperature fluctuated between 25°C (in storage) and 4°C (during acquisition). Both samples experienced identical temperatures for equivalent time frames.

2D NMR was performed to further investigate the 1D ^{1}H NMR signals observed. The ^{1}H ^{1}H TOCSY 2D spectrum (Figure 6.11) confirmed the relationship between each of the ^{1}H pairings within the poses observed in Figure 6.10. Each of the

cross peaks (peaks away from the diagonal) represent direct correlations between the three pairs of aromatic protons.
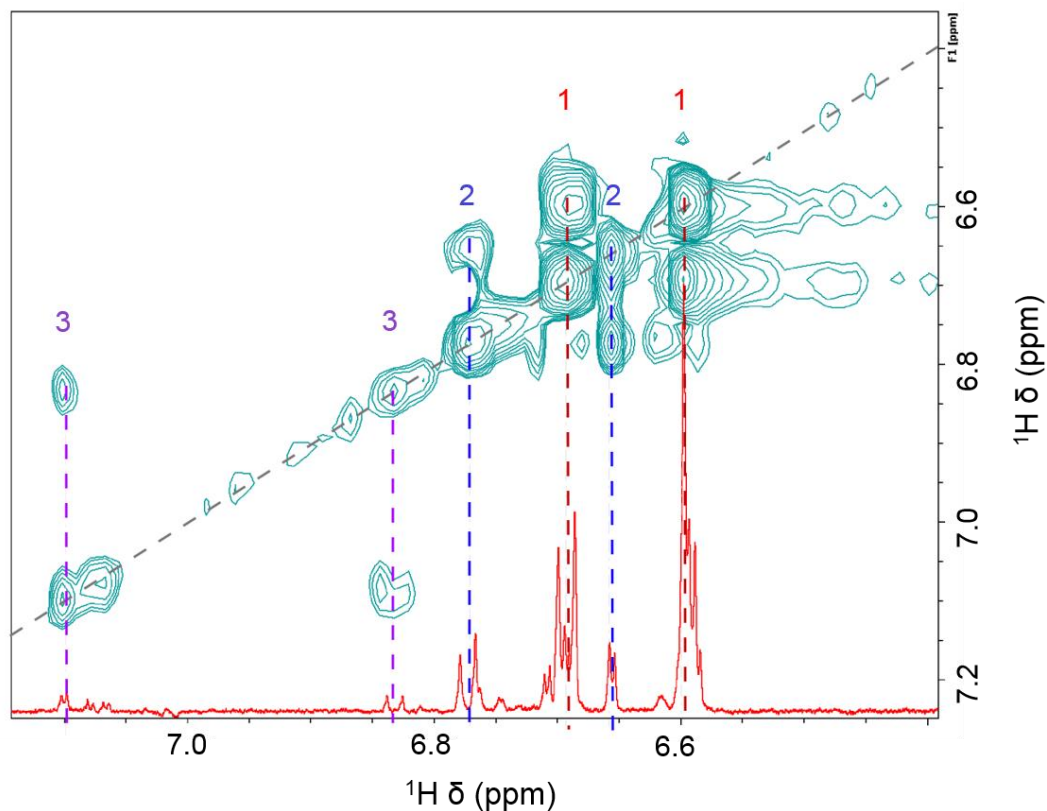
Figure 6.11. 2D NMR confirms the three different temporal conformations of HGA (i). ^{1}H^{1}H 2D TOCSY (total correlation spectroscopy; signal shown in teal) confirms the three different conformations of the HGA aromatic resonance by presence of cross peaks demonstrating that peaks can be attributed to the three sets of resonances (three conformations/poses indicated by red, blue and magenta dashed lines). 1D spectra shown in red.

Heteronuclear 2D NMR (^{1}H^{13}C HSQC) confirmed that the ^{1}H signals observed from HGA/pigment in the 3 conformational pairings were from protons attached to an aromatic ring. As shown in Figure 6.12, the ^{1}H signals were paired with aromatic carbons (aromatic ^{13}C chemical shift range shown). Here the signals

were also more intense for ^{1}H-^{13}C pairings for pose 1 (highlighted by red shading) at 0 days (blue 2D spectrum) compared with 7 days (red 2D spectrum).
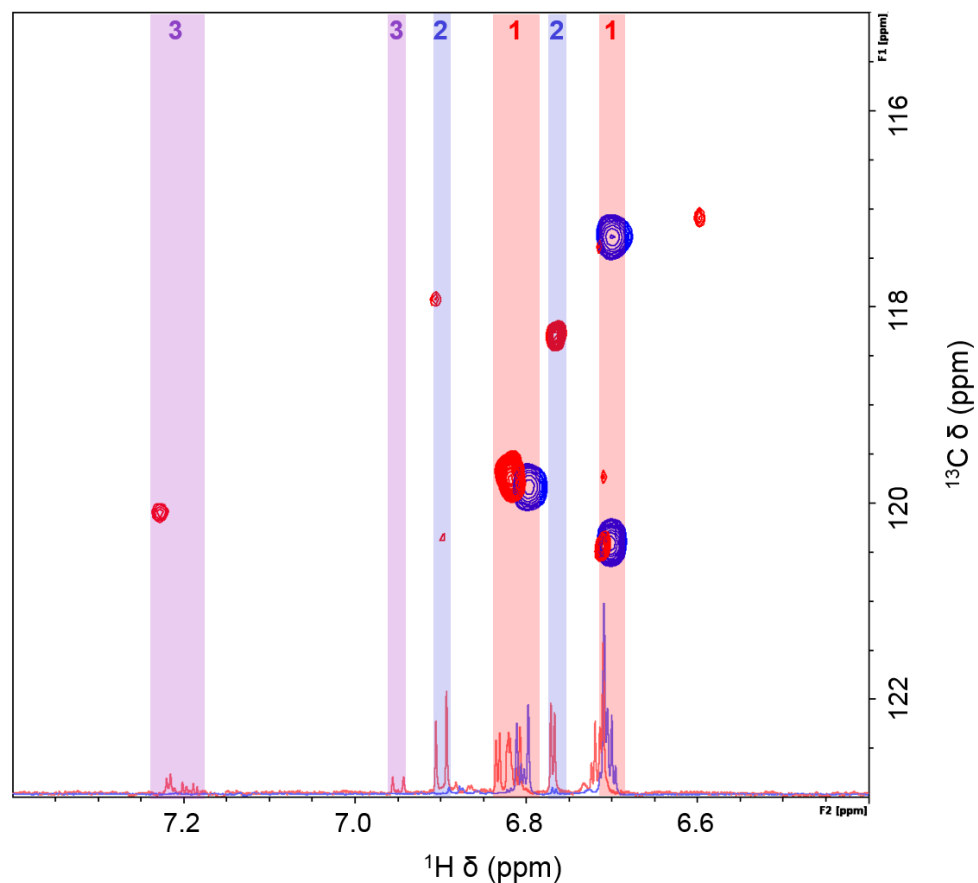
Figure 6.12. 2D NMR confirms the three different temporal conformations of HGA (ii). ^{1}H^{13}C 2D HSQC for HGA without cysteine at time 0 (blue) and time 7 days (red) overlaid. 2D spectra show ^{1}H pairings with between aromatic carbons (aromatic ^{13}C chemical shift range shown), confirming attachment of the protons to an aromatic ring. Equivalent 1D ^{1}H spectra are overlaid.

6.4   Discussion

The data presented in this chapter are fundamental to understanding ochronosis, the primary cause of pathophysiological manifestations in AKU. Generation of pigment formed from HGA *in vitro* enabled the study of ochronosis in pure form

in the absence of complex biological matrix. The most striking results from these studies were from MS analyses: a) mass spectra obtained from HGA-derived pigment were dominated by low molecular weight ions in the *m/z* range 100-400 including HGA and fragment ions derived directly from HGA; b) the first demonstration of an aggregate structure from HGA-derived pigment by MS.

The spectra obtained from HGA-derived pigment showed marked contrast to that of synthetic melanin formed by oxidation of tyrosine by hydrogen peroxide. While HGA-derived pigment showed a general broadband increase in absorbance across the UV-visual range (Figure 6.02B) similar to that described previously for melanin (160), its mass spectra are not consistent with a polymeric structure. In contrast, melanin gave a mass spectrum comprising ions at regularly repeating mass intervals indicating a highly-organised, well-defined structure.

The mass spectra of HGA-derived pigment were unexpected for a number of reasons. First, the marked colour change and change in UV-visual absorbance profiles upon alkalinisation or formation of BQA via nitrous acid method indicated oxidation to benzoquinone (MS signals indicated dominance of HGA, *i.e.* the reduced form). Second, ultrafiltration of HGA-derived pigment solution indicated partial presence of species >10 kDa. The final reason is the widespread assumption in the literature that ochronotic pigment is a polymer (see Chapter 1, section 1.3.1).

Closer inspection of the spectra acquired by ESI MS (by FIA) revealed the first aggregate structure obtained from HGA-derived pigment by MS. Targeted

fragmentation of this ion (*m/z* 375.0529) revealed constituent fragments comprising a smaller aggregate (*m/z* 209.0259), HGA and BQA. The data therefore reveal that these fragment ions are part of the oxidation/aggregation process resulting in pigment. This also suggests that pigment is formed by a combination of HGA and BQA aggregation.

In the literature there is scant evidence to support the widespread assumption that ochronotic pigment is a polymer. One piece of evidence comes from the fascinating case of an Egyptian mummy diagnosed with AKU following observation of ochronosis in hip cartilage and osteoarthropathy indicated by joint space narrowing and extensive intervertebral disk calcification (329). Presence of methylene groups and six-seven membered ring ketones was deduced from the absorbance spectra of ochronotic pigment extracted from cartilage. These data were interpreted as indication of an oxidised HGA polymer. However, these findings are confounded by the cartilage matrix from which the pigment was extracted; it is not certain whether the compound observed was from ochronotic pigment or the cartilage matrix. This underlies the value of comparatively simpler experiments on solutions of purified HGA and HGA-derived pigment formed under controlled conditions in the laboratory, as discussed here.

It was interesting in the present analyses that the fragment ion identified as BQA was not readily detectable by MS analysis (even in solutions of BQA), but observed upon breaking the bonds of the observed aggregate structure by application of 20 V collision energy. One potential explanation is that MS ionisation via ESI and APCI caused reduction of benzoquinone species

300

(proposed intermediates in the classic HGA polymerisation scheme) back to HGA in the HGA-derived pigment samples. This phenomenon has been described previously in ESI MS for quinones including 1,4-benzoquinone, methyl-p-benzoquinone, 1,4- and 1,2-naphthoquinone and 1,4-anthraquinone (333) and attributed to conditions that bring about severe corona discharge. Additional APCI MS analysis was performed in order to rule out this potential explanation, but the pigment spectra obtained were still dominated by the reduced form, HGA. It is worth noting that reduction may also have been induced in-source in APCI; in APCI, analytes are not subjected to the electron-rich environment as in ESI sources, but the ionisation is by corona discharge. Another potential explanation is that fragmentation of a high molecular-weight polymer was being induced in-source. However, ESI and APCI are known as soft ionisation techniques, and decreasing the source fragmentor voltage (120 V to 80 V) had little to no observable effect on the spectra. Alternatively, a more general explanation is that HGA-derived pigment is an analytically challenging compound due to the inherent tendency of benzoquinone to undergo redox cycling (benzoquinone to hydroquinone) (174,334).

The NMR studies confirmed that the formation of pigment from HGA is accompanied by significant structural conformation change. The three different conformations observed represented different proton environments associated with an aromatic ring structure. This conformational change related to the carbons on an aromatic ring structure, which is likely to reflect the oxidation/aggregation of HGA and BQA that was informed by the MS experiments.

The MS data also indicated interaction between HGA and the amino acids cysteine, glycine, and hydroxylysine, and also for BQA with glycine and cysteine. These findings are in line with previous work demonstrating that BQA binds to amino acids via free amino or thiol groups (177). A glycine-HGA adduct is also previously reported in AKU urine; this adduct was predicted to represent 25% of total urinary glycine in this patient (335). The interactions with glycine and hydroxyproline are particularly interesting, as these amino acids are major components of type-II cartilage collagen, as therefore could present potential binding sites for HGA-derived species in cartilage. Such binding sites could explain the high affinity that HGA or its derivatives have for cartilage; cartilage is the major accumulation site for ochronotic pigment. It is important to show that these interactions can occur in pure solutions *in vitro,* as identification of specific binding sites is complicated *in vivo* due to the complexity of the cartilage extracellular matrix.

In summary, these data fundamentally challenge the widely-held idea that ochronotic pigment is a classic polymer comprising covalently-bound and regularly repeating sub-structures. The data confirm that HGA undergoes significant conformational change in forming pigment, but this change is more consistent with non-covalent aggregation of a number of species including HGA and BQA. These self-aggregation mechanisms appear to be more complex and less organised than those involved in the formation of melanin formed by oxidation of tyrosine.

Acknowledgements

7.0   STUDY OF HGA AND OCHRONOTIC PIGMENT FORMED *IN VIVO*

7.1 Introduction

The exact mechanism by which ochronotic pigment interacts with collagen is not yet fully elucidated. The exposed collagen hypothesis of ochronosis predicts that natural, age-related alteration to cartilage type-II collagen renders it susceptible to binding by HGA and any derived products in AKU (9). In AKU, ochronotic cartilage shows different bulk mechanical properties compared to non-ochronotic cartilage (306), therefore it was predicted that there are structural changes at a molecular and even atomic lengthscale in the collagen proteins that underlie and predate the onset of arthritic symptoms. Despite the lifelong elevation in plasma HGA, ochronosis and joint degeneration is not evident until later life in AKU (18); therefore, the preceding changes to the cartilage matrix that facilitate binding of pigment are likely to be age-related and of broader interest in understanding other cartilage diseases.

In the current model of pigmentation, HGA is first oxidised to BQA, then undergoes aggregation or polymerisation to yield a final pigment species that would include conjugated carbon-carbon double bonds containing benzoquinone moieties, giving rise to the dark colour (42). If the pigment becomes polymerised extensively and intertwined in the ECM, it would potentially provide a mechanism of the irreversibility of the pigmentation process, though low molecular-mass pigments are also known (42) and the presence of colouration is not concrete proof of a polymer species (16). As shown in Chapter 6, ochronotic pigment (HGA-derived pigment) can be generated in the laboratory, but it is unclear how

similar pigment derived from purified synthetic HGA is to pigment observed in tissue matrices *in vivo*.

ssNMR provides a means to attempt structural characterisation of the pigmented tissue without extracting the pigment from its native environment. However, a previous ssNMR study on AKU cartilage tissue only revealed a global increase in disorder in pigmented cartilage compared to non-pigmented cartilage from the same patient (336). No distinct signal was identified as the HGA-derived pigment species specifically, possibly due to the low level of pigment species relative to the rest of the collagen matrix, thus preventing the observation of a pigment species due to low sensitivity. With the recent developments in DNP in MAS ssNMR (193), whereby the observed signal can be enhanced by factors of one to two orders of magnitude, it is attractive to revisit the ssNMR approach for studying intact AKU tissues.

It can be challenging to obtain sufficient sensitivity on non-isotopically enriched organic and biological samples. However, DNP-enhanced ssNMR has been successfully applied to non-isotope enriched (unlabelled) organic materials (337,338) and crystalline biological materials (339,340) to yield well-resolved 2D spectra. In terms of application to heterogeneous biomaterials, DNP-enhanced ssNMR has enabled ^{1}H-^{13}C HETCOR experiments on bone (341). I therefore expected to obtain 2D spectra on natural abundance cartilage samples in the following studies, which could potentially resolve signals that are normally overlapping in 1D spectra.

The aim in the following studies was therefore to employ ssNMR in order to better understand ochronotic pigment as it occurs *in vivo* in cartilage tissues. The specific objectives were as follows: to investigate how HGA or the HGA-derived pigment binds into the collagen matrix, to determine the chemical identity of the pigment species, and to understand the effect of such binding on the collagen matrix at an atomic level. Identifying the lengthscale at which the HGA-derived pigment induces disruption of the collagen matrix, whether at the fibre, fibril or triple helical level, will be key to understanding the mechanism by which HGA binds to and reacts with tissue, and will improve our understanding of disease progression in AKU. Understanding the mechanism of binding may illuminate new treatment strategies, not only to prevent or arrest the ochronotic process, but also potentially to find alternative therapeutic approaches for patients already with long-term exposure to high levels of HGA and a high degree of cartilage ochronosis.

Since the most debilitating symptom of AKU is a chronic and accelerated form of OA, potential structural changes to collagen within OA cartilage matrix were also investigated, in order to determine whether such disruptions occur in a similar manner. This comparison was designed to test the hypothesis that AKU is an extreme form of OA in which similar molecular changes occur. This insight is essential in supporting the study of AKU as a model for common idiopathic OA, potentially leading to the discovery of novel biomarkers of OA, for which there is a very pressing need, or suggesting methods to ameliorate and potentially improve the prognosis for patients with AKU and OA.

7.2 Methods

7.2.1 Human cartilage tissue collection

Human joint tissue was obtained at the time of surgery from the knee of a 55-year-old male previously diagnosed with AKU, and the knee of a 71 year old female with OA. Both patients underwent a complete knee replacement due to osteoarthropathy. All cartilage tissue was obtained as surgical waste under ethical approval from the Liverpool Research Ethics Committee with informed consent from patients with diagnosis of AKU or OA undergoing joint replacement. For the AKU specimen, macroscopically pigmented and non-pigmented sections of articular cartilage were dissected from the same tissue sample. For the OA specimen, samples were taken from regions of articular cartilage that were visually worn, 'damaged' tissue, and also from regions of articular cartilage that were not visually damaged, referred to as 'control' tissue. Samples were dissected using a No. 10 disposable scalpel (Swann-Morton, UK) and washed with sterile PBS before storage at -80°C. Further details are provided in Table 7.01.

Table 7.01. Details of human cartilage samples analysed by ssNMR.

NMR sample	Pigmentation	Fixation	Soaking after fixation	Disease	Gender	Age (years)
1	Pigmented	Formalin	No			
2	Pigmented	Formalin	Yes	AKU	Male	55
3	Non-pigmented	Formalin	No			
4	Non-pigmented	Formalin	Yes			
5	Non-pigmented	Formalin	Yes	OA	Female	71

7.2.2   Mouse tissue collection

The distal condyles of the femur, which are covered by cartilage, were taken from *HGD*$^{-/-}$ and wild type mice using a No. 10 disposable scalpel (Swann-Morton, UK). Due to the small size of the condyles, separation of cartilage from the underlying bone was not attempted. Cartilage of the *HGD*$^{-/-}$ mice is microscopically pigmented, while cartilage of wild type mice is not pigmented. Tissue from BALB/c and C57BL/6J mouse strains were combined in both of the above groups to give sufficient material for rotor packing. Samples were washed in sterile PBS and stored at -80°C unless intended for subsequent fixation. Further details are provided in Table 7.02.

Table 7.02. Details of mouse tissue samples analysed by ssNMR.

NMR sample	Genotype	Fixation	Strain	Sex	Age (months)
6	Wild type	None	BALB/c	Female	7
			BALB/c	Female	7
7	Wild type	Formalin	C57BL/6J	Female	15
			BALB/c	Female	7
8	*HGD*-/-	None	C57BL/6J	Male	9
			C57BL/6J	Male	9
9	*HGD*-/-	Formalin	C57BL/6J	Male	9
			BALB/c	Male	16

7.2.3 Tissue fixation

Mouse tissues underwent fixation immediately following dissection. Human tissues were removed from -80°C storage and thawed prior to fixation. Tissues were soaked in twice the tissue volume of 10% formalin (pH 7.4) for 48 h with gentle agitation. The formalin was changed after 24 h. After 48 h, tissues were removed from the formalin and stored short-term at 4°C in tubes containing paper towel soaked in formalin to prevent drying.

In order to remove the large NMR signal from the formalin fixative, some of the fixed cartilage was washed by rinsing in distilled water, then soaked in distilled water with one change of water for up to one week prior to the NMR experiments.

7.2.4 Formation of pigment from synthetic HGA

A solution of pigment (referred to as HGA-derived pigment) was formed from synthetic HGA as described in Chapter 6 (section 6.2.1). In this study, 30 µL NaOH (5 mol/L) was added to raise the pH, and an immediate darkening of the solution was observed which progressed over the total incubation period of approximately 4 months. A brittle brown solid was obtained following drying under a stream of dry nitrogen gas.

7.2.5 Rotor packing

No grinding or cryomilling procedures were used on the biological samples. The mouse condyles were used without further dissection, while the cartilage samples were cut into smaller pieces prior to rotor filling. The samples were gently dried by a combination of two techniques: brief blotting using a paper towel, or brief pressing between two layers of parafilm to expel water, which was then removed with a fine pipette. No freeze drying was used. For all NMR experiments, 3.2 mm outer diameter zirconia rotors with Vespel caps were used. The small pieces of human cartilage were placed directly into the rotor with slight compaction if necessary. For the mouse tissue, condyles from the different mice had to be pooled within each group to provide sufficient sample mass.

Due to the difficulty of transferring small amounts of the synthetic HGA-derived brown solid into the rotor, the solid was dissolved in a small amount of water. The brown solution was transferred to the rotor, then a stream of dry nitrogen gas was used to remove the water. For DNP experiments, AMUPol (342), bcTol (343), and bcTol-M (344) were prepared as solutions in D_2O only, to a concentration of 20 mmol/L for AMUPol and 40 mmol/L for bcTol and bcTol-M. A range of radicals was used due to development of new and improved radicals over the course of the project. Depending on the tissue mass, 5-7 µL of the radical solution was added to the filled rotor. No glycerol was used as it was previously shown to bind to native collagen and modify its stability (345). No H_2O was added; as the samples were well-hydrated, a sufficient amount of H_2O was expected to remain even after the gentle drying process. The procedure was based on the incipient wetness impregnation technique that is used widely in DNP NMR studies of various materials (346,347).

After adding the radical solution, the packed rotors were centrifuged to distribute the radical solution throughout the sample, and then left to equilibrate overnight at 4°C before the start of DNP experiments.

7.2.6 ssNMR spectroscopy experiments

NMR experiments were performed on Bruker (Germany) Avance III wide-bore spectrometers operating at static magnetic fields of 9.4 T and 14 T, with ^1H Larmor frequencies of 400 and 600 MHz respectively. The 600 MHz spectrometer was equipped with a 3.2 mm (^1H-^{13}C-^{15}N) triple resonance probe, with the MAS rate

set to 13333 Hz, and the sample temperature set to 265 K (actual temperature found to be 285 K after calibration) using a Bruker cooling unit with a gas flow of 1200 L per h. The 400 MHz spectrometer was equipped for DNP enhancement experiments, with a gyrotron operating at 9.7 T and a 3.2 mm (^{1}H-^{13}C-^{15}N) triple resonance low temperature MAS (LTMAS) probe. The MAS rate was set to 8889 Hz, and the sample temperature set to 100 K (actual temperature estimated to be 110 K) using the LTMAS cooling cabinet with a variable temperature gas flow of 2000 L per h.

NMR parameters on the 400 MHz DNP spectrometer were: ^{1}H 90° pulse length 2.5 μs, ^{13}C 90° pulse length 3.8 μs, 88 kHz ^{1}H decoupling. DNP-enhanced experiments were carried out under 20-30 mA microwave irradiation. NMR parameters on the 600 MHz spectrometer were: ^{1}H 90° pulse length 3.25 μs, ^{13}C 90º pulse length 5 μs, 73 kHz ^{1}H decoupling.

^{1}H-^{13}C CP experiments are widely used in ssNMR. This technique is based on dipolar coupling and therefore specifically detects immobile and solid species. 600 μs ^{1}H-^{13}C CP contact time was used on the 400 MHz DNP spectrometer. 1000 μs ^{1}H-^{13}C CP contact time was used on the 600 MHz spectrometer. A ramped shape was used on the carbon channel during the CP contact to broaden the matching condition (348).

^{1}H-^{13}C refocused insensitive nuclei enhanced by polarisation transfer (INEPT) (349,350) experiments are a common technique for spectral editing in ssNMR, specifically for detecting mobile and soluble species (351) as the transfer is based

on J-coupling, which is not averaged to zero under MAS, and the relatively long delays lead to broadening of solid signals from T_2 relaxation. This experiment was only carried out on the 600 MHz spectrometer at 285 K and not carried out at DNP conditions. The INEPT delay was 1.5 ms and the refocusing delay was 0.9 ms.

^{1}H-^{13}C 2D HETCOR spectra were carried out with an FSLG sequence during ^{1}H evolution (352,353) to improve resolution, with a square shape applied on both ^{1}H and ^{13}C during the CP contact step. On the 400 MHz spectrometer, under DNP conditions, CP contact times of 20–100 µs were recorded. For some samples additional contact times of 10 and 500 µs were also recorded (as detailed in the results section). Each HETCOR experiment was signal averaged for 5-10 h. On the 600 MHz spectrometer, for each human cartilage sample, two experiments with contact times of 500 µs and 1000 µs were recorded, with each experiment taking 27-47 h.

7.2.7 NMR spectra processing

For 1D ^{13}C spectra, 1024 points of the experimental FID (12.5 ms) was zero filled to 8192 points. A Gaussian window function was applied, with 10 Hz line broadening and a Gaussian max position (GB in the Bruker Topspin program) of 0.1, prior to Fourier transform. Baseline correction was applied to the whole spectral range of -100 to 300 ppm. ^{13}C spectra were externally referenced to tetramethylsilane to enable comparison to the chemical literature of polyquinones. However, this referencing method leads to a small (2.7 ppm) discrepancy of the

chemical shift values of any protein signals from those found in the literature of other protein NMR studies, which are often referenced to 2,2-dimethylsilapentane-5-sulfonic acid DSS. Previous collagen NMR chemical shifts were reported by referencing to TMS.

For 2D ^{1}H-^{13}C spectra, 1024 points of the direct dimension (^{13}C, 12.5 ms) was zero filled to 8192 points, and 96 points of the indirect dimension (^{1}H, 5 ms) was zero filled to 512 points. As is common for FSLG experiments, the ^{1}H dimension scaling factor was empirically determined. A Gaussian window function was applied to the direct dimension, with 30 kHz line broadening and a Gaussian max position (GB in Bruker Topspin) of 0.1. For the indirect dimension, two processing regimes were used: for plots of the carbon aromatic region, a Gaussian window function was applied with 20 Hz line broadening and a Gaussian max position of 0.1; for plots of the glycine Cα region, a sine bell squared function was applied with a sine bell shift of 4. Baseline correction was applied to the full spectral range.

7.3 Results

7.3.1 *Enhancement of ssNMR signals from cartilage samples by DNP*

Employment of DNP successfully achieved enhancement of the NMR signals obtained from both human and mouse cartilage. Enhancement was observed for both formalin-fixed and unfixed tissue. Under DNP, the 1D ^{13}C CP ssNMR spectra of human cartilage sample showed enhancement (ε) values of 22.5 - 35.4, and mouse cartilage spectra showed ε values of 12.5 - 20.3 (Table 7.03).

Table 7.03. Samples investigated in this study by DNP-enhanced ssNMR, and enhancement values (microwave on/off intensity ratio) obtained from CP-MAS ssNMR ^{13}C 1D spectra. 256 transients were acquired per 1D spectrum for human samples, while 3072 transients were acquired for mouse samples. The choice of radicals used, including AMUPol (342), bcTol (343) and bcTol-M (344) was based on improved radicals being available through the course of the study.

Sample source	Tissue	Treatment	Radical	Enhancement
Human AKU patient	Non-pigmented articular cartilage	Formalin fixation	bcTol	22.5 ± 1.8
Human AKU patient	Pigmented articular cartilage	Formalin fixation	bcTol	35.4 ± 5.4
Human OA patient	Osteoarthritic articular cartilage	Formalin fixation	bcTol-M	25.5 ± 3.5
Mouse WT (unpigmented)	Femoral condyles	No treatment	bcTol	12.5 ± 5.2
Mouse HGD$^{-/-}$ (pigmented)	Femoral condyles	No treatment	AMUPol	14.3 ± 5.3
Mouse WT (unpigmented)	Femoral condyles	Formalin fixation	bcTol	20.3 ± 12.3
Mouse HGD$^{-/-}$ (pigmented)	Femoral condyles	Formalin fixation	AMUPol	16.1 ± 4.5
HGA-derived pigment	-	Drying	bcTol	0

7.3.2   The effects of fixation on ssNMR spectra of cartilage tissues

Analyses were performed on cartilage tissue from a patient with AKU and from another patient with idiopathic OA. The osteoarthritic cartilage served firstly as a non-pigmented control sample, and secondly enabled comparison to AKU pigmented cartilage at a molecular level. All human tissues were formalin-fixed to

reduce possible hazards of transmissible diseases. However, tissue fixation may introduce structural changes that could potentially confound any effects due to pigmentation, though previous reports suggest that protein secondary structure is not affected by formalin fixation (354). Therefore, a comparison of mouse tissue from wild type and $HGD^{-/-}$ mice (as a model for AKU) (50), with and without fixation, was first carried out. The aim of these experiments was to distinguish between the effects of fixation on NMR spectra and any effects of cartilage pigmentation.

The main effect of fixation on ^{13}C NMR spectra acquired under DNP conditions appears to be a large signal at 82.2 ppm, and a smaller signal at 84.5 ppm. These signals can be clearly observed in Figure 7.01A, which shows the comparison of fresh, unfixed mouse tissue with fixed mouse and human tissue. The occurrence of these signals in the 70-100 ppm region associated with hydroxylated carbons is due to the formation of methylene glycol and polyoxymethylene glycol from formaldehyde under aqueous conditions (355,356). As shown in Figure 7.01B, the signals from residual methylene glycol were eliminated following soaking of the samples in distilled water in order to remove excess fixative.

A 1D ^{13}C DNP NMR, 8889 Hz MAS, 110 K

mouse, fresh
-non-pigmented
-pigmented

fixative signals

mouse, fixed
-non-pigmented
-pigmented

fixative signals

human, fixed, not soaked
-non-pigmented
-pigmented

180 160 140 120 100 80 60 40 20 ppm

B 1D ^{13}C NMR, non-pigmented human cartilage

soluble fixative @ 285K

INEPT @ 285K
soluble/mobile
-not soaked
-soaked

adsorbed fixative

CP @ 285K
solid/immobile
-not soaked
-soaked

adsorbed fixative   frozen fixative

CP DNP @ 110K
-not soaked
-soaked

180 160 140 120 100 80 60 40 20 ppm

C

OH
1
2
6 7
3   8 OH
5
4   O
OH

homogentisic acid (HGA)
colourless

oxidation →

HGA-derived pigment
dark brown

1D ^{13}C NMR, autoxidized HGA-derived pigment
and pigmented human cartilage

C_2, C_3, C_5

C_1, C_4

C_6

C_7

C_8

CP @ 110K
HGA-derived
pigment

180 160 140 120 100 80 60 40 20 ppm

D

formaldehyde

methylene glycol/
polyoxymethylene glycol

$H_2O + n$

No NMR signal
observed

84.5 ppm (dissolved)
82.2 ppm (adsorbed)

E

adsorbed fixative

CP @ 285K
-not soaked
-soaked

frozen fixative

adsorbed fixative

CP DNP @ 110K
-not soaked
-soaked

180 160 140 120 100 80 60 40 20 ppm

318

Figure 7.01. ^{13}C 1D spectra of various samples analysed in the study, illustrating the effect of fixation. Signals arising from the fixative are highlighted by yellow boxes. **A**: DNP-enhanced CP-MAS ssNMR spectra of fresh/unfixed mouse tissue (top), fixed mouse tissue (middle) and fixed human tissue (bottom). Black spectra represent non-pigmented tissue, while red spectra represent tissue that was pigmented macroscopically (human tissue) or microscopically (mouse tissue). **B**: non-pigmented human cartilage NMR spectra at 285K (no DNP enhancement, 600 MHz spectrometer) and 110K (with DNP enhancement, 400 MHz spectrometer). Of the six spectra, the top two were INEPT experiments, the bottom four were CP experiments. Cartilage samples that were not soaked (black traces) showed a signal arising from the fixative, which disappeared after soaking (green traces). **C**: spectrum from autoxidized synthetic HGA-derived pigment (right) with each signal assigned to a carbon on the HGA structure (left). **D**: reaction of formaldehyde with water to give methylene glycol and polyoxymethylene glycol. **E**: pigmented human cartilage NMR spectra at 285K (no DNP enhancement, 600 MHz spectrometer) and 110K (with DNP enhancement, 400 MHz spectrometer).

From the spectra shown in Figures 7.01B and 7.01E, it is evident that significant inhomogeneous line broadening occurred under DNP conditions, to the extent that there appears to be little to no difference between the NMR spectra of pigmented and non-pigmented cartilage. Comparing the CP spectra obtained at 285K (Figures 7.01B and 7.01E), the pigmented and non-pigmented cartilage do show a difference in linewidth under conventional ssNMR conditions as previously demonstrated (336), and this difference was maintained after soaking both samples. This observation indicates a loss of motional averaging and dynamics of the cartilage tissue under DNP conditions.

It was important to show next that methylene glycol signals generated by formaldehyde did not interfere with the potential observation of any signals representing HGA-derived pigment in human AKU cartilage. For this purpose, the ssNMR spectra of pigment formed from synthetic HGA was studied. As shown in Figure 7.01C, it is clear that the signals that arise from fixation are well clear of the ^{13}C signals found in the pigment. Therefore, it was concluded that the presence of the methylene glycol signals would not hinder the potential observation of pigment signals in these analyses.

7.3.3 SsNMR observation of HGA-derived pigment in cartilage

Both the structure of HGA and the spectra of the synthetic HGA-derived pigment (Figure 7.01C) suggests that any pigment signals should be most intense in the aromatic region, from 100-150 ppm in the ^{13}C NMR spectrum. As shown in Figure 7.02, this spectral region in the DNP-enhanced ^{1}H -^{13}C HETCOR 2D experiments was also the least crowded. In these spectra, the

signals represent ^1H and ^{13}C nuclei whose spins are associated; *i.e.* indicative of ^1H and ^{13}C nuclei that are directly or indirectly bound within a chemical structure. Consistent with the above predictions, clear differences were observed between the spectra of pigmented and non-pigmented human AKU cartilage in the aromatic region, also indicated in Figure 7.02.
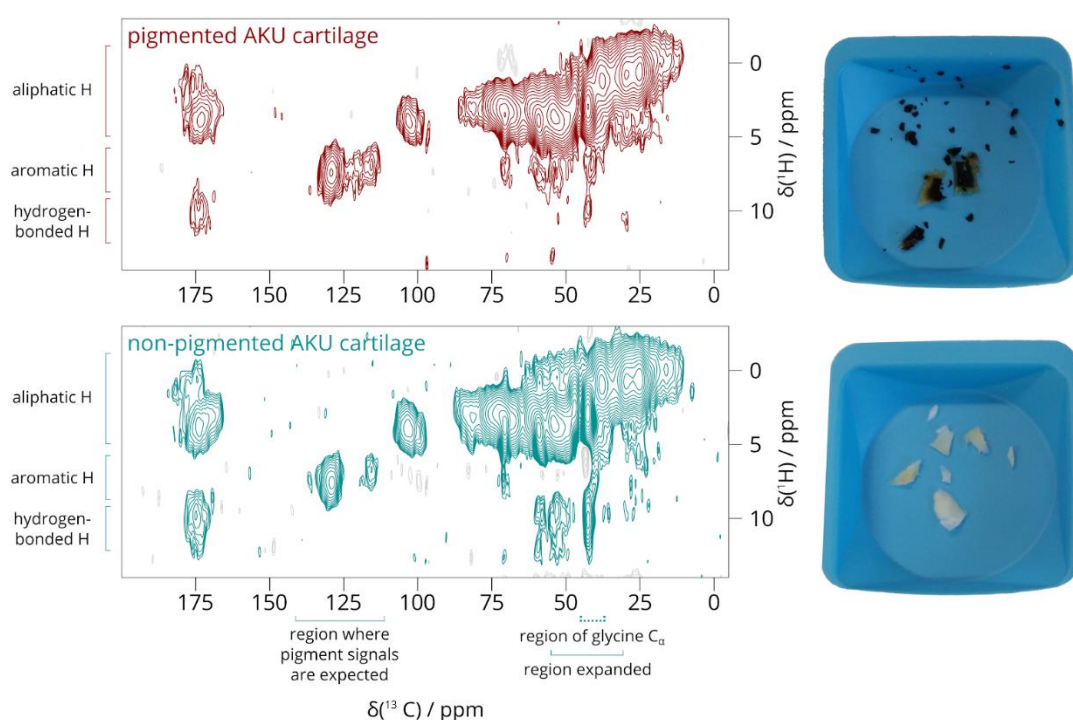
Figure 7.02. 2D ^1H-^{13}C FSLG HETCOR DNP-enhanced ssNMR spectrum of pigmented (brown spectrum) and non-pigmented (green spectrum) human AKU cartilage. The ^{13}C NMR range of 110-140 ppm is the region where pigment signals are expected (expanded in Figure 7.03). The glycine C_α ^{13}C signal is centred at 42.5 ppm (expanded in Figure 7.04). Photographs of the respective cartilage samples from which these data were obtained are shown to the right of each spectrum; note the striking dark appearance of the pigmented sample. Both spectra were acquired with a contact time of 50 µs.

For further analysis of HGA and its derivatives in the human AKU cartilage samples, a series of ^{1}H-^{13}C HETCOR experiments with different CP contact times was performed (Figures 7.03 - 7.05). Increasing CP contact time led to an increase in the number of signals, with signals from more distal ^{1}H and ^{13}C pairs emerging at longer CP contact times (Figure 7.05).
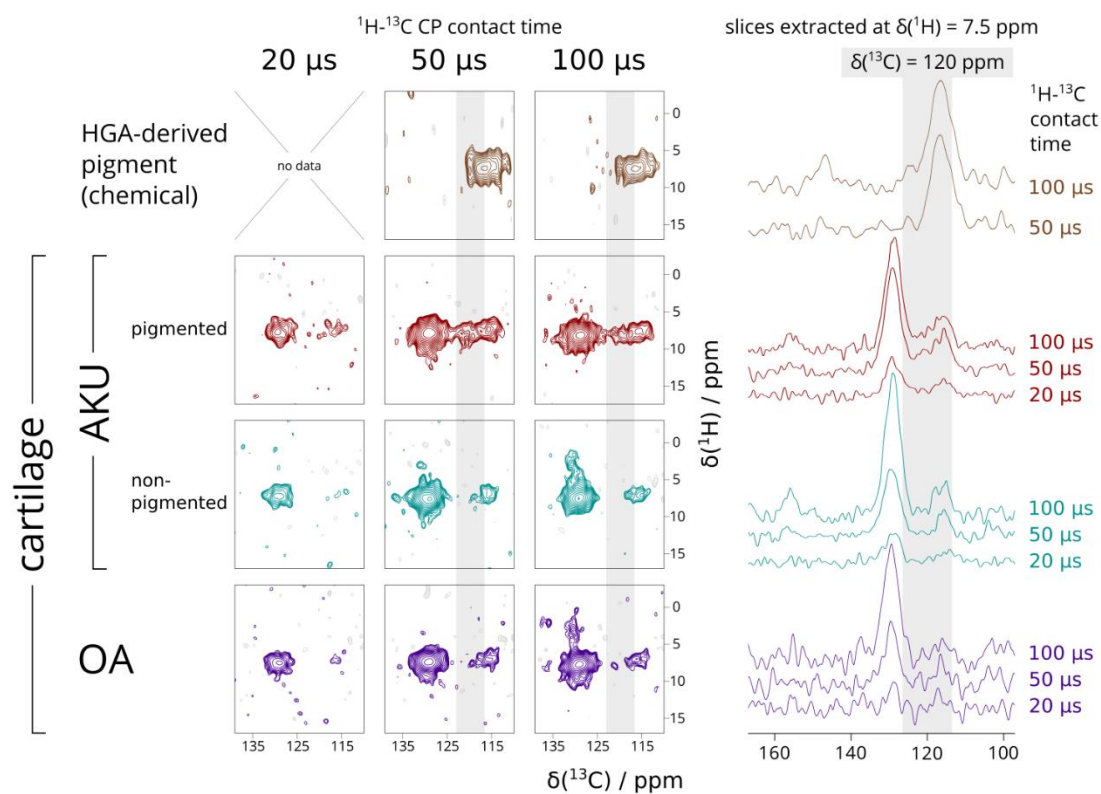
Figure 7.03. Similarity between the 2D ¹H-¹³C FSLG HETCOR ssNMR spectra obtained from synthetic HGA-derived pigment and pigmented AKU human cartilage. These spectra were obtained under DNP conditions at 20-100 µs contact times, and the 110-140 ppm region is shown. Grey shading highlights the aromatic ¹³C signals at 120 ppm, which were present in the spectra from synthetic HGA-derived pigment (brown, top row) and pigmented AKU human cartilage (red, second row) but not those from non-pigmented AKU human cartilage (green, third row), and OA human cartilage (purple, bottom row). On the right, slices extracted from the 2D spectra highlight (grey) where the HGA-derived pigment and pigmented AKU cartilage show signals in the same ¹³C chemical shift range, which are weaker in non-pigmented AKU cartilage and not observable above noise in OA cartilage. An overlay of the extracted slices (50 µs) is shown in Figure 7.04.
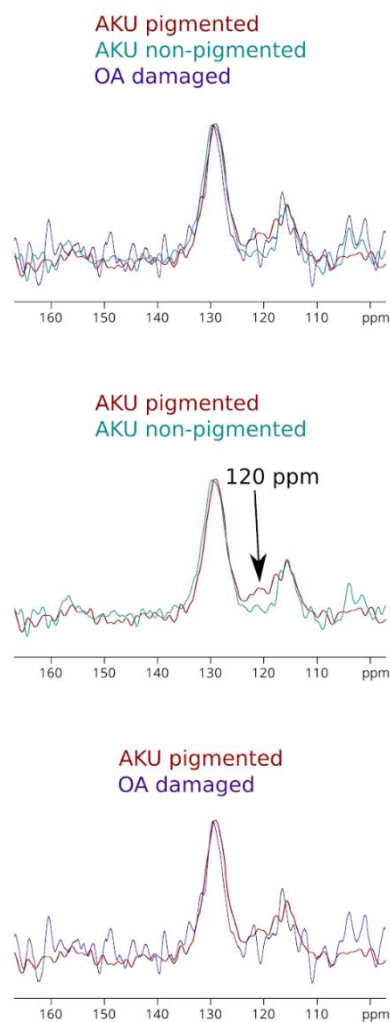
Figure 7.04. ^{13}C spectra extracted from the 2D ssNMR ^{1}H-^{13}C FSLG HETCOR DNP-enhanced ssNMR spectra. Spectra highlight the increased signal at 120 ppm that was attributable to ochronotic pigment in the pigmented AKU cartilage sample. The spectra shown were from pigmented AKU human cartilage (red), non-pigmented AKU human cartilage (green) and OA human cartilage (purple) at 50 μs contact times. To aid visual comparison, the top shows all three overlaid, the middle and bottom show just two out of the three same spectra overlaid. The vertical height of the largest signal in each spectrum were scaled to match each other. The AKU pigmented cartilage spectrum shows a broadening/smearing of signal intensity at 120 ppm that is clearly absent from the AKU non-pigmented cartilage, and is also weak or absent from the OA damaged cartilage though it is harder to gauge due to the poor signal to noise of the OA cartilage spectrum even with DNP enhancement.

In Figure 7.03 it is evident that the 20 μs contact time HETCOR spectra of all cartilage samples (furthest left column) have a broad ^{13}C signal at 130 ppm that corresponds to aromatic residues such as tyrosine, phenylalanine and histidine. In pigmented cartilage (Figure 7.03, second row), there is an additional ^{13}C signal at 120 ppm that is not present in the non-pigmented AKU cartilage nor the non-AKU OA cartilage. This 120-ppm pigmented cartilage signal corresponds to signal intensity that is also present in the spectrum of the synthetic HGA-derived pigment (Figure 7.03, top row), and this signal was therefore assigned to the aromatic carbons in the HGA-derived pigment. Despite the differences in this area of the spectra, overall, the spectra from cartilage are still dominated by collagen signals.

In the HETCOR experimental series expanded in Figure 7.05, the contact time was increased incrementally from 50-500 μs. These experiments were designed to probe the different chemical structures represented by the signals, as their intensity over differing contact times reflects the distance between the ^{1}H and ^{13}C pairings. As the contact time increased, more signals were observed in the spectra from synthetic HGA-derived pigment (dark brown). The signal at ^{13}C 116.8 ppm correspond to carbons directly bonded to a proton in the aromatic ring and is present in all spectra. In the 100 and 500 μs contact time spectrum (Figure 7.05, 2 furthest-right rows), an additional signal emerges at ^{13}C 148.5 ppm, corresponding to carbons that are directly bonded to a hydroxyl group (i.e. a phenyl carbon), where the rate of signal build-up is distinctly slower than the 116.8 ppm signal as the distance to the nearest proton is greater (Figure 7.05, lower portion). The signal at 148.5 ppm therefore indicates presence of a hydroquinone structure in the synthetic HGA-

derived pigment as opposed to an oxidised quinone, for which the ^{1}H and ^{13}C pairs would be less distal.
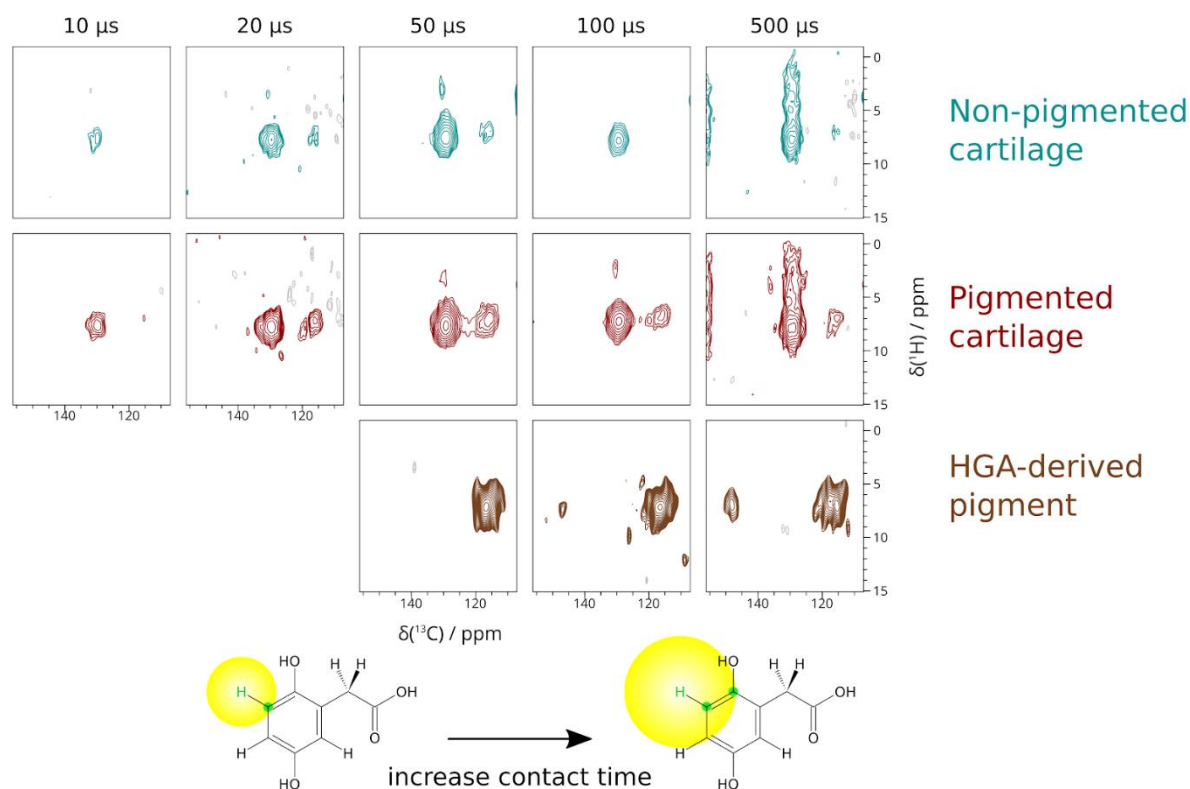
Figure 7.05. ^{1}H-^{13}C FSLG HETCOR ssNMR spectra, showing emergence of an additional signal at 148.5 ppm for pigment derived from synthetic HGA. The spectra were obtained from unsoaked pigmented AKU human cartilage (top), non-pigmented AKU human cartilage (middle) and pigment derived from synthetic HGA (bottom) under DNP conditions at various contact times from 10 µs - 500 µs; the signal at 148.5 ppm emerges with increased contact time. Only the [13]C 105-155 ppm range is shown.

7.3.4 Change in collagen structure of pigmented cartilage

It is interesting to consider the effects of pigmentation on cartilage collagen structure. Even though pigment levels are low, there are profound material differences between non-pigmented and pigmented AKU cartilage. Apart from its dark colour, ochronotic cartilage is stiffer and resistant to enzymatic degradation (306), suggesting changes in collagen structure throughout the cartilage tissue. Particular attention was therefore paid to the glycine Cα NMR signals as a marker of collagen structural integrity more generally. These signals were combined into a single ^{13}C peak at 42.7 ppm, well-separated from other Cα signals. In NMR spectra of tissues with a high proportion of collagen, such as cartilage, this signal is one of the most intense due to the fact that glycines occur every third residue along each collagen chain. In Figure 7.06A, this spectral region is expanded from the ^{1}H-^{13}C HETCOR spectra acquired at different contact times from the human cartilage samples.
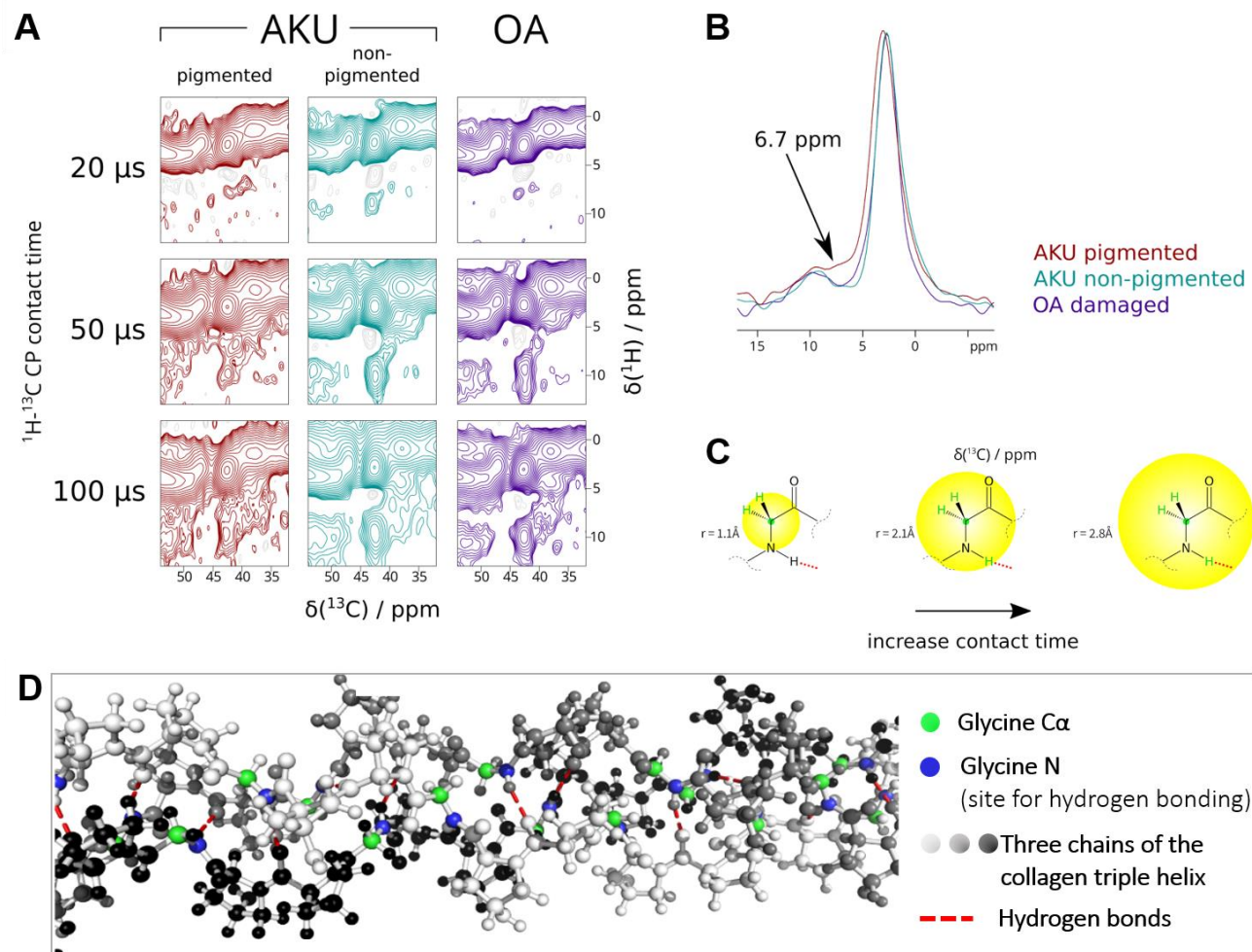
Figure 7.06. Spectra from human cartilage, indicating loss of collagen integrity in AKU pigmented cartilage by alteration to the type-II collagen glycine signal. A: 2D ^{1}H-^{13}C FSLG HETCOR obtained from pigmented AKU human cartilage (brown, left column), non-pigmented AKU human cartilage (green, middle column) and OA human cartilage (purple, right column) under DNP conditions at 20-100 µs contact times, showing only the ^{13}C 30-55 ppm range. Alteration, or 'smearing' of the signals obtained at 42.7 ppm in the ^{13}C dimension is evident for the AKU pigmented cartilage sample. The signal at 42.7 ppm corresponds to the collagen glycine C_α. The signals at 3.3 ppm and 9.6 ppm in the ^{1}H dimension are correlated with the glycine C_α signal and correspond to directly bonded glycine H_α and the and the protons on the amide H_N ('C_α-H_N' signal), respectively. Alteration of these signals in the pigmented sample therefore reflects disruption to the inter-chain hydrogen bonds between glycine residues in the three chains of the type-II collagen triple helix (Figure 7.01D); these changes indicate loss of structural integrity of type-II collagen in ochronotic cartilage more generally. This experiment was repeated on fixed and unsoaked AKU human cartilage samples with 10-100 µs contact times with similar results (Figure 7.07). **B**: overlay of extracted slices from the 2D spectra in Figure 7.06A (from the ^{1}H dimension obtained at 100 µs contact time), highlighting broadening/smearing of the signal at 6.7 ppm observed for the AKU pigmented cartilage sample only. **C**: chemical structure of glycine, highlighting the sites of hydrogens for which build-up of NMR signal was observed with increasing CP contact time (C_α-H_α and C_α-H_N; see section Figure 7.06A). **D**: graphical representation of the type-II collagen triple helix, highlighting the inter-chain hydrogen bonds (red dashed lines) between the amide (N) and C_α of neighbouring glycine residues. These bonds are critical for maintaining the integrity of the triple helix that defines the collagen protein structure.

In Figure 7.06A, for all cartilage spectra obtained at short contact times (top row), the glycine C_α signal at 42.7 ppm mainly correlates to a ^1H signal centered at 3.3 ppm, which correspond to the directly bonded glycine H_α (1.1 Å). As the contact time is increased to 50 μs and beyond, a second clear correlation emerges, centered at 9.6 ppm (^1H) in the spectra of the non-pigmented AKU cartilage (green, middle column) and the OA cartilage (purple, right column). This signal can be attributed to correlations between glycine C_α and the respective protons on the amide H_N (2.1 Å); 'C_α-H_N' signal. This proton occurs at a much higher chemical shift value of 9.6 ppm due to the interchain hydrogen bonding that is a defining feature of all collagen triple helices.

Looking at the same signals in the pigmented cartilage spectra (Figure 7.06A, left column, middle and bottom rows), the glycine C_α-H_N signal is not as well-defined, and appears to be smeared towards lower ^1H chemical shift (ppm) values, which can also be seen in the ^1H spectra extracted at ^{13}C 42.5 ppm (Figure 7.06B). This indicates that the hydrogen bonds that the glycine H_N are participating in are disrupted, suggesting at least partial dissociation of the collagen triple helix in the pigmented AKU cartilage sample (Figure 7.06D).

Finally, as the experiments acquired under DNP conditions exhibited inhomogeneous line broadening, it was thought that better spectral resolution in the glycine region might be obtained by conducting the ^1H-^{13}C HETCOR experiments at room temperature. For this, standard ssNMR was performed at 287 K on a 600 MHz spectrometer. The displayed spectra were acquired over a much longer experiment time (47-h instead of 10-h per experiment) but still showed reduced sensitivity without DNP-enhancement, therefore no clear improvement in spectral line width was achieved. However, in confirmation of

the results from DNP ssNMR, a glycine C_α-amide H_N signal can be observed in non-pigmented cartilage, but is absent from pigmented cartilage (Figure 7.07).
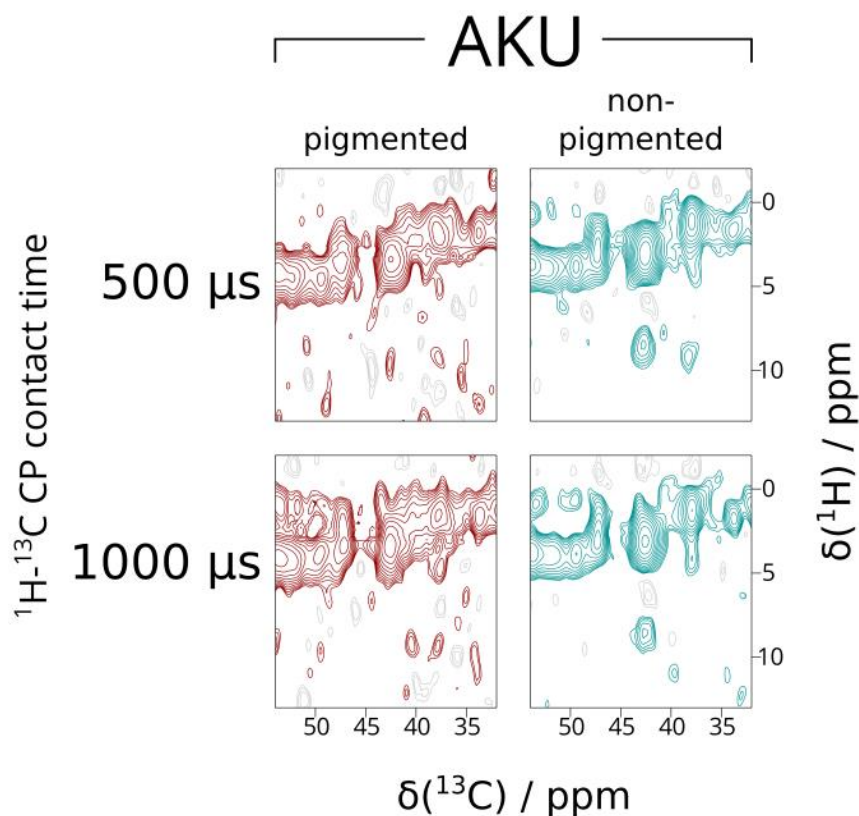
Figure 7.07. ^{1}H-^{13}C FSLG HETCOR ssNMR spectra obtained on human cartilage without DNP-enhancement, corroborating the finding of disruption to the type-II collagen glycine signal. Standard ssNMR was performed at 285 K on a 600 MHz spectrometer in an attempt to improve spectral resolution. These samples were human AKU cartilage; pigmented (brown spectrum) and non-pigmented (green spectrum). The ^{13}C ppm range shown is 30-55 ppm, as with Figure 7.06A. In agreement with the data shown in Figure 7.06A, alteration is observed to the glycine C_α-H_N signal, at 42.7 ppm in the ^{13}C dimension and 9.6 ppm in the ^{1}H dimension; here there is complete absence of this signal from the pigmented sample. These data further support widespread disruption to the type-II collagen protein structure in ochronotic cartilage as indicated by the glycine NMR signal.

7.4. Discussion

7.4.1 *Potential of DNP-enhanced ssNMR for studying fixed tissue samples*

Using a variety of human and mouse tissue samples the data presented here show, to my knowledge, for the first time that 1D and 2D NMR spectra with DNP-enhancement values of 10-40 can be reliably obtained from formalin-fixed tissue using a range of water-soluble biradicals. As chemical fixation is a ubiquitous technique for preserving tissue samples across biological disciplines, these studies highlight the further opportunities for NMR analyses of archival specimens. The structural capabilities of NMR would complement traditional histological approaches and other analytical techniques that have been described for studying fixed tissues, such as magnetic resonance imaging (357–359) and various types of MS (360–364).

Importantly, these results show that fixation did not significantly alter the NMR signals from the protein component of the cartilage tissue, especially under the low temperature DNP enhancement conditions where there is significant inhomogeneous broadening. The main effect in the case of formalin fixation, where no attempt is made to remove the fixative, is the detection of a variable intensity ^{13}C signal at 82.2 ppm due to frozen methylene glycol, and a generally weak signal at 84.5 ppm due to adsorbed methylene glycol. These signals can potentially interfere in the case where the aim is to observe GAG species in the sample, but present little interference for the HGA-derived pigment signals, which was the primary interest in this study. Moreover, it was shown that these methylene glycol ^{13}C signals could be easily removed by rinsing and soaking in water (Figure 7.01B & E).

7.4.2 Similarities between the HGA-derived pigment and polyquinones

One advantage of ssNMR is its ability to provide specific insight into the chemical composition of heterogeneous samples and polymers, such as the solid brown pigment derived from benchtop oxidation of synthetic HGA, and to also provide direct comparison to other known/related substances in the literature (Table 7.04). The synthetic HGA pigment sample had a relatively simple ^{13}C NMR spectrum dominated by four signals (Figure. 7.01C), therefore it seems that the structure of the HGA-derived pigment is simpler than the L-dopa-derived melanins. By comparing the NMR spectra in the current study with previous ssNMR analyses on melanin in fungal cell wall (365), hair samples (366), and squid ink (367), it is clear that the structure of HGA-derived pigment is different from L-dopa and derivatives (such as eumelanin and pheomelanin) but bears similarities to HGA. Since HGA is known to form pyomelanin, the present data are in agreement with prior work suggesting that the AKU pigment is pyomelanin (42). Pyomelanin is produced mainly from bacteria and has thus far, to my knowledge, not been subjected to NMR analysis.

In order to gain insight into the chemical structure of the pigment, the NMR literature was surveyed for likely candidates that align with the observations reported here. Formation of lactone, which was previously reported as an HGA derivative in AKU urine (368), can be ruled out, as this would be expected to give a ^{13}C chemical shift less than 178 ppm. The 181-ppm chemical shift observed here for the synthetic HGA-derived pigment is close to that observed in solution-state NMR at 183.4 ppm (Table 7.04) suggesting that the carboxylate functionality of HGA had not undergone significant change.

The broad signals centred at 116.8 ppm and 148.5 ppm in the ^{13}C ssNMR spectrum of the brown pigment sample are characteristic of carbon atoms situated in phenol-like rings, in agreement with hydroquinone-like or HGA-like ring structures (369–371). If benzoquinone rings are present in the pigment, ^{13}C signals should be observed at 136 ppm and 181 ppm. The former signal was clearly not observed. While a signal at the latter position was observed, it is likely to correspond to the carboxylic acid functionality in HGA rather than the carbonyl of benzoquinone. It cannot be ruled out that benzoquinone-like structures were present within the substance at low quantities, as this sample failed to show any DNP signal enhancement, presence of a minority of benzoquinone-like rings can be expected from the colour of the sample alone. However, it is clear that hydroquinone rings were much more prevalent than benzoquinone rings in the HGA-derived pigment sample. Although it may seem like a conflict that the colour of the substance indicated presence of benzoquinone-like species and the NMR spectra indicated a dominance of hydroquinone-like species, this is in agreement with the common observation that impurities with intense absorption may dominate the overall colour of a substance.

Table 7.04. ^{13}C NMR chemical shifts of the structures relevant to ochronotic pigment, as obtained from literature.

Chemical entity	Carbon (by position)	Respective ^{13}C chemical shift (ppm)
Homogentisic acid (HGA)*	1, 4 2, 3, 5 6 7 8	150.4, 151.6 117.2, 119.8, 120.4 128.3 43.2 183.4
Benzoquinone acetic acid (BQA)†	1, 4 2, 3, 5 6 7 8	187 130-136 136+ 40-50 183
Hydroquinone‡	1,2	151.5, 118.5
Benzoquinone‡	1,2	187.0, 136.4

* Bruker-format data of HGA in aqueous solution was downloaded from the Madison Metabolomics Consortium Database (MMCD) and the referencing taken as is. Note that the chemical shift values reported on the HGA page of the MMCD (http://mmcd.nmrfam.wisc.edu/test/cqsearch.py?cqid=cq_00389) do not seem to match the values from the spectrum downloaded (http://mmcd.nmrfam.wisc.edu/rawnmr/expnmr_00219_3.tar), but the chemical shifts of the spectrum downloaded showed reasonable agreement with the solid-state ^{13}C NMR spectrum of the HGA-derived pigment in this study.

† Experimental values were not found for BQA, likely due to its instability. Expected chemical shift values were derived with reference to benzoquinone and HGA values in the literature.

‡ In addition, similar values can be found in the following references: (370–372).

The following discussion therefore focuses on known benzoquinone- and hydroquinone-related multimeric structures and their properties. Oligomeric benzoquinones and hydroquinones (dimers and trimers) are unstable and frequently undergo further polymerization (372,373). Stable, well-defined dimers and trimers have only been synthesized with the protection of bulky t-butyl groups that prevent further polymerisation (372). By referring to previous solid-state ^{13}C NMR study of benzoquinones and hydroquinones (371) and their well-characterized oligomers (372), the NMR chemical shifts obtained here from the pigment are in agreement with those obtained from hydroquinone rings in monomers and oligomers. From the literature cited here, it is expected that under the benchtop conditions in which the pigment sample was formed from HGA standard solution, the final product is unlikely to be strictly restricted to low molecular weight compounds such as dimers and trimers.

Hydroquinone and benzoquinone are known for their redox properties (374); these chemicals and their polymeric forms have been used as inhibitors against radical polymerization, *e.g.* as antioxidants in rubber (375). Polyhydroquinone, polyquinone and polybenzoquinone are increasingly oxidised polymeric forms, with the intermediate polyquinone being a mix of hydroquinone and benzoquinone rings, and susceptible to both oxidation and reduction. While pure polyhydroquinone can be formed as a white solid by reduction of polyquinone, it is easily reoxidised on exposure to oxygen in air (376). The redox activity of polyquinone would trap and quench any added radical species, thus explaining the inability here to obtain any DNP enhancement on the synthetic HGA-derived pigment sample. The reducing

capability of HGA also explains the negative interference observed in samples with elevated HGA concentration on a range of routine clinical assays that are based on the production of hydrogen peroxide (377). In this respect, the pigment is also chemically similar to bacterial pyomelanin, which has been reported to scavenge free radicals (378). Based on the data discussed here, I propose that the HGA-derived pigment is polyquinone-like, substituted with alkyl carboxylic acids.

7.4.3 Identification of an NMR signal representing HGA-derived pigment in human tissue samples

A fundamental question in these studies was whether the NMR spectra of pigment derived from synthetic HGA bears any resemblance to that of the pigment species present in ochronotic AKU cartilage. Despite the black appearance and brittle texture of ochronotic cartilage (306), the 1D ^{13}C NMR spectra obtained from pigmented and non-pigmented cartilage here and in previous studies (336) are very similar; all significant NMR signals can be attributed to collagen proteins. Clear differences only became apparent with the enhanced sensitivity enabled by DNP in this study, which made it possible to acquire a series of 2D ^{1}H-^{13}C correlation experiments on cartilage samples from patients. In the case of the fixed tissues, when searching for specific signals representing the pigment, it was important to consider potential spectral interference from either residual fixative or the cartilage tissue.

Importantly, the formalin fixative could be eliminated as a source of interference, as it was shown that no fixative signals occurred in the aromatic region of interest at 110-160 ppm (Figure 7.01), where signals representing

HGA or its derivatives would be expected. Of the natural amino acids, phenylalanine and histidine side chain signals would be expected to to overlap with those arising from benzoquinone/BQA, while tyrosine side chain signals would be expected to overlap with hydroquinone/HGA. Fortunately, the type-II collagen of cartilage has a low proportion of aromatic residues: phenylalanine (1.4% of the triple helix), histidine (0.2%), and tyrosine (0.2%). These side chains in the aromatic region should therefore be minor signals in the spectra. Tryptophan is not part of the triple helix and is therefore excluded. Moreover, the non-pigmented AKU cartilage and OA cartilage samples were used as controls to establish the expected relative signal intensities arising from these minor aromatic side chains prior to extracting information specific to the pigment species.

A ^{13}C signal centered at 120 ppm was observed from both the synthetic HGA-derived pigment and the pigmented AKU cartilage and therefore attributable to pigmentation. I acknowledge that in all cartilage samples, ^{13}C signals were observed at 115 ppm; these are likely to be primarily attributable to tyrosine side chains from collagen and other ECM proteins. Since the hydroquinone functionality is similar in structure to the phenol ring in tyrosine, the HGA-derived pigment also shows this signal. The fact that the pigment signal occurs centered at 120 ppm strongly supports the presence of hydroquinone functionality, but not the benzoquinone functionality.

7.4.4 Investigating the sites of disruption in AKU cartilage

Having established the chemical similarities of the synthetic HGA-derived pigment and the pigmented cartilage, I briefly consider the process by which

HGA and derived species can bind to collagen. One possible point in time is during secretion of collagen proteins from the cell. The cytoplasmic environment is reducing (379), while the extracellular space is oxidizing (380), which is likely to promote non-enzymatic oxidation of BQA and other polyquinone species. It is therefore possible that collagen triple helices are vulnerable to binding and disruption only when secreted into the extracellular space, prior to or during fibrillization. The local redox environment may explain observations that pigmentation is localized to individual chondrocytes and their pericellular matrices in the calcified zone of articular cartilage (50) and that early pigmentation follows the periodicity of collagen fibrils (381,382). That said, cartilage is a slow turnover tissue and the osteoarthropathy in AKU is not from birth but only after 30-40 years of life, so it seems unlikely for pigmentation during secretion to be the dominant mechanism in advanced ochronosis.

There are similarities between ochronotic cartilage and the normal ageing of cartilage; in both processes, cartilage becomes stiffer and more resistant to enzymatic degradation. In the normal ageing process, these changes are associated with loss of proteoglycans, and increases in both enzymatic and non-enzymatic cross-linking of the collagen matrix (383–385). In OA, increased ECM stiffness is thought to be an early pathophysiologic event, rendering cartilage susceptible to further damage by loading (385,386). Presumably, the altered physico-mechanical properties of ochronotic cartilage exacerbate progression of osteoarthropathy in AKU, as the altered cartilage no longer supports normal loading due to aberrant transmission. The data

reported in this chapter pinpoint a mechanism of structural disruption that can lead to this type of material change.

The ^{1}H-^{13}C HETCOR spectra showed remarkable changes in triple helix integrity between non-pigmented and pigmented AKU cartilage. The observation that the ^{1}H chemical shift of the glycine amide moves to lower values is particularly interesting, indicating a lengthening and weakening of these hydrogen bonds (387,388). In collagen, where these interchain hydrogen bonds are pervasive, this level of molecular structure change is likely key to the mechanical and functional failures of pigmented cartilage. From the structure of the collagen triple helix, it is expected that glycine residues are buried in the centre of the collagen triple helix and naturally protected against disruptive bindings. I hypothesise that the binding of HGA and/or derived species occurs in regions where the triple helix is under strain. Such strain leads to partial uncoiling or unfolding of the triple helix, exposing the glycine amide site for binding. In the case of AKU, possibly due to the high binding affinity of benzoquinone rings, the usually temporary uncoiling or unfolding of the triple helix becomes exacerbated by the presence of ochronotic pigment and perhaps even leads to an "unzipping" of the triple helix. A similar process could also apply in general OA, where temporary strain or change to the triple helix becomes permanent damage.

While the spectra obtained from OA cartilage was broadly similar to that from non-pigmented AKU cartilage, in the OA cartilage spectrum the most intense part of the glycine H_N signal showed a drift to lower chemical shift (Figure 7.06). A possible interpretation is that in OA cartilage, the collagen triple helix degradation is more than that in non-pigmented AKU cartilage, but less than

that in pigmented AKU cartilage. This supports the idea that common, age-related cartilage degeneration not only precedes ochronosis in AKU but also acts as a pathophysiological event that renders cartilage susceptible to ochronosis at specific sites (9). However, without ^{13}C enrichment, it is difficult to correlate pigment and collagen proteins signals to directly identify a potential specific binding site for AKU-derived species within the cartilage matrix.

7.5   Chapter 7 conclusion

The data presented in this chapter support that the process of ochronosis in AKU causes widespread molecular disorder to the cartilage collagen matrix. This fundamental alteration to the triple helical structure of collagen is likely to underlie its mechanical failings responsible for severe osteoarthropathy in AKU. Moreover, the comparison of AKU to OA cartilage shows that similar, but milder molecular alteration occurs in OA. This suggests that the onset of ochronosis in AKU, which is not usually observed until the third-fourth decades, is facilitated by common, age-related osteoarthritic degradation of cartilage. Currently, there is still no biomarker that accurately predicts the onset of OA. In AKU, development of osteoarthropathy is both predictable and inevitable, therefore providing a unique opportunity to model early-stage osteoarthritic changes to cartilage. Together, the data highlight that the study of a rare disease can provide valuable lessons about widespread, degenerative ageing processes.

Acknowledgements

8.0   OVERALL CONCLUSION

The work presented in this thesis represents the most comprehensive biochemical investigation of AKU to date. Metabolomic analyses revealed for the first time the wider biochemical consequences of a) targeted *HGD* gene deletion in mice and b) nitisinone treatment beyond the tyrosine metabolic pathway. Studies on ochronotic pigment suggested for the first time a) a specific compound structure from ochronotic pigment formed by aggregation of HGA-related species and b) alteration to a specific site in the cartilage matrix, which is likely to explain susceptibility to ochronosis.

The biochemical alterations observed in the metabolomics experiments reflect a complex, inter-connected network of metabolites across multiple metabolic pathways. Many of these changes were unexpected, both at the level of individual metabolites and the pathways involved. These data therefore offer an array of novel insights into AKU pathology, including potential biomarkers of OA progression. Mechanistic evaluation of the metabolite changes observed is beyond the scope of this thesis, but the wealth of novel insights that these experiments have provided will stimulate a plethora of further research projects in AKU. At the heart of these projects will be the clinical implications of the changes observed.

The analytical data acquired from ochronotic pigment have significantly advanced our knowledge of its chemical properties. Ochronotic pigment represents the major knowledge gap in AKU at present. Ochronosis is central to the devastating consequences of the disease, yet its chemical structure and the mechanism through which it binds to cartilage is not known. The analyses

carried out on pigment produced from HGA *in vitro* fundamentally challenge the idea that ochronotic pigment is a polymer, and instead support the aggregation of a combination of HGA and BQA by weaker, non-covalent interactions. The data from in *vivo* studies indicate that ochronosis is accompanied by widespread alteration to the structure and organisation of the cartilage matrix. The observed alteration to the glycine signal in ochronotic cartilage, an abundant amino acid residue in type-II cartilage collagen, potentially pinpoints the specific structural alteration that enables pigment deposition.

8.1   Next steps

The metabolomic analysis workflow described here is a system for comprehensive metabolic phenotyping more generally. The strategy will be an invaluable tool to AKU researchers going forward and has great potential to be used in a clinical context more broadly. For AKU, the SONIA-II clinical trial (139), which finished recently, is a particularly exciting opportunity to further deploy this technique. SONIA II is an international, multicentre randomised control clinical trial including the largest AKU patient cohort studied to date. This cohort provides an opportunity to independently validate the metabolite changes observed here on nitisinone and the metabolite correlations with clinical parameters, for example measures of osteoarthropathy disease severity. Furthermore, the increased size of the SONIA II patient cohort compared with the NAC cohort studied here will enhance statistical power, potentially revealing further associations between metabolite profiles and clinical parameters.

Further work will also aim to gain mechanistic insights into the metabolomic alterations observed. There are a number of potential hypotheses that could explain the surprising widespread network of changes observed here as a consequence of both nitisinone treatment and targeted *HGD* disruption in untreated AKU. A series of further metabolic flux experiments have been planned in mice using stable isotopically labelled compounds (HGA and tyrosine), with the aim of identifying the whole range of metabolites that are increased due to being derived from the elevated tyrosine and HGA observed in treated and untreated AKU respectively.

An alternative hypothesis is that some of the metabolite changes observed are due to the indirect effects of compounds that are markedly elevated in treated or untreated AKU. For example, it has been shown that the enzyme tryptophan transaminase, which converts tryptophan to indoxyl sulfate (decreased here in mice on nitisinone) via indolepyruvate, has increased affinity for tryptophan in the presents of keto acids such as HPLA which are markedly increased on nitisinone (389). In contrast, it has been suggested from HT-I that increased tyrosine, as a result of nitisinone therapy, inhibits the activity of tryptophan hydroxylase, resulting reduced conversion of tryptophan to serotonin (236). It is therefore conceivable that elevated HGA in untreated AKU, and elevated tyrosine and/or keto acids in nitisinone-treated AKU impact upon the activity of other enzymes in tyrosine and neighbouring metabolic pathways. The potential for these markedly elevated compounds to act upon specific enzymes within neighbouring metabolic pathways could explain the complex network of metabolite alteration beyond those increases that are likely to be derived directly from HGA or tyrosine.

Enzyme inhibition or enhancement attributable to increased concentrations of such metabolites could be studied in a series of *in vitro* experiments. These experiments could involve tissue homogenates, cell cultures or isolated enzymes in the presence of varying concentrations of the above metabolites, with enzyme activity indicated by the rates of metabolite substrate conversion and product formation. A 'systems' approach, whereby metabolomics data are combined with data from other 'omics modalities such as transcriptomics and proteomics, could be particularly informative. Such an approach could investigate potential associations between metabolite levels and the expression of individual enzymes of tyrosine metabolism and neighbouring metabolic pathways.

Nitisinone is a very promising treatment for AKU, but it is important to note that its mechanism of action is not ideal; its inhibition of HPPD essentially shifts the metabolic 'block' observed in the tyrosine pathway in AKU further upstream of HGA. The gold-standard therapy for AKU is restoration of HGD activity in the liver by enzyme replacement or gene therapy, for example; such approaches have shown promise in other inborn errors of metabolism, including monogenic disorders of the liver (390,391). It is most likely a question of 'when' and not 'if' these approaches are adopted as standard treatments in AKU, and the metabolomics strategy developed as part of this thesis could be invaluable in assessing the wider metabolic impact or off-target effects of future therapies compared with nitisinone.

The studies reported here on ochronotic pigment and cartilage could be extended in a series of further NMR analyses using stable isotopically labelled HGA for enhancement of signals in the ^{13}C dimension. These analyses would

be performed on pigment formed *in vitro* from $^{13}C_6$ HGA and on cartilage tissue explants incubated with $^{13}C_6$ HGA; non-ochronotic cartilage explants have been shown to develop ochronosis macroscopically in *ex vivo* cultures over a period of weeks-months (63). Enhancement of NMR ^{13}C signals in these studies would improve signal-to-noise in 2D studies aimed at further probing the molecular conformation of ochronotic pigment and its association with the cartilage extracellular matrix. It would also be of interest to study the effect of specific sites of damage to the cartilage matrix, induced by natural ageing or under controlled conditions in the laboratory by chemical/mechanical means, on the interaction of HGA, pigment or associated species with cartilage. NMR analyses on samples from these studies could help to elucidate a specific binding site for HGA or pigment within the type-II collagen matrix. This information will be key to the development of future treatment strategies that directly target cartilage in AKU and potentially in other diseases of cartilage more generally; identification of an agent with a specific affinity for cartilage would also be of major interest in common OA.

Other analytical techniques should be explored in order to solve the chemical structure of ochronotic pigment. Raman spectroscopy has emerged as a sensitive technique for detection of ochronosis *ex vivo* (183) and is therefore worth considering as a means of obtaining more detailed chemical information on the pigment itself. MALDI-MS could offer a more suitable mass spectrometric technique for testing the presence of polymer species in ochronotic pigment due to its ability to generate singly-charged polymer ion distributions with minimal fragmentation (392,393).

The indication of co-existing hydroquinone & quinone functionalities in the data acquired from ochronotic pigment is interesting and provides novel insights into its chemistry. These data are in line with the observation that pigment, an oxidised quinone product of HGA, is able to simultaneously act as a reducing agent, (*e.g.* it reduces of Schmorl's stain to result in a blue colouration (62)), which indicates co-presence of reduced hydroquinone electron-donating groups. This is an important and previously under-recognised feature of ochronotic pigment. As discussed in Chapter 6, detection of oxidised quinone products of HGA poses a significant challenge in mass spectrometric analyses of ochronotic pigment. Electrochemical techniques could be particularly informative in understanding the oxidised quinone-like sub-component of ochronotic pigment, as oxidation of hydroquinone-containing species such as HGA produces free electrons. In summary, a range of further studies employing a range of analytical techniques will be required to probe the diverse chemical properties of ochronotic pigment in order to finally elucidate its complex chemical composition.

8.2   Concluding remarks

The data reported in this thesis, particularly from metabolomic analyses, highlight several important points beyond AKU. These findings demonstrate the power of metabolomic analyses combining high-sensitivity, high-resolution modern instruments such as LC-QTOF-MS with chemometric data analyses. Second is that targeted deletion of an enzyme can result in a complex pattern of metabolite alteration which extends to other unexpected pathways, including those that were previously thought to operate independently. Single-

gene disorders of metabolism, such as AKU, therefore present unique windows into metabolism that can uncover associations between metabolite pathways in physiology more generally. In conclusion, this work therefore serves as a reminder of the value of studying rare diseases as gateways to understanding common conditions and human physiology (9). As stated elegantly by William Harvey:

"*Nature is nowhere accustomed more openly to display her secret mysteries than in cases where she shows tracings of her workings apart from the beaten paths; nor is there any better way to advance the proper practice of medicine than to give our minds to the discovery of the usual law of nature, by careful investigation of cases of rarer forms of disease*" (394).

9.0   REFERENCES

1. Knox WE. Sir Archibald Garrod's Inborn Errors of Metabolism. IV. Pentosuria. Am J Hum Genet. 1958;10(4):385–97.
2. Garrod AE. The incidence of alkaptonuria: a study in chemical individuality. Lancet. 1902;2:1616–20.
3. O'Brien WM, La Du BN, Bunim JJ. Biochemical, pathologic and clinical aspects of alcaptonuria, ochronosis and ochronotic arthropathy. Am J Med. 1963;34(6):813–38.
4. Galdston M, Steele JM, Dobriner K. Alcaptonuria and ochronosis; with a report of three patients and metabolic studies in two. Am J Med. 1952;13(4):432–52.
5. Zatkova A. An update on molecular genetics of Alkaptonuria (AKU). J Inherit Metab Dis. 2011;34(6):1127–36.
6. Gallagher JA, Ranganath LR, Zatkova A. Alkaptonuria. In: Molloy S, Hughes K, editors. Brenner's Encyclopedia of Genetics. 2nd ed., San Diego, CA: Academic Press; 2013. p. 71–5.
7. Uhlén M, Fagerberg L, Hallström BM, Lindskog C, Oksvold P, Mardinoglu A, Sivertsson Å, Kampf C, Sjöstedt E, Asplund A, Olsson IM, Edlund K, Lundberg E, Navani S, Szigyarto CAK, Odeberg J, Djureinovic D, Takanen JO, Hober S, Alm T, Edqvist PH, Berling H, Tegel H, Mulder J, Rockberg J, Nilsson P, Schwenk JM, Hamsten M, Von Feilitzen K, Forsberg M, Persson L, Johansson F, Zwahlen M, Von Heijne G, Nielsen J, Pontén F.Tissue-based map of the human proteome. Science. 2015;347(6220):1260419.
8. Zatkova A, Sedlackova T, Radvansky J, Polakova H, Nemethova M, Aquaron R, Dursun I, Usher JL, Kadasi L. Identification of 11 novel homogentisate 1,2 dioxygenase variants in Alkaptonuria patients and establishment of a novel LOVD-based HGD mutation database. JIMD Rep. 2012;4:55–65.
9. Gallagher JA, Dillon JP, Sireau N, Timmis O, Ranganath LR. Alkaptonuria: An example of a "fundamental disease"- A rare disease with important lessons for more common disorders. Semin Cell Dev Biol. 2016;52:53–7.
10. La Du BN, Zannoni VG, Laster L, Seegmiller JE. The nature of the defect in tyrosine metabolism in alcaptonuria. J Biol Chem. 1958;230(1):251–60.
11. Ranganath LR, Milan AM, Hughes AT, Dutton JJ, Fitzgerald R, Briggs MC, Bygott H, Psarelli EE, Cox TF, Gallagher JA, Jarvis JC, van Kan C, Hall a. K, Laan D, Olsson B, Szamosi J, Rudebeck M, Kullenberg T, Cronlund A, Svensson L, Junestrand C, Ayoob H, Timmis OG, Sireau N, Le Quan Sang KH, Genovese F, Braconi D, Santucci A, Nemethova M, Zatkova A, McCaffrey J, Christensen P, Ross G, Imrich R, Rovensky J. Suitability Of Nitisinone In Alkaptonuria 1 (SONIA 1): an international, multicentre, randomised, open-label, no-treatment controlled, parallel-group, dose-response study to investigate the effect of once daily nitisinone on 24-h urinary homogentisic acid. Ann Rheum Dis. 2016;75(2):362–7.
12. Milan AM, Hughes AT, Davison AS, Devine J, Usher J, Curtis S, Khedr M, Gallagher JA, Ranganath LR. The effect of nitisinone on homogentisic acid and tyrosine: a two-year survey of patients attending the National Alkaptonuria Centre, Liverpool. Ann Clin Biochem. 2017;54(3):323–30.
13. Deutsch JC, Santhosh-Kumar CR. Quantitation of homogentisic acid in normal human plasma. J Chromatogr B Biomed Sci Appl. 1996;677(1):147–51.
14. Introne WJ, Perry MB, Troendle J, Tsilou E, Kayser MA, Suwannarat P, O'Brien KE, Bryant J, Sachdev V, Reynolds JC, Moylan E, Bernardini I, Gahl WA. A 3-year randomized therapeutic trial of nitisinone in alkaptonuria. Mol Genet Metab. 2011;103:307–14.

15. Zannoni VG, Lomtevas N, Goldfinger S. Oxidation of homogentisic acid to ochronotic pigment in connective tissue. Biochim Biophys Acta - Gen Subj. 1969;177(1):94–105.
16. Ranganath LR, Norman BP, Gallagher JA. Ochronotic pigmentation is caused by homogentisic acid and is the key event in Alkaptonuria leading to the destructive consequences of the disease – a review. J Inherit Metab Dis. 2019;42:776-92.
17. Ranganath LR, Jarvis JC, Gallagher JA. Recent advances in management of alkaptonuria (invited review; best practice article). J Clin Pathol. 2013;66(5):367–73.
18. Phornphutkul C, Introne WJ, Perry MB, Bernardini I, Murphey MD, Fitzpatrick DL, Anderson PD, Huizing M, Anikster Y, Gerber LH, Gahl WA. Natural history of alkaptonuria. N Engl J Med. 2002;347:2111–21.
19. Ranganath LR, Cox TF. Natural history of alkaptonuria revisited: analyses based on scoring systems. J Inherit Metab Dis. 2011;34(6):1141–51.
20. Ranganath LR, Milan A, Hughes A, Khedr M, Davison A, Schweidi E, Norman BP, Hughes J, Bygott H, Luangrath E, Fitzgerald R, Psarelli E, van Kan C, Laan D, Olsson B, Rudebeck M, Mankowitz L, Sireau N, Arnoux J, Le Quan Sang KH, Jarvis JC, Genovese F, Braconi D, Santucci A, Zatkova A, Glasova H, Stančík R, Imrich R, Rhodes NP, Gallagher JA. Homogentisic acid is not only eliminated by glomerular filtration and tubular secretion but also produced in the kidney in alkaptonuria. J Inherit Metab Dis. doi: 10.1002/jimd.12181.
21. Davison AS, Milan AM, Gallagher JA, Ranganath LR. Acute fatal metabolic complications in alkaptonuria. J Inherit Metab Dis. 2016;39(2):203–10.
22. Introne WJ, Phornphutkul C, Bernardini I, McLaughlin K, Fitzpatrick D, Gahl WA. Exacerbation of the ochronosis of alkaptonuria due to renal insufficiency and improvement after renal transplantation. Mol Genet Metab. 2002;77:136–42.
23. Helliwell TR, Gallagher J a., Ranganath L. Alkaptonuria - a review of surgical and autopsy pathology. Histopathology. 2008;53(5):503–12.
24. Mannoni A, Selvi E, Lorenzini S, Giorgi M, Airó P, Cammelli D, Andreotti L, Marcolongo R, Porfirio B. Alkaptonuria, ochronosis, and ochronotic arthropathy. Semin Arthritis Rheum. 2004;33(4):239–48.
25. Perry MB, Suwannarat P, Furst GP, Gahl WA, Gerber LH. Musculoskeletal findings and disability in alkaptonuria. J Rheumatol. 2006;33(11):2280–5.
26. Pettit SJ, Fisher M, Gallagher JA, Ranganath LR. Cardiovascular manifestations of Alkaptonuria. J Inherit Metab Dis. 2011 Dec 20;34(6):1177–81.
27. Vijaikumar M, Thappa DM, Srikanth S, Sethuraman G, Nadarajan S. Alkaptonuric ochronosis presenting as palmoplantar pigmentation. Clin Exp Dermatol. 2000;25:305–7.
28. Srsen S, Srsnová K, Lányi A. [Clinical manifestation of alkaptonuria in relation to age (author's transl)]. Bratisl Lek Listy. 1982;77(6):662–9.
29. Siekert RG, Gibilisco JA. Discoloration of the teeth in alkaptonuria (ochronosis) and parkinsonism. Oral Surgery, Oral Med Oral Pathol. 1970;29:197–9.
30. Taylor AM, Boyde A, Wilson PJM, Jarvis JC, Davidson JS, Hunt JA, Ranganath LR, Gallagher JA. The role of calcified cartilage and subchondral bone in the initiation and progression of ochronotic arthropathy in alkaptonuria. Arthritis Rheum. 2011;63:3387–96.
31. Ranganath L. Unpublished observations.
32. Lindner M, Bertelmann T. On the ocular findings in ochronosis: a systematic review of literature. BMC Ophthalmol. 2014;14(1):12.
33. Steven RA, Kinshuck AJ, McCormick MS, Ranganath LR. ENT

manifestations of alkaptonuria: report on a case series. J Laryngol Otol. 2015;129(10):1004–8.

34. Pau HW. Involvement of the tympanic membrane and ear ossicle system in ochronotic alkaptonuria [German]. Laryngol Rhinol Otol (Stuttg). 1984;63(10):541–4.

35. Cox TF, Ranganath L. A quantitative assessment of alkaptonuria: Testing the reliability of two disease severity scoring systems. J Inherit Metab Dis. 2011;34:1153–62.

36. Manoj Kumar R V, Rajasekaran S. Spontaneous tendon ruptures in alkaptonuria. J Bone Joint Surg Br. 2003;85(6):883–6.

37. Ranganath LR, Cox TF. Natural history of alkaptonuria revisited: Analyses based on scoring systems. J Inherit Metab Dis. 2011;34:1141–51.

38. Taylor AM, Wilson PJM, Ingrams DR, Helliwell TR, Gallagher JA, Ranganath LR. Calculi and intracellular ochronosis in the submandibular tissues from a patient with alkaptonuria. J Clin Pathol. 2010;63:186–8.

39. Mullan A, Cocker D, Taylor G, Millar C, Ranganath L. Fatal oxidative haemolysis and methaemoglobinaemia in a patient with alkaptonuria and acute kidney injury. Clin Kidney J. 2015;8:109–12.

40. Davison AS, Milan AM, Gallagher JA, Ranganath LR. Acute fatal metabolic complications in alkaptonuria. J Inherit Metab Dis. 2016;39:203–10.

41. Braconi D, Millucci L, Bernardini G, Santucci A. Oxidative stress and mechanisms of ochronosis in alkaptonuria. Free Radic Biol Med. 2015;88:70–80.

42. Roberts NB, Curtis SA, Milan AM, Ranganath LR. The pigment in alkaptonuria relationship to melanin and other coloured substances: A review of metabolism, composition and chemical analysis. JIMD Rep. 2015;24:51–66.

43. Lindstedt S, Holme E, Lock EA, Hjalmarson O, Strandvik B. Treatment of hereditary tyrosinaemia type I by inhibition of 4-hydroxyphenylpyruvate dioxygenase. Lancet. 1992;340:813–7.

44. McKiernan PJ. Nitisinone for the treatment of hereditary tyrosinemia type I. Expert Opin Orphan Drugs. 2013;1:491–7.

45. McKiernan PJ. Nitisinone in the treatment of hereditary tyrosinaemia type 1. Drugs. 2006;66(6):743–50.

46. McKiernan PJ, Preece MA, Chakrapani A. Outcome of children with hereditary tyrosinaemia following newborn screening. Arch Dis Child. 2015;100(8):738–41.

47. de Laet C, Dionisi-Vici C, Leonard J V, McKiernan P, Mitchell G, Monti L, de Baulny HO, Pintos-Morell G, Spiekerkötter U. Recommendations for the management of tyrosinaemia type 1. Orphanet J Rare Dis rare Dis. 2013;8:8.

48. Anikster Y, Nyhan WL, Gahl WA. NTBC and Alkaptonuria. Am J Hum Genet. 1998;63(3):920–1.

49. Ranganath LR, Khedr M, Milan AM, Davison AS, Hughes AT, Usher JL, Taylor S, Loftus N, Daroszewska A, West E, Jones A, Briggs M, Fisher M, McCormick M, Judd S, Vinjamuri S, Griffin R, Psarelli EE, Cox TF, Sireau N, Dillon JP, Devine JM, Hughes G, Harrold J, Barton GJ, Jarvis JC, Gallagher JA. Nitisinone arrests ochronosis and decreases rate of progression of Alkaptonuria: Evaluation of the effect of nitisinone in the United Kingdom National Alkaptonuria Centre. Mol Genet Metab. 2018;125:127–34.

50. Preston AJ, Keenan CM, Sutherland H, Wilson PJ, Wlodarski B, Taylor AM, Williams DP, Ranganath LR, Gallagher JA, Jarvis JC. Ochronotic osteoarthropathy in a mouse model of alkaptonuria, and its inhibition by nitisinone. Ann Rheum Dis. 2014;73:284–9.

51. Davison AS, Milan AM, Hughes AT, Dutton JJ, Ranganath LR. Serum concentrations and urinary excretion of homogentisic acid and tyrosine in

normal subjects. Clin Chem Lab Med. 2015;53(3):e81-3.

52. Davison AS, Strittmatter N, Sutherland H, Hughes AT, Hughes J, Bou-Gharios G, Milan AM, Goodwin RJA, Ranganath LR, Gallagher JA. Assessing the effect of nitisinone induced hypertyrosinaemia on monoamine neurotransmitters in brain tissue from a murine model of alkaptonuria using mass spectrometry imaging. Metabolomics. 2019;15(5):81.

53. Braconi D, Laschi M, Amato L, Bernardini G, Millucci L, Marcolongo R, Cavallo G, Spreafico A, Santucci A. Evaluation of anti-oxidant treatments in an in vitro model of alkaptonuric ochronosis. Rheumatology. 2010;49(10):1975–83.

54. Braconi D, Bianchini C, Bernardini G, Laschi M, Millucci L, Spreafico A, Santucci A. Redox-proteomics of the effects of homogentisic acid in an in vitro human serum model of alkaptonuric ochronosis. J Inherit Metab Dis. 2011;34(6):1163–76.

55. Millucci L, Spreafico A, Tinti L, Braconi D, Ghezzi L, Paccagnini E, Bernardini G, Amato L, Laschi M, Selvi E, Galeazzi M, Mannoni A, Benucci M, Lupetti P, Chellini F, Orlandini M, Santucci A. Alkaptonuria is a novel human secondary amyloidogenic disease. Biochim Biophys Acta - Mol Basis Dis. 2012;1822(11):1682–91.

56. Braconi D, Laschi M, Taylor AM, Bernardini G, Spreafico A, Tinti L, Gallagher JA, Santucci A. Proteomic and redox-proteomic evaluation of homogentisic acid and ascorbic acid effects on human articular chondrocytes. J Cell Biochem. 2010;111(4):922–32.

57. Braconi D, Bernardini G, Bianchini C, Laschi M, Millucci L, Amato L, Tinti L, Serchi T, Chellini F, Spreafico A, Santucci A. Biochemical and proteomic characterization of alkaptonuric chondrocytes. J Cell Physiol. 2012;227(9):3333–43.

58. Mistry JB, Jackson DJ, Bukhari M, Taylor AM. A role for interleukins in ochronosis in a chondrocyte in vitro model of alkaptonuria. Clin Rheumatol. 2016;35(7):1849–56.

59. Tinti L, Spreafico A, Braconi D, Millucci L, Bernardini G, Chellini F, Cavallo G, Selvi E, Galeazzi M, Marcolongo R, Gallagher JA, Santucci A. Evaluation of antioxiodant drugs for the treatment of ochronotic alkaptonuria in an in vitro human cell model. J Cell Physiol. 2010;225(1):84–91.

60. Spreafico A, Millucci L, Ghezzi L, Geminiani M, Braconi D, Amato L, Chellini F, Frediani B, Moretti E, Collodel G, Bernardini G, Santucci A. Antioxidants inhibit SAA formation and pro-inflammatory cytokine release in a human cell model of alkaptonuria. Rheumatology. 2013;52(9):1667–73.

61. Millucci L, Giorgetti G, Viti C, Ghezzi L, Gambassi S, Braconi D, Marzocchi B, Paffetti A, Lupetti P, Bernardini G, Orlandini M, Santucci A. Chondroptosis in alkaptonuric cartilage. J Cell Physiol. 2015;230(5):1148–57.

62. Tinti L, Taylor AM, Santucci A, Wlodarski B, Wilson PJ, Jarvis JC, Fraser WD, Davidson JS, Ranganath LR, Gallagher JA. Development of an in vitro model to investigate joint ochronosis in alkaptonuria. Rheumatology. 2011;50:271–7.

63. Tinti L, Spreafico A, Chellini F, Galeazzi M, Santucci A. A novel ex vivo organotypic culture model of alkaptonuria-ochronosis. Clin Exp Rheumatol. 2011;29(4):693–6.

64. Bernardini G, Leone G, Millucci L, Consumi M, Braconi D, Spiga O, Galderisi S, Marzocchi B, Viti C, Giorgetti G, Lupetti P, Magnani A, Santucci A. Homogentisic acid induces morphological and mechanical aberration of ochronotic cartilage in alkaptonuria. J Cell Physiol. 2019;234(5):6696–708.

65. Braconi D, Millucci L, Bernardini G, Santucci A. Oxidative stress and mechanisms of ochronosis in alkaptonuria. Free Radic Biol Med. 2015;88(Pt A):70–80.

66. Taylor AM, Preston AJ, Paulk NK, Sutherland H, Keenan CM, Wilson PJM, Wlodarski B, Grompe M, Ranganath LR, Gallagher JA, Jarvis JC. Ochronosis in a murine model of alkaptonuria is synonymous to that in the human condition. Osteoarthr Cartil. 2012;20:880–6.

67. Lewis R. Consequences and prevention of elevated circulating tyrosine during nitisinone therapy in alkaptonuria [PhD thesis]. Liverpool John Moores University; 2018.

68. Taylor AM, Preston AJ, Paulk NK, Sutherland H, Keenan CM, Wilson PJM, Wlodarski B, Grompe M, Ranganath LR, Gallagher JA, Jarvis JC. Ochronosis in a murine model of alkaptonuria is synonymous to that in the human condition. Osteoarthr Cartil. 2012;20(8):880–6.

69. Keenan CM, Preston AJ, Sutherland H, Wilson PJ, Psarelli EE, Cox TF, Ranganath LR, Jarvis JC, Gallagher JA. Nitisinone arrests but does not reverse ochronosis in alkaptonuric mice. JIMD Rep. 2015;24:45–50.

70. Hughes JH, Liu K, Plagge A, Wilson PJM, Sutherland H, Norman BP, Hughes AT, Keenan CM, Milan AM, Sakai T, Ranganath LR, Gallagher JA, Bou-Gharios G. Conditional targeting in mice reveals that hepatic homogentisate 1,2-dioxygenase activity is essential in reducing circulating homogentisic acid and for effective therapy in the genetic disease alkaptonuria. Hum Mol Genet. 2019;28(23):3928-39.

71. Justice MJ, Noveroske JK, Weber JS, Zheng B, Bradley A. Mouse ENU mutagenesis. Hum Mol Genet. 1999;8(10):1955–63.

72. Oliver SG, Winson MK, Kell DB, Baganz F. Systematic functional analysis of the yeast genome. Trends Biotechnol. 1998;16(9):373–8.

73. Fell DA. Understanding the Control of Metabolism. London: Portland Press; 1996.

74. Viant MR, Kurland IJ, Jones MR, Dunn WB. How close are we to complete annotation of metabolomes? Curr Opin Chem Biol. 2017;36:64–9.

75. Patti GJ, Yanes O, Siuzdak G. Innovation: Metabolomics: the apogee of the omics trilogy. Nat Rev Mol Cell Biol. 2012;13(4):263–9.

76. Hollywood K, Brison DR, Goodacre R. Metabolomics: Current technologies and future trends. Proteomics. 2006;6(17):4716–23.

77. Holmes E, Wilson ID, Nicholson JK. Metabolic phenotyping in health and disease. Cell. 2008;134(5):714–7.

78. van der Greef J, van Wietmarschen H, van Ommen B, Verheij E. Looking back into the future: 30 years of metabolomics at TNO. Mass Spectrom Rev. 2013;32(5):399–415.

79. Williams RJ. Biochemical individuality and its implications. Chem Eng News. 1947;25(16):1112–3.

80. Putri SP, Nakayama Y, Matsuda F, Uchikata T, Kobayashi S, Matsubara A, Fukusaki E. Current metabolomics: Practical applications. J Biosci Bioeng. 2013;115(6):579–89.

81. Goodacre R, Vaidyanathan S, Dunn WB, Harrigan GG, Kell DB. Metabolomics by numbers: acquiring and understanding global metabolite data. Trends Biotechnol. 2004;22(5):245–52.

82. Weckwerth W. Metabolomics in systems biology. Annu Rev Plant Biol. 2003;54(1):669–89.

83. Kell DB, Oliver SG. The metabolome 18 years on: a concept comes of age. Metabolomics. 2016;12(9):148.

84. Jellum E. Profiling of human body fluids in healthy and diseased states using gas chromatography and mass spectrometry, with special reference to organic acids. J Chromatogr. 1977;143(5):427–62.

85. Politzer IR, Dowty BJ, Laseter JL. Use of gas chromatography and mass spectrometry to analyze underivatized volatile human or animal constituents of clinical interest. Clin Chem. 1976;22(11):1775–88.

86. Dettmer K, Aronov PA, Hammock BD. Mass spectrometry-based metabolomics. Mass Spectrom Rev. 2007;26(1):51–78.

87. Makarov A, Denisov E, Kholomeev A, Balschun W, Lange O, Strupat K, Horning S. Performance evaluation of a hybrid linear ion trap/orbitrap mass spectrometer. Anal Chem. 2006;78(7):2113–20.

88. Zelena E, Dunn WB, Broadhurst D, Francis-McIntyre S, Carroll KM, Begley P, O'Hagan S, Knowles JD, Halsall A, Wilson ID, Kell DB, Kell DB. Development of a robust and repeatable UPLC−MS method for the long-term metabolomic study of human serum. Anal Chem. 2009;81(4):1357–64.

89. O'Hagan S, Dunn WB, Brown M, Knowles JD, Kell DB. Closed-loop, multiobjective optimization of analytical instrumentation: gas chromatography/time-of-flight mass spectrometry of the metabolomes of human serum and of yeast fermentations. Anal Chem. 2005;77(1):290–303.

90. Dunn WB, Broadhurst D, Begley P, Zelena E, Francis-McIntyre S, Anderson N, Brown M, Knowles JD, Halsall A, Haselden JN, Nicholls AW, Wilson ID, Kell DB, Goodacre R. Procedures for large-scale metabolic profiling of serum and plasma using gas chromatography and liquid chromatography coupled to mass spectrometry. Nat Protoc. 2011;6(7):1060–83.

91. Dunn WB, Broadhurst DI, Atherton HJ, Goodacre R, Griffin JL. Systems level studies of mammalian metabolomes: the roles of mass spectrometry and nuclear magnetic resonance spectroscopy. Chem Soc Rev. 2011;40(1):387–426.

92. Wishart DS. Advances in metabolite identification. Bioanalysis. 2011;3(15):1769–82.

93. Fiehn O. Metabolomics - The link between genotypes and phenotypes. Plant Mol Biol. 2002;48(1–2):155–71.

94. Dunn WB, Erban A, Weber RJM, Creek DJ, Brown M, Breitling R, Hankemeier T, Goodacre R, Neumann S, Kopka J, Viant MR. Mass appeal: Metabolite identification in mass spectrometry-focused untargeted metabolomics. Metabolomics. 2013;9(1):44–66.

95. Worley B, Powers R. Multivariate analysis in metabolomics. Curr Metabolomics. 2012;1(1):92–107.

96. Trygg J, Holmes E, Lundstedt T. Chemometrics in metabonomics. J Proteome Res. 2007;6(2):469–79.

97. Nicholson JK, Lindon JC, Holmes E. "Metabonomics": understanding the metabolic responses of living systems to pathophysiological stimuli via multivariate statistical analysis of biological NMR spectroscopic data. Xenobiotica. 1999;29(11):1181–9.

98. Fonville JM, Richards SE, Barton RH, Boulange CL, Ebbels TMD, Nicholson JK, Holmes E, Dumas M-E. The evolution of partial least squares models and related chemometric approaches in metabonomics and metabolic phenotyping. J Chemom. 2010;24(11–12):636–49.

99. Lee MY, Hu T. Computational methods for the discovery of metabolic markers of complex traits. Metabolites. 2019;9(4).

100. Weber RJM, Lawson TN, Salek RM, Ebbels TMD, Glen RC, Goodacre R, Griffin JL, Haug K, Koulman A, Moreno P, Ralser M, Steinbeck C, Dunn WB, Viant MR. Computational tools and workflows in metabolomics: An international survey highlights the opportunity for harmonisation through Galaxy. Metabolomics. 2017;13(2).

101. Gromski PS, Muhamadali H, Ellis DI, Xu Y, Correa E, Turner ML, Goodacre R. A tutorial review: Metabolomics and partial least squares-discriminant analysis – a marriage of convenience or a shotgun wedding. Anal Chim Acta. 2015;879:10–23.

102. Dunn WB, Broadhurst DI, Edison A, Guillou C, Viant MR, Bearden DW, Beger RD. Quality assurance and quality control processes: summary of a

metabolomics community questionnaire. Metabolomics. 2017;13(5).

103. Rubtsov D V., Jenkins H, Ludwig C, Easton J, Viant MR, Günther U, Griffin JL, Hardy N. Proposed reporting requirements for the description of NMR-based metabolomics experiments. Metabolomics. 2007;3(3):223–9.

104. Morrison N, Bearden D, Bundy JG, Collette T, Currie F, Davey MP, Haigh NS, Hancock D, Jones OAH, Rochfort S, Sansone S-A, Štys D, Teng Q, Field D, Viant MR. Standard reporting requirements for biological samples in metabolomics experiments: environmental context. Metabolomics. 2007;3(3):203–10.

105. Griffin JL, Nicholls AW, Daykin CA, Heald S, Keun HC, Schuppe-Koistinen I, Griffiths JR, Cheng LL, Rocca-Serra P, Rubtsov D V., Robertson D. Standard reporting requirements for biological samples in metabolomics experiments: mammalian/in vivo experiments. Metabolomics. 2007;3(3):179–88.

106. Goodacre R, Broadhurst D, Smilde AK, Kristal BS, Baker JD, Beger R, Bessant C, Connor S, Capuani G, Craig A, Ebbels T, Kell DB, Manetti C, Newton J, Paternostro G, Somorjai R, Sjöström M, Trygg J, Wulfert F. Proposed minimum reporting standards for data analysis in metabolomics. Metabolomics. 2007;3(3):231–41.

107. Sumner LW, Amberg A, Barrett D, Beale MH, Beger R, Daykin CA, Fan TWM, Fiehn O, Goodacre R, Griffin JL, Hankemeier T, Hardy N, Harnly J, Higashi R, Kopka J, Lane AN, Lindon JC, Marriott P, Nicholls AW, Reily MD, Thaden JJ, Viant MR. Proposed minimum reporting standards for chemical analysis: Chemical Analysis Working Group (CAWG) Metabolomics Standards Initiative (MSI). Metabolomics. 2007;3(3):211–21.

108. Stanstrup J, Broeckling CD, Helmus R, Hoffmann N, Mathé E, Naake T, Nicolotti L, Peters K, Rainer J, Salek RM, Schulze T, Schymanski EL, Stravs MA, Thévenot EA, Treutler H, Weber RJM, Willighagen E, Witting M, Neumann S. The metaRbolomics Toolbox in Bioconductor and beyond. Metabolites. 2019;9(10):200.

109. Salek RM, Haug K, Conesa P, Hastings J, Williams M, Mahendraker T, Maguire E, González-Beltrán AN, Rocca-Serra P, Sansone S-A, Steinbeck C. The MetaboLights repository: curation challenges in metabolomics. Database. 2013;bat029.

110. Sud M, Fahy E, Cotter D, Azam K, Vadivelu I, Burant C, Edison A, Fiehn O, Higashi R, Nair KS, Sumner S, Subramaniam S. Metabolomics Workbench: An international repository for metabolomics data and metadata, metabolite standards, protocols, tutorials and training, and analysis tools. Nucleic Acids Res. 2016;44(D1):D463–70.

111. Mendez KM, Pritchard L, Reinke SN, Broadhurst DI. Toward collaborative open data science in metabolomics using Jupyter Notebooks and cloud computing. Metabolomics. 2019;15(10):125.

112. Dunn WB, Bailey NJC, Johnson HE. Measuring the metabolome: Current analytical technologies. Analyst. 2005;130(5):606–25.

113. Zhou B, Xiao JF, Tuli L, Ressom HW. LC-MS-based metabolomics. Mol Biosyst. 2012;8(2):470.

114. Xian F, Hendrickson CL, Marshall AG. High resolution mass spectrometry. Anal Chem. 2012;84(2):708–19.

115. Garcia-Reyes JF, Moreno-Gonzalez D, Nortez-Méndez R, López BG, Molina-Díaz A. High resolution mass spectrometry: hardware and software. In: Romero-González R, Frenich AG, editors. Applications in High Resolution Mass Spectrometry. 1st ed. Amsterdam: Elsevier; 2017.

116. Capati A, Ijare OB, Bezabeh T. Diagnostic applications of nuclear magnetic resonance-based urinary metabolomics. Magn Reson Insights. 2017;10:1178623X17694346.

117. Cameron AE, Eggers DF. An Ion "Velocitron." Rev Sci Instrum.

1948;19(9):605–7.

118. Shevchenko A, Chernushevich I, Ens W, Standing KG, Thomson B, Wilm M, Mann M. Rapid 'de novo' peptide sequencing by a combination of nanoelectrospray, isotopic labeling and a quadrupole/time-of-flight mass spectrometer. Rapid Commun Mass Spectrom. 1997;11(9):1015–24.

119. Agilent Technologies Inc. Time-of-flight mass spectrometry: technical overview. 2011.

120. Agilent Technologies Inc. Agilent 6200 series TOF and 6500 series Q-TOF LC/MS system concepts guide. 2014. p. 1–132.

121. Yamashita M, Fenn JB. Electrospray ion source. Another variation on the free-jet theme. J Phys Chem. 1984;88(20):4451–9.

122. Karas M, Hillenkamp F. Laser desorption ionization of proteins with molecular masses exceeding 10,000 daltons. Anal Chem. 1988;60(20):2299–301.

123. Bruins AP. Mechanistic aspects of electrospray ionization. J Chromatogr A. 1998;794(1–2):345–57.

124. Konermann L, Ahadi E, Rodriguez AD, Vahidi S. Unraveling the mechanism of electrospray ionization. Anal Chem. 2013;85(1):2–9.

125. Kruve A, Kaupmees K, Liigand J, Leito I. Negative electrospray ionization via deprotonation: predicting the ionization efficiency. Anal Chem. 2014;86(10):4822–30.

126. Annesley TM. Ion suppression in mass spectrometry. Clin Chem. 2003;49(7):1041–4.

127. Horvath CG, Lipsky SR. Peak capacity in chromatography. Anal Chem. 1967;39(14):1893.

128. Grushka E. Chromatographic peak capacity and the factors influencing it. Anal Chem. 1970;42(11):1142–7.

129. Waters Corporation. Bands, peaks and band spreading [Internet]. [cited 2019 Sep 27]. Available from: https://www.waters.com/waters/en_GB/Chromatographic-Bands%2C-Peaks-and-Band-Spreading/nav.htm?cid=134803614&locale=en_GB

130. Banoei MM, Donnelly SJ, Mickiewicz B, Weljie A, Vogel HJ, Winston BW. Metabolomics in critical care medicine: a new approach to biomarker discovery. Clin Invest Med. 2014;37(6):E363-76.

131. Wilson ID. Drugs, bugs, and personalized medicine: Pharmacometabonomics enters the ring. Proc Natl Acad Sci. 2009;106(34):14187–8.

132. Nicholson JK. Global systems biology, personalized medicine and molecular epidemiology. Mol Syst Biol. 2006;2(52):Epub.

133. Nicholson JK, Holmes E, Kinross JM, Darzi AW, Takats Z, Lindon JC. Metabolic phenotyping in clinical and surgical environments. Nature. 2012;491(7424):384–92.

134. Beger RD, Dunn W, Schmidt MA, Gross SS, Kirwan JA, Cascante M, Brennan L, Wishart DS, Oresic M, Hankemeier T, Broadhurst DI, Lane AN, Suhre K, Kastenmüller G, Sumner SJ, Thiele I, Fiehn O, Kaddurah-Daouk R, Initiative for "Precision M and PTG-MS. Metabolomics enables precision medicine: "A White Paper, Community Perspective." Metabolomics. 2016;12(9):149.

135. Hampe MH, Panaskar SN, Yadav AA, Ingale PW. Gas chromatography/mass spectrometry-based urine metabolome study in children for inborn errors of metabolism: An Indian experience. Clin Biochem. 2017;50(3):121–6.

136. Sandlers Y. The future perspective: metabolomics in laboratory medicine for inborn errors of metabolism. Transl Res. 2017;189:65–75.

137. Hughes AT, Milan AM, Davison AS, Christensen P, Ross G, Gallagher JA, Dutton JJ, Ranganath LR. Serum markers in alkaptonuria: simultaneous analysis of homogentisic acid, tyrosine and nitisinone by liquid chromatography tandem mass spectrometry. Ann Clin Biochem.

2015;52(5):597–605.

138. Hughes A, Milan A, Christensen P, Ross G, Davison A, JA G, Dutton J, Ranganath L. Urine homogentisic acid and tyrosine: Simultaneous analysis by liquid chromatography tandem mass spectrometry. J Chromatogr B Anal Technol Biomed Life Sci. 2014;963:106–12.

139. ClinicalTrials.gov. Suitability of Nitisinone in Alkaptonuria 2 (SONIA 2): identifier NCT01916382 [Internet]. [cited 2019 Sep 29]. Available from: https://clinicaltrials.gov/ct2/show/NCT01916382

140. Ranganath LR, Khedr M, Milan AM, Davison AS, Hughes AT, Usher JL, Taylor S, Loftus N, Daroszewska A, West E, Jones A, Briggs M, Fisher M, McCormick M, Judd S, Vinjamuri S, Griffin R, Psarelli EE, Cox TF, Sireau N, Dillon JP, Devine JM, Hughes G, Harrold J, Barton GJ, Jarvis JC, Gallagher JA. Nitisinone arrests ochronosis and decreases rate of progression of Alkaptonuria: Evaluation of the effect of nitisinone in the United Kingdom National Alkaptonuria Centre. Mol Genet Metab. 2018;125(1–2):127–34.

141. Davison AS, Hughes A, Milan AM, Sireau N, Gallagher J, Ranganath L. ANNALS EXPRESS: Alkaptonuria - many questions answered, further challenges beckon. Ann Clin Biochem Int J Lab Med. 2019;4563219879957.

142. Meissner T, Betz RC, Pasternack SM, Eigelshoven S, Ruzicka T, Kruse R, Laitenberger G, Mayatepek E. Richner–Hanhart syndrome detected by expanded newborn screening. Pediatr Dermatol. 2008;25(3):378–80.

143. Stewart RMK, Briggs MC, Jarvis JC, Gallagher JA, Ranganath L. Reversible keratopathy due to hypertyrosinaemia following intermittent low-dose nitisinone in alkaptonuria: a case report. JIMD Rep. 2014;17:1–6.

144. Khedr M, Judd S, Briggs MC, Hughes AT, Milan AM, Stewart RMK, Lock EA, Gallagher JA, Ranganath LR. Asymptomatic corneal keratopathy secondary to hypertyrosinaemia following low dose nitisinone and a literature review of tyrosine keratopathy in alkaptonuria. JIMD Rep. 2018;40:31–7.

145. Lynch JJ, Van Vleet TR, Mittelstadt SW, Blomme EAG. Potential functional and pathological side effects related to off-target pharmacological activity. J Pharmacol Toxicol Methods. 2017 Sep;87:108–26.

146. Zannoni VG, Lomtevas N, Goldfinger S. Oxidation of homogentisic acid to ochronotic pigment in connective tissue. Biochim Biophys Acta. 1969;177(1):94–105.

147. Zannoni VG, Malawista SE, La Du BN. Studies on ochronosis. II. Studies on benzoquinoneacetic acid, a probable intermediate in the connective tissue pigmentation of alcaptonuria. Arthritis Rheum. 1962;5(6):547–56.

148. Taylor AM, Kammath V, Bleakley A. Tyrosinase, could it be a missing link in ochronosis in alkaptonuria? Med Hypotheses. 2016;91:77–80.

149. Hunter RC, Newman DK. A putative ABC transporter, hatABCDE, is among molecular determinants of pyomelanin production in Pseudomonas aeruginosa. J Bacteriol. 2010;192(22):5962–71.

150. Martin JP, Batkoff B. Homogentisic acid autoxidation and oxygen radical generation: implications for the etiology of alkaptonuric arthritis. Free Radic Biol Med. 1987;3(4):241–50.

151. Milch RA, Titus ED, Loo TL. Atmospheric oxidation of homogentisic acid: Spectrophotometric studies. Science. 1957;126(3266):209–10.

152. Consden R, Forbes HAW, Glynn LE, Stainer WM. Observations on the oxidation of homogentisic acid in urine. Biochem J. 1951;50(2):274–8.

153. McNaught AD, Wilkinson A. IUPAC. Compendium of Chemical Terminology 2nd ed. (the "Gold Book"). Oxford: Blackwell Scientific Publications; 2014.

154. Allcock H, Lampe F, Mark J. Contemporary Polymer Chemistry. 3rd ed. New Jersey: Prentice Hall; 2003.

155. Taylor AM, Vercruysse KP. Analysis of melanin-like pigment synthesized from homogentisic acid, with or without tyrosine, and its implications in

alkaptonuria. JIMD Rep. 2017;35:79–85.

156. Mekala LP, Mohammed M, Chinthalapati S, Chinthalapati VR. Pyomelanin production: Insights into the incomplete aerobic L-phenylalanine catabolism of a photosynthetic bacterium, Rubrivivax benzoatilyticus JA2. Int J Biol Macromol. 2019;126:755–64.

157. Mekala LP, Mohammed M, Chintalapati S, Chintalapati VR. Stable isotope-assisted metabolic profiling reveals growth mode dependent differential metabolism and multiple catabolic pathways of L-phenylalanine in rubrivivax benzoatilyticus JA2. J Proteome Res. 2018;17(1):189–202.

158. Tran ML, Powell BJ, Meredith P. Chemical and structural disorder in eumelanins: A possible explanation for broadband absorbance. Biophys J. 2006;90(3):743–52.

159. Riesz J. The spectroscopic properties of melanin [PhD thesis]. University of Queensland; 2007.

160. Chen CT, Chuang C, Cao J, Ball V, Ruch D, Buehler MJ. Excitonic effects from geometric order and disorder explain broadband optical absorption in eumelanin. Nat Commun. 2014;5:3859.

161. Tokuhara Y, Shukuya K, Tanaka M, Mouri M, Ohkawa R, Fujishiro M, Takahashi T, Okubo S, Yokota H, Kurano M, Ikeda H, Yamaguchi S, Inagaki S, Ishige-Wada M, Usui H, Yatomi Y, Shimosawa T. Detection of novel visible-light region absorbance peaks in the urine after alkalization in patients with alkaptonuria. PLoS One. 2014;9(1):e86606.

162. Tokuhara Y, Shukuya K, Tanaka M, Sogabe K, Ejima Y, Hosokawa S, Ohsaki H, Morinishi T, Hirakawa E, Yatomi Y, Shimosawa T. Absorbance measurements of oxidation of homogentisic acid accelerated by the addition of alkaline solution with sodium hypochlorite pentahydrate. Sci Rep. 2018;8(1):11364.

163. Li Y, Liu J, Wang Y, Chan HW, Wang L, Chan W. Mass spectrometric and spectrophotometric analyses reveal an alternative structure and a new formation mechanism for melanin. Anal Chem. 2015;87:7958–63.

164. Damian LO, Felea I, Boloşiu C, Botar-Jid C, Fodor D, Rednic S. A case of alkaptonuria - ultrasonographic findings. Med Ultrason. 2013;15:321–5.

165. Fisher AA, Davis MW. Alkaptonuric ochronosis with aortic valve and joint replacements and femoral fracture: a case report and literature review. Clin Med Res. 2004;2:209–15.

166. Gaines JJ, Tom GD, Khankhanian N. An ultrastructural and light microscopic study of the synovium in ochronotic arthropathy. Hum Pathol. 1987;18:1160–4.

167. Hagg R, Bruckner P, Hedbom E. Cartilage fibrils of mammals are biochemically heterogeneous: differential distribution of decorin and collagen IX. J Cell Biol. 1998;142(1):285–94.

168. Nicolae C, Ko Y-P, Miosge N, Niehoff A, Studer D, Enggist L, Hunziker EB, Paulsson M, Wagener R, Aszodi A. Abnormal collagen fibrils in cartilage of matrilin-1/matrilin-3-deficient mice. J Biol Chem. 2007;282(30):22163–75.

169. Taylor AM, Wlodarski B, Prior IA, Wilson PJM, Jarvis JC, Ranganath LR, Gallagher JA. Ultrastructural examination of tissue in a patient with alkaptonuric arthropathy reveals a distinct pattern of binding of ochronotic pigment. Rheumatology. 2010;49:1412–4.

170. Taylor AM, Hsueh MF, Ranganath LR, Gallagher JA, Dillon JP, Huebner JL, Catterall JB, Kraus VB. Cartilage biomarkers in the osteoarthropathy of alkaptonuria reveal low turnover and accelerated ageing. Rheumatology. 2017;56:156–64.

171. Taylor AM, Batchelor TJP, Adams VL, Helliwell TR, Gallagher JA, Ranganath LR. Ochronosis and calcification in the mediastinal mass of a patient with alkaptonuria. J Clin Pathol. 2011;64:935–6.

172. Chow WY, Taylor AM, Reid DG, Gallagher JA, Duer MJ. Collagen atomic scale molecular disorder in ochronotic cartilage from an alkaptonuria patient, observed by solid state NMR. J Inherit Metab Dis. 2011;34:1137–40.

173. Gottardi R, Hansen U, Raiteri R, Loparic M, Düggelin M, Mathys D, Friederich NF, Bruckner P, Stolz M. Supramolecular organization of collagen fibrils in healthy and osteoarthritic human knee and hip joint cartilage. PLoS One. 2016;11(10):e0163552.

174. Williams DP, Lawrence A, Meng X. Pharmacological and toxicological considerations of homogentisic acid in alkaptonuria. Pharmacologia. 2012;3:61–74.

175. Milch RA, Murray RA. Studies of alcaptonuria: Adsorption of homogentisic acid solutions on collagen chromatographic columns. Arthritis Rheum. 1961;4(3):268–74.

176. Milch RA, Murray RA. Studies of alcaptonuria : Mechanisms of swelling of homogentisic acid-collagen preparations. Arthritis Rheum. 1961;4(3):253–67.

177. Lustberg TJ, Schulman JD, Seegmiller JE. The preparation and identification of various adducts of oxidized homogentisic acid and the development of a new sensitive colorimetric assay for homogentisic acid. Clin Chim Acta. 1971;35(2):325–33.

178. Stoner R, Blivaiss BB. Reaction of quinone of homogentisic acid with biological amines. Arthritis Rheum. 1967;10(1):53–60.

179. Zannoni VG, Malawista SE, La Du BN. Studies on ochronosis. II. Studies on benzoquinoneacetic acid, a probable intermediate in the connective tissue pigmentation of alcaptonuria. Arthritis Rheum. 1962;5(6):547-56.

180. Milch RA. Studies of alcaptonuria: infra-red spectra of deuterated homogentisic acid solutions. Arthritis Rheum. 1965;8(5):1002–5.

181. Karavaggelis A, Young C, Attia R. Black heart at surgery - primary diagnosis of Alkaptonuria at surgery. J Cardiol Curr Res. 2017;9(5):00335.

182. Cox T, Psarelli E, Taylor S, Shepherd H, Robinson M, Barton G, Mistry A, Genovese F, Braconi D, Giustarini D, Rossi R, Santucci A, Khedr M, Hughes A, Milan A, LF T, West E, Sireau N, Dillon J, Rhodes NP, Gallagher JA, Ranganath LR. Subclinical ochronosis features in Alkaptonuria: a cross-sectional study. BMJ Innov. [in press.]

183. Taylor AM, Jenks DD, Kammath VD, Norman BP, Dillon JP, Gallagher JA, Ranganath LR, Kerns JG. Raman Spectroscopy identifies differences in ochronotic and non-ochronotic cartilage; a potential novel technique for monitoring ochronosis. Osteoarthr Cartil. 2019;S1063-4584.

184. Taylor AM, Wlodarski B, Wilson PJ, Jarvis J, Ranganath LR, Boyde A, Gallagher JA. Deposition of ochronotic pigment in articular cartilage in alkaptonuria is initiated near the tidemark and progresses to the articular surface. Bone. 2010;47:S79.

185. Smith CA, O'Maille G, Want EJ, Qin C, Trauger SA, Brandon TR, Custodio DE, Abagyan R, Siuzdak G. METLIN: a metabolite mass spectral database. Ther Drug Monit. 2005;27(6):747–51.

186. Hufsky F, Scheubert K, Böcker S. Computational mass spectrometry for small-molecule fragmentation. TrAC Trends Anal Chem. 2014 Jan 1;53:41–8.

187. Hill AW, Mortishire-Smith RJ. Automated assignment of high-resolution collisionally activated dissociation mass spectra using a systematic bond disconnection approach. Rapid Commun Mass Spectrom. 2005;19(21):3111–8.

188. Holden NE, Coplen TB, Böhlke JK, Tarbox L V., Benefield J, de Laeter JR, Mahaffy PG, O'Connor G, Roth E, Tepper DH, Walczyk T, Wieser ME, Yoneda S. IUPAC periodic table of the elements and isotopes (IPTEI) for the education community (IUPAC technical report). Pure Appl Chem. 2018;90(12):1833–2092.

189. Keeler J. Understanding NMR spectroscopy. Chichester, UK: Wiley; 2005.
190. Compound Interest. A guide to 1H and 13C NMR chemical shift values [Internet]. [cited 2019 Aug 1]. Available from: www.compoundchem.com
191. van Rossum B-J, Förster H, de Groot HJM. High-field and high-speed CP-MAS13C NMR heteronuclear dipolar-correlation spectroscopy of solids with frequency-switched Lee–Goldburg homonuclear decoupling. J Magn Reson. 1997;124(2):516–9.
192. Duer MJ. Essential techniques for spin-1/2 nuclei. In: Duer MJ, editor. Solid-state NMR spectroscopy principles and applications. Oxford: Blackwell Science; 2002. p. 73–110.
193. Lilly Thankamony AS, Wittmann JJ, Kaushik M, Corzilius B. Dynamic nuclear polarization for sensitivity enhancement in modern solid-state NMR. Prog Nucl Magn Reson Spectrosc. 2017;102–103:120–95.
194. McLafferty FW, Stauffer DA, Loh SY, Wesdemiotis C. Unknown identification using reference mass spectra. quality evaluation of databases. J Am Soc Mass Spectrom. 1999;10(12):1229–40.
195. Milman BL. Identification of chemical compounds. TrAC - Trends Anal Chem. 2005;24(6):493–508.
196. Kind T, Fiehn O. Advances in structure elucidation of small molecules using mass spectrometry. Bioanal Rev. 2010;2(1–4):23–60.
197. Stein S. Mass spectral reference libraries: An ever-expanding resource for chemical identification. Anal Chem. 2012;84(17):7274–82.
198. Vinaixa M, Schymanski EL, Neumann S, Navarro M, Salek RM, Yanes O. Mass spectral databases for LC/MS- and GC/MS-based metabolomics: State of the field and future prospects. TrAC - Trends Anal Chem. 2016;78:23–35.
199. Wishart DS, Feunang YD, Marcu A, Guo AC, Liang K, Vázquez-Fresno R, Sajed T, Johnson D, Li C, Karu N, Sayeeda Z, Lo E, Assempour N, Berjanskii M, Singhal S, Arndt D, Liang Y, Badran H, Grant J, Serra-Cayuela A, Liu Y, Mandal R, Neveu V, Pon A, Knox C, Wilson M, Manach C, Scalbert A. HMDB 4.0: The human metabolome database for 2018. Nucleic Acids Res. 2018;46(D1):D608–17.
200. Milman BL, Zhurkovich IK. Mass spectral libraries: A statistical review of the visible use. TrAC - Trends Anal Chem. 2016;80:636–40.
201. Norman BP, Davison A, Ross GA, Milan AM, Hughes AT, Sutherland H, Jarvis JC, Roberts NB, Gallagher JA, Ranganath L. Three accurate mass retention time (AMRT) databases generated from IROA Technologies Metabolite Library of Standards by LC-QTOF-MS analysis. figshare. Collection. 2019. Available from: http://doi.org/10.6084/m9.figshare.c.4378235.v2
202. King R, Bonfiglio R, Fernandez-Metzler C, Miller-Stein C, Olah T. Mechanistic investigation of ionization suppression in electrospray ionization. J Am Soc Mass Spectrom. 2000;11(11):942–50.
203. Pitt JJ. Principles and applications of liquid chromatography-mass spectrometry in clinical biochemistry. Clin Biochem Rev. 2009;30(1):19–34.
204. Matuszewski BK, Constanzer ML, Chavez-Eng CM. Matrix effect in quantitative LC/MS/MS analyses of biological fluids: a method for determination of finasteride in human plasma at picogram per milliliter concentrations. Anal Chem. 1998;70(5):882–9.
205. Ogata H, Goto S, Sato K, Fujibuchi W, Bono H, Kanehisa M. KEGG: Kyoto encyclopedia of genes and genomes. Nucleic Acids Res. 1999;27:29–34.
206. Horai H, Arita M, Kanaya S, Nihei Y, Ikeda T, Suwa K, Ojima Y, Tanaka K, Tanaka S, Aoshima K, Oda Y, Kakazu Y, Kusano M, Tohge T, Matsuda F, Sawada Y, Hirai MY, Nakanishi H, Ikeda K, Akimoto N, Maoka T, Takahashi H, Ara T, Sakurai N, Suzuki H, Shibata D, Neumann S, Iida T, Tanaka K, Funatsu K, Matsuura F, Soga T, Taguchi R, Saito K, Nishioka T. MassBank:

A public repository for sharing mass spectral data for life sciences. J Mass Spectrom. 2010;45(7):703–14.

207. Weckwerth W, Morgenthal K. Metabolomics: from pattern recognition to biological interpretation. Drug Discov Today. 2005;10(22):1551–8.

208. Chen J, Xu F. What can we do to refine the redundant data in LC–MS and GC–MS based metabolomics? Bioanalysis. 2017;9(3):235–8.

209. Broadhurst DI, Kell DB. Statistical strategies for avoiding false discoveries in metabolomics and related experiments. Metabolomics. 2007;2(4):171–96.

210. Castillo S, Gopalacharyulu P, Yetukuri L, Orešič M. Algorithms and tools for the preprocessing of LC-MS metabolomics data. Chemom Intell Lab Syst. 2011;108(1):23–32.

211. Agilent Technologies Inc. MassHunter Profinder software quick start giude. 2016. p. 1–40.

212. Warrack BM, Hnatyshyn S, Ott K-H, Reily MD, Sanders M, Zhang H, Drexler DM. Normalization strategies for metabonomic analysis of urine samples. J Chromatogr B. 2009;877(5–6):547–52.

213. Bouatra S, Aziat F, Mandal R, Guo AC, Wilson MR, Knox C, Bjorndahl TC, Krishnamurthy R, Saleem F, Liu P, Dame ZT, Poelzer J, Huynh J, Yallou FS, Psychogios N, Dong E, Bogumil R, Roehring C, Wishart DS. The human urine metabolome. PLoS One. 2013;8(9):e73076.

214. Miller RC, Brindle E, Holman DJ, Shofer J, Klein NA, Soules MR, O'Connor KA. Comparison of specific gravity and creatinine for normalizing urinary reproductive hormone concentrations. Clin Chem. 2004;50(5):924–32.

215. van den Berg RA, Hoefsloot HCJ, Westerhuis JA, Smilde AK, van der Werf MJ. Centering, scaling, and transformations: improving the biological information content of metabolomics data. BMC Genomics. 2006;7:142.

216. Kvalheim OM, Brakstad F, Liang Y. Preprocessing of analytical profiles in the presence of homoscedastic or heteroscedastic noise. Anal Chem. 1994;66(1):43–51.

217. Suwannarat P, O'Brien K, Perry MB, Sebring N, Bernardini I, Kaiser-Kupfer MI, Rubin BI, Tsilou E, Gerber LH, Gahl WA. Use of nitisinone in patients with alkaptonuria. Metabolism. 2005;54(6):719–28.

218. Lindstedt S, Holme E, Lock EA, Hjalmarson O, Strandvik B. Treatment of hereditary tyrosinaemia type I by inhibition of 4-hydroxyphenylpyruvate dioxygenase. Lancet. 1992;340(8823):813–7.

219. Olsson B, Cox TF, Psarelli EE, Szamosi J, Hughes AT, Milan AM, Hall AK, Rovensky J, Ranganath LR. Relationship between serum concentrations of nitisinone and its effect on homogentisic acid and tyrosine in patients with alkaptonuria. JIMD Rep. 2015;24:21–7.

220. Davison AS, Norman BP, Smith EA, Devine J, Usher J, Hughes AT, Khedr M, Milan AM, Gallagher JA, Ranganath LR. Serum amino acid profiling in patients with alkaptonuria before and after treatment with nitisinone. JIMD Rep. 2018;41:109–17.

221. Masurel-Paulet A, Poggi-Bach J, Rolland MO, Bernard O, Guffon N, Dobbelaere D, Sarles J, Baulny OOH, Touati G. NTBC treatment in tyrosinaemia type I: Long-term outcome in French patients. J Inherit Metab Dis. 2008;31(1):81–7.

222. De Laet C, Terrones Munoz V, Jaeken J, François B, Carton D, Sokal EM, Dan B, Goyens PJ. Neuropsychological outcome of NTBC-treated patients with tyrosinaemia type 1. Dev Med Child Neurol. 2011;53(10):962–4.

223. Thimm E, Richter-Werkle R, Kamp G, Molke B, Herebian D, Klee D, Mayatepek E, Spiekerkoetter U. Neurocognitive outcome in patients with hypertyrosinemia type I after long-term treatment with NTBC. J Inherit Metab Dis. 2012;35(2):263–8.

224. Bendadi F, De Koning TJ, Visser G, Prinsen HCMT, De Sain MGM,

Verhoeven-Duif N, Sinnema G, Van Spronsen FJ, Van Hasselt PM. Impaired cognitive functioning in patients with tyrosinemia type i receiving nitisinone. J Pediatr. 2014;164(2):398–401.

225. Vorkas PA, Isaac G, Anwar MA, Davies AH, Want EJ, Nicholson JK, Holmes E. Untargeted UPLC-MS profiling pipeline to expand tissue metabolome coverage: Application to cardiovascular disease. Anal Chem. 2015;87(8):4184–93.

226. Gertsman I, Barshop BA, Panyard-Davis J, Gangoiti JA, Nyhan WL. Metabolic effects of increasing doses of nitisinone in the treatment of alkaptonuria. JIMD Rep. 2015;24:13-20.

227. Taylor LF. Assessment of disease progression in the rare disease alkaptonuria by quantitative image analysis [PhD thesis]. University of Liverpool; 2018.

228. Kobayashi N, Inaba Y, Yukizawa Y, Ike H, Kubota S, Inoue T, Saito T. Use of 18 F-fluoride positron emission tomography as a predictor of the hip osteoarthritis progression. Mod Rheumatol. 2015;25(6):925–30.

229. Beck AT, Ward CH, Mendelson M, Mock J, Erbaugh J. An inventory for measuring depression. Arch Gen Psychiatry. 1961;4(6):561.

230. Beck AT, Steer RA, Ball R, Ranieri WF. Comparison of Beck Depression Inventories-IA and-II in psychiatric outpatients. J Pers Assess. 1996;67(3):588–97.

231. Usher JL, Ascher DB, Pires DE V., Milan AM, Blundell TL, Ranganath LR. Analysis of HGD gene mutations in patients with alkaptonuria from the United Kingdom: identification of novel mutations. JIMD Rep. 2014;24:3–11.

232. Zasadny KR, Wahl RL. Standardized uptake values of normal tissues at PET with 2-[fluorine-18]-fluoro-2-deoxy-D-glucose: variations with body weight and a method for correction. Radiology. 1993;189(3):847–50.

233. David JC, Dairman W, Udenfriend S. Decarboxylation to tyramine: a major route of tyrosine metabolism in mammals. Proc Natl Acad Sci. 1974;71(5):1771–5.

234. Davison AS, Norman BP, Ross GA, Hughes AT, Khedr M, Milan AM, Gallagher JA, Ranganath LR. Evaluation of the serum metabolome of patients with alkaptonuria before and after two years of treatment with nitisinone using LC-QTOF-MS. JIMD Rep. 2019;48(1):67–74.

235. Davison AS, Norman BP, Milan AM, Hughes AT, Khedr M, Rovensky J, Gallagher JA, Ranganath LR. Assessment of the effect of once daily nitisinone therapy on 24-h urinary metadrenalines and 5-hydroxyindole acetic acid excretion in patients with alkaptonuria after 4 weeks of treatment. JIMD Rep. 2018;41:1–10.

236. Thimm E, Herebian D, Assmann B, Klee D, Mayatepek E, Spiekerkoetter U. Increase of CSF tyrosine and impaired serotonin turnover in tyrosinemia type I. Mol Genet Metab. 2011;102(2):122–5.

237. Gertsman I, Gangoiti JA, Nyhan WL, Barshop BA. Perturbations of tyrosine metabolism promote the indolepyruvate pathway via tryptophan in host and microbiome. Mol Genet Metab. 2015;114(3):431–7.

238. Hillgartner MA, Coker SB, Koenig AE, Moore ME, Barnby E, MacGregor GG. Tyrosinemia type I and not treatment with NTBC causes slower learning and altered behavior in mice. J Inherit Metab Dis. 2016;39(5):673–82.

239. Harding CO, Winn SR, Gibson KM, Arning E, Bottiglieri T, Grompe M. Pharmacologic inhibition of L-tyrosine degradation ameliorates cerebral dopamine deficiency in murine phenylketonuria (PKU). J Inherit Metab Dis. 2014;37(5):735–43.

240. Pratt OE. Transport inhibition in the pathology of phenylketonuria and other inherited metabolic diseases. J Inherit Metab Dis. 1982;5(S2):75–81.

241. Réus GZ, Jansen K, Titus S, Carvalho AF, Gabbay V, Quevedo J.

Kynurenine pathway dysfunction in the pathophysiology and treatment of depression: Evidences from animal and human studies. J Psychiatr Res. 2015;68:316–28.

242. Vécsei L, Szalárdy L, Fülöp F, Toldi J. Kynurenines in the CNS: recent advances and new questions. Nat Rev Drug Discov. 2013;12(1):64–82.

243. Guo S, Vecsei L, Ashina M. The L-kynurenine signalling pathway in trigeminal pain processing: a potential therapeutic target in migraine? Cephalalgia. 2011;31(9):1029–38.

244. Curto M, Lionetto L, Negro A, Capi M, Fazio F, Giamberardino MA, Simmaco M, Nicoletti F, Martelletti P. Altered kynurenine pathway metabolites in serum of chronic migraine patients. J Headache Pain. 2016;17(1):47.

245. Fazio F, Lionetto L, Curto M, Iacovelli L, Copeland CS, Neale SA, Bruno V, Battaglia G, Salt TE, Nicoletti F. Cinnabarinic acid and xanthurenic acid: two kynurenine metabolites that interact with metabotropic glutamate receptors. Neuropharmacology. 2017;112(Pt B):365–72.

246. Perkins MN, Stone TW. An iontophoretic investigation of the actions of convulsant kynurenines and their interaction with the endogenous excitant quinolinic acid. Brain Res. 1982;247(1):184–7.

247. Stone TW, Perkins MN. Quinolinic acid: a potent endogenous excitant at amino acid receptors in CNS. Eur J Pharmacol. 1981;72(4):411–2.

248. Schwarcz R, Whetsell WO, Mangano RM. Quinolinic acid: an endogenous metabolite that produces axon-sparing lesions in rat brain. Science. 1983;219(4582):316–8.

249. Okuda S, Nishiyama N, Saito H, Katsuki H. 3-Hydroxykynurenine, an endogenous oxidative stress generator, causes neuronal cell death with apoptotic features and region selectivity. J Neurochem. 1998;70(1):299–307.

250. Ho M-L, Tsai T-N, Chang J-K, Shao T-S, Jeng Y-R, Hsu C. Down-regulation of N-methyl D-aspartate receptor in rat-modeled disuse osteopenia. Osteoporos Int. 2005;16(12):1780–8.

251. Elefteriou F. Neuronal signaling and the regulation of bone remodeling. Cell Mol Life Sci. 2005;62(19–20):2339–49.

252. Itzstein C, Cheynel H, Burt-Pichat B, Merle B, Espinosa L, Delmas PD, Chenu C. Molecular identification of NMDA glutamate receptors expressed in bone cells. J Cell Biochem. 2001;82(1):134–44.

253. Forrest CM, Kennedy A, Stone TW, Stoy N, Darlington LG. Kynurenine and neopterin levels in patients with rheumatoid arthritis and osteoporosis during drug treatment. Adv Exp Med Biol. 2003;527:287–95.

254. Forrest CM, Mackay GM, Oxford L, Stoy N, Stone TW, Darlington LG. Kynurenine pathway metabolism in patients with osteoporosis after 2 years of drug treatment. Clin Exp Pharmacol Physiol. 2006;33(11):1078–87.

255. Igari T, Tsuchizawa M, Shimamura T. Alteration of tryptophan metabolism in the synovial fluid of patients with rheumatoid arthritis and osteoarthritis. Tohoku J Exp Med. 1987;153(2):79–86.

256. Maiuolo J, Oppedisano F, Gratteri S, Muscoli C, Mollace V. Regulation of uric acid metabolism and excretion. Int J Cardiol. 2016;213:8–14.

257. Albrecht E, Waldenberger M, Krumsiek J, Evans AM, Jeratsch U, Breier M, Adamski J, Koenig W, Zeilinger S, Fuchs C, Klopp N, Theis FJ, Wichmann H-E, Suhre K, Illig T, Strauch K, Peters A, Gieger C, Kastenmüller G, Doering A, Meisinger C. Metabolite profiling reveals new insights into the regulation of serum urate in humans. Metabolomics. 2014;10(1):141–51.

258. Jin M, Yang F, Yang I, Yin Y, Luo JJ, Wang H, Yang X-F. Uric acid, hyperuricemia and vascular diseases. Front Biosci. 2012;17:656–69.

259. Busso N, So A. Mechanisms of inflammation in gout. Arthritis Res Ther. 2010;12(2):206.

260. Zhang ZH, Wei F, Vaziri ND, Cheng XL, Bai X, Lin RC, Zhao YY.

Metabolomics insights into chronic kidney disease and modulatory effect of rhubarb against tubulointerstitial fibrosis. Sci Rep. 2015;5:14472.

261. Men L, Pi Z, Zhou Y, Liu Y, Wei M, Song F, Liu Z. Metabolomics insights into diabetes nephropathy and protective effects of Radix Scutellariae on rats using ultra-high performance liquid chromatography coupled with quadrupole time-of-flight mass spectrometry. RSC Adv. 2017;7(27):16494–504.

262. Kand'ár R, Žáková P. Allantoin as a marker of oxidative stress in human erythrocytes. Clin Chem Lab Med. 2008;46(9):1270–4.

263. Yardim-Akaydin S, Sepici A, Özkan Y, Şimşek B, Sepici V. Evaluation of allantoin levels as a new marker of oxidative stress in Behçet's disease. Scand J Rheumatol. 2006;35(1):61–4.

264. Masuda T, Shingai Y, Takahashi C, Inai M, Miura Y, Honda S, Masuda A. Identification of a potent xanthine oxidase inhibitor from oxidation of caffeic acid. Free Radic Biol Med. 2014;69:300–7.

265. Chan WS, Wen PC, Chiang HC. Structure-activity relationship of caffeic acid analogues on xanthine oxidase inhibition. Anticancer Res. 1995;15(3):703–7.

266. Chung T-W, Moon S-K, Chang Y-C, Ko J-H, Lee Y-C, Cho G, Kim S-H, Kim J-G, Kim C-H. Novel and therapeutic effect of caffeic acid and caffeic acid phenyl ester on hepatocarcinoma cells: complete regression of hepatoma growth and metastasis by dual mechanism. FASEB J. 2004;18(14):1670–81.

267. Anwar J, Spanevello RM, Pimentel VC, Gutierres J, Thomé G, Cardoso A, Zanini D, Martins C, Palma HE, Bagatini MD, Baldissarelli J, Schmatz R, Leal CAM, da Costa P, Morsch VM, Schetinger MRC. Caffeic acid treatment alters the extracellular adenine nucleotide hydrolysis in platelets and lymphocytes of adult rats. Food Chem Toxicol. 2013;56:459–66.

268. Ban JY, Cho SO, Koh SB, Song K-S, Bae K, Seong YH. Protection of amyloid beta protein (25-35)-induced neurotoxicity by methanol extract of Smilacis chinae rhizome in cultured rat cortical neurons. J Ethnopharmacol. 2006;106(2):230–7.

269. Chen JH, Ho C. Antioxidant activities of caffeic acid and its related hydroxycinnamic acid compounds. J Agric Food Chem. 1997;45(7):2374–8.

270. Kaddurah-Daouk R, Zhu H, Sharma S, Bogdanov M, Rozen SG, Matson W, Oki NO, Motsinger-Reif AA, Churchill E, Lei Z, Appleby D, Kling MA, Trojanowski JQ, Doraiswamy PM, Arnold SE. Alterations in metabolic pathways and networks in Alzheimer's disease. Transl Psychiatry. 2013;3:e244.

271. Kaddurah-Daouk R, Rozen S, Matson W, Han X, Hulette CM, Burke JR, Doraiswamy PM, Welsh-Bohmer KA. Metabolomic changes in autopsy-confirmed Alzheimer's disease. Alzheimer's Dement. 2011;7(3):309–17.

272. Lord RS, Bralley JA. Clinical applications of urinary organic acids. Part I: Detoxification markers. Altern Med Rev. 2008;13(3):205–15.

273. Banerjee R, Zou C-G. Redox regulation and reaction mechanism of human cystathionine-beta-synthase: a PLP-dependent hemesensor protein. Arch Biochem Biophys. 2005;433(1):144–56.

274. Doorn J, Leusink M, Groen N, van de Peppel J, van Leeuwen JPTM, van Blitterswijk CA, de Boer J. Diverse effects of cyclic AMP variants on osteogenic and adipogenic differentiation of human mesenchymal stromal cells. Tissue Eng Part A. 2012;18(13–14):1431–42.

275. Ifegwu OC, Awale G, Rajpura K, Lo KW-H, Laurencin CT. Harnessing cAMP signaling in musculoskeletal regenerative engineering. Drug Discov Today. 2017;22(7):1027–44.

276. Poesen R, Mutsaers HAM, Windey K, van den Broek PH, Verweij V, Augustijns P, Kuypers D, Jansen J, Evenepoel P, Verbeke K, Meijers B, Masereeuw R. The influence of dietary protein intake on mammalian tryptophan and phenolic metabolites. PLoS One. 2015;10(10):e0140820.

277. Chaudhary K, Malhotra K, Sowers J, Aroor A. Uric Acid - key ingredient in the recipe for cardiorenal metabolic syndrome. Cardiorenal Med. 2013;3(3):208–20.

278. Andrew R, Watson DG, Best SA, Midgley JM, Wenlong H, Petty RKH. The determination of hydroxydopamines and other trace amines in the urine of Parkinsonian patients and normal controls. Neurochem Res. 1993;18(11):1175–7.

279. Brun L, Ngu LH, Keng WT, Ch'ng GS, Choy YS, Hwu WL, Lee WT, Willemsen MAAP, Verbeek MM, Wassenberg T, Regal L, Orcesi S, Tonduti D, Accorsi P, Testard H, Abdenur JE, Tay S, Allen GF, Heales S, Kern I, Kato M, Burlina A, Manegold C, Hoffmann GF, Blau N. Clinical and biochemical features of aromatic L-amino acid decarboxylase deficiency. Neurology. 2010;75(1):64–71.

280. Hakim AA, Thiele KA. Conversion of tryptophan to kynurenine, and serotonin to kynuramine. Biochem Biophys Res Commun. 1960;2(3):242–7.

281. Raynor W, Houshmand S, Gholami S, Emamzadehfard S, Rajapakse CS, Blomberg BA, Werner TJ, Høilund-Carlsen PF, Baker JF, Alavi A. Evolving role of molecular imaging with 18F-sodium fluoride PET as a biomarker for calcium metabolism. Curr Osteoporos Rep. 2016;14(4):115–25.

282. Kobayashi N, Inaba Y, Tateishi U, Yukizawa Y, Ike H, Inoue T, Saito T. New application of 18F-fluoride PET for the detection of bone remodeling in early-stage osteoarthritis of the hip. Clin Nucl Med. 2013;38(10):e379-83.

283. Corciulo C, Lendhey M, Wilder T, Schoen H, Cornelissen AS, Chang G, Kennedy OD, Cronstein BN. Endogenous adenosine maintains cartilage homeostasis and exogenous adenosine inhibits osteoarthritis progression. Nat Commun. 2017;8(1):15019.

284. Cutolo M, Sulli A, Pizzorni C, Seriolo B, Straub RH. Anti-inflammatory mechanisms of methotrexate in rheumatoid arthritis. Ann Rheum Dis. 2001;60(8):729–35.

285. Morgan SL, Oster RA, Lee JY, Alarcón GS, Baggott JE. The effect of folic acid and folinic acid supplements on purine metabolism in methotrexate-treated rheumatoid arthritis. Arthritis Rheum. 2004;50(10):3104–11.

286. Mistry D, Chambers MG, Mason RM. The role of adenosine in chondrocyte death in murine osteoarthritis and in a murine chondrocyte cell line. Osteoarthr Cartil. 2006;14(5):486–95.

287. Johnson SM, Patel S, Bruckner FE, Collins DA. 5'-Nucleotidase as a marker of both general and local inflammation in rheumatoid arthritis patients. Rheumatology. 1999;38(5):391–6.

288. Win AZ, Aparici CM. Normal SUV values measured from NaF18- PET/CT bone scan studies. PLoS One. 2014;9(9):e108429.

289. Ferro DR, Provasoli A, Ragazzi M, Casu B, Torri G, Bossennec V, Perly B, Sinaÿ P, Petitou M, Choay J. Conformer populations of L-iduronic acid residues in glycosaminoglycan sequences. Carbohydr Res. 1990;195(2):157–67.

290. Buckwalter JA, Mankin HJ. Articular cartilage: tissue design and chondrocyte-matrix interactions. Instr Course Lect. 1998;47:477–86.

291. Gauza-Włodarczyk M, Kubisz L, Włodarczyk D. Amino acid composition in determination of collagen origin and assessment of physical factors effects. Int J Biol Macromol. 2017;104:987–91.

292. Klein BY, Gal I, Segal D. Selection of malonate-resistant stromal cell-derived osteoprogenitor cells in vitro. J Cell Biochem. 1993 Feb;51(2):190–7.

293. Kim KS. Sex difference in histamine metabolism in rats. Am J Physiol Content. 1959;197(6):1258–60.

294. Netter KJ, Cohn VH, Shore PA. Sex difference in histamine metabolism in the rat. Am J Physiol Content. 1961;201(2):224–6.

295. Said HM, Ross AC. Roboflavin. In: Ross AC, Caballero B, Cousins RJ, Tucker KL, Ziegler TR, editors. Modern Nutrition in Health and Disease. 11th edition. Baltimore, MD: Lippincott Williams & Wilkins; 2014. p. 325–30.

296. Choi JY, Kim Y-N, Cho Y-O. Evaluation of riboflavin intakes and status of 20-64-year-old adults in South Korea. Nutrients. 2014;7(1):253–64.

297. Caldwell J, Anthony A, Cotgreave IA, Sangster SA, Sutton JD, Bernard BK, Ford RA. Influence of dose and sex on the disposition and hepatic effects of cinnamyl anthranilate in the B6C3F1 mouse. Food Chem Toxicol. 1985;23(6):559–66.

298. Sumi S, Kidouchi K, Kondou M, Hayashi K, Dobashi K, Kouwaki M, Togari H, Wada Y. Possible prediction of adverse reactions to fluorouracil by the measurement of urinary dihydrothymine and thymine. Int J Mol Med. 1998;2(4):477–82.

299. Zhang A, Sun H, Wu X, Wang X. Urine metabolomics. Clin Chim Acta. 2012;414:65–9.

300. Eisenhofer G, Kopin IJ, Goldstein DS. Catecholamine metabolism: a contemporary view with implications for physiology and medicine. Pharmacol Rev. 2004;56(3):331–49.

301. Boobis AR, Murray S, Jones DH, Reid JL, Davies DS. Urinary conjugates of 4-hydroxy-3-methoxyphenylethylene glycol do not provide an index of brain amine turnover in man. Clin Sci. 1980;58(4):311–6.

302. Hughes JH, Liu K, Sutherland H, Wilson PJ, Hughes AT, Milan AM, Ranganath LR, Gallagher JA, Bou-Gharios G. Generation and phenotyping of a targeted mouse model of alkaptonuria. Osteoarthr Cartil. 2018;26:S89-90.

303. Norman BP, Davison AS, Ross GA, Milan AM, Hughes AT, Sutherland H, Jarvis JC, Roberts NB, Gallagher JA, Ranganath LR. A comprehensive LC-QTOF-MS metabolic phenotyping strategy: Application to alkaptonuria. Clin Chem. 2019;65(4):530–9.

304. Laboratory Animal Science Association. Good practice guidelines: collection of blood samples (rat, mouse, guinea pig, rabbit). 1998.

305. Garrod A. The Croonian lectures on inborn errors of metabolism. Lancet. 1908;172(4428):73–9.

306. Taylor AM, Boyde A, Wilson PJM, Jarvis JC, Davidson JS, Hunt JA, Ranganath LR, Gallagher JA. The role of calcified cartilage and subchondral bone in the initiation and progression of ochronotic arthropathy in alkaptonuria. Arthritis Rheum. 2011;63(12):3887–96.

307. Hegedus ZL, Nayak U. Homogentisic acid and structurally related compounds as intermediates in plasma soluble melanin formation and in tissue toxicities. Arch Physiol Biochem. 1994;102(3):175–81.

308. Wolff TF, Jordan RA. Basic concepts in drug metabolism: Part I. J Clin Pharmacol. 1987;27(1):15–7.

309. Williams RT. The metabolism of certain drugs and food chemicals in man. Ann N Y Acad Sci. 1971;179(1):141–54.

310. Tephly TR, Burchell B. UDP-glucuronosyltransferases: a family of detoxifying enzymes. Trends Pharmacol Sci. 1990;11(7):276–9.

311. Pirmohamed M, Park BK. Mechanisms of adverse drug reactions. In: Mann D, Andrews EB, editors. Pharmacovigilance, 2nd ed. New Jersey: John Wiley and Sons; 2007. p. 85–103.

312. Tukey RH, Strassburg CP. Human UDP-glucuronosyltransferases: metabolism, expression, and disease. Annu Rev Pharmacol Toxicol. 2000;40(1):581–616.

313. Burchell B, Nebert DW, Nelson DR, Bock KW, Iyanagi T, Jansen PL, Lancet D, Mulder GJ, Chowdhury JR, Siest G, Tephly TR, Mackenzie PI. The UDP glucuronosyltransferase gene superfamily: suggested nomenclature based on evolutionary divergence. DNA Cell Biol. 1991;10(7):487–94.

314. Pirmohamed M, Madden S, Park BK. Idiosyncratic drug reactions. Metabolic bioactivation as a pathogenic mechanism. Clin Pharmacokinet. 1996;31(3):215–30.

315. Pirmohamed M, Kitteringham NR, Park BK. The role of active metabolites in drug toxicity. Drug Saf. 1994;11(2):114–44.

316. Hughes TB, Miller GP, Swamidass SJ. Modeling epoxidation of drug-like molecules with a deep machine learning network. ACS Cent Sci. 2015;1(4):168–80.

317. Obach RS, Kalgutkar AS. Reactive electrophiles and metabolic activation. In: McQueen CA, editor. Comprehensive Toxicology. Amsterdam: Elsevier Science; 2010. p. 309–47.

318. Davison AS, Harrold JA, Hughes G, Norman BP, Devine J, Usher J, Hughes AT, Khedr M, Gallagher JA, Milan AM, J.C.G. H, Ranganath LR. Clinical and biochemical assessment of depressive symptoms in patients with Alkaptonuria before and after two years of treatment with nitisinone. Mol Genet Metab. 2018;125(1–2):135–43.

319. Weng R, Shen S, Tian Y, Burton C, Xu X, Liu Y, Chang C, Bai Y, Liu H. Metabolomics approach reveals integrated metabolic network associated with serotonin deficiency. Sci Rep. 2015;5:11864.

320. Kang CW, Han YE, Kim J, Oh JH, Cho YH, Lee EJ. 4-Hydroxybenzaldehyde accelerates acute wound healing through activation of focal adhesion signalling in keratinocytes. Sci Rep. 2017;7(1):14192.

321. O'Meara A, Tormey W, FitzGerald RJ, Fitzgibbon M, Kenny D. Interpretation of random urinary catecholamines and their metabolites in neuroblastoma. Acta Paediatr. 1994;83(1):88–92.

322. Verly IRN, van Kuilenburg ABP, Abeling NGGM, Goorden SMI, Fiocco M, Vaz FM, van Noesel MM, Zwaan CM, Kaspers GL, Merks JHM, Caron HN, Tytgat GAM. Catecholamines profiles at diagnosis: Increased diagnostic sensitivity and correlation with biological and clinical features in neuroblastoma patients. Eur J Cancer. 2017;72:235–43.

323. Maiuolo J, Oppedisano F, Gratteri S, Muscoli C, Mollace V. Regulation of uric acid metabolism and excretion. Int J Cardiol. 2016;213:8–14.

324. Kristal BS, Vigneau-Callahan KE, Moskowitz AJ, Matson WR. Purine catabolism: Links to mitochondrial respiration and antioxidant defenses? Arch Biochem Biophys. 1999;370(1):22–33.

325. Yao JK, Dougherty GG, Reddy RD, Keshavan MS, Montrose DM, Matson WR, McEvoy J, Kaddurah-Daouk R. Homeostatic imbalance of purine catabolism in first-episode neuroleptic-naïve patients with schizophrenia. PLoS One. 2010;5(3):e9508.

326. Hallan S, Afkarian M, Zelnick LR, Kestenbaum B, Sharma S, Saito R, Darshi M, Barding G, Raftery D, Ju W, Kretzler M, Sharma K, de Boer IH. Metabolomics and gene expression analysis reveal down-regulation of the citric acid (TCA) cycle in non-diabetic CKD patients. EBioMedicine. 2017;26:68–77.

327. Wolff F, Biaou I, Koopmansch C, Bossche M Vanden, Pozdzik A, Roumeguère T, Cotton F. Renal and prostate stones composition in alkaptonuria: a case report. Clin Nephrol. 2015;84(6):339–42.

328. Zuckerman JM, Assimos DG. Hypocitraturia: pathophysiology and medical management. Rev Urol. 2009;11(3):134–44.

329. Stenn F, Milgram J, Lee S, Weigand R, Veis A. Biochemical identification of homogentisic acid pigment in an ochronotic egyptian mummy. Science. 1977;197(4303):566–8.

330. Wilke T, Schneider M, Kleinermanns K. 1,4-Hydroquinone is a hydrogen reservoir for fuel cells and recyclable via photocatalytic water splitting. Open J Phys Chem. 2013;03(02):97–102.

331. Seitzer PM, Searle BC. Incorporating in-source fragment information improves metabolite identification accuracy in untargeted LC–MS data sets. J Proteome Res. 2019;18(2):791–6.

332. Pei J, Hsu CC, Wang Y, Yu K. Corona discharge-induced reduction of quinones in negative electrospray ionization mass spectrometry. RSC Adv. 2017;7(69):43540–5.

333. Pei J, Hsu CC, Wang Y, Yu K. Corona discharge-induced reduction of quinones in negative electrospray ionization mass spectrometry. RSC Adv. 2017;7(69):43540–5.

334. Eslami M, Namazian M, Zare HR. Electrooxidation of homogentisic acid in aqueous and mixed solvent solutions: experimental and theoretical studies. J Phys Chem B. 2013;117(9):2757–63.

335. Skarżyński B, Sarnecka-Keller M, Frendo J. The bound form of homogentisic acid in alkaptonuric urine. Clin Chim Acta. 1962;7(2):243–7.

336. Chow WY, Taylor AM, Reid DG, Gallagher JA, Duer MJ. Collagen atomic scale molecular disorder in ochronotic cartilage from an alkaptonuria patient, observed by solid state NMR. J Inherit Metab Dis. 2011;34(6):1137–40.

337. Märker K, Paul S, Fernández-De-Alba C, Lee D, Mouesca JM, Hediger S, De Paëpe G. Welcoming natural isotopic abundance in solid-state NMR: probing π-stacking and supramolecular structure of organic nanoassemblies using DNP. Chem Sci. 2017;8(2):974–87.

338. Märker K, Pingret M, Mouesca J-M, Gasparutto D, Hediger S, De Paëpe G. A new tool for NMR crystallography: complete 13 C/ 15 N assignment of organic molecules at natural isotopic abundance using DNP-enhanced solid-state NMR. J Am Chem Soc. 2015;137(43):13796–9.

339. Takahashi H, Hediger S, De Paëpe G. Matrix-free dynamic nuclear polarization enables solid-state NMR 13C–13C correlation spectroscopy of proteins at natural isotopic abundance. Chem Commun. 2013;49(82):9479.

340. Takahashi H, Lee D, Dubois L, Bardet M, Hediger S, De Paëpe G. Rapid natural-abundance 2D ^{13} C- ^{13} C correlation spectroscopy using dynamic nuclear polarization enhanced solid-state NMR and matrix-free sample preparation. Angew Chemie Int Ed. 2012;51(47):11766–9.

341. Singh C, Rai RK, Aussenac F, Sinha N. Direct evidence of imino acid-aromatic interactions in native collagen protein by DNP-enhanced solid-state NMR spectroscopy. J Phys Chem Lett. 2014;5(22):4044–8.

342. Sauvée C, Rosay M, Casano G, Aussenac F, Weber RT, Ouari O, Tordo P. Highly efficient, water-soluble polarizing agents for dynamic nuclear polarization at high frequency. Angew Chemie - Int Ed. 2013;125(41):11058–61.

343. Jagtap AP, Geiger MA, Stöppler D, Orwick-Rydmark M, Oschkinat H, Sigurdsson ST. BcTol: A highly water-soluble biradical for efficient dynamic nuclear polarization of biomolecules. Chem Commun. 2016;52(43):7020–3.

344. Geiger MA, Jagtap AP, Kaushik M, Sun H, Stöppler D, Sigurdsson ST, Corzilius B, Oschkinat H. Efficiency of water-soluble nitroxide biradicals for dynamic nuclear polarization in rotating solids at 9.4 T: bcTol-M and cyolyl-TOTAPOL as new polarizing agents. Chem - A Eur J. 2018;24(51):13485–94.

345. Na GC. Interaction of calf skin collagen with glycerol: linked function analysis. Biochemistry. 1986;25(5):967–73.

346. Rossini AJ, Zagdoun A, Lelli M, Lesage A, Copéret C, Emsley L. Dynamic nuclear polarization surface enhanced NMR spectroscopy. Acc Chem Res. 2013;46(9):1942–51.

347. Lesage A, Lelli M, Gajan D, Caporini MA, Vitzthum V, Miéville P, Alauzun J, Roussey A, Thieuleux C, Mehdi A, Bodenhausen G, Copéret C, Emsley L. Surface enhanced NMR spectroscopy by dynamic nuclear polarization. J Am Chem Soc. 2010;132(44):15459–61.

348. Metz G, Wu X, Smith SO. Ramped-amplitude cross polarization in magic-angle-spinning NMR. J Magn Reson - Ser A. 1994;110(2):219–27.

349. Burum DP, Ernst RR. Net polarization transfer via a J-ordered state for signal enhancement of low-sensitivity nuclei. J Magn Reson. 1980;39(1):163–8.

350. Morris GA, Freeman R. Enhancement of nuclear magnetic resonance signals by polarization transfer. J Am Chem Soc. 1979;101(3):760–2.

351. Nowacka A, Bongartz NA, Ollila OHS, Nylander T, Topgaard D. Signal intensities in 1H-13C CP and INEPT MAS NMR of liquid crystals. J Magn Reson. 2013;230:165–75.

352. Bielecki A, Kolbert AC, Levitt MH. Frequency-switched pulse sequences: Homonuclear decoupling and dilute spin NMR in solids. Chem Phys Lett. 1989;155(4-5):341–6.

353. Van Rossum BJ, Förster H, De Groot HJM. High-Field and High-Speed CP-MAS 13 C NMR Heteronuclear Dipolar-Correlation Spectroscopy of Solids with Frequency-Switched Lee-Goldburg Homonuclear Decoupling. J Magn Reson. 1997;124(2):516–9.

354. Mason JT, O'Leary TJ. Effects of formaldehyde fixation on protein secondary structure: A calorimetric and infrared spectroscopic investigation. J Histochem Cytochem. 1991;39(2):225–9.

355. Fox CH, Johnson FB, Whiting J, Roller PP. Formaldehyde fixation. J Histochem Cytochem. 1985;33(8):845–53.

356. Thavarajah R, Mudimbaimannar V, Rao U, Ranganathan K, Elizabeth J. Chemical and physical basics of routine formaldehyde fixation. J Oral Maxillofac Pathol. 2012;16(3):400–5.

357. Fishbein KW, Gluzband YA, Kaku M, Ambia-Sobhan H, Shapses SA, Yamauchi M, Spencer RG. Effects of formalin fixation and collagen cross-linking on T2 and magnetization transfer in bovine nasal cartilage. Magn Reson Med. 2007;57(6):1000–11.

358. Macura S, Mishra PK, Gamez JD, Pirko I. MR microscopy of formalin fixed paraffin embedded histology specimens. Magn Reson Med. 2014;71(6):1989–94.

359. Thickman D, Kundel H, Wolf G. Nuclear magnetic resonance characteristics of fresh and fixed tissue: the effect of elapsed time. Radiology. 1983;148(1):183–5.

360. Yuan M, Breitkopf SB, Yang X, Asara JM. A positive/negative ion–switching, targeted mass spectrometry–based metabolomics platform for bodily fluids, cells, and fresh and fixed tissue. Nat Protoc. 2012;7(5):872–81.

361. Kelly AD, Breitkopf SB, Yuan M, Goldsmith J, Spentzos D, Asara JM. Metabolomic profiling from formalin-fixed, paraffin-embedded tumor tissue using targeted LC/MS/MS: Application in sarcoma. PLoS One. 2011;6(10):e25357.

362. Buck A, Ly A, Balluff B, Sun N, Gorzolka K, Feuchtinger A, Janssen KP, Kuppen PJK, Van De Velde CJH, Weirich G, Erlmeier F, Langer R, Aubele M, Zitzelsberger H, Aichler M, Walch A. High-resolution MALDI-FT-ICR MS imaging for the analysis of metabolites from formalin-fixed, paraffin-embedded clinical tissue samples. J Pathol. 2015;237(1):123–32.

363. Wojakowska A, Marczak Ł, Jelonek K, Polanski K, Widlak P, Pietrowska M. An optimized method of metabolite extraction from formalin-fixed paraffin-embedded tissue for GC/MS analysis. PLoS One. 2015 Sep 8;10(9):e0136902.

364. Cacciatore S, Zadra G, Bango C, Penney KL, Tyekucheva S, Yanes O, Loda M. Metabolic profiling in formalin-fixed and paraffin-embedded prostate cancer tissues. Mol Cancer Res. 2017;15(4):439–47.

365. Chatterjee S, Prados-Rosales R, Itin B, Casadevall A, Stark RE. Solid-state NMR reveals the carbon-based molecular architecture of Cryptococcus

neoformans fungal eumelanins in the cell wall. J Biol Chem. 2015;290(22):13779–90.

366. Thureau P, Ziarelli F, Thévand A, Martin RW, Farmer PJ, Viel S, Mollica G. Probing the motional behavior of eumelanin and pheomelanin with solid-state NMR spectroscopy: New insights into the pigment properties. Chem - A Eur J. 2012;18(34),(34):10689–700.

367. Hervé M, Hirschinger J, Granger P, Gilard P, Deflandre A, Goetz N. A 13C solid-state NMR study of the structure and auto-oxidation process of natural and synthetic melanins. Biochim Biophys Acta - Protein Struct Mol Enzymol. 1994;1204(1):19–27.

368. Fišer-Herman M, Petrovački M. Reduzierende substanzen aus alkaptonurischem harn. Clin Chim Acta Int J Clin Chem. 1958;3(3):248–52.

369. Kalinowski H-O, Berger S, Braun S, Griffiths L. Carbon-13 NMR spectroscopy. Anal Chim Acta1. 1989;221:367–8.

370. Madkour TM. Chemistry of polymerization products of p-benzoquinone. 13C NMR and molecular dynamics study. Polym J. 1997;29(8):670–7.

371. Scheffer JR, Wong YF, Patil AO, Curtin DY, Paul IC. CPMAS (cross polarization magic angle spinning) carbon-13 NMR spectra of quinones, hydroquinones, and their complexes. Use of CMR to follow a reaction in the solid state. J Am Chem Soc. 1985;107(17):4898–904.

372. Hayashi N, Yoshikawa T, Ohnuma T, Higuchi H, Sako K, Uekusa H. Synthesis, structure, and properties of benzoquinone dimer and trimers bearing t-Bu substituents. Org Lett. 2007;9(26):5417–20.

373. Erdtman H, Granath M, Schultz G, Sörensen NA. Studies on humic acids. VI. Triquinone and tetraquinone. Acta Chem Scand. 1954;8:1442–50.

374. Guin PS, Das S, Mandal PC. Electrochemical Reduction of Quinones in Different Media: A Review. Int J Electrochem. 2011;1–22.

375. Sabaa MW, Madkour TM, Yassin AA. Polymerization products of p-benzoquinone as bound antioxidants for SBR. Part II - the antioxidizing efficiency. Polym Degrad Stab. 1988;22(3):205–22.

376. Sabaa MW, Madkour TM, Yassin AA. Polymerization products of p-benzoquinone as bound antioxidants for styrene-butadiene rubber: Part I - preparation of quinone polymers. Polym Degrad Stab. 1988;22(3):195–203.

377. Curtis SL, Norman BP, Milan AM, Gallagher JA, Olsson B, Ranganath LR, Roberts NB. Interference of hydroxyphenylpyruvic acid, hydroxyphenyllactic acid and tyrosine on routine serum and urine clinical chemistry assays; implications for biochemical monitoring of patients with alkaptonuria treated with nitisinone. Clin Biochem. 2019;71:24–30.

378. Keith KE, Killip L, He P, Moran GR, Valvano MA. Burkholderia cenocepacia C5424 produces a pigment with antioxidant properties using a homogentisate intermediate. J Bacteriol. 2007;189(24):9057–65.

379. López-Mirabal HR, Winther JR. Redox characteristics of the eukaryotic cytosol. Biochim Biophys Acta - Mol Cell Res. 2008;1783(4):629–40.

380. Ottaviano FG, Handy DE, Loscalzo J. Redox regulation in the extracellular environment. Circ J. 2008;72(1):1–16.

381. Taylor AM, Wlodarski B, Prior IA, Wilson PJM, Jarvis JC, Ranganath LR, Gallagher JA. Ultrastructural examination of tissue in a patient with alkaptonuric arthropathy reveals a distinct pattern of binding of ochronotic pigment. Rheumatology. 2010;49(7):1412–14.

382. Melis M, Onori P, Aliberti G, Vecci E, Gaudio E. Ochronotic arthropathy: Structural and ultrastructural features. Ultrastruct Pathol. 1994;18(5):467–71.

383. Bank RA, Bayliss MT, Lafeber FPJG, Maroudas A, Tekoppele JM. Ageing and zonal variation in post-translational modification of collagen in normal human articular cartilage: The age-related increase in non-enzymatic glycation affects biomechanical properties of cartilage. Biochem J.

1998;330(1):345–51.

384. Jørgensen AEM, Kjær M, Heinemeier KM. The effect of aging and mechanical loading on the metabolism of articular cartilage. J Rheumatol. 2017;44(4):410–7.

385. Wen CY, Wu CB, Tang B, Wang T, Yan CH, Lu WW, Pan H, Hu Y, Chiu KY. Collagen fibril stiffening in osteoarthritic cartilage of human beings revealed by atomic force microscopy. Osteoarthr Cartil. 2012;20(8):916–22.

386. Kim J-H, Lee G, Won Y, Lee M, Kwak J-S, Chun C-H, Chun J-S. Matrix cross-linking-mediated mechanotransduction promotes posttraumatic osteoarthritis. Proc Natl Acad Sci. 2015;112(30):9424–29.

387. Yamauchi K, Kuroki S, Ando I. The amide proton NMR chemical shift and hydrogen-bonded structure of glycine-containing peptides and polypeptides in the solid state as studied by multi-pulse-associated high-speed MAS 1H NMR. J Mol Struct. 2002;602–603:9–16.

388. Yamauchi K, Kuroki S, Fujii K, Ando I. The amide proton NMR chemical shift and hydrogen-bonded structure of peptides and polypeptides in the solid state as studied by high-frequency solid-state 1H NMR. Chem Phys Lett. 2000;324(5–6):435–439.

389. Lees GJ, Weiner N. Transaminations between amino acids and keto acids elevated in phenylketonuria and maple syrup urine disease. J Neurochem. 1973;20(2):389–403.

390. Harding CO. Gene and cell therapy for inborn errors of metabolism. In: Hoffmann GF, Zschocke J, Nyhan WL, editors. Inherited metabolic diseases. Berlin, Heidelberg: Springer; 2017. p. 155–71.

391. Brunetti-Pierri N. Gene therapy for inborn errors of liver metabolism: Progress towards clinical applications. Italian Journal of Pediatrics. 2008.

392. Weidner SM, Trimpin S. Mass spectrometry of synthetic polymers. Anal Chem. 2010;82(12):4811–29.

393. Montaudo G, Samperi F, Montaudo MS. Characterization of synthetic polymers by MALDI-MS. Prog Polym Sci. 2006;31(3):277–357.

394. Wills R. The works of W. Harvey (transl.). London: Sydenham Society; 1847. 616 p.

10.0   SUPPORTING MATERIAL

Appendix 1

Method 1: urine unique AMRT matches obtained (negative polarity)

Name	Formula	RT (Tgt)	Mass (Tgt)
2,3-Dihydroxybenzoic acid	C7 H6 O4	3.23	154.0266
2-Hydroxyphenylacetic acid	C8 H8 O3	4.258	152.0473
2-METHYLMALEATE	C5 H6 O4	1.168	130.0266
3-hydroxy-3-methyl-Glutaric acid	C6 H10 O5	1.218	162.0528
4-acetamidobutanoate	C6 H11 N O3	1.339	145.0739
4-Hydroxybenzaldehyde	C7 H6 O2	3.746	122.0368
4-Hydroxyphenyllactic acid	C9 H10 O4	2.928	182.0579
4-Pyridoxic acid	C8 H9 N O4	1.305	183.0532
Acetoacetic acid	C4 H6 O3	0.947	102.0317
Caffeic Acid	C9 H8 O4	3.836	180.0423
Citramalic acid	C5 H8 O5	1.052	148.0372
Citric acid	C6 H8 O7	0.8	192.027
Ethylmalonic acid	C5 H8 O4	1.946	132.0423
Hippuric acid	C9 H9 N O3	3.674	179.0582
Homogentisic acid	C8 H8 O4	1.619	168.0423
Homovanillic acid	C9 H10 O4	4.193	182.0579
Hypoxanthine	C5 H4 N4 O	0.847	136.0385
L-Phenylalanine	C9 H11 N O2	2.045	165.079
L-Tryptophan	C11 H12 N2 O2	2.82	204.0899
L-Tyrosine	C9 H11 N O3	0.917	181.0739
Methyl N-(a-methylbutyryl)glycine	C9 H16 O4	6.455	188.1049
m-Salicylic acid	C7 H6 O3	3.749	138.0317
N-Acetyl-L-phenylalanine	C11 H13 N O3	5.14	207.0895
N-Acetylserotonin	C12 H14 N2 O2	3.639	218.1055
Oxoadipic acid	C6 H8 O5	0.863	160.0372
Pantothenic Acid	C9 H17 N O5	2.731	219.1107
Phenylpyruvic acid	C9 H8 O3	3.546	164.0473
p-Hydroxyphenylacetic acid	C8 H8 O3	3.585	152.0473
Pimelic acid	C7 H12 O4	4.179	160.0736
p-Salicylic acid	C7 H6 O3	3.116	138.0317
Pyroglutamic acid	C5 H7 N O3	0.897	129.0426
Salicylic acid	C7 H6 O3	4.994	138.0317
Suberic acid	C8 H14 O4	5.39	174.0892
Theophylline	C7 H8 N4 O2	3.366	180.0647
trans-Aconitate	C6 H6 O6	0.982	174.0164
Uridine	C9 H12 N2 O6	0.932	244.0695
Vanillylmandelic acid	C9 H10 O5	1.399	198.0528

Xanthine   C5 H4 N4 O2   0.903   152.0334

Method 1: urine unique AMRT matches obtained (positive polarity)

Name	Formula	RT (Tgt)	Mass (Tgt)
(+)-α-Tocopherol	C29 H50 O2	11.652	430.3811
(±)-Mevalonolactone	C6 H10 O3	1.472	130.063
12-Hydroxydodecanoic acid	C12 H24 O3	9.022	216.1725
1-Methyladenosine	C11 H15 N5 O4	2.719	281.1124
2,3-Pyridinedicarboxylic acid	C7 H5 N O4	0.825	167.0219
2-Aminophenol	C6 H7 N O	0.975	109.0528
2-Hydroxyphenylacetic acid	C8 H8 O3	4.258	152.0473
3-(2-Hydroxyphenyl)propionic acid	C9 H10 O3	5.322	166.063
3,4-Dihydroxybenzoic acid	C7 H6 O4	2.211	154.0266
3-Amino-4-hydroxybenzoic acid	C7 H7 N O3	0.975	153.0426
3-hydroxy-3-methyl-Glutaric acid	C6 H10 O5	1.218	162.0528
3-Indoleacetic Acid	C10 H9 N O2	5.531	175.0633
3-Methoxytyramine	C9 H13 N O2	1.549	167.0946
3-O-Methyl-L-DOPA	C10 H13 N O4	1.357	211.0845
4-acetamidobutanoate	C6 H11 N O3	1.339	145.0739
4-Hydroxybenzaldehyde	C7 H6 O2	3.746	122.0368
4-Hydroxyphenyllactic acid	C9 H10 O4	2.928	182.0579
4-Pyridoxic acid	C8 H9 N O4	1.305	183.0532
4-Quinolinecarboxylic acid	C10 H7 N O2	1.399	173.0477
5′-Deoxy-5′-(methylthio)adenosine	C11 H15 N5 O3 S	3.674	297.0896
Acetoin	C4 H8 O2	9.57	88.0524
Acetyl-L-Cysteine	C5 H9 N O3 S	1.367	163.0303
Acetyl-L-Leucine	C8 H15 N O3	5.058	173.1052
Caffeic Acid	C9 H8 O4	3.836	180.0423
Caffeine	C8 H10 N4 O2	4.155	194.0804
cAMP	C10 H12 N5 O6 P	1.318	329.0525
Citramalic acid	C5 H8 O5	1.052	148.0372
Citric acid	C6 H8 O7	0.8	192.027
Cortisol	C21 H30 O5	7.716	362.2093
d-Dethiobiotin	C10 H18 N2 O3	5.288	214.1317
Deoxyadenosine	C10 H13 N5 O3	1.489	251.1018
Deoxyguanosine	C10 H13 N5 O4	1.382	267.0968
Desmosterol	C27 H44 O	13.159	384.3392
DL-3,4-Dihydroxymandelic acid	C8 H8 O5	0.837	184.0372
Dopamine	C8 H11 N O2	0.804	153.079
ferulic acid	C10 H10 O4	5.153	194.0579
Guanosine	C10 H13 N5 O5	1.194	283.0917
Hippuric acid	C9 H9 N O3	3.674	179.0582
Homogentisic acid	C8 H8 O4	1.619	168.0423
Hypoxanthine	C5 H4 N4 O	0.847	136.0385
Indoleacetaldehyde	C10 H9 N O	2.43	159.0684

Kynurenine	C10 H12 N2 O3	1.672	208.0848
Linoleic acid	C18 H32 O2	12.207	280.2402
L-Nicotine	C10 H14 N2	0.999	162.1157
L-Phenylalanine	C9 H11 N O2	2.045	165.079
LTB4	C20 H32 O4	11.378	336.2301
L-Tryptophan	C11 H12 N2 O2	2.82	204.0899
L-Tyrosine	C9 H11 N O3	0.917	181.0739
Lumazine	C6 H4 N4 O2	1.141	164.0334
Lumichrome	C12 H10 N4 O2	6.514	242.0804
Methyl jasmonate	C13 H20 O3	8.787	224.1412
Methyl N-(a-methylbutyryl)glycine	C9 H16 O4	6.455	188.1049
MG(18:1(9Z)/0:0/0:0)[rac]	C21 H40 O4	12.212	356.2927
mono-Ethyl malonate	C5 H8 O4	2.854	132.0423
Monomethyl glutaric acid	C6 H10 O4	3.643	146.0579
N,N-Dimethyl-1,4-phenylenediamine	C8 H12 N2	1.038	136.1
N6-(DELTA2-ISOPENTENYL)-ADENINE	C10 H13 N5	5.995	203.1171
N-Acetyl-DL-methionine	C7 H13 N O3 S	3.263	191.0616
N-Acetyl-L-alanine	C5 H9 N O3	1.134	131.0582
N-Acetyl-L-phenylalanine	C11 H13 N O3	5.14	207.0895
Niacin (Nicotinic acid)	C6 H5 N O2	0.894	123.032
Niacinamide	C6 H6 N2 O	1.003	122.048
Nitrotyrosine	C9 H10 N2 O5	2.225	226.059
Palmitic acid	C16 H32 O2	12.36	256.2402
Pantothenic Acid	C9 H17 N O5	2.731	219.1107
Phenylethylamine	C8 H11 N	2.617	121.0891
p-Hydroxyphenylacetic acid	C8 H8 O3	3.585	152.0473
Pimelic acid	C7 H12 O4	4.179	160.0736
Quinaldic acid	C10 H7 N O2	3.815	173.0477
RAC-GLYCEROL 1-MYRISTATE	C17 H34 O4	11.572	302.2457
Resorcinol monoacetate	C8 H8 O3	5.251	152.0473
salsolinol	C10 H13 N O2	0.931	179.0946
Suberic acid	C8 H14 O4	5.39	174.0892
Theobromine	C7 H8 N4 O2	2.75	180.0647
trans-Aconitate	C6 H6 O6	0.982	174.0164
Tryptamine	C10 H12 N2	3.197	160.1
Tyramine	C8 H11 N O	1.069	137.0841
Uridine	C9 H12 N2 O6	0.932	244.0695
Xanthine	C5 H4 N4 O2	0.903	152.0334
Xanthosine	C10 H12 N4 O6	1.548	284.0757

Method 2: urine unique AMRT matches obtained (negative polarity)

Name	Formula	RT (Tgt)	Mass (Tgt)
2-METHYLMALEATE	C5 H6 O4	2.836	130.0266
4-Hydroxybenzaldehyde	C7 H6 O2	5.843	122.0368
4-Hydroxyphenyllactic acid	C9 H10 O4	4.751	182.0579
4-Pyridoxic acid	C8 H9 N O4	3.219	183.0532
Acetoacetic acid	C4 H6 O3	2.19	102.0317
Adipic acid	C6 H10 O4	4.615	146.0579
Citramalic acid	C5 H8 O5	2.412	148.0372
Citric acid	C6 H8 O7	1.924	192.027
Ethylmalonic acid	C5 H8 O4	3.825	132.0423
Galactaric acid	C6 H10 O8	1.346	210.0376
Gentisic acid	C7 H6 O4	4.668	154.0266
Gulonolactone	C6 H10 O6	1.38	178.0477
Hippuric acid	C9 H9 N O3	5.476	179.0582
Homogentisic acid	C8 H8 O4	3.542	168.0423
Hypoxanthine	C5 H4 N4 O	2.065	136.0385
Isocitrate	C6 H8 O7	1.625	192.027
Lithocholic acid	C24 H40 O3	13.732	376.2977
L-Tyrosine	C9 H11 N O3	2.169	181.0739
Maleic acid	C4 H4 O4	2.007	116.011
Malic acid	C4 H6 O5	1.577	134.0215
Malonic acid	C3 H4 O4	1.706	104.011
Methyl N-(a-methylbutyryl)glycine	C9 H16 O4	8.124	188.1049
Methylglyoxal	C3 H4 O2	2.195	72.0211
m-Salicylic acid	C7 H6 O3	5.702	138.0317
N-Acetyl-DL-serine	C5 H9 N O4	1.674	147.0532
N-Acetyl-L-phenylalanine	C11 H13 N O3	6.897	207.0895
N-Acetylneuraminic Acid	C11 H19 N O9	1.466	309.106
N-Acetylserotonin	C12 H14 N2 O2	5.38	218.1055
N-ALPHA-ACETYL-L-ASPARAGINE	C6 H10 N2 O4	1.606	174.0641
Oxoglutaric acid	C5 H6 O5	1.671	146.0215
Pantothenic Acid	C9 H17 N O5	4.337	219.1107
Phenyl acetate	C8 H8 O2	4.751	136.0524
p-Hydroxyphenylacetic acid	C8 H8 O3	5.314	152.0473
p-Salicylic acid	C7 H6 O3	5.312	138.0317
Pyroglutamic acid	C5 H7 N O3	2.119	129.0426
Quinic acid	C7 H12 O6	1.469	192.0634
Sarcosine	C3 H7 N O2	1.304	89.0477
Succinic acid	C4 H6 O4	2.391	118.0266
tartaric acid	C4 H6 O6	1.461	150.0164
Taurine	C2 H7 N O3 S	1.267	125.0147
Theophylline	C7 H8 N4 O2	5.226	180.0647

| Uric acid | C5 H4 N4 O3 | 1.902 | 168.0283 |
| Xanthine | C5 H4 N4 O2 | 2.266 | 152.0334 |

Method 2: urine unique AMRT matches obtained (positive polarity)

Name	Formula	RT (Tgt)	Mass (Tgt)
(±)-Mevalonolactone	C6 H10 O3	3.125	130.063
10-hydroxy capric acid	C10 H20 O3	9.269	188.1412
12-Hydroxydodecanoic acid	C12 H24 O3	10.523	216.1725
1-Methyladenosine	C11 H15 N5 O4	4.389	281.1124
2-Aminophenol	C6 H7 N O	2.162	109.0528
2-hydroxypyridine	C5 H5 N O	3.114	95.0371
3,4-Dihydroxybenzoic acid	C7 H6 O4	4.312	154.0266
3-Hydroxy-DL-kynurenine	C10 H12 N2 O4	2.193	224.0797
3-Methoxytyramine	C9 H13 N O2	3.197	167.0946
3-Methyladenine	C6 H7 N5	1.703	149.0701
3-O-Methyl-L-DOPA	C10 H13 N O4	3.129	211.0845
3-Ureidopropionic acid	C4 H8 N2 O3	1.717	132.0535
4-acetamidobutanoate	C6 H11 N O3	3.142	145.0739
4-Hydroxybenzaldehyde	C7 H6 O2	5.843	122.0368
4-Pyridoxic acid	C8 H9 N O4	3.219	183.0532
4-Quinolinecarboxylic acid	C10 H7 N O2	3.229	173.0477
5′-Deoxy-5′-(methylthio)adenosine	C11 H15 N5 O3 S	5.335	297.0896
5-Hydroxytryptophan	C11 H12 N2 O3	3.079	220.0848
5-Methylcytosine	C5 H7 N3 O	1.391	125.0589
Acetoin	C4 H8 O2	1.339	88.0524
Acetyl-L-Cysteine	C5 H9 N O3 S	3.112	163.0303
Arachidic Acid	C20 H40 O2	14.297	312.3028
Benzaldehyde	C7 H6 O	9.367	106.0419
Butanal	C4 H8 O	8.759	72.0575
Caffeic Acid	C9 H8 O4	6.168	180.0423
Caffeine	C8 H10 N4 O2	5.967	194.0804
cAMP	C10 H12 N5 O6 P	3.474	329.0525
Carnitine	C7 H15 N O3	1.228	161.1052
Cortisol	C21 H30 O5	9.27	362.2093
Creatine	C4 H9 N3 O2	1.321	131.0695
Creatinine	C4 H7 N3 O	1.211	113.0589
Cytidine	C9 H13 N3 O5	1.449	243.0855
Cytosine	C4 H5 N3 O	1.444	111.0433
Dehydroascorbic acid	C6 H6 O6	1.367	174.0164
Deoxyadenosine	C10 H13 N5 O3	3.404	251.1018
Deoxycytidine	C9 H13 N3 O4	1.583	227.0906
D-Glucosamine 6-phosphate	C6 H14 N O8 P	1.218	259.0457
D-Glucuronic acid	C6 H10 O7	1.328	194.0427
Diacetyl	C4 H6 O2	9.478	86.0368
Dihydrouracil	C4 H6 N2 O2	1.321	114.0429
Dihydroxyphenylacetic acid	C8 H8 O4	4.432	168.0423
Dioctyl phthalate	C24 H38 O4	13.507	390.277

Dopamine	C8 H11 N O2	1.663	153.079
D-Saccharic acid	C6 H10 O8	1.423	210.0376
Elaidic Acid	C18 H34 O2	13.64	282.2559
Gluconic acid	C6 H12 O7	1.365	196.0583
Glucosaminic acid	C6 H13 N O6	1.251	195.0743
Glutamine	C5 H10 N2 O3	1.262	146.0691
Glyceraldehyde	C3 H6 O3	1.399	90.0317
Glyceryl phosphate	C3 H9 O6 P	1.426	172.0137
Guanidineacetic acid	C3 H7 N3 O2	1.388	117.0538
Guanidinosuccinic Acid	C5 H9 N3 O4	1.325	175.0593
Guanine	C5 H5 N5 O	1.724	151.0494
Gulonolactone	C6 H10 O6	1.38	178.0477
Hippuric acid	C9 H9 N O3	5.476	179.0582
Homocysteine thiolactone	C4 H7 N O S	1.339	117.0248
Hypoxanthine	C5 H4 N4 O	2.065	136.0385
Inosine	C10 H12 N4 O5	2.956	268.0808
Isocitrate	C6 H8 O7	1.625	192.027
Itaconic acid	C5 H6 O4	3.444	130.0266
Kynurenic acid	C10 H7 N O3	5.313	189.0426
Kynurenine	C10 H12 N2 O3	3.505	208.0848
L-2-Aminoadipic acid	C6 H11 N O4	1.375	161.0688
L-Asparagine	C4 H8 N2 O3	1.253	132.0535
L-Aspartic Acid	C4 H7 N O4	1.276	133.0375
L-Cystine	C6 H12 N2 O4 S2	1.234	240.0238
L-Dopa	C9 H11 N O4	1.768	197.0688
lipoamide	C8 H15 N O S2	8.158	205.0595
L-Methionine	C5 H11 N O2 S	1.852	149.051
L-Phenylalanine	C9 H11 N O2	3.881	165.079
L-Serine	C3 H7 N O3	1.246	105.0426
L-Tryptophan	C11 H12 N2 O2	4.603	204.0899
L-Tyrosine	C9 H11 N O3	2.169	181.0739
Maleamic acid	C4 H5 N O3	1.87	115.0269
Maleimide	C4 H3 N O2	1.873	97.0164
Methyl jasmonate	C13 H20 O3	10.209	224.1412
Methyl N-(a-methylbutyryl)glycine	C9 H16 O4	8.124	188.1049
N-acetylaspartate	C6 H9 N O5	1.893	175.0481
N-Acetyl-DL-serine	C5 H9 N O4	1.674	147.0532
N-Acetyl-L-glutamic acid	C7 H11 N O5	2.262	189.0637
N-Acetylneuraminic Acid	C11 H19 N O9	1.466	309.106
N-Acetylputrescine	C6 H14 N2 O	1.374	130.1106
N-ALPHA-ACETYL-L-ASPARAGINE	C6 H10 N2 O4	1.606	174.0641
Niacinamide	C6 H6 N2 O	2.39	122.048
N-Methylhistamine	C6 H11 N3	1.257	125.0953
Norepinephrine (noradrenaline)	C8 H11 N O3	1.23	169.0739

Normetanephrine	C9 H13 N O3	1.606	183.0895
Octopamine (p-Hydroxyphenylethanolamine)	C8 H11 N O2	1.387	153.079
Oleic Acid	C18 H34 O2	13.101	282.2559
O-Phosphorylethanolamine	C2 H8 N O4 P	1.25	141.0191
Orotic acid	C5 H4 N2 O4	1.76	156.0171
O-Succinyl-L-homoserine	C8 H13 N O6	1.763	219.0743
p-Aminobenzoic acid	C7 H7 N O2	4.405	137.0477
Pantothenic Acid	C9 H17 N O5	4.337	219.1107
Phenylethylamine	C8 H11 N	4.163	121.0891
p-Hydroxyphenylacetic acid	C8 H8 O3	5.314	152.0473
Picolinic acid	C6 H5 N O2	1.786	123.032
Pimelic acid	C7 H12 O4	5.903	160.0736
Pterine	C6 H5 N5 O	2.826	163.0494
Purine	C5 H4 N4	2.754	120.0436
Pyridoxine (Vitamin B6)	C8 H11 N O3	1.794	169.0739
Pyrocatechol	C6 H6 O2	4.698	110.0368
Pyroglutamic acid	C5 H7 N O3	2.119	129.0426
Quinaldic acid	C10 H7 N O2	5.751	173.0477
Quinic acid	C7 H12 O6	1.469	192.0634
R-(-)-Mandelic acid	C8 H8 O3	4.852	152.0473
Riboflavin (Vitamin B2)	C17 H20 N4 O6	6.495	376.1383
salsolinol	C10 H13 N O2	2.153	179.0946
Suberic acid	C8 H14 O4	7.096	174.0892
Taurine	C2 H7 N O3 S	1.267	125.0147
Theobromine	C7 H8 N4 O2	4.547	180.0647
Thymine	C5 H6 N2 O2	3.052	126.0429
Thyrotropin releasing hormone	C16 H22 N6 O4	2.121	362.1703
Trigonelline	C7 H7 N O2	1.397	137.0477
Tyramine	C8 H11 N O	2.291	137.0841
Uracil	C4 H4 N2 O2	1.811	112.0273
Urocanic acid	C6 H6 N2 O2	1.613	138.0429
Valeric acid	C5 H10 O2	9.487	102.0681
Xanthine	C5 H4 N4 O2	2.266	152.0334
Xanthosine	C10 H12 N4 O6	3.493	284.0757
Xanthurenic acid	C10 H7 N O4	5.029	205.0375

Method 3: urine unique AMRT matches obtained (negative polarity)

Name	Formula	RT (Tgt)	Mass (Tgt)
2-METHYLMALEATE	C5 H6 O4	3.995	130.0266
3-hydroxy-3-methyl-Glutaric acid	C6 H10 O5	3.881	162.0528
3-Hydroxybenzaldehyde	C7 H6 O2	1.549	122.0368
4-Methylcatechol	C7 H8 O2	1.779	124.0524
Allantoin	C4 H6 N4 O3	5.489	158.044
cAMP	C10 H12 N5 O6 P	7.801	329.0525
Citramalic acid	C5 H8 O5	4.344	148.0372
Creatinine	C4 H7 N3 O	4.994	113.0589
D-(+)-GALACTURONIC ACID	C6 H10 O7	7.122	194.0427
D-Glucuronic acid	C6 H10 O7	6.803	194.0427
DL-3,4-Dihydroxymandelic acid	C8 H8 O5	4.341	184.0372
Gluconic acid	C6 H12 O7	6.946	196.0583
Glutamine	C5 H10 N2 O3	7.307	146.0691
Glyceric acid	C3 H6 O4	4.309	106.0266
Gulonolactone	C6 H10 O6	4.77	178.0477
Hypoxanthine	C5 H4 N4 O	5.117	136.0385
Indoxylsulfuric acid	C8 H7 N O4 S	4.217	213.0096
L-Phenylalanine	C9 H11 N O2	5.681	165.079
L-Tryptophan	C11 H12 N2 O2	5.681	204.0899
L-Tyrosine	C9 H11 N O3	6.262	181.0739
N-acetylaspartate	C6 H9 N O5	4.539	175.0481
N-Acetyl-DL-serine	C5 H9 N O4	4.389	147.0532
N-Acetyl-L-glutamic acid	C7 H11 N O5	4.403	189.0637
N-Acetylneuraminic Acid	C11 H19 N O9	7.019	309.106
N-ALPHA-ACETYL-L-ASPARAGINE	C6 H10 N2 O4	4.876	174.0641
N-Formylmethionine	C6 H11 N O3 S	3.439	177.046
Pantothenic Acid	C9 H17 N O5	3.983	219.1107
p-Coumaric acid	C9 H8 O3	1.723	164.0473
Pyridoxal (Vitamin B6)	C8 H9 N O3	4.317	167.0582
Pyroglutamic acid	C5 H7 N O3	4.136	129.0426
Salicylic acid	C7 H6 O3	1.7	138.0317
Taurine	C2 H7 N O3 S	6.791	125.0147
trans-Aconitate	C6 H6 O6	4.441	174.0164
Uracil	C4 H4 N2 O2	4.05	112.0273
Uric acid	C5 H4 N4 O3	6.615	168.0283
Vanillylmandelic acid	C9 H10 O5	3.535	198.0528
Xanthine	C5 H4 N4 O2	5.448	152.0334

Method 3: urine unique AMRT matches obtained (positive polarity)

Name	Formula	RT (Tgt)	Mass (Tgt)
1-Methyladenosine	C11 H15 N5 O4	6.182	281.1124
1-methylhistidine	C7 H11 N3 O2	7.636	169.0851
2-Aminophenol	C6 H7 N O	1.744	109.0528
3-O-Methyl-L-DOPA	C10 H13 N O4	6.137	211.0845
4-Guanidinobutanoate	C5 H11 N3 O2	4.957	145.0851
4-Hydroxy-L-proline	C5 H9 N O3	6.981	131.0582
5′-Deoxy-5′-(methylthio)adenosine	C11 H15 N5 O3 S	4.466	297.0896
5-Hydroxytryptophan	C11 H12 N2 O3	6.27	220.0848
5-Methylcytosine	C5 H7 N3 O	5.524	125.0589
6-Deoxy-L-galactose	C6 H12 O5	5.77	164.0685
Acetyl-L-Leucine	C8 H15 N O3	1.715	173.1052
Acetyllysine	C8 H16 N2 O3	6.06	188.1161
Adenosine	C10 H13 N5 O4	5.374	267.0968
Caffeine	C8 H10 N4 O2	1.659	194.0804
cAMP	C10 H12 N5 O6 P	7.801	329.0525
Carnitine	C7 H15 N O3	4.492	161.1052
Citrulline	C6 H13 N3 O3	7.46	175.0957
Creatine	C4 H9 N3 O2	5.838	131.0695
Creatinine	C4 H7 N3 O	4.994	113.0589
Cytidine	C9 H13 N3 O5	6.459	243.0855
Cytosine	C4 H5 N3 O	5.842	111.0433
DL-pipecolic acid	C6 H11 N O2	6.246	129.079
D-MANNOSAMINE	C6 H13 N O5	7.18	179.0794
D-Tagatose	C6 H12 O6	6.317	180.0634
Glucosamine	C6 H13 N O5	7.553	179.0794
Glucosaminic acid	C6 H13 N O6	8.308	195.0743
Glutamine	C5 H10 N2 O3	7.307	146.0691
Guanidineacetic acid	C3 H7 N3 O2	6.214	117.0538
Guanidinosuccinic Acid	C5 H9 N3 O4	6.401	175.0593
Guanine	C5 H5 N5 O	6.058	151.0494
Guanosine	C10 H13 N5 O5	6.394	283.0917
Histamine	C5 H9 N3	7.093	111.0796
Hypoxanthine	C5 H4 N4 O	5.117	136.0385
Indole-3-methyl acetate	C11 H11 N O2	1.706	189.079
Inosine	C10 H12 N4 O5	5.735	268.0808
Kynurenic acid	C10 H7 N O3	5.023	189.0426
Kynurenine	C10 H12 N2 O3	5.793	208.0848
L-Arginine	C6 H14 N4 O2	7.983	174.1117
L-Asparagine	C4 H8 N2 O3	7.501	132.0535
L-Cystathionine	C7 H14 N2 O4 S	8.802	222.0674
L-Cystine	C6 H12 N2 O4 S2	8.987	240.0238
L-Histidine	C6 H9 N3 O2	8.075	155.0695

L-Leucine	C6 H13 N O2	5.417	131.0946
L-Serine	C3 H7 N O3	7.372	105.0426
L-Tryptophan	C11 H12 N2 O2	5.681	204.0899
L-Tyrosine	C9 H11 N O3	6.262	181.0739
Melibiose	C12 H22 O11	8.325	342.1162
N-Acetylneuraminic Acid	C11 H19 N O9	7.019	309.106
N-Acetylputrescine	C6 H14 N2 O	5.281	130.1106
Niacinamide	C6 H6 N2 O	3.514	122.048
N-Methylhistamine	C6 H11 N3	6.779	125.0953
Nε,Nε,Nε-Trimethyllysine	C9 H20 N2 O2	7.377	188.1525
Pantothenic Acid	C9 H17 N O5	3.983	219.1107
Paraxanthine	C7 H8 N4 O2	3.528	180.0647
Pyroglutamic acid	C5 H7 N O3	4.138	129.0426
Raffinose	C18 H32 O16	8.653	504.169
Retinoic acid	C20 H28 O2	1.499	300.2089
Riboflavin (Vitamin B2)	C17 H20 N4 O6	5.598	376.1383
S-(Carboxymethyl)-L-cysteine	C5 H9 N O4 S	6.989	179.0252
S-Adenosylhomocysteine	C14 H20 N6 O5 S	7.664	384.1216
salsolinol	C10 H13 N O2	4.423	179.0946
Serotonin	C10 H12 N2 O	4.512	176.095
Stachyose	C24 H42 O21	9.382	666.2219
Taurine	C2 H7 N O3 S	6.791	125.0147
Tryptamine	C10 H12 N2	1.736	160.1
Tyramine	C8 H11 N O	3.417	137.0841
Uric acid	C5 H4 N4 O3	6.615	168.0283
Xanthine	C5 H4 N4 O2	5.448	152.0334
Xanthosine	C10 H12 N4 O6	5.958	284.0757
β-Alanine	C3 H7 N O2	6.098	89.0477

Non-targeted feature extraction parameters (LC-QTOF-MS data)

Molecular feature extraction (MFE) parameters for 'recursive non-targeted feature extraction' were as follows: peak height >5000 counts, charge state range 1-2, minimum ion count = 1, binning and alignment tolerances - RT window = 0.3 min and mass window = 50 ppm, MFE score >70 in 60% of samples in at least one sample group. 'Find by ion' parameters (as part of the recursive feature extraction workflow) were as follows: mass score = 100, isotope abundance score = 60, isotope spacing score = 50, RT score = 100, expected data variation = 50ppm (15%) and RT 0.3 min, integration - 'Agile 2' algorithm, EIC smoothed before integration, smoothing function - Gaussian, peaks filtered by height (absolute height >1000 counts), spectra to include - average scans >40% of peak height, exclude if >20% of saturation, centroiding - maximum spike width = 2, required valley = 0.7, 'find by ion' filters - score >60 in at least one sample group. Allowed ion species and sample groupings were the same as described for targeted data mining.

Appendix 3

Table S4.01. Summary of unidentified chemical entities with differential abundance in urine taken from patients with AKU baseline vs on nitisinone treatment, from non-targeted feature extraction. Peak areas for these entities showed statistically significant differences (Benjamini-Hochberg false-discovery rate p <0.05, fold change >2) between baseline and each sampling time point (3, 12 and 24 months) on nitisinone. Putative compound identifications were not obtained for these compounds using the MassHunter METLIN metabolite PCD/PCDL accurate mass database (build 07.00, Agilent), with a match criteria of accurate mass ±5 ppm. Fold changes are for baseline vs 24 months on nitisinone time points.

Neutral mass	RT (min)	Polarity	Direction of alteration on nitisinone	Fold change	p-value
264.085	3.2	(-)	↑	>100	<0.0001
407.12	4.7	(-)	↑	>100	<0.0001
326.088	5.1	(-)	↑	>100	0.025
378.116	4.8	(-)	↑	>100	<0.0001
386.099	4.6	(+)	↑	29.6	<0.0001
343.2	4.4	(+)	↑	21.2	<0.0001
204.055	3.9	(+)	↑	20.8	<0.0001
326.115	5	(+)	↑	9.6	<0.0001
120.058	2.3	(+)	↑	9.5	<0.0001
293.102	3.4	(+)	↑	6.7	<0.0001
115.1	2.3	(+)	↑	5.7	<0.0001
295.143	4.3	(+)	↑	5.6	<0.0001
433.135	4.8	(+)	↑	3.7	<0.0001
260.091	2.6	(+)	↑	2.6	<0.0001
312.134	2.1	(+)	↑	2.6	<0.01
189.055	3.4	(+)	↓	24.6	<0.0001
185.068	3.4	(+)	↓	16.6	<0.0001

150.032	3.4	(+)	↓		7.9	<0.0001
208.091	3.5	(+)	↓		6.9	<0.0001
205.999	3.4	(+)	↓		5.1	<0.01
303.122	5.7	(+)	↓		3.9	<0.0001
212.081	7.1	(+)	↓		2.9	<0.001
144.043	4.6	(+)	↓		2.7	<0.0001
304.079	4.2	(+)	↓		2.4	<0.01

Table S4.02. Summary of unidentified urinary chemical entities with differential abundance in urine taken from AKU mice at baseline vs 1 week on nitisinone, from non-targeted feature extraction. Putative compound identifications were not obtained for these compounds using the MassHunter METLIN metabolite PCD/PCDL accurate mass database (build 07.00, Agilent), with match criterion of accurate mass ±5 ppm. P-values are adjusted by Benjamini-Hochberg false-discovery rate.

Neutral mass	RT (min)	Polarity	Direction of alteration post-nitisinone	Fold change	p-value
232.005	4.1	(-)	↑	76.7	<0.0001
182.058	4.7	(-)	↑	53.7	<0.0001
261.067	4.7	(-)	↑	31.5	<0.0001
508.141	3.9	(-)	↑	17.6	<0.0001
514.125	4.8	(-)	↑	16	<0.0001
174.029	5.3	(-)	↑	13.7	<0.0001
367.001	4.7	(-)	↑	11.3	<0.0001
235.979	2.5	(-)	↑	10.1	<0.0001
417.049	4.7	(-)	↑	10	<0.01
127.051	1.6	(+)	↑	7.8	0.036
461.133	5.6	(-)	↑	5.4	<0.01
472.009	4.2	(-)	↑	5	<0.0001
242.983	3.8	(-)	↑	4.6	<0.0001
444.165	5.6	(-)	↑	4.1	<0.0001
508.18	6.6	(-)	↑	3.7	<0.001
253.956	4.3	(-)	↑	3.6	0.035
229.071	6.9	(+)	↑	3.1	<0.0001
274.116	3.6	(-)	↑	3	<0.01
140.001	4.6	(+)	↑	2.9	<0.001
298.11	4	(-)	↑	2.8	<0.0001
240.009	5.3	(-)	↑	2.7	<0.0001
225.073	4.7	(+)	↑	2.7	<0.0001
155.014	1.4	(+)	↑	2.6	<0.0001
170.092	1.6	(+)	↑	2.5	<0.0001
199.122	8.9	(+)	↑	2.3	<0.001
229.168	10.4	(-)	↑	2.2	<0.0001
386.194	6.9	(+)	↑	2.2	0.036
300.06	9.7	(-)	↑	2.1	<0.0001
191.092	4.8	(+)	↑	2.1	<0.01
246.138	7.5	(+)	↑	2.1	<0.01
191.098	4.8	(+)	↑	2.1	<0.01
199.125	8.5	(-)	↑	2	<0.0001

184.146	7.6	(-)		↑	2	0.021
56.0624	4.7	(-)		↑	2	<0.001
195.071	4.1	(+)		↑	2	<0.001
243.037	2.5	(-)		↓	16	<0.0001
509.209	4.2	(-)		↓	16	<0.0001
182.005	4.9	(-)		↓	14.5	<0.01
403.158	3.7	(-)		↓	11.1	<0.0001
432.221	3.6	(+)		↓	3.9	<0.0001
110.003	6	(-)		↓	3.1	<0.01
96.0213	2.3	(-)		↓	2.8	<0.0001
309.081	4.7	(+)		↓	2.4	<0.0001
287.993	4.6	(-)		↓	2.2	<0.0001
324.084	3.4	(+)		↓	2.2	<0.01
324.093	3.4	(+)		↓	2.2	<0.01
397.998	5.7	(-)		↓	2.1	<0.01
217.005	4.4	(-)		↓	2.1	<0.001
121.032	3.4	(+)		↓	2.1	0.024

Appendix 4

Table S4.03. Summary of unidentified urinary chemical entities with differential abundance between patients with high (score >10) vs low (score <10) ^{18}F-NaF uptake to intervertebral disks. Putative compound identifications were not obtained for these compounds using the MassHunter METLIN metabolite PCD/PCDL accurate mass database (build 07.00, Agilent), with match criterion of accurate mass ±5 ppm. P-values are adjusted by Benjamini-Hochberg false-discovery rate.

[†] Entity with statistically significant difference between patients with high vs low uptake score <u>at baseline</u>

[‡] Entity with statistically significant difference between patients with high vs low uptake score <u>at 24 months on nitisinone</u>

Neutral mass	RT (min)	Polarity	Direction of alteration in 'high' ^{18}F-NaF uptake group	Fold change	p-value
618.1348 [†]	4.1	(-)	↑	>100	0.026
370.0571 [†]	4.8	(-)	↑	>100	<0.0001
277.0071 [‡]	5.0	(-)	↑	>100	0.013
151.0877 [†‡]	4.2	(+)	↑	54.4	0.039
231.0212 [‡]	3.6	(+)	↑	29.3	0.012
231.0212 [†]	3.6	(+)	↑	18.7	0.026
351.0947 [†]	6.0	(-)	↓	16.0	<0.0001
145.0753 [†]	2.2	(+)	↓	16.0	<0.0001
289.0952 [†]	8.8	(+)	↓	3.2	<0.0001
174.0897 [†]	3.5	(+)	↓	2.5	<0.0001
250.0497 [†]	1.4	(-)	↓	2.1	<0.0001

Table S4.04. Summary of unidentified urinary chemical entities with differential abundance between patients with high (score >10) vs low (score <10) ^{18}F-NaF uptake to lumbar/thoracic vertebrae. Putative compound identifications were not obtained for these compounds using the MassHunter METLIN metabolite PCD/PCDL accurate mass database (build 07.00, Agilent), with match criterion of accurate mass ±5 ppm. P-values are adjusted by Benjamini-Hochberg false-discovery rate.

[†] Entity with statistically significant difference between patients with high vs low uptake score <u>at baseline</u>

[‡] Entity with statistically significant difference between patients with high vs low uptake score <u>at 24 months on nitisinone</u>

Neutral mass	RT (min)	Polarity	Direction of alteration in 'high' ^{18}F-NaF uptake group	Fold change	p-value
188.0151 [‡]	5.9	(-)	↓	41.1	<0.0001
285.1268 [†]	4.6	(+)	↑	2.8	<0.0001

Appendix 5

Table S5.01. Compounds appended to databases for targeted feature extraction (Chapter 5). Includes compounds of potential interest in AKU and potential biotransformation products of HGA.

Name	Formula	Mass	CAS	METLIN	KEGG	HMP
1-(2-Carboxyphenylamino)-1'-deoxy-D-ribulose 5'-phosphate	C12H16NO9P	349.056268	5962-18-5	3492	C01302	
2,5-Dihydroxybenzaldehyde	C7H6O3	138.031694	1194-98-5	44699	C05585	
2-aminomuconic acid semialdehyde	C6H7NO3	141.042593		45872	C03824	HMDB01280
2-Dehydro-3-deoxy-D-arabino-heptonate 7-phosphate (DAHP)	C7H13O10P	288.024633	2627-73-8	3414	C04691	
3,4-Dihydroxymandelaldehyde	C8H8O4	168.042259	13023-73-9	58374	C05577	HMDB06242
3-Dehydroquinic acid	C7H10O6	190.047738	10534-44-8	3327	C00944	HMDB12710
3-Hydroxytyrosol	C8H10O3	154.06299				
3-Hydroxytyrosol sulfate	C8H10O6S	234.01981				
3-Methoxy-4-hydroxymandelate	C9H9O5	197.0449984			C05584	
3-Methoxy-4-hydroxyphenylglycolaldehyde	C9H10O4	182.057909	17592-23-3	58200	C05583	HMDB04061
4-Hydroxyphenylacetaldehyde	C8H8O2	136.052429	7339-87-9	63506	C03765	
4-Hydroxyphenylpyruvic acid	C9H8O4	180.042259	156-39-8	3315	C01179	
5-O-(1-Carboxyvinyl)-3-phosphoshikimate	C10H13O10P	324.024633	74708-67-1	3454	C01269	
5-phosphoribosyl-1-diphosphate	C5H13O14P3	389.951815	7540-64-9	158	C00119	HMDB00280
Acetic acid	C2H4O2	60.021129	64-19-7	3206	C00033	HMDB00042
Acetyl-CoA	C23H38N7O17P3S	809.125773			C00024	HMDB01206
Acetyl-L-tyrosine	C11H13NO4	223.084458	537-55-3	5827		HMDB00866
Adenosine triphosphate (ATP)	C10H16N5O13P3	506.995745	987-65-5	95	C00002	HMDB00538
Aminomuconic acid	C6H7NO4	157.037508		6440	C02220	

Chorismic acid	C10H10O6	226.047738	617-12-9	368	C00251	HMDB12199
Coenzyme A (CoA)	C21H36N7O16P3S	767.115208	85-61-0	242	C00010	HMDB01423
Dehydro-HGA	C8H6O4	166.02661				
D-Erythrose 4-phosphate	C4H9O7P	200.008589	585-18-2	355	C00279	HMDB01321
D-Glyceraldehyde 3-phosphate	C3H7O6P	169.998024	591-57-1	3294	C00118	
D-Pantetheine 4'-phosphate	C11H23N2O7PS	358.096358	2226-71-3	398	C01134	HMDB01416
Glutamyl-Tyrosine	C14H18N2O6	310.116486		85745		HMDB28831
Hipposudoric acid	C16H8O8	328.02192				
Indole	C8H7N	117.057849	120-72-9	286	C00463	HMDB00738
Indoleglycerol phosphate	C11H14NO6P	287.055874	4220-97-7	3412	C03506	
Metanephrine	C10H15NO3	197.105193	5001-33-2	65	C05588	HMDB04063
N-(5-Phospho-D-ribosyl)anthranilate	C12H16NO9P	349.056268	27695-85-8	3493	C04302	
NADH	C21H29N7O14P2	665.1247717			C00004	HMDB01487
NADPH	C21H30N7O17P3	745.091102	2646-71-1	3691	C00005	
N-Methyltyramine	C9H13NO	151.099714	370-98-9	58167	C02442	HMDB03633
Norhipposudoric acid	C15H8O6	284.03209				
Oxaloacetate	C4H4O5	132.005873	328-42-7	123	C00036	HMDB00223
Phenylpyruvic acid	C9H8O3	164.047344	156-06-9	328	C00166	HMDB31629
Phloretic acid	C9H10O3	166.062994	501-97-3	4148	C01744	HMDB02199
Prephenic acid	C10H10O6	226.047738	126-49-8	369	C00254	HMDB12283
Pyrophosphate	H4O7P2	177.943225	03/09/2466	3306	C00013	HMDB00250
Pyruvate	C3H4O3	88.016044	127-17-3	117	C00022	HMDB00243
S-Acetylphosphopantetheine	C13H25N2O8PS	400.106923		66032	C03725	
S-Adenosylmethionine	C15H23N6O5S	399.145064	29908-03-0	3289	C00019	
Salicyluric acid	C9H9NO4	195.05316				
Shikimate-3-phosphate	C7H11O8P	254.019154	63959-45-5	3384	C03175	
Tyramine glucuronide	C14H19NO7	313.116152	27972-85-6	61650	C03033	HMDB10328
Tyramine-O-sulfate	C8H11NO4S	217.040879	30223-92-8	58408		HMDB06409

Tyrosol	C8H10O2	138.06808		
Tyrosol 4-sulfate	C8H10O5S	218.024894	96137	HMDB41785

			Phase
HGA_Hydroxylation + Glucuronide	C14H16O11	360.06926	II
HGA_Acetylation	C10H10O5	210.05282	II
HGA_Alkenes to Dihydrodiol	C8H10O6	202.04774	I
HGA_Demethylation and Methylene to Ketone	C7H4O5	168.00587	I
HGA_Demethylation	C7H6O4	154.02661	I
HGA_1,4-Dihydropyridines to Pyridines	C8H6O4	166.02661	I
HGA_Ethyl Ether to Acid	C6H2O5	153.99022	I
HGA_Hydroxylation and Methylation	C9H10O5	198.05282	II
HGA_Glycine Conjugation	C10H11NO5	225.06372	II
HGA_Cysteine Conjugation	C11H15NO6S	289.06201	II
HGA_Cysteine Glycine Conjugation	C13H18N2O7S	346.08347	II
HGA_(O, N, S) Methylation	C9H10O4	182.05791	II
HGA_N-Acetylcysteine Conjugation and Desaturation	C13H15NO7S	329.05692	II
HGA_Alcohols Dehydration	C8H6O3	150.03169	I
HGA_Glutamine Conjugation	C13H16NO6	282.09776	II
HGA_Demethylation and Hydroxylation	C7H6O5	170.02152	I
HGA_Demethylation and two Hydroxylations	C7H6O6	186.01644	I
HGA_Ethyl to Carboxylic Acid	C7H4O6	184.00079	I
HGA_Hydration, Hydrolysis (Internal)	C8H10O5	186.05282	I
HGA_Hydroxylation and Desaturation	C8H6O5	182.02152	I
HGA_Hydroxymethylene Loss	C7H6O3	138.03169	I
HGA_Decarboxylation	C7H8O2	124.05243	I

HGA_Cysteine Conjugation and Desaturation	C11H13NO6S	287.04636	II
HGA_Ketone to Alcohol	C8H10O4	170.05791	I
HGA_Taurine Conjugation	C10H13NO6S	275.04636	II
HGA_Hydroxylation and Sulfation	C8H8O8S	263.99399	II
HGA_Sulfate Conjugation	C8H8O7S	247.99907	II
HGA_2x Sulfate Conjugation	C8H8O10S2	327.95589	II
HGA_Glucuronide Conjugation	C14H16O10	344.07435	II
HGA_Sulfate Conjugation	C8H8O7S	247.99907	II
HGA_N-Acetylcysteine Conjugation	C13H17NO7S	331.07257	II
HGA_Deethylation	C6H4O4	140.01096	I
HGA_Alkene to Epoxide	C8H8O5	184.03717	I
HGA_Tert-Butyl to Acid	C5O6	155.96949	I
HGA_1,4-Dihydropyridines to Pyridines	C8H6O4	166.02661	I
HGA_Isopropyl to Acid	C6H2O6	169.98514	I

[Author accepted manuscript]

Title: Ochronotic pigmentation is caused by homogentisic acid and is the key event in Alkaptonuria leading to the destructive consequences of the disease – a review

Authors: Ranganath LR[1,2*], Norman BP[2], Gallagher JA[2].

Royal Liverpool University Hospital, Prescot Street, Liverpool, L7 8XP[1]; Musculoskeletal Biology I, Institute of Ageing & Chronic Disease, William Henry Duncan Building, University of Liverpool, Liverpool, UK[2]

Corresponding author: *LR Ranganath, Department of Clinical Biochemistry and Metabolic Medicine, Royal Liverpool University Hospital, Prescot Street, Liverpool, L7 8XP

e-mail: lrang@liv.ac.uk
Tel no: 0441517064256

Word count: Manuscript: 6549   **References:** 95   **Abstract:** 143

Figures: 5   **Tables** 4

Abstract

Ochronosis is the process in Alkaptonuria (AKU) that causes all the debilitating morbidity. The process involves selective deposition of homogentisic acid-derived pigment in tissues altering the properties of these tissues, leading to their failure. Some tissues like cartilage are more easily affected by ochronosis while others such as the liver and brain are unaffected for reasons that are still not understood. *In vitro* and mouse models of ochronosis have confirmed the dose relationships between homogentisic acid and ochronosis and also their modulation by HPPD inhibition. Ochronosis cannot be fully reversed and is a key factor in influencing treatment decisions. Earlier detection of ochronosis preferably by non-invasive means is desirable. A cause-effect relationship between HGA and ochronosis is discussed. The similarity in AKU and familial hypercholesterolaemia is explored, and lessons learnt. More research is needed to more fully understand the crucial nature of ochronosis.

Introduction: Archibald Garrod is the father of inborn errors of metabolism and studied four disorders, namely alkaptonuria, pentosuria, cystinuria and albinism. Pentosuria (OMIM#260800) is a defect in l-xylulose reductase, necessary for xylitol metabolism, leading to overproduction of pentose sugars and pentosuria, but is otherwise harmless as it does not accumulate in the body or produce a disease process of its own (Knox 1958). This is unlike alkaptonuria (AKU) (Figure 1) (OMIM#203500) where ochronosis is the key pathophysiological event (Garrod 1902; Galdston et al. 1952; O'Brien et al. 1963). Ochronosis-like processes have also been described in non-AKU states and termed pseudo-ochronosis or exogenous ochronosis (as opposed to endogenous ochronosis in AKU). Exogenous ochronosis has been the subject of previous reviews and will not be discussed further here (Levin and Maibach 2001; Bhattar et al. 2015).

Ochronosis is the term used to describe a process by which a yellowish (ochre) discoloration develops (Figure 2A), due to the deposition of pigment. First described by Rudolf Virchow in 1866, ochronosis was observed in miscroscopic examination of connective tissues (Virchow 1866). The tissues affected by ochronosis macroscopically, however, appear to be blue grey or black when large amounts of pigment are present. Virchow noticed that the pigment also specially accumulated in damaged or inflamed sites, such as irritated joint synovia and arteriosclerotic plaques. He hypothesised that non-crystalline colouring matter, whose nature was then unknown, slowly saturates the cartilage, a special target in AKU.

It is now known that AKU is a condition characterised by the absence of homogentisate dioxygenase enzyme (HGD) (EC:1.13.11.5) (Figure 1) leading to an inability to metabolise homogentisic acid (HGA) (La Du et al. 1958). HGA, an intermediary in the phenylalanine/tyrosine pathway, is normally completely and rapidly metabolised to yield fumarate and acetoacetate (Phornphutkul et al. 2002). The degradation of phenylalanine (an essential amino acid) and tyrosine (a non-essential amino acid) proceeds to fumarate and acetoacetate through homogentisic acid so rapidly that in normal subjects there is no increase in circulating HGA and very little in urine (Davison et al. 2015). The HGD deficiency in AKU causes excessive HGA production, causing homogentisicaciduria, while also increasing HGA concentrations within body tissues. Alkaptonuria is an autosomal recessive disorder with a delayed slowly-progressive

multisystemic damage for which there is still no approved disease modifying therapy, even though a drug called nitisinone decreases HGA (Introne et al. 2011). Although HGA by itself causes morbidity such as renal, prostate, gall bladder and salivary stones, the main disease process leading to tissue destruction and debilitating clinical sequelae is ochronosis, the focus of this review.

Formation of ochronotic pigment (OP): *Overproduction of HGA:* With an inability to metabolise HGA, there is an increase in whole body HGA despite massively increased renal excretion of HGA (Figure 1). The most important pathophysiological consequence in tyrosine pathway in AKU is conversion of accumulating HGA to OP. HGA is a reducing agent and especially under alkaline conditions it is itself rapidly oxidised via benzoquinone acetic acid turning solutions and tissues black (Zannoni et al. 1969). Slow spontaneous blackening of urine is usual in AKU, but adding alkali instantly turns urine black; acidifying the black urine does not return the black colour to normal, suggesting a potential irreversible change.

Apparently slow formation of pigment: Despite the metabolic defect existing from birth, pigmentation of eyes and ears is slow to develop externally, taking up to two or three decades to appear, presumably a reflection of the gradual accretion of the pigment. Similarly, back and knee pain, due to development of critical ochronosis in these sites, are also apparent in the second and third decades (Cox and Ranganath 2011). Formation of OP is markedly increased in renal failure, consistent with increased retention of HGA, and accelerates severe morbidity and debility (Introne et al. 2002).

Factors supporting molecular interaction of HGA with cartilage matrix (Figure 2B, C): The molecular mechanism through which HGA interacts with collagen matrix is unknown. Ultrastructural examination of pigmented cartilage showed that initial pigmentation is closely associated with the periodicity of collagen fibres. A periodic banding pattern of OP was observed on individual collagen fibres, with very early pigmentation appearing as small granules on the surface of fibres (Taylor et al. 2010b). These findings suggest that a nucleation-like event underlies ochronosis, where initial granular deposits on individual collagen fibres is followed by further rapid pigment deposition (Gallagher et al. 2016). The data also suggest that collagen fibrils provide specific binding sites for pigment.

Close observation of tissues obtained from AKU patients (Taylor et al. 2011b, 2012) and mice (Taylor et al. 2012; Preston et al. 2014) and *in vitro* (Tinti et al. 2011b) models

of AKU show that cartilage is initially resistant to pigmentation. It is proposed that biomechanical and biochemical changes, such as those that occur in cartilage as part of the natural ageing process, render tissues susceptible to ochronosis. The 'exposed collagen hypothesis' theorises that binding sites become available for HGA following the loss of protective molecules such as proteoglycans and glycosaminoglycans (Figure 2C) (GAGs) (Gallagher et al. 2016). In support of this hypothesis, AKU cartilage is shown to have lower levels of extractable GAGs and oligomeric matrix protein than osteoarthritic and non-osteoarthritic cartilage (Taylor et al. 2017). There is also evidence that the structure and maturity of the collagen matrix can influence pigment deposition. Newly-synthesised aberrant matrix proteins in scar tissue also appear to pigment rapidly, as reported in the case of a mediastinal mass from an AKU patient (Taylor et al. 2011a). Solid-state NMR analysis of AKU articular joint cartilage also observed spectra indicative of marked collagen disorder at the atomic level (Chow et al. 2011), further supporting the idea that disruption to the collagen matrix supports OP deposition.

Further work is required to elucidate a specific binding site for OP within the collagen matrix. It is also not known whether the initial binding occurs as HGA, the oxidised intermediate benzoquinoneacetic acid (BQA) or OP.

Reversibility of pigment: It is currently believed that the OP process is not fully reversible, although there are indications of some reversal of pigment; the mechanism of such reversal is currently unknown (Ranganath et al. 2018). It is not known if OP process is dynamic i.e. formation and removal co-existing, although macrophages with pigment has been described in case reports (Gaines et al. 1987; Fisher and Davis 2004; Damian et al. 2013).

Diet and OP: Tyrosine and phenylalanine are not limiting amino acids in the diet and are consumed in vast excess of requirement. The attempts to retard progression of AKU by reducing the amounts phenylalanine and tyrosine flux is in keeping with the belief that amounts of protein intake could influence AKU morbidity.

Generation of OP in vitro and *ex vivo (Figure 2D):* Model systems have been developed for investigating various effects of HGA exposure under controlled conditions *in vitro* and *ex vivo. In vitro* models include human serum (Braconi et al. 2010a, 2011) and osteoblastic (cell line) and chondrocytic (cell line and isolated from human cartilage) cell cultures (Braconi et al. 2010b; Tinti et al. 2010, 2011b; Braconi

et al. 2012; Millucci et al. 2012; Spreafico et al. 2013; Millucci et al. 2015b; Mistry et al. 2016). Incubation of cell cultures with HGA leads to OP formation by four weeks. The amount of OP formed is proportionate to the concentrations of HGA in the medium (Tinti et al. 2011b). *Ex vivo* 'organotypic' approaches have modelled AKU by studying the effect of HGA on human cartilage explants (Tinti et al. 2011a; Millucci et al. 2012; Bernardini et al. 2019). In these studies, pericellular pigmentation of cartilage explants is visible after two months of incubation with HGA (Tinti et al. 2011a).

Development of these model systems has not only been useful in studying the development of OP under controlled conditions, but also a number of other pathophysiological events closely associated with ochronosis including secondary amyloidosis (Millucci et al. 2012; Spreafico et al. 2013; Millucci et al. 2015b), perturbed redox homeostasis (Braconi et al. 2010a, b, 2011, 2012, 2015; Tinti et al. 2010; Spreafico et al. 2013) and disorder of the cartilage extracellular matrix, including proteoglycan loss and collagen fibril re-arrangement (Tinti et al. 2010; Millucci et al. 2015b; Bernardini et al. 2019). Oxidative stress is a frequent and major consequence of HGA exposure in these studies, as indicated by direct oxidative modification to proteins and depletion of endogenous antioxidants such as glutathione (Braconi et al. 2010a, b, 2011, 2012, 2015; Spreafico et al. 2013). *In vitro* studies have also been important in firmly establishing the pro-oxidant role of ascorbic acid in the presence of HGA. These studies indicate that ascorbic acid is not efficacious as a mediator of HGA-induced oxidative stress and secondary amyloidosis, unless combined with other reducing agents (Tinti et al. 2010).

Animal model of OP formation: In addition, mouse models of AKU, previously only considered to be models of AKU biochemistry, have been shown to develop OP (Taylor et al. 2012; Preston et al. 2014). Despite a marked elevation in plasma and urinary HGA, these mice do not develop the striking macroscopic ochronosis observed in adult humans, except in the kidneys after 13 months (Taylor et al. 2012). Several potential explanations for the lack of widespread macroscopic pigmentation in AKU mice have been postulated. We suggest that the most likely explanations are a) reduced joint-loading of quadrupedal mice, b) faster cellular turnover for removal of pigment and c) that the shorter lifespan of mice compared with humans might be insufficient for gross pigment deposition (Preston et al. 2014). However, pigmentation is observed in AKU mice microscopically; chondrocytes within the articular cartilage begin to show pigmentation at 15 weeks (Preston et al. 2014). This pigmentation

progressively increases over the lifespan of the AKU mouse. Pigmentation of individual chondrocytes in AKU mice mirrors the focal initiation of ochronosis observed in human AKU, in which pigment deposition begins within single chondrocytes of the calcified cartilage and eventually results in widespread extracellular pigmentation (Taylor et al. 2011b). An exact serum HGA concentration threshold at which OP develops is not known, but a threshold range of 40-60 µM/L has been reported in mice (Lewis 2018).

In AKU mice administration of p-hydroxyphenylpyruvate dioxygenase (EC:1.13.11.27) (HPPD) inhibitors, such as nitisinone, soon after birth resulted in failure to pigment, while a similar administration of HPPD inhibitors during the middle of the life span of the mouse, when pigmentation has already developed, prevented progression, but not reversal, of pigmentation (Keenan et al. 2015). This suggests that HPPD inhibitors, by decreasing HGA, can modify the OP process in AKU mice.

Structure of OP: *The classic view of OP formation:* A fundamental question in understanding the process of ochronosis and the development of therapeutic interventions aimed at its prevention or potential reversal concerns the structure of OP itself. The prevailing view on the mechanism by which HGA produces OP in AKU largely comes from work carried out by Zannoni and colleagues in the 1960's. These authors stated that guinea pig cartilage and skin contain enzymes, namely HGA polyphenol oxidases, which were shown to catalyse the *in vitro* oxidation of HGA into a dark, ochronotic-like pigment (Zannoni et al. 1969). BQA was demonstrated as an intermediate in the *in vitro* enzymatic oxidation of HGA, and it was proposed that BQA may polymerize to form OP in AKU (Zannoni et al. 1962, 1969). However, it is important to note the lack of evidence that polyphenol oxidase enzymes are expressed in human or other mammals (Taylor et al. 2016). Enzymes that can oxidise HGA, 'HGA-oxidases', are observed in various species of bacteria known to produce 'pyomelanin' pigment derived from HGA (Hunter and Newman 2010; Roberts et al. 2015). In these species, HGA-derived pyomelanin is thought to serve various adaptive functions including resistance to environmental stress such as UV light and oxidising agents. Subsequent research showed that HGA oxidation can occur non-enzymatically between pH 6.8 and 9.5 in the presence of oxygen (Martin and Batkoff 1987); it is still widely-stated in the literature that oxidation of HGA to benzoquinone intermediates occurs spontaneously in AKU and results in a polymeric pigment structure.

An alternative view of OP chemical structure: Given the general assumption that OP is a product of HGA oxidative polymerisation, it is important to define the term 'polymer'. The accepted definition of a polymer is a large molecule, or 'macromolecule', composed of multiple repeating subunits of a relatively lower molecular weight monomer (McNaught and Wilkinson 2014). The classic concept of a polymer is that the monomer subunits are covalently bound (Allcock et al. 2003).

Roberts et al. (2015) question the widely-held assumption that OP is polymeric. These authors cite the lack of conclusive evidence for this in the literature, and make the point that a polymeric structure is not necessarily required to produce the visual properties of a dark pigment; there are numerous examples of low molecular weight biological pigments. More recent analysis of synthetically-derived OP solutions using size exclusion chromatography suggest that OP has greater molecular weight than HGA, as indicated by a peak at shorter retention time. However, visually pigmented solutions could be formed from HGA (over a shorter period of time; 10 days as opposed to 2 years) without evidence of the peak corresponding to OP and no observed decrease in the HGA peak (Taylor and Vercruysse 2017). This suggests that the visual darkening of the solution due to increased pH can be due to presence of the low molecular weight oxidised form of HGA over a relatively shorter period.

OP has been referred to as 'melanin-like' in the literature (Roberts et al. 2015); largely because melanin is another dark biological pigment derived from tyrosine and classically considered to be formed by polymerisation. However there is a growing body of evidence that melanin and OP are not covalently bound structures, and that their macromolecules do not comprise regularly repeating monomers. A recent study reports data from physicochemical analyses on the 'pyomelanin' pigment produced in the bacteria *rubrivivax benzoatilyticus* (strain JA2) (Mekala et al. 2019). This bacteria mirrors the conditions of OP pigment production in AKU; absence of the *HGD* gene causes accumulation of HGA, resulting in a brown pigment under aerobic conditions and in the presence of phenylalanine (Mekala et al. 2018). Fourier-transform infrared spectroscopy (FTIR) showed a range of band stretching vibrations indicating various chemical groups (aromatic and aliphatic C-H stretches, phenolic C-O stretches, aromatic ring C=C bonds, and C=O stretches due to –COOH groups) characteristic of a pigment structure derived from HGA. X-ray diffraction spectra of the pigment showed similar characteristics to that of melanin, with broad diffraction indicating an 'amorphous' compound structure. The absorbance spectra showed a broad band at

wavelengths 280-350 nm, with increased general absorbance across the UV-visual range without distinct peaks. Similar absorbance properties have been reported previously for OP (Roberts et al. 2015; Taylor and Vercruysse 2017) and also melanin (eumelanin derived from 5,6,-dihydroxyindole-2-carboxylic acid; Tran et al. 2006), and are thought to reflect a chemically heterogeneous structure of oligomers formed by a range of different bonding mechanisms, also referred to as 'chemical disorder' (Riesz 2007). These absorbance properties might also account for the physical appearance of some pigmented substances, which in the case of melanin is thought to provide its physiologically important optical characteristics, *i.e.* its ability to absorb UV light (Chen et al. 2014; Roberts et al. 2015). In contrast, specific, well-defined chemical signatures have been obtained for BQA, the proposed low molecular weight oxidation product of HGA. Specific visual-range absorbance peaks corresponding to BQA have been reported in a series of publications by Tokuhara and colleagues. These peaks were observed at 406 and 430 nm from analysis of solutions of HGA or AKU urine following alkalinisation (Tokuhara et al. 2014, 2018) although other authors have been unable to replicate these findings (Roberts et al. 2015). Specific BQA signals were also recently reported by Tokuhara et al. (2018) from LC-QTOF-MS and NMR analyses.

Computational analyses of eumelanin support a structure formed by stacked eumelanin protomolecules with random-like arrangement; in other words loosely-bound aggregates as opposed to a covalent polymer (Chen et al. 2014). This aggregate structure of melanin is further supported by mass spectrometric and spectrophotometric analyses indicating a possible formation mechanism by self-aggregation of L-dopa by a combination of non-covalent mechanisms including hydrogen bonds, π- π stacking and ionic bonds (Li et al. 2015). The same aggregation was observed for other structurally similar catecholamines which, like HGA, are derived from tyrosine. In support of this data, a recent ultrafiltration study on solutions of pigment derived from synthetic HGA in our laboratory (Norman [unpublished data]) using a 10,000 Da molecular-weight filter suggested a heterogeneous mixture of compounds of varying molecular weight. The filtered solution was still pigmented, although visually lighter in appearance, suggesting presence of some OP molecules >10,000 Da (Figure 5E). Together, these more recent data are inconsistent with the idea that OP or melanin are polymers comprised of regularly repeating units with distinct chemical signatures. The term 'polymer' therefore does not appear to accurately describe the chemical nature of OP or melanin, as currently understood.

Chemical characterisation of OP in human AKU tissue: Another approach employed to study the chemical structure of OP is chemical analysis of heavily pigmented AKU cartilage samples. Solid-state NMR analyses revealed remarkably similar spectra between deeply pigmented and non-pigmented cartilage from AKU patients (Chow et al. 2011; Norman et al. 2018). The spectra were dominated by clear amino acid signals attributable to collagen protein. The absence of a specific NMR signal attributable to OP or related structures could be due to the reduced sensitivity of NMR compared with other analytical platforms such as mass spectrometry. However, with NMR signal enhancement achieved by recent developments using dynamic nuclear polarisation, Norman et al. (2018) observed the first NMR signal attributable to OP in cartilage. A signal at 116.8 ppm in the ^{13}C spectrum (^{1}H-^{13}C FSLG HETCOR 2D NMR experiment) was observed from analysis of pure, dried OP derived from synthetic HGA (3-month incubation at 37°C, as above) and heavily pigmented AKU cartilage. The same signal was observed in non-ochronotic AKU cartilage but much weaker, and even weaker in non-AKU osteoarthritic cartilage. The 2D NMR spectra were more indicative of a poly-hydroquinone versus poly-benzoquinone structure, suggesting that the final structure of OP may not be formed simply by polymerisation of the BQA intermediate.

Raman spectroscopy has recently been identified as a promising analytical technique for studying the nature of OP. Studies of cartilage samples from AKU patients revealed distinct spectra for ochronotic versus non-ochronotic cartilage. Ochronotic samples were highly fluorescent and, unlike non-ochronotic samples, provided limited to no discernible Raman spectral peaks. A novel peak was obtained from the ochronotic sample that was also observed in the spectra from pigment derived from synthetic HGA (Taylor et al. 2019). The ability of Raman spectroscopy to clearly distinguish between ochronotic and non-ochronotic tissue warrants further study, as the technique appears to have potential to provide fundamental information on the chemical nature of OP. Furthermore, there is potential for the technique to be employed as an *in vivo* tool for measuring and monitoring ochronosis progression in a clinical setting (Taylor et al. 2019).

Distribution of ochronosis is not uniform (Table 1; Figures 3, 4): At post-mortem, pigment is patchy and present in areas of stress/damage in non-cartilaginous tissues, suggesting that 'damaged' tissue undergoes pigmentation whereas 'undamaged' tissue is resistant to pigmentation; all cartilage is more consistently and highly pigmented, but also unevenly (Helliwell et al. 2008). Possible factors involved in

'damaging' tissues and producing pigment are shown in Table 1. It is not known if the OP structure is similar in nature in these diverse areas. Despite the constant exposure of all body tissues to HGA in AKU, most tissues are resistant to pigmentation.

Cartilages are especially affected, and these include highly loaded tissues such as articular cartilage, fibrocartilage (intervertebral discs, pubic symphysis), costochondral, as well as less stressed cartilage in the pinna and nose, and the trachea-bronchial system (Helliwell et al. 2008). Highly stressed tissues such as joints, spine, tendons and ligament are pigmented. OP has been noticed in the skin, especially at the interface of the palmar and dorsal skin of the hands, possibly sites of greater stress (Vijaikumar et al. 2000). Pigmentation of the nails of the hands and feet has been described. Secretory glands in the axilla, groin and eyelids pigment (Srsen et al. 1982). Ocular tissues, the conjunctivum, cornea and sclera, pigment in the second and third decade. Ochronosis has been observed in the tympanic membrane and ear wax. Cardiac valves are affected by OP more in left side of heart, and aortic more than mitral valve (Helliwell et al. 2008); the pulmonary vascular system and the right side of the heart is much less pigmented. Pigmentation can be observed in the arterial system, mostly around branch points and tributaries, areas of greater stress (Helliwell et al. 2008). There is currently no information on pigmentation in the venous system. Renal parenchyma, especially the medulla and pyramidal tissues, show OP. The periosteum has been shown to pigment but not the bone itself, raising the possibility of bone mineral binding to collagen preventing HGA-derived molecules from binding. Teeth enamel have been said to pigment, but needs more evidence given that bone does not pigment (Siekert and Gibilisco 1970). It is generally held that muscle, liver, lung (excluding the bronchial system) and brain are not affected by OP. There is no convincing description of gastrointestinal and genital pigmentation. There is no description of pigment in pancreas, an organ where alkaline secretions are produced unlike the salivary gland. It is not known if there is biliary excretion of HGA, even though pigmented gall stones make this likely. Paucicellular tissues such as cartilage, tendon and ligament appear to pigment easily compared to highly cellular tissues such as liver.

It is likely that HGA is increased in the CSF, tears and saliva although direct evidence is lacking. Circulating HGA is well characterised but tissue HGA, i.e. intracellular HGA, has not been directly assessed. It is likely HGA (and other metabolite acids) is protein-bound but requires characterisation.

Detection of ochronosis: Gross ochronosis is easily visible as blue-black pigment in tissues such as the cartilage of the ears and sclera of the eyes. Photographs of the eyes and ears have been used to follow the OP process, both in terms of understanding the natural history as well as to study the effect of reducing HGA concentrations on OP using HPPD inhibitors, and may be the most efficacious way to follow pigment change (Ranganath et al. 2018).

However, to detect small amounts of pigment, more sensitive techniques are required. Such techniques can employ the property of HGA as a reducing agent and can be used in tissues *in vitro* and *ex vivo* to more easily detect the OP; one such approach employs Schmorl's stain, a ferricyanide reduction method (Figure 5B) for tissue reducing substances, to detect microscopic OP (Tinti et al. 2011b); this supports the idea that OP originated from HGA. Biopsy of tissues such as ear cartilage reveals OP even when it is not visible externally visually through the intact skin (Vijaikumar et al. 2000).

Investigations such as arthroscopy and bronchoscopy can reveal OP and suggest diagnosis of AKU for the first time. Diagnosis due to black aorta at open heart surgery for aortic valve replacement has been made (Karavaggelis et al. 2017). Direct detection of OP *in vivo* by Raman spectroscopy is possible. Such a technique has been validated in *ex vivo* tissue (Cox et al. [in press]) and is being adapted as an *in vivo* technique using ear cartilage and Achilles tendon, both tissues with little subcutaneous tissue (Taylor et al. 2019).

Ochronosis can be seen in *ex vivo* samples especially from joints and these can vary in extent and degree of pigment. The earliest pigment is articular cartilage is seen in calcified cartilage cells at the junction between the calcified cartilage and subchondral bone (Taylor et al. 2010c). Similarly, OP is found in intervertebral discs and adjacent articular cartilage (Helliwell et al. 2008). Investigations such as arthroscopy and bronchoscopy can reveal OP and suggest diagnosis of AKU for the first time. Diagnosis due to black aorta at open heart surgery for aortic valve replacement has been made (Cox et al. [in press]).

It would be an advantage to quantify the pigmentation *in vivo* but such a technique is not available at present. Availability of techniques to monitor changes in whole body pigment, increase or decrease or no change over time, would be highly informative.

Magnitude of flux in OP: It can be difficult to know how much metabolite flux is taking place in the phenylalanine/tyrosine pathway as these amino acids are normally fully degraded and utilised. Justus Von Liebig's theory of the minimum applied to dietary amino acid consumption suggests that amino acids lysine, threonine, methionine, and tryptophan are limiting in diet (Liebig 1831). Tyrosine and phenylalanine are not limiting and are therefore consumed to excess. Since only 5% of consumed phenylalanine and tyrosine are needed to meet normal needs, these surplus aromatic amino acids require degradation via HGA. An attempt has been made to quantify the flux in the pigment pathway following HPPD inhibitor therapy (Milan et al. [under review]); data suggests a large flux in the ochronotic pathway.

Effect of ochronosis (Table 2): The deposition of OP in tissue including cartilage alters its material properties leading to the tissue becoming hard and brittle. The Young's modulus is altered depending upon the degree of pigmentation in AKU. Our group has proposed a model of joint failure based on initiation of ochronosis in calcified cartilage before progressing to involve the entire cartilage and spiralling into joint failure, requiring joint replacement (Taylor et al. 2011b).

This process of ochronotic stiffening also compromises intervertebral discs and adjoining bone leading to severe spinal disease characterised by severe pain as well as kyphosis and scoliosis; spinal cord compression by involved discs can require decompressive spinal surgery. Spinal fusion can ensue in the latter stages resulting in loss of mobility and flexibility in all parts of the spine, but especially in the cervical and lumbar regions.

Although all cardiac valves can show ochronosis, it is the aortic valve and the aortic root that are subject to more pigmentation resulting in severe aortic stenosis requiring valve replacement surgery. Valve replacement surgery is often hazardous due to the friable ochronotic aortic root and valve.

Scleral ochronosis can distort the corneal curvature and result in astigmatism (Ranganath [unpublished observations]; Lindner and Bertelmann 2014). Ochronotic ear cartilage may be associated with pain in pinna of the ear (Ranganath [unpublished observations]). Hearing loss especially to high frequency is a feature of AKU (Pau 1984; Steven et al. 2015). Rigid articular cartilage can lead to subchondral osteopenia. Generalised osteoporosis is also noted in AKU and associated with increased fracture (Cox and Ranganath 2011). The failure of ochronotic connective tissue results in tendon ruptures, especially of the Achilles tendon, but other tendon

ruptures have also been observed such as flexor and extensor foot, patellar, bicipital and gluteal regions (Manoj Kumar and Rajasekaran 2003; Ranganath and Cox 2011). Similarly ruptures of ochronotic ligaments have been described. Muscle rupture has also been reported more frequently in AKU. Stone formation featuring OP has been found in kidney, prostate, gall bladder and salivary gland, resulting in symptomatic obstruction of these organs (Taylor et al. 2010a). Renal failure can ensue both due to obstruction and renal parenchymal ochronosis, sometimes leading to fatal intractable haemolytic anaemia (Mullan et al. 2015; Davison et al. 2016). Interestingly, the frequency of atheromatous vascular disease is not increased, despite pigmentation of atherosclerosis plaques. Neo-angiogenesis has been reported in the synovia of AKU patients (Millucci et al. 2016). Millucci and colleagues propose that in AKU angiogenesis and inflammation are inter-related pathological manifestations. These authors postulate that newly-formed blood vessels provide inflammatory cells with oxygen and nutrients, resulting in the release of pro-inflammatory cytokines which support angiogenesis. In this way angiogenesis may contribute to the progression from acute to chronic inflammation in AKU, as is the case in the chronic rheumatic disease synovitis (Millucci et al. 2016). In AKU, the exact association between angiogenesis and OP specifically is not fully understood. However, it is plausible that neo-vascularisation increases access of circulating HGA to the synovium, thereby further propagating ochronosis and associated inflammation.

Despite a sedentary lifestyle imposed by the morbidity of AKU, overweight and obesity is not more common in AKU, and also consequently, there is no increased prevalence of diabetes. It is debatable whether this relates to loss of nutrient, i.e. HGA, in the urine from birth, comparable to the use of inhibitors of renal glucose transport, to induce nutrient loss, in diabetes management.

Mechanism of ochronotic joint and spine disease (Figure 5C): Ageing causes changes in the composition and organization of the extracellular matrix. These include loss of proteoglycans and disruption of collagen fibrils. Trauma can exacerbate these changes. Reactive molecules attack collagen fibres lacking protective proteoglycans. In AKU, homogentisic acid is the culprit leading to ochronosis (Taylor et al. 2010b). HGA-pigment modified collagen fibres become stiffened and less resistant to mechanical loading, leading to a downward spiral of

structural damage. In AKU this cascade is initiated in calcified cartilage and spreads throughout the hyaline cartilage to the articular surface. Ochronosis initiates in calcified articular cartilage, beginning with the deposition of pigment in individual chondrocytes and their territorial matrix in calcified cartilage, spreading to other chondrons in the calcified matrix, then proliferating throughout the hyaline cartilage. Ochronotic cartilage shields the underlying bone from normal mechanical loading, leading to aggressive resorption of the subchondral plate, including calcified cartilage and bone, leading to catastrophic failure of the cartilage and joint itself, with fragments of cartilage escaping into the synovial space as well as impacting into underlying trabecular bone and bone marrow (Taylor et al. 2011b).

It is postulated that a similar sequence of events may also take place in the spine with OP build up in the intervertebral disc and adjacent articular cartilage, leading to increased stress in the vertebral bodies and bone loss. Direct experimental evidence for this is at present lacking.

OP and amyloid: Amyloidosis involves the accumulation of normally soluble proteins into insoluble fibrillar aggregate structures. A growing literature suggests that it is a secondary effect of ochronosis *in vitro* (Millucci et al. 2012; Spreafico et al. 2013; Braconi et al. 2017) and *in vivo* (Millucci et al. 2012, 2014, 2016, 2017). A number of observations report presence of amyloid A protein aggregates and fibrils in AKU serum and ochronotic tissue from a number of locations including cartilage, synovia, heart, periumbrical abdominal, articular fat and labial salivary gland. Amyloidosis is well-recognised in the chronic inflammatory condition rheumatoid arthritis (Obici et al. 2009), and in AKU it is also proposed to result from oxidative protein modification due to presence of reactive oxygen species (Millucci et al. 2015a). The co-localisation of amyloid with shards of OP suggests a close association with ochronosis (Millucci et al. 2015a), although the clinical significance of amyloidosis in AKU remains unknown. In addition, the observation that OP can reverse (Ranganath et al. 2018) is difficult to explain on the basis of amyloid being the major component of pigment.

Linking HGA, ochronosis and damage: The genetic defect and the disease manifestations can be linked as follows. The genetic defect in AKU leads to an increase in HGA. Ochronosis is the result of increased HGA. Ochronotic tissue breaks down causing the multisystem damage in AKU. Conversely, lowering HGA should decrease ochronosis, in turn reducing damage and tissue breakdown.

Incubating osteoblastic (Figure 2D) and chrondrocytic cell lines with HGA leads to development of OP within 3 weeks (Tinti et al. 2011b). OP is seen both within and outside cells. The amount of OP formed is directly proportional to the concentration of HGA in the medium. In the AKU mouse model, OP develops by around 15 weeks increasing progressively in articular joints around the body of the mouse (Taylor et al. 2012; Preston et al. 2014). Decreasing HGA in mice by employing HPPD inhibitors prevents ochronosis when started soon after birth and arrests ochronosis when started later (Taylor et al. 2012; Preston et al. 2014). In humans, the OP process is accelerated when renal failure supervenes and this also leads to more rapid clinical deterioration (Introne et al. 2002).

HGA and causation of Alkaptonuria: Despite the fact that HGA is normal in those without AKU, regulators still consider HGA a biomarker of disease rather than a cause of AKU in their decisions regarding approval of drugs. Debates about causative agents in disease have raged for a long time. In the nineteenth century bacteria were blamed for causing all sorts of diseases, including alkaptonuria. A bacterium in the intestine was blamed as the culprit of the black urine of affected patients that was said to convert tyrosine to homogentisic acid. In the midst of this chaos, Robert Koch established objective rules through which a causative role could be attributed to an organism/agent/factor (Koch 1876; Brown and Goldstein 1992). We have applied Koch's postulates to AKU as seen in Table 3.

Koch's postulates and AKU (Table 3): Over time it came to be recognized that Koch's postulates did not comply with all situations to establish a causal relationship and other criteria have been proposed such as the Bradford-Hill criteria or Hill's criteria for causation (Hill 1965). We have applied Hill's criteria for HGA in AKU also as shown in Table 3. It is clear that HGA conforms to the requirements for a causative molecule.

Lessons for AKU from the cholesterol and atherosclerosis and cardiovascular disease (Table 4): What can AKU learn from other disorders? There are strong similarities between familial hypercholesterolaemia (FH) and AKU; both are inherited diseases that form a template for the more common conditions cardiovascular disease (CVD) and osteoarthritis, respectively. AKU and FH are present from birth but the effects are delayed. HGA is the culprit molecule in AKU, while cholesterol is the molecule in CVD. Ochronosis is the process by which HGA causes the morbidity, while

the comparable process in CVD is atherosclerosis. HPPD inhibitors decrease HGA production, while HMG-CoA reductase inhibitors (statins) decrease cholesterol production. Recent data suggests partial reversibility of external ochronosis by HPPD inhibition, while data from imaging studies have shown a similar partial reversal of atherosclerosis by reducing cholesterol (Nissen et al. 1991). A study has been reported showing a slower progression of AKU using HPPD inhibition, similar to statin trials on CVD progression (Ranganath et al. 2018). However it is worth noting differences in the inheritance of AKU and FH; AKU is autosomal recessive with prevalence of 1:250,000 worldwide, whereas FH is autosomal dominant (Nordestgaard et al. 2013) and relatively more common with estimated worldwide prevalence of 1:200-300 (Vallejo-Vaz and Ray 2018). Table 4 shows the comparison between AKU and FH.

It is worth emphasizing that modification of the disease process (atherosclerosis) by imaging has been used as outcomes in a number of studies in CVD. This is consistent with a proposal to similarly approve the use of modification of the OP process in AKU as acceptable outcomes in clinical interventional studies. Statins have revolutionised the management of cardiovascular disease, and in FH it has been shown that earlier use of statins at low dose is superior in prevention terms compared to use of high dose statins later on in the natural history (Nordestgaard et al. 2013). Mouse studies indicate that nitisinone treatment from soon after birth completely prevents ochronosis, the main pathophysiological process in AKU (Preston et al. 2014).

Therapy and ochronosis:

Antioxidant therapy by preventing oxidation of HGA to BQA, is expected to decrease ochronosis but for reasons which are unclear antioxidant ascorbic acid therapy has no clear benefit in AKU, a condition with a proposed oxidant-damage hypothesis in terms of ochronosis formation (Roberts et al. 2015); this is analogous to the lack of efficacy of antioxidant strategies in coronary artery disease studies despite the well-validated oxidised LDL theory of atherogenesis (Ranganath et al. 2013).

Lower protein intakes should also in theory be associated with less ochronosis and less morbidity in AKU. However, except occasional case reports in childhood, there is no evidence that restricting dietary protein decreases ochronosis. Anecdotally vegetarians and vegans are noted to have less ochronosis and lower morbidity but no systematic evidence exists to support low protein diet in AKU (Ranganath [unpublished observations]). All other approved therapies used in clinical practice are

supportive and palliative and do not address the HGA and its effects (Ranganath et al. 2013).

The use of HPPD inhibition employing nitisinone to decrease flux in the tyrosine pathway has revolutionised the treatment of hereditary tyrosinaemia (Lindstedt et al. 1992; McKiernan 2013). Nitisinone was first suggested as a treatment for HGA in AKU in 1998 (Anikster et al. 1998), and is now shown to be highly effective in reducing HGA despite not being approved in AKU (Ranganath et al. 2018). Nitisinone not only decreases HGA but also arrests ochronosis in mice (Preston et al. 2014). A recent publication shows photographic evidence of partial reversal of ochronosis in sclera and ear cartilage in AKU patients (Ranganath et al. 2018).

The NIH carried out a nitisinone interventional study in AKU between 2005 and 2009 and showed a sustained and marked decrease in urine HGA over the 3-year duration of the study (Davison et al. 2015). However, their agreement with the regulatory agency, the FDA, to approve nitisinone for AKU, required them to show a difference in range of motion at the hip between the nitisinone-treated and untreated groups; statistical significance was not found for the range of motion comparison and the study reported inconclusive.

The dose of nitisinone used in the NIH study was 2 mg daily. This same dose is being used in the NAC, as mentioned in a recent publication (Ranganath et al. 2018). Further the NAC data in the publication confirmed the biochemical efficacy of nitisinone in AKU in terms of urinary HGA, but also showed slower progression of morbidity. Finally, the NAC data showed overt partial reversal of ochronosis, the primary pathogenetic event in AKU.

It is important to note that nitisinone, while shown to be a highly effective therapy for preventing ochronosis, is not a perfect therapy in AKU. It is well recognised that nitisinone treatment results in siginificant hyper-tyrosinaemia in AKU (Phornphutkul et al. 2002) and hereditary tyrosinaemia type-I (HT-I) (Lindstedt et al. 1992; McKiernan 2013). The consequences of hyper-tyrosinaemia are not fully understood in AKU, but concerns have been raised that it may contribute to the neurodevelopmental delay observed in infants with HT-I on nitisinone therapy (McKiernan 2013). Recent analyses did not find changes to monoamine neurotransmitters in brain tissue from nitisinone-treated mice (Davison et al. 2019), although more data are required to fully ascertain the impact of hyper-tyrosinaemia on central nervous system homoeostasis. Other potential future therapies for AKU may employ approaches to directly restore HGD

activity for example by enzyme replacement therapy and gene therapy. However research into these approaches is still in very early stages, and they are also not without their own concerns and challenges, for example potential off-target effects.

We believe that it is time that the scientific community recognised the fundamental role played by pigmentation due to ochronosis in AKU and to take this into consideration when assessing the disease as well as effectiveness of treatments for the disease. The accumulation of HGA in AKU is comparable to the increase in xylulose in pentosuria, but unlike pentosuria where there is no syndrome associated with pentose sugar accumulation, the situation is different in AKU, a condition dominated by ochronosis. While much has been learnt about ochronosis, there still remains unanswered questions.

References

Allcock H, Lampe F, Mark J (2003) Contemporary Polymer Chemistry, 3rd edn. Prentice Hall, New Jersey

Anikster Y, Nyhan WL, Gahl WA (1998) NTBC and Alkaptonuria. Am J Hum Genet 63:920–921. doi: 10.1086/302027

Bernardini G, Leone G, Millucci L, et al (2019) Homogentisic acid induces morphological and mechanical aberration of ochronotic cartilage in alkaptonuria. J Cell Physiol 234:6696–708. doi: 10.1002/jcp.27416

Bhattar P, Zawar V, Godse K, et al (2015) Exogenous Ochronosis. Indian J Dermatol 60:537–42. doi: 10.4103/0019-5154.169122

Braconi D, Bernardini G, Bianchini C, et al (2012) Biochemical and proteomic characterization of alkaptonuric chondrocytes. J Cell Physiol 227:3333–43. doi: 10.1002/jcp.24033

Braconi D, Bianchini C, Bernardini G, et al (2011) Redox-proteomics of the effects of homogentisic acid in an in vitro human serum model of alkaptonuric ochronosis. J Inherit Metab Dis 34:1163–76. doi: 10.1007/s10545-011-9377-6

Braconi D, Laschi M, Amato L, et al (2010a) Evaluation of anti-oxidant treatments in an in vitro model of alkaptonuric ochronosis. Rheumatology 49:1975–83. doi: 10.1093/rheumatology/keq175

Braconi D, Laschi M, Taylor AM, et al (2010b) Proteomic and redox-proteomic evaluation of homogentisic acid and ascorbic acid effects on human articular chondrocytes. J Cell Biochem 111:922–32. doi: 10.1002/jcb.22780

Braconi D, Millucci L, Bernardini G, Santucci A (2015) Oxidative stress and mechanisms of ochronosis in alkaptonuria. Free Radic Biol Med 88(Pt A):70–80. doi: 10.1016/j.freeradbiomed.2015.02.021

Braconi D, Millucci L, Bernini A, et al (2017) Homogentisic acid induces aggregation and fibrillation of amyloidogenic proteins. Biochim Biophys acta Gen Subj 1861:135–146. doi: 10.1016/j.bbagen.2016.11.026

Brown MS, Goldstein JL (1992) Koch's postulates for cholesterol. Cell 71:187–188. doi: 10.1016/0092-8674(92)90346-E

Chen CT, Chuang C, Cao J, et al (2014) Excitonic effects from geometric order and disorder explain broadband optical absorption in eumelanin. Nat Commun 5:3859. doi: 10.1038/ncomms4859

Chow WY, Taylor AM, Reid DG, et al (2011) Collagen atomic scale molecular disorder in ochronotic cartilage from an alkaptonuria patient, observed by solid state NMR. J Inherit Metab Dis 34:1137–40. doi: 10.1007/s10545-011-9373-x

Cox T, Psarelli E, Taylor S, et al Subclinical ochronosis features in Alkaptonuria: a cross-sectional study. BMJ Innov in press:

Cox TF, Ranganath L (2011) A quantitative assessment of alkaptonuria: Testing the reliability of two disease severity scoring systems. J Inherit Metab Dis 34:1153–62. doi: 10.1007/s10545-011-9367-8

Damian LO, Felea I, Boloşiu C, et al (2013) A case of alkaptonuria - ultrasonographic findings. Med Ultrason 15:321–25. doi: 10.11152/mu.2013.2066.154.lod2

Davison AS, Milan AM, Gallagher JA, Ranganath LR (2016) Acute fatal metabolic complications in alkaptonuria. J Inherit Metab Dis 39:203–10. doi: 10.1007/s10545-015-9902-0

Davison AS, Milan AM, Hughes AT, et al (2015) Serum concentrations and urinary excretion of homogentisic acid and tyrosine in normal subjects. Clin Chem Lab Med 53:e81-3. doi: 10.1515/cclm-2014-0668

Davison AS, Strittmatter N, Sutherland H, et al (2019) Assessing the effect of

nitisinone induced hypertyrosinaemia on monoamine neurotransmitters in brain tissue from a murine model of alkaptonuria using mass spectrometry imaging. Metabolomics 15:81. doi: 10.1007/s11306-019-1531-4

Fisher AA, Davis MW (2004) Alkaptonuric ochronosis with aortic valve and joint replacements and femoral fracture: a case report and literature review. Clin Med Res 2:209–15. doi: 10.3121/cmr.2.4.209

Gaines JJ, Tom GD, Khankhanian N (1987) An ultrastructural and light microscopic study of the synovium in ochronotic arthropathy. Hum Pathol 18:1160–4. doi: 10.1016/S0046-8177(87)80385-4

Galdston M, Steele JM, Dobriner K (1952) Alcaptonuria and ochronosis; with a report of three patients and metabolic studies in two. Am J Med 13:432–52

Gallagher JA, Dillon JP, Sireau N, et al (2016) Alkaptonuria: An example of a "fundamental disease"-A rare disease with important lessons for more common disorders. Semin Cell Dev Biol 52:52–57. doi: 10.1016/j.semcdb.2016.02.020

Garrod AE (1902) The incidence of alkaptonuria: a study in chemical individuality. Lancet 2:1616–20. doi: 10.1016/S0140-6736(01)41972-6

Helliwell TR, Gallagher J a., Ranganath L (2008) Alkaptonuria - A review of surgical and autopsy pathology. Histopathology 53:503–512. doi: 10.1111/j.1365-2559.2008.03000.x

Hill AB (1965) The Environment and Disease: Association or Causation? Proc R Soc Med 58:295

Hunter RC, Newman DK (2010) A putative ABC transporter, hatABCDE, is among molecular determinants of pyomelanin production in Pseudomonas aeruginosa. J Bacteriol 192:5962–71. doi: 10.1128/JB.01021-10

Introne WJ, Perry MB, Troendle J, et al (2011) A 3-year randomized therapeutic trial of nitisinone in alkaptonuria. Mol Genet Metab 103:307–14. doi: 10.1016/j.ymgme.2011.04.016

Introne WJ, Phornphutkul C, Bernardini I, et al (2002) Exacerbation of the ochronosis of alkaptonuria due to renal insufficiency and improvement after renal transplantation. Mol Genet Metab 77:136–142. doi: 10.1016/S1096-7192(02)00121-X

Karavaggelis A, Young C, Attia R (2017) Black Heart at Surgery - Primary Diagnosis of Alkaptonuria at Surgery. J Cardiol Curr Res 9:00335. doi: 10.15406/jccr.2017.09.00336

Keenan CM, Preston AJ, Sutherland H, et al (2015) Nitisinone arrests but does not reverse ochronosis in alkaptonuric mice. J Inherit Metab Dis 73:284-9.

Knox WE (1958) Sir Archibald Garrod's Inborn Errors of Metabolism. IV. Pentosuria. Am J Hum Genet 10:385–97

Koch R (1876) Investigations into bacteria: V. The etiology of anthrax, based on the ontogenesis of Bacillus anthracis [in German]. Cohns Beitrage zur Biol der Pflanz 2:227–310

La Du BN, Zannoni VG, Laster L, Seegmiller JE (1958) The nature of the defect in tyrosine metabolism in alcaptonuria. J Biol Chem 230:251–60

Levin CY, Maibach H (2001) Exogenous ochronosis: An update on clinical features, causative agents and treatment options. Am J Clin Dermatol 2:213–17. doi: 10.2165/00128071-200102040-00002

Lewis R (2018) Consequences and prevention of elevated circulating tyrosine during nitisinone therapy in alkaptonuria. Liverpool John Moores University

Li Y, Liu J, Wang Y, et al (2015) Mass Spectrometric and Spectrophotometric Analyses Reveal an Alternative Structure and a New Formation Mechanism for

Melanin. Anal Chem 87:7958–63. doi: 10.1021/acs.analchem.5b01837

Liebig J (1831) Ueber einen neuen Apparat zur Analyse organischer Körper, und über die Zusammensetzung einiger organischen Substanzen. Ann Phys 21:1–47. doi: 10.1002/andp.18310970102

Lindner M, Bertelmann T (2014) On the ocular findings in ochronosis: a systematic review of literature. BMC Ophthalmol 14:12. doi: 10.1186/1471-2415-14-12

Lindstedt S, Holme E, Lock EA, et al (1992) Treatment of hereditary tyrosinaemia type I by inhibition of 4-hydroxyphenylpyruvate dioxygenase. Lancet 340:813–17. doi: 10.1016/0140-6736(92)92685-9

Manoj Kumar R V, Rajasekaran S (2003) Spontaneous tendon ruptures in alkaptonuria. J Bone Joint Surg Br 85:883–6

Martin JP, Batkoff B (1987) Homogentisic acid autoxidation and oxygen radical generation: implications for the etiology of alkaptonuric arthritis. Free Radic Biol Med 3:241–50

McKiernan PJ (2013) Nitisinone for the treatment of hereditary tyrosinemia type I. Expert Opin Orphan Drugs 1:491–7. doi: 10.1517/21678707.2013.800807

McNaught AD, Wilkinson A (2014) IUPAC. Compendium of Chemical Terminology 2nd ed. (the "Gold Book"). Blackwell Scientific Publications, Oxford

Mekala LP, Mohammed M, Chintalapati S, Chintalapati VR (2018) Stable Isotope-Assisted Metabolic Profiling Reveals Growth Mode Dependent Differential Metabolism and Multiple Catabolic Pathways of l -Phenylalanine in Rubrivivax benzoatilyticus JA2. J Proteome Res 17:189–202. doi: 10.1021/acs.jproteome.7b00500

Mekala LP, Mohammed M, Chinthalapati S, Chinthalapati VR (2019) Pyomelanin production: Insights into the incomplete aerobic L-phenylalanine catabolism of a photosynthetic bacterium, Rubrivivax benzoatilyticus JA2. Int J Biol Macromol 126:755–64. doi: 10.1016/j.ijbiomac.2018.12.142

Milan A, Hughes A, Davison A, et al Quantification of tyrosine flux in the ochronotic pathway during nitisinone treatment of Alkaptonuria. (under review)

Millucci L, Bernardini G, Marzocchi B, et al (2016) Angiogenesis in alkaptonuria. J Inherit Metab Dis 39:801–06. doi: 10.1007/s10545-016-9976-3

Millucci L, Bernardini G, Spreafico A, et al (2017) Histological and Ultrastructural Characterization of Alkaptonuric Tissues. Calcif Tissue Int 101:50–64. doi: 10.1007/s00223-017-0260-9

Millucci L, Braconi D, Bernardini G, et al (2015a) Amyloidosis in alkaptonuria. J Inherit Metab Dis 38:797–805. doi: 10.1007/s10545-015-9842-8

Millucci L, Ghezzi L, Braconi D, et al (2014) Secondary amyloidosis in an alkaptonuric aortic valve. Int J Cardiol 172:e121-3. doi: 10.1016/j.ijcard.2013.12.117

Millucci L, Giorgetti G, Viti C, et al (2015b) Chondroptosis in alkaptonuric cartilage. J Cell Physiol 230:1148–57. doi: 10.1002/jcp.24850

Millucci L, Spreafico A, Tinti L, et al (2012) Alkaptonuria is a novel human secondary amyloidogenic disease. Biochim Biophys Acta - Mol Basis Dis 1822:1682–91. doi: 10.1016/j.bbadis.2012.07.011

Mistry JB, Jackson DJ, Bukhari M, Taylor AM (2016) A role for interleukins in ochronosis in a chondrocyte in vitro model of alkaptonuria. Clin Rheumatol 35:1849–56. doi: 10.1007/s10067-015-3091-y

Mullan A, Cocker D, Taylor G, et al (2015) Fatal oxidative haemolysis and methaemoglobinaemia in a patient with alkaptonuria and acute kidney injury. Clin Kidney J 8:109–12. doi: 10.1093/ckj/sfu121

Nissen SE, Gurley JC, Grines CL, et al (1991) Intravascular ultrasound assessment of lumen size and wall morphology in normal subjects and patients with coronary artery disease. Circulation 84:1087–99. doi: 10.1161/01.CIR.84.3.1087

Nordestgaard BG, Chapman MJ, Humphries SE, et al (2013) Familial hypercholesterolaemia is underdiagnosed and undertreated in the general population: guidance for clinicians to prevent coronary heart disease: Consensus Statement of the European Atherosclerosis Society. Eur Heart J 34:3478–3490. doi: 10.1093/eurheartj/eht273

Norman B Unpublished data

Norman BP, Ying CW, Sutherland H, et al (2018) Disruption of collagen triple helix hydrogen bonding in ochronotic human cartilage in alkaptonuria observed by dynamic nuclear polarisation-enhanced solid-state nuclear magnetic resonance spectroscopy. Osteoarthr Cartil 26:S99. doi: 10.1016/j.joca.2018.02.214

O'Brien WM, La Du BN, Bunim JJ (1963) Biochemical, pathologic and clinical aspects of alcaptonuria, ochronosis and ochronotic arthropathy - Review of world literature (1584-1962). Am J Med 34:813–38. doi: 10.1385/1-59259-763-7

Obici L, Raimondi S, Lavatelli F, et al (2009) Susceptibility to AA amyloidosis in rheumatic diseases: a critical overview. Arthritis Rheum 61:1435–40. doi: 10.1002/art.24735

Pau HW (1984) Involvement of the tympanic membrane and ear ossicle system in ochronotic alkaptonuria [German]. Laryngol Rhinol Otol (Stuttg) 63:541–4

Phornphutkul C, Introne WJ, Perry MB, et al (2002) Natural History of Alkaptonuria. N Engl J Med 347:2111–21. doi: 10.1056/nejmoa021736

Preston AJ, Keenan CM, Sutherland H, et al (2014) Ochronotic osteoarthropathy in a mouse model of alkaptonuria, and its inhibition by nitisinone. Ann Rheum Dis 73:284–9. doi: 10.1136/annrheumdis-2012-202878

Ranganath L Unpublished observations

Ranganath LR, Cox TF (2011) Natural history of alkaptonuria revisited: Analyses based on scoring systems. J Inherit Metab Dis 34:1141–51. doi: 10.1007/s10545-011-9374-9

Ranganath LR, Jarvis JC, Gallagher JA (2013) Recent advances in management of alkaptonuria (invited review; best practice article). J Clin Pathol 66:367–373. doi: 10.1136/jclinpath-2012-200877

Ranganath LR, Khedr M, Milan AM, et al (2018) Nitisinone arrests ochronosis and decreases rate of progression of Alkaptonuria: Evaluation of the effect of nitisinone in the United Kingdom National Alkaptonuria Centre. Mol Genet Metab 125:127–34. doi: 10.1016/j.ymgme.2018.07.011

Riesz J (2007) The spectroscopic properties of melanin. University of Queensland

Roberts NB, Curtis SA, Milan AM, Ranganath LR (2015) The pigment in alkaptonuria relationship to melanin and other coloured substances: A review of metabolism, composition and chemical analysis. JIMD Rep 24:51–66. doi: 10.1007/8904_2015_453

Siekert RG, Gibilisco JA (1970) Discoloration of the teeth in alkaptonuria (ochronosis) and parkinsonism. Oral Surgery, Oral Med Oral Pathol 29:197–99. doi: 10.1016/0030-4220(70)90081-2

Spreafico A, Millucci L, Ghezzi L, et al (2013) Antioxidants inhibit SAA formation and pro-inflammatory cytokine release in a human cell model of alkaptonuria. Rheumatol (United Kingdom) 52:1667–73. doi: 10.1093/rheumatology/ket185

Srsen S, Srsnová K, Lányi A (1982) [Clinical manifestation of alkaptonuria in relation to age (author's transl)]. Bratisl Lek Listy 77:662–9

Steven RA, Kinshuck AJ, McCormick MS, Ranganath LR (2015) ENT manifestations of alkaptonuria: report on a case series. J Laryngol Otol 129:1004–8. doi: 10.1017/S0022215115002315

Taylor AM, Batchelor TJP, Adams VL, et al (2011a) Ochronosis and calcification in the mediastinal mass of a patient with alkaptonuria. J Clin Pathol 64:935–6. doi: 10.1136/jcp.2011.090126

Taylor AM, Boyde A, Wilson PJM, et al (2011b) The role of calcified cartilage and subchondral bone in the initiation and progression of ochronotic arthropathy in alkaptonuria. Arthritis Rheum 63:3387–96. doi: 10.1002/art.30606

Taylor AM, Hsueh MF, Ranganath LR, et al (2017) Cartilage biomarkers in the osteoarthropathy of alkaptonuria reveal low turnover and accelerated ageing. Rheumatology (Oxford) 56:156–64. doi: 10.1093/rheumatology/kew355

Taylor AM, Jenks DD, Kammath VD, et al (2019) Raman Spectroscopy identifies differences in ochronotic and non-ochronotic cartilage; a potential novel technique for monitoring ochronosis. Osteoarthr Cartil S1063-4584:. doi: 10.1016/j.joca.2019.04.012

Taylor AM, Kammath V, Bleakley A (2016) Tyrosinase, could it be a missing link in ochronosis in alkaptonuria? Med Hypotheses 91:77–80. doi: 10.1016/j.mehy.2016.04.001

Taylor AM, Preston AJ, Paulk NK, et al (2012) Ochronosis in a murine model of alkaptonuria is synonymous to that in the human condition. Osteoarthr Cartil 20:880–886. doi: 10.1016/j.joca.2012.04.013

Taylor AM, Vercruysse KP (2017) Analysis of Melanin-like Pigment Synthesized from Homogentisic Acid, with or without Tyrosine, and Its Implications in Alkaptonuria. JIMD Rep 35:79–85. doi: 10.1007/8904_2016_27

Taylor AM, Wilson PJM, Ingrams DR, et al (2010a) Calculi and intracellular ochronosis in the submandibular tissues from a patient with alkaptonuria. J Clin Pathol 63:186–8. doi: 10.1136/jcp.2009.071365

Taylor AM, Wlodarski B, Prior IA, et al (2010b) Ultrastructural examination of tissue in a patient with alkaptonuric arthropathy reveals a distinct pattern of binding of ochronotic pigment. Rheumatology 49:1412–14. doi: 10.1093/rheumatology/keq027

Taylor AM, Wlodarski B, Wilson PJ, et al (2010c) Deposition of ochronotic pigment in articular cartilage in alkaptonuria is initiated near the tidemark and progresses to the articular surface. Bone 47:S79. doi: 10.1016/j.bone.2010.04.156

Tinti L, Spreafico A, Braconi D, et al (2010) Evaluation of antioxiodant drugs for the treatment of ochronotic alkaptonuria in an in vitro human cell model. J Cell Physiol 225:84–91. doi: 10.1002/jcp.22199

Tinti L, Spreafico A, Chellini F, et al (2011a) A novel ex vivo organotypic culture model of alkaptonuria-ochronosis. Clin Exp Rheumatol 29:693–6

Tinti L, Taylor AM, Santucci A, et al (2011b) Development of an in vitro model to investigate joint ochronosis in alkaptonuria. Rheumatology 50:271–7. doi: 10.1093/rheumatology/keq246

Tokuhara Y, Shukuya K, Tanaka M, et al (2014) Detection of Novel Visible-Light Region Absorbance Peaks in the Urine after Alkalization in Patients with Alkaptonuria. PLoS One 9:e86606. doi: 10.1371/journal.pone.0086606

Tokuhara Y, Shukuya K, Tanaka M, et al (2018) Absorbance measurements of oxidation of homogentisic acid accelerated by the addition of alkaline solution with sodium hypochlorite pentahydrate. Sci Rep 8:11364. doi: 10.1038/s41598-018-29769-w

Tran ML, Powell BJ, Meredith P (2006) Chemical and structural disorder in eumelanins: A possible explanation for broadband absorbance. Biophys J 90:743–52. doi: 10.1529/biophysj.105.069096

Vallejo-Vaz AJ, Ray KK (2018) Epidemiology of familial hypercholesterolaemia: Community and clinical. Atherosclerosis 277:289–297. doi: 10.1016/j.atherosclerosis.2018.06.855

Vijaikumar M, Thappa DM, Srikanth S, et al (2000) Alkaptonuric ochronosis presenting as palmoplantar pigmentation. Clin Exp Dermatol 25:305–7. doi: 10.1046/j.1365-2230.2000.00649.x

Virchow R (1866) Ein Fall von allgemeiner Ochronose der Knorpel und knorpelähnlichen Theile. Arch für Pathol Anat und Physiol und für Klin Med 37:121–19. doi: 10.1007/bf01935634

Zannoni VG, Lomtevas N, Goldfinger S (1969) Oxidation of homogentisic acid to ochronotic pigment in connective tissue. Biochim Biophys Acta 177:94–105

Zannoni VG, Malawista SE, La Du BN (1962) Studies on ochronosis. II. Studies on benzoquinoneacetic acid, a probable intermediate in the connective tissue pigmentation of alcaptonuria. Arthritis Rheum 5:547–556. doi: 10.1002/art.1780050603

Tables

Table 1. Factors influencing tissue ochronotic pigmentation

Tissue	Damaging factor
Spine intervertebral discs	Weight bearing stress
Joints	Weight bearing stress, movement damage
Tendons	Tensile stresses
Ligaments	Tensile stresses
Aortic valve	Systemic blood pressure, expansile stress
Aortic root	Systemic blood pressure, expansile stress
Arterial tree branch points	Shear stress, Bernoulli effect
Airways cartilage	Expansion, lengthening and contracting
Ear cartilage	Pressure on ears during sleep
Sclera and conjunctiva	UV light damage, stress arising from muscle contraction

Table 2. The effect of ochronosis on tissues

Tissue with ochronosis	Effect
Articular cartilage	Resorption of calcified articular cartilage and subchondral bone and joint failure
Articular cartilage	Osteopenia due to resorption activated by stiff pigment
Intervertebral discs	Spondylosis, fracture, osteopenia, scoliosis, kyphosis, cord compression, radiculopathy
Ligament	Rupture
Aortic valve and root	Aortic stenosis and Aortic rigidity
Tendons (e.g.Achilles, foot flexor and extensor, patellar, gluteal, biceps)	Rupture
Sclera	Distortion in corneal curvature; astigmatism, glaucoma
Ear cartilage	Possible pain in ear
Middle/inner ear	Hearing loss, ear ossicles disorder leading to conductive deafness and high frequency hearing loss

Table 3. Koch's postulates and its application to Alkaptonuria

Koch's postulates rules	Relevance to AKU
The agent must be found in abundance in all organisms suffering from the disease, but should not be found in healthy organisms	HGA found in abundance in AKU but not in non-AKU
The microorganism must be isolated from a diseased organism and grown in pure culture.	HGA can be isolated from AKU patients by preparative HPLC techniques to yield pure HGA.
The cultured microorganism should cause disease when introduced into a healthy organism.	HGA can be incubated with cell cultures and shown to produce the features of AKU namely pigment.
The microorganism must be re-isolated from the inoculated, diseased experimental host and identified as being identical to the original specific causative agent.	Producing a mutation of the HGD gene in a normal mouse can produce an AKU mouse showing high HGA.

Hill's Criteria and HGA and cause in AKU	
Bradford-Hill criteria for causation	**How does AKU fit in**
Strength: A small association does not mean that there is not a causal effect, though the larger the association, the more likely that it is causal	Strong relationship between HGA and disease severity.

Consistency: Consistent findings observed by different persons in different places with different samples strengthens the likelihood of an effect.	Link between HGA in AKU and disease by various dispersed researchers.
Specificity: Causation is likely if there is a very specific population at a specific site and disease with no other likely explanation. The more specific an association between a factor and an effect is, the bigger the probability of a causal relationship	Only HGA causes the disease process characteristic of AKU
Temporality: The effect has to occur after the cause (and if there is an expected delay between the cause and expected effect, then the effect must occur after that delay)	In mouse models where the mutation of HGD has been produced, HGA levels increase after the mutation. The disease process however develops after slight delay both in mouse and humans.
Biological Gradient: Greater exposure should generally lead to greater incidence of the effect. However, in some cases, the mere presence of the factor can trigger the effect. In other cases, an inverse proportion is observed: greater exposure leads to lower incidence	Places with high frequency of HGD mutations have high HGA and AKU disease.

Plausibility: A plausible mechanism between cause and effect is helpful (but Hill noted that knowledge of the mechanism is limited by current knowledge)	HGA can reproduce the disease process ochronosis *in vivo* and is therefore plausible.
Coherence: Coherence between epidemiological and laboratory findings increases the likelihood of an effect.	The relationship between mutations of the HGD gene, high HGA, increasing ochronosis over time and increasing morbidity is coherent.
Experiment: Occasionally it is possible to appeal to experimental evidence	Tissues exposed to HGA produce a similar disease process to naturally occurring condition AKU
Analogy: The effect of similar factors may be considered	HGA in AKU is similar to cholesterol in atheromatous coronary disease.

Table 4. Comparison between Familial Hypercholesterolaemia and Alkaptonuria

	Familial hypercholesterolaemia	Alkaptonuria
Frequency	Heterozygote 1 in 500 Homozygote 1 in 1,000,000	Heterozygote 1 in 500 Homozygote 1 in 1,000,000
Causative agent	Cholesterol (LDL)	Homogenstic acid
Condition	Familial hypercholesterolaemia	Alkaptonuria
Genetic defect	LDL receptor, Apo-B defects, PCSK9 mutations	HGD mutations
Disease process	Atherosclerosis	Ochronosis
Main consequence	Atherosclerosis	Ochronosis
Main disease	Myocardial infarction	Spondyloarthropathy
Childhood	Minimal disease	Minimal disease
Latent period	40 -50 years in heterozygous 20 years in homozygous	20-30y
Prevention	Yes, lifestyle and statins	Nitisinone
Lifestyle factors	Yes, hypertension, physical activity, Smoking, diabetes, others	Diet, activity, occupation
Disease modifying therapy	HMG CoA inhibitors (Statins), others	HPPD inhibitors, Nitisinone
Reversibility	Partial	Partial

Earlier treatment beneficial	Yes	Probably yes, mouse data

Figure legends

Figure 1. Tyrosine metabolic pathway – highlighting (1) the metabolic fate of tyrosine in health, (2) site of the enzyme defect observed in Alkaptonuria, *homogentisate dioxygenase* (*HGD* EC 1.13.11.5) and Hereditary Tyrosinaemia type 1, maleylacetoacetate isomerase (*MAI* EC 5.2.1.2), and (3) the site where nitisinone inhibits *4-hydroxyphenylpyruvate dioxygenase* (*HPPD* EC 1.13.11.27) activity

Figure 2. Ochronotic pigmentation features (1). A. Ochre or yellowish discoloration due to ochronotic pigment in the unstained ear cartilage in two patients with alkaptonuria; less in left panel, more in right panel. B. TEM image of ochronotic ligamentous capsule. Collagen fibres in longitudinal section show a distinct electron-dense pigment on their surface. Not all fibres present with pigment deposition. Numerous pigment shards can be seen on single fibres. Gradient of pigmentation can be seen running left (no pigment on fibres) to right (large electron-dense shards replacing fibres). Arrows indicate a distinct periodic binding pattern associated with pigment granules on a single fibre. C. Schematic representation of the development of ochronosis. Genetic lack of homogentisate 1,2-dioxygenase leads to an increase in concentration of homogentisic acid (HGA). HGA, its oxidation product, benzoquinone acetic acid or the final product ochronotic pigment binds to collagenous matrices. Initially, matrix is resistant to pigmentation, but following loss or breakdown of specific constituents, including proteoglycans (PG), HGA-associated compounds access binding sites which are associated with the ultrastructural periodicity of the collagen fibrils. It is proposed that the initial binding event initiates ochronosis and that the process of widespread joint pigmentation occurs over time. Pigmentation increases the stiffness of the collagen fibres, which leads to further biomechanical and biochemical damage and a downward spiral of ochronosis and tissue destruction. D. Schmorl's staining of HGA-derived pigment deposits in cultures of SaOS-2 cells. E. Samples obtained during ultrafiltration of solutions of ochronotic pigment formed via incubation of an aqueous solution of HGA (10 mmol/L) at 37 °C for 3 months. Solutions 1 and 2 are samples pre- and post-centrifugation (10 min at 2500 x g) respectively. The pigment solution was filtered using an Amicon Ultracel 10 K filter (Merck Millipore) by centrifugation (10 min at 2500 x g). Solution 3 is taken from the portion of the solution that had not passed the filter after 10 min centrifugation. Solution 4 is a sample of the filtrate and is visibly lighter in colour, indicating that the 10 K filter had retained some of the pigment.

Figure 3. Ochronotic pigmentation features (2). A. Dark urine of AKU. B. External ear cartilage pigmentation. C. Dark pigmentation of temporal aspect of sclera in right eye with vessels coursing superficial to pigmentation. D. Unstained cut section of femur condyle showing little pigment on left side progressing to full thickness on the right side. E. Longitudinal cut section of abdominal aorta and common iliac bifurcation, showing more pigment at bifurcation and branch point orifices. F. Pigment at junction of palmer and dorsal skin of hands.

Figure 4. Ochronotic pigmentation features (3). A. Markedly ochronotic bulging intervertebral discs and vertebral body seen from within abdominal cavity. B. Spine seen from posterior or dorsal aspect showing ochronotic pigment. C. Dark

pigmentation seen in arthroscopy of knee joint showing fibrillar blackened cartilage. D. Appearance of pigmented head of femur showing uniform pigmentation. E. Hip joint showing rim of cartilage and marked cartilage loss with exposure of underlying bone. F. Low pressure pulmonary trunk and valve showing little pigment. G. Aortic root and valve with marked pigmentation.

Figure 5. Presence of ochronotic pigment in the knee joint and its progression in AKU. A. Ochronotic pigment in the joint capsule of the knee of a patient with AKU. Pigment deposits are seen as brown granules in the ECM and within fibroblasts. Section stained with nuclear fast red. B. Near serial section of the joint capsule with Schmorl's stain. Ochronotic deposits are stained green. Bar: 50 µm. C. Schematic representation of the progression of ochronosis in articular cartilage from initiation in calcified cartilage to eventual destruction of the joint: a) ochronosis begins with the deposition of pigment in individual chondrocytes and their territorial matrix in calcified cartilage. Pigmentation leads to focal increases of stiffness altering the load distribution and inducing stress risers; b) ochronosis spreads to other chondrons in the calcified matrix, then c) proliferates throughout the hyaline cartilage; d) ochronotic cartilage shields the underlying bone from normal mechanical loading, leading to aggressive resorption of the subchondral plate, including calcified cartilage and bone; e) despite the increased stiffness, the pigmented shell of the remaining articular cartilage fails catastrophically. Pigmented cartilage becomes impacted on the underlying trabecular bone and embedded in the marrow space.
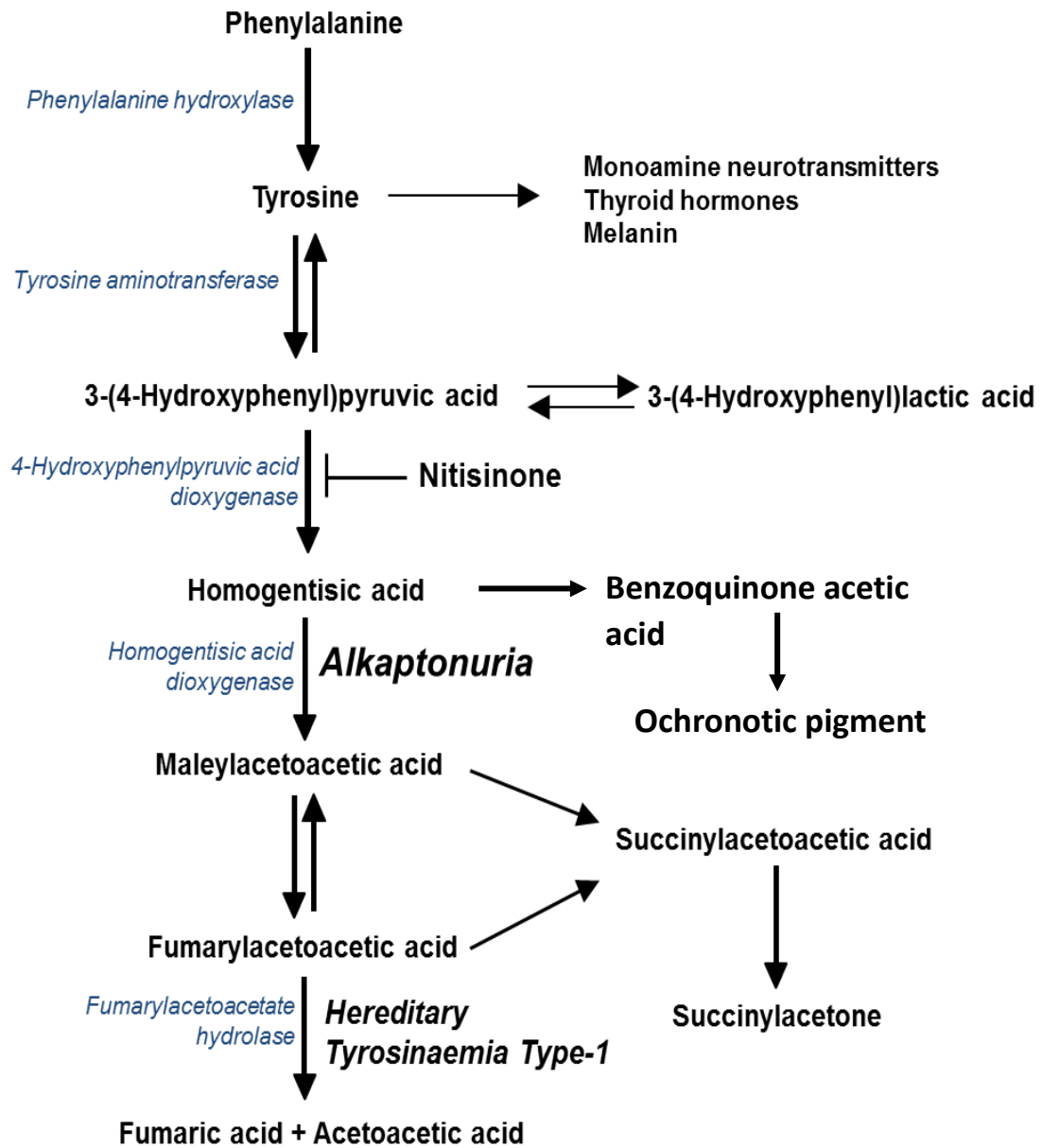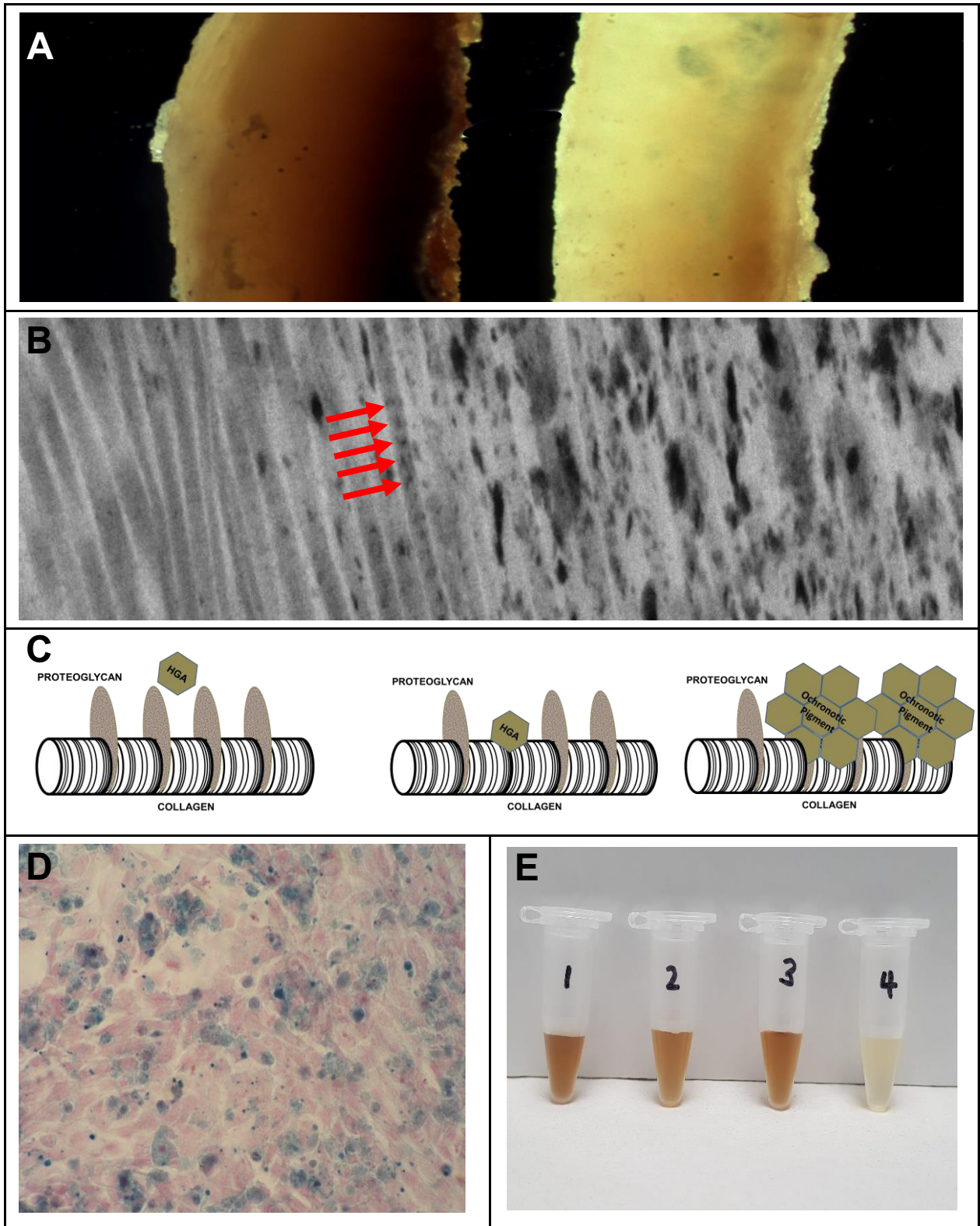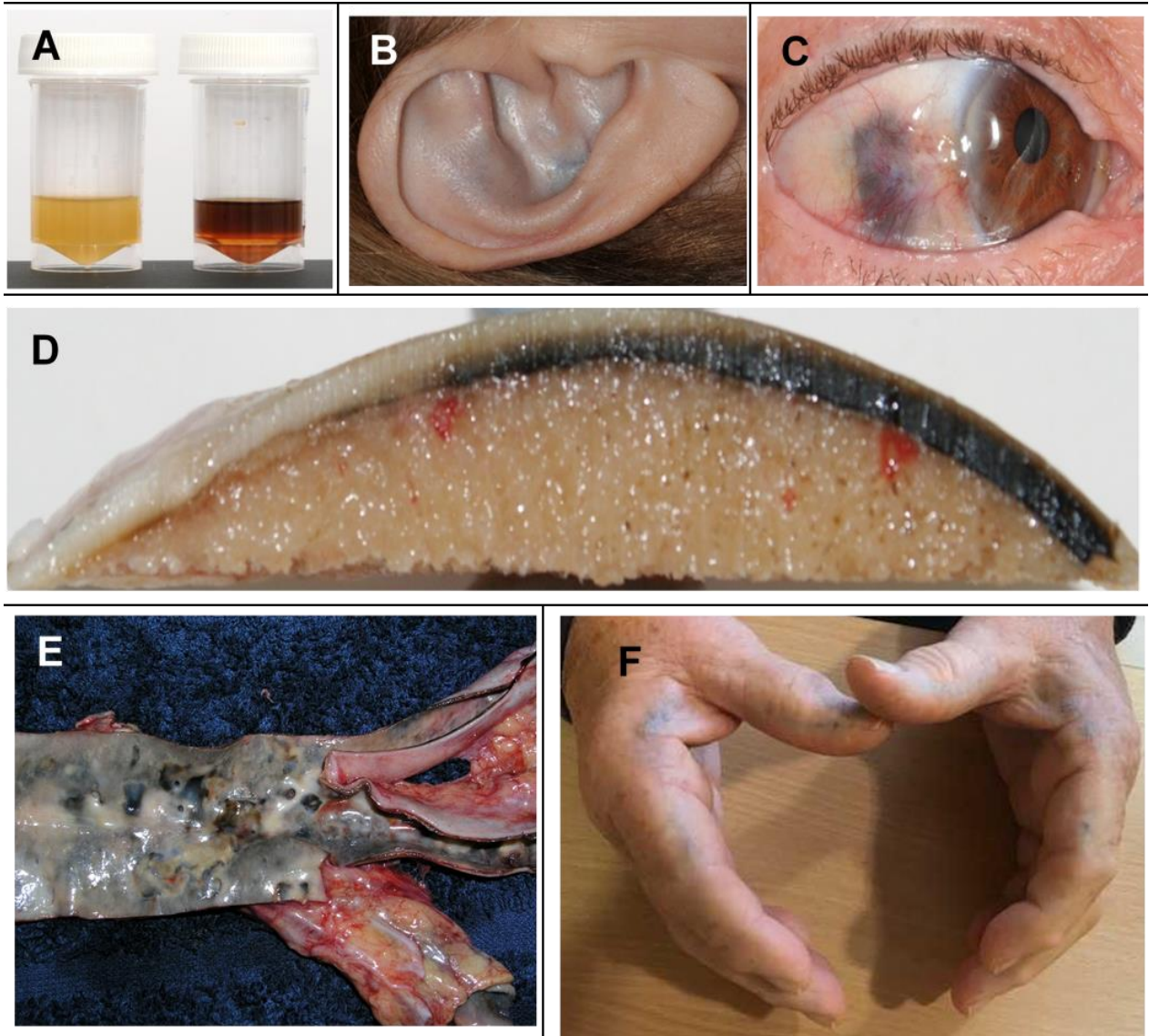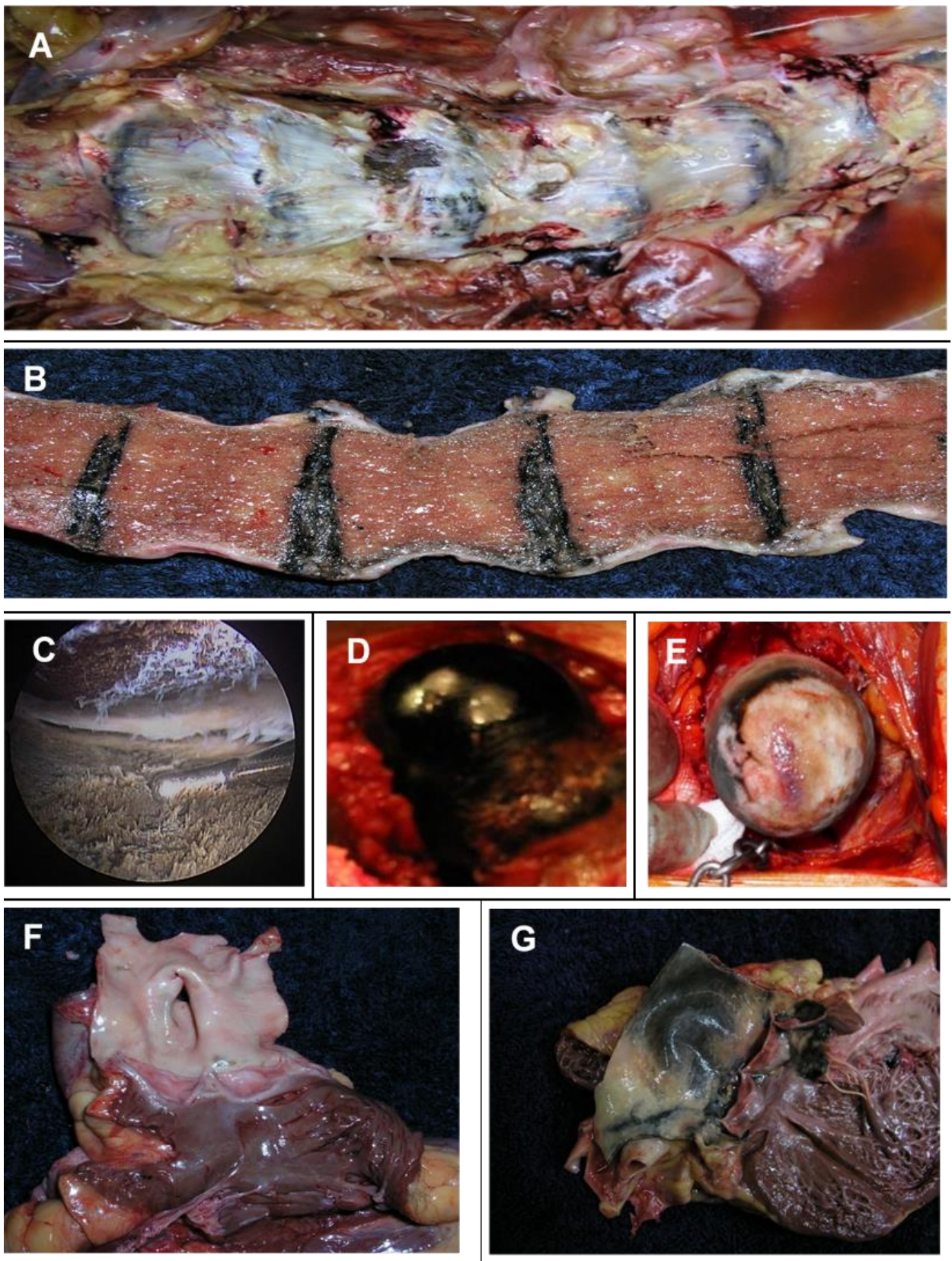
Figure 1

Figure 2

Figure 3
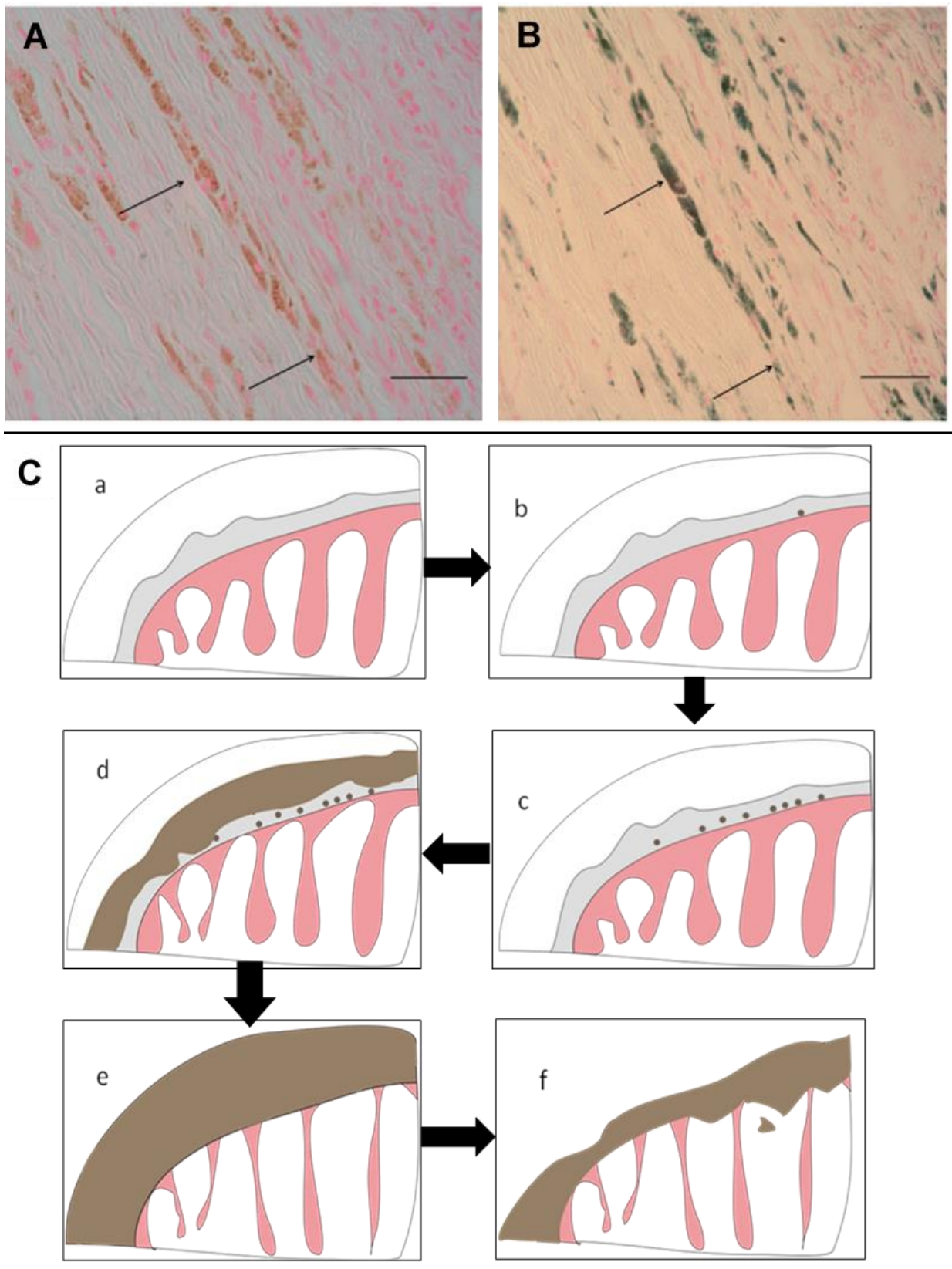
Figure 4

Figure 5

[Author accepted manuscript]

Title: A comprehensive LC-QTOF-MS metabolic phenotyping strategy: application to alkaptonuria

Running Head: Comprehensive metabolic phenotyping by LC-QTOF-MS

Authors: Brendan P Norman[1,†,*], Andrew S Davison[1,2,†], Gordon A Ross[3], Anna M Milan[1,2], Andrew T Hughes[1,2], Hazel Sutherland [1,4], Jonathan C Jarvis[4], Norman B Roberts[1], James A Gallagher[1,‡], Lakshminarayan R Ranganath[1,2,‡]

[1]Musculoskeletal Biology I, Institute of Ageing & Chronic Disease, William Henry Duncan Building, University of Liverpool, Liverpool, UK; [2]Department of Clinical Biochemistry and Metabolic Medicine, Liverpool Clinical Laboratories, Royal Liverpool University Hospitals Trust, Liverpool, UK; [3]Agilent Technologies UK Ltd., Cheadle, UK; [4]School of Exercise Science, Liverpool John Moores University, Liverpool, UK

*Corresponding Author: Brendan P Norman, Musculoskeletal Biology I, Institute of Ageing & Chronic Disease, William Duncan Building, University of Liverpool, Liverpool, UK, L7 8TX; Telephone: (0)151 794 9063; Email: bnorman@liverpool.ac.uk

Key words: alkaptonuria, LC-QTOF-MS, metabolomics, metabolite identification, accurate mass, retention time

Number of tables and figures: 5 figures (3 supplemental figures), 1 table

Abbreviations: LC – liquid chromatography; QTOF-MS – quadrupole time-of-flight mass spectrometry; AMRT - accurate mass/retention time; MS – mass spectrometry; AKU – alkaptonuria; HGD – homogentisate 1,2-dioxygenase; HGA – homogentisic acid; NAC – National Alkaptonuria Centre; RT – retention time; QC - quality control; CV – coefficient of variation; FC – fold change; HT-1 - hereditary tyrosinaemia type-1

List of human genes: *HGD* - homogentisate 1,2-dioxygenase

---

† Joint first author

‡ JAG & LRR contributed equally to this work

Abstract

Background: Identification of unknown chemical entities is a major challenge in metabolomics. To address this challenge, we developed a comprehensive targeted profiling strategy, combining three complementary liquid chromatography (LC) quadrupole time-of-flight mass spectrometry (QTOF-MS) techniques and in-house accurate mass retention time (AMRT) databases established from commercial standards. This strategy was used to evaluate the effect of nitisinone on the urinary metabolome of patients and mice with alkaptonuria (AKU). Because hypertyrosinemia is a known consequence of nitisinone therapy, we investigated the wider metabolic consequences beyond hypertyrosinemia.

Methods: 619 standards (molecular weight 45-1354 Da) covering a range of primary metabolic pathways were analyzed using three LC methods, two reversed phase and one normal phase, coupled to QTOF-MS. Separate AMRT databases were generated for the three methods, comprising chemical name, formula, theoretical accurate mass and measured retention time. Databases were used to identify chemical entities acquired from non-targeted analysis of AKU urine; match window theoretical accurate mass ±10 ppm and retention time ±0.3 min.

Results: Application of the AMRT databases to data acquired from analysis of urine from 25 patients with AKU (pre-treatment and after 3, 12 and 24 mo on nitisinone) and 18 $HGD^{-/-}$ mice (pre-treatment and after one week on nitisinone) revealed 31 previously unreported statistically significant changes in metabolite patterns and abundance, indicating alterations to tyrosine, tryptophan and purine metabolism post-nitisinone.

Conclusions: The comprehensive targeted profiling strategy described here has the potential of enabling discovery of novel pathways associated with pathogenesis and treatment of AKU.

Metabolic profiling has potential to advance knowledge of disease beyond established biochemical pathways and will play a major role in precision medicine (1-5). However, metabolite identification is a major challenge in untargeted profiling studies using mass spectrometry (MS) (6,7) since they generate many chemical signals representing 'unknowns'. To address this challenge, we generated three accurate mass/retention time (AMRT) databases from 619 metabolite standards using liquid chromatography-quadrupole time-of-flight-mass spectrometry (LC-QTOF-MS). We applied this strategy to the inborn error of metabolism alkaptonuria (AKU, OMIM 203500). In AKU, bi-allelic mutations in the homogentisate 1,2-dioxygenase (*HGD*) gene result in a lack of homogentisate 1,2-dioxygenase (HGD, E.C.1.12.11.5) (8), leading to increased homogentisic acid (HGA), a metabolite of tyrosine catabolism (Fig. 1). HGA accumulates in connective tissue, particularly cartilage, where it is deposited as a dark pigment, a process termed ochronosis. Ochronosis underlies a range of clinical features in AKU, the most debilitating of which is severe, early-onset osteoarthropathy that alters the physico-mechanical properties of cartilage (9,10).

Nitisinone has emerged as a promising therapeutic agent in AKU because it reduces circulating HGA concentrations (11–15) by reversibly inhibiting hydroxyphenylpyruvic acid dioxygenase (E.C. 1.13.11.27). Nitisinone completely inhibits ochronosis in an *HGD*[-/-] mouse model of AKU (16). Although not currently licensed for AKU, nitisinone is being used to treat patients attending the National Alkaptonuria Centre (NAC) in the UK. One of the major metabolic consequences of nitisinone is hypertyrosinaemia (11,13–15,17–21) (Fig. 1). The objective of our study was to apply the developed profiling strategy to urine from patients with AKU and HGD[-/-] mice to understand the wider metabolic consequences of nitisinone treatment.

Materials and Methods

Metabolite standards library preparation

619 standards (IROA Technology MS metabolite library of standards; molecular weight 45-1354 Da) were from Sigma-Aldrich (UK) in seven deepwell plates. Each compound was present at 5 μg/well (>95% purity). Plates were stored at -80°C. The standards covered a broad range of primary and intermediary metabolism, including

the following compound classes: carboxylic acids, amino acids, biogenic amines, polyamines, nucleotides, coenzymes, vitamins, carbohydrates, fatty acids, lipids, steroids and hormones (Supplemental Files 1 and 2).

Prior to analysis, plates were thawed at ambient temperature (18°C) and compounds reconstituted in 15 µL methanol (LC/MS grade, Sigma-Aldrich) followed by 285 µL deionized water (DIRECT-Q 3UV Millipore water purification system). Plates 1-5 were left to stand for 1 h following addition of water and methanol. Plates 6 and 7 (primarily non-polar compounds) were left to stand for 2 h at room temperature following addition of methanol to ensure solubilization; water was then added. Plates were agitated on a plate shaker (MTS 2/4m IKA, Germany) at 600 rpm for 10 min. Compounds were pooled across rows of each plate (12 wells/row). 20 µL was removed from each well and pooled into one well (12 compounds analyzed per injection). Fifty-six injections were performed in positive and negative polarity across all plates (112 injections for each LC-QTOF-MS method; Supplemental Fig. 1). The total volume of each pool was 240 µL with each compound at 14 mg/dL (1-31 µmol/L).

Evaluation of LC-QTOF-MS strategy and effect of nitisinone on the metabolome using AKU urine

24 h urine was obtained from 25 patients with AKU attending the NAC (12 male, mean(±SD) age 51±15 y). Samples were collected before nitisinone then at 3 (2 mg every other day), 12 and 24 (2 mg daily) mo on nitisinone and stored at -80°C. Samples were diluted 1:3 with deionized water and stored at -80°C as three separate aliquots for analysis by methods 1-3. Metabolomic analysis was part of the diagnostic service to patients being seen at the NAC and with approval from the Royal Liverpool and Broadgreen University Hospital Trusts Audit Committee (audit no. ACO3836).

Mouse urine was from 18 (9 male, mean age 27±12 wk) $HGD^{-/-}$ BALB/c mice (16) bred from laboratory stocks at the University of Liverpool. Samples were obtained on a single-collection basis before treatment then after 1 wk on nitisinone, administered in all drinking water (0.4 mg/dL), supplied *ad libitum*. Mouse urine was collected onto cling film, pipetted into sample tubes and stored at -80°C. Analysis was performed

following dilution of 1:9 with deionized water. Mouse breeding and dosing was authorized under the Animals (Scientific Procedures) Act, UK.

Samples were pooled for quality assurance in profiling experiments. For each sampling time point, a representative pool was created by pooling 20 µL of each individual urine sample. An additional overall pool was created separately for human and mouse experiments by pooling equal proportions of the pooled urine groups detailed above. Pooled human and mouse samples were stored and treated as above prior to analysis.

Analytical conditions

Chemicals

Water for mobiles phases was purified as above. Methanol, acetonitrile, isopropanol (Sigma-Aldrich), formic acid (Biosolve, Netherlands) and ammonium formate (Fisher Scientific, Germany) were LC/MS grade.

Equipment

Sample analysis was performed on an Agilent 1290 Infinity LC coupled to an Agilent 6550 QTOF-MS equipped with a dual AJS electrospray ionization source (Agilent, UK).

Chromatographic conditions

Three chromatographic methods were designed to separate different compound classes. Method 1: non-polar compounds. Method 2: a range of polar/non-polar compounds. Method 3: polar compounds. Metabolite standards were analyzed using all three LC-QTOF-MS techniques.

Method 1

A Zorbax Eclipse Plus C_{18} column (2.1x100 mm, 1.8 µm, Agilent) was maintained at 60°C (flow rate 0.4 mL/min). Mobile phases were (A) water and (B) methanol both containing 31.6 mg/dL ammonium formate and 0.1% formic acid. The elution gradient

started at 5% B at 0-1 min increasing linearly to 100% by 12 min, held at 100% B until 14 min, returning to 95% A for 5 min.

Method 2

Method 2 employed the same conditions and elution gradient as method 1, but with an Atlantis dC$_{18}$ column (3x100 mm, 3 µm, Waters, UK).

Method 3

A BEH amide column (3x150 mm, 1.7 µm, Waters) was maintained at 40°C (flow rate 0.6 mL/min). Mobile phases were (A) water and (B) acetonitrile (both containing 0.1% formic acid). The elution gradient started at 99% B decreasing linearly to 30% from 1-12 min, held at 30% B until 12.6 min, returning to 99% B for 3.4 min.

Sample injection volume was 1 µL for metabolite standards and human urine, and 2 µL for mouse urine. The autosampler was maintained at 4°C and the needle was washed with a solution of water:methanol:isopropanol (45:45:10 v/v) between injections.

Design of urine metabolic profiling experiments

Human and mouse urine was analyzed separately. Human samples were analyzed batch-wise using all 3 methods, negative followed by positive polarity. Mouse samples were analyzed by method 2 only due to limited sample volumes, in one batch comprising both polarities.

The analytical sequence of each profiling batch was designed according to published guidance (22). Each run commenced with 20 replicate injections of the overall pooled sample to condition the system. The order of individual samples was randomized computationally. Pooled samples were interspersed throughout the analytical sequence, every 10th injection. Injections of each sample group pool and the overall pooled sample were also placed at the start (post-conditioning) and end of each analytical sequence.

Data acquisition and processing

Data acquisition and processing were performed using the MassHunter suite (Agilent). Data were acquired with Data Acquisition (build 06.00). Metabolite standards and urine samples were analyzed in both polarities, mass range 50-1700 (Supplemental File 3). Quality checks and processing of raw data were performed with Qualitative Analysis (build 07.00). Extracted ion chromatograms of reference masses were performed to check mass accuracy remained <5 ppm throughout the run. Binary pump pressure curves for injections across each analytical sequence were overlaid to check chromatographic reproducibility.

Compound signals were extracted from the standards data in Qualitative Analysis (build 07.00) by molecular formula using the 'find by formula' algorithm; mass window of theoretical accurate mass (calculated from molecular formula) ±5 ppm. Allowed ion species were: H^+ and Na^+ and additionally NH_4^+ for methods 1 and 2 (positive polarity); and H^- and additionally CHO_2^- for methods 1 and 2 (negative polarity). Charge state range was 1-2. Dimers were allowed. Separate AMRT databases were created for each method using PCDL Manager (build 08.00). Theoretical accurate mass, retention time (RT), molecular formula and METLIN chemical name were entered into each AMRT database for compounds that were retained and detected (databases publicly available via https://figshare.com/collections/_/4378235/0 (23)).

Urine profiling data were mined for signals matching AMRT database compounds using 'targeted feature extraction' (Profinder, build 08.00). Targeted feature extraction uses the molecular formulae to extract and group spectral signals (*i.e.* adducts, multimers and isotopes), corresponding to individual database compounds. Feature extraction window; theoretical accurate mass ±10 ppm and database RT ±0.3 min. Allowed ion species were the same as specified above.

Detection of non-AMRT database compounds in urine

Urine data were also mined using Profinder for compounds not from generated AMRT databases but of interest for their predicted role in AKU/nitisinone metabolism. These compounds were associated with a) increased tyrosine: acetyl-L-tyrosine, γ-glutamyl-L-tyrosine and tyramine-sulfate; b) ochronotic pigment: 2,5-dihydroxybenzaldehyde,

hipposudoric acid and norhipposudoric acid; or c) nitisinone metabolism: hydroxy-nitisinone, nitisinone and 2-nitro-4-trifluoromethylbenzoic acid. For non-AMRT compounds, the same RT (range <0.3 min) was required across samples. Compound identifications were based on theoretical accurate mass ±10 ppm.

Data quality control and statistical analyses

Several quality control (QC) filters were applied to AMRT-matched entities from each profiling batch. First, entities were retained if observed in at least two samples per experimental group (*i.e.* sampling time point). Data files were then exported from Profinder and imported into Mass Profiler Professional (build 14.5). 24 h creatinine excretions (Roche Diagnostics, Germany) were used as external scalar for individual human urine samples (average values across each sampling time point used for pooled samples). Creatinine values were not available for mouse samples; the signal identified as creatinine by matched AMRT was used as external scalar for each sample using the peak area of the ^{13}C $[M+H]^+$ ion (^{13}C because ^{12}C creatinine signal was saturated) calculated in MassHunter Quantitative analysis (build 06.00).

Entities were then further filtered in Mass Profiler Professional based on data from pooled samples from each experiment. Entities were retained if observed in 100% of replicate injections for at least one sample group pool, and with peak area coefficient of variation (CV) <25% across replicate injections of all sample group pools.

Statistical analyses were performed in Mass Profiler Professional based on peak area. Human urine profiles were compared at 3, 12 and 24 mo on nitisinone with pre-nitisinone by one-way repeated-measures ANOVA. Mouse urine profiles were compared pre-nitisinone vs 1 wk on nitisinone by paired t-test. Benjamini-Hochberg false-discovery rate adjustment was used in all statistical significance testing. Fold changes (FC's) were calculated based on raw peak area. Principal component analyses employing four-component models were also performed on each filtered dataset.

Results

Analysis of metabolite standards

Signals representing the chemical formulae for standards in each injection were extracted from the data by theoretical accurate mass ±5 ppm. Only standards with RT >0.3 min after the column void volume were considered. 519/619 (83.8%) of the standards were retained sufficiently to be detected by at least one method. 116/619 (18.7%) and 226/619 (36.5%) of the compounds were retained sufficiently and detected by one or two methods only, respectively, demonstrating the utility of combining data from the three chromatographic methods (Figure 2A). RT, accurate mass and charge state were entered into an AMRT metabolite database for each method for matching unknowns against.

Figure 2B-D shows the mass/RT distribution for compounds detected by each method. Figure 2B-D also highlights differences in selectivity for compounds from three example compound groups with different chemical properties: carbohydrates, amino acids and lipids (Supplemental File 4). Method 3 retained and separated highly polar compounds such as carbohydrates (20/26 detected, RT range 5-9.5 min). In methods 1 and 2, the same carbohydrates were weakly retained; for method 1 all carbohydrates detected eluted within the initial column void (24/26 detected, RT range 0.58-0.68 min) and for method 2 all carbohydrates detected eluted close to the initial void (23/26 detected, RT range 1.3-1.4 min). Method 1 showed some evidence of improved suitability for analysis of lipid-like compounds (11/11 detected, RT range 7.7-13.7 min) compared to methods 2 (8/11 detected, RT range 9.3-14.3 min) and 3 (7/11 detected, RT range 1.5-8.3 min), which detected fewer of these compounds within a narrower RT range overall. All three methods enabled detection of the amino acids, with method 3 appearing to show the most useful chromatographic resolution and retention (19/21 detected, RT range 5.4-8.1 min) compared to methods 1 (21/21 detected, RT range 0.6-2.8 min) and 2 (21/21 detected, RT range 1.0-4.6 min).

Identification of metabolites in pooled urine *by AMRT*

Unknown chemical entities were matched against the respective AMRT database generated for each method: accurate mass ±10 ppm, RT ±0.3 min. For data presented here, analysis comprised two replicate injections of the overall pooled urine from patients with AKU from the start and end of each analytical run. Only unknowns obtained from both injections in positive or negative polarity and with single AMRT compound matches were retained in this analysis. Additional QC filters were then applied to ensure reproducibility across the run: CV <25% for peak area and RT between the two replicate injections. Table 1 summarizes matches retained for each method. No compounds were filtered out due to RT CV >25%. The maximum RT CV% change between the replicate injections across all analytical runs (each comprised 205 injections) was 4.1%.

Figure 3A shows that 203 unique compound matches were obtained from urine. Supplemental File 5 shows the matches obtained by methods 1-3. Fourteen AMRT matches were obtained by all three methods. An additional 61 AMRT matches were obtained by two methods, and 128 matches by only one of the methods, further supporting the increased coverage from the three methods in combination. Figure 3B-D shows the mass/RT distribution of the AMRT matches.

Table 1 and Supplemental Figure 2 show the number of AMRT matches obtained with narrower AMRT-matching windows. For each method, >50% of matches obtained with ±10 ppm and ±0.3 min post-QC remained with ±5 ppm and ±0.15 min.

Application of strategy to AKU: effect of nitisinone therapy on the urine metabolome

Figure 4A,B shows the overall study design and a representative example of AMRT compound signals extracted from the data respectively. Representative principal components analysis plots (Figure 4C) show clear separation between the AMRT-matched profiles of urine pre- vs post-nitisinone for human and mouse, showing that our strategy captured key metabolite changes.

Thirty-five metabolites showed statistically significant changes (p<0.05, FC>2) in abundance after 3, 12 or 24 mo on nitisinone in humans or 1 wk on nitisinone in mice

(Figure 5). An FC >2 was chosen in order to limit false discovery and focus on clear changes. In patients with AKU, 13 metabolites increased and 14 decreased. In

$HGD^{-/-}$ mice, 12 metabolites increased and 6 decreased. Ten metabolites changed in the same direction in humans and mice (Figure 5). Together, the majority of these metabolite changes could be categorized into three main metabolic pathways: those of tyrosine, tryptophan and purine.

Interestingly, mouse data also showed clear separation by gender (principal component 1; x-axis Figure 4C). Histamine was the primary driver of this separation in positive polarity (principal component 1 loading; 0.42), which captured the effect of gender and explained 43% of the variation in the dataset. Histamine was significantly increased in female mice ($p<0.0001$, FC=16).

Alterations to non-AMRT database compounds post-nitisinone

Data showed alteration ($p<0.05$, FC>2) to metabolites with a predicted association to AKU/nitisinone metabolism that were not from AMRT databases (Supplemental Figure 3). 2,5-dihydroxybenzaldehyde, probably associated with ochronotic pigment, was decreased post-nitisinone. The tyrosine metabolites acetyl-L-tyrosine and γ-glutamyl-L-tyrosine and the nitisinone metabolite hydroxy-nitisinone were increased. These changes were observed in human and mouse urine.

Discussion

We have developed a strategy for comprehensive LC-QTOF-MS profiling with compound identification by three AMRT databases generated from metabolite standards. Application of this strategy enabled: a) identification of unknown chemical entities in complex biological matrix by AMRT; and b) identification of previously unreported changes to urinary metabolites and metabolic pathways following nitisinone treatment in AKU.

A limitation of LC-MS compound identification by public databases (24–28) is that data were acquired using different analytical techniques and parameters. The Metabolomics Standards Initiative has established levels of metabolite identification confidence (29,30). Only identification strategies, as reported here, that compared two or more orthogonal chemical properties (*e.g.* AMRT) with an authentic standard under identical analytical conditions achieved the highest identification confidence level. Identifications by public databases in which data were acquired under different analytical conditions can only achieve the second level of confidence ('putative identification'), even with two matched orthogonal chemical properties.

Combining data from three LC techniques enhanced the number of unique urine AMRT matches obtained, improving coverage of the metabolome. Methods 1 and 2 showed some similarity in analyte retention (Figure 2B-D). Method 2 was used because the Atlantis dC_{18} column was previously shown to chromatographically resolve metabolites of the tyrosine pathway in patients with AKU (31-32). It also provided greater overall retention of polar metabolites compared to a standard C_{18} column due to endcapping of bi-functionally bonded C_{18} stationary phase. Method 1 was included here to improve peak capacity owing to the increased theoretical plates provided by the smaller 2.1 µm column particle size. Moreover, methods 1 and 2 yielded 39 and 57 AMRT matches respectively that were exclusive to each method (Figure 3A).

The congenital defect that causes AKU directly affects tyrosine catabolism. However, AKU is multi-systemic (9) and the wider metabolic consequences of the disease and nitisinone treatment have not been systematically studied. The strategy was applied to AKU, but given the range of primary and intermediary metabolism covered by the AMRT compounds, it could be applied to study metabolism in any disease. The data show that nitisinone alters tyrosine and tryptophan metabolism, and support alteration to the purine metabolic pathway. The changes observed comprise increased and decreased concentrations within the same pathways, suggesting that nitisinone has complex, wide-ranging effects on metabolism. Nitisinone is licensed to treat patients with hereditary tyrosinaemia type-1 (HT-1, OMIM 276700), another congenital disease of tyrosine metabolism, and is a promising HGA-reducing agent in AKU. In nitisinone-treated HT-1 and AKU it is established that circulating tyrosine increases markedly

(11,13–15,17–21). Tyrosine is the precursor for catecholamines, thyroid hormones and melanin (Figure 1), suggesting that increased tyrosine substrate has potential to dysregulate these pathways. In HT-1 there are concerns that hypertyrosinaemia may contribute to neurodevelopmental delay (19,33-36).

Application of our profiling strategy identified a number of previously unreported metabolite changes post-nitisinone. Increased 3-(4-hydroxyphenyl)lactic acid is probably directly related to elevated tyrosine and/or its clearance in urine. Changes a) homovanillic acid and the trace amine tyramine and b) L-tryptophan represent further alterations to dopamine and tryptophan-serotonin metabolism (respectively) than those previously reported. L-tryptophan is the precursor for serotonin, and a post-nitisinone increase in the serotonin metabolite 5-hydroxyindoleacetic acid was previously reported in HT-1 cerebrospinal fluid (37) and AKU urine (20), although not observed here. Xanthurenic acid and L-kynurenine were increased and decreased respectively, indicating for the first time that the kynurenine pathway, originating from tryptophan, is altered by nitisinone. Also, indoxyl sulfate, a metabolite of the indolepyruvate pathway from tryptophan, was decreased. Alterations to indolepyruvate metabolism are previously reported in nitisinone-treated AKU plasma; carboxaldehyde, indole-lactate and indole-pyruvate increased, and *in vitro* studies indicated this as a direct consequence of increased 3-(4-hydroxyphenyl)pyruvic acid (38). This is the first report of alterations to purine metabolism following treatment with nitisinone. Decreased 3,5-cyclic-AMP and xanthosine were unexpected but replicated in mice.

The remaining metabolite changes have been previously reported following nitisinone treatment; decreased HGA and increased tyrosine, acetyl-L-tyrosine, γ-glutamyl-tyrosine, 3-(4-hydroxyphenyl)pyruvic acid and 3-methoxytyramine. Decreased HGA and increased tyrosine are well-known consequences of nitisinone, and observation of these changes in each analytical run (tyrosine in positive and HGA in negative polarity) supports the analysis, data extraction workflow, and validity of the data. Increased acetyl-L-tyrosine and γ-glutamyl-tyrosine have been previously reported in nitisinone-treated AKU and were proportional to tyrosine elevation (39). Increased 3-methoxytyramine was previously reported in AKU urine (20); it is a metabolite of

447

dopamine metabolism and derived from tyrosine via decarboxylation of dihydroxyphenylalanine to dopamine.

Ten AMRT-matched metabolite changes were observed in human and mouse: 4-hydroxyphenylacetic acid, 3,(4-hydroxyphenyl)lactic acid, 3-(4-hydroxyphenyl)pyruvic acid, L-tyrosine, phenylacetic acid and tyramine increased post-nitisinone; 2-hydroxy-4-(methylthio)butyric acid, 3,5-cyclic AMP, HGA and xanthosine decreased post-nitisinone. The concordance between human and mouse data supports the approach and validity of the data. For mice, it was possible to control potentially confounding factors that could affect metabolism, such as diet and genetic diversity. This increases likelihood that observed changes are attributable to nitisinone. Reduced phenotypic variation could explain the prominent gender difference observed for mouse only; histamine was particularly elevated in urine from females, as previously reported in the literature for rats (40,41) but not mice to our knowledge.

The limitations of this study are as follows. First is the relatively small sample size for a clinical metabolomics study (25 patients). However, AKU is a rare disease and the repeated-measures design enhanced statistical power. Second, data represent changes observed in urine not serum, which more closely reflects internal homeostasis. The urine metabolome is a composite of products from endogenous metabolism, diet, drugs and the gut microbiome. Urine does however give a valuable indication of the metabolic fate of the increased circulating tyrosine post-nitisinone. Further metabolomic analyses on fluids and tissues from other biological locations are required to achieve more detailed, compartment-specific data, for example cerebrospinal fluid to directly study the impact of nitisinone-induced tyrosinaemia on the central nervous system. Third, three chromatographic methods were used, however this may not always be feasible. For large-scale studies, it may be pragmatic to use two of the methods to reduce analytical and processing time.

In conclusion, we have developed a targeted LC-QTOF-MS strategy for comprehensive coverage of the metabolome with compound structure identification using three AMRT databases, which are publicly available. Application of the approach to AKU has advanced our knowledge of the wider metabolic consequences of nitisinone, demonstrating the potential of our method as a metabolic phenotyping strategy more generally.

Acknowledgements

References

1. Holmes E, Wilson ID, Nicholson JK. Metabolic phenotyping in health and disease. Cell 2008;134:714–7.

2. Baker M. Metabolomics: from small molecules to big ideas. Nat Methods 2011;8:117–21.

3. Fiehn O. Metabolomics -- the link between genotypes and phenotypes. Plant Mol Biol 2002;48:155-71.

4. Beger RD, Dunn W, Schmidt MA, Gross SS, Kirwan JA, Cascante M, et al. Metabolomics enables precision medicine: "A White Paper, Community Perspective." Metabolomics 2016;12:149.

5. Nicholson JK, Holmes E, Kinross JM, Darzi AW, Takats Z, Lindon JC. Metabolic phenotyping in clinical and surgical environments. Nature 2012;491:384–92.

6. Dunn WB, Broadhurst DI, Atherton HJ, Goodacre R, Griffin JL. Systems level studies of mammalian metabolomes: the roles of mass spectrometry and nuclear magnetic resonance spectroscopy. Chem Soc Rev 2011;40:387–426.

7. Wishart DS. Advances in metabolite identification. Bioanalysis 2011;3:1769–82.

8. Zatkova A. An update on molecular genetics of Alkaptonuria (AKU). J Inherit Metab Dis 2011;34:1127–36.

9. Ranganath LR, Jarvis JC, Gallagher JA. Recent advances in management of alkaptonuria (invited review; best practice article). J Clin Pathol 2013;66:367–73.

10. Taylor AM, Boyde A, Wilson PJM, Jarvis JC, Davidson JS, Hunt JA, et al. The role of calcified cartilage and subchondral bone in the initiation and progression of ochronotic arthropathy in alkaptonuria. Arthritis Rheum 2011;63:3887–96.

11. Introne WJ, Perry MB, Troendle J, Tsilou E, Kayser MA, Suwannarat P, et al. A 3-year randomized therapeutic trial of nitisinone in alkaptonuria. Mol Genet Metab 2011;103:307–14.

12. Phornphutkul C, Introne WJ, Perry MB, Bernardini I, Murphey MD, Fitzpatrick DL, et al. Natural history of alkaptonuria. N Engl J Med 2002;347:2111–21.

13. Suwannarat P, O'Brien K, Perry MB, Sebring N, Bernardini I, Kaiser-Kupfer MI, et al. Use of nitisinone in patients with alkaptonuria. Metabolism 2005;54:719–28.

14. Milan AM, Hughes AT, Davison AS, Devine J, Usher J, Curtis S, et al. The effect of nitisinone on homogentisic acid and tyrosine: a two-year survey of patients attending the National Alkaptonuria Centre, Liverpool. Ann Clin Biochem 2017;54:323–30.

15. Ranganath LR, Milan AM, Hughes AT, Dutton JJ, Fitzgerald R, Briggs MC, et al. Suitability Of Nitisinone In Alkaptonuria 1 (SONIA 1): an international, multicentre, randomised, open-label, no-treatment controlled, parallel-group, dose-response study to investigate the effect of once daily nitisinone on 24-h urinary homogentisic acid excretion in patients with alkaptonuria after 4 weeks of treatment. Ann Rheum Dis 2016;75:362–7.

16. Preston AJ, Keenan CM, Sutherland H, Wilson PJ, Wlodarski B, Taylor AM, et al. Ochronotic osteoarthropathy in a mouse model of alkaptonuria, and its inhibition by nitisinone. Ann Rheum Dis 2014;73:284–9.

17. Lindstedt S. Treatment of hereditary tyrosinaemia type I by inhibition of 4-hydroxyphenylpyruvate dioxygenase. Lancet 1992;340:813–7.

18. Olsson B, Cox TF, Psarelli EE, Szamosi J, Hughes AT, Milan AM, et al. Relationship Between Serum Concentrations of Nitisinone and Its Effect on Homogentisic Acid and Tyrosine in Patients with Alkaptonuria. JIMD Rep 2015;24:21–7.

19. McKiernan PJ, Preece MA, Chakrapani A. Outcome of children with hereditary tyrosinaemia following newborn screening. Arch Dis Child 2015;100:738–41.

20. Davison AS, Norman B, Milan AM, Hughes AT, Khedr M, Rovensky J, et al. Assessment of the effect of once daily nitisinone therapy on 24-h urinary metadrenalines and 5-hydroxyindole acetic acid excretion in patients with alkaptonuria after 4 weeks of treatment. JIMD Rep [Internet] 2017; Available from: http://dx.doi.org/10.1007/8904_2017_72

21. Davison AS, Harrold JA, Hughes G, Norman BP, Devine J, Usher J, et al. Clinical and biochemical assessment of depressive symptoms in patients with Alkaptonuria before and after two years of treatment with nitisinone. Mol Genet Metab 2018;125:135–43.

22. Vorkas PA, Isaac G, Anwar MA, Davies AH, Want EJ, Nicholson JK, et al. Untargeted UPLC-MS profiling pipeline to expand tissue metabolome coverage: application to cardiovascular disease. Anal Chem 2015;87:4184–93.

23. Norman BP, Davison AS, Ross GA, Milan AM, Hughes AT, Sutherland H, et al. Three accurate mass retention time (AMRT) databases generated from IROA Technology Metabolite Library of Standards by LC-QTOF-MS analysis [Internet]. figshare; 2019 [cited 2019Jan25].. Available from: https://figshare.com/collections/_/4378235/0

24. Smith CA, Maille GO, Want EJ, Qin C, Trauger SA, Brandon TR, et al. METLIN. Ther Drug Monit 2005;27:747–51.

25. Wishart DS, Feunang YD, Marcu A, Guo AC, Liang K, Vázquez-Fresno R, et al. HMDB 4.0: the human metabolome database for 2018. Nucleic Acids Res 2018;46:D608–17.

26. Milman BL, Zhurkovich IK. Mass spectral libraries: A statistical review of the visible use. Trends Analyt Chem 2016;80:636–40.

27. Horai H, Arita M, Kanaya S, Nihei Y, Ikeda T, Suwa K, et al. MassBank: a public repository for sharing mass spectral data for life sciences. J Mass Spectrom 2010;45:703–14.

28. Kanehisa M, Goto S. KEGG: kyoto encyclopedia of genes and genomes. Nucleic Acids Res 2000;28:27–30.

29. Sumner LW, Amberg A, Barrett D, Beale MH, Beger R, Daykin CA, et al. Proposed minimum reporting standards for chemical analysis Chemical Analysis Working Group (CAWG) Metabolomics Standards Initiative (MSI). Metabolomics 2007;3:211–21.

30. Dunn WB, Erban A, Weber RJM, Creek DJ, Brown M, Breitling R, et al. Mass appeal: metabolite identification in mass spectrometry-focused untargeted metabolomics. Metabolomics 2012;9:44–66.

31. Hughes AT, Milan AM, Davison AS, Christensen P, Ross G, Gallagher JA, et al. Serum markers in alkaptonuria: simultaneous analysis of homogentisic acid, tyrosine and nitisinone by liquid chromatography tandem mass spectrometry. Ann Clin Biochem 2015;52:597–605.

32. Hughes AT, Milan AM, Christensen P, Ross G, Davison AS, Gallagher JA, et al. Urine homogentisic acid and tyrosine: simultaneous analysis by liquid chromatography tandem mass spectrometry. J Chromatogr B Analyt Technol Biomed Life Sci 2014;963:106–12.

33. Masurel-Paulet A, Poggi-Bach J, Rolland M-O, Bernard O, Guffon N, Dobbelaere D, et al. NTBC treatment in tyrosinaemia type I: long-term outcome in French patients. J Inherit Metab Dis 2008;31:81–7.

34. De Laet C, Munoz VT, Jaeken J, François B, Carton D, Sokal EM, et al. Neuropsychological outcome of NTBC-treated patients with tyrosinaemia type 1. Dev Med Child Neurol 2011;53:962–4.

35. Thimm E, Richter-Werkle R, Kamp G, Molke B, Herebian D, Klee D, et al. Neurocognitive outcome in patients with hypertyrosinemia type I after long-term treatment with NTBC. J Inherit Metab Dis 2012;35:263–8.

36. Bendadi F, de Koning TJ, Visser G, Prinsen HCMT, de Sain MGM, Verhoeven-Duif N, et al. Impaired cognitive functioning in patients with tyrosinemia type I receiving nitisinone. J Pediatr 2014;164:398–401.

37. Thimm E, Herebian D, Assmann B, Klee D, Mayatepek E, Spiekerkoetter U. Increase of CSF tyrosine and impaired serotonin turnover in tyrosinemia type I. Mol Genet Metab 2011;102:122–5.

38. Gertsman I, Gangoiti JA, Nyhan WL, Barshop BA. Perturbations of tyrosine metabolism promote the indolepyruvate pathway via tryptophan in host and microbiome. Mol Genet Metab 2015;114:431–7.

39. Gertsman I, Barshop BA, Panyard-Davis J, Gangoiti JA, Nyhan WL. Metabolic Effects of Increasing Doses of Nitisinone in the Treatment of Alkaptonuria. JIMD Rep 2015;24:13–20.

40. Kim KS. Sex difference in histamine metabolism in rats. American Journal of Physiology-Legacy Content 1959;197:1258–60.

41. Netter KJ, Cohn VH Jr, Shore PA. Sex difference in histamine metabolism in the rat. Am J Physiol 1961;201:224–6.

Table 1. Number of accurate mass retention time (AMRT) matches obtained by the three methods and retained following quality control filtering and decreasing AMRT window size during feature extraction.

Feature extraction window	Filtering step	Number of urine AMRT matches					
		Method 1		Method 2		Method 3	
		Positive	Negative	Positive	Negative	Positive	Negative
Accurate mass 10 ppm, RT 0.3 min	None	274	153	248	105	111	74
	(1) Unique AMRT's only, manual curation	80	38	121	43	70	37
	(2) Abundance QC: peak area CV <25% between replicates	75	36	107	35	59	30
	(3) RT shift QC: RT CV <25% between replicates	75	36	107	35	59	30
Accurate mass 5 ppm, RT 0.15 min	Steps (1) - (3)	44	31	65	25	38	22
Accurate mass 2.5 ppm, RT 0.075 min	Steps (1) - (3)	17	21	22	12	14	13

RT, retention time; Positive, positive ionization mode; Negative, negative ionization mode.

Figure Legends

Figure 1. Tyrosine catabolic pathway with reference to its defect in alkaptonuria and treatment with nitisinone. Lack of the enzyme homogentisate 1,2-dioxygenase increases concentrations of circulating homogentisic acid by preventing its conversion to maleylacetoacetic acid. It is previously established that nitisinone has the upstream consequence of increased concentrations of tyrosine, 3-(4-hydroxyphenyl)pyruvic acid and 3-(4-hydroxyphenyl)lactic acid as a result of the metabolic block induced by nitisinone.

Figure 2. Summary of the data acquired from analysis of metabolite standards. A, Venn diagram summarizing the number of compounds retained and detected by methods 1-3, both alone (non-overlapping sections) and in combination (overlapping sections). B-D, Mass/RT scatter plots for metabolite standards detected with the three analytical methods, showing the different selectivities of the methods for compounds from different chemical classes.

Figure 3. Summary of compounds identified from analysis of urine by accurate mass retention time (ARMT). A, Venn diagram summarizing the number of AMRT matches obtained for methods 1-3. B-D, Mass/RT scatter plots for AMRT compound matches obtained from analysis of urine by the three analytical methods.

Figure 4. Application of the profiling strategy to alkaptonuria (AKU). A, Urine was obtained from patients with AKU and $HGD^{-/-}$ mice pre- then post-nitisinone therapy. B, Representative urine profiles (top) and extracted signals (bottom) for compounds identified by accurate mass retention time. C, Principle component analysis showing alteration to human (left) and mouse (right) urine metabolomes post-nitisinone. x, y and z axes represent components 1, 2 and 3 respectively.

Figure 5. Urinary metabolite changes identified post-nitisinone in alkaptonuria by application of the profiling strategy. Red and blue indicate increases and decreases respectively. Fold changes (FC's) are indicated in brackets and were calculated from raw peak area.
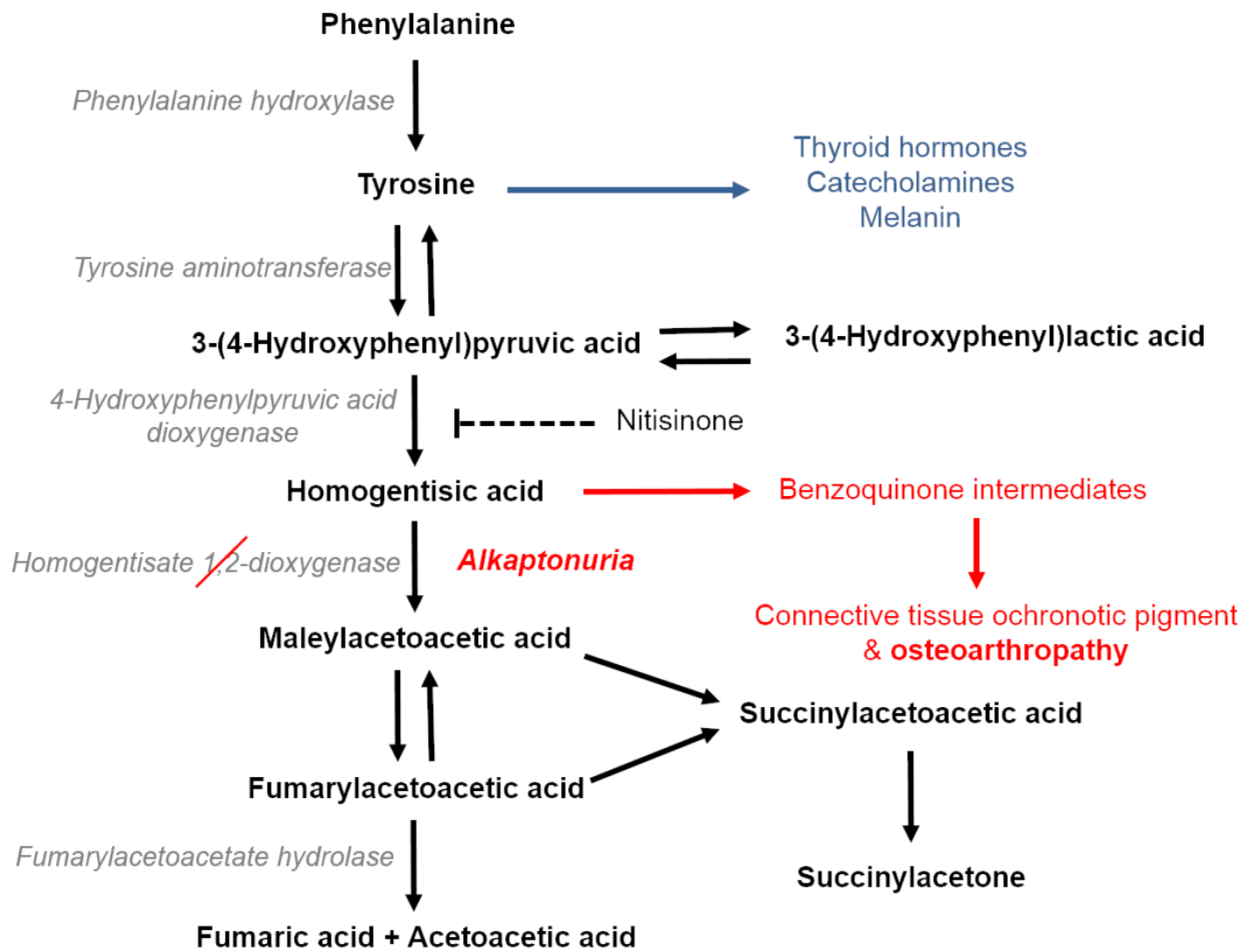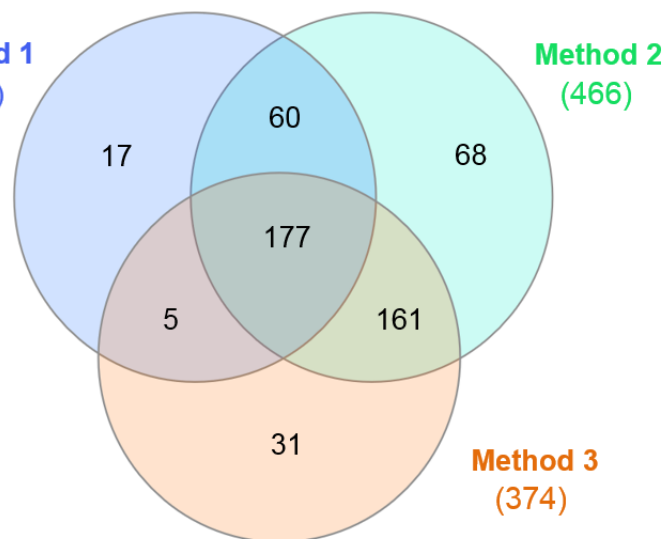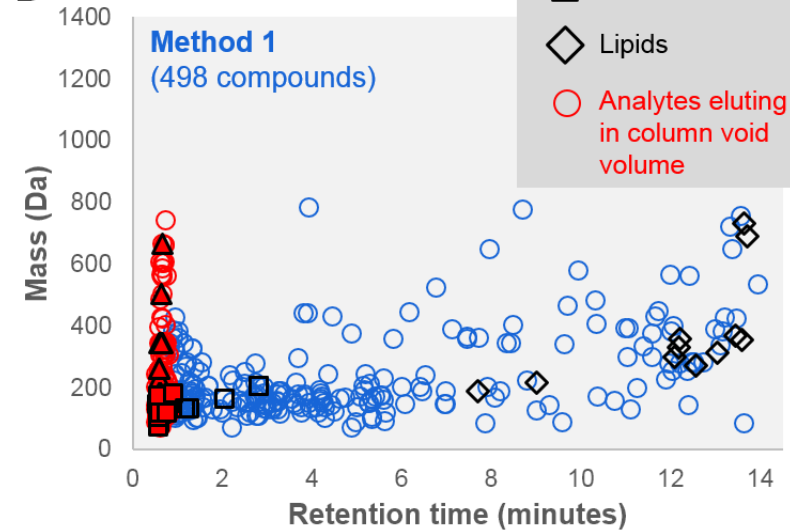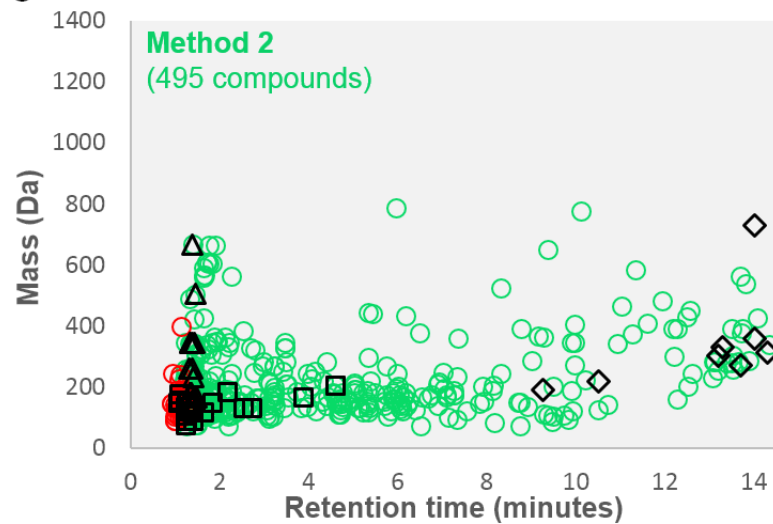
455
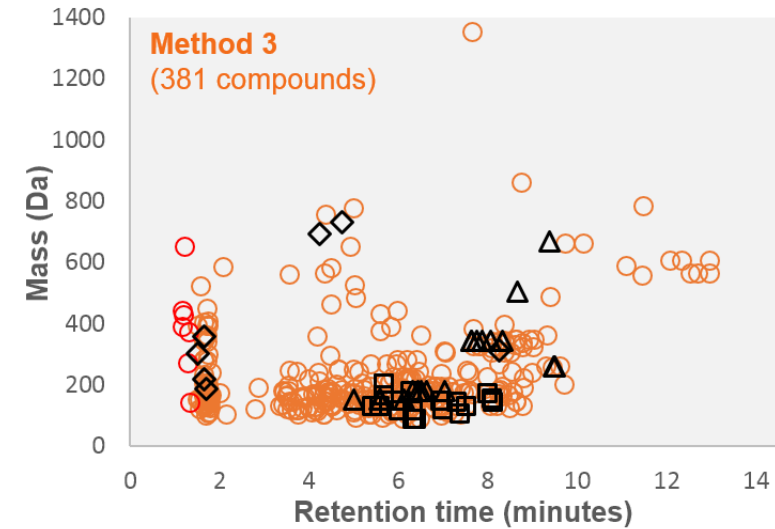
Figure 1
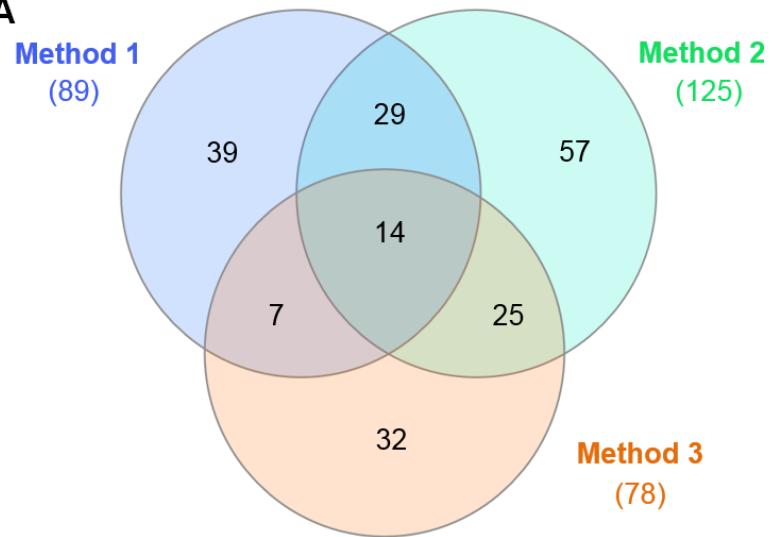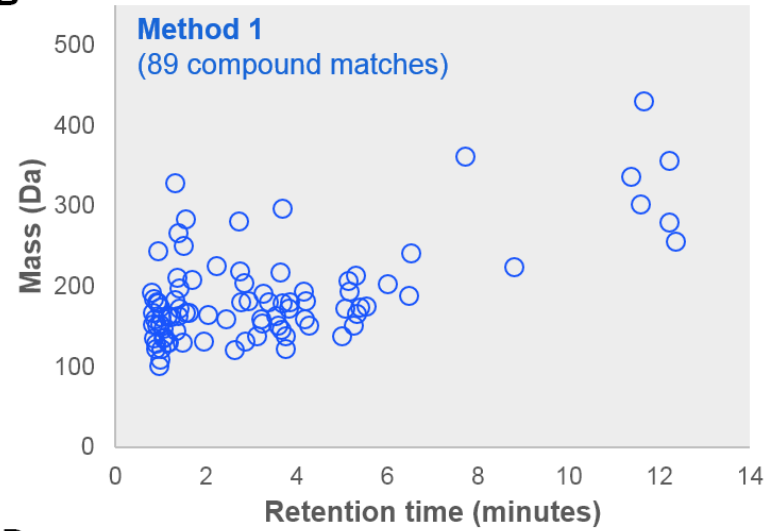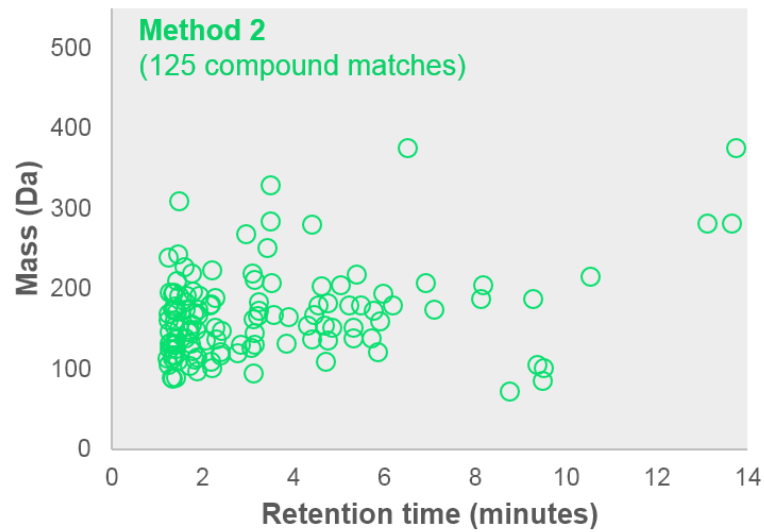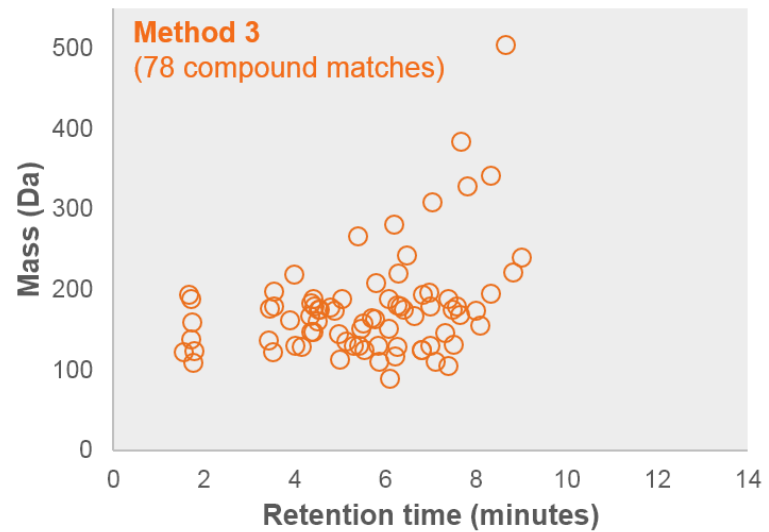
Figure 2
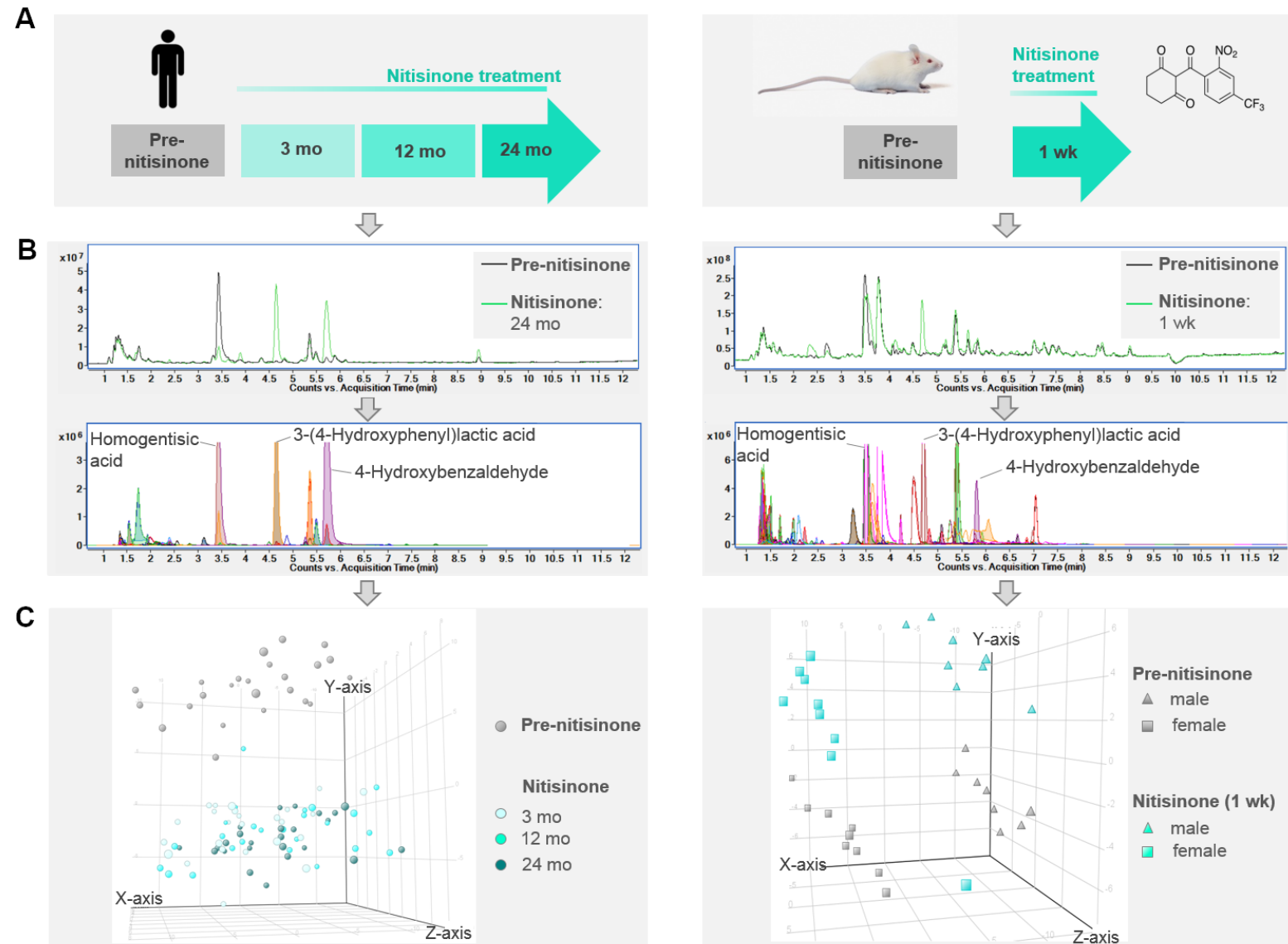
Figure 3

Figure 4

Figure 5

	Tyrosine		
	3-(4-Hydroxyphenyl)lactic acid (84, p<0.0001)	**L-Tyrosine*** (100+, p<0.0001)	*3,4-Dihydroxyphenyl glycol* (2.4, p<0.0001)
	3-(4-Hydroxyphenyl)pyruvic acid* (16, p<0.0001)	Mandelic acid (7, p<0.001)	**Homogentisic acid*** (5.3, p<0.0001)
	3-Methoxytyramine* (3.3, p<0.0001)	**Phenylacetic acid** (16, p<0.0001)	Homovanillic acid (4.2, p<0.0001)
	4-Coumarate (37, p<0.0001)	**Tyramine** (9.1, p<0.0001)	
	4-Hydroxybenzaldehyde (58, p<0.0001)		
	4-Hydroxyphenylacetic acid (2.1, p<0.0001)		

	Tryptophan	
	4-Quinolinecarboxylic acid (4.8, p<0.0001)	Indoxyl sulfate (11, p<0.0001)
	Xanthurenic acid (2.4, p<0.0001)	L-Kynurenine (4.8, p<0.0001)
		L-Tryptophan (2.2, p<0.0001)

	Purine	
	Caffeate (2.3, p<0.0001)	**3,5-Cyclic AMP** (4.3, p<0.0001)
		3-Ureidopropionate (100+, p<0.0001)
		Adenine (4.6, p<0.0001)
		Allantoin (100+, p<0.0001)
		Xanthosine (4.6, p<0.0001)

	Other		
	5-Valeroclactone (6.6, p<0.0001)	*Isocitric acid* (3, p=0.002)	2-Hydroxybutyric acid (1.7, p<0.0001)
	Alpha-ketoglutaric acid (2.2, p=0.0014)	*Maleimide* (2.8, p=0.0041)	**2-Hydroxy-4-(methylthio) butyric acid** (100+, p<0.0001)
	Creatine (2.1, p=0.032)		L-Threonine (7.3, p=0.022)
	Ethylmalonic acid (5.2, p<0.0001)		Methyl vanillate (21, p<0.0001)
			N-Acetylglycine (2.4, p<0.0001)

* Finding previously reported

FC's (brackets) are at 24 mo nitisinone for patients with AKU unless change observed in mouse only, in which case FC is at 1 wk nitisinone in mouse

Bold: Change observed in human and mouse data

Italics: *Change observed in mouse data only*

460

Supplemental Figure 1
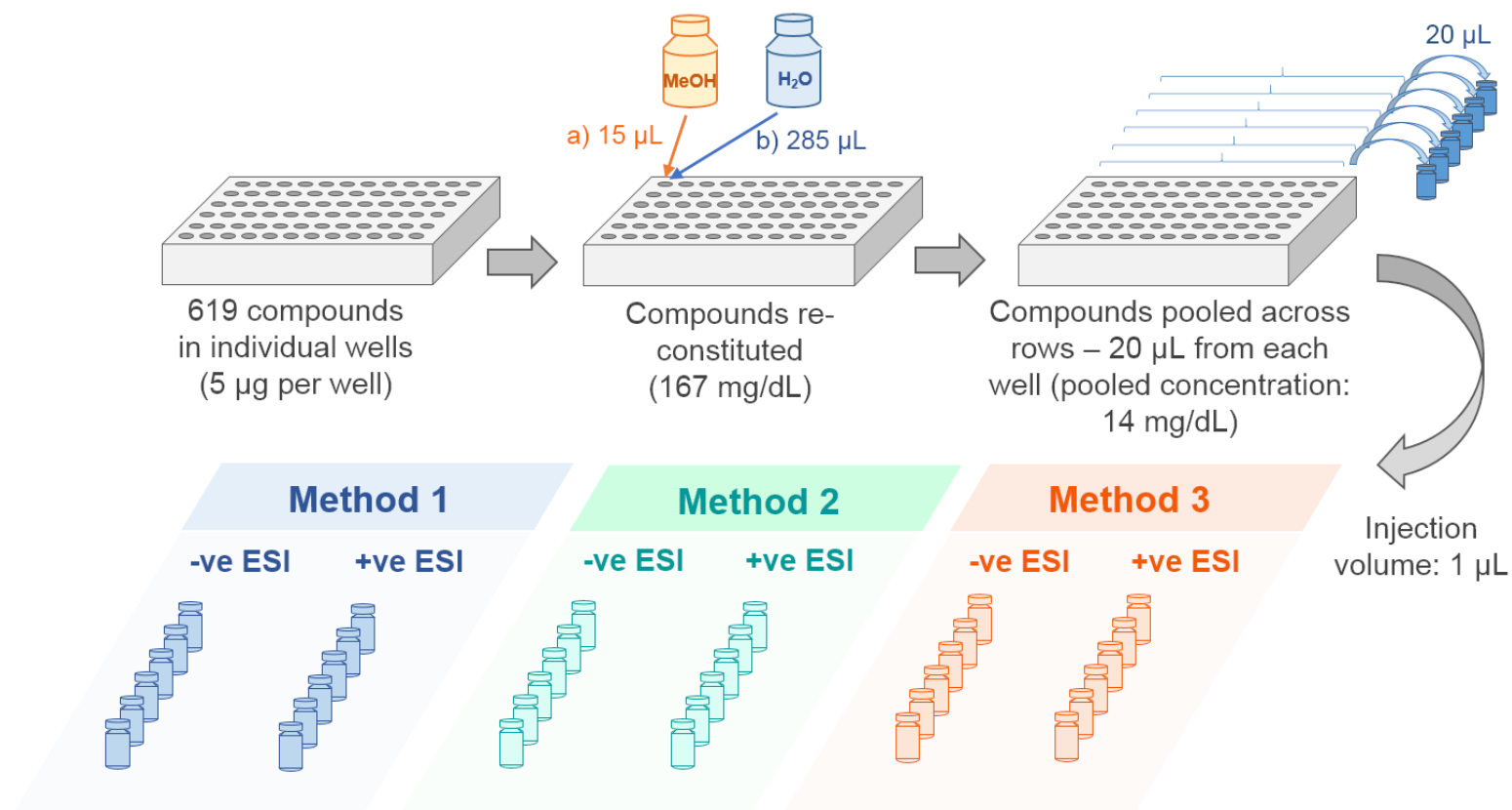
Supplemental Figure 1. Summary of experiment workflow for preparation and LC-QTOF-MS analysis of the 619 metabolite standards. Compounds were pooled across rows (A1-A12, B1-12 *etc.*) in each of seven 96-well plates. Analysis comprised 112 injections of compound pools incorporating positive (+ve ESI) and negative (-ve ESI) ionization modes for methods 1, 2 and 3.

Supplemental Figure 2

AM 10 ppm, RT 0.3 min

Method 1 (89), Method 2 (125), Method 3 (78)
39, 29, 57, 14, 7, 25, 32

AM 5 ppm, RT 0.15 min

Method 1 (69), Method 2 (81), Method 3 (53)
32, 22, 39, 8, 7, 12, 26

AM 2.5 ppm, RT 0.075 min

Method 1 (35), Method 2 (31), Method 3 (25)
20, 6, 18, 3, 6, 4, 12
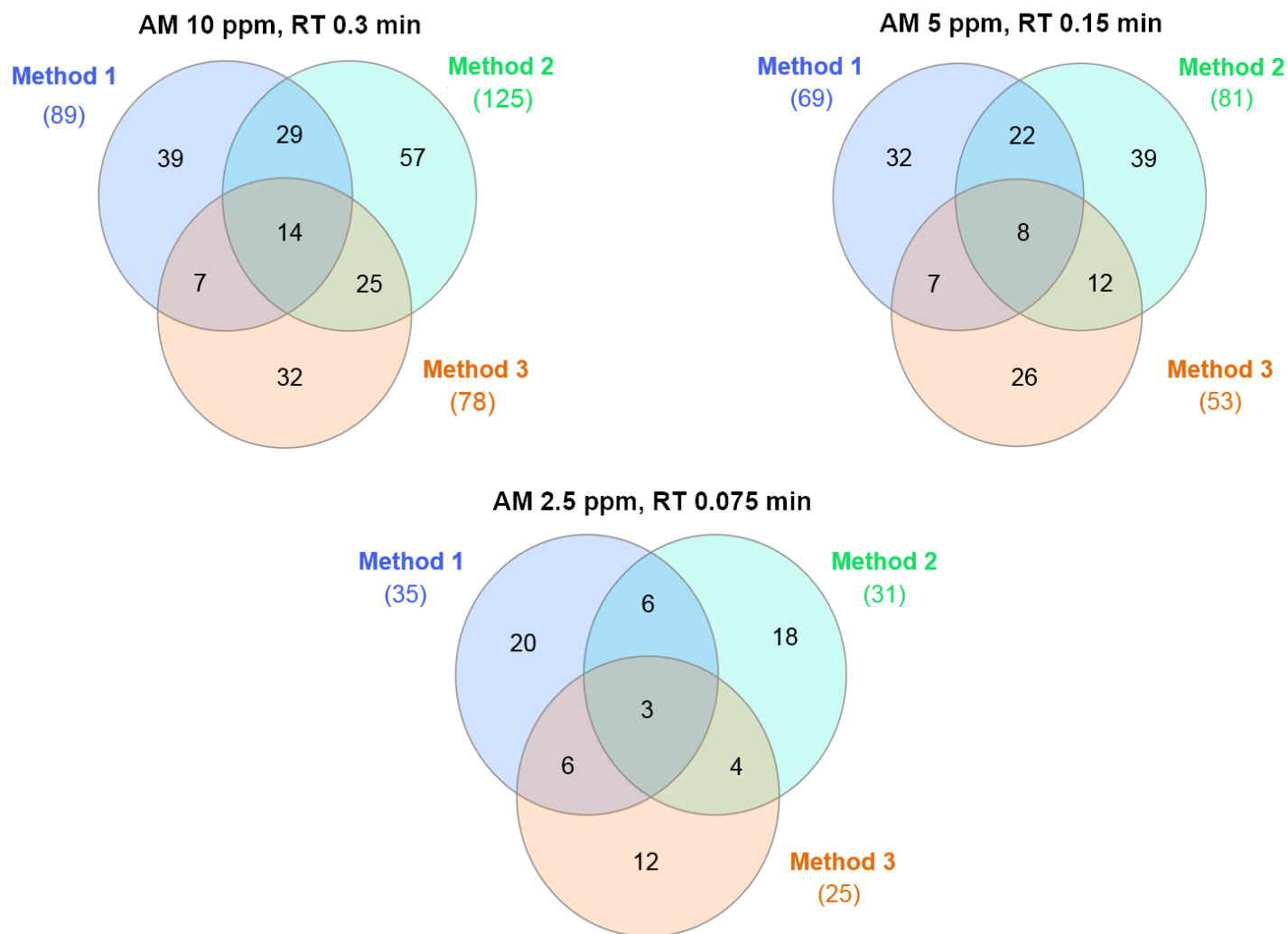
Supplemental Figure 2. Summary of the number of accurate mass retention time matches obtained by methods 1-3 with accurate mass (AM) and retention time (RT) windows of decreasing size.

Supplemental Figure 3
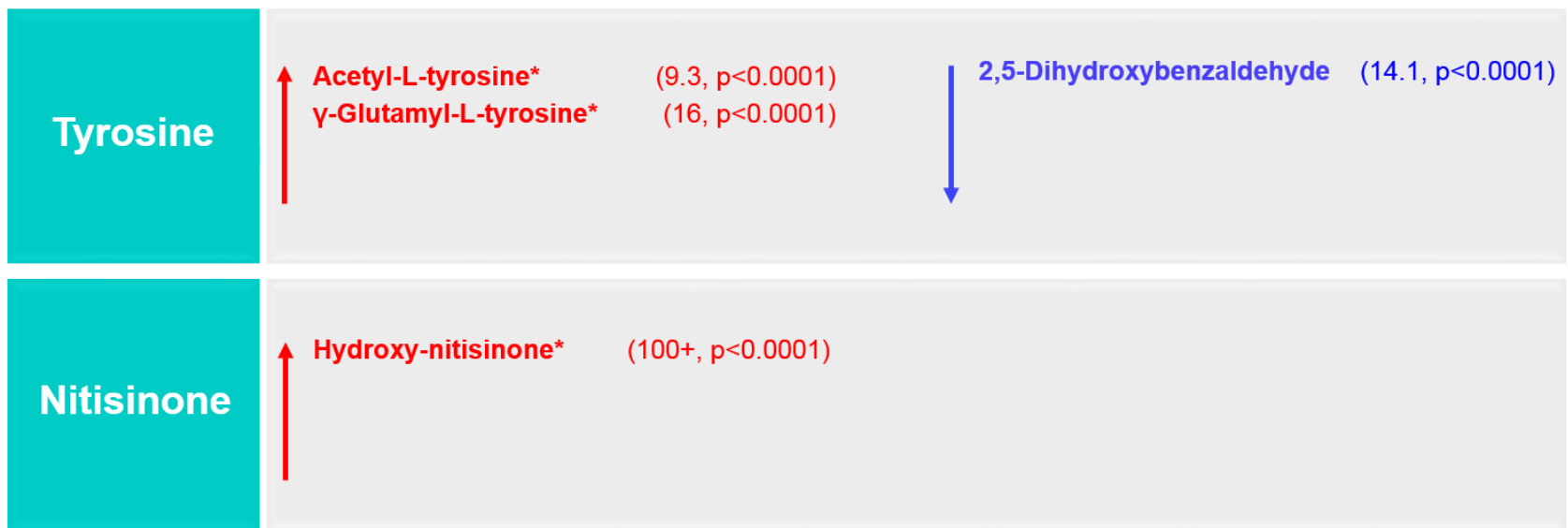
Bold: Change observed in human and mouse data
* Finding previously reported

Supplemental Figure 3. Non-AMRT urinary metabolite changes identified post-nitisinone. Red and blue indicate increases and decreases respectively. Fold changes (FC's; indicated in brackets) are at 24 mo for patients with AKU and were calculated from raw peak area. Compounds were identified by theoretical accurate mass match <10 ppm.

[Author accepted manuscript]

Title: Evaluation of the serum metabolome of patients with Alkaptonuria before and after two years of treatment with nitisinone using LC-QTOF-MS

Authors: Davison AS[1,2,*,†], Norman BP[2,†], Ross GA,[3] Hughes AT[1,2], Khedr M[1,2], Milan AM[1,2], Gallagher JA[2] Ranganath LR[1,2]

Department of Clinical Biochemistry and Metabolic Medicine, Liverpool Clinical Laboratories, Royal Liverpool University Hospitals Trust, Liverpool, members of Liverpool Health Partners, UK[1]; Musculoskeletal Biology I, Institute of Ageing and Chronic Disease, University of Liverpool, Liverpool, UK[2]; Agilent Technologies, Lakeside Drive, Cheadle, UK[3]

[†]Joint first authors
***Corresponding Author:** Andrew S Davison
Corresponding Author Address: Department of Clinical Biochemistry and Metabolic Medicine, Liverpool Clinical Laboratories, Royal Liverpool University Hospitals Trust, Liverpool, L7 8XP, UK; Telephone: 0151 706 4011; Fax: 0151 706 4250; e-mail: andrew.davison@rlbuht.nhs.uk

Word count: 3238 – main text only, including reference; Summary – 248

Number of: Tables – 1 (2 supplementary); Figures – 2 (1 supplementary, coloured images may be used for front cover)

Supplementary files: 4 (Tables – 2; text – 1; Figure – 1)

Abbreviations:
AKU – Alkaptonuria; HGA – homogentisic acid; HGD – homogentisate-1,2-dioxygenase; FAH – fumarylacetoacetate hydrolase; HPPD – hydroxyphenylpyruvic acid dioxygenase; HT1 – Hereditary Tyrosinaemia type 1; 3-(4-HPPA) – 3-(4-hydroxyphenyl)pyruvic acid; 3-(4-HPLA) – 3-(4-hydroxyphenyl)lactic acid; I-3-P – indole-3-pyruvate; I-3-L – indole-3-lactate; NAC – National Alkaptonuria Centre; AMRT – accurate mass retention time; QC – quality control; MS/MS – tandem mass spectrometry; MPP – mass profiler professional; 4-HPA – 4-hydroxyphenylacetate; 4-HBA – 4-hydroxybenzaldehyde

Summary

Background: The homogentisic acid-lowering therapy nitisinone is being evaluated for the treatment of Alkaptonuria (AKU) at the National Centre for AKU. Beyond hypertyrosinaemia the wider metabolic consequences of its use are largely unknown. The aim of this work was to evaluate the impact of nitisinone on the serum metabolome of patients with AKU after 12 and 24 months of treatment.

Methods: Deproteinised serum from 25 patients with AKU (mean age(±SD) 51.1±14.9y, 12 male) was analysed using the 1290 Infinity II LC system coupled to a 6550 QTOF-MS (Agilent, UK). Raw data were processed using a batch targeted feature extraction algorithm and an accurate mass retention time database containing 469 intermediary metabolites (mw 72-785). Matched entities (±10ppm theoretical accurate mass and ±0.3mins retention time window) were filtered based on their frequency and variability (<25% CV) in group QC samples, and repeated measures statistical significance analysis with Benjamini-Hochberg false discovery rate adjustment was used to assess changes in metabolite abundance.

Results: Eight metabolites increased in abundance ($\log_2$ fold change 2.1-15.2, $p<0.05$); 7 of 8 entities were related to tyrosine metabolism, and 13 decreased in abundance ($\log_2$ fold change 1.5-15.5, $p<0.05$); including entities related to tyrosine (n=2); tryptophan (n=3); xanthine (n=2) and citric acid cycle metabolism (n=2).

Conclusions: Evaluation of the serum metabolome of patients with AKU showed a significant difference in the abundance of several metabolites following treatment with nitisinone, including a number that have not been previously reported; several of these were not related to the tyrosine metabolic pathway.

Take home message

Nitisinone therapy has a significant impact on several metabolites beyond the tyrosine metabolic pathway, several of which appear to be related to the redox state of the cell**.**

Introduction

Alkaptonuria (AKU, OMIM: 203500) is a rare autosomal recessive disorder of tyrosine metabolism resulting from a defect in homogentisate-1,2-dioxygenase (HGD, E.C.1.12.11.5), which leads to a marked increase in the circulating concentration of homogentisic acid (HGA) (Phornphutkul et al. 2002). The pathological hallmark of AKU is 'ochronosis', which is a consequence of the deposition of a dark pigment in connective tissue, mainly cartilage, which alters its physiomechanical properties. The exact composition and structure of this pigment is unknown, but is known to result from the accumulation of HGA (Figure S1) (Ranganath et al. 2013).

Treatment options for AKU are conservative in large focusing on supportive and palliative measures (Ranganath et al. 2013). The HGA lowering agent nitisinone (Figure S1) has been shown to completely prevent ochronosis in AKU mice (Preston et al. 2014) and is being evaluated as a potential treatment in AKU patients (Phornphutkul et al. 2002; Suwannarat et al. 2005; Introne et al. 2011; Ranganath et al. 2016, 2018; Milan et al. 2017); it is however not without its own challenges as it is well documented to result in significant hypertyrosinaemia in AKU (Phornphutkul et al. 2002; Suwannarat et al. 2005; Introne et al. 2011; Olsson et al. 2015; Ranganath et al. 2016, 2018; Milan et al. 2017; Davison et al. 2018a) and Hereditary Tyrosinaemia type 1 (HT1) (Lindstedt et al. 1992; Lock 2017; van Ginkel et al. 2017). The consequences of this are largely unknown in AKU; in HT1 it has been suggested this may contribute to the neurodevelopmental delay that is frequently observed in children treated with nitisinone (McKiernan et al. 2015). Several mechanisms have been proposed for this including altered metabolism of the monoamine neurotransmitters (Thimm et al. 2011). Davison et al.(2018b; 2018c) demonstrated in a cohort of AKU patients that nitisinone therapy resulted in altered urinary excretion of dopaminergic and serotoninergic neurotransmitter metabolites. However these findings are limited as they are not a direct reflection of neurotransmitter metabolism in the central nervous system. A recent study in an animal model of HT1 suggested that the disease itself and not treatment with nitisinone is likely to be responsible for slower learning and altered behavior in mice (Hillgartner et al. 2016).

Recently, changes in the urine metabolome of an AKU mouse model and patients with AKU treated with nitisinone were reported (Norman et al. 2019). This study demonstrated novel changes in the tyrosine metabolic pathway, and unexpectedly in tryptophan and purine metabolism. While these changes in the urine metabolome provide insight into how nitisinone alters metabolism, one may postulate that changes observed in the serum are more relevant as they are a more direct reflection of internal homeostasis. Gertsman et al.(2015a) reported on the impact of nitisinone therapy on the serum metabolome of patients with AKU. In addition to the expected decrease in HGA and increase in tyrosine, significant increases in acetyl- and γ-glutamyltyrosine were also observed. In a separate publication (Gertsman et al. 2015b) in the same cohort of patients, novel disturbances in tryptophan metabolism were reported.

Herein for the first time we report the impact of nitisinone (2mg daily) therapy on the serum metabolome in the largest cohort of AKU patients to date over a two year period at the National AKU Centre (NAC) in the UK.

Materials and methods
Reagents

Water for mobiles phases was purified in-house (DIRECT-Q 3UV Millipore water purification system). Methanol, acetonitrile and isopropanol were purchased from Sigma Aldrich (Dorset, UK). Formic acid and ammonium formate were obtained from Biosolve, (Netherlands) and Fisher Scientific (Germany), respectively. All reagents were LC-MS grade. Acetyl-tyrosine was purchased from Sigma Aldrich, UK.

Patients and serum sample collection
Ethical Approval

Data collection and sample analyses at the NAC has approval from the Royal Liverpool and Broadgreen University Hospital Trusts Audit Committee (Audit no. ACO3836), this approval includes the use of patient data and biological material for metabolomics evaluation. As data and samples were collected as part of the clinical service, ethical approval was not required. Patients are informed verbally and through patient information leaflets about the clinical and research activities of the NAC and are informed that data may be used for publication. For a detailed protocol see Milan et al.(2017).

Sample collection

Serum samples (S-monovette, Sarstedt, Germany) were collected from patients after an overnight fast (≥8h). Patients' dietary intake of protein was managed through a 7-day food diary by a combination of lower protein in diet and phenylalanine/tyrosine-free meal exchanges. Samples were centrifuged at 1500xg for 10 min at 4°C; and then deproteinised with perchloric acid (60% 5.8M; 1:10, perchloric acid:serum), vortexed and centrifuged at 1500×g for 10 min. Supernatant was stored at -20°C until analysis.

Patient and quality control sample preparation

Patient samples were prepared by diluting 150μL serum with 450μL deionized water (DIRECT-Q 3UV Millipore water purification system). Diluted samples were then transferred into a 96 well plate which was then agitated on a plate shaker (MTS 2/4m IKA, Germany) at 600rpm for 10 min.

Patient group quality control (QC) samples were produced by adding 50μL of each patient sample into a single pool. In total four group QC pools were made [(i) baseline, (ii) 12-months, (iii) 24-months and (iv) overall – containing serum from all patients and visits; this acted as a system QC]. QC samples were prepared as per patient samples.

The analytical sequence of samples was performed as per published guidance (Vorkas et al. 2015). Each run commenced with 20 replicate injections of the overall pooled sample to condition the system. The order of individual samples was randomised computationally. Pooled samples were interspersed throughout the analytical sequence every tenth injection.

Analytical method

Chromatographic conditions

Liquid chromatography (LC) was performed on an Agilent 1290 Infinity II LC system. An Atlantis dC$_{18}$ column (3.0x100mm, 3μm, Waters, UK) was maintained at 60°C with a flow rate of 0.4mL/min. Mobile phases were (A) water and (B) methanol both containing 5mmol/L ammonium formate and 0.1% formic acid. The elution gradient started at 5% mobile phase B at 0-1 min increasing linearly to 100% B by 12 min, held at 100% B until 14 min, returning to 95% A for 5 min to recondition the column. Injection volume was 1μL.

Quadrupole time-of-flight mass spectrometry (QTOF-MS) conditions

An Agilent 6550 QTOF-MS equipped with a dual jet stream electrospray ionisation source was operated in 2GHz mode, over the mass range of 50-1700, in negative and positive polarities. A reference mass correction solution was continually infused at a flow rate of 0.5 mL/min via an external isocratic pump (Agilent, UK) for constant mass correction (see supplementary information 1 for additional details of QTOF-MS operating parameters and composition of reference ion solution).

Metabolite identification, data quality control and statistical analyses

Metabolite identification was carried out using an established accurate mass retention time (AMRT) database to match chemical entities (Norman et al. 2019). The database included theoretical accurate mass, measured retention time, and empirical formula. This was modified to include acetyl-tyrosine, γ-glutamyl-tyrosine and indole-3-lactate (I-3-L). AMRT data for acetyl-tyrosine were verified following the analysis of an analytical standard. AMRT data for γ-glutamyl-tyrosine and I-3-L were based on the elution time associated with theoretical monoisotopic mass of each compound and were not verified using an analytical standard. Data quality control and statistical analysis were performed using the MassHunter software suite (Agilent, UK). For additional details on data quality control and statistical analyses see supplementary information 1.

Results

Twenty-five patients [13 female, mean age (±standard deviation) 55.3(15.3) years (range 22-75); 12 male, mean age 44.2(15.8) years (range 22-70)] were included at baseline and after treatment with nitisinone at 12 and 24 months.

Raw data from LC-QTOF-MS analysis showed that retention time and accurate mass ranges were 1.15-13.93min and 75.0318-730.5955Da, in positive and 1.03-14.67min and 75.0318-722.6247Da in negative polarity, respectively. Principal components analysis (Figure 1) showed clear separation between the AMRT matched profiles of AKU patients pre- vs. post-nitisinone therapy.
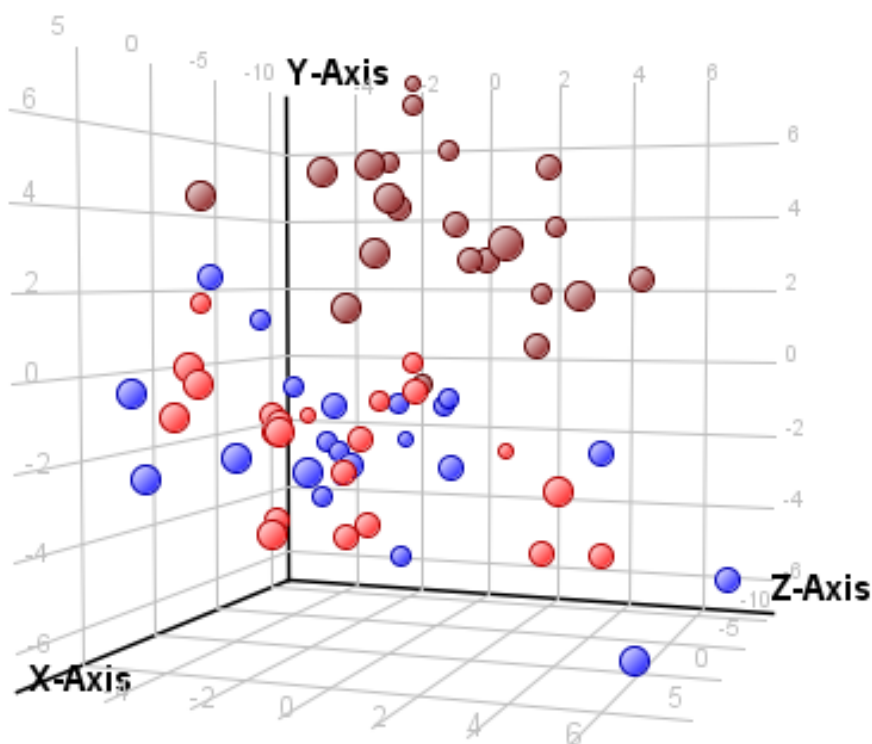
Figure 1. Principal component analysis of raw data from LC-QTOF-MS profiling experiments of serum samples from patients at baseline (pre-nitisinone, brown circle), 12-months (2mg daily nitisinone, red circle) and 24-months (2mg daily nitisinone, blue circle) x-axis – component 1, 22.09%; y-axis – component 2, 9.66%; z-axis – component 3, 7.78%.

151/469 and 249/469 metabolites were aligned across all samples in positive and negative polarities (for matched compounds see Table S2), respectively at baseline, and after nitisinone treatment at 12- and 24-months. After filtering entities based on their frequency and variability across replicate injections of pooled QC samples from each experimental group 123 and 209 entities were retained from positive and negative polarity profiling experiments, respectively. Of these, 60 and 121 respectively, were shown to be significantly different (p<0.05) following nitisinone therapy. Those with a $\log_2$ fold change (FC) >2 at 12 and or 24-months are summarised in Table 1. Applying this cut-off, eight (6.5%) entities were considered to increase in abundance and 13 (10.7%) decrease in abundance.

Table 1. Serum metabolite changes identified post-nitisinone therapy at 12- and 24-months in patients with AKU using an in-house AMRT database. Regulation expressed as $\log_2$ FC compared to baseline (pre-nitisinone treatment). $\log_2$ FC included if >2 at 12- and or 24-months.

Compound	$\log_2$ FC		p value	Regulation		Metabolic process/pathway affected
	12-months	24-months		Down	Up	
Glycocholate	12.7	12.7	<0.001	√		Bile acid
Succinic acid	11.4	11.4	<0.001	√		Citric acid cycle
α-Ketoglutaric acid	10.0	10.0	<0.001	√		Citric acid cycle
Trans-4-hydroxyproline	15.5	15.5	<0.001	√		Collagen
Homoserine	12.5	12.5	<0.001	√		Methionine/Cysteine
Mevalolactone	10.7	10.7	<0.001	√		Steroid
L-Allothreonine	11.8	11.8	0.02	√		Threonine
Trigonelline	12.8	12.8	<0.001	√		Tryptophan
4-Quinolinecarboxylic acid	8.3	8.3	<0.001	√		Tryptophan
Indole-3-lactate	2.1	2.3	<0.001		√	Tryptophan
4-Hydroxyphenylacetate	4.2	5.4	<0.001		√	Tyrosine
Benzaldehyde	6.9	6.9	<0.001	√		Tyrosine
Homogentisate	4.0	5.0	<0.001	√		Tyrosine
4-Hydroxybenzaldehyde	14.4	14.4	<0.001		√	Tyrosine
L-N-Acetyl-Tyrosine	15.5	15.2	<0.001		√	Tyrosine
γ-L-Glutamyl-L-tyrosine	3.2	3.0	<0.001		√	Tyrosine
3-(4-Hydroxyphenyl)lactate	6.4	6.4	<0.001		√	Tyrosine
L-Tyrosine	2.7	2.7	<0.001		√	Tyrosine
Mandelic acid	12.6	12.0	<0.001		√	Tyrosine
Inosine	10.8	10.8	<0.001	√		Purine
Uridine	2.1	1.5	<0.001	√		Pyrimidine

Nine (43%) of the 21 metabolites that were affected following treatment with nitisinone relate to tyrosine metabolism. Many of the other metabolites that had altered regulation following nitisinone therapy did not follow a clear theme apart from tryptophan, citric acid cycle and purine/pyrimidine metabolism.

Discussion

Over the last two decades there have been several reports on the use of nitisinone to treat AKU. Its inhibition of HPPD (Figure S1) has been shown to dramatically reduce the circulating concentration of HGA (Phornphutkul et al. 2002; Suwannarat et al. 2005; Introne et al. 2011; Ranganath et al. 2016, 2018; Milan et al. 2017), but leads to marked hypertyrosinaemia (Phornphutkul et al. 2002; Suwannarat et al. 2005; Introne et al. 2011; Olsson et al. 2015; Ranganath et al. 2016, 2018; Milan et al. 2017; Davison et al. 2018a). Beyond hypertyrosinaemia there is very little reported on the biochemical consequences of nitisinone therapy. Herein we report the impact of nitisinone therapy on the serum metabolome, using LC-QTOF-MS and a validated strategy to identify metabolites using an AMRT database developed in-house (Norman et al. 2019). This study is unique as it includes the largest cohort of patients with AKU to date, taking a 2mg daily dose of nitisinone over 24-months. In addition as it is based on the analysis of serum it is a better reflection of the impact of nitisinone on internal homeostasis.

This study confirms previous reports that nitisinone treatment in AKU results in a marked reduction in serum HGA and increase in tyrosine. In addition, a significant increase in 3-(4-hydroxyphenyl)lactate (3-(4-HPLA) was observed. Unexpectedly 3-(4-hydroxyphenyl)pyruvate (3-(4-HPPA), the metabolite immediately proximal to the site of action of nitisinone, was not increased (Figure S1). The marked increase in HPLA is an expected consequence of nitisinone therapy, but has not been reported in serum. There is an equilibrium between 3-(4-HPPA) and 3-(4-HPLA) (Figure S1) and it is likely that the reason for only observing significant increases in latter is that the sample pH shifted the equilibrium position to favour its formation. This pattern has been previously reported (Norman et al. 2019) in a study that reported on the urine metabolome of patients with AKU treated with nitisinone; a 84 and 16 fold (raw FC) increase in 3-(4-HPLA) and 3-(4-HPPA) were reported, respectively.

Marked increases in the tyrosine conjugates acetyl-tyrosine and γ-glutamyl-tyrosine were also observed. The $\log_2$ FC observed was very similar for tyrosine and γ-glutamyl-tyrosine, but that of acetyl-tyrosine was markedly higher. The latter suggests that an equilibrium shift between tyrosine and acetyl- tyrosine may exist, but not for tyrosine and γ-glutamyl-tyrosine. Norman et al.(2019) showed that urinary tyrosine was significantly higher than acetyl- and γ-glutamyltyrosine in a cohort of AKU patients on 2mg daily of nitisinone supporting that an equilibrium shift between tyrosine and tyrosine conjugates exists. In contrast, Gertsman et al.(2015a) reported a proportional increase in plasma tyrosine, acetyl- and γ-glutamyl-tyrosine following nitisinone suggesting there is no equilibrium shift. While direct comparisons between the magnitudes of FC cannot be made as herein we report $\log_2$ FC, the proportions of metabolites are clearly different between the two studies. This may be due to the small number of patients in the study reported by Gertsman et al (2015a) and that a 2mg dose was used only in our study. Nonetheless our findings support that non-traditional metabolic pathways are active in the face of tyrosine excess.

The elevated γ-glutamyl-tyrosine suggests that glutathione metabolism and the redox state of cell (Griffith et al. 1979; Zhang and Forman 2009) are altered following treatment with nitisinone. This is of particular importance to AKU as elevated HGA is thought to lead to a pro-oxidant environment where 'soluble melanins' are formed (Davison et al. 2016). Increasing evidence to support HGA induced oxidative stress has been reported *in vitro* in serum (Braconi et al. 2011) and cellular models (Braconi et al. 2010), and in patients with AKU (Braconi et al. 2014). For a recent review on oxidative stress and its contribution to the mechanisms of the ochronotic process see Braconi *et al* 2015. One may postulate that treatment with nitisinone reduces the burden on the glutathione cycle improving glutathione availability. In turn this may enable transfer of the glutamyl moiety from glutathione to tyrosine, via the action of γ-glutamyl-transpeptidase (Griffith et al. 1979; Zhang and Forman 2009) to form g-glutamyl-tyrosine. Interestingly, this study also revealed a marked decrease in the citric acid cycle metabolites α-ketoglutaric and succinic acid, the keto-acids of α-ketoglurarate and succinate, following nitisinone. In this study glutamine levels did not change, thus one may hypothesise that glutamate formed from glutamine is

preferentially converted to glutathione and not α-ketoglutaric and succinic acid (Figure 2). The significance of this is not understood and requires further investigation.
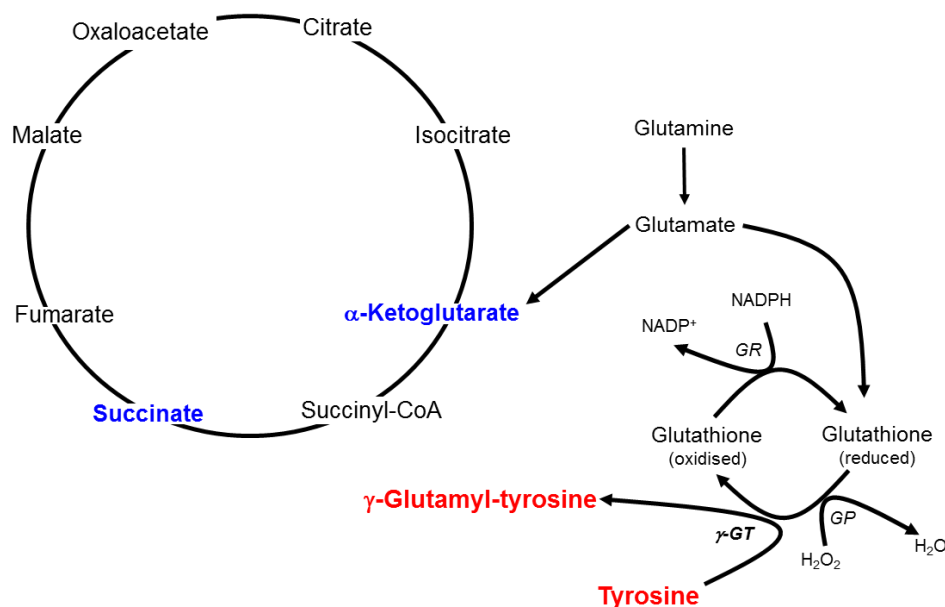
Figure 2. Proposed mechanism for the formation of γ-glutamyl-tyrosine in AKU following nitisinone therapy. Entities in **red** and **blue** represent an increase and decrease in metabolite abundance, respectively. GR – glutathione reductase; GP – glutathione peroxidase; GT - γ-glutamyltranspeptidase

The significance of increased acetyl-tyrosine is unknown, but has been reported in serum (Gertsman et al. 2015a) and urine (Norman et al. 2019) of AKU patients treated with nitisinone. It has also been observed in the urine of patients with Hereditary Tyrosinaemia type 2 (OMIM: 276600) (Macsai et al. 2001). One may postulate that acetyl tyrosine represents a more efficient way to eliminate the excess tyrosine from the body as it is more water soluble than tyrosine. Interestingly acetyl tyrosine has also been reported as an additive to foodstuffs given to patients on nutritional support (Hoffer et al. 2003).

4-Hydroxyphenylacetic (HPA) and 4-hydroxybenzaldehyde (4-HBA) were increased and benzaldehyde decreased following treatment with nitisinone; these have not been previously reported in the serum metabolome of patients with AKU following nitisinone treatment. HPA has been reported in urine following treatment with nitisinone (Norman

et al. 2019a). 4-HPA is generated from gut microbiota (Liu et al. 2016); one may hypothesise that the increase observed herein resulted from less oxidative stress following nitisinone, which increased microbiotic metabolism. This is the first report in humans, but has been seen previously observed in a rat model of oxidative stress (Liu et al. 2016).

The increase in 4-HBA has been reported in the urine of patients with AKU following nitisinone (Norman et al. 2019), but the decrease in benzaldehyde has never been observed. The significance of both entities is uncertain. 4-HBA is a naturally occurring compound that originates from the saprophytic perennial *Gastrodia elata* (Kang et al. 2017). Recent studies have reported the therapeutic effects of benzaldehydes in a number of areas including wound healing, cancer, vascular disease, and renal disease (Lee et al. 2010, 2014; Moon et al. 2012; Kong et al. 2014). In the context of AKU the significance of 4-HBA and benzaldehyde are unknown and require further investigation. One may speculate that they relate to the benzoquinones, proposed as intermediates in the formation of ochronotic pigment observed in AKU (Norman et al. 2019).

Mandelic acid was also increased following treatment with nitisinone, which has been reported previously (Norman et al. 2019). The significance of this is uncertain, but urinary elevation has previously been observed in patients with PKU on a phenylalanine restricted diet (Rampini et al. 1974).

The 'off target effects' of any drug are essential when considering its suitability in treating a patient (Lynch et al. 2017). Nitisinone therapy in AKU and HT1 have long been associated with altered tyrosine metabolism, and so can be considered a 'targeted effect'. More recently off targets effects of nitisinone have been reported due to its impact on tryptophan metabolism. A decrease in 5-hydroxyindoleacetic acid (serotonin metabolite) has been reported in the cerebral spinal fluid and urine of patients with HT1 (Thimm et al. 2011) and AKU (Davison et al. 2018b; 2018c), respectively. Serum tryptophan itself has been shown not to change following treatment with nitisinone (Davison et al. 2018a); in contrast urinary tryptophan has been shown to decrease (Norman et al. 2019). These differences may be explained by the fact that tryptophan is highly protein bound (~90-95%) (Cervenka et al. 2017)

and measurement in serum reflects total tryptophan and urinary tryptophan reflects free tryptophan. The biologically active fraction is the free fraction, which is not typically measured in serum, and the majority is metabolized via the kynurenine pathway (Cervenka et al. 2017). Herein I-3-L, 4-quinolinecarboxylic acid and trigonelline were the only tryptophan-related metabolites that were changed following nitisinone. The significance of the latter two metabolites is uncertain, but reinforces that downstream metabolism of kynurenine and niacin may be altered following nitisinone treatment. Gertsman et al (2015b) also demonstrated an increase in plasma I-3-L, but additionally indole-3-pyruvate (I-3-P). It has been proposed that the increased activity in the indole pyruvate pathway results from tryptophan aminotransferase having a higher affinity for tryptophan in the presence of keto-acids (e.g. 4-HPPA, 4-HPLA) (Lees and Weiner 1973). The reason for not observing I-3-P is uncertain, however one may speculate that I-3-L and not I-3-P was increased due to reaction conditions favoring its formation. In addition, differences may in part be explained by different approaches to sample preparation, chromatographic and mass spectrometry conditions, patient cohorts and nitisinone doses used. In addition a separate study showed urinary xanthurenic acid, and L-kynurenine and indoxyl-sulfate were increased and decreased following nitisinone therapy, respectively (Norman et al. 2019) The significance of these changes are unknown.

Beyond tryptophan metabolism there are limited reports on the off target effects of nitisinone on the metabolome. Herein we have shown that trans-4-hydroxyproline decreases significantly following treatment. This is of particular relevance to AKU as it suggests that there may be decreased collagen breakdown. This requires further investigation as previous authors (Taylor et al. 2017) have demonstrated a very low cartilage turnover state in AKU patients. In contrast, a different study reported that cartilage degradation, as well as bone resorption markers were elevated in AKU patients compared to controls (Genovese et al. 2015).

In addition, significant decreases were observed in the purine and pyrimidine precursor's inosine and uridine, respectively. In a previous study decreased excretion of the purine metabolites adenine and allantoin were reported in urine from AKU patients and an AKU mouse treated with nitisinone (Norman et al. 2019). Patients included in this study and previous (Norman et al. 2019) were not on uric acid-lowering

medication, however were on a protein-restricted diet which may contribute to this change. A plausible explanation for the observations in serum and in urine may also relate to changes in the oxidative state of the cell. Allantoin has previously been suggested as a marker of oxidative stress (Marrocco et al. 2017), one may postulate that the decreases observed in purine and pyrimidine metabolites are a reflection of reduced oxidation stress in the face of lower HGA concentrations (Marrocco et al. 2017).

Four unrelated entities were also shown to be decreased following nitisinone therapy; glycocholate, homoserine, mevalolactone and allothreonine. The reasons for this remain of unknown significance and requires further investigation.

It is important to highlight several limitations in this study. Firstly, while the number of patients included in this study is the largest reported to date, metabolites identified need validating in a larger cohort or patients. Moreover, an untreated group of AKU patients would provide greater credibility that the findings presented herein are a consequence of nitisinone therapy. In addition while we believe the AMRT database used in this study is comprehensive, the targeted evaluation of the serum metabolome precludes the identification of novel changes in metabolites. Furthermore, the use of non-selective sample preparation to gain broad coverage of the metabolome may have limited metabolite detection if present at a low concentration or if changes were not reproducible due to ion suppression.

Conclusions

Evaluation of the impact of nitisinone treatment on the serum metabolome of patients with AKU revealed a number of novel changes including a number that are not directly related to the tyrosine metabolic pathway. Some of these changes can be explained by the impact of nitisinone therapy on the cellular redox state and the wider impact of metabolites on enzyme activity. Further work is required to provide greater insight into the changes observed in the serum metabolome following nitisinone therapy.

References

Braconi D, Laschi M, Taylor AM et al (2010) Proteomic and redox-proteomic evaluation of homogentisic acid and ascorbic acid effects on human articular chondrocytes. *J Cell Biochem* 111:922e32

Braconi D, Bianchini C, Bernardini G et al (2011) Redox-proteomics of the effects of homogentisic acid in an in vitro human serum model of alkaptonuric ochronosis. *J Inherit Metab Dis* 34:1163e76

Millucci L, Ghezzi L, Paccagnini E et al (2014) Amyloidosis, inflammation, and oxidative stress in the heart of an alkaptonuric patient. *Mediators Inflamm* 2014:258471

Braconi D, Millucci L, Bernardini G, et al (2015) Oxidative stress and mechanisms of ochronosis in alkaptonuria. *Free Radic Biol Med* 88(Pt A):70-80

Cervenka I, Agudelo LZ, Ruas JL (2017) Kynurenines: Tryptophan's metabolites in exercise, inflammation, and mental health. *Science* 357:eaaf9794

Davison AS, Milan AM, Gallagher JA, Ranganath LR (2016) Acute fatal metabolic complications in alkaptonuria. *J Inherit Metab Dis* 39:203–210

Davison AS, Norman BP, Smith EA, et al (2018a) Serum Amino Acid Profiling in Patients with Alkaptonuria Before and After Treatment with Nitisinone. *JIMD Rep* 41:109-117

Davison AS, Norman B, Milan AM, et al (2018b) Assessment of the Effect of Once Daily Nitisinone Therapy on 24-h Urinary Metadrenalines and 5-Hydroxyindole Acetic Acid Excretion in Patients with Alkaptonuria After 4 Weeks of Treatment. *JIMD Rep* 41:1-10

Davison AS, Harrold JA, Hughes G, et al (2018c) Clinical and biochemical assessment of depressive symptoms in patients with Alkaptonuria before and after two years of treatment with nitisinone. *Mol Genet Metab* 125:135–143

Genovese F, Siebuhr AS, Musa K, et al (2015) Investigating the Robustness and Diagnostic Potential of Extracellular Matrix Remodelling Biomarkers in Alkaptonuria. *JIMD Rep* 24:29–37

Gertsman I, Barshop BA, Panyard-Davis J, et al (2015a) Metabolic Effects of Increasing Doses of Nitisinone in the Treatment of Alkaptonuria. *JIMD Rep* 24:13–20

Gertsman I, Gangoiti JA, Nyhan WL, Barshop BA (2015b) Perturbations of tyrosine metabolism promote the indolepyruvate pathway via tryptophan in host and microbiome. *Mol Genet Metab* 114:431–437

Griffith OW, Bridges RJ, Meister A (1979) Transport of gamma-glutamyl amino acids: role of glutathione and gamma-glutamyl transpeptidase. *Proc Natl Acad Sci USA* 76:6319–6322

Hillgartner MA, Coker SB, Koenig AE, et al (2016) Tyrosinemia type I and not treatment with NTBC causes slower learning and altered behavior in mice. *J Inherit Metab Dis* 39:673–682

Hoffer LJ, Sher K, Saboohi F, et al (2003) N-acetyl-L-tyrosine as a tyrosine source in adult parenteral nutrition. *J Parenter Enteral Nutr* 27:419–422

Introne WJ, Perry MB, Troendle J, et al (2011) A 3-year randomized therapeutic trial of nitisinone in alkaptonuria. *Mol Genet Metab* 103:307–314

Kang CW, Han YE, Kim J, et al (2017) 4-Hydroxybenzaldehyde accelerates acute wound healing through activation of focal adhesion signalling in keratinocytes. *Sci Rep* 7(1):14192

Kong BS, Cho YH, Lee EJ (2014) G Protein-Coupled Estrogen Receptor-1 Is Involved in the Protective Effect of Protocatechuic Aldehyde against Endothelial Dysfunction. *PLoS ONE* 9:e113242

Lee HJ, MiRanSeo, Lee EJ (2014) Salvianolic Acid B Inhibits Atherogenesis of Vascular Cells through Induction of Nrf2-dependent Heme Oxygenase-1. *Curr Med Chem* 21:3095–3106

Lee J-H, Thanigaimalai P, Lee K-C, et al (2010) Novel Benzo[d]imidazole-2(3H)-thiones as Potent Inhibitors of the α-Melanocyte Stimulating Hormone Induced Melanogenesis in Melanoma B16 Cells. *Chem Pharm Bull* 58:918–921

Lees GJ, Weiner N (1973) Transaminations between amino acids and keto acids elevated in phenylketonuria and maple syrup urine disease. *J Neurochem* 20:389–403

Lindstedt S, Holme E, Lock EA, et al (1992) Treatment of hereditary tyrosinaemia type I by inhibition of 4-hydroxyphenylpyruvate dioxygenase. *Lancet* 340:813–817

Liu G, Cao W, Fang T, et al (2016) Urinary Metabolomics Approach Provides New Insights into Distinct Metabolic Profiles of Glutamine and N-Carbamylglutamate Supplementation in Rats. *Nutrients* 8:478

Lock EA (2017) From Weed Killer to Wonder Drug. *Adv Exp Med Biol* 959:175–185

Lynch JJ, Van Vleet TR, Mittelstadt SW, Blomme EAG (2017) Potential functional and pathological side effects related to off-target pharmacological activity. *J Pharmacol Toxicol Methods* 87:108–126

Macsai MS, Schwartz TL, Hinkle D, et al (2001) Tyrosinemia type II: nine cases of ocular signs and symptoms. *Am J Ophthalmol* 132:522–527

Marrocco I, Altieri F, Peluso I (2017) Measurement and Clinical Significance of Biomarkers of Oxidative Stress in Humans. *Oxid Med Cell Longev* 2017:6501046

McKiernan PJ, Preece MA, Chakrapani A (2015) Outcome of children with hereditary tyrosinaemia following newborn screening. *Arch Dis Child* 100:738–741

Milan AM, Hughes AT, Davison AS, et al (2017) The effect of nitisinone on homogentisic acid and tyrosine: a two-year survey of patients attending the National Alkaptonuria Centre, Liverpool. *Ann Clin Biochem* 54:323–330

Moon CY, Ku CR, Cho YH, Lee EJ (2012) Protocatechuic aldehyde inhibits migration and proliferation of vascular smooth muscle cells and intravascular thrombosis. *Biochem Biophys Res Comm* 423:116–121

Norman BP, Davison AS, Ross GA, et al (2019) A Comprehensive LC-QTOF-MS Metabolic Phenotyping Strategy: Application to Alkaptonuria. *Clin Chem* 2019;65(4):530-539

Olsson B, Cox TF, Psarelli EE, et al (2015) Relationship Between Serum Concentrations of Nitisinone and Its Effect on Homogentisic Acid and Tyrosine in Patients with Alkaptonuria. *JIMD Rep* 24:21–27

Phornphutkul C, Introne WJ, Perry MB, et al (2002) Natural History of Alkaptonuria. *N Engl J Med* 347:2111–2121

Preston AJ, Keenan CM, Sutherland H, et al (2014) Ochronotic osteoarthropathy in a mouse model of alkaptonuria, and its inhibition by nitisinone. *Ann Rheum Dis* 73:284–289

Rampini S, Völlmin JA, Bosshard HR, et al (1974) Aromatic acids in urine of healthy infants, persistent hyperphenylalaninemia, and phenylketonuria, before and after phenylalanine load. *Pediatr Res* 8:704–709

Ranganath LR, Jarvis JC, Gallagher JA (2013) Recent advances in management of alkaptonuria. *J Clin Pathol* 66:367–373

Ranganath LR, Khedr M, Milan AM, et al (2018) Nitisinone arrests ochronosis and decreases rate of progression of Alkaptonuria: Evaluation of the effect of nitisinone in the United Kingdom National Alkaptonuria Centre. *Mol Genet Metab* 125:127–134

Ranganath LR, Milan AM, Hughes AT, et al (2016) Suitability Of Nitisinone In Alkaptonuria 1 (SONIA 1): an international, multicentre, randomised, open-label, no-treatment controlled, parallel-group, dose-response study to investigate the effect of once daily nitisinone on 24-h urinary homogentisic acid excretion in patients with alkaptonuria after 4 weeks of treatment. *Ann Rheum Dis* 75:362–367

Suwannarat P, O'Brien K, Perry MB, et al (2005) Use of nitisinone in patients with alkaptonuria. *Metabolism* 54:719–728

Taylor AM, Hsueh M-F, Ranganath LR, et al (2017) Cartilage biomarkers in the osteoarthropathy of alkaptonuria reveal low turnover and accelerated ageing. *Rheumatology* 56:156–164

Thimm E, Herebian D, Assmann B, et al (2011) Increase of CSF tyrosine and impaired serotonin turnover in tyrosinemia type I. *Mol Genet Metab* 102:122–125

van Ginkel WG, Jahja R, Huijbregts SCJ, van Spronsen FJ (2017) Neurological and Neuropsychological Problems in Tyrosinemia Type I Patients. *Adv Exp Med Biol* 959:111–122

Vorkas PA, Isaac G, Anwar MA, et al (2015) Untargeted UPLC-MS profiling pipeline to expand tissue metabolome coverage: application to cardiovascular disease. *Anal Chem* 87:4184–4193

Zhang H, Forman HJ (2009) Redox Regulation of γ-Glutamyl Transpeptidase. *Am J Respir Cell Mol Biol* 41:509–515